新規試験法提案書

皮膚感作性試験代替法（rLLNA 法）

平成25年 1 月

国立医薬品食品衛生研究所

新規試験法提案書

平成 25 年 1 月 20 日
No. 2012-01

皮膚感作性試験代替法 Reduced Local Lymph Node Assay (rLLNA)の提案

平成 24 年 7 月 4 日に東京、国立医薬品食品衛生研究所にて開催された新規試験法評価会議（通称：JaCVAM 評価会議）において以下の提案がなされた。

提案内容：皮膚感作性試験代替法 Reduced Local Lymph Node Assay (rLLNA)は 偽陰性（最高用量での反応低下等）の可能性を理解して適切に試験を実施し、結果を評価するとき、行政上のプログラムあるいは関係官庁が広く対象としている化学物質や製品の皮膚感作性の有無を科学的に評価できる。

この提案書は、米国Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)による第三者評価資料およびOECD Test Guideline (TG) 429の成立時により準備された資料をもとに、皮膚感作性試験代替法評価委員会によりまとめられた文書を用いてJaCVAM評価会議が評価および検討した結果、その有用性が確認されたことから作成された。

以上の理由により、行政当局の安全性評価方法として「皮膚感作性試験代替法rLLNA」の使用を提案するものである。

吉田武美

JaCVAM 評価会議　議長

西川秋佳

JaCVAM 運営委員会　委員長

JaCVAM 評価会議

吉田武美（日本毒性学会）：座長
浅野哲秀（日本環境変異原学会）
五十嵐良明（国立医薬品食品衛生研究所　生活衛生化学部）
大島健幸（日本化学工業協会）
小笠原弘道（独立行政法人　医薬品医療機器総合機構）
小野寺博志（独立行政法人　医薬品医療機器総合機構）
黒澤　努（日本動物実験代替法学会）
杉山真理子（日本化粧品工業連合会）
西川秋佳（国立医薬品食品衛生研究所　安全性生物試験研究センター）
長谷川隆一（独立行政法人　製品評価技術基盤機構）
牧　栄二（日本免疫毒性学会）
増田光輝（座長推薦）
横関博雄（日本皮膚アレルギー・接触皮膚炎学会）
吉田　緑　（国立医薬品食品衛生研究所　安全性生物試験研究センター　病理部）
吉村　功（座長推薦）
渡部一人（日本製薬工業協会）

任期：平成 24 年 4 月 1 日～平成 26 年 3 月 31 日

JaCVAM運営委員会

西川秋佳（国立医薬品食品衛生研究所　安全性生物試験研究センター）：委員長
大野泰雄（国立医薬品食品衛生研究所）
小川久美子（国立医薬品食品衛生研究所　安全性生物試験研究センター　病理部）
菅野　純（国立医薬品食品衛生研究所　安全性生物試験研究センター　毒性部）
斎藤和幸（独立行政法人　医薬品医療機器総合機構）
佐々木正広（厚生労働省　医薬食品局　化学物質安全対策室）
関野祐子（国立医薬品食品衛生研究所　安全性生物試験研究センター　薬理部）
高木篤也（国立医薬品食品衛生研究所　安全性生物試験研究センター　毒性部　動物管理室）
長谷部和久（厚生労働省　医薬食品局　化学物質安全対策室）
広瀬明彦　国立医薬品食品衛生研究所　安全性生物試験研究センター　総合評価研究室）
本間正充（国立医薬品食品衛生研究所　安全性生物試験研究センター　変異遺伝部）
光岡俊成（厚生労働省　医薬食品局　審査管理課）
小島　肇（（国立医薬品食品衛生研究所　安全性生物試験研究センター　薬理部　新規試験表評価室）：事務局

JaCVAM statement
on the Reduced Local Lymph Node Assay (rLLNA)
for skin sensitization assay

At the meeting concerning the above method, held on 4 July 2012 at the National Institute of Health Sciences (NIHS), Tokyo, Japan, the members of the Japanese Center for the Validation of Alternative Methods (JaCVAM) Regulatory Acceptance Board unanimously endorsed the following statement:

The reduced Local Lymph Node Assay (rLLNA) can be used as a screening test to distingish between sensitizer or non-sensitizer applied under proper conditions considering the possibility of false negative.

Following the review of the results of the ICCVAM (Interagency Coordinating Committee on theValidation of Alternative Methods, USA) Background Review Document, Evaluation Report, and OECD (Organisation for Economic Co-operation and Development) Test Guideline revised No. 429, it is concluded that the Reduced Local Lymph Node Assay (rLLNA) for skin sensitization assay is clearly beneficial.

The JaCVAM Regulatory Acceptance Board has been regularly kept informed of the progress of the study, and this endorsement is based on an assessment of various documents, including, in particular, the evaluation report prepared by the JaCVAM ad hoc peer review panel for skin sensitization assay.

Takemi Yoshida
Chairperson
JaCVAM Regulatory Acceptance Board

Akiyoshi Nishikawa
Chairperson
JaCVAM Steering Committee

20 January, 2013

The JaCVAM Regulatory Acceptance Board was established by the JaCVAM Steering Committee, and is composed of nominees from the industry and academia.

Mr. Takemi Yoshida (Japanese Society of Toxicology): Chairperson
Mr. Norihide Asano (Japanese Environmental Mutagen Society)
Mr. Yoshiaki Ikarashi (National Institute of Health Sciences: NIHS)
Mr. Takeyuki Oshima (Japan Chemical Industry Association)
Mr. Hiromichi Ogasawara (Pharmaceuticals and Medical Devices Agency)
Mr. Hiroshi Onodera (Pharmaceuticals and Medical Devices Agency)
Mr. Tsutomu Miki Kurosawa (Japanese Society for Animal Experimentation)
Ms. Mariko Sugiyama (Japan Cosmetic Industry Association)
Mr. Akiyoshi Nishikawa (Biological Safety Research Center: BSRC, NIHS)
Mr. Ryuichi Hasegawa (National Institute of Technology and Evaluation)
Mr. Eiji Maki (Japanese Society of Immunotoxicology)
Mr. Mitsuteru Masuda(nominee by Chairperson)
Mr. Hiroo Yokozeki (Japanese Society for Dermatoallergology and Contact Dermatitis)
Ms. Midori Yoshida (BSRC, NIHS)
Mr. Isao Yoshimura (nominee by Chairperson)
Mr. Kazuto Watanabe (Japan Pharmaceutical Manufacturers Association)

This statement was endorsed by the following members of the JaCVAM steering Committee after receiving the report from JaCVAM Regulatory Acceptance Board:

Mr. Akiyoshi Nishikawa (BSRC, NIHS): Chairperson
Mr. Yasuo Ohno (NIHS)
Ms. Kumiko Ogawa (Division of Pathology, BSRC, NIHS)
Mr. Jun Kanno (Division of Cellular and Molecular Toxicology, BSRC, NIHS)
Mr. Kazuyuki Saito (Pharmaceutical & Medical Devices Agency)
Mr. Masahiro Sasaki (Ministry of Health, Labour and Welfare)
Ms. Yuko Sekino (Division of Pharmacology, BSRC, NIHS)
Mr. Atsuya Takagi (Animal Management Section of the Division of Cellular and Molecular Toxicology, BSRC, NIHS)
Mr. Kazuhisa Hasebe (Ministry of Health, Labour and Welfare)
Mr. Akihiko Hirose (Division of Risk Assessment, BSRC, NIHS)
Mr. Masamitsu Honma (Division of Genetics and Mutagenesis, BSRC, NIHS)
Mr. Toshinari Mitsuoka (Ministry of Health, Labour and Welfare)
Mr. Hajime Kojima (Section for the Evaluation of Novel Methods, Division of Pharmacology, BSRC, NIHS):Secretary

皮膚感作性試験代替法 rLLNA

目次

皮膚感作性試験代替法 Reduced Local Lymph Node Assay (rLLNA)の

評価会議報告書

JaCVAM 評価会議

平成 24 年（2012 年）7 月 4 日

JaCVAM 評価会議

吉田武美（日本毒性学会）：座長

浅野哲秀（日本環境変異原学会）

五十嵐良明（国立医薬品食品衛生研究所　生活衛生化学部）

大島健幸（日本化学工業協会）

小笠原弘道（独立行政法人　医薬品医療機器総合機構）

小野寺博志（独立行政法人　医薬品医療機器総合機構）

黒澤　努（日本動物実験代替法学会）

杉山真理子（日本化粧品工業連合会）

西川秋佳（国立医薬品食品衛生研究所　安全性生物試験研究センター）

長谷川隆一（独立行政法人　製品評価技術基盤機構）

牧　栄二（日本免疫毒性学会）

増田光輝（座長推薦）

横関博雄（日本皮膚アレルギー・接触皮膚炎学会）

吉田　緑　（国立医薬品食品衛生研究所　安全性生物試験研究センター　病理部）

吉村　功（座長推薦）

渡部一人（日本製薬工業協会）

任期：平成 24 年 4 月 1 日～平成 26 年 3 月 31 日

以上

皮膚感作性試験代替法 Reduced Local Lymph Node Assay (rLLNA)について、第三者評価委員会からの報告を受け[1)]、以下の 9 項目について審議した。本 2～8 項目は OECD ガイダンス文書　No. 34 に示された検討項目である[2)]。なお、本動物実験代替法の利用にあたっては、適用範囲を十分に配慮した上で使用されるべきである。

＜審議内容＞

１．検討対象の試験法は、日本のどの法規制やガイドラインに関係しているか。

Reduced local lymph node assay (rLLNA) は、化学物質の皮膚感作性を評価する試験法である。rLLNAの原理、試験手順及び判定基準は、既にOECDテストガイドラインに収載されているLocal lymph node assay (LLNA) と同一で、一用量のみで試験する点だけが異なる。

rLLNAは、医薬品、医療機器、医薬部外品又は化粧品の原料、その他の物質又は製品に関する法規制又はガイドラインが求めている皮膚感作性に関係する。

２．検討対象の試験法とその妥当性を示すデータは、透明で独立な評価を受けているか。

ICCVAM は、第三者評価委員会を組織し、11 施設で実施された LLNA の試験成績を基に、rLLNA の試験法及び妥当性を評価した。その組織及び評価結果は ICCVAM のホームページで公表されている。

また、我が国の rLLNA 第三者評価委員会は、上記 ICCVAM の結論を評価し、その結果を評価会議に提出した。この結果も、評価会議での審議が終了後公開される予定である。

よって、rLLNA は透明で独立な評価を受けていると判断する。

３．当該試験法で得られるデータは、対象毒性を十分に評価あるいは予測できるものであるか。データは、当該試験法と従来の試験法の、代替法としての繋がりを示しているか。あるいは（同時に）そのデータは、当該試験法と、対象としているあるいはモデルとしている動物種についての影響との繋がりを示しているか。

皮膚感作性物質は、適用部位の所属リンパ節を活性化する。従来のLLNAはこの原理に基づき、マウスに種々の濃度の被験物質を適用し、所属リンパ節細胞の増殖反応を[^{3}H-methyl] thymidineの取り込みを指標として求め、皮膚感作性の有無を評価する。rLLNAも同様の原理に基づく試験法である。

LLNAでは適切な用量範囲で試験した場合、反応強度との間に用量相関性があるとされている。当該試験法で得られるデータは、最も反応が強く表れると想定される一用量での作用をとらえるものであるが、適切な用量が設定されれば、LLNAと同様に、ヒトに対する被験物質の皮膚感作性を十分に評価あるいは予測できるものである。

４．当該試験法は、ハザードあるいはリスク、あるいはその両方を評価するのに有用であるか。

rLLNAは、被験物質の皮膚感作性の有無のみを評価するものであり、適用用量と皮膚感作性の程度との関係を調べるものでは無い。

５．当該試験法とその妥当性を示すデータは、その試験法で安全性を保証しようとする、行政上のプログラムあるいは関係官庁が対象としている化学物質や製品を、十分広く対象としたものとなっ

ているか。当該試験法が適用できる条件及び適用できない条件が明確であるか。

rLLNA試験法とその妥当性を示すデータは、様々な用途で使用されている被験物質465物質について行われた11施設でのLLNAの試験成績を元にICCVAMが作成したものであり、その整合性を我が国のrLLNA第三者評価委員会で評価したものである。これらのデータによりrLLNAの皮膚感作性試験としての妥当性が示されている。よって、上記製品又は原料を幅広くカバーするデータが提供されていると見なせる。

当該試験法においては、過度の局所刺激性や明白な全身毒性を示さない最高用量で試験するが、LLNAと比較して1.9%（6/315）の偽陰性がみられている。その原因としては、設定された最高用量での反応性が低下したことが考えられる。

偽陰性の疑いがある場合は被験物質の毒性や物性等の情報を集め、また用量反応情報が必要な場合はLLNAを実施すべきである。

６．当該試験法は、プロトコルの微細な変更に対して十分頑健で、適切な訓練経験を持つ担当者と適切な設備のある施設において、技術習得が容易なものであるか。

LLNA法と原理的に同じであることから、rLLNAの精度、施設内及び施設間再現性及び頑健性はLLNAと同等と考えられるが、一用量のみを用いることから、その用量の選定や被験物質の溶解性等により評価がばらつく可能性がある。よって本試験法は、ICCVAMが推奨する最新の試験計画書に準拠する必要がある。

適切な訓練経験を持つ担当者と適切な設備のある施設において、上記の留意点以外においては、rLLNAはLLNAと比べて技術的に困難になるところはなく、技術習得は容易である。

７．当該試験法は、時間的経費的に有用性があり、行政上で用いられやすいものであるか。

rLLNAは、LLNAと同じ試験手順をとることから、LLNAに対して時間的な優位性は小さいが、使用動物数を最大40%まで削減でき、経費的には優位である。

化学物質の皮膚感作性の有無を区別するために、まずrLLNAを実施し、偽陰性の可能性がある場合は被験物質の毒性や物性等の情報を集め、用量反応情報を得るためのLLNAの実施を考慮すべきである。

陽性反応が得られた場合において、行政上、その用量依存性が求められたときは追加試験が必要となる。

８．当該試験法は、従来の試験法と比べて、科学的・倫理的・経済的に、新しい試験法あるいは改訂試験法であることが正当化されているか。

rLLNAは、被験物質を一用量のみで試験することにより従来のLLNAに比べて使用動物数を削減でき、かつ同等の精度を有する改訂試験法であることが、ICCVAM及び我が国のrLLNA第三者評価委員会で評価されている。

９．安全性評価のための行政的資料として、 受け入れ可能な試験法であるか。

rLLNAは、医薬品、医療機器、医薬部外品又は化粧品の原料、並びにその他の物質又は製品の皮膚感作性の有無を評価するために有用である。

rLLNAでは用量反応情報は得られないため、リスクアセスメントを行うための評価用資料として利用するには不十分であるが、rLLNAで陰性と判定された物質は、他の科学的情報から皮膚感作性が疑われる可能性がある場合を除き、これ以上の皮膚感作性に関する試験を求めないとすることも可能である。

以上の審議の結果、JaCVAM 評価会議は、皮膚感作性試験代替法 rLLNA について以下のように結論した。

本試験法は、従来の LLNA を一用量で行い使用動物数を削減する改訂試験法であり，LLNA と同等の検出感度を有する。

本試験法は、最新の ICCVAM が推奨する試験計画書に準拠し、偽陰性（最高用量での反応低下等）の可能性を理解して適切に試験を実施し、結果を評価するとき、行政上のプログラムあるいは関係官庁が広く対象としている化学物質や製品の皮膚感作性の有無を科学的に評価できる。

参考文献

１．皮膚感作性試験代替法 Reduced Local Lymph Node Assay (rLLNA)の第三者評価報告書

２．OECD (2005) OECD Series on testing and assessment Number 34, Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment, ENJ/JM/MONO(2005) 14

３．Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), National Toxicology Program (NTP), et al. Background Review Document:“*in vitro* Cytotoxicity Test Methods for Estimating Acute Oral systemic Toxicity”, NIH Publication No: 07-4518

Reduced Local Lymph Node Assay (rLLNA、単群局所リンパ節試験*)の第三者評価報告書

（* 試験法の邦文名は、提案名称）

平成 23 年 11 月 19 日

皮膚感作性試験代替法評価委員会

委員長　金澤由基子（医薬品機構）
委員　　牧栄二（元・安評センター）

皮膚感作性試験法としては、皮膚外用剤として用いる医薬品ならびに化粧品原料を含む化学物質等の皮膚での接触感作性のリスクを動物で予測する試験法としてモルモットを用いる Maximization Test（GPMT）法があり、その試験成績は皮膚外用剤の臨床での皮膚感作性と良い相関性を有することが示されている。また、GPMT の代替試験法として、マウスを用いる Local Lymph Node Assay (LLNA，局所リンパ節試験)があり、その予測率は、GPMT に劣らないとされ、国際的に認知されている。両者の違いは、GPMT が感作誘発期の皮膚反応を肉眼的に判定するのに対して、LLNA は感作に基づく耳介リンパ節の細胞増殖反応を[^{3}H-methyl]-thymidine（^{3}H-TdR）の取り込み量を測定することで定量的かつ客観的に判定するところにある。また、2010 年に改定された OECD テストガイドライン（TG）429 には、LLNA と Reduced Local Lymph Node Assay (rLLNA)が記載されている。rLLNA は、その手法において元の LLNA の試験法を踏襲するものであり、異なる点は、LLNA が 3 用量の被験物質で検討されるのに対し、rLLNA は 1 用量の被験物質で皮膚感作性を検討するところにある。また、rLLNA は、動物福祉の 3R の原則（動物実験代替法の活用（replacement）、使用動物数の削減（reduction）、苦痛の軽減（refinement））に沿ったものであり、使用動物数においては、元の LLNA と比較して、最大（注解）40%まで動物数が削減される（注解：陽性対照群が設定しない場合）。

本報告では、本邦において未だ検証されていない rLLNA の試験法について、その有用性と限界に関する ICCVAM の勧告（2008）を基に編纂を行う。

１．試験法

1-1　原理

rLLNA の試験法の原理は、LLNA の試験法の原理と同じである。即ち、皮膚感作性を有する低分子化合物が経皮投与されると、皮膚組織中のタンパク質と結合し、感作抗原として皮膚の樹状細胞に認識される。その後、樹状細胞は活性化しながら皮膚から所属リンパ節へ遊走し、抗原提示を行い抗原特異的な T リンパ球細胞の増殖を誘導する。この一連の生体応答が感作誘導期である。LLNA および rLLNA では、感作誘導期のリンパ節における抗原特異的な T リンパ球細胞の増殖を、^{3}H-TdR の核酸成分への取り込みを指標として検出する。LLNA と rLLNA の試験法における唯一の違いは、被験物質の試験に用いる用量数にある。LLNA においては、最大溶解度で過度の局所刺激性を示さず、更には、全身毒性を示さない濃度を最高用量とし、各被験物質に対して少なくとも 3 用量が試験に使用されるが、rLLNA においては、LLNA で示す最高用量の 1 用量のみが試験に使用される。

1-2　試験手順および判定

1-2-1　試験手順

8-12 週齢の CBA/J あるいは CBA/Ca 系の雌マウスを用い、マウスの両耳に被験物質を各々25μL、3 日間塗布し、その 3 日後に ^{3}H-TdR を尾静脈投与する。その 5 時間後に耳介

リンパ節を摘出し、その耳介リンパ節細胞中に取り込まれた ^{3}H-TdR 量を測定する。

1-2-2　判定

溶媒対照群に対する被験物質投与群の ^{3}H-TdR 取り込み量の比（Stimulation index：SI）が 3 倍を越えた際に、陽性と判定する。判定においては、陰性対照群および陽性対照群の反応も考慮する。LLNA においては、結果が明確でない場合、用量相関性の強さ、統計学的有意差も考慮されるが、rLLNA では実施できない。

２．rLLNA の精度

潜在的に皮膚感作性を有する物質を確認するための rLLNA の精度は、LLNA で実施された 11 施設の試験成績を基に、retrospective に評価された。

LLNA では 471 被験物質のうち、318 被験物質が陽性であり、153 被験物質が陰性であった。或る被験物質が同じ溶媒で 1 回以上試験されている場合を 1 試験として数えると、465 被験物質が評価の対象となり、その中の 315 被験物質が感作性として分類され、150 被験物質が非感作性として分類された。6 被験物質は、LLNA において最高用量以外の用量の SI 値が 3 以上を示したため陽性と評価された。換言すれば、rLLNA においては偽陰性となる。rLLNA は、最高用量でのみ試験され、評価されるため、これら 6 被験物質（2-Methyl-2H-isothiazole-3-one、C19-azlactone、Azithromycin、Camphorquinone、Nickel sulfate および Non-ionic surfactant 2）は、LLNA と比較すると、非感作性物質として位置づけられた。

３．rLLNA の施設間再現性

rLLNA の施設間再現性は、2 或いは 3 施設において、次の 5 被験物質について同じ溶媒を用いて個々に試験され、LLNA と比較して評価された。

- Dinitrochlorobenzene (DNCB)
- Hexyl cinnamic aldehyde (HCA)
- Linalool alcohol
- Methyl salicylate
- Potassium dichromate

各施設における全ての試験は、DNCB と potassium dichromate を感作性物質として、methyl salicylate を非感作性物質として分類した（一致率 100%）。

HCA と linalool alcohol は、2 施設で試験され、LLNA 試験では感作性物質として分類され、rLLNA では非感作性物質として分類された。これら 2 試験法における一致しない結果を検証してみると、試験された最高用量での両試験における結果の違いによるものであった。しかしながら、rLLNA と LLNA は同一の試験計画書を使用しており、両者の精度を評価するために使用した一連のデータは類似しているため、rLLNA の施設内および施設間信頼性は、LLNA に類似するものと考えられる。

４．試験法の有用性と限界

rLLNA の科学的な価値は、十分に評価されており、最新の ICCVAM が推奨した LLNA 試験計画書（ICCVAM 2008a）に準拠して行われたとき、rLLNA の成績は、用量-反応情報を必要としない場合には、皮膚感作性物質と非感作性物質を区別するに十分であると、ICCVAM は結論している。

- LLNA と比較すると、rLLNA は、各々の試験に対して最大 40%まで動物数を減らすことができる。
- 化学物質や製品のアレルギー性接触皮膚炎（ACD）の潜在性を調べるために LLNA を行う前に、先ず rLLNA を実施することを ICCVAM は勧めている。
  - rLLNA において陽性と評価された物質は、感作性物質として分類できる。
  - 用量-反応情報が必要な場合、用量段階を持つ LLNA において試験されなければならない。
  - ACD を誘発すると考えられる物質については、rLLNA よりむしろ LLNA を最初から実施して評価すべきである。
- LLNA と比較して rLLNA に 1.9% (6/318)の偽陰性の結果を示す可能性がある。
  - この情報は rLLNA の結果を評価するとき考慮されるべきであり、陰性の結果は常に補足情報（例えば、高用量での反応低下の可能性、類似物質による試験結果、タンパク結合活性、分子量、その他の試験データ）を統合して評価することを促すべきである。
  - もし偽陰性の結果が示唆されるならば、LLNA 或いは他の公認の皮膚感作性試験法による確認試験が考慮されるべきである。

５．試験計画書

rLLNA は、ICCVAM が推奨する最新の LLNA 試験計画書に従って実施されることを ICCVAM は勧めている。その重要な点は、以下の通りである。

- 使用される最高濃度は、過度の局所刺激性や明白な全身毒性を示さない、最大溶解濃度であること。
- 動物の成績は、個々に収集されること。
  - これは、偽陰性や偽陽性の結果を引き起こす外れ値の識別や棄却に備えるためである。
  - 動物の成績を個々に収集することは（合算して得た成績と比較して）、被験物質の反応が溶媒対照群の反応と有意に異なっているか否かを決めるための統計解析に備えるためである。
- 投与群には最低 4 匹の動物が使用されること。
  - OECD TG 429 の LLNA では、投与群には少なくとも 4 匹の動物が必要であるとし

ている。

－ 統計解析によれば、投与群の動物数が 4 匹でも、統計解析の結果に有意な影響はないようである。

・ 陽性対照物質は、試験毎に使用されること。

－ これは、この試験の実施手順や試験系の全てが、陽性反応を生ずるために適切に反応しているかを明らかにする。

６．今後の検討事項

ICCVAM は、潜在的な皮膚感作性を確認するための rLLNA を更に特徴付け、その有用性と応用性を高めるために下記の追加の検討を推奨している。

・ 今回の評価において、LLNA と比較し、rLLNA に偽陰性の結果を生じさせた 6 被験物質の異常な用量反応の原因を明らかにするために努力すること。

－ この情報は、LLNA と比較して rLLNA の精度を高める方法を見つけるために役立つ。

－ 市販後調査や職業上の暴露に関する情報も含め、LLNA において異常な用量反応を示すこれらの物質やその他の物質のモルモットやヒトにおける成績も一緒に含めて考えるよう努めること。

・ LLNA および rLLNA の試験で、動物の成績は個々に収集すること。

－ 個別に集めれば外れ値があった場合、棄却することができる。

・ 試験法の精度を低下させずに投与群当たりの動物数を減らすため、動物からの成績は個々に集められ、解析されるべきである。

－ ICCVAM が推奨する最新の LLNA 試験計画書には、そのような判断のために必要な統計解析手法が組み込まれている。

－ また、最新の試験計画書には、一緒に設定する陽性対照群の動物数を、その施設の陽性対照の背景データを評価することにより、減らすことができるか決めるためのガイダンスも盛り込まれている。

７．rLLNA の試験計画書

Day 1 – 3：投与毎に 4 匹のマウスの各々の両耳の背面に適当な溶媒を用いた被験物質（LLNA で試験される最高用量）或いは対照物質の 25 μL を塗布する。

↓

Day 4 – 5：無処置

↓

Day 6：各々の動物の尾静脈に 20 μCi ^{3}H-TdR 若しくは 2 μCi ^{125}I-iododeoxyuridine と 10^{-5} M fluorodeoxyuridine を投与する。

↓

5 時間後：耳介リンパ節を個別に摘出し、押しつぶして単離細胞懸濁液を調製する。

↓

単離細胞懸濁液をリン酸緩衝生理食塩水で 2 回洗浄し、4℃で 18 時間、5% trichloroacetic acid で DNA を沈殿させる。

↓

Trichloroacetic acid にペレットを再懸濁し、^{3}H 測定にはシンチレーション溶液を加え、^{125}I 測定には再懸濁したペレットをガンマー線測定チューブに加え、放射能測定に備える。

↓

放射能測定：対照群と処置群の平均 dpm（毎分崩壊数）を測定する。

取り込み量（Stimulation index: SI）は、次の式より算出；

処置群の平均 dpm／対照群の平均 dpm

SI $\geq$ 3：被験物質を感作性物質として分類

SI < 3：被験物質を非感作性物質として分類

8．まとめ

- rLLNA は、最新の ICCVAM が推奨する LLNA 試験計画書に準拠して行われるとき、皮膚感作性物質と非感作性物質を十分に区別することができると、ICCVAM は結論している。
- 化学物質や製品の ACD の潜在性を調べるために、LLNA を行う前に rLLNA を実施することを ICCVAM は勧めている。
- LLNA と比較して、rLLNA では使用動物数を各試験当たり最大 40%まで削減できる。
- rLLNA には 1.9% (6/318)の偽陰性の結果を示す可能性がある。もし偽陰性の結果が示唆されるならば、LLNA 或いは他の公認の皮膚感作性試験法による確認試験を考慮すべきである。
- 用量-反応情報が必要な場合は、用量段階を持つ LLNA において試験されなければならない。

9．rLLNA 編纂委員会としての結論

- rLLNA は、化学物質の皮膚感作性を評価することが可能な試験方法である。

- rLLNA の結果が２ ＜SI＜ ３の場合、その判定に当たっては、被験物質に関する情報（例えば、最高用量での反応低下の可能性、タンパク結合性、類似化合物における皮膚感作性など）に基づいて慎重に評価する必要がある。
- rLLNA の結果が SI ≥ ３であれば、これ以上の皮膚感作性試験は不要である。
- LLNA や GPMT により非感作性物質と評価された原薬（化学物質）で製造された製品の評価には、rLLNA のみの実施でよい。
- rLLNA の邦文名として「単群局所リンパ節試験」を提案する。

１０． 文献

ICCVAM 1999. The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemical/Compounds. NIH Publication No. 99-4494. Research Triangle Park, NC: National Institute of Environmental Health Sciences.
Available: http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf

ICCVAM 2008a. Recommended Performance Standards. Murine Local Lymph Node Assay. NIH Publication No. 07-4518. Research Triangle Park. NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/

ICCVAM 2008b. The Reduced Murine Local Lymph Node Assay: An Alternative Test Method Using Fewer Animals to Assess the Allergic Contact Dermatitis Potential of Chemicals and Products. Research Triangle Park. NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/

Kimber I, Deaman RJ, Betts CJ, Gerberick GF, Ryan CA, Kem PS, et al. 2006. The local lymph node assay and skin sensitization: a cut-down screen to reduce animal requirements? Contact Dermatitis. 54: 181-185

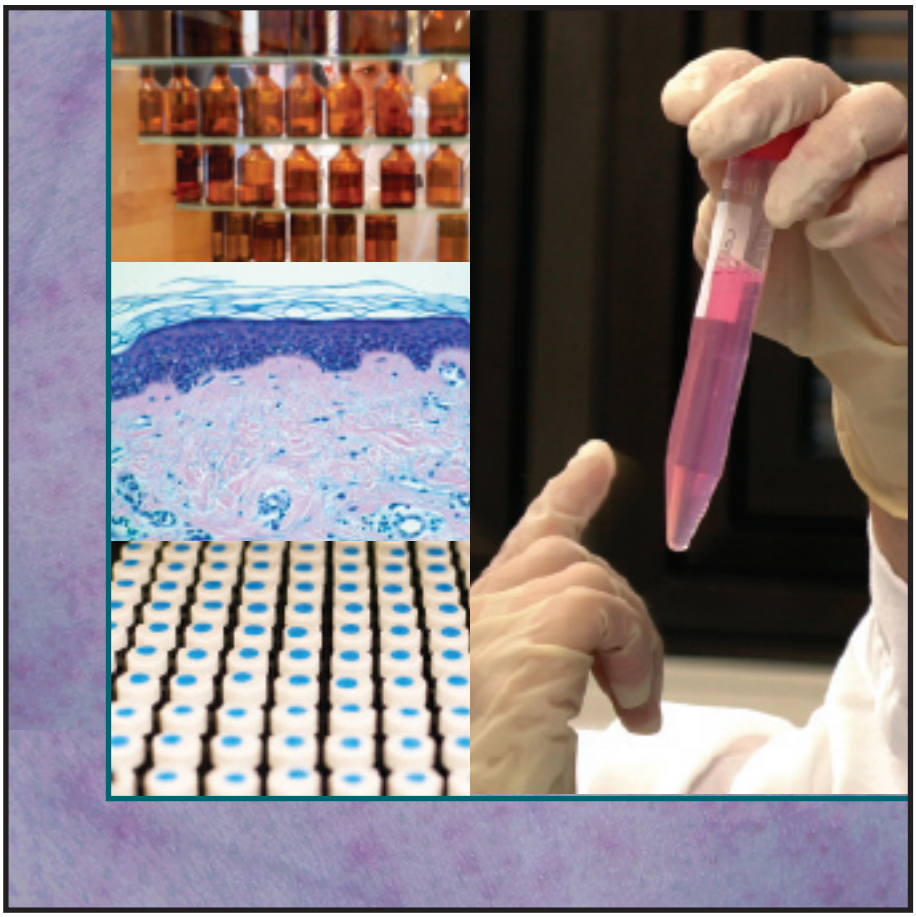
OECD 2002: Test guideline 429. Skin Sensitization: Local Lymph Node Assay, adopted April 24, 2002. In: OECD Guidelines for Testing of Chemicals. Paris: OECD. Available:
http://lysander.sourceoecd.org/vl=1706714/cl=11/nw=1/rpsv/cw/vhosts/oecdjournals/1607310x/v1n4/contp1-1.htm

OECD 2010: Test guideline 429. Skin Sensitization: Local Lymph Node Assay, adopted July 22, 2010. In: OECD Guidelines for Testing of Chemicals. Available:

http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/OECD/OECD-TG429-2010.pdf

Wind M, Matheson J, Jacobs A, Tice R, Stokes W: ICCVAM Test Method Recommendations for the Reduced LLNA (rLLNA): An Alternative Test Method Using Fewer Animals to Assess the Allergic Contact Dermatitis Potential of Chemicals and Products. Available: http://ICCVAM.niehs.nih.gov/meetings/SOT09/posters/Wind.pdf

NIH Publication Number 09-6439



ICCVAM TEST METHOD EVALUATION REPORT

The Reduced Murine Local Lymph Node Assay: An Alternative Test Method Using Fewer Animals to Assess the Allergic Contact Dermatitis Potential of Chemicals and Products

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Institute of Environmental Health Sciences
National Institutes of Health
U.S. Public Health Service
Department of Health and Human Services

About the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)
and
The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

In 1997, the National Institute of Environmental Health Sciences (NIEHS), one of the National Institutes of Health, established ICCVAM to:

- Coordinate interagency technical reviews of new and revised toxicological test methods, including alternative test methods that reduce, refine, or replace the use of animals
- Coordinate cross-agency issues relating to validation, acceptance, and national and international harmonization of new, modified, and alternative toxicological test methods

On December 19, 2000, the ICCVAM Authorization Act (Public Law 106-545, 42 U.S.C. 285*l*-3) established ICCVAM as a permanent interagency committee of NIEHS under NICEATM.

ICCVAM consists of representatives from 15 U.S. Federal regulatory and research agencies that use, generate, or disseminate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability. ICCVAM promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety or hazards of chemicals and products and that refine (decrease or eliminate pain and distress), reduce, and/or replace animal use. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities. More information about ICCVAM and NICEATM can be found on the NICEATM-ICCVAM website (http://iccvam.niehs.nih.gov) or obtained by contacting NICEATM (telephone: [919] 541-2384, e-mail: niceatm@niehs.nih.gov).

The following Federal regulatory and research agencies are ICCVAM members:

- Consumer Product Safety Commission
- Department of Agriculture
- Department of Defense
- Department of Energy
- Department of Health and Human Services
  - Centers for Disease Control and Prevention
    - Agency for Toxic Substances and Disease Registry
    - National Institute of Occupational Safety and Health
  - Food and Drug Administration
  - National Institutes of Health
    - Office of the Director
    - National Cancer Institute
    - National Institute of Environmental Health Sciences
    - National Library of Medicine
- Department of the Interior
- Department of Labor
  - Occupational Safety and Health Administration
- Department of Transportation
- Environmental Protection Agency

The NICEATM-ICCVAM graphic symbolizes the important role of new and alternative toxicological methods in protecting and advancing the health of people, animals, and our environment.

ICCVAM Test Method Evaluation Report

The Reduced Murine Local Lymph Node Assay: An Alternative Test Method Using Fewer Animals to Assess the Allergic Contact Dermatitis Potential of Chemicals and Products

Interagency Coordinating Committee on the Validation of Alternative Methods

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods

National Institute of Environmental Health Sciences
National Institutes of Health
U.S. Public Health Service
Department of Health and Human Services

2009
NIH Publication No. 09-6439

National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709

This document is available electronically at
http://iccvam.niehs.nih.gov/methods/immunotox/LLNA-LD/TMER.htm

When referencing this document, please cite as follows:

Interagency Coordinating Committee on the Validation of Alternative Methods. 2009. ICCVAM Test Method Evaluation Report. The Reduced Murine Local Lymph Node Assay: An Alternative Test Method Using Fewer Animals to Assess the Allergic Contact Dermatitis Potential of Chemicals and Products. NIH Publication Number 09-6439. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

Table of Contents

List of Tables

List of Abbreviations and Acronyms

ACD	Allergic contact dermatitis
AOO	Acetone: olive oil (4:1 by volume)
BAuA	Federal Institute for Occupational Safety and Health (Germany)
BRD	Background review document
CPSC	U.S. Consumer Product Safety Commission
DMSO	Dimethyl sulfoxide
DNCB	Dinitrochlorobenzene
dpm	Disintegrations per minute
EC	Emulsion concentrate
EC3	Estimated concentration needed to produce a stimulation index of 3
ECPA	European Crop Protection Association
ECVAM	European Centre for the Validation of Alternative Methods
EFfCI	European Federation for Cosmetic Ingredients
EPA	U.S. Environmental Protection Agency
EPR	Epoxy resin
ESAC	ECVAM Scientific Advisory Committee
FR	*Federal Register*
GP	Guinea pig
GPMT	Guinea Pig Maximization Test
^{3}H	Tritiated
HCA	Hexyl cinnamic aldehyde
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ILS	Integrated Laboratory Systems
^{125}IU	^{125}I-iododeoxyuridine
IWG	Immunotoxicity Working Group
JaCVAM	Japanese Center for the Validation of Alternative Methods
LLNA	Murine local lymph node assay
LNC	Lymph node cells
NC	Not calculated
NICEATM	National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
OECD	Organisation for Economic Co-operation and Development
PBS	Phosphate-buffered saline
rLLNA	Reduced murine local lymph node assay

SACATM	Scientific Advisory Committee on Alternative Toxicological Methods
SC	Suspension concentrate
SD	Standard deviation
SI	Stimulation index
TCA	Trichloroacetic acid
TG	Test Guideline
U.K.	United Kingdom
U.S.	United States
U.S.C.	United States Code

Interagency Coordinating Committee on the Validation of Alternative Methods: Agency Representatives

Agency for Toxic Substances and Disease Registry
• Moiz Mumtaz, Ph.D.

Consumer Product Safety Commission
• Marilyn L. Wind, Ph.D. (Chair)
◊ Kristina Hatlelid, Ph.D.
Joanna Matheson, Ph.D.

Department of Agriculture
• Jodie Kulpa-Eddy, D.V.M. (Vice-Chair)
◊ Elizabeth Goldentyer, D.V.M.

Department of Defense
• Robert E. Foster, Ph.D.
◊ Patty Decot
Peter J. Schultheiss, D.V.M., D.A.C.L.A.M.
Harry Salem, Ph.D.

Department of Energy
• Michael Kuperberg, Ph.D.
◊ Marvin Stodolsky, Ph.D.

Department of the Interior
• Barnett A. Rattner, Ph.D.
◊ Sarah Gerould, Ph.D.

Department of Transportation
• George Cushmac, Ph.D.
◊ Steve Hwang, Ph.D.

Environmental Protection Agency
Office of Science Coordination and Policy
• Karen Hamernik, Ph.D.
Office of Research and Development
◊ Julian Preston, Ph.D.
Suzanne McMaster, Ph.D.
Office of Pesticide Programs
Amy Rispin, Ph.D.
Deborah McCall
OECD Test Guidelines Program
Jerry Smrchek, Ph.D.

• Principal agency representative
◊ Alternate principal agency representative

Food and Drug Administration
Office of Science
• Suzanne Fitzpatrick, Ph.D., D.A.B.T.
Center for Drug Evaluation and Research
◊ Abigail C. Jacobs, Ph.D.
Paul C. Brown, Ph.D.
Center for Devices and Radiological Health
Melvin E. Stratmeyer, Ph.D.
Vasant G. Malshet, Ph.D., D.A.B.T.
Center for Biologics Evaluation and Research
Richard McFarland, Ph.D., M.D.
Ying Huang, Ph.D.
Center for Food Safety and Nutrition
David G. Hattan, Ph.D.
Robert L. Bronaugh, Ph.D.
Center for Veterinary Medicine
Devaraya Jagannath, Ph.D.
M. Cecilia Aguila, D.V.M.
National Center for Toxicological Research
William T. Allaben, Ph.D.
Paul Howard, Ph.D.
Office of Regulatory Affairs
Lawrence D'Hoostelaere, Ph.D.

National Cancer Institute
• Alan Poland, M.D.
◊ T. Kevin Howcroft, Ph.D.

National Institute of Environmental Health Sciences
• William S. Stokes, D.V.M., D.A.C.L.A.M
◊ Raymond R. Tice, Ph.D.
Rajendra S. Chhabra, Ph.D., D.A.B.T.
Jerrold J. Heindel, Ph.D.

National Institute for Occupational Safety and Health
• Paul Nicolaysen, V.M.D.
◊ K. Murali Rao, M.D., Ph.D.

National Institutes of Health
• Margaret D. Snyder, Ph.D.

National Library of Medicine
• Pertti (Bert) Hakkinen, Ph.D.
◊ Jeanne Goshorn, M.S.

Occupational Safety and Health Administration
• Surender Ahir, Ph.D.

[This Page Intentionally Left Blank]

Acknowledgements

The following individuals are acknowledged for their contributions to the evaluation of the reduced LLNA.

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Immunotoxicity Working Group (IWG)

U.S. Consumer Product Safety Commission
Joanna Matheson, Ph.D. (IWG Co-Chair)
Marilyn Wind, Ph.D.

U.S. Environmental Protection Agency
Office of Pesticide Programs
Masih Hashim, D.V.M., Ph.D.
Marianne Lewis
Deborah McCall
Timothy McMahon, Ph.D.
Amy Rispin, Ph.D.
Office of Prevention, Pesticides, and Toxic Substances
Ronald Ward, Ph.D.
Office of Research and Development
Marsha Ward, Ph.D.
Office of Science Coordination and Policy
Karen Hamernik, Ph.D.

U.S. Food and Drug Administration
Center for Devices and Radiological Health
Daniel Lyle, Ph.D.
Vasant G. Malshet, Ph.D., D.A.B.T.
Jeffrey Toy, Ph.D.
Center for Drug Evaluation and Research
Paul Brown, Ph.D.
Abigail Jacobs, Ph.D. (IWG Co-Chair)
Jiaqin Yao, Ph.D.
Center for Veterinary Medicine
Ruth Barratt, Ph.D., D.V.M.

National Institute of Environmental Health Sciences
Dori Germolec, Ph.D.
William S. Stokes, D.V.M., D.A.C.L.A.M.
Raymond R. Tice, Ph.D.

National Institute for Occupational Safety and Health
B. Jean Meade, D.V.M., Ph.D.

European Centre for the Validation of Alternative Methods – Liaison
Silvia Casati, Ph.D.

Japanese Center for the Validation of Alternative Methods – Liaison
Hajime Kojima, Ph.D.

Murine Local Lymph Node Assay
Independent Scientific Peer Review Panel
(March 4–6, 2008)

Michael Luster, Ph.D. (Panel Chair)
Senior Consultant to the National Institute for Occupational Safety and Health
Health Effects Laboratory
Morgantown, WV

Nathalie Alépée, Ph.D.
Associate Research Fellow
Pfizer PDRD
MCT Laboratory
Amboise, France

Anne Marie Api, Ph.D.
Vice President, Human Health Sciences
Research Institute for Fragrance Materials
Woodcliff Lake, NJ

Nancy Flournoy, M.S., Ph.D.
Professor and Chair
Department of Mathematics and Statistics
University of Missouri – Columbia
Columbia, MO

Thomas Gebel, Ph.D.
Regulatory Toxicologist
Federal Institute for Occupational Safety & Health
Dortmund, Germany

Sidney Green, Ph.D.*
Professor of Pharmacology
College of Medicine
Howard University
Washington, DC

Kim Headrick, B.Admin., B.Sc.
International Harmonization and Senior Policy Advisor
Policy and Programme Service Office
Health Canada
Ottawa, Ontario, Canada

Dagmar Jírová, M.D., Ph.D.
Toxicologist, Research Manager
Head of Reference Center for Cosmetics and Reference Laboratory for Experimental Immunotoxicology
National Institute of Public Health
Prague, Czech Republic

David Lovell, Ph.D.
Reader in Medical Statistics
Postgraduate Medical School
University of Surrey
Guildford, Surrey, U.K.

Howard Maibach, M.D.
Professor, Department of Dermatology
University of California – San Francisco
San Francisco, CA

James McDougal, Ph.D.
Professor and Director of Toxicology Research
Department of Pharmacology and Toxicology
Boonshoft School of Medicine
Wright State University
Dayton, OH

Michael Olson, Ph.D.
Director of Occupational Toxicology
Corporate Environment, Health and Safety
GlaxoSmithKline
Research Triangle Park, NC

Raymond Pieters, Ph.D.
Associate Professor
Immunotoxicology Group Leader
Institute for Risk Assessment Sciences
Utrecht University
Utrecht, The Netherlands

Jean Regal, Ph.D.
Professor, Department of Pharmacology
University of Minnesota Medical School
Duluth, MN

Jonathan Richmond, MB ChB, FRCSEd*
Head, Animal (Scientific) Procedures Division
Home Office
London, U.K.

Peter Theran, V.M.D.
Consultant
Massachusetts Society for the Prevention of Cruelty to Animals
Novato, CA

Stephen Ullrich, Ph.D.
Professor of Immunology
Graduate School of Biomedical Sciences
University of Texas
M.D. Anderson Cancer Center – Houston
Houston, TX

Michael Woolhiser, Ph.D.
Technical Leader – Immunotoxicology
Senior Toxicology Specialist
Dow Chemical
Midland, MI

Takahiko Yoshida, M.D., Ph.D.
Professor, Department of Health Science
Asahikawa Medical College
Hokkaido, Japan

* Drs. Green and Richmond were unable to attend the public meeting on March 4-6, 2008. However, they were involved in the peer review of the background review documents and concur with the conclusions and recommendations included in the *Independent Scientific Peer Review Panel Report - Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products.*

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Institute of Environmental Health Sciences

William Stokes, D.V.M., D.A.C.L.A.M.
Director; Project Officer

Raymond Tice, Ph.D.
Deputy Director

Deborah McCarley
Special Assistant; Assistant Project Officer

NICEATM Support Contract Staff (Integrated Laboratory Systems [ILS], Inc.)

David Allen, Ph.D.
Thomas Burns, M.S.
Linda Litchfield
Gregory Moyer, M.B.A.
Michael Paris
Eleni Salicru, Ph.D.
Catherine Sprankle
Judy Strickland, Ph.D., D.A.B.T.
Linda Wilson

Statistical Consultant for ILS, Inc.
Joseph Haseman, Ph.D.

Other Acknowledgements

ICCVAM and NICEATM gratefully acknowledge the following individuals and institutions that submitted data to NICEATM for the evaluation of the reduced LLNA.

David Basketter, Ph.D.
Unilever Safety and Environmental Assurance Centre
Sharnbrook, U.K.

Phil Botham, Ph.D.
European Crop Protection Association
Brussels, Belgium

Eric Debruyne, Ph.D.
Bayer CropScience SA, Sophia Antipolis
Cedex, France

G. Frank Gerberick, Ph.D.
Procter and Gamble Company
Cincinnati, OH

Dori Germolec, Ph.D.
National Toxicology Program
Research Triangle Park, NC

Ian Kimber, Ph.D.
Syngenta Central Toxicology Laboratory
Macclesfield, U.K.

Michael J. Olson, Ph.D.
GlaxoSmithKline
Research Triangle Park, NC

K. Skirda
TNO Quality of Life
Delft, Netherlands

Peter Ungeheuer, Ph.D.
European Federation for Cosmetic Ingredients
Frankfurt, Germany

Hans Werner Vohr, Ph.D.
Bayer HealthCare
Wuppertal-Elberfeld, Germany

[This Page Intentionally Left Blank]

Preface

Allergic contact dermatitis (ACD) is an adverse health effect that frequently develops in workers and consumers exposed to skin-sensitizing chemicals and products. ACD results in lost workdays[1] and can significantly diminish quality of life (Hutchings et al. 2001; Skoet et al. 2003). To minimize the occurrence of ACD, regulatory authorities require testing to identify substances that may cause it. Sensitizing substances must be labeled with a description of the potential hazard and the precautions necessary to avoid development of ACD.

Skin sensitization testing has typically required the use of guinea pigs (Buehler 1965; Magnusson and Kligman 1970). However, in 1998, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) evaluated an alternative known as the murine (mouse) local lymph node assay (traditional LLNA[2]). ICCVAM concluded that the traditional LLNA provided several advantages over the commonly accepted guinea pig test methods, including elimination of potential pain and distress, use of fewer animals, less time to perform, and availability of dose-response information. ICCVAM recommended the LLNA as an alternative test method for assessing the skin sensitization potential of most types of substances. United States and international regulatory agencies subsequently accepted the traditional LLNA as a valid alternative test method for ACD testing.

In 2007, the U.S. Consumer Product Safety Commission requested that ICCVAM evaluate several modifications of the traditional LLNA, [3] including the "reduced LLNA" (rLLNA), also referred to as the "cut-down" or "limit dose" LLNA. ICCVAM assigned this activity a high priority after considering comments from the public and ICCVAM's Scientific Advisory Committee on Alternative Toxicological Methods (SACATM). As part of their ongoing collaboration with ICCVAM, scientists from the European Centre for Validation of Alternative Methods and the Japanese Center for Validation of Alternative Methods served as liaisons to the ICCVAM Immunotoxicity Working Group (IWG). A detailed timeline of the rLLNA test method evaluation is included with this report.

This Test Method Evaluation Report provides ICCVAM's recommendations regarding the usefulness and limitations of the rLLNA for assessing the ACD potential of substances. When deemed appropriate for use, the rLLNA can reduce by 40% the number of animals used for each test compared to the traditional LLNA. The report also provides the updated ICCVAM-recommended LLNA test method protocol, which addresses the rLLNA procedure. The database of substances used to validate the rLLNA is discussed and summarized.

ICCVAM carefully compiled and assessed all available data and arranged an independent scientific peer review. ICCVAM and the IWG solicited and considered public comments and stakeholder involvement throughout the rLLNA evaluation process. The National Toxicology Program Interagency Center for the Evaluation of Alternative Methods (NICEATM), ICCVAM, and the IWG began the process by preparing a draft background review document (BRD) describing the validation status of the rLLNA test method, including its reliability and

---

[1] http://www.bls.gov/IIF

[2] The "traditional LLNA" refers to the validated ICCVAM-recommended LLNA test method protocol, which measures lymphocyte proliferation based on incorporation of tritiated thymidine into the cells of the draining auricular lymph nodes.

[3] Available at http://iccvam.niehs.nih.gov/methods/immunotox/llnadocs/CPSC_LLNA_nom.pdf

accuracy for the substances evaluated, and draft test method recommendations for usefulness and limitations. ICCVAM released these documents to the public for comment on January 8, 2008, at which time ICCVAM also announced a meeting of the international independent scientific peer review panel (Panel) (*Federal Register* 73 FR 1360[4]).

The Panel met in public session on March 4–6, 2008, to review the ICCVAM draft BRD for completeness and accuracy. The Panel then evaluated (1) the extent to which the draft BRD addressed established validation and acceptance criteria and (2) the extent to which the BRD supported ICCVAM's draft test method recommendations. Before concluding their deliberations, the Panel considered written comments and comments made at the meeting by public stakeholders. The final Panel report was made available to the public for comment on May 20, 2008.[5]

ICCVAM provided SACATM with the draft BRD and draft Test Method Evaluation Report, the Panel report, and all public comments for discussion at their meeting on June 18-19, 2008, where public stakeholders were given another opportunity to comment.

After SACATM's meeting, ICCVAM considered the SACATM comments, the Panel report, and all public comments before finalizing the ICCVAM Test Method Evaluation Report and Background Review Document, which is provided as an appendix to this report. The consolidated document will be provided to U.S. Federal regulatory agencies for consideration and be made available to the public. The ICCVAM Authorization Act requires that Federal agencies respond to ICCVAM within 180 days after receiving the ICCVAM test method recommendations. Agency responses will be posted on the NICEATM–ICCVAM website[6] as they become available.

We gratefully acknowledge the many individuals who contributed to the preparation, review, and revision of this report. We especially recognize the Panel members for their thoughtful evaluations and generous contributions of time and effort. Special thanks are extended to Dr. Michael Luster for serving as the Panel Chair and to Dr. Michael Woolhiser, Dr. Michael Olson, and Ms. Kim Headrick for their service as Evaluation Group Chairs. We thank the IWG for assuring a meaningful and comprehensive review. We especially thank Dr. Joanna Matheson (Consumer Product Safety Commission) and Dr. Abigail Jacobs (U.S. Food and Drug Administration Center for Drug Evaluation and Research) for serving as Co-Chairs of the IWG. Integrated Laboratory Systems, Inc., the NICEATM support contractor, provided excellent scientific and operational support, for which we thank Dr. David Allen, Mr. Thomas Burns, Ms. Linda Litchfield, Mr. Michael Paris, Dr. Eleni Salicru, Ms. Catherine Sprankle, Dr. Judy Strickland, and Ms. Linda Wilson; and Dr. Joseph Haseman, ILS consulting statistician, for statistical support. We also acknowledge Dr. Raymond Tice, Deputy Director of NICEATM, for his efforts. Finally, we thank Dr. Silvia Casati and Dr. Hajime Kojima, the IWG liaisons from the European Centre for the Validation of Alternative Methods and the Japanese Center for the Validation of Alternative Methods, respectively, for their participation and contributions.

[4] Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_25553.pdf

[5] Announced in 73 FR 29136 (http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR-E8-11195.pdf); available at http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

[6] http://iccvam.niehs.nih.gov/methods/immunotox/rLLNA.htm

This comprehensive ICCVAM evaluation of the rLLNA should facilitate regulatory agency decisions on the acceptability of the method. Following regulatory acceptance, use of the method by industry can be expected to significantly reduce the number of animals required for ACD testing while continuing to support the protection of human health.

William S. Stokes, D.V.M., D.A.C.L.A.M
Rear Admiral/Assistant Surgeon General, U.S. Public Health Service
Director, NICEATM
Executive Director, ICCVAM

Marilyn Wind, Ph.D.
Deputy Associate Executive Director
Directorate for Health Sciences
U.S. Consumer Product Safety Commission
Chair, ICCVAM

[This Page Intentionally Left Blank]

Executive Summary

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) evaluated the validation status of the reduced murine local lymph node sssay (rLLNA), a test method for assessing the potential of substances to cause allergic contact dermatitis (ACD). ACD is an allergic skin reaction characterized by redness, swelling, and itching that can result from contact with a sensitizing chemical or product. This Test Method Evaluation Report provides ICCVAM's recommendations regarding the usefulness and limitations of the rLLNA as an alternative to the traditional murine local lymph node assay (LLNA). When deemed appropriate for use, the rLLNA can reduce by 40% the number of animals used for each test compared to the traditional LLNA. This report also includes the updated ICCVAM-recommended LLNA test method protocol, the final rLLNA background review document (BRD), and recommendations for future studies and performance standards.

The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), ICCVAM, and the ICCVAM Immunotoxicity Working Group prepared a draft BRD and draft test method recommendations, which were provided to an international independent scientific peer review panel (Panel) and the public for comment. The BRD evaluated data from 471 traditional LLNA studies, including the 211 substances from the 1998 ICCVAM evaluation of the traditional LLNA (ICCVAM 1999), and 246 from the peer-reviewed literature and submissions to NICEATM in response to a May 17, 2007, *Federal Register* request for comments (72 FR 27815[7]). A detailed timeline of the rLLNA test method evaluation is included with this report.

The Panel met in public session on March 4–6, 2008, to discuss their peer review of the ICCVAM draft BRD and to provide conclusions and recommendations on the current validation status of the rLLNA test method. The Panel also reviewed how well the information contained in the draft BRD supported ICCVAM's draft test method recommendations. In finalizing this Test Method Evaluation Report and the BRD, which is included as an appendix, ICCVAM considered the conclusions and recommendations of the Panel and comments from ICCVAM's Scientific Advisory Committee on Alternative Toxicological Methods and the public.

ICCVAM Recommendations: Test Method Usefulness and Limitations

ICCVAM concludes that the scientific validity of the rLLNA has been adequately evaluated and that the performance of the rLLNA, when conducted in accordance with the updated ICCVAM-recommended LLNA protocol, is sufficient to distinguish between skin sensitizers and non-sensitizers in cases that do not require dose-response information. ICCVAM also concludes that, compared to the traditional LLNA, the rLLNA will reduce animal use by 40% for each test. Accordingly, ICCVAM recommends that the rLLNA test method should be used routinely to determine the ACD potential of chemicals and products before conducting the traditional LLNA. Negative substances can be classified as non-sensitizers, and positive substances can be classified as sensitizers.

In cases that require dose-response information, positive substances must be tested in the traditional multiple-dose LLNA. Therefore, if dose-response information is required for a

[7] Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf

substance that, after consideration of all available information, is also suspected of having the potential to produce ACD, it should be evaluated initially using the traditional LLNA.

There is a small possibility of a false negative result (1.9% [6/318]) in the rLLNA compared to the traditional LLNA. This information should be considered when evaluating results from the rLLNA, and negative results should always prompt a weight-of-evidence evaluation of supplemental information (e.g., possibility of downturn in response at the high dose, test results with similar substances, peptide-binding activity, molecular weight, other testing data). If false negative results are suggested, confirmatory testing in the traditional LLNA or another accepted skin sensitization test method should be considered.

ICCVAM Recommendations: Test Method Protocol

The updated LLNA test method protocol recommended by ICCVAM is included as an appendix to this report. In the traditional LLNA, at least three dose levels of each test substance are evaluated. The rLLNA evaluates only the highest dose of the test substance along with the concurrent vehicle- and positive-control groups. ICCVAM recommends testing only the highest concentration, defined as the maximum soluble concentration that does not induce excessive local irritation and/or overt systemic toxicity.

ICCVAM recommends that individual animal data should be collected in order to permit identification and exclusion of outlier values that could cause false negative or false positive results. Collection of individual animal data (versus pooled) also allows for statistical analysis to determine whether the test-substance response is significantly different from that of the vehicle control.

The ICCVAM-recommended LLNA test method protocol has been revised to require a minimum of four animals per dose group. Data analysis indicated that reducing dose groups from five animals to four is unlikely to significantly affect the results of an LLNA study. Organisation for Economic Co-operation and Development (OECD) Test Guideline (TG) 429 for the LLNA currently requires at least five animals per dose group if individual animal data are collected but only four animals in each dose group if lymph nodes from all animals in the group are pooled into one sample for data collection (OECD 2002). To determine if these requirements could be harmonized without diminishing accuracy, NICEATM evaluated data from 83 LLNA studies (275 dose groups) from six different laboratories. This revision is important because many national regulations and policies require that the minimum number of animals be used for studies. Therefore, once TG 429 is updated with the revision, the collection of individual animal data will be consistent with this requirement.

ICCVAM also recommends including a positive-control substance with each test to ensure that all protocol procedures are conducted properly and all aspects of the test system work properly such that they can produce a positive response. However, similar to OECD TG 429, the updated ICCVAM-recommended LLNA test method protocol states that laboratories that conduct the LLNA at least once per month and that have a history of and a documented proficiency for obtaining consistent results with positive controls may consider testing positive control substances at intervals of no more than six months.

ICCVAM Recommendations: Future Studies

ICCVAM recommends additional studies to further characterize and potentially improve the usefulness and applicability of the rLLNA for identifying potential skin sensitizers.

- Additional efforts should be made to understand the basis for abnormal dose responses for six substances in this evaluation that would have resulted in false negative results using the rLLNA compared to the traditional LLNA. This information should help identify ways to improve the accuracy of the rLLNA compared to the traditional LLNA. Efforts should also be made to identify data from guinea pigs and humans for substances that exhibit abnormal dose responses in the traditional LLNA. Information from post-marketing surveillance and/or occupational exposures should be collected and assessed.
- All future traditional LLNA and rLLNA studies should collect individual animal data. This will allow detection of outliers and avoidance of false negative results that can occur from pooling data that include one or more abnormally low values. Existing LLNA studies using data pooled from all animals in a dose group, such as four of the six false negative rLLNA results in this evaluation, should be evaluated further with data obtained from individual animals within each dose group to determine if pooling of data may have led to false negative outcomes.
- Data from individual animals should be collected and analyzed to identify opportunities to use fewer animals per dose group without compromising test method accuracy. The updated ICCVAM-recommended LLNA test method protocol incorporates statistical procedures necessary for such determinations. This includes evaluating the laboratory's historical positive-control database to determine if the number of animals in the concurrent positive-control group can be reduced.

ICCVAM Recommendations: Performance Standards

The ICCVAM-recommended test method performance standards for the traditional LLNA[8] may be used to evaluate the performance of modified test methods, including the rLLNA, that are functionally and mechanistically similar to the traditional LLNA. Modified protocols for the rLLNA that adhere to the traditional LLNA performance standards would be considered acceptable for hazard identification purposes.

Validation Status of the rLLNA Test Method

ICCVAM (1999) compared the accuracy and reliability of traditional LLNA results to results from guinea pig tests (EPA 2003) and results obtained from the human maximization test and sensitizing substances included in human patch test allergen panels. ICCVAM concluded that the LLNA was a valid alternative to currently accepted guinea pig test methods for most testing situations and that the LLNA reduces the number of animals required for testing while also refining the procedure by eliminating animal pain and distress. The LLNA was subsequently accepted by U.S. regulatory agencies as an alternative to the guinea pig tests (e.g., Guinea Pig Maximization Test and Buehler Test) for assessing the potential of substances to cause ACD.

[8] Available at http://iccvam.niehs.nih.gov/methods/immunotox/PerfStds/llna-ps.htm

The only difference between the test method protocols for the traditional LLNA and the rLLNA is the number of dose levels tested for a test substance. In the traditional LLNA, at least three dose levels are tested for each substance, with the highest dose based on maximum solubility and the avoidance of excessive local irritation and/or systemic toxicity. In contrast, only the highest dose of a substance is tested in the rLLNA (Kimber et al. 2006). Because the criteria for choosing the highest dose in the traditional LLNA and in the rLLNA are the same, the maximum dose level tested in the traditional LLNA and that tested in the rLLNA should be the same. Thus, the accuracy and reliability of the rLLNA test method should be similar for the same substances tested in the traditional LLNA, although the accuracy was slightly different based on available data described below.

Accuracy and Reliability of the rLLNA

The accuracy of the rLLNA for identifying potential skin sensitizers was compared to that of the traditional LLNA. In the 471 traditional LLNA studies, 318 results were positive and 153 were negative. When studies in which substances were tested more than once in the same vehicle were combined to yield an overall skin sensitization classification, 465 studies with unique combinations of substances and vehicles were evaluated, with 315 classified as sensitizers and 150 classified as non-sensitizers.

As shown in **Table 1**, compared to the traditional LLNA, the rLLNA has an accuracy of 98.7% (465/471), a sensitivity of 98.1% (312/318), a specificity of 100% (153/153), a false positive rate of 0% (0/153), and a false negative rate of 1.9% (6/318). When only unique combinations of substances and vehicles are considered, the rLLNA has an accuracy of 98.7% (459/465), a sensitivity of 98.1% (309/315), a specificity of 100% (150/150), a false positive rate of 0% (0/150), and a false negative rate of 1.9% (6/315).

Table 1 Performance of the rLLNA in Predicting Skin Sensitizers Compared to the Traditional LLNA

Data	N	Accuracy	Sensitivity	Specificity	False Positive	False Negative
Kimber et al. (2006)	211	98.6% (208/211)	98.2% (166/169)	100% (42/42)	0% (0/42)	1.8% (3/169)
rLLNA	471	98.7% (465/471)	98.1% (312/318)	100% (153/153)	0% (0/153)	1.9% (6/318)
rLLNA (substances repeated in the same vehicle considered together)	465	98.7% (459/465)	98.1% (309/315)	100% (150/150)	0% (0/150)	1.9% (6/315)

Abbreviation: N = number of tests
Accuracy = the percentage of correct outcomes (positive and negative) of a test method
Sensitivity = the percentage of all positive substances that are classified as positive
Specificity = the percentage of all negative substances that are classified as negative
False positive rate = the percentage of all negative substances that are falsely identified as positive
False negative rate = the percentage of all positive substances that are falsely identified as negative

Interlaboratory reproducibility of the rLLNA was assessed with traditional LLNA data for five substances tested independently in the same vehicle at two or three laboratories: dinitrochlorobenzene (DNCB), hexyl cinnamic aldehyde (HCA), linalool alcohol, methyl

salicylate, and potassium dichromate. All studies classified DNCB, methyl salicylate, and potassium dichromate as sensitizers or non-sensitizers (i.e., 100% concordance). HCA and linalool alcohol, which were tested independently in two laboratories, were classified as sensitizers by one traditional LLNA study and as non-sensitizers by the other study. Review of these two studies indicates that the discordant results were due to differences in the highest dose levels tested. However, because the rLLNA and traditional LLNA use identical protocols and the data sets used to evaluate their accuracy are similar, the intra- and interlaboratory reliability of the rLLNA is deemed to be similar to that of the traditional LLNA.

[This Page Intentionally Left Blank]

1.0 Introduction

The murine local lymph node assay (traditional LLNA[9]) is an alternative skin sensitization test method that requires fewer animals and less time than currently accepted guinea pig tests (e.g., the Guinea Pig Maximization Test and the Buehler Test). It can also eliminate animal pain and distress. The LLNA measures cell proliferation in the draining auricular lymph nodes of the mouse by analyzing incorporation of a radioactive marker into newly synthesized DNA. The LLNA was the first alternative test method evaluated and recommended by the U.S. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). International regulatory authorities have now recognized the traditional LLNA as an acceptable alternative to guinea pig tests for most testing situations.

The reduced murine local lymph node assay (rLLNA), also referred to as the "cut-down" or "limit dose" LLNA, was one of several modified versions of the LLNA nominated by the U.S. Consumer Product Safety Commission (CPSC) for evaluation by ICCVAM and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM).[10] (The term "reduced LLNA" has been adopted in this document to be consistent with the terminology used for this test method in Europe.) The proposed rLLNA could reduce the number of animals used for skin sensitization testing by 40% for each test compared to the traditional LLNA.

The ICCVAM Authorization Act of 2000 (Public Law 106-545, 42 United States Code 285*l*-3) charged ICCVAM with coordinating the technical evaluations of new, revised, and alternative test methods with regulatory applicability. After considering comments from the public and ICCVAM's advisory committee, the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM), ICCVAM members unanimously agreed that the rLLNA should have a high priority for evaluation. A detailed timeline of the rLLNA test method evaluation is provided in **Appendix A**. The updated ICCVAM-recommended LLNA test method protocol, accompanying statistical evaluation, and final rLLNA background review document (BRD) are provided in **Appendices B, C, and D**, respectively.

The ICCVAM Immunotoxicity Working Group (IWG) was formed to work with NICEATM in evaluating the test methods. Dr. Silvia Casati was the European Centre for the Validation of Alternative Methods (ECVAM) liaison, and Dr. Hajime Kojima was the Japanese Center for the Validation of Alternative Methods (JaCVAM) liaison to the IWG.

To facilitate peer review of the validation status of the rLLNA, the IWG and NICEATM, which administers ICCVAM and provides scientific support for ICCVAM activities, prepared a comprehensive BRD that provided information and data from validation studies and scientific literature. A May 17, 2007, *Federal Register* (*FR*) notice (72 FR 27815[11]) requested data and information on these test methods and nominations of individuals to serve on an international independent scientific peer review panel (Panel). The request was also disseminated via the ICCVAM electronic mailing list and through direct requests to over 100 stakeholders. Eight

[9] The "traditional LLNA" refers to the validated ICCVAM-recommended LLNA protocol (ICCVAM 1999; Dean et al. 2001), which measures lymphocyte proliferation based on incorporation of tritiated thymidine into the cells of the draining auricular lymph nodes.

[10] Available at http://iccvam.niehs.nih.gov/methods/immunotox/llnadocs/CPSC_LLNA_nom.pdf

[11] Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf

individuals submitted data and three individuals or organizations nominated members to the Panel.

ICCVAM examined data from 471 traditional LLNA studies (318 sensitizers and 153 non-sensitizers) representing 457 unique substances. ICCVAM built on a recent assessment of this procedure by the ECVAM Scientific Advisory Committee (ESAC; ESAC 2007), which used data from 211 traditional LLNA studies (211 unique substances) (Kimber et al. 2006). In an April 2007 statement, ESAC concluded "that the peer reviewed and published information is of a quality and nature to support the use of the rLLNA within tiered-testing strategies to reliably distinguish between chemicals that are skin sensitisers and non-sensitisers…" (**Appendix E**)

On January 8, 2008, ICCVAM announced the availability of the ICCVAM draft BRD and a public Panel meeting to review the validation status of the rLLNA (and other modifications to the traditional LLNA) (73 FR 1360[12]). The ICCVAM draft BRD and draft test method recommendations were posted on the NICEATM–ICCVAM website.[13] All of the information provided to the Panel and all public comments received prior to the Panel meeting were made available on the NICEATM–ICCVAM website.

The Panel met in public session on March 4–6, 2008, to review the rLLNA's validation status and the completeness and accuracy of the ICCVAM draft BRD. The Panel evaluated (1) the extent to which the draft BRD addressed established validation and acceptance criteria and (2) the extent to which the BRD supported ICCVAM's draft proposed test method uses, recommended protocols, draft test method performance standards, and proposed future studies. Interested stakeholders from the public were provided opportunities to comment at the Panel meeting. The Panel considered these comments as well as those submitted prior to the meeting before concluding their deliberations. On May 20, 2008, ICCVAM posted a report of the Panel's recommendations[14] (see **Appendix F**) on the NICEATM–ICCVAM website for public review and comment (announced in 73 FR 29136[15]).

ICCVAM provided SACATM with the draft BRD and draft test method recommendations, the Panel report, and all public comments for discussion at their meeting on June 18–19, 2008, where public stakeholders were given another opportunity to comment.

ICCVAM and the IWG considered the SACATM comments, the Panel report, and all public comments when finalizing the test method recommendations provided in this report. As required by the ICCVAM Authorization Act, ICCVAM will make this Test Method Evaluation Report and the accompanying final BRD available to the public and to U.S. Federal agencies for consideration. Federal agencies must respond to ICCVAM within 180 days after receiving ICCVAM test method recommendations. Agency responses will be made available to the public on the NICEATM–ICCVAM website as they are received.

[12] Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_25553.pdf
[13] http://iccvam.niehs.nih.gov/methods/immunotox/llna-panelDocs.htm
[14] Available at http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf
[15] Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR-E8-11195.pdf

2.0 ICCVAM Recommendations for the rLLNA Test Method

ICCVAM evaluated the validation status of the rLLNA test method as a reduction alternative to the traditional LLNA. The rLLNA should be used for the hazard identification of skin-sensitizing substances if dose-response information is not needed (e.g., for a compound presumed to be a strong sensitizer), provided there is adherence to all other LLNA protocol specifications as described in the updated ICCVAM-recommended LLNA test method protocol (available in **Appendix B** and at the NICEATM–ICCVAM website[16]). To further reduce animal use, the rLLNA should be used routinely as an initial test to determine allergic contact dermatitis (ACD) potential of chemicals and products before conducting the traditional LLNA. Negative substances can be classified as non-sensitizers, and positive substances can be classified as sensitizers.

Where dose-response information is required (e.g., for a compound presumed to be a weak or borderline sensitizer), positive substances must be tested in the traditional multidose LLNA. Accordingly, those substances for which dose-response information will be required and that are also suspected of having allergic contact dermatitis potential following consideration of all available information should be initially evaluated using the traditional LLNA.

2.1 ICCVAM Recommendations: Test Method Usefulness and Limitations

NICEATM and ICCVAM conducted a retrospective evaluation of rLLNA data to determine the test method's ability to distinguish between skin sensitizers and non-sensitizers. The performance assessment for the 465 unique substance and vehicle combinations evaluated in the study is provided in **Section 3.0**. Based on a review of the available data and comparison with the traditional LLNA, the scientific validity of the rLLNA has been adequately evaluated. ICCVAM concluded that, when conducted in accordance with the updated ICCVAM-recommended LLNA test method protocol specifications included in **Appendix B,** the rLLNA's performance is sufficient to distinguish between skin sensitizers and non-sensitizers when dose-response information is not required. This recommendation is based on its performance compared to that of the traditional LLNA. ICCVAM also concludes that use of the rLLNA can reduce by 40% the number of animals used for each test.

There is a small possibility of a false negative result (1.9% [6/318]) when compared to the traditional LLNA. This information should be considered when evaluating results from the rLLNA, and negative results should always be subjected to a weight-of-evidence evaluation of supplemental information (e.g., possibility of downturn in response at the high dose, test results with similar substances, peptide-binding activity, molecular weight, other testing data). If false negative results are suggested, confirmatory testing in the traditional LLNA or another accepted skin sensitization test method should be considered.

All of the testing limitations that apply to the traditional LLNA apply to the rLLNA also. For example, the rLLNA may not be suitable for use with certain types of test substances, such as nickel salts, mixtures, high-molecular weight compounds that cannot penetrate the stratum corneum, strong dermal irritants, or chemicals whose pharmacodynamic activity is to release dermal cytokines that cause local lymph node proliferation (e.g., certain pharmaceuticals such

[16] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAprotocol2008.pdf

as imiquimod [Gaspari 2007]). Additionally, the rLLNA may not be suitable for test substances that do not adhere for an acceptable period of time when applied to the dorsum of the ear.

Independent Peer Review Panel Conclusions and Recommendations

The Panel agreed that the available data support ICCVAM's draft recommendation that the rLLNA should be routinely recommended for hazard identification when dose-response information is not required. The Panel also agreed that to further reduce animal use the rLLNA should be routinely recommended as the initial test to identify sensitizers even if dose-response information *is* required, because negative results would not require additional testing. This is applicable in the occupational and public health setting in which obtaining hazard information is of critical importance. Subsequent traditional LLNA testing of substances that were positive in the rLLNA will provide dose-response information to assure detection of hazardous substances and allow potency estimates. The benefits of screening out the negatives, which do not require dose-response information, are clear; however, the animal welfare gains will depend on the proportion of test substances in any class that turn out to be non-sensitizers. The possible consequences of delays from another round of testing of those materials identified as sensitizers should also be considered.

The Panel agreed that the draft test method recommendations adequately addressed the low false negative rate by giving cautionary and weight-of-evidence consideration to the negative substances (and any possible false positive results). Furthermore, the Panel concluded that interspecies differences between the animal model and humans would probably make the false negative rate unimportant.

2.2 ICCVAM Recommendations: Test Method Protocol

ICCVAM recommends basing the protocol for rLLNA testing on the updated ICCVAM-recommended LLNA protocol, which addresses the rLLNA procedure (**Appendix B)**. The only difference between the traditional LLNA and the rLLNA test methods is that the middle- and low-dose groups are omitted in the rLLNA. On the basis of Panel comments, ICCVAM updated the traditional LLNA test method protocol to provide guidance on identifying the appropriate maximum dose for testing. In the rLLNA, in addition to the concurrent vehicle and positive-control groups, each test substance is tested at only one dose level (the high dose), whereas in the traditional LLNA each test substance is tested at a minimum of three dose levels. The test substance concentration should be the highest soluble concentration that does not induce overt systemic toxicity and/or excessive local irritation. Any other approach, such as one based on a pre-established threshold dose level, is inappropriate. For example, Kimber et al. (2006) proposed a 10% threshold concentration at which all negative results would be considered valid. However, 51 (16% [51/315]) of the test substances evaluated were non-sensitizers at concentrations of at least 10%[17] but were sensitizers at higher concentrations.

In the traditional LLNA test method protocol, a stimulation index (SI) is calculated as the ratio of the mean incorporation of ^{3}H-thymidine or ^{125}I-iododeoxyuridine by the auricular lymph nodes of the treated animals and that of the vehicle control animals. In the rLLNA, as in the traditional LLNA, the threshold for classifying a substance as a skin sensitizer is an SI $\geq$ 3.

[17] An initial dose was tested at 10% or greater and resulted in a stimulation index (SI) < 3, while a subsequent higher dose resulted in an SI $\geq$ 3.

In the updated LLNA test method protocol (**Appendix B**), ICCVAM recommends collecting individual animal data in order to allow identification and exclusion of outlier values that could result in false negative or false positive results. This is especially important to help avoid false negative results for weaker sensitizers (i.e., substances that induce an SI just above 3). The U.S. Environmental Protection Agency (EPA) Health Effects Test Guideline 870.2600 (EPA 2003) also requires the collection of individual animal data for the assessment of interanimal variability and a statistical comparison of test- and control-group measurements. While the Organisation for Economic Co-operation and Development (OECD) Test Guideline (TG) 429 (OECD 2002) allows for both the collection of individual animal measurements and the pooling of the lymph nodes for each treatment group, the latter eliminates any measure of interanimal variability and/or identification of outlier values, as well as statistical identification of a positive/negative response.

OECD TG 429 requires that each dose group consist of at least four animals if pooled animal data are collected and a minimum of five animals if individual animal data are collected (OECD 2002). To determine if the required number of animals for individual animal data collection could be the same as the required number for pooled data without diminishing accuracy, NICEATM evaluated data from 83 LLNA studies (275 dose groups) from six different laboratories **(Appendix C)**. This is important because most animal-use regulations require that the minimum number of animals be used in studies, which currently results in many countries collecting only pooled data because doing so requires fewer animals. This evaluation indicated that a reduction in the sample size from five to four animals per group is unlikely to have a significant impact on the results of an LLNA study; therefore, the ICCVAM-recommended LLNA test method protocol (**Appendix B**) was revised to require a minimum of four animals per dose group.

The updated ICCVAM-recommended LLNA test method protocol (**Appendix B**) also recommends that each test include a concurrent positive-control substance. Use of a positive-control substance can ensure that all protocol procedures are conducted properly and that all aspects of the test system work properly such that they produce a positive response. However, similar to OECD TG 429 (OECD 2002), the updated ICCVAM-recommended test method protocol states that testing of the positive-control substance at intervals of no more than six months may be considered in laboratories that conduct the LLNA at least once per month and that have a history and a documented proficiency for obtaining consistent results with positive controls.

Users should be aware that the decision to include a positive control only periodically instead of concurrently could affect the adequacy and acceptability of negative study results generated without a concurrent positive control. For example, if a false negative result is obtained in the periodic positive-control test, all negative test-substance results obtained since the last acceptable periodic positive-control test and the unacceptable periodic positive-control test could be questioned. In order to demonstrate that the prior negative test-substance results are acceptable, a laboratory could be expected to repeat all negative tests, which would require additional expense and increased animal use.

Independent Peer Review Panel Conclusions and Recommendations

The Panel agreed with ICCVAM's draft test method recommendations and recommended adherence to the ICCVAM-recommended LLNA protocol (with modifications omitting the

middle- and low-dose groups) for future rLLNA testing. The Panel also advised collecting individual animal data for future studies because it would allow an estimate of interanimal variability and conducting a statistical analysis to determine if the test substance is significantly different from the control substance.

The Panel agreed that the current recommendation to select a maximum applied dose for the rLLNA based on the absence of overt systemic toxicity and/or excessive local irritation is appropriate. The Panel also agreed that the data did not support establishment of a uniform concentration threshold for the maximum concentration to be tested. Thus, it seemed justifiable that preliminary experimentation (as would be typically performed during a dose range-finding study) should be conducted for vehicle selection, test substance solubility, and stability in the vehicle.

2.3 ICCVAM Recommendations: Future Studies

ICCVAM recommends additional studies to further characterize and potentially improve the usefulness and applicability of the rLLNA for identifying potential skin sensitizers. For instance, to improve the predictive performance of the rLLNA compared to the traditional LLNA, ICCVAM recommends investigating the basis for abnormal dose responses for six substances that would have resulted in false negative results using the rLLNA rather than the traditional LLNA. This information should help identify ways to improve the accuracy of the rLLNA compared to the traditional LLNA.

Efforts should also be made to identify data from guinea pigs and humans for substances like these that exhibit abnormal dose responses in the traditional LLNA. Information from post-marketing and/or occupational exposures should be collected and assessed.

ICCVAM recommends that all future LLNA studies should collect and analyze individual animal data. This will allow detection of outliers and avoidance of false negative results that can occur from pooling data that include one or more abnormally low values. Existing LLNA studies using data pooled from all animals in a dose group, such as four of the six false negative rLLNA results in this evaluation, should be evaluated further with data obtained from individual animals within each dose group to determine if data pooling may have led to false negative outcomes.

ICCVAM also recommends that users identify opportunities to use fewer animals per dose group without compromising test method accuracy. Thus, laboratories conducting the LLNA should collect and analyze data from individual animals. The updated ICCVAM-recommended LLNA test method protocol includes statistical procedures necessary for such determinations (**Appendix B**). This includes evaluating the laboratory's historical positive-control database to determine if the number of animals in the concurrent positive-control group can be reduced.

Independent Peer Review Panel Conclusions and Recommendations

The Panel indicated that, though limited in scope, the available data supported ICCVAM's draft test method recommendations for additional studies. The Panel agreed that attempts should be made to investigate if maximum solubility was achieved (e.g., use of chemical-specific methods to document solubility). For hazard assessment, it was troublesome that there were so many vehicle choices, because the vehicle could have a significant effect on whether (and how much) a test substance penetrated the skin barrier. Observed vehicle effects may relate to dermal

penetration as well as to immunomodulation. The Panel considered it desirable to follow the hierarchy of vehicles recommended in the ICCVAM-recommended LLNA protocol. The Panel suggested that it might be informative to test both known mild and severe sensitizers concurrently in all recommended vehicles to evaluate whether a specific vehicle choice(s) might influence the results.

2.4 ICCVAM Recommendations: Performance Standards

ICCVAM developed performance standards for the traditional LLNA, which may in turn be applied to the rLLNA.[18] These test method performance standards are proposed to evaluate modified LLNA test methods that are functionally and mechanistically similar to the traditional LLNA. Thus, modified rLLNA test method protocols that adhere to the LLNA performance standards would be considered acceptable for hazard identification purposes.

[18] Available at http://iccvam.niehs.nih.gov/methods/immunotox/PerfStds/llna-ps.htm

[This Page Intentionally Left Blank]

3.0 Validation Status of the rLLNA Test Method

The following is a synopsis of the information in the final ICCVAM BRD (**Appendix D**), which reviews the available data and information for the rLLNA test method. The ICCVAM BRD describes the current validation status of the rLLNA test method, including what is known about its reliability and accuracy, the scope of the substances tested, and standardized protocols used for the validation study.

3.1 Test Method Description

The purpose of the rLLNA test method is to identify potential skin sensitizers by quantifying lymphocyte proliferation. The mechanistic basis is identical to that of the traditional LLNA, which measures the magnitude of lymphocyte proliferation, which in turn correlates with the extent to which sensitization develops after a topical induction exposure to a skin-sensitizing substance.

With one exception, the technical aspects of the rLLNA are identical to those of the traditional LLNA (ICCVAM 1999). The traditional LLNA tests three dose levels of each test substance for skin-sensitizing activity. In the rLLNA, only one dose of the test substance is tested: the concentration that provides maximum solubility without causing overt systemic toxicity and/or excessive skin irritation (Kimber et al. 2006). Guidance for evaluating local irritation and systemic toxicity in the LLNA is provided in the updated ICCVAM-recommended LLNA protocol (**Appendix B**).

3.1.1 General Test Method Procedures

The rLLNA measures lymphocyte proliferation after topical exposure to a potential skin-sensitizing substance. The test substance is administered topically on three consecutive days to the ears of mice at a concentration that provides maximum solubility of the test substance without systemic toxicity and/or excessive local irritation. Two days after the final application of the test substance, ^{3}H-thymidine or ^{125}I-iododeoxyuridine (in phosphate-buffered saline; 250 µL/mouse) is administered via the tail vein. Five hours later the draining auricular lymph nodes are excised, and a single-cell suspension from the lymph nodes of each animal is prepared for quantifying the incorporation of radioactivity, which correlates with lymph node cell proliferation.

The incorporation of radioactive ^{3}H-thymidine or ^{125}I-iododeoxyuridine for each mouse is expressed in disintegrations per minute (dpm). The SI is calculated as the ratio of the mean dpm/mouse for each treatment group against the mean dpm/mouse for the vehicle control group. The threshold for a positive response is an SI ≥ 3.

3.1.2 Similarities and Differences between the Protocols for the Traditional LLNA and the rLLNA

As mentioned above, the only difference between the traditional LLNA (ICCVAM 1999) and the rLLNA is that only one test substance dose is included in rLLNA, while three doses are tested in the traditional LLNA. All other procedures are identical.

3.2 Validation Database

Data were obtained from 11 different sources, including published reports and unpublished data submitted to NICEATM in response to a May 17, 2007, *FR* notice (72 FR 27815[19]). The rLLNA database consisted of the results for the highest doses tested in these studies.

The resulting database consisted of 457 unique substances tested in a total of 471 traditional LLNA studies (**Table 3-1**), 211 of which were included in the original ICCVAM evaluation of the traditional LLNA (ICCVAM 1999). Fourteen of the 457 unique substances[20] were repeated from two to five times in different LLNA studies. Specifically, nine of the 14 substances were evaluated two to five times in different vehicles, and five of the 14 substances were evaluated two to five times in the same vehicle. Two of the five substances evaluated in the same vehicle (hexyl cinnamic aldehyde [HCA] and potassium dichromate) were also tested using different vehicles (one study for HCA and two studies for potassium dichromate). Due to the small number of repeated studies (5% of total studies), all studies were treated independently for the purpose of this accuracy evaluation. When the studies for the substances repeated in the same vehicle were considered together to yield an overall skin sensitization classification, there were 465 studies with unique substance and vehicle combinations.

Table 3-1 provides the chemical class information for these test substances. The table distinguishes the chemical classifications of the 211 substances included in the original evaluation of the rLLNA (Kimber et al. 2006; ESAC 2007) and the chemical classifications of the additional substances received in response to the *FR* notice. Of the 211 substances initially evaluated by Kimber et al. (2006), the chemical classes with the greatest number of substances were carboxylic acids (29) and halogenated hydrocarbons (27). Of the additional 246 substances included in this evaluation, the chemical classes with the greatest number of substances tested were pharmaceutical chemicals (125), carboxylic acids (15), and lipids (14). Of the substances included in this evaluation, 10 were formulations. Seventy substances could not be assigned to a specific chemical class due to incomplete available information (e.g., the lack of a Chemical Abstracts Service Registry Number or structure).

[19] Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf

[20] Some substances were tested in more than one vehicle. In such instances, each substance–vehicle combination was considered separately, thus a total of 465 unique substance–vehicle combinations were used in the performance evaluation.

Table 3-1 Chemical Classes[1] Represented in the Current Traditional LLNA Database

Chemical Class	Number of Substances - Original[2]	Number of Substances - Additional[2]	Chemical Class	Number of Substances - Original	Number of Substances - Additional
Alcohols	9	4	Inorganic Chemicals	0	2
Aldehydes	21	4	Isocyanates	1	0
Amides	4	0	Ketones	5	0
Amidines	1	0	Lactones	2	2
Amines	14	7	Lipids	7	14
Anhydrides	1	0	Macromolecular substances[3]	0	5
Carbohydrates	3	2	Nitriles	1	1
Carboxylic acids	29	15	Nitro compounds	2	0
Esters	3	0	Nitroso compounds	3	0
Ethers	14	2	Onium compounds	1	0
Formulations[3]	0	10	Pharmaceutical chemicals[4]	0	125
Heterocyclic compounds	18	4	Phenols	18	2
Hydrocarbons, Acyclic	2	1	Polycyclic compounds	5	3
Hydrocarbons, Cyclic	14	7	Quinones	1	1
Hydrocarbons, halogenated	27	1	Sulfur compounds	20	2
Hydrocarbons, other	7	8	Urea	3	0
Imines	0	1	Unknown	28	42

[1] Total number of substances assigned to chemical classes does not equal the total number of substances evaluated because some substances were assigned to more than one class and some substances were not assigned to a specific chemical class.

[2] Number of substances - Original represents the substances evaluated in Kimber et al. (2006).
Number of substances - Additional represents the substances received in response to the released *Federal Register* notice (72 FR 27815, May 17, 2007) available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf.

[3] No chemical class could be assigned. The terms "formulation" and "macromolecular substances" were used to classify these substances.

[4] The chemical classification of "pharmaceutical chemicals" for the GlaxoSmithKline (GSK) substances was suggested by Dr. Michael Olson of GSK to capture three types of pharmaceutical substances (actives, intermediates, and starting materials).

3.3 Reference Test Method Data

The traditional LLNA data used for evaluation of the rLLNA include the results for all tested doses of each substance. In addition to calculated SI values for each of the tested doses, the vehicles tested and EC3 values (estimated concentration needed to produce an SI value of 3) for substances classified as sensitizers were provided in Gerberick et al. (2005). The data received in response to the May 2007 *FR* notice included calculated SI values for the vehicle and each of

the tested doses. If EC3 values were not included in the data source, they were calculated, where possible, using either interpolation or extrapolation (Dearman et al. 2007). This information and the complete database (by each source) are provided in **Annex III** of the BRD (**Appendix D**).

3.4 Test Method Accuracy

The ability of the rLLNA to correctly identify potential skin sensitizers was compared to that of the traditional LLNA. Of the 471 studies, 318 detected skin sensitizers, and 153 detected non-sensitizers. When studies of the substances tested more than once in the same vehicle were considered together to yield an overall skin sensitization classification, 465 unique substance–vehicle combinations resulted. Of these, 315 were identified as sensitizers and 150 as non-sensitizers.

Based on the available study data, the rLLNA has an accuracy of 98.7% (465/471), a sensitivity of 98.1% (312/318), a specificity of 100% (153/153), a false positive rate of 0% (0/153), and a false negative rate of 1.9% (6/318) when compared to the traditional LLNA (**Table 3-2**). When substances tested more than once in the same vehicle were considered together, the resulting 465 studies had an accuracy of 98.7% (459/465), a sensitivity of 98.1% (309/315), a specificity of 100% (150/150), a false positive rate of 0% (0/150), and a false negative rate of 1.9% (6/315).

This analysis of the rLLNA yielded six false negative results. A review of the data for these six substances indicates that the traditional LLNA classification of the substances as skin sensitizers was based not on the highest tested dose but on a low- or mid-dose level that produced an SI $\geq$ 3, while the highest dose tested produced an SI < 3. Because the rLLNA tests substances at only the highest dose level, all six substances would be incorrectly identified as non-sensitizers (i.e., false negatives). Four of the six substances that resulted in false negatives using the rLLNA compared to the traditional LLNA came from LLNA studies that used pooled data. There were no patterns of consistency for these substances with regard to physicochemical properties.

Table 3-2 Evaluation of the Performance of the rLLNA in Predicting Skin Sensitizers Compared to the Traditional LLNA

Data	N	Accuracy	Sensitivity	Specificity	False Positive	False Negative
Kimber et al. (2006)	211	98.6% (208/211)	98.2% (166/169)	100% (42/42)	0% (0/42)	1.8% (3/169)
rLLNA	471	98.7% (465/471)	98.1% (312/318)	100% (153/153)	0% (0/153)	1.9% (6/318)
rLLNA approach (substances repeated in the same vehicle considered together)	465	98.7% (459/465)	98.1% (309/315)	100% (150/150)	0% (0/150)	1.9% (6/315)

Abbreviation: N = number of tests
Accuracy = the percentage of correct outcomes (positive and negative) of a test method
Sensitivity = the percentage of all positive substances that are classified as positive
Specificity = the percentage of all negative substances that are classified as negative
False positive rate = the percentage of all negative substances that are falsely identified as positive
False negative rate = the percentage of all positive substances that are falsely identified as negative

3.5 Test Method Reliability

The BRD assessed interlaboratory reproducibility of the rLLNA with traditional LLNA data for five substances that had been tested independently in the same vehicle at multiple laboratories. These five substances were dinitrochlorobenzene (DNCB), HCA, linalool alcohol, methyl salicylate, and potassium dichromate. **Table 3-3** summarizes the responses obtained by the rLLNA. All studies classified DNCB, methyl salicylate, and potassium dichromate (3/5 = 60%) as sensitizers or non-sensitizers (i.e., 100% concordance). HCA and linalool alcohol, which were tested independently in two laboratories, were each classifed as a sensitizer by one traditional LLNA study and as a non-sensitizer by the other traditional LLNA study. Review of the studies indicates that the discordant results were due to differences in the highest dose levels tested. However, because the rLLNA and traditional LLNA use identical protocols and the data sets used to evaluate their accuracy are similar, the intra- and interlaboratory reliability of the rLLNA is deemed to be similar to that of the traditional LLNA (see ICCVAM 1999 for these statistics).

Table 3-3 Interlaboratory Reproducibility of Skin sensitization Outcome for the rLLNA

Substance	Data Source	Vehicle	rLLNA Dose (%)/SI	rLLNA Classification[1]
1-Chloro-2-dinitrobenzene	Gerberick et al. (2005)	AOO	0.25/38.00	+
	Data submitted by D. Germolec		0.25/7.10	+
Hexyl cinnamic aldehyde	Gerberick et al. (2005)	AOO	50/17.00	+
	Data submitted by H.W. Vohr		10/2.84	–
Linalool alcohol	Gerberick et al. (2005)	AOO	100/8.30	+
	Data submitted by D. Basketter, I. Kimber, and F. Gerberick		30/1.30	–
Methyl salicylate	Gerberick et al. (2005)	AOO	20/0.90	–
	Data submitted by D. Germolec		20/1.72	–
Potassium dichromate	Gerberick et al. (2005)	DMSO	0.5/16.10	+
	Data submitted by D. Germolec		0.25/3.39	+
	Ryan et al. (2002)		0.5/10.10	+

Abbreviations: AOO = Acetone: olive oil; DMSO = Dimethyl sulfoxide; rLLNA = Reduced murine local lymph node assay; SI = stimulation index

[1] – = non-sensitizer, + = sensitizer

3.6 Animal Welfare Considerations: Reduction, Refinement, and Replacement

Compared to the traditional LLNA, the rLLNA will reduce the number of animals used to assess skin sensitization. Becuse the rLLNA tests only the highest dose level of the test substance in addition to the concurrent control groups, the number of animals tested would be decreased by at least 40% for each test. Ryan et al. (2008) described the impact of reducing the number of animals per group from five to two on the performance of the rLLNA and concluded that such a small number of animals per group was inadequate for hazard identification of skin sensitizers.

4.0 ICCVAM Consideration of Public and SACATM Comments

ICCVAM received 27 public comments in response to four *FR* notices released between May 2007 and May 2008 (see **Appendix G**). Comments received in response to or related to the *FR* notices are also available on the NICEATM–ICCVAM website.[21] The following sections, delineated by *FR* notice, briefly discuss the public comments received.

4.1 Public Comments in Response to 72 FR 27815 (May 17, 2007): The Murine Local Lymph Node Assay: Request for Comments, Nominations of Scientific Experts, and Submission of Data

NICEATM requested the following:

1. Public comments on the appropriateness and relative priority of evaluation of the validation status of
    a. The LLNA as a stand-alone assay for determining potency (including severity) for the purpose of hazard classification
    b. The rLLNA approach
    c. Non-radioactive LLNA methods
    d. The use of the LLNA for testing mixtures, aqueous solutions, and metals
    e. The current applicability domain
2. Nominations of expert scientists to consider as members of a possible peer review panel
3. Submission of data for the LLNA and/or modified versions of the LLNA

In response to this *FR* notice, NICEATM received 17 comments. Six comments included additional data and information, while two others offered data and information upon request. Three nominated four potential panelists for consideration. Three commenters suggested reference publications for consideration during the Panel evaluation. NICEATM provided the data and suggested references to the Panel for evaluation.

Three comments remarked specifically on the rLLNA.

One commenter suggested rearranging the priority sequence of test method evaluation from most to least pressing: a, e, d, b, and c (see list above). ICCVAM did not establish a relative priority for these activities because they were all considered to be high-priority activities. Accordingly, all LLNA-related activities described above were discussed at the March 2008 Panel meeting.

Another commenter noted that ESAC issued a statement supporting the use of the rLLNA "within tiered-testing strategies to reliably distinguish between chemicals that are skin sensitisers and non-sensitisers" (**Appendix E**), thereby reducing animal use by as much as 50%. The ESAC statement also notes the following limitations: "the test results provided by the rLLNA do not allow the determination of the potency of a sensitising chemical" and "negative

[21] Available at http://ntp-apps.niehs.nih.gov/iccvampb/searchPubCom.cfm

test results associated with testing using concentrations of less than 10% should undergo further evaluation." The commenter states that ICCVAM should (1) expeditiously review and endorse the ESAC peer review and circulate harmonized testing recommendations regarding this assay to U.S. agencies before the end of the year, and (2) NICEATM should collaborate with ECVAM to address the question of concentration threshold.

As indicated in **Section 1.0**, ICCVAM and NICEATM collaborated with liaisons from ECVAM and JaCVAM to update with 260 additional LLNA studies the Kimber et al. (2006) evaluation upon which the ESAC statement was based. This comprehensive evaluation was expedited for inclusion in the publicly transparent ICCVAM peer review process, which included the March 2008 Panel meeting.

A third commenter stated that ESAC considered the rLLNA to be scientifically validated but only when used as a screening test to distinguish between sensitizers and non-sensitizers and with due regard to the conditions set forth in the official ESAC statement of April 27, 2007. This statement was based on the outcome of a review of LLNA data for 211 chemicals (Kimber et al. 2006). The review of existing and newly provided LLNA data proposed by NICEATM–ICCVAM therefore presents an ideal opportunity to assess further the validity of the rLLNA for screening purposes. The ICCVAM test method recommendations detailed in **Section 2.0** describe the usefulness and limitations of the rLLNA based on the comprehensive ICCVAM evaluation of an expanded database of 471 LLNA studies.

4.2 Public Comments in Response to 72 FR 52130 (September 12, 2007): Draft Performance Standards for the Murine Local Lymph Node Assay: Request for Comments

NICEATM requested public comments on the initial ICCVAM-recommended draft LLNA performance standards developed to facilitate evaluation of modified LLNA protocols with regard to the traditional LLNA. In response to this *FR* notice, NICEATM received four comments, two of which suggested clarifications to the text. Another recommended that test substances chosen for testing in the various LLNA methods should be pure, with conclusive structures, and should not be mixtures.

The ICCVAM review of the rLLNA, in which only the highest dose is used to assign a positive/negative result for a test substance, was a retrospective evaluation of available LLNA studies with which to compare the outcome of the traditional protocol (in which all doses are considered and any positive result, regardless of concentration, can be used to establish a sensitizing substance). Therefore, although the validation status of the LLNA for testing mixtures is still under review, ICCVAM and NICEATM considered it appropriate to include all available data in the evaluation of the rLLNA.

The fourth commenter addressed the rLLNA in general. The commenter supported the development of performance standards that expedite the validation of new protocols similar to previously validated methods but was disappointed that NICEATM–ICCVAM has chosen to develop performance standards for such a narrow scope of applicability (i.e., modifications of the standard LLNA that involve incorporation of non-radioactive methods of detecting lymphocyte proliferation). The commenter suggested that limited resources available to NICEATM and ICCVAM would be better spent on activities that would have greater impact on

the reduction, refinement, or replacement of animal use, such as evaluating the use of human cell lines or one of the available *in vitro* skin models as a replacement for the LLNA.

ICCVAM considered the comment and concluded that the proposed modifications to the LLNA protocol and expanded applications have significant potential to further reduce and refine animal use. ICCVAM is also interested in *in vitro* models and non-animal approaches for assessing allergic contact dermatitis; however, no *in vitro* replacements for the LLNA have yet been nominated or submitted to ICCVAM for evaluation.

4.3 Public Comments in Response to 73 FR 1360 (January 8, 2008): Announcement of an Independent Scientific Peer Review Panel Meeting on the Murine Local Lymph Node Assay; Availability of Draft Background Review Documents; Request for Comments

NICEATM requested public comments on the draft BRDs, draft ICCVAM test recommendations, draft test method protocols, and revised draft LLNA performance standards for an international independent scientific peer review panel meeting to evaluate modifications and new applications for the LLNA. NICEATM received six comments in response to this *FR* notice. Four commenters focused on the traditional LLNA and two commenters provided comments specific to the rLLNA.

One commenter agreed with ICCVAM's recommendation of the rLLNA for hazard identification purposes, noting that Kimber et al. (2006) did not propose a 10% concentration threshold as the absolute cutoff for defining non-sensitizing chemicals. Gerberick et al. (2005) showed that for some compounds tested the highest concentration was at least 20% and did not induce a positive response at any concentration tested; these compounds were categorized as non-sensitizing. Cockshott et al. (2006) reported that a negative result obtained with the highest concentration tested at 10% would be considered a valid result if the positive control, a mild to moderate sensitizer, gave a positive response (i.e., a chemical that is negative at a top concentration of 10% does not represent a significant human sensitization hazard). This is similar to the definition of a non-sensitizing chemical in the Guinea Pig Maximization Test (GPMT) or Buehler Test as one that induces responses lower than 30% or 15%, respectively. Therefore, if a chemical elicits positive responses in 20% or 25% of the test animals in a GPMT, it would be considered a non-sensitizer from a regulatory perspective.

ICCVAM and the Panel agreed that the maximum applied dose for the rLLNA should be based on the absence of overt systemic toxicity and/or excessive local irritation. The available data did not support establishment of a uniform concentration threshold for the maximum concentration to be tested.

Another commenter's response referred first to the April 2007 ESAC statement:

> "…supporting the use of the rLLNA 'within tiered-testing strategies to reliably distinguish between chemicals that are skin sensitisers and non-sensitisers,' thereby reducing animal use by as much as 50%. In spite of the ESAC recommendation, ICCVAM conducted its own data call-in and data review. The reviewed database is comprehensive and contains a broad cross-section of the chemical universe. The performance characteristics were all above 95% (false negative and positive rates are very low or zero). Even though this additional review was largely unnecessary, [the

commenter was] pleased that ICCVAM's draft recommendations concluded favorably for the rLLNA procedure…"

The commenter urged the Panel to concur. As reflected in the Independent Scientific Review Panel Assessment (**Appendix F**), the Panel generally agreed with ICCVAM's test method recommendations for the rLLNA, which have been updated to reflect comments from the Panel, SACATM, and the public.

4.4 Public Comments in Response to 73 FR 29136 (May 20, 2008): Peer Review Panel Report on the Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay (LLNA): A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products: Notice of Availability and Request for Public Comments

NICEATM requested submission of written public comments on the Independent Scientific Peer Review Panel Assessment. No comments were received in response to this *FR* notice.

4.5 Public and SACATM Comments: SACATM Meeting on June 18-19, 2008

The June 18-19, 2008, SACATM meeting included a discussion of the ICCVAM review of the rLLNA test method.

There were no public comments specific to the rLLNA.

One SACATM member concurred with the recommendation that the rLLNA protocol should discuss how to determine the maximum dose if only a single dose is to be used in a screening process. An investigator must be able to define excessive irritation; otherwise, the testing may produce a bell-shaped response curve. In response to this comment and the Panel's recommendation, ICCVAM added to the updated LLNA test method protocol specific guidance on how to determine the maximum concentration to be tested so as to avoid overt systemic toxicity and/or excessive local irritation (**Appendix B, Annex III**).

Another SACATM member suggested that the rLLNA appeared favorable because 100% (153/153) of the non-sensitizing agents and 98.1% (312/318) of the sensitizing agents were correctly predicted. ICCVAM agrees that this high level of agreement between the traditional LLNA and the rLLNA supports routine use of the rLLNA as recommended by ICCVAM.

5.0 References

Buehler EV. 1965. Delayed contact hypersensitivity in the guinea pig. Arch Dermatol 91:171-177.

Cockshott A, Evans P, Ryan CA, Gerberick GF, Betts CJ, Dearman RJ, et al. 2006. The local lymph node assay in practice: A current regulatory perspective. Hum Exp Toxicol 25:387–394.

Dean J, Twerdok L, Tice R, Sailstad D, Hattan D, Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay. Conclusions and recommendations of an independent scientific peer review panel. Regul Toxicol Pharmacol 34:258–273.

Dearman RJ, Betts CJ, Farr C, McLaughlin J, Berdasco N, Wiench K, et al. 2007. Comparative analysis of skin sensitization potency of acrylates (methyl acrylate, ethyl acrylate, butyl acrylate, and ethylhexyl acrylate) using the local lymph node assay. Contact Dermatitis 57:242–247.

EPA. 2003. Health Effects Test Guideline, OPPTS 870.2600. Skin Sensitization EPA 712–C–03–197. Washington, DC: U.S. Environmental Protection Agency. Available: http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/870-2600.pdf.

ESAC. 2007. Statement on the Reduced Local Lymph Node Assay (rLLNA). European Commission Directorate General, Joint Research Centre, Institute for Health and Consumer Protection, European Centre for the Validation of Alternative Methods. April 2007. Available: http://ecvam.jrc.it/ft_doc/ESAC26_statement_rLLNA_20070525-1.pdf.

Gaspari AA. 2007. Mechanism of action and other potential roles of an immune response modifier. Cutis 79:36–45.

Gerberick GF, Ryan CA, Kern PS, Schlatter H, Dearman RJ, Kimber I, et al. 2005. Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. Dermatitis 16:157–202.

Hutchings CV, Shum KW, Gawkrodger DJ. 2001. Occupational contact dermatitis has an appreciable impact on quality of life. Contact Dermatitis 45:17–20.

ICCVAM 1999. The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemical/Compounds. NIH Publication No. 99-4494. Research Triangle Park, NC: National Toxicology Program. Available: http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf.

ICCVAM Authorization Act. 2000. Public Law 106-545, 42 U.S.C. 285*l*-3.

Kimber I, Dearman RJ, Betts CJ, Gerberick GF, Ryan CA, Kern PS, et al. 2006. The local lymph node assay and skin sensitization: a cut-down screen to reduce animal requirements? Contact Dermatitis 54:181–185.

Magnusson B, Kligman AM. 1970. Allergic Contact Dermatitis in the Guinea Pig. Springfield, IL: Charles C. Thomas.

OECD. 2002. Test Guideline 429. Skin Sensitisation: Local Lymph Node Assay, adopted April 24, 2002. In: OECD Guidelines for Testing of Chemicals. Paris: OECD. Available:

http://titania.sourceoecd.org/vl=7033968/cl=16/nw=1/rpsv/cw/vhosts/oecdjournals/1607310x/v1n4/contp1-1.htm

Ryan CA, Chaney JG, Kern PS, Patlewicz GY, Basketter DA, Betts CJ, et al. 2008. The reduced local lymph node assay: the impact of group size. J Appl Toxicol 28:518–523.

Ryan CA, Cruse LW, Skinner RA, Dearman RJ, Kimber I, Gerberick GF. 2002. Examination of a vehicle for use with water soluble materials in the murine local lymph node assay. Food Chem Toxicol 40:1719–1725.

Skoet R, Zachariae R, Agner T. 2003. Contact dermatitis and quality of life: a structured review of the literature. Br J Dermatol 149:452–456.

Appendix A

ICCVAM Evaluation of the Reduced Local Lymph Node Assay Test Method – Timeline

[This Page Intentionally Left Blank]

January 10, 2007	ICCVAM receives a letter from the CPSC nominating six LLNA review activities for evaluation,[22] including the reduced murine local lymph node assay (rLLNA).
January 2007	The ICCVAM IWG is re-established to work with NICEATM to carry out LLNA evaluations.
January 24, 2007	ICCVAM endorses the six CPSC-nominated LLNA review activities, including evaluation of the rLLNA.
May 17, 2007	*Federal Register* notice (72 FR 27815) – The Murine Local Lymph Node Assay: Request for Comments, Nominations of Scientific Experts, and Submission of Data
June 12, 2007	SACATM endorses with high priority the six CPSC-nominated LLNA review activities, including evaluation of the rLLNA.
November 12–13, 2007	ECVAM Workshop on Alternative Methods (Reduction, Refinement, Replacement)
January 8, 2008	*Federal Register* notice (73 FR 1360) – Announcement of an Independent Scientific Peer Review Panel Meeting on the Murine Local Lymph Node Assay; Availability of Draft Background Review Documents; Request for Comments
March 4–6, 2008	Independent Scientific Peer Review Panel holds a public meeting, with opportunity for oral public comments, at CPSC Headquarters in Bethesda, MD, to discuss LLNA review activities, including the rLLNA. The Panel was charged with reviewing the current validation status of the rLLNA and commenting on the extent to which the information in the draft BRD supported the draft ICCVAM test method recommendations.
May 20, 2008	*Federal Register* notice (73 FR 29136) – Announcement of the Peer Review Panel Report on the Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay (LLNA): A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products: Notice of Availability and Request for Public Comments[23]
June 18–19, 2008	SACATM public meeting for comments on the Panel report
October 29, 2008	ICCVAM endorses the TMER for the rLLNA test method, which includes the final rLLNA BRD.

Abbreviations: BRD = Background review document; CPSC = U.S. Consumer Product Safety Commission; ECVAM = European Centre for the Validation of Alternative Methods; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; IWG = ICCVAM Immunotoxicity Working Group; LLNA = murine local lymph node assay; NICEATM = National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods; rLLNA = Reduced murine local lymph node assay;

[22] http://iccvam.niehs.nih.gov/methods/immunotox/llnadocs/CPSC_LLNA_nom.pdf
[23] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

SACATM = Scientific Advisory Committee on Alternative Toxicological Methods; TMER = Test method evaluation report

Appendix B

ICCVAM-Recommended Protocol (Updated 2008)
The Murine Local Lymph Node Assay:[24] A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products

[24] Based on ICCVAM (1999) and Dean et al. (2001)

[This Page Intentionally Left Blank]

Preface

The murine local lymph node assay (LLNA) is a test method developed to assess whether a chemical has the potential to induce allergic contact dermatitis (ACD) in humans. In 1998, the LLNA was submitted to the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) for evaluation as an alternative (i.e., stand-alone) test method to the guinea pig (GP) sensitization tests accepted by U.S. regulatory agencies. In 1999, based on a comprehensive evaluation of the LLNA by an independent scientific peer review panel (Panel),[25] ICCVAM concluded that the LLNA is an acceptable alternative to the GP test methods to assess the ACD hazard potential of most substances (Dean et al. 2001). The Panel also concluded that the LLNA offers animal welfare advantages compared to use of the traditional GP methods, in that it provides for animal use refinement (i.e., elimination of distress and pain) and reduces the total number of animals required. An ICCVAM Immunotoxicity Working Group (IWG) reviewed the 1999 Panel report and developed recommendations applicable to the regulatory use of the LLNA. The IWG then worked with the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) to produce a recommended test method protocol (ICCVAM 2001)[26] that would accurately reflect the ICCVAM and Panel recommendations (ICCVAM 1999).

In March 2008, ICCVAM and NICEATM convened an independent scientific peer review panel (Panel) to evaluate new versions and applications of the LLNA. The Panel provided conclusions and recommendations in their report, many of which were applicable to the traditional LLNA test method protocol.[27] ICCVAM subsequently considered the Panel's conclusions and recommendations, as well as comments from the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) and public, and updated the 2001 ICCVAM-recommended LLNA test method protocol. The updated ICCVAM-recommended LLNA test method protocol will be forwarded with the Panel's report to agencies for their consideration.

The updated ICCVAM-recommended test method protocol for the LLNA is based on evaluation of previous experience and scientific data. It is provided to Federal agencies for their consideration as a standardized test method protocol recommended for generation of data for regulatory purposes. Prior to conducting a LLNA test to meet a regulatory requirement, it is recommended that the appropriate regulatory agency be contacted for their current guidance on the conduct and interpretation of this assay. Additional information on the ICCVAM LLNA review process and deliberations of the Panel can be found at the ICCVAM website (http://iccvam.niehs.nih.gov) or in the Panel report (ICCVAM 2008a).

We want to express our sincere appreciation to the ICCVAM IWG for their careful deliberations and efforts in updating the LLNA test method protocol, and especially appreciate the efforts of the Working Group Co-Chairs, Abigail Jacobs, Ph.D., from the U.S. Food and Drug Administration and Joanna Matheson, Ph.D., from the U.S. Consumer Products Safety Commission. We also want to acknowledge the outstanding support

[25] http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf
[26] http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/LLNAProt.pdf
[27] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

provided by NICEATM and the Integrated Laboratory Systems, Inc., support staff. Lastly, we appreciate the efforts of the Panel members for their diligent review, and the comments provided by SACATM and numerous stakeholders, including the public.

William S. Stokes, D.V.M., D.A.C.L.A.M.
Rear Admiral/Assistant Surgeon General, U.S. Public Health Service
Director, NICEATM
Executive Director, ICCVAM

Marilyn Wind, Ph.D.
Deputy Associate Executive Director
Directorate for Health Sciences
U.S. Consumer Product Safety Commission
Chair, ICCVAM

1.0 General Principle of Detection of Skin Sensitization Using the Local Lymph Node Assay

The basic principle underlying the murine local lymph node assay (LLNA) is that sensitizers induce proliferation of lymphocytes in the lymph node draining the site of substance application. Under appropriate test conditions, this proliferation is proportional to the dose applied, and provides a means of obtaining an objective, quantitative measurement of sensitization. The test measures cellular proliferation as a function of *in vivo* radioisotope incorporation into the DNA of dividing lymphocytes. The LLNA assesses this proliferation in the draining lymph nodes proximal to the application site (see **Annex I**). This effect occurs as a dose response in which the proliferation in test groups is compared to that in the concurrent vehicle-treated control group. A concurrent positive control is added to each assay to provide an indication of appropriate assay performance.

2.0 Description of the Local Lymph Node Assay

2.1 Sex and strain of animals

Young adult female mice (nulliparous and non-pregnant) of the CBA/Ca or CBA/J strain are recommended.[28] Females are used because most data in the existing database were generated using mice of this gender. At the start of the study, mice should be age 8–12 weeks. All mice should be age matched (preferably within a one-week time frame). Weight variations between the mice should not exceed 20% of the mean weight.

2.2 Preparation of animals

The temperature of the experimental animal room should be 21°C (±3°C) and the relative humidity 30%–70%. When artificial lighting is used, the light cycle should be 12 hours light: 12 hours dark. For feeding, an unlimited supply of standard laboratory mouse diets and drinking water should be used. The mice should be acclimatized for at least five days prior to the start of the test (ILAR 1996). Mice should be housed in small groups unless adequate scientific rationale for housing mice individually is provided (ILAR 1996). Healthy mice are randomly assigned to the control and treatment groups. The mice are uniquely identified prior to being placed in the study. The method used to mark the mice should not involve identification via the ear (e.g., marking, clipping, or punching of the ear). All mice should be examined prior to the initiation of the test to ensure that there are no skin lesions present.

2.3 Preparation of doses

Solid test substances should be dissolved in appropriate solvents or vehicles and diluted, if appropriate, prior to dosing of the mice. Liquid test substances may be dosed directly (i.e., applied neat) or diluted prior to dosing. Fresh preparations of the test substance should be prepared daily unless stability data demonstrate the acceptability of storage.

[28] Male mice or other strains of mice may be used if it is sufficiently demonstrated that these animals perform as well as female CBA mice in the LLNA.

2.4 Test Conditions

2.4.1 Solvent/vehicle

The selected solvent/vehicle must not interfere with or bias the test result and should be selected on the basis of maximizing the test concentrations while producing a solution/suspension suitable for application of the test substance. In order of preference, recommended solvents/vehicles are acetone: olive oil (4:1 v/v), *N,N*-dimethylformamide, methyl ethyl ketone, propylene glycol, and dimethyl sulfoxide, but others may be used (Kimber and Basketter 1992). Particular care should be taken to ensure that hydrophilic materials are incorporated into a vehicle system that wets the skin and does not immediately run off. Thus, wholly aqueous vehicles may need to be avoided. It may be necessary for regulatory purposes to test the substance in the clinically relevant solvent or product formulation.

2.4.2 Controls

Concurrent negative (solvent/vehicle) controls should be included in each test to ensure that the test system is functioning properly and that the specific test is valid. In some circumstances (e.g., when using a solvent/vehicle not recommended in **Section 2.4.1**), it may be useful to include a naïve control. Except for treatment with the test substance, the mice in the negative control groups should be handled in an identical manner to the mice of the treatment groups.

Concurrent positive controls are used to ensure the appropriate performance of the assay by demonstrating that the test method is responding with adequate and reproducible sensitivity to a sensitizing substance for which the magnitude of the response is well characterized. Inclusion of a concurrent positive control is also important since it can confirm technical competence in performing the test and can demonstrate intra- and interlaboratory reproducibility and comparability. The positive control should produce a positive LLNA response (i.e., a stimulation index [SI] ≥ 3 over the negative control group). In particular, for negative LLNA studies, the concurrent positive control must induce a SI ≥ 3 relative to its vehicle-treated control. The positive control dose should be chosen such that the induction is reproducible but not excessive (i.e., SI > 20). Preferred positive control substances are hexyl cinnamic aldehyde or mercaptobenzothiazole. There may be circumstances where, given adequate justification, other positive control substances may be used.

Although the positive control substance should be tested in the same vehicle as the test substance, there may be certain regulatory situations where it is necessary to test the positive control substance in both a standard and a non-standard vehicle (e.g., a clinically/chemically relevant formulation) to test for possible interactions.

Inclusion of a positive control with each test is recommended to ensure that all test method protocol procedures are being conducted properly and that all aspects of the test system are working properly such that they are capable of producing a positive response. However, periodic testing (i.e., at intervals ≤6 months) of the positive control substance may be considered in laboratories that conduct the LLNA regularly (i.e., conduct the LLNA at a frequency of no less than once per month) and that have a history and a documented proficiency for obtaining consistent results with positive controls. Adequate proficiency with the LLNA can be successfully demonstrated by generating consistent results with the positive control in at least 10 independent tests conducted within a reasonable period of time (i.e., less than one

year). A positive control group should always be included when there is a procedural change to the LLNA (i.e., change in trained personnel, change in test method materials and/or reagents, change in test method equipment, change in source of test animals, etc.), and such changes should be documented in laboratory reports. Consideration should be given to the impact of these changes on the adequacy of the previously established historical database in determining the necessity for establishing a new historical database to document consistency in the positive control results. Users should be aware that the decision to only include a positive control on a periodic basis instead of concurrently will have ramifications on the adequacy and acceptability of negative study results generated without a concurrent positive control during the interval between each periodic positive control study. For example, if a false negative result is obtained in the periodic positive control study, all negative test substance results obtained in the interval between the last acceptable periodic positive control study and the unacceptable periodic positive control study will be questioned. In order to demonstrate that the prior negative test substance study results are acceptable, a laboratory would be expected to repeat all negative studies, which would require additional expense and increased animal use. These implications should be carefully considered when determining whether to include concurrent positive controls or to only conduct periodic positive controls. Consideration should also be given to using fewer animals in the concurrent positive control group when this is scientifically justified, as discussed below and in **Annex II**.

Benchmark controls may be useful to demonstrate that the test method is functioning properly for detecting the skin sensitization potential of substances of a specific chemical class or a specific range of responses, or for evaluating the relative skin sensitization potential of a test substance. Appropriate benchmark controls should have the following properties:

- Structural and functional similarity to the class of the substance being tested
- Known physical/chemical characteristics
- Supporting data on known effects in animal models
- Known potency for sensitization response

2.5 Methodology

A minimum of four animals per dose group is recommended. The collection of lymph nodes from individual mice is necessary in order to identify if any of the individual animal responses are outliers (e.g., in accordance with statistical tests such as Dixon's test). This will aid in avoiding false negative results for weaker sensitizers (i.e., substances that normally would induce an SI just above 3 might be incorrectly classified as negative due to a low outlier value, because the resulting mean SI may be less than 3 if an outlier is not identified and excluded). Individual animal measurements allow for the assessment of interanimal variability, a statistical comparison of the difference between test substance and vehicle control group measurements, and the evaluation of statistical power for different group sizes. Finally, evaluating the possibility of reducing the number of mice in the positive control group is only feasible when individual animal data are collected.

As noted above, concurrent negative and positive control groups should be included, unless a laboratory can demonstrate adequate proficiency that would support the use of a periodic positive control study. The number of mice in the concurrent positive control group might be

reduced compared to the vehicle and test substance groups, if the laboratory demonstrates, based on laboratory-specific historical data,[29] that fewer mice can be used without substantially increasing the frequency with which studies will need to be repeated. An example of how to reduce the number of mice in the concurrent positive control group is provided in **Annex II**.

Test substance treatment dose levels should be based on the recommendations given in Kimber and Basketter (1992) and in the ICCVAM Panel Report (ICCVAM 1999). Dose levels are selected from the concentration series 100%, 50%, 25%, 10%, 5%, 2.5%, 1%, 0.5%, etc. The maximum concentration tested should be the highest achievable level while avoiding excessive local irritation and overt systemic toxicity (**Annex III**). Efforts should be made to identify existing information that may aid in selecting the appropriate maximum test substance dose level. In the absence of such information, an initial prescreen test, conducted under identical experimental conditions except for not conducting an assessment of lymph node proliferative activity, may be necessary. In order to have adequate information on which to select a maximum dose level to use in the definitive test and to identify a dose-response relationship, data should be collected on at least three test substance dose levels with two mice per dose group, in addition to the concurrent solvent/vehicle control group.

The LLNA experimental procedure is performed as follows:

> **Day 1.** Identify and record the weight of each mouse before applying the test substance. Apply 25 μL/ear of the appropriate dilution of the test substance, or the positive control, or the solvent/vehicle only, to the dorsum of both ears of each mouse.
>
> **Days 2 and 3.** Repeat the application procedure as carried out on Day 1.
>
> **Days 4 and 5**. No treatment.
>
> **Day 6**. Record the weight of each mouse. Inject 250 μL of sterile phosphate-buffered saline (PBS) containing 20 μCi of tritiated (^{3}H)-methyl thymidine or 250 μL PBS containing 2 μCi of ^{125}I-iododeoxyuridine (^{125}IU) and 10^{-5} M fluorodeoxyuridine into each mouse via the tail vein (Kimber et al. 1995; Loveless et al. 1996). Five hours later, each mouse is euthanized and the draining ("auricular") lymph nodes of both ears are collected and placed in PBS (one container per mouse). Both bilateral draining lymph nodes must be collected (see diagram and description of dissection in **Annex I**). A single-cell suspension of lymph node cells (LNC) is prepared for each individual mouse. The single-cell suspension is prepared in PBS by either gentle mechanical separation through 200-mesh stainless steel gauze or another acceptable technique for generating a single-cell suspension. LNC are washed twice with an excess of PBS and the DNA precipitated with 5% trichloroacetic acid (TCA) at 4°C for approximately 18 hours.
>
> For the ^{3}H-methyl thymidine method, pellets are resuspended in 1 mL TCA and transferred to 10 mL of scintillation fluid. Incorporation of ^{3}H-methyl thymidine

[29] A robust historical dataset should include at least 10 independent tests, conducted within a reasonable period of time (i.e., less than one year), with a minimum of four mice per negative and positive control groups.

is measured by β-scintillation counting as disintegrations per minute (dpm) for each mouse and expressed as dpm/mouse. For the ^{125}IU method, the 1 mL TCA pellet is transferred directly into gamma-counting tubes. Incorporation of ^{125}IU is determined by gamma counting and also expressed as dpm/mouse.

2.6 Observations

Mice should be carefully observed for any clinical signs, either of local irritation at the application site or of systemic toxicity (**Annex III**). Weighing mice prior to treatment and at the time of necropsy will aid in assessing systemic toxicity. All observations are systematically recorded and records maintained for each individual mouse. Animal monitoring plans must include criteria to promptly identify mice exhibiting systemic toxicity or excessive irritation or corrosion of skin for euthanasia.

3.0 Calculation of Results

Results for each treatment group are expressed as the mean SI. Each SI is the ratio of the mean dpm/mouse within each test-substance treatment group or the positive control treated group against the mean dpm/mouse for the solvent/vehicle treated control group. However, the investigator should be alert to possible outlier responses for individual mice within a group that may necessitate analysis both with and without the outlier.

In addition to a formal assessment of the magnitude of the SI, a statistical analysis for presence and degree of dose response may be conducted, which is possible only with the use of individual animals. Any statistical assessment should include an assessment of the dose-response relationship as well as suitably adjusted comparisons of test groups (e.g., pair-wise dosed group versus concurrent solvent/vehicle control comparisons). Analyses may include, for instance, linear regression, William's test to assess dose-response trends, or Dunnett's test for pairwise comparisons. In choosing an appropriate method of statistical analysis, the investigator should be aware of possible inequality of variances and other related problems that may necessitate a data transformation or a non-parametric statistical analysis.

4.0 Evaluation and Interpretation of Results

In general, when the SI for any single treatment dose group is ≥ 3, the test substance is regarded as a skin sensitizer (Kimber et al. 1994; Basketter et al. 1996; ICCVAM 1999) and a test substance not meeting this criterion is considered a non-sensitizer in this test. However, the magnitude of the observed SI should not be the sole factor used in determining the biological significance of a skin sensitization response. Additional factors that could be considered include the outcomes of statistical analyses, the strength of the dose-response relationship, chemical toxicity, and solubility. For instance, a quantitative assessment may be performed by statistical analysis of individual mouse data and may provide a more complete evaluation of the test substance's ability to act as a sensitizer (see **Section 3.0**). Equivocal results (e.g., the SI does not reach 3, but it is near 3 and there is a positive dose-response relationship) should be clarified by performing statistical analysis, and by considering structural relationships, available toxicity information, and dose selection.

5.0 Data and Reporting

5.1 Data

Individual animal dpm data should be presented in tabular form, along with the group mean dpm/mouse, its associated error term, and the mean SI (and associated error term) for each dose group compared against the concurrent solvent/vehicle control group.

5.2 Test Report

The test report should contain the following information:

Test Substances and Control Substances

- Identification data and Chemical Abstracts Service Registry Number, if known
- Physical nature and purity
- Physiochemical properties relevant to the conduct of the study
- Stability of the test substance, if known
- Lot number of the test substance

Solvent/Vehicle:

- Justification for choice of solvent/vehicle
- Solubility and stability of the test substance in the solvent/vehicle

Test Animals:

- Strain of mice used
- Number, age, and sex of mice
- Source, housing conditions, diet, etc.
- Individual weight of the mice at the start and end of the test, including body weight range, as well as mean and associated error term for each group
- Microbiological status of the mice

Test Conditions:

- Concurrent and historical positive and negative (solvent/vehicle) control data
- Data from range-finding study, if conducted
- Rationale for dose-level selection
- Details of test substance preparation
- Details of the administration of the test substance
- Details of food and water quality
- Detailed description of treatment and sampling schedules
- Methods for measurement of toxicity
- Criteria for considering studies as positive, negative, or equivocal

Results:

- Signs of systemic toxicity and/or local irritation
- Values for dpm/mouse for each mouse within each treatment group
- Mean and associated error term for dpm/mouse for each treatment group and the results of outliner analysis for each dose group should be provided
- Calculated SI and an appropriate measure of variability that takes into account the interanimal variability in both the test substance dosed and control groups
- Dose-response relationship
- Statistical analyses and method applied
- Concurrent and historical positive and negative (solvent/vehicle) control data as established in the test laboratory
- Concurrent positive control data or, if not done, the date and laboratory report for the most recent periodic positive control and a report detailing the historical positive control data for the laboratory justifying the basis for not conducting a concurrent positive control.

Discussion of the Results

Conclusion

A Quality Assurance Statement for GLP-compliant Studies

- This statement should indicate all inspections made during the study and the dates any results were reported to the Study Director. This statement should also confirm that the final report reflects the raw data.

6.0 References

Basketter DA, Gerberick GF, Kimber I, Loveless SE. 1996. The local lymph node assay – a viable alternative to currently accepted skin sensitisation tests. Food Chem Toxicol 34:985-997.

Dean JH, Twerdok LE, Tice RR, Sailstad DM, Hattan DG, Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay. Conclusions and recommendations of an independent scientific peer review panel. Regul Toxicol Pharmacol 34:258-273.

Ehling G, Hecht M, Heusener A, Huesler J, Gamer AO, van Loveren H, et al. 2005. A European inter-laboratory validation of alternative endpoints of the murine local lymph node assay: first round. Toxicology 212:60-68.

EPA. 1998. Health Effects Test Guidelines: OPPTS 870.1200 - Acute Dermal Toxicity. EPA 712–C–98–192. Washington, DC: U.S. Environmental Protection Agency. Available: http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/870-1200.pdf.

Hayes BB, Gerber PC, Griffey SS, Meade BJ. 1998. Contact hypersensitivity to dicyclohexylcarbodiimide and diisopropylcarbodiimide in female B6C3F1 mice. Drug Chem Toxicol 21:195-206.

Hayes BB, Meade BJ. 1999. Contact sensitivity to selected acrylate compounds in B6C3F1 mice: relative potency, cross reactivity, and comparison of test methods. Drug Chem Toxicol 22:491-506.

Homey B, von Schilling C, Blumel J, Schuppe H-C, Ruzicka T, Jürgen AH, et al. 1998. An integrated model for the differentiation of chemical-induced allergic and irritant skin reactions. Toxicol Appl Pharmacol 153:83-94.

ICCVAM. 1999. The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals/Compounds. The Results of an Independent Peer Review Evaluation Coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication No. 99-4494. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf.

ICCVAM. 2001. Protocol: Murine Local Lymph Node Assay (LLNA). Purpose: Assessment of Allergic Contact Dermatitis Potential. Recommended by: ICCVAM Immunotoxicity Working Group – based on an Independent Expert Peer Review Panel Evaluation of the LLNA. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/LLNAProt.pdf.

ICCVAM. 2008a. Independent Scientific Peer Review Panel Report. Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf.

ICCVAM. 2008b. Non-radioactive Murine Local Lymph Node Assay: Flow Cytometry Test Method Protocol (LLNA: BrdU-FC) Draft Background Review Document. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/methods/immunotox/llna-BRD.htm.

ICCVAM. In press. Report on the ICCVAM-NICEATM/ECVAM/JaCVAM Scientific Workshop on Acute Chemical Safety Testing: Advancing In Vitro Approaches and Humane Endpoints for Systemic Toxicity Evaluations. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/methods/acutetox/Tox_workshop.htm.

Institute of Laboratory Animal Research (ILAR). 1996. Guide for the Care and Use of Laboratory Animals. 7th ed. Washington, DC: National Academies Press.

Kimber I, Basketter DA. 1992. The murine local lymph node assay; collaborative studies and new directions: a commentary. Food Chem Toxicol 30:165-169.

Kimber I, Dearman RJ, Scholes EW, Basketter DA. 1994. The local lymph node assay: developments and applications. Toxicology 93:13-31.

Kimber I, Hilton J, Dearman RJ, Gerberick GF, Ryan CA, Basketter DA, et al. 1995. An international evaluation of the murine local lymph node assay and comparison of modified procedures. Toxicology 103:63-73.

Loveless SE, Ladics GS, Gerberick GF, Ryan CA, Basketter DA, Scholes EW, et al. 1996. Further evaluation of the local lymph node assay in the final phase of an international collaborative trial. Toxicology 108:141-152.

OECD. 1987. Guideline For Testing of Chemicals – Test Guideline 402: Acute Dermal Toxicity. Paris: Organisation for Economic Co-operation and Development. Available: http://www.oecd.org/document/40/0,2340,en_2649_34377_37051368_1_1_1_1,00.html.

Patterson RM, Noga E, Germolec D. 2007. Lack of evidence for contact sensitization by Pfiesteria extract. Environ Health Perspect 115:1023-1028.

Reeder MK, Broomhead YL, DiDonato L, DeGeorge GL. 2007. Use of an enhanced local lymph node assay to correctly classify irritants and false positive substances. Toxicologist 96 (S-1):235.

Tilney NL. 1971. Patterns of lymphatic drainage in the adult laboratory rat. J Anat 109:369-383.

Vohr, HW, Jürgen AH. 2005. The local lymph node assay being too sensitive? Arch Toxicol 79:721-728.

Woolhiser MR, Hayes BB, Meade BJ. 1998. A combined murine local lymph node and irritancy assay to predict sensitization and irritancy potential of chemicals. Toxicol Meth 8: 245-256.

[This Page Intentionally Left Blank]

.

Annex I:
An Approach to Dissection and Identification of the Draining ("Auricular") Lymph Nodes

1.0 Background

Although minimal technical training of the murine local lymph node assay (LLNA) is required, extreme care must be taken to ensure appropriate and consistent dissection of the lymph nodes. It is recommended that technical proficiency in the dissection and identification of the lymph nodes draining the ear be achieved by practice on mice that have been (a) injected with a colored agent (dye) and/or (b) sensitized with a strong positive sensitizer. Brief descriptions of these practice dissections are provided below. Recognizing that nodes from vehicle-treated and naïve mice are smaller, laboratories performing the LLNA must also gain proficiency in the dissection of these nodes. It may be helpful for laboratories inexperienced in this procedure to request guidance from laboratories that have successfully performed the LLNA.

2.0 Training and Preparation for Node Identification

2.1 Identification of the Draining Node – Dye Treatment

There are several methods that can be used to provide color identification of the draining nodes. These techniques may be helpful for initial identification and should be performed to ensure proper isolation of the appropriate node. Examples of such treatments are listed below. It should be noted that other such protocols might be used effectively.

Evan's Blue Dye treatment:

Inject approximately 0.1 mL of 2% Evan's Blue Dye (prepared in sterile saline) intradermally into the pinnae of an ear. Euthanize the mouse after several minutes and continue with the dissection as noted below.

Colloidal carbon and other dye treatments:

Colloidal carbon and India ink are examples of other dye treatments that may be used (Tilney 1971).

2.2 Identification of the Draining Node – Application of Strong Sensitizers

For the purpose of node identification and training, a strong sensitizer is recommended. This agent should be applied in the standard acetone: olive oil vehicle (4:1). Suggested sensitizers for this training exercise include 0.1% oxazolone, 0.1% (w/v) 2,4-dinitrochlorobenzene, and 0.1% (v/v) dinitrofluorobenzene. After treating the ear with a strong sensitizer, the draining node will dramatically increase in size, thus aiding in identification and location of the node.

Using a procedure similar to that described in the test method protocol, apply the agent to the dorsum of both ears (25 μL/ear) for 3 consecutive days. On the fourth day, euthanize the mouse. Identification and dissection (listed below) of the node should be performed in these animals prior to practice in non-sensitized or vehicle-treated mice, where the node is significantly smaller.

Please note: Due to the exacerbated response, the suggested sensitizers are not recommended as controls for assay performance. They should only be used for training and node identification purposes.

3.0 Dissection Approach

3.1 Lateral Dissection (Figure B-1)

Although lateral dissection is not the conventional approach used to obtain the nodes draining the ear, it may be helpful as a training procedure when used in combination with the ventral dissection. Perform this approach bilaterally (on both sides of the mouse). After euthanizing the mouse, place it in a lateral position. Wet the face and neck with 70% ethanol. Use scissors and forceps to make an initial cut from the neck area slightly below the ear. Carefully extend the incision toward the mouth and nose. Angle the tip of the scissors slightly upward during this procedure to prevent the damage of deeper tissue. Gently retract the glandular tissue in the area using the forceps. Using the masseter muscle, facial nerves, blood vessels, and the bifurcation of the jugular vein as landmarks, isolate and remove the draining node (**Figure B-1**). The draining node ("auricular") will be positioned adjacent to the masseter muscle and proximal to and slightly above the jugular bifurcation.

3.2 Ventral Dissection (Figure B-2)

The most commonly used dissection approach is from the ventral surface of the mouse. This approach allows both right and left draining nodes to be obtained without repositioning the mouse. With the mouse ventrally exposed, wet the neck and abdomen with 70% ethanol. Use scissors and forceps to carefully make the first incision across the chest and between the arms. Make a second incision up the midline perpendicular to the initial cut, and then cut up to the chin area. Reflect the skin to expose the external jugular veins in the neck area. Take care to avoid salivary tissue at the midline and nodes associated with this tissue. The nodes draining the ear ("auricular") are located distal to the masseter muscle, away from the midline, and near the bifurcation of the jugular veins.

4.0 Accuracy in Identification

The nodes can be distinguished from glandular and connective tissue in the area by the uniformity of the nodal surface and a shiny translucent appearance. Application of sensitizing agents (especially the strong sensitizers used in training) will cause enlargement of the node size. If a dye is injected for training purposes, the node will take on the tint of the dye.

Figure B-1 Lateral Dissection

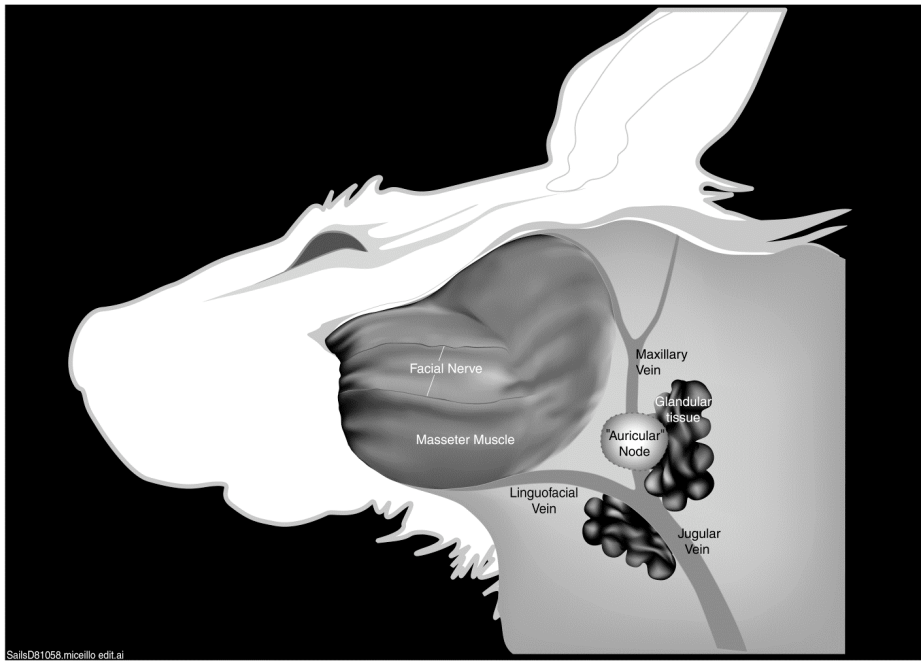
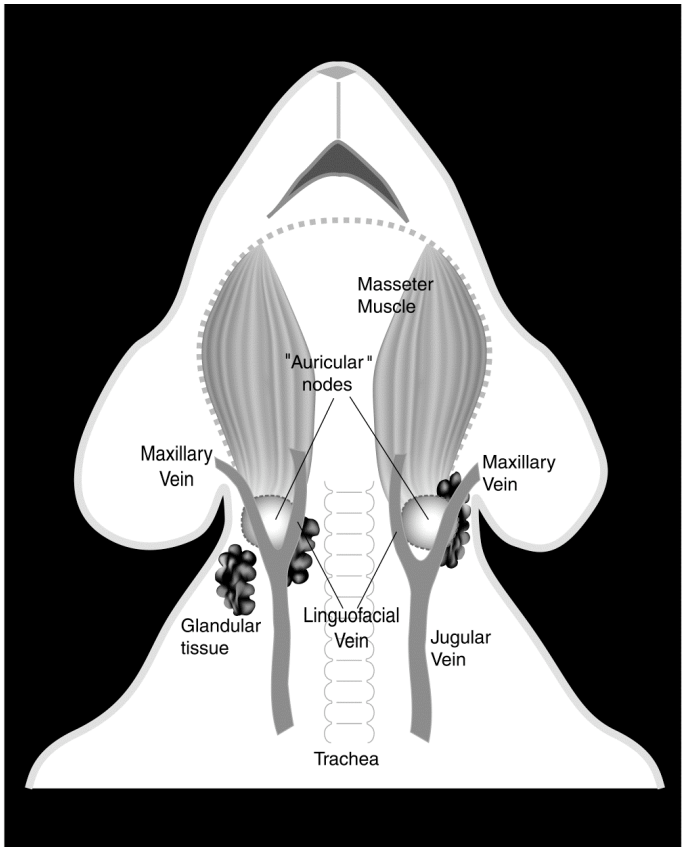


Figure B-2 Ventral Dissection



[This Page Intentionally Left Blank]

Annex II:
An Example of How to Reduce the Number of Animals in the Concurrent Positive Control Group of the Local Lymph Node Assay

As stated in the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Murine Local Lymph Node Assay (LLNA) test method protocol (**Section 2.4.2** of **Appendix B**), a concurrent positive control is recommended to ensure the appropriate performance of the assay. Appropriate performance is demonstrated when the test method responds with adequate and reproducible sensitivity to a sensitizing substance for which the magnitude of the response is well characterized. The number of mice in the concurrent positive control group may possibly be reduced if the laboratory demonstrates, based on laboratory-specific historical data, that fewer mice can be used without compromising the integrity of the study (i.e., positive control results should be always be positive compared to the vehicle control results). As illustrated in the example and accompanying explanation below, reducing the number of animals in the positive control group is only feasible when individual animal data are collected.

The stimulation index (SI) results for each positive control test can be used to generate mean SI values for every possible combination of SI values for as few as two animals. The mean SI values for every combination of numbers for each group size can then be used to calculate the failure rate of the positive control for each group size (i.e., the percentage of the combinations for which the mean SI < 3). **Table B-1** provides an example of positive control results from four tests in one laboratory of 30% hexyl cinnamic aldehyde (HCA) using six CBA/J mice per group. In these tests, with six animals, HCA produced "borderline" positive results (i.e., the mean SI values were marginally greater than 3). To determine whether the number of animals can be reduced, sample size reductions (i.e., N = 5, 4, 3, or 2) can be evaluated by taking all possible samples from the six values for each test given in **Table B-1**, which can occur in the following ways: N = 2 (15 samples), N = 3 (20 samples), N = 4 (15 samples), and N = 5 (six samples).

Table B-1 Example of SI Results from Four Local Lymph Node Assay Positive Control Studies with 30% HCA

Test	1	2	3	4
Animal 1	2.13	3.56	4.68	0.78
Animal 2	4.55	1.54	4.44	9.16
Animal 3	3.64	3.00	5.41	6.66
Animal 4	1.98	3.87	3.32	3.02
Animal 5	3.09	3.79	2.89	2.32
Animal 6	3.77	3.96	1.81	2.91
Mean SI	**3.19**	**3.29**	**3.76**	**4.14**

Abbreviations: HCA = hexyl cinnamic aldehyde; SI = stimulation index

The failure rate of the positive control was then calculated using the SI results for each group of two, three, four, or five values to determine the likelihood of obtaining a mean SI < 3. The

results for these four "borderline" HCA tests were then added to the results from an additional 12 robust positive control tests included in this laboratory's historical database to determine the overall likelihood of obtaining a mean SI < 3 for the positive control substance (**Table B-2**). The failure rate reflects the frequency with which a positive control test will fail, which would result in retesting the positive control and any concurrent test substances. Each laboratory is encouraged to determine the lowest number of animals to use in the positive control group based on the highest failure rate considered acceptable by the laboratory.

Table B-2 Example of Positive Control Failure Rate for 30% HCA Based on Data Collected in Single Laboratory

Number of Animals	HCA Test 1	HCA Test 2	HCA Test 3	HCA Test 4	Results from Other Tests[1]	Overall Likelihood of a Mean SI < 3
5	17% (1/6)	0% (0/6)	0% (0/6)	0% (0/6)	0% (0/72)	1% (1/96)
4	27% (4/15)	13% (2/15)	0% (0/15)	7% (1/15)	0% (0/180)	3% (7/240)
3	40% (8/20)	30% (6/20)	5% (1/20)	20% (4/20)	0% (0/240)	6% (19/320)
2	47% (7/15)	33% (5/15)	13% (2/15)	40% (6/15)	1% (1/180)	9% (21/240)

Abbreviations: HCA = hexyl cinnamic aldehyde; SI = stimulation index

[1] These represent 12 positive control studies in the same laboratory where all mice in the positive control groups treated with 30% HCA produced an SI ≥ 3.

Annex III: Evaluating Local Irritation and Systemic Toxicity in the Local Lymph Node Assay

As noted in the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Murine Local Lymph Node Assay (LLNA) protocol, at least three dose levels of a test substance should be evaluated. The highest dose level tested should be a concentration of 100% (i.e., neat substance for liquid substances) or the maximum soluble concentration (for solids), unless available information suggests that this concentration induces systemic toxicity or excessive local irritation after topical application.

In the absence of such information, a prescreen test should be performed using three dose levels of the test substance, in order to define the appropriate dose level to test in the LLNA. Six mice (two per concentration) are used, and the prescreen is conducted under identical conditions as the main LLNA study, except there is no assessment of lymph node proliferation. All mice will be observed daily for any clinical signs of systemic toxicity or local irritation at the application site. For example, observations might occur before and after treatment on Days 1, 2, and 3. Body weights are recorded pre-test and prior to termination (Day 6). Both ears of each mouse are observed for erythema (and scored using **Table B-3**). Ear thickness measurements are taken using a thickness gauge (e.g., digital micrometer or Peacock Dial thickness gauge) on Day 1 (pre-dose), Day 3 (approximately 48 hours after the first dose), and Day 6.

Excessive local irritation is indicated by an erythema score ≥3 and/or ear swelling of ≥25%.

Table B-3 Erythema Scores

Observation	Value
No visual effect	0
Slight erythema (barely perceptible)	1
Well-defined erythema	2
Moderate to severe erythema (beet redness)	3
Eschar (i.e., piece of dead tissue that is cast off from the surface of the skin)	4

A 25% increase in ear swelling has been used as an initial step to identify substances that cause a skin reaction due to an irritant response rather than sensitization (Reeder et al. 2007; ICCVAM 2008b). A statistically significant difference from control animals has also been used to delineate irritants from non-irritants in the LLNA (Hayes et al. 1998; Homey et al. 1998; Woolhiser et al. 1998; Hayes and Meade 1999; Ehling et al. 2005; Vohr and Jürgen 2005; Patterson et al. 2007). While these statistical differences often occur when ear swelling is less than 25%, they have not been associated specifically with excessive irritation (Woolhiser et al. 1998; Ehling et al. 2005; Vohr and Jürgen 2005; Patterson et al. 2007). Additionally, an adequately robust statistical comparison would require that a vehicle control group be included and that more than two animals per group be tested. Both of these requirements would substantially increase the number of animals used for this prescreen test. For this reason, a threshold increase in ear swelling above pre-dosing levels is recommended for this prescreen test.

Test guidelines for assessing acute systemic toxicity recommend a number of clinical observations for assessing systemic toxicity (OECD 1987; EPA 1998). The following observations, which are based on test guidelines and current practices (ICCVAM in press), may indicate systemic toxicity when used as part of an integrated assessment and therefore may indicate that the maximum dose recommended for the LLNA has been exceeded:

- Clinical signs:
  - Changes in nervous system function (e.g., piloerection, ataxia, tremors, and convulsions)
  - Changes in behavior (e.g., aggressiveness, change in grooming activity, marked change in activity level)
  - Changes in respiratory patterns (i.e., changes in frequency and intensity of breathing such as dyspnea, gasping, and rales)
  - Changes in food and water consumption
  - Lethargy and/or unresponsiveness
  - Any clinical signs of more than slight or momentary pain and distress
- Reduction in body weight >10% from Day 1 to Day 6
- Mortality

Appendix C

Evaluating the Impact of Reducing the Sample Size from Five to Four Animals per Group on the Performance of the Ratio Rule of SI > 3 in LLNA Testing

[This Page Intentionally Left Blank]

1.0 Introduction

Test Guideline 429 issued by the Organisation for Economic Co-operation and Development (OECD; OECD 2002) states that "A minimum of four animals is used per dose group, with a minimum of three concentrations of the test substance, plus a negative control group treated only with the vehicle for the test substance, and a positive control, as appropriate. *In those cases in which individual animal data are to be collected, a minimum of five animals per dose group are used*." This analysis was undertaken to determine if the number of animals required for individual animal data collection could be harmonized with that required for pooled data without diminishing accuracy. This is important because most animal-use regulations require that the minimum number of animals be used in studies, which currently results in only pooled data being collected in many countries because it currently requires fewer animals.

Therefore, the issue under investigation in the evaluation that follows is the impact of modifying the murine local lymph node assay (LLNA) test method protocol by reducing the number of individual animals per group from 5 to 4. More specifically, the evaluation considers how often this reduction in animal usage would have an impact on the overall LLNA outcome when the decision criterion used to determine a sensitizer from a non-sensitizer is a stimulation index (SI) greater than or equal to 3 (i.e., the "Ratio Rule"). Since the "true" underlying sensitizer status for individual substances is generally not known, this investigation will focus on the degree of disagreement rather than on which observed outcome is the "correct" one. This evaluation focused primarily on the Ratio Rule, although the possible use of a formal statistical test will also be considered.

The results of the following analyses indicate that a reduction in the sample size from 5 to 4 animals per group is unlikely to have any significant impact on the results of the LLNA test when using the Ratio Rule. If using statistics, the power for detecting LLNA effects will be reduced slightly when using 4 animals per group relative to using 5 animals per group. However, the practical impact of this power difference may be minimal, in that the power difference appears to be small for detecting effects above the Ratio Rule cutoff point of SI = 3. Importantly, this analysis also indicates that a statistical test based on 4 animals per group will identify more sensitizers than using the Ratio Rule based on 5 animals per group.

2.0 Methods

The database evaluated includes three different strains of animals: CBA, BALB/c, and B6C3F1. This report evaluates in detail only the CBA database; the data from the other two strains are summarized (**Section 4.0** and **Table C-7**) and may be evaluated more definitively in due course. The CBA database consists of 83 individual studies, each with three or four dosed groups and a control group. There are not 83 distinct substances, because some substances are tested in multiple studies. The number of individual animals per group in these studies ranged from 2 to 9. There were a total of 277 dosed groups, two of which were excluded from the agreement-disagreement analysis since there were only 2 or 3 animals per group. Study results were evaluated on a dose-by-dose basis as well as on a study-by-study basis, recognizing that the doses within a study used a common control group. Also, for certain labs, a common control group was used for multiple substances.

For each study having 5 animals per group (i.e., N = 5), all possible random samples of size 4 (responses measured as disintegrations per minute [dpm] of a radiolabeled tracer compound)

were taken from both the control and experimental groups (25 possible combinations), and the results of the Ratio Rule were compared for each of the samples with that of the full data set of 5 animals. The level of agreement was then determined.

For those studies having more than 5 animals per group, a similar procedure was applied, but in this case random samples were taken for both the N = 5 and N = 4 protocols, and there were far more combinations of samples to deal with (8100 rather than 25). Once again, the level of agreement between the N = 5 and N = 4 protocols were determined.

3.0 Results

Using the Ratio Rule criterion, the CBA mouse database consisted of a mix of sensitizers (49 studies) and non-sensitizers (33 studies), with one study (discussed in more detail below) producing a borderline effect. **Table C-1** shows the frequency of the various SI values in the 275 usable (for agreement-disagreement analysis) dosed groups, together with the average agreement seen between samples of N = 5 and N = 4. As can be seen in the table, the disagreement in study results is limited to SIs in the 2.1 to 4.7 range, with the disagreement increasing as the SI approaches 3. The overall average agreement between N = 4 and N = 5 studies is quite good: 97.5%. Moreover, as discussed in more detail below, the disagreement in outcome is due primarily to the inherent variability in the data (and the closeness of the SI to 3), not to the reduction in sample size.

The individual study results for the CBA strain are summarized in **Annex I**.

Although the primary focus of this evaluation is on the Ratio Rule (i.e., SI > 3), it is possible that a formal statistical test may be used in addition to (or possibly even in place of) the Ratio Rule. For this reason, a simple Student's *t* test (based on the logged dpm data) was also used to compare each dosed group with its concurrent control. The results of this analysis are summarized in **Table C-2**. It is clear that using a formal statistical test will identify far more "positives" than the Ratio Rule, i.e., statistical significance ($p < 0.05$) was achieved for some dosed groups producing an SI well below 3. This matter is discussed in more detail below.

Table C-1 Breakdown of Individual Dosed Group SIs: CBA Strain

SI	Frequency	Agreement between N = 5 and N = 4 samples
<2.1	154	100.00%
2.1 – 2.5	16	90.10%
2.6	2	85.00%
2.7	3	73.30%
2.8	2	64.00%
3.1	1	56.00%
3.2	2	55.50%
3.3	4	73.50%
3.4	1	88.00%
3.5	1	68.00%
3.6	1	84.00%
3.7	1	90.00%
3.8	1	100.00%
4.0 – 4.7	16	97.90%
>4.7	70	100.00%
Total	**275**	**97.50%**

Abbreviations: N = number of animals per dose group; SI = stimulation index

Table C-2 Distribution of Statistically Significant ($p < 0.05$) SIs: CBA Strain

SI	Frequency	Percentage of statistically significant ($p < 0.05$) SIs
<1.7	131	0.00%
1.7 – 1.9	23	52.20%
2.0 – 2.5	17	88.00%
2.6 – 3.0	7	85.70%
> 3.0	1	100.00%
Total	**277**	

Abbreviation: SI = stimulation index

4.0 Discussion

It was known in advance that the reduction in sample size from N = 5 to N = 4 would have essentially no impact on study results for "strong sensitizers" and for "clear non-sensitizers," and this is confirmed in **Table C-1**. What was not known was (1) how frequently such outcomes are seen in practice; (2) the specific range of SI values in which some impact on study outcome may be evident; (3) the magnitude of the impact for those studies having an SI close to 3; and (4) whether the disagreement in study outcome was due primarily to the reduction in sample size or to the inherent variability in the data (and the closeness of the SI to 3). The current investigation addresses all of these issues.

With regard to the first issue, for the CBA mouse database, only 34 of the 275 dosed groups (12%) had less than 100% agreement between N = 5 and N = 4 outcomes. Thus, for most dosed groups, the reduced sample size will not even be an issue when using the Ratio Rule.

Moreover, the reduced sample size becomes an issue only for a relatively narrow range of SI values. The range of SI values in this database producing less than 100% agreement was 2.1 to 4.7, but this may be somewhat misleading in that many studies in this range produced 100% agreement (see **Table C-1** and **Annex I**).

As the SI approaches 3, the disagreement between a sample of N = 5 and N = 4 increases notably (**Table C-1**). However, and this may be the single most important "take home" message of this entire analysis, the disagreement is far more a function of the animal-to-animal variability than it is to the reduction in sample size. That is, a second sample of 5 animals would show almost the same level of disagreement with the first sample of 5 animals, as would a sample of 4 animals. Thus, the reduction in sample size is a relatively small contributor to this difference. This important concept is illustrated below with two examples from the CBA mouse database, the first showing an SI of 2.8 **(Table C-3)**, just below the Ratio Rule threshold of SI = 3, the second showing an SI of 3.2 **(Table C-4)**, just above the Ratio Rule threshold.

The first example is the high dose of the third hexyl cinnamic aldehyde study, which had an SI of 2.8 for N = 6 **(Table C-3)**. This is the one study noted above with a borderline effect. Since N = 6, this required selection of samples of size 5 from both the control and dosed groups, and some of these samples did not give the same result as that seen for the full six animal sample. The results are summarized below and compared with the N = 4 strategy.

Table C-3 Example Showing Effect of Sample Size on Agreement of Results for a Test Substance with SI = 2.8

	Two N = 5 samples	**One N = 5 sample and one N = 4 sample**
Agreement (SI > 3)	7.7% (10/36) (10/36)	10.5% (10/36) (85/225)
Agreement (SI < 3)	52.2% (26/36) (26/36)	44.9% (26/36) (140/225)
Disagreement (one SI > 3; one SI < 3)	40.1% (by subtraction)	44.6% (by subtraction)

Abbreviations: N = number of animals per dose group; SI = stimulation index

As can be seen from these calculations (see also **Annex I**), the agreement between N = 5 and N = 4 strategies is "only" 55%. However, the disagreement is *not* due primarily to a reduction in sample size, since the agreement is very similar to that found for two N = 5 samples (60%). In other words, only 4.5% of the observed 45% disagreement is due to the reduction in sample size. The rest is due to the inherent variability among animals (and the closeness of the SI to 3) that would be evident even if a second sample of size 5 were used.

The second example is the mid-dose of the dipropylene triamine study, which had an SI of 3.2 also for N = 6 **(Table C-4)**. The results are summarized below and compared with the N = 4 strategy.

Table C-4 Example Showing Effect of Sample Size on Agreement of Results for a Test Substance with SI = 3.2

	Two N = 5 samples	**One N = 5 sample and one N = 4 sample**
Agreement (SI > 3)	56.25% (27/36) (27/36)	50.67% (27/36) (152/225)
Agreement (SI < 3)	6.25% (9/36) (9/36)	8.11% (9/36) (73/225)
Disagreement (one SI > 3; one SI < 3)	37.50% (by subtraction)	41.22% (by subtraction)

Abbreviations: N = number of animals per dose group; SI = stimulation index

The results are very similar to those of the first example, in that most of the 41% disagreement between the N = 4 sample and the N = 5 sample is due to the inherent variability of the data and the closeness of the SI to 3, not to the reduction in sample size.

Another point that should be noted: in the instances in which there is disagreement, the N = 4 strategy may actually have a higher likelihood of producing an SI > 3 result than using a sample of size 5. This occurs when the underlying SI is close to but below 3. For instance, consider the first example given above in which the observed SI = 2.8. A sample of size 4 would have a 38% chance (85/225) of producing an SI > 3 compared with only 28% (10/36) when using N = 5. In that sense, N = 4 could be regarded as having greater "power" than N = 5 for these data.

However, use of the Ratio Rule implicitly assumes that an SI less than 3 is biologically unimportant and thus should not be detected. Thus, the increased likelihood of exceeding the Ratio Rule criterion using N = 4 in the example above could be regarded as an increase in the false positive rate, rather than an increase in power. Importantly, as N increases, the likelihood of detecting SI = 2.8 by the Ratio Rule approaches zero, with maximum "power" occurring for N = 1.

However, some investigators may regard an SI of 2.8 as biologically important, especially if seen at the top dose, as was the case in this study. Consequently, these investigators might actually prefer the performance of N = 4 rather than N = 5 in this example. Of course, if SI < 3 responses are considered important, it would make far more sense to carry out a formal statistical test to detect them rather than using the Ratio Rule, which will likely not detect them. Although not detected by the Ratio Rule, the SI = 2.8 effect noted above in the high dose hexyl cinnamic aldehyde study is highly significant ($p < 0.01$) by Student's *t* test.

Moreover, it is likely that this particular SI = 2.8 is a "real" effect, not only because it is highly significant statistically, but also because in four other studies with this compound, the SIs produced for this dose were 2.2, 4.1, 4.2, and 6.6, with higher doses producing even greater effects (see **Annex I**). Without these additional studies, it is possible that this effect would be "missed" since SI = 2.8 does not satisfy the Ratio Rule criterion of SI > 3, and without individual animal data, it would not be possible to determine whether or not this effect was statistically significant. This is another illustration of the value of individual animal data and also the value of using a formal statistical test. It also shows that in some cases a sample of N = 4 is actually more likely to produce the "correct" conclusion than N = 5 when using the Ratio Rule.

As can be seen in **Table C-2**, a formal statistical test will identify as statistically significant ($p < 0.05$) many responses that would not be detected by the Ratio Rule. In some cases, statistical significance is achieved for SI values as low as 1.7 (see **Annex I** and **Table C-2**). Normally, this "increased power" would be considered very desirable, but apparently it is possible that certain SIs in the 1.7 to 3.0 range, while truly different from controls, may be reflecting "irritation" rather than a true sensitizing effect, and thus may not be indicative of a meaningful human risk. Discussion of this matter is beyond the scope of this investigation, but it is logical to assume that since the Ratio Rule is widely used for LLNA data, while a formal statistical test is not, there must be concern that a formal statistical test will produce too many "significant effects" for SIs in the 2 to 3 range. That is, SIs below 3 may be statistically significant and reflect "real" dosed group effects, but responses in this range are considered biologically unimportant. As can be seen in **Table C-2**, most of the SIs in the 2 to 3 range are in fact statistically significant. Use of the Ratio Rule also implicitly assumes that false positives are more important than false negatives.

Any consideration of statistical power must take into account the variability in response among animals. To illustrate this, consider the 17 CBA mouse studies carried out at BASF (see **Table C-11** in **Annex I**). The mean control dpm response across these 17 studies was 552.3. The mean standard deviation (SD; based on the logged dpm responses) among the control animals was 0.4077. Based on this information, we can carry out a power calculation, which is summarized in **Table C-5**.

To explain further: Power is primarily a function of (1) the magnitude of the difference between the dosed and control groups, (2) the underlying variability among animals, and (3) the sample size. In the table below, "difference" is the size (on a log scale) of the "fold increase" that is to be detected. The SD is the assumed underlying standard deviation among animals (on a log scale) as determined by the data from BASF (see **Table C-11** in **Annex I**). This SD is assumed to be the same in the dosed and control groups, an assumption consistent with the data from multiple labs obtained to date. Delta is the standardized (by SD) difference to be detected and is the key input variable into the power calculation program. The power calculations given below are based on a two-sided Student's *t* test, and assume an underlying normal distribution for the logged data. The specific power calculations were taken from http://www.danielsoper.com/statcalc/calc49.aspx. In this program "Cohen's *d*" is just the standardized difference, Delta. This is a very simple program to use, and alternative power calculations can easily be made.

Table C-5 Post-hoc Power Calculations Based on the BASF Control Data

	Dosed Group Increase Relative to Controls			
	3.5-fold	**3-fold**	**2.5-fold**	**2-fold**
Assumed control response	552.3	552.3	552.3	552.3
Log (Control response)	6.314	6.314	6.314	6.314
Dosed group response	1933.05	1656.90	1380.75	1104.60
Log (Dosed group response)	7.567	7.413	7.230	7.007
Difference (log scale)	1.253	1.099	0.916	0.693
Assumed SD (log scale)	0.4077	0.4077	0.4077	0.4077
Delta = Difference/SD	3.07	2.70	2.25	1.70
Power for N = 5	99.0%	96.4%	87.9%	65.8%
Power for N = 4	95.7%	89.8%	76.8%	53.0%

Abbreviations: N = number of animals per dose group; SD = standard deviation

From these calculations, the conclusion is that if the underlying variability among control animals is similar to that seen in an average BASF study, then there is an excellent chance that an underlying SI of 2.5 will be detected as statistically significant ($p < 0.05$), although this likelihood is higher for N = 5 (87.9%) than for N = 4 (76.8%). This power calculation is also consistent with the empirical results summarized in **Table C-2**. An underlying SI of 2.5 would almost certainly not be detected by the Ratio Rule, nor would one want it to be detected, since use of the Ratio Rule implicitly assumes that such an effect is of no consequence, as noted earlier.

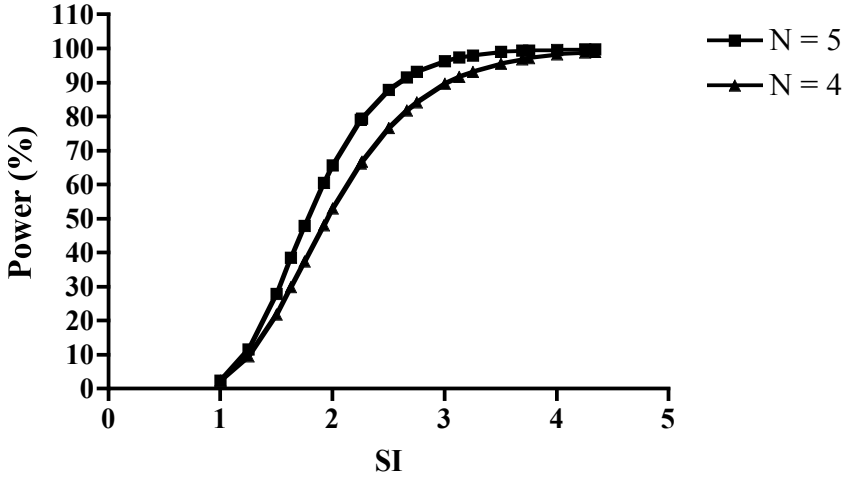
From the website given above, a general power curve can be constructed for N = 5 and N = 4 by specifying different values of Delta, which could reflect different "-fold increases (i.e., SI values)," different underlying variabilities, or a combination of these two factors. Such power comparisons are summarized below in **Table C-6** and **Figure C-1** and include the four from **Table C-5**.

Table C-6 Selected Power Comparisons for N = 5 and N = 4 Samples Based on BASF Control Data

SI	Delta	N = 5	N = 4
4.34	3.60	99.9%	99.1%
4.25	3.55	99.9%	98.9%
4.00	3.40	99.7%	98.3%
3.75	3.24	99.5%	97.2%
3.69	3.20	99.4%	96.9%
3.50	3.07	99.0%	95.7%
3.25	2.89	98.0%	93.3%
3.13	2.80	97.4%	91.8%
3.00	2.70	96.4%	89.8%
2.75	2.48	93.2%	84.3%
2.66	2.40	91.6%	81.9%
2.50	2.25	87.9%	76.8%
2.26	2.00	79.5%	66.8%
2.25	1.99	79.1%	66.3%
2.00	1.70	65.8%	53.0%
1.92	1.60	60.5%	48.2%
1.75	1.37	47.9%	37.4%
1.63	1.20	38.6%	30.0%
1.50	0.99	28.0%	21.9%
1.25	0.55	11.6%	9.7%
1.00	0.00	2.5%	2.5%

Abbreviations: N = number of animals per dose group; SI = stimulation index

Figure C-1 Power Curve for N = 5 and N = 4 Samples Based on BASF Control Data



Abbreviations: N = number of animals per dose group; SI = stimulation index

Although these particular "Deltas" could result from different combinations of –fold increases and assumed variability, the power calculations for the BASF data indicate that the most notable differences in power between N = 5 and N = 4 occur for SIs below 3, a range for which detection of an effect is apparently viewed as a "false positive" as discussed earlier. That is, the Ratio Rule implicitly assumes that SIs less than 3 should not be detected, so the fact that samples of N = 5 are more likely than samples with N = 4 to detect significant effects for SIs below 3 could be viewed as a disadvantage rather than an advantage of a larger sample size. For SI = 3.5 (at least for the BASF data), the power is high and similar for N = 5 and N = 4 (99.0% vs. 95.7%).

Note also from **Table C-6** that there will be far more sensitizers identified by a statistical test based on 4 animals per group than would be identified by the Ratio Rule using 5 animals per group. For example, a formal statistical test with N = 4 would have approximately 90% power for detecting an SI = 3, compared with only 50% power by using the Ratio Rule (regardless of N).

Although this report focuses on the large CBA mouse database, there are two smaller LLNA databases involving BALB/c and B6C3F1 mice. Although these other databases were not evaluated in detail, the pattern of LLNA response seen in these two strains was very similar to that seen in the CBA database. This comparison is summarized in **Table C-7** below. In this table, the percentage of positive studies is the percentage of studies having SI > 3 in at least one dosed group. As can be seen in **Table C-7**, there is little evidence of a strain difference in the pattern of LLNA response, and thus there is very little likelihood that a detailed evaluation of these other two strains would change the conclusions of this report.

Table C-7 Comparison of CBA, BALB/c, and B6C3F1 Databases

Strain	No. of Studies	No. of Doses	% Positive Studies	Distribution of SIs				
				<1.7	1.7 – 1.9	2.0 – 2.5	2.6 – 3.0	> 3.0
CBA	83	277	59 (49/83)	131 (47%)	23 (8%)	17 (6%)	7 (3%)	99 (36%)
BALB/c	41	133	63 (26/41)	67 (50%)	12 (9%)	8 (6%)	6 (5%)	40 (30%)
B6C3F1	10	28	70 (7/10)	15 (54%)	1 (4%)	1 (4%)	2 (7%)	9 (32%)

Abbreviation: No. = number; SI = stimulation index

There is one B6C3F1 mouse study that deserves special mention: the National Toxicology Program 2,4,5-trichlorophenoxyacetic acid study, which used a sample size of 6 animals per group. The top dose in this study produced a mean SI response of 3.03, which is the weakest "Ratio Rule positive" of any study in the three databases (control dpm responses were 63-69-75-90-119-133 compared with 213-229-244-249-325-405 in the top dosed group). The impact of reducing the sample size from 6 to 5 or 4 animals per group is summarized below.

Table C-8 Example Showing Effect of Sample Size on Agreement of Results for a Test Substance with SI = 3.03

	Two N = 5 samples	One N = 5 sample and one N = 4 sample
Agreement (SI > 3)	25.0% (18/36) (18/36)	26.4% (18/36) (119/225)
Agreement (SI < 3)	25.0% (18/36) (18/36)	23.6% (18/36) (106/225)
Disagreement (one SI > 3; one SI < 3)	50.0% (by subtraction)	50.0% (by subtraction)

Abbreviations: N = number of animals per dose group; SI = stimulation index

For these data, there is 50% disagreement between samples of size 4 and samples of size 5, but there is also 50% disagreement between two samples of size 5. This is a somewhat extreme example of the point made earlier, namely that most of the disagreement in Ratio Rule results observed between samples of size 5 and samples of size 4 shown in **Table C-1** is not due to the reduction in sample size, but rather due to the variability in response among animals and the closeness of the SI to the cutoff point of 3.

Finally, it is important to understand that **Table C-1** is not measuring accuracy; it is measuring agreement. That is, **Table C-1** assesses the reliability of N = 5 and N = 4 samples to produce the same classification outcome using the Ratio Rule; it does not assess the ability of N = 5 and N = 4 samples to produce the correct sensitizer classification (which for most substances is not known in any case). As illustrated in this report, as SI approaches 3, different samples may produce different classifications using the Ratio Rule, regardless of sample size, because of naturally occurring variability among animals. Importantly, most of the discordance between N = 5 and N = 4 samples shown in **Table C-1** is *not* due to the reduction in sample size.

With regard to accuracy of classification using the Ratio Rule, for 90% (75/83) of the CBA studies, there is no difference in accuracy using N = 5 and N = 4, based on the top dose group SI response. For eight studies, each with a top dose SI close to 3, there are slight differences in agreement, as shown in **Table C-9**.

Table C-9 Likelihood of SI > 3 for All CBA Studies Showing Less than Complete Agreement for the Top Dose Response Using N = 5 and N = 4 Samples

Substance	Top Dose SI	Likelihood of SI > 3 (%)	
		N = 5	N = 4
Formulation 54	2.3	0 (0/36)	7 (16/225)
Hexyl cinnamic aldehyde	2.8	28 (10/36)	38 (85/225)
Formulation 39	3.3	92 (33/36)	78 (175/225)
Bakelite EPR 161	3.5	83 (30/36)	77 (174/225)
Formulation 55	3.7	100 (36/36)	90 (202/225)
Potassium dichromate	4.1	100 (1/1)	92 (23/25)
Formulation 51	4.5[1]	100 (36/36)	96 (215/225)
1,6-(Bis(2-3-epoxypropoxy)hexane	4.7	100 (36/36)	94 (211/225)

Abbreviations: N = number of animals per dose group; SI = stimulation index
[1]Maximum response seen at mid-dose rather than top dose.

It is not known with certainty whether or not these eight substances are truly sensitizers. The one exception may be hexyl cinnamic aldehyde, which was confirmed in four other studies to be positive, with three showing SI > 4 at this dose. Thus, for this one compound the N = 4 sample may actually be more likely to be "accurate" than the N = 5 sample using the Ratio Rule.

If we assume that the Ratio Rule classifies all other substances correctly, and thus all six substances in **Table C-9** with SI > 3 are sensitizers, then there is a small loss in power by reducing the sample size per group from 5 to 4. However, this difference in power is small, and for all six substances, the likelihood is still quite high (77% - 96%) that the substance will be identified as a sensitizer using a sample of size 4. Recall also that these are "worst cases" and that for 90% of the CBA studies there is no difference in power at all between samples of N = 5 and N = 4. Thus, not only does the reduction in sample size from N = 5 to N = 4 have little impact on reliability using the Ratio Rule, it also appears to have little impact on the accuracy of classification.

5.0 Conclusion

For strong sensitizers and for obvious non-sensitizers, the reduction in sample size from 5 to 4 will have essentially no impact on the observed study outcome using the Ratio Rule. For those substances having an SI between (approximately) 2 and 4, the outcomes may be different, especially as SI approaches 3, but any such differences reflect primarily the inherent variability among animals and the closeness of the SI to 3 rather than the impact of reducing the sample size. Empirical examination of data from 83 CBA LLNA studies confirms that it is very unlikely that a reduction in sample size from 5 to 4 animals per group would have any impact on the overall interpretation of study results using the Ratio Rule.

Although the BALB/c and B6C3F1 databases were not evaluated in detail, the pattern of LLNA response seen in these strains is very similar to that seen in the larger CBA database, so a more definitive analysis of these other two strains would almost certainly not change the conclusions of this report. We conclude that a reduction in the sample size from 5 to 4 animals per group is unlikely to significantly impact the results of the LLNA test when using the Ratio Rule.

If a formal statistical test is used rather than (or in addition to) the Ratio Rule, the effect of reducing the sample size from N = 5 to N = 4 is to decrease the power slightly. However, for SI > 3, the power differences between samples of N = 5 and N = 4 are minimal. Moreover, a statistical test based on 4 animals per group will identify more sensitizers than using the Ratio Rule based on 5 animals per group. Thus, even if a formal statistical test is used rather than (or in addition to) the Ratio Rule, the practical impact of reducing the sample size from 5 to 4 animals per group on the interpretation of experimental results appears to be minimal.

Annex I:
Summary of Study Results – CBA Mouse Database

Table C-10 Experiments Conducted at ECPA Laboratories

Study[1]	Control N	Control Mean	Control SD	Experi-mental N	Experi-mental Mean	Experi-mental SD	SI	Agreement (%)[2]
Dincocap EC 0.8	5	175	50	5	471	198	2.7[3]	88 (22/25)
Dincocap EC 4.0	5	175	50	5	4007	1578	22.9[3]	100
Dincocap EC 10.0	5	175	50	4	7088	1863	40.5[3]	100[4]
Formaldehyde-1 1.0	5	163	59	5	125	12	0.8	100
Formaldehyde-1 5.0	5	163	59	5	208	147	1.3	100
Formaldehyde-1 20.0	5	163	59	5	781	439	4.8[3]	100
Formaldehyde-2 1.0	5	844	513	5	838	737	1.0	100
Formaldehyde-2 5.0	5	844	513	5	1824	1341	2.2	92 (23/25)
Formaldehyde-2 20.0	5	844	513	5	5188	2845	6.1[3]	100
HCA-1 3.0	5	430	154	5	571	153	1.3	100
HCA-1 10.0	5	430	154	5	955	368	2.2[3]	100
HCA-1 30.0	5	430	154	5	1870	376	4.3[3]	100
HCA-2 3.0	5	708	172	5	1353	649	1.9[3]	100
HCA-2 10.0	5	708	172	5	2981	1422	4.2[3]	100
HCA-2 30.0	5	708	172	5	6525	4014	9.2[3]	100
Oxyfluorfen EC 1	5	192	117	5	238	67	1.2	100
Oxyfluorfen EC 7	5	192	117	5	234	162	1.2	100
Oxyfluorfen EC 33	5	192	117	5	1043	311	5.4[3]	100
Potassium dichromate 0.02	5	153	84	5	260	139	1.7	100
Potassium dichromate 0.10	5	153	84	5	234	135	1.5	100
Potassium dichromate 0.50	5	153	84	5	626	390	4.1[3]	92 (23/25)
Quinoxyfen/ cyproconazole 7	5	226	86	5	283	102	1.3	100
Quinoxyfen/ cyproconazole 33	5	226	86	5	1470	276	6.5[3]	100
Quinoxyfen/ cyproconazole 100	5	226	86	5	3075	621	13.6[3]	100
Trifluralin EC 7	5	194	46	5	357	163	1.8[3]	100
Trifluralin EC 33	5	194	46	5	1585	349	8.2[3]	100
Trifluralin EC 100	5	194	46	5	3965	1456	20.5[3]	100

Abbreviations: EC = emulsion concentrate; ECPA = European Crop Protection Association; HCA = hexyl cinnamic aldehyde; N = number of animals per dose group; SD = standard deviation; SI = stimulation index

[1] Test substance and dose tested (%)

[2] Agreement (%) between N = 5 and N = 4 for the Ratio Rule. When agreement is less than 100%, numbers in parentheses indicate the proportion of the total number of N = 4 and N = 5 dose group combinations that agree with respect to whether SI < 3 or SI > 3. This is calculated by multiplying the proportion of N = 5 dose groups yielding SI > 3 with the proportion of N = 4 dose groups yielding SI > 3 and then adding the product of the proportion of N = 5 dose groups yielding SI < 3 with the proportion of N = 4 dose groups yielding SI < 3.

[3] These SIs are significantly different ($p < 0.05$) from 1 based on a Student's *t* test applied to the logged disintegrations per minute data.

[4] Although N = 4 for the experimental group, the responses in this particular group clearly would have shown 100% concordance between the outcomes for N = 5 and N = 4.

Table C-11 Experiments Conducted at BASF Laboratories

Study[1]	Control N	Control Mean	Control SD	Experi-mental N	Experi-mental Mean	Experi-mental SD	SI	Agreement (%)[2]
SC-1 3	6	626	216	6	511	124	0.8	100
SC-1 10	6	626	216	6	789	245	1.3	100
SC-1 30	6	626	216	6	1168	414	1.9[3]	100
HCA-3 2.5	6	1322	465	6	1479	161	1.1	100
HCA-3 5	6	1322	465	6	1571	921	1.2	100
HCA-3 10	6	1322	465	6	3749	1791	2.8[3]	55[4]
HCA-4 3	6	703	197	5	3209	1479	4.6[3]	100
HCA-4 10	6	703	197	6	4659	1409	6.6[3]	100
HCA-4 30	6	703	197	6	6929	1187	9.9[3]	100
HCA-5 10	5	176	26	5	711	240	4.1[3]	100
HCA-5 30	5	176	26	5	1362	611	7.8[3]	100
HCA-5 50	5	176	26	5	849	422	4.8[3]	100
1,6-Bis(2,3-epoxypropoxy)hexane 0.3	6	967	454	6	913	81	0.9	100
1,6-Bis(2,3-epoxypropoxy)hexane 1.0	6	967	454	6	1611	584	1.7	100
1,6-Bis(2,3-epoxypropoxy)hexane 3.0	6	967	454	6	4500	3061	4.7[3]	94 (211/225)
m-Phenylenebis (methylamine) 0.3	6	468	154	6	900	440	1.9[3]	100
m-Phenylenebis (methylamine) 1.0	6	468	154	6	4256	1298	9.1[3]	100
m-Phenylenebis (methylamine) 3.0	6	468	154	6	20691	6436	44.2[3]	100
Oxirane, mono((C12-14-alkyloxy) methyl) derivs 0.3	6	218	96	6	512	218	2.3[3]	92 (208/225)
Oxirane, mono((C12-14-alkyloxy) methyl) derivs 1.0	6	218	96	6	908	598	4.2[3]	92 (206/225)
Oxirane, mono((C12-14-alkyloxy) methyl) derivs 3.0	6	218	96	6	4963	1861	22.7[3]	100
1,2-Diaminocyclohexane 0.1	5	446	327	6	528	114	1.2	100
1,2-Diaminocyclohexane 0.3	5	446	327	6	810	290	1.8	100
1,2-Diaminocyclohexane 1.0	5	446	327	6	3736	1982	8.4[3]	100
Trimethylhexamine diamine 1.0	6	742	448	6	1599	400	2.2[3]	88[5]
Trimethylhexamine diamine 3.0	6	742	448	6	2972	1191	4.0[3]	93 (209/225)
Trimethylhexamine diamine 10.0	6	742	448	6	6581	1250	8.9[3]	100
1-(2,3-epoxypropoxy)-2,2-bis[(2,3-epoxypropoxy) methylbutane 1.0	6	388	310	6	797	392	2.1[3]	81[6]
1-(2,3-epoxypropoxy)-2,2-bis[(2,3-epoxypropoxy) methylbutane 3.0	6	388	310	6	2531	1812	6.5[3]	100

Study[1]	Control N	Control Mean	Control SD	Experi-mental N	Experi-mental Mean	Experi-mental SD	SI	Agreement (%)[2]
1-(2,3-epoxypropoxy)-2,2-bis[(2,3-epoxypropoxy) methylbutane 10.0	6	388	310	6	4644	2150	12.0[3]	100
3-Aminomethyl-3,5,5-trimethylcyclohexylamine 0.3	6	309	85	6	384	134	1.2	100
3-Aminomethyl-3,5,5-trimethylcyclohexylamine 1.0	6	309	85	6	806	248	2.6[3]	86[7]
3-Aminomethyl-3,5,5-trimethylcyclohexylamine 3.0	6	309	85	6	6597	1867	21.4[3]	100
Dipropylene triamine 0.3	6	349	101	6	753	228	2.2[3]	100
Dipropylene triamine 1.0	6	349	101	6	1106	254	3.2[3]	59[8]
Dipropylene triamine 3.0	6	349	101	6	4344	1350	12.4[3]	100
N-(2-Hydroxyethyl)-ethylendiamine 3.0	6	445	179	6	891	277	2.0[3]	100
N-(2-Hydroxyethyl)-ethylendiamine 10.0	6	445	179	6	766	230	1.7[3]	100
N-(2-Hydroxyethyl)-ethylendiamine 30.0	6	445	179	6	2937	626	6.6[3]	100
p-tert-Butylphenyl 1-(2,3-epoxy)propyl ether 0.1	6	406	83	6	553	148	1.4	100
p-tert-Butylphenyl 1-(2,3-epoxy)propyl ether 0.3	6	406	83	6	681	230	1.7[3]	100
p-tert-Butylphenyl 1-(2,3-epoxy)propyl ether 1.0	6	406	83	6	5780	3279	14.2[3]	100
Bakelite EPR 161 0.1	6	770	189	6	789	108	1	100
Bakelite EPR 161 0.3	6	770	189	6	1825	733	2.4[3]	99 (222/225)
Bakelite EPR 161 1.0	6	770	189	6	2694	1652	3.5[3]	68[9]
Bakelite EPR 162 0.3	6	591	251	6	6225	3285	10.5[3]	100
Bakelite EPR 162 1.0	6	591	251	6	11790	4292	19.9[3]	100
Bakelite EPR 162 3.0	6	591	251	6	23583	3469	39.9[3]	100
Bakelite EPR 164 0.3	6	463	208	6	2920	1049	6.3[3]	100
Bakelite EPR 164 1.0	6	463	208	6	8427	1833	18.2[3]	100
Bakelite EPR 164 3.0	6	463	208	6	10387	7000	22.4[3]	100

Abbreviations: EPR = epoxy resin; N = number of animals per dose group; SC = suspension concentrate; SD = standard deviation; SI = stimulation index

[1] Test substance and dose tested (%)

[2] Agreement (%) between N = 5 and N = 4 for the Ratio Rule. When agreement is less than 100%, numbers in parentheses or footnoted indicate the proportion of the total number of N = 4 and N = 5 dose group combinations that agree with respect to whether SI < 3 or SI > 3. This is calculated by multiplying the proportion of N = 5 dose groups yielding SI > 3 with the proportion of N = 4 dose groups yielding SI > 3 and then adding the product of the proportion of N = 5 dose groups yielding SI < 3 with the proportion of N = 4 dose groups yielding SI < 3.

[3] These SIs are significantly ($p < 0.05$) different from 1 based on a Student's *t* test applied to the logged disintegrations per minute data.

[4] 55% = (26/36 x 140/225) + (10/36 x 85/225)

[5] 88% = (35/36 x 204/225) + (1/36 x 21/225)

[6] 81% = (33/36 x 195/225) + (3/36 x 30/225)

[7] 86% = (35/36 x 198/225) + (1/36 x 27/225)

[8] 59% = (27/36 x 152/225) + (9/36 x 73/225)

[9] 68% = (30/36 x 174/225) + (6/36 x 51/225)

Table C-12 Experiments Conducted at DuPont Laboratories

Study[1]	Control N	Control Mean	Control SD	Experimental N	Experimental Mean	Experimental SD	SI	Agreement (%)[2]
DU-1A 5	5	506	185	5	284	122	0.6	100
DU-1A 25	5	506	185	5	596	166	1.2	100
DU-1A 50	5	506	185	5	354	198	0.7	100
DU-1A 100	5	506	185	5	526	313	1.0	100
DU-1B 1	5	1067	301	5	635	202	0.6	100
DU-1B 5	5	1067	301	5	1165	386	1.1	100
DU-1B 10	5	1067	301	5	1413	1145	1.3	100
DU-1B 25	5	1067	301	5	1144	388	1.1	100
DU-1C 5	5	617	265	5	419	156	0.7	100
DU-1C 25	5	617	265	4	883	517	1.4	100[3]
DU-1C 50	5	617	265	5	1075	432	1.7	100
DU-1C 100	5	617	265	4	779	262	1.3	100[3]
DU-1D 5	5	1067	301	5	755	196	0.7	100
DU-1D 10	5	1067	301	5	1019	266	1.0	100
DU-1D 25	5	1067	301	5	1337	493	1.3	100
DU-1D 50	5	1067	301	4	1086	281	1.0	100[3]
DU-2A 5	5	992	446	5	4132	815	4.2[4]	100
DU-2A 25	5	992	446	5	5422	939	5.5[4]	100
DU-2A 50	5	992	446	5	6604	1282	6.7[4]	100
DU-2A 100	5	992	446	5	6482	724	6.5[4]	100
DU-2E 5	5	452	219	5	433	169	1.0	100
DU-2E 25	5	452	219	5	370	142	0.8	100
DU-2E 50	5	452	219	5	509	285	1.1	100
DU-2E 100	5	452	219	5	623	200	1.4	100
DU-3 5	5	917	533	5	531	231	0.6	100
DU-3 10	5	917	533	5	720	306	0.8	100
DU-3 25	5	917	533	5	699	174	0.8	100
DU-3 50	5	917	533	5	538	179	0.6	100
DU-4 5	5	516	114	5	439	203	0.9	100
DU-4 25	5	516	114	5	505	257	1.0	100
DU-4 50	5	516	114	5	500	200	1.0	100
DU-4 100	5	516	114	5	538	65	0.9	100
DU-5A 5	5	589	317	5	1576	504	2.7[4]	76 (19/25)
DU-5A 25	5	589	317	5	903	534	1.5	100
DU-5A 50	5	589	317	5	915	223	1.6	100
DU-5A 100	5	589	317	5	499	230	0.8	100
DU-5B 5	5	1057	256	5	835	406	0.8	100
DU-5B 25	5	1057	256	5	1168	352	1.1	100
DU-5B 50	5	1057	256	5	1087	200	1.0	100
DU-5B 100	5	1057	256	5	1200	394	1.1	100

Study[1]	Control N	Control Mean	Control SD	Experimental N	Experimental Mean	Experimental SD	SI	Agreement (%)[2]
DU-5C 1	5	354	140	5	491	136	1.4	100
DU-5C 5	5	354	140	5	692	313	2.0[4]	100
DU-5C 25	5	354	140	5	429	195	1.2	100
DU-5C 100	5	354	140	5	312	124	0.9	100
DU-6 5	4	468	290	5	503	300	1.1	100[3]
DU-6 25	4	468	290	5	381	106	0.8	100[3]
DU-6 50	4	468	290	5	400	176	0.9	100[3]
DU-6 80	4	468	290	5	440	211	0.9	100[3]
DU-7 5	5	721	191	5	1394	1154	1.9	100
DU-7 25	5	721	191	5	846	331	1.2	100
DU-7 50	5	721	191	5	817	286	1.1	100
DU-7 80	5	721	191	5	915	249	1.3	100
DU-8A 1	9	486	186	4	680	178	1.4	100[3]
DU-8A 10	9	486	186	5	658	261	1.4	100
DU-8A 50	9	486	186	4	391	184	0.8	100[3]
DU-8A 100	9	486	186	5	473	263	1.0	100
DU-8B 5	5	786	312	5	916	460	1.2	100
DU-8B 25	5	786	312	5	1515	621	1.9	100
DU-8B 50	5	786	312	5	1121	764	1.4	100
DU-8B 100	5	786	312	5	1422	921	1.8	100
DU-9A 5	5	677	307	5	2405	1569	3.6[4]	84 (21/25)
DU-9A 25	5	677	307	5	3354	1463	5.0[4]	100
DU-9A 50	5	677	307	5	5975	773	8.8[4]	100
DU-9A 100	5	677	307	5	9118	3211	13.5[4]	100
DU-9B 5	5	1049	285	5	809	362	0.8	100
DU-9B 25	5	1049	285	5	822	195	0.8	100
DU-9B 50	5	1049	285	5	622	242	0.6	100
DU-9B 100	5	1049	285	5	493	88	0.5	100
DU-10 0.5	5	177	67	5	174	25	1.0	100
DU-10 1.0	5	177	67	5	230	73	1.3	100
DU-10 2.5	5	177	67	5	265	55	1.5	100
DU-10 5.0	5	177	67	3	289	122	1.6	NC[5]
DU-11B 5	5	984	210	5	1362	561	1.4	100
DU-11B 25	5	984	210	5	639	449	0.6	100
DU-11B 50	5	984	210	5	651	531	0.7	100
DU-11B 100	5	984	210	5	1016	1032	1.0	100
DU-11C 5	5	769	310	5	1168	472	1.5	100
DU-11C 25	5	769	310	5	871	217	1.1	100
DU-11C 50	5	769	310	5	719	133	0.9	100
DU-11C 100	5	769	310	5	1113	300	1.4	100
DU-12 1	5	617	265	5	479	132	0.8	100

Study[1]	Control N	Control Mean	Control SD	Experimental N	Experimental Mean	Experimental SD	SI	Agreement (%)[2]
DU-12 5	5	617	265	5	749	378	1.2	100
DU-12 25	5	617	265	5	477	253	0.8	100
DU-12 50	5	617	265	5	872	497	1.4	100
DU-13A 5	5	621	455	5	284	67	0.5	100
DU-13A 25	5	621	455	5	276	93	0.4	100
DU-13A 50	5	621	455	5	322	167	0.5	100
DU-13A 100	5	621	455	5	370	56	0.6	100
DU-13B 1	5	578	161	5	703	450	1.2	100
DU-13B 10	5	578	161	5	551	179	1.0	100
DU-13B 50	5	578	161	5	413	117	0.7	100
DU-13B 100	5	578	161	5	376	201	0.7	100

Abbreviations: DU = DuPont; N = number of animals per dose group; NC = not calculated; SD = standard deviation; SI = stimulation index

[1] Test substance and dose tested (%)

[2] Agreement (%) between N = 5 and N = 4 for the Ratio Rule. When agreement is less than 100%, numbers in parentheses indicate the proportion of the total number of N = 4 and N = 5 dose group combinations that agree with respect to whether SI < 3 or SI > 3. This is calculated by multiplying the proportion of N = 5 dose groups yielding SI > 3 with the proportion of N = 4 dose groups yielding SI > 3 and then adding the product of the proportion of N = 5 dose groups yielding SI < 3 with the proportion of N = 4 dose groups yielding SI < 3.

[3] Although N = 4 for the experimental group, the responses in this particular group clearly would have shown 100% concordance between the outcomes for N = 5 and N = 4.

[4] These SIs are significantly ($p < 0.05$) different from 1 based on a Student's *t* test applied to the logged disintegrations per minute data.

[5] Agreement could not be assessed, since N < 4.

Table C-13 Experiments Conducted at EFfCI Laboratories

Study[1]	Control N	Control Mean	Control SD	Experimental N	Experimental Mean	Experimental SD	SI	Agreement (%)[2]
Fumaric Acid 5	5	327	85	5	419	126	1.3	100
Fumaric Acid 10	5	327	85	5	742	284	2.3[3]	100
Fumaric Acid 25	5	327	85	5	479	201	1.5	100
Linoleic Acid 10	5	223	133	5	326	176	1.5	100
Linoleic Acid 25	5	223	133	5	1567	303	7.0[3]	100
Linoleic Acid 50	5	223	133	5	2025	601	9.1[3]	100
Linoleic Acid 10	5	223	133	5	699	301	3.1[3]	56 (14/25)
Linoleic Acid 25	5	223	133	5	2075	344	9.3[3]	100
Linoleic Acid 50	5	223	133	5	2290	1174	10.3[3]	100
Maleic Acid 10	5	327	85	5	2186	934	6.7[3]	100
Maleic Acid 25	5	327	85	5	5262	686	16.1[3]	100
Maleic Acid 50	5	327	85	5	5244	2304	16.0[3]	100
Octinol 10	5	1120	512	5	6327	1446	5.6[3]	100
Octinol 25	5	1120	512	5	9833	2523	8.8[3]	100
Octinol 50	5	1120	512	4	12594	1250	11.2[3]	100[4]
Oleic Acid 10	5	223	133	5	581	408	2.6[3]	84 (21/25)
Oleic Acid 25	5	223	133	5	3336	1688	14.9[3]	100
Oleic Acid 50	5	223	133	5	1550	897	6.9[3]	100
Squalene 10	5	223	133	5	839	245	3.8[3]	100
Squalene 25	5	223	133	5	1536	209	6.9[3]	100
Squalene 50	5	223	133	5	1821	327	8.2[3]	100
Succinic Acid 5	5	327	85	5	376	146	1.1	100
Succinic Acid 10	5	327	85	5	407	113	1.2	100
Succinic Acid 25	5	327	85	5	420	243	1.3	100
Undecylenic Acid 10	5	223	133	5	556	140	2.5[3]	80 (20/25)
Undecylenic Acid 25	5	223	133	5	736	250	3.3[3]	84 (21/25)
Undecylenic Acid 50	5	223	133	5	991	149	4.4[3]	100

Abbreviations: EFfCI = European Federation for Cosmetics Ingredients; N = number of animals per dose group; SD = standard deviation; SI = stimulation index

[1] Test substance and dose tested (%)

[2] Agreement (%) between N = 5 and N = 4 for the Ratio Rule. When agreement is less than 100%, numbers in parentheses indicate the proportion of the total number of N = 4 and N = 5 dose group combinations that agree with respect to whether SI < 3 or SI > 3. This is calculated by multiplying the proportion of N = 5 dose groups yielding SI > 3 with the proportion of N = 4 dose groups yielding SI > 3 and then adding the product of the proportion of N = 5 dose groups yielding SI < 3 with the proportion of N = 4 dose groups yielding SI < 3.

[3] These SIs are significantly ($p < 0.05$) different from 1 based on a Student's *t* test applied to the logged disintegrations per minute data.

[4] Although N = 4 for the experimental group, the responses in this particular group clearly would have shown 100% concordance between the outcomes for N = 5 and N = 4.

Table C-14 Experiments Conducted at BAuA Laboratories

Study[1]	Control N	Control Mean	Control SD	Experi-mental N	Experi-mental Mean	Experi-mental SD	SI	Agreement (%)[2]
Yellow E-JD 3442 1	5	70	21	5	70	19	1.0	100
Yellow E-JD 3442 3	5	70	21	5	52	9	0.8	100
Yellow E-JD 3442 9	5	70	21	5	60	32	0.9	100
Yellow E-JD 3442 15	5	70	21	5	61	16	0.9	100
CI Reactive Red 231 1	5	70	21	5	334	147	4.8[3]	100
CI Reactive Red 231 3	5	70	21	5	234	78	3.4[3]	88 (22/25)
CI Reactive Red 231 9	5	70	21	5	305	121	4.4[3]	100
CI Reactive Red 231 15	5	70	21	5	317	105	4.6[3]	100
P-46 1	5	70	21	5	167	86	2.4[3]	100
P-46 3	5	70	21	5	175	73	2.5[3]	96 (24/25)
P-46 9	5	70	21	5	135	39	1.9[3]	100
P-46 15	5	70	21	5	175	45	2.5[3]	100
CI Reactive Yellow 174 1	5	70	21	5	288	62	4.1[3]	100
CI Reactive Yellow 174 3	5	70	21	5	231	70	3.3[3]	80 (20/25)
CI Reactive Yellow 174 9	5	70	21	5	385	242	5.5[3]	100
CI Reactive Yellow 174 15	5	70	21	5	539	114	7.8[3]	100
Navy 14 08 723 1	5	70	21	5	353	54	5.1[3]	100
Navy 14 08 723 3	5	70	21	5	335	116	4.8[3]	100
Navy 14 08 723 9	5	70	21	5	398	102	5.7[3]	100
Navy 14 08 723 15	5	70	21	5	361	90	5.2[3]	100
Dispersionsrot 2754 1	5	70	21	5	68	27	1.0	100
Dispersionsrot 2754 3	5	70	21	5	65	19	0.9	100
Dispersionsrot 2754 9	5	70	21	5	67	40	1.0	100

Abbreviations: BAuA = Federal Institute for Occupational Safety and Health (Germany); N = number of animals per dose group; SD = standard deviation; SI = stimulation index

[1] Test substance and dose tested (%)

[2] Agreement (%) between N = 5 and N = 4 for the Ratio Rule. When agreement is less than 100%, numbers in parentheses indicate the proportion of the total number of N = 4 and N = 5 dose group combinations that agree with respect to whether SI < 3 or SI > 3. This is calculated by multiplying the proportion of N = 5 dose groups yielding SI > 3 with the proportion of N = 4 dose groups yielding SI > 3 and then adding the product of the proportion of N = 5 dose groups yielding SI < 3 with the proportion of N = 4 dose groups yielding SI < 3.

[3] These SIs are significantly ($p < 0.05$) different from 1 based on a Student's *t* test applied to the logged disintegrations per minute data.

Table C-15 Experiments Conducted at Dow AgroSciences Laboratories

Study[1]	Control N	Control Mean	Control SD	Experimental N	Experimental Mean	Experimental SD	SI	Agreement (%)[2]
Formulation 29 5	6	567	305	6	1036	663	1.8	100
Formulation 29 25	6	567	305	6	913	200	1.6	100
Formulation 29 100	6	567	305	6	823	373	1.5	100
Formulation 30 5	6	536	258	6	947	253	1.8[3]	100
Formulation 30 25	6	536	258	6	3839	736	7.2[3]	100
Formulation 30 100	6	536	258	6	7269	1014	13.6[3]	100
Formulation 31 5	6	385	121	5	393	223	1.0	100
Formulation 31 25	6	385	121	5	724	215	1.9[3]	100
Formulation 31 100	6	385	121	6	696	262	1.8[3]	100
Formulation 32 5	6	332	346	6	2136	737	6.5[3]	100
Formulation 32 25	6	332	346	6	14833	6139	44.7[3]	100
Formulation 32 100	6	332	346	6	22965	5480	69.3[3]	100
Formulation 33 5	6	672	249	6	479	194	0.7	100
Formulation 33 25	6	672	249	6	913	496	1.4	100
Formulation 33 100	6	672	249	6	843	303	1.3	100
Formulation 34 5	6	385	121	6	713	331	1.9	100
Formulation 34 25	6	385	121	6	528	227	1.4	100
Formulation 34 100	6	385	121	6	581	216	1.5	100
Formulation 35 5	6	332	346	6	360	294	1.1	100
Formulation 35 25	6	332	346	6	383	158	1.2	100
Formulation 35 100	6	332	346	6	412	317	1.3	100
Formulation 37 1	6	744	359	6	1008	525	1.4	100
Formulation 37 5	6	744	359	6	1999	1687	2.7	56[4]
Formulation 37 15	6	744	359	6	5586	4162	7.5[3]	100
Formulation 38 5	6	889	520	6	960	515	1.1	100
Formulation 38 25	6	889	520	6	4098	1541	4.6[3]	100
Formulation 38 100	6	889	520	6	11232	2102	12.7[3]	100
Formulation 39 1	6	627	256	6	1076	268	1.7[3]	100
Formulation 39 5	6	627	256	6	1551	650	2.5[3]	84[5]
Formulation 39 25	6	627	256	6	2083	259	3.3[3]	73[6]
Formulation 40 1	5	821[7]	263	6	1481	621	1.8	100
Formulation 40 5	5	821[7]	263	6	2316	401	2.8[3]	73 (55/75)
Formulation 40 25	5	821[7]	263	6	4646	1833	5.7[3]	100
Formulation 41 5	6	1017	325	6	1936	1024	1.9[3]	100
Formulation 41 25	6	1017	325	6	1891	1133	1.9	100
Formulation 41 100	6	1017	325	5	5653[7]	2750	5.6[3]	100
Formulation 49 5	5	626[7]	298	6	442	250	0.7	100

Study[1]	Control N	Control Mean	Control SD	Experimental N	Experimental Mean	Experimental SD	SI	Agreement (%)[2]
Formulation 49 25	5	626[7]	298	6	880	444	1.4	100
Formulation 49 100	5	626[7]	298	5	2958	489	4.7[3]	100
Formulation 50 5	6	1208	882	6	796	183	0.7	100
Formulation 50 25	6	1208	882	6	786	436	0.7	100
Formulation 50 100	6	1208	882	6	9439	4239	7.8[3]	100
Formulation 51 5	6	863	526	6	1346	537	1.6	100
Formulation 51 25	6	863	526	6	3893	2120	4.5[3]	96 (215/225)
Formulation 51 100	6	863	526	6	2084	1725	2.4	66[8]
Formulation 53 2.5	5	392[7]	159	6	596	317	1.5	100
Formulation 53 7.5	5	392[7]	159	6	1240	987	3.2[3]	52[9]
Formulation 53 15	5	392[7]	159	4	2609	1494	6.7[3]	100[10]
Formulation 54 5	6	438	143	6	551	357	1.3	100
Formulation 54 25	6	438	143	6	502	262	1.2	100
Formulation 54 100	6	438	143	6	1016	583	2.3	93 (209/225)
Formulation 55 5	6	529	238	6	781	602	1.5	100
Formulation 55 25	6	529	238	6	1348	947	2.5[3]	68[11]
Formulation 55 100	6	529	238	6	1972	758	3.7[3]	90 (202/225)
Formulation 56 5	6	529	238	6	1726	831	3.3[3]	57[12]
Formulation 56 25	6	529	238	6	3217	1996	6.1[3]	100
Formulation 56 100	6	529	238	2	2064	21	3.9[3]	NC[13]

Abbreviations: N = number of animals per dose group; NC = not calculated; SD = standard deviation; SI = stimulation index

[1] Test substance and dose tested (%)

[2] Agreement (%) between N = 5 and N = 4 for the Ratio Rule. When agreement is less than 100%, numbers in parentheses or footnoted indicate the proportion of the total number of N = 4 and N = 5 dose group combinations that agree with respect to whether SI < 3 or SI > 3. This is calculated by multiplying the proportion of N = 5 dose groups yielding SI > 3 with the proportion of N = 4 dose groups yielding SI > 3 and then adding the product of the proportion of N = 5 dose groups yielding SI < 3 with the proportion of N = 4 dose groups yielding SI < 3.

[3] These SIs are significantly ($p < 0.05$) different from 1 based on a Student's *t* test applied to the logged disintegrations per minute data.

[4] 56% = (26/36 x 142/225) + (10/36 x 83/225)

[5] 84% = (35/36 x 194/225) + (1/36 x 31/225)

[6] 73% = (33/36 x 175/225) + (3/36 x 50/225)

[7] Data reflects elimination of one control outlier (4258) in Formulation 40, one dosed group outlier (428) in Formulation 41, one control outlier (3) and one dosed group outlier (6273) in Formulation 49, and one control outlier (3172) in Formulation 53.

[8] 66% = (29/36 x 172/225) + (7/36 x 53/225)

[9] 52% = (4/6 x 42/75) + (2/6 x 33/75)

[10] Although N = 4 for the experimental group, the responses in this particular group clearly would have shown 100% concordance between the outcomes for N = 5 and N = 4.

[11] 68% = (31/36 x 168/225) + (5/36 x 57/225)

[12] 57% = (26/36 x 150/225) + (10/36 x 75/225)

[13] Agreement could not be assessed, since N < 4.

Appendix D

Final Background Review Document: Reduced Murine Local Lymph Node Assay

[This Page Intentionally Left Blank]

Background Review Document
Reduced Murine Local Lymph Node Assay

Interagency Coordinating Committee on the Validation of Alternative Methods

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods

National Institute of Environmental Health Sciences
National Institutes of Health
U.S. Public Health Service
Department of Health and Human Services

2009

National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709

[This Page Intentionally Left Blank]

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

ACD	Allergic contact dermatitis
ACE	Acetone
AOO	Acetone: olive oil (4:1 by volume)
BGIA	Berufsgenossenschaftliches Institut fur Arbeitsschutz (German Institute for Occupational Safety and Health)
BRD	Background review document
CASRN	Chemical Abstracts Service Registry Number
CESIO	Comite Europeen des Agents de Surface et de Leurs Intermediaires Organiques (European Committee of Surfactants and Their Organic Intermediates)
Conc.	Concentration tested
CPSC	U.S. Consumer Product Safety Commission
DBP	Dibutyl phosphate
DEP	Diethyl phthalate
DMF	Dimethyl formamide
DMSO	Dimethyl sulfoxide
DNCB	Dinitrochlorobenzene
EC3	Estimated concentration needed to produce a stimulation index of 3
ECPA	European Crop Protection Association
ECVAM	European Centre for the Validation of Alternative Methods
EFfCI	European Federation for Cosmetic Ingredients
EPA	U.S. Environmental Protection Agency
ESAC	European Centre for the Validation of Alternative Methods Scientific Advisory Committee
FDA	U.S. Food and Drug Administration
FR	*Federal Register*
GLP	Good Laboratory Practice
GPMT	Guinea Pig Maximization Test
HCA	Hexyl cinnamic aldehyde
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ILS	Integrated Laboratory Systems
ISO	International Organization for Standardization
IWG	Immunotoxicity Working Group
K_{ow}	Octanol-water partition coefficient
LLNA	Murine local lymph node assay

MEK	Methyl ethyl ketone
NA	Not applicable
NC	Not calculated
ND	No data
NICEATM	National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
NIEHS	National Institute of Environmental Health Sciences
NTP	National Toxicology Program
OECD	Organisation for Economic Co-operation and Development
OPPTS	Office of Prevention, Pesticides and Toxic Substances
PA	Pluronic acid
PG	Propylene glycol
RIFM	Research Institute for Fragrance Materials
rLLNA	Reduced murine local lymph node assay
SACATM	Scientific Advisory Committee on Alternative Toxicological Methods
SI	Stimulation index
SOT	Society of Toxicology
TG	Test guideline
TNO	TNO Nutrition and Food Research Institute (Netherlands)
U.K.	United Kingdom
U.N.	United Nations
U.S.	United States
w/v	Weight-to-volume ratio

[This Page Intentionally Left Blank]

Interagency Coordinating Committee on the Validation of Alternative Methods: Agency Representatives

Agency for Toxic Substances and Disease Registry
• Moiz Mumtaz, Ph.D.

Consumer Product Safety Commission
• Marilyn L. Wind, Ph.D. (Chair)
◊ Kristina Hatlelid, Ph.D.
Joanna Matheson, Ph.D.

Department of Agriculture
• Jodie Kulpa-Eddy, D.V.M. (Vice-Chair)
◊ Elizabeth Goldentyer, D.V.M.

Department of Defense
• Robert E. Foster, Ph.D.
◊ Patty Decot
Peter J. Schultheiss, D.V.M., D.A.C.L.A.M.
Harry Salem, Ph.D.

Department of Energy
• Michael Kuperberg, Ph.D.
◊ Marvin Stodolsky, Ph.D.

Department of the Interior
• Barnett A. Rattner, Ph.D.
◊ Sarah Gerould, Ph.D.

Department of Transportation
• George Cushmac, Ph.D.
◊ Steve Hwang, Ph.D.

Environmental Protection Agency
Office of Science Coordination and Policy
• Karen Hamernik, Ph.D.
Office of Research and Development
◊ Julian Preston, Ph.D.
Suzanne McMaster, Ph.D.
Office of Pesticide Programs
Amy Rispin, Ph.D.
Deborah McCall
OECD Test Guidelines Program
Jerry Smrchek, Ph.D.

• Principal agency representative
◊ Alternate principal agency representative

Food and Drug Administration
Office of Science
• Suzanne Fitzpatrick, Ph.D., D.A.B.T.
Center for Drug Evaluation and Research
◊ Abigail C. Jacobs, Ph.D.
Paul C. Brown, Ph.D.
Center for Devices and Radiological Health
Melvin E. Stratmeyer, Ph.D.
Vasant G. Malshet, Ph.D., D.A.B.T.
Center for Biologics Evaluation and Research
Richard McFarland, Ph.D., M.D.
Ying Huang, Ph.D.
Center for Food Safety and Nutrition
David G. Hattan, Ph.D.
Robert L. Bronaugh, Ph.D.
Center for Veterinary Medicine
Devaraya Jagannath, Ph.D.
M. Cecilia Aguila, D.V.M.
National Center for Toxicological Research
William T. Allaben, Ph.D.
Paul Howard, Ph.D.
Office of Regulatory Affairs
Lawrence D'Hoostelaere, Ph.D.

National Cancer Institute
• Alan Poland, M.D.
◊ T. Kevin Howcroft, Ph.D.

National Institute of Environmental Health Sciences
• William S. Stokes, D.V.M., D.A.C.L.A.M
◊ Raymond R. Tice, Ph.D.
Rajendra S. Chhabra, Ph.D., D.A.B.T.
Jerrold J. Heindel, Ph.D.

National Institute for Occupational Safety and Health
• Paul Nicolaysen, V.M.D.
◊ K. Murali Rao, M.D., Ph.D.

National Institutes of Health
• Margaret D. Snyder, Ph.D.

National Library of Medicine
• Pertti (Bert) Hakkinen, Ph.D.
◊ Jeanne Goshorn, M.S.

Occupational Safety and Health Administration
• Surender Ahir, Ph.D.

[This Page Intentionally Left Blank]

Acknowledgements

The following individuals are acknowledged for their contributions to the evaluation of the reduced LLNA.

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Immunotoxicity Working Group (IWG)

U.S. Consumer Product Safety Commission
Joanna Matheson, Ph.D. (IWG Co-Chair)
Marilyn Wind, Ph.D.

U.S. Environmental Protection Agency
Office of Pesticide Programs
Masih Hashim, D.V.M., Ph.D.
Marianne Lewis
Deborah McCall
Timothy McMahon, Ph.D.
Amy Rispin, Ph.D.
Office of Prevention, Pesticides, and Toxic Substances
Ronald Ward, Ph.D.
Office of Research and Development
Marsha Ward, Ph.D.
Office of Science Coordination and Policy
Karen Hamernik, Ph.D.

U.S. Food and Drug Administration
Center for Devices and Radiological Health
Daniel Lyle, Ph.D.
Vasant G. Malshet, Ph.D., D.A.B.T.
Jeffrey Toy, Ph.D.
Center for Drug Evaluation and Research
Paul C. Brown, Ph.D.
Abigail Jacobs, Ph.D. (IWG Co-Chair)
Jiaqin Yao, Ph.D.
Center for Veterinary Medicine
Ruth Barratt, Ph.D., D.V.M.

National Institute of Environmental Health Sciences
Dori Germolec, Ph.D.
William S. Stokes, D.V.M., D.A.C.L.A.M.
Raymond R. Tice, Ph.D.

National Institute for Occupational Safety and Health
B. Jean Meade, D.V.M., Ph.D.

European Centre for the Validation of Alternative Methods — Liaison
Silvia Casati, Ph.D.

Japanese Center for the Validation of Alternative Methods — Liaison
Hajime Kojima, Ph.D.

Reduced Murine Local Lymph Node Assay
Independent Scientific Peer Review Panel
(March 4–6, 2008)

Michael Luster, Ph.D. (Panel Chair)
Senior Consultant to the National Institute for Occupational Safety and Health
Health Effects Laboratory
Morgantown, WV

Nathalie Alépée, Ph.D.
Associate Research Fellow
Pfizer PDRD
MCT Laboratory
Amboise, France

Anne Marie Api, Ph.D.
Vice President, Human Health Sciences
Research Institute for Fragrance Materials
Woodcliff Lake, NJ

Nancy Flournoy, M.S., Ph.D.
Professor and Chair
Department of Mathematics and Statistics
University of Missouri – Columbia
Columbia, MO

Thomas Gebel, Ph.D.
Regulatory Toxicologist
Federal Institute for Occupational Safety & Health
Dortmund, Germany

Sidney Green, Ph.D.*
Professor of Pharmacology
College of Medicine
Howard University
Washington, DC

Kim Headrick, B.Admin., B.Sc.
International Harmonization and Senior Policy Advisor
Policy and Programme Service Office
Health Canada
Ottawa, Ontario, Canada

Dagmar Jírová, M.D., Ph.D.
Toxicologist, Research Manager
Head of Reference Center for Cosmetics and Reference Laboratory for Experimental Immunotoxicology
National Institute of Public Health
Prague, Czech Republic

David Lovell, Ph.D.
Reader in Medical Statistics
Postgraduate Medical School
University of Surrey
Guildford, Surrey, U.K.

Howard Maibach, M.D.
Professor, Department of Dermatology
University of California – San Francisco
San Francisco, CA

James McDougal, Ph.D.
Professor and Director of Toxicology Research
Department of Pharmacology and Toxicology
Boonshoft School of Medicine
Wright State University
Dayton, OH

Michael Olson, Ph.D.
Director of Occupational Toxicology
Corporate Environment, Health and Safety
GlaxoSmithKline
Research Triangle Park, NC

Raymond Pieters, Ph.D.
Associate Professor
Immunotoxicology Group Leader
Institute for Risk Assessment Sciences
Utrecht University
Utrecht, The Netherlands

Jean Regal, Ph.D.
Professor, Department of Pharmacology
University of Minnesota Medical School
Duluth, MN

Jonathan Richmond, MB ChB, FRCSEd*
Head, Animal (Scientific) Procedures Division
Home Office
London, U.K.

Peter Theran, V.M.D.
Consultant
Massachusetts Society for the Prevention of Cruelty to Animals
Novato, CA

Stephen Ullrich, Ph.D.
Professor of Immunology
Graduate School of Biomedical Sciences
University of Texas
M.D. Anderson Cancer Center – Houston
Houston, TX

Michael Woolhiser, Ph.D.
Technical Leader – Immunotoxicology
Senior Toxicology Specialist
Dow Chemical
Midland, MI

Takahiko Yoshida, M.D., Ph.D.
Professor, Department of Health Science
Asahikawa Medical College
Hokkaido, Japan

* Drs. Green and Richmond were unable to attend the public meeting on March 4–6, 2008; however, they took part in the scientific peer review and concur with the conclusions and recommendations included in the *Independent Scientific Peer Review Panel Report - Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products.*

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Institute of Environmental Health Sciences

William Stokes, D.V.M., D.A.C.L.A.M.
Director; Project Officer

Raymond Tice, Ph.D.
Deputy Director

Deborah McCarley
Special Assistant; Assistant Project Officer

NICEATM Support Contract Staff (Integrated Laboratory Systems [ILS], Inc.)

David Allen, Ph.D.

Thomas Burns, M.S.

Linda Litchfield

Michael Paris

Eleni Salicru, Ph.D.

Catherine Sprankle

Judy Strickland, Ph.D., D.A.B.T.

Other Acknowledgements

ICCVAM and NICEATM gratefully acknowledge the following individuals and institutions that submitted data to NICEATM used for the evaluation of the rLLNA.

David Basketter, Ph.D.
Unilever Safety and Environmental Assurance Centre
Sharnbrook, U.K.

Phil Botham, Ph.D.
European Crop Protection Association
Brussels, Belgium

Eric Debruyne, Ph.D.
Bayer CropScience SA, Sophia Antipolis
Cedex, France

G. Frank Gerberick, Ph.D.
Procter and Gamble Company
Cincinnati, OH, U.S.

Dori Germolec, Ph.D.
National Toxicology Program
Research Triangle Park, NC, U.S.

Ian Kimber, Ph.D.
Syngenta Central Toxicology Laboratory
Macclesfield, U.K.

Michael J. Olson, Ph.D.
GlaxoSmithKline
Research Triangle Park, NC, U.S.

K. Skirda
TNO Quality of Life
Delft, Netherlands

Peter Ungeheuer, Ph.D.
European Federation for Cosmetic Ingredients
Frankfurt, Germany

Hans Werner Vohr, Ph.D.
Bayer HealthCare
Wuppertal-Elberfeld, Germany

[This Page Intentionally Left Blank]

Preface

In 1998, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) evaluated the validation status of the murine local lymph node assay (traditional LLNA) as an alternative to guinea pig test methods (e.g., the Guinea Pig Maximization Test and the Buehler Test) for assessing the allergic contact dermatitis (ACD) potential of substances. ICCVAM subsequently recommended that the LLNA could be used as a valid substitute for the accepted guinea pig test methods in most ACD testing situations (ICCVAM 1999).

Based on the ICCVAM recommendations, the ICCVAM member agencies that require regulatory submission of ACD data accepted the LLNA, with identified limitations, as an alternative to guinea pig tests for assessing the potential of substances to cause ACD. In 2002, the LLNA was adopted as Test Guideline 429 by the 30 member countries of the Organisation for Economic Co-operation and Development (OECD; OECD 2002).

The reduced murine local lymph node assay (rLLNA), also referred to as the "cut-down" or "limit dose" LLNA, was one of several modified versions of the LLNA nominated by the U.S. Consumer Product Safety Commission (CPSC) for evaluation by ICCVAM.[30] (The term "reduced LLNA" has been adopted in this document to be consistent with the terminology used for this test method in Europe.) The proposed rLLNA could reduce the number of animals for skin sensitization testing by 40% for each test compared with the traditional LLNA. ICCVAM assigned this activity a high priority following consideration of comments from the public and ICCVAM's advisory committee, the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM).

The ICCVAM Immunotoxicity Working Group (IWG) and NICEATM (1) prepared a draft background review document (BRD) that described the validation status of the rLLNA test method, including its reliability and accuracy, the substances evaluated, and the availability of a standardized protocol and (2) developed draft test method recommendations based on this evaluation. An international independent scientific peer review panel (Panel) met on March 4–6, 2008, to assess the current validation status of the rLLNA. The Panel also reviewed the completeness and accuracy of the draft ICCVAM BRD and the extent to which the information therein supported the ICCVAM draft test method recommendations for proposed test method uses, recommended protocol, test method performance standards, and future studies.

ICCVAM considered the conclusions and recommendations of the Panel, as well as comments received from the public and SACATM, when finalizing ICCVAM's BRD and test method recommendations on the usefulness and limitations of the rLLNA.

We gratefully acknowledge the organizations and scientists who provided data and information for this BRD. We would also like to recognize the efforts of the individuals who contributed to its preparation, review, and revision. We especially recognize the Panel members for their thoughtful evaluations and generous contributions of time and effort. Special thanks are extended to Dr. Michael Luster for serving as the Panel Chair and to Dr. Michael Woolhiser, Dr. Michael Olson, and Ms. Kim Headrick for their service as Evaluation Group Chairs. We

[30] Available at http://iccvam.niehs.nih.gov/methods/immunotox/llnadocs/CPSC_LLNA_nom.pdf

thank the IWG for assuring a meaningful and comprehensive review. We especially thank Dr. Joanna Matheson (CPSC) and Dr. Abigail Jacobs (U.S. Food and Drug Administration Center for Drug Evaluation and Research) for serving as Co-Chairs of the IWG, as well as the IWG members and ICCVAM representatives who subsequently reviewed the BRD and provided comments.

Integrated Laboratory Systems, Inc., the NICEATM support contractor, provided excellent scientific and operational support for which we thank Dr. David Allen, Mr. Thomas Burns, Ms. Linda Litchfield, Mr. Michael Paris, Dr. Eleni Salicru, Ms. Catherine Sprankle, and Dr. Judy Strickland. We also acknowledge Dr. Raymond Tice, Deputy Director of NICEATM, for his contributions to this project. Finally, we want to thank Dr. Silvia Casati and Dr. Hajime Kojima, the IWG liaisons from the European Centre for the Validation of Alternative Methods and the Japanese Center for the Validation of Alternative Methods, respectively for their participation.

William S. Stokes, D.V.M., D.A.C.L.A.M.
Rear Admiral/Assistant Surgeon General, U.S. Public Heath Service
Director, NICEATM
Executive Director, ICCVAM

Marilyn Wind, Ph.D.
Deputy Associate Executive Director
Directorate for Health Sciences
U.S. Consumer Product Safety Commission
Chair, ICCVAM

Executive Summary

In 1999, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) recommended the murine local lymph node assay (traditional LLNA) as a valid substitute for currently accepted guinea pig test methods to assess allergic contact dermatitis (ACD) potential of substances in most ACD testing situations. The recommendation was based on a comprehensive evaluation that included an independent scientific peer review panel (Panel) assessment of the validation status of the LLNA. The Panel report and the ICCVAM recommendations (ICCVAM 1999) are available at the NICEATM–ICCVAM website.[31]

ICCVAM forwarded to U.S. Federal agencies its recommendation that the traditional LLNA should be considered for regulatory acceptance or other non-regulatory applications for assessing the ACD potential of substances, while recognizing that some testing situations would still require the use of traditional guinea pig test methods (ICCVAM 1999). The LLNA was subsequently incorporated into national and international test guidelines for the assessment of skin sensitization (International Organization for Standardization [ISO] 10993-10: Tests for Irritation and Sensitization [ISO 2002]; Organisation for Economic Co-operation and Development Test Guideline [TG] 429 [OECD 2002]; U.S. Environmental Protection Agency Health Effects Test Guideline OPPTS 870.2600: Skin Sensitization [EPA 2003]).

In 2007, the U.S. Consumer Product Safety Commission (CPSC) nominated the rLLNA (also referred to as the "cut-down" or "limit dose" LLNA) as one of several modified versions of the LLNA for evaluation by ICCVAM. The proposed rLLNA could reduce the number of animals for skin sensitization testing by 40% per test compared with the traditional LLNA. The term "reduced LLNA" has been adopted in this document to be consistent with the terminology used for this test method in Europe.

ICCVAM assigned this activity a high priority; and the National Toxicology Program Interagency Committee on the Evaluation of Alternative Methods (NICEATM), along with the ICCVAM Immunotoxicity Working Group (IWG), collaborated closely with liaisons from the European Centre for the Validation of Alternative Methods and the Japanese Center for the Validation of Alternative Methods to facilitate the evaluations requested by the CPSC. NICEATM and the ICCVAM IWG prepared this background review document (BRD), which summarizes the current validation status of the rLLNA for assessing the skin sensitization potential of substances. It includes detailed information about the reliability and relevance of the rLLNA, and the scope of the substances that were evaluated. It provides a comprehensive review of available data and information on the use of the rLLNA for hazard classification.

This information summarized in this BRD is from a retrospective review of traditional LLNA data. The database considered was obtained from 12 different sources and included 457 unique substances[32] tested in a total of 471 traditional LLNA studies. ICCVAM had considered 211 of the substances during its 1998 evaluation of the traditional LLNA (ICCVAM 1999). An additional 246 substances were obtained from the peer-reviewed literature published after that evaluation and from data submitted to NICEATM in response to a 2007 *Federal Register* (*FR*)

[31] Available at http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf

[32] Some substances were tested in more than one vehicle. In such instances, each substance-vehicle combination was considered separately, and thus there were a total of 465 unique substance-vehicle combinations that were used in the performance evaluation.

notice (72 FR 27815, May 17, 2007[33]). Specifically, three sources were published journal articles and eight were responses to the May 2007 *FR* notice. Due to the small number of repeated studies (5% of total studies), all studies were treated independently for the purpose of this accuracy evaluation.

The 1999 ICCVAM-recommended LLNA protocol accepted by U.S. regulatory agencies is consistent with procedures described in OECD TG 429 and was used as the basis for development of the OECD test guideline. Still, TG 429 allows for more procedural variation than the 1999 ICCVAM-recommended protocol (ICCVAM 1999). The protocol for the rLLNA is identical to that for the traditional LLNA (ICCVAM 1999), except that the traditional LLNA tests a substance at three dose levels, with the highest dose level being that which does not induce systemic toxicity and/or excessive skin irritation. In the rLLNA, a substance is tested at only a single dose level, which is the highest dose level that would have been tested in the traditional LLNA. As in the traditional LLNA, the threshold for classifying a substance as a skin sensitizer in the rLLNA is a stimulation index (SI) ≥ 3.

Information on chemical classes for each substance was retrieved from the National Library of Medicine's ChemIDplus® database or assigned for each test substance using a standard classification scheme based on the National Library of Medicine Medical Subject Headings classification system.[34] Chemical class information is included to indicate the variety of structural elements in the evaluated substances. One hundred and twenty-five complex substances were identified simply as pharmaceuticals. Ten substances were formulations. Seventy substances could not be assigned to a specific chemical class due to incomplete information (e.g., no Chemical Abstracts Service Registry Number or structure provided).

The ability of the rLLNA to correctly identify potential skin sensitizers was compared to that of the traditional LLNA. In the 471 studies, 318 detected skin sensitizers, and 153 detected non-sensitizers. When studies for substances tested more than once in the same vehicle (i.e., 465 unique substance and vehicle combinations) were considered together to yield an overall skin sensitization classification, 315 were classified as sensitizers, and 150 were classified as non-sensitizers.

Based on the data available from the 471 studies, the rLLNA has an accuracy of 98.7% (465/471), a sensitivity of 98.1% (312/318), a specificity of 100% (153/153), a false positive rate of 0% (0/153), and a false negative rate of 1.9% (6/318) when compared to the traditional LLNA. Based on the 465 unique substance and vehicle combinations, the rLLNA has an accuracy of 98.7% (459/465), a sensitivity of 98.1% (309/315), a specificity of 100% (150/150), a false positive rate of 0% (0/150), and a false negative rate of 1.9% (6/315).

Six substances yielded false negative results in the rLLNA (i.e., the substances were classified as sensitizers in the traditional LLNA but as non-sensitizers in the rLLNA). A review of the data for these six substances indicates that the traditional LLNA classification of the substances as skin sensitizers was based not on the highest dose level tested, which induced an SI < 3 but on a low- or mid-dose level that produced an SI ≥ 3. Because the rLLNA only tests substances at the highest dose level, all six substances would be incorrectly identified as non-sensitizers (i.e., false negatives). Four of the six substances that resulted in false negatives using the

[33] Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf
[34] Available at http://www.nlm.nih.gov/mesh/meshhome.html

rLLNA compared to the traditional LLNA came from LLNA studies that used pooled data. There were no patterns of consistency for these substances with regard to physicochemical properties.

Interlaboratory reproducibility of the rLLNA was assessed with data for five substances tested independently in the same vehicle at multiple laboratories. Among these five substances, three (60%) were classified as sensitizers or non-sensitizers in all studies (i.e., 100% concordance). Each of the other two substances, tested independently in two laboratories, was classified as a sensitizer by one traditional LLNA study and as a non-sensitizer by the other traditional LLNA study. Review of the studies indicates that the discordant results were due to differences in the highest dose levels tested. However, because the traditional LLNA and the rLLNA use identical protocols and the data sets used to evaluate their accuracy are similar, the reliability of the two methods would be expected to be similar. That is, the intra- and interlaboratory reliability of the rLLNA would be expected to be the same as that of the traditional LLNA (see ICCVAM 1999 for these statistics).

A review of published literature on the rLLNA revealed only one published report in addition to that of Kimber et al. (2006). Ryan et al. (2008) described the impact of reducing the number of animals per group from five to two on the performance of the rLLNA and concluded that the sensitivity is inadequate for hazard identification of skin sensitizers.

Compared to the traditional LLNA, the rLLNA will reduce the number of animals used to assess skin sensitization. Because the rLLNA tests only the highest dose level of the test substance in addition to the concurrent control groups, the number of animals tested would decrease by at least 40% for each test.

The database included in this BRD will be updated as additional information becomes available during future use of the traditional LLNA and the rLLNA.

[This Page Intentionally Left Blank]

1.0 Introduction and Rationale for the Proposed Use of the Reduced Murine Local Lymph Node Assay (rLLNA) to Identify Skin Sensitizers

1.1 Introduction

1.1.1 Historical Background

In 1999, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) recommended the murine local lymph node assay (traditional LLNA[35]) as a valid substitute for currently accepted guinea pig test methods to assess allergic contact dermatitis (ACD) potential of most types of substances. ICCVAM based its recommendation on a comprehensive evaluation that included an independent scientific peer review panel (Panel) assessment of the validation status of the LLNA. The Panel report and the ICCVAM recommendations (ICCVAM 1999) are available at the NICEATM–ICCVAM website.[36]

ICCVAM forwarded to U.S. Federal agencies its recommendation that the traditional LLNA should be considered for regulatory acceptance or other non-regulatory applications for assessing the ACD potential of substances, while recognizing that some testing situations would still require the use of traditional guinea pig test methods (ICCVAM 1999). The LLNA was subsequently incorporated into national and international test guidelines for the assessment of skin sensitization (International Organization for Standardization [ISO] 10993-10: Tests for Irritation and Sensitization [ISO 2002]; Organisation for Economic Co-operation and Development Test Guideline [TG] 429 [OECD 2002]; U.S. Environmental Protection Agency [EPA] Health Effects Test Guideline OPPTS 870.2600: Skin Sensitization [EPA 2003]).

1.1.2 Allergic Contact Dermatitis

ACD is a frequent occupational health problem. According to the U.S. Department of Labor Bureau of Labor Statistics, in 2005, 980 cases of ACD involved days away from work.[37]

ACD develops in two phases, induction and elicitation. The induction phase occurs when a susceptible individual is exposed topically to a skin-sensitizing substance. Induction depends on the substance passing through the epidermis, where it forms a hapten complex with dermal proteins. Langerhans cells, the resident antigen-presenting cells in the skin, process the hapten complex. The processed hapten complex then migrates to the draining lymph nodes. Antigen presentation to T-lymphocytes follows, which leads to the clonal expansion of these cells. At this point, the individual is sensitized to the substance (Basketter et al. 2003; Jowsey et al. 2006). Studies have shown that the magnitude of lymphocyte proliferation correlates with the extent to which sensitization develops (Kimber and Dearman 1991, 1996).

During the elicitation phase, the individual is again topically exposed to the substance. As in the induction phase, the substance penetrates the epidermis, is processed by the Langerhans cells, and is presented to circulating T-lymphocytes. The T-lymphocytes are then activated, which

[35] The "traditional LLNA" refers to the validated ICCVAM-recommended LLNA (ICCVAM 1999), which measures lymphocyte proliferation based on incorporation of tritiated thymidine into the cells of the draining auricular lymph nodes.
[36] Available at http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf
[37] Available at http://www.bls.gov/IIF

causes release of cytokines and other inflammatory mediators. This release produces a rapid dermal immune response that can lead to ACD (ICCVAM 1999; Basketter et al. 2003; Jowsey et al. 2006).

1.1.3 U.S. Consumer Product Safety Commission (CPSC) Nomination

On January 10, 2007, the CPSC formally requested that ICCVAM and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) evaluate several activities related to the LLNA.[38] The nominated activities included the following:

- The LLNA as a stand-alone assay for potency determination (including severity) for classification purposes
- Non-radioactive LLNA protocols
- The reduced LLNA (rLLNA) (also known as the ‘‘cut-down’’ or ‘‘limit dose’’ LLNA procedure)
- The use of the LLNA to test mixtures, aqueous solutions, and metals

ICCVAM unanimously agreed that the nominated activities should have a high priority for evaluation. ICCVAM's advisory committee, the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM), also recommended that the nominated activities be undertaken with a high priority.

As ICCVAM and NICEATM collaborate closely with the European Centre for the Validation of Alternative Methods (ECVAM) and the Japanese Center for the Validation of Alternative Methods, both organizations identified liaisons to the ICCVAM Immunotoxicity Working Group to facilitate the evaluations requested by the CPSC.

1.1.4 Description of the Reduced Murine Local Lymph Node Assay

Kimber and colleagues initially discussed the rLLNA in a 2006 publication (Kimber et al. 2006). The rLLNA was also discussed in two posters (Basketter et al. 2007; Chaney et al. 2007, subsequently published as Ryan et al. 2008) and one platform presentation (Basketter 2007) at the Society of Toxicology (SOT) Annual Meeting in Charlotte, NC, on March 25–29, 2007.

The protocol for the rLLNA is identical to that of the traditional LLNA (as described in the 1999 ICCVAM-recommended protocol) with one exception. In the traditional LLNA, three dose levels of each test substance are tested, while in the rLLNA only the highest dose level that does not induce systemic toxicity and/or excessive skin irritation is tested for skin-sensitizing activity (Kimber et al. 2006).

The term “limit dose,” sometimes used to refer to the rLLNA, accurately depicts a modified LLNA that tests only the highest dose level that does not induce local irritation and/or systemic toxicity. The terms “cut-down” and “reduced” LLNA also accurately describe the reduction in the number of doses tested and emphasize the reduction in the number of animals used to perform the test. For consistency with the terminology presented in the publications that first described this version of the LLNA, the term “reduced LLNA” (rLLNA) will be used.

[38] Available at http://iccvam.niehs.nih.gov/methods/immunotox/llnadocs/CPSC_LLNA_nom.pdf

1.1.5 Results of an ECVAM Peer Review of the rLLNA

The ECVAM Scientific Advisory Committee (ESAC) established a review panel to retrospectively analyze the published LLNA data to determine if limiting the number of test substance dose levels to only the highest dose level could successfully reduce the number of animals used per test. The review was based on the evaluation published by Kimber et al. (2006). At its semi-annual meeting on April 26–27, 2007, ESAC reviewed the rLLNA.

The ESAC statement on the rLLNA, dated April 27, 2007 (**Annex I**), states that:

> "… the peer reviewed and published information is of a quality and nature to support the use of the rLLNA within tiered-testing strategies to reliably distinguish between chemicals that are skin sensitisers and non-sensitisers, and that animal use can be minimised providing:
>
> - The concentration used to evaluate sensitisation potential is the maximum consistent with solubility and the need to avoid local and other systemic adverse effects, and that this principle rather than strict adherence to the specific recommended absolute concentrations as in OECD TG 429 should be used.
> - Negative test results associated with testing using concentrations of less than 10% should undergo further evaluation.
> - Positive and negative (vehicle) control groups are used, as appropriate, per OECD TG 429.
> - The full LLNA should be performed when it is known that an assessment of sensitisation potency is required."

The ESAC statement also recommends "that further work should be undertaken to determine if the 10% concentration threshold referenced above is optimal."

1.2 Regulatory Rationale and Applicability of the rLLNA

Current regulatory testing requires assessment of the potential skin sensitization hazard of regulated substances/products. The rLLNA is being considered for use in identifying skin sensitizers in a weight-of-evidence strategy such as that proposed in the United Nations Globally Harmonised System of Classification and Labelling of Chemicals (U.N. 2005). Unlike the traditional LLNA, the rLLNA evaluates the ability of a substance to be a sensitizer based on testing a single, highest-testable dose level; therefore, dose-response information is not generated. Thus, the rLLNA is being proposed for "yes/no" identification of sensitization hazards.

1.3 Scientific Basis for the rLLNA

1.3.1 Purpose and Mechanistic Basis

The purpose of the rLLNA is to identify potential skin sensitizers by quantifying lymphocyte proliferation in the draining auricular lymph nodes after application of a test substance to the ears of a mouse. The mechanistic basis is identical to that of the traditional LLNA (see **Section 1.1.2**).

1.3.2 Applicability Domain

The applicability domain of the rLLNA should be identical to that of the traditional LLNA. The traditional LLNA was not recommended for the testing of metals, mixtures/extracts, pharmaceuticals, or strong dermal irritants (ICCVAM 1999).

1.4 Test Method Validation

The ICCVAM Authorization Act of 2000 (Sec. 4(c)) mandates that "[e]ach Federal Agency … shall ensure that any new or revised … test method … is determined to be valid for its proposed use prior to requiring, recommending, or encouraging [its use]" (Public Law 106-545, 42 United States Code 285*l*-3).

Validation is the process by which the reliability and relevance of an assay for a specific purpose are established (ICCVAM 1997). *Relevance* is the extent to which an assay will correctly predict or measure the biological effect of interest (ICCVAM 1997). For the rLLNA, relevance is determined by how well the assay identifies (1) substances capable of producing skin sensitization in humans and (2) substances that should be assessed using a diverse set of substances that represent both of the types of chemical and product classes to be tested and the range of responses to be identified.

Reliability is the reproducibility of a test method within and among laboratories. The validation process provides data and information that allow U.S. Federal agencies to develop guidance on the use of test methods in evaluating the skin sensitization potential of substances.

The first stage in this evaluation is the preparation of a draft background review document (BRD) that comprehensively reviews the relevant data and information about a test method, including its mechanistic basis, proposed uses, reliability, and performance characteristics (ICCVAM 1997). The draft BRD is made available to the public and an independent scientific peer review panel (Panel) for review and comment. ICCVAM considers these comments and those of SACATM as they finalize the BRD. ICCVAM provides the final BRD to regulatory agencies for consideration as part of the ICCVAM Test Method Evaluation Report.

1.5 Selection of Citations for the rLLNA BRD

The test method data summarized in this BRD were obtained from the original LLNA evaluation (ICCVAM 1999), peer-reviewed scientific literature, the 2007 SOT Annual Meeting, and responses to a *Federal Register* (*FR*) notice requesting such data (72 FR 27815, May 17, 2007[39]). The terms "reduced LLNA," "cut-down LLNA," "limit dose LLNA," and "limit test LLNA" were used to search MEDLINE®, TOXLINE®, and Web of Science® for publications relevant to the rLLNA test method. A review of these databases through December 2007 revealed two published reports (Kimber et al. 2006; Ryan et al. 2008 [published online ahead of print as Ryan et al. 2007]). The rLLNA was also represented at the 2007 SOT Annual Meeting in two posters (Basketter et al. 2007; Chaney et al. 2007, subsequently published as Ryan et al. 2008) and one platform presentation (Basketter 2007).

[39]Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf

2.0 rLLNA Protocol Components

2.1 Overview

The technical aspects of the rLLNA are identical to those of the traditional LLNA; the two methods differ only in the number of test substance dose levels tested (Kimber et al. 2006). In the traditional LLNA, each test substance is tested at a minimum of three dose levels. The highest dose level is the maximum soluble concentration that does not cause systemic toxicity and/or excessive local irritation (ICCVAM 1999). In the rLLNA, in addition to the concurrent vehicle-control group, each test substance is tested at only the highest testable dose level (Kimber et al. 2006).

A Stimulation Index (SI) is calculated as the ratio of radioactivity incorporated into the cells of draining auricular lymph nodes of the treated animals to that of the vehicle-control animals. In both the traditional LLNA and the rLLNA, the threshold for classifying a substance as a skin sensitizer is an SI $\geq$ 3.

2.2 Basis for Test Method Selection

The rLLNA was proposed by Kimber et al. (2006) in an effort to reduce the number of animals used for skin sensitization testing and as a means of streamlining the LLNA for testing that will be required under the Registration, Evaluation and Authorisation of Chemicals regulations (Kimber et al. 2006).

2.3 Proprietary Test Method Components

The rLLNA does not employ any proprietary components.

2.4 Basis for the Number of Mice per Dose Group

The basis for the number of mice per dose group in the rLLNA is the same as that for the traditional LLNA (ICCVAM 1999).

2.5 Study Acceptance Criteria

Similar to the traditional LLNA, in order for an rLLNA study to be considered acceptable, the positive control must yield an SI $\geq$ 3 (ICCVAM 1999).

2.6 Basis for Selection of the Test Substance Dose

As noted in **Section 2.1**, the rLLNA tests each substance at only the highest testable dose level, in addition to the concurrent vehicle control. Consistent with the criteria for selecting the highest dose level in the traditional LLNA (ICCVAM 1999), the dose level used to evaluate sensitization potential in the rLLNA should be the maximum soluble concentration that does not cause systemic toxicity and/or excessive local irritation (ICCVAM 1999).

[This Page Intentionally Left Blank]

3.0 Substances Used for Validation of the rLLNA

3.1 Rationale for the Substances or Products Included in the Evaluation

Data from 471 LLNA studies were obtained from 12 sources (**Table D-1**), including published reports and unpublished data submitted to NICEATM in response to 72 FR 27815.[40]

3.2 Rationale for the Number of Substances Included in the Evaluation

The database from the 471 traditional LLNA studies included 457 unique substances,[41] 211 of which were included in the original ICCVAM evaluation of the traditional LLNA (ICCVAM 1999). Fourteen of the 457 unique substances were tested two to five times each in different LLNA studies. Specifically, nine of the 14 substances were evaluated two to five times in different vehicles, and five of the 14 substances were evaluated two to five times in the same vehicle. Two of the five substances evaluated in the same vehicle (hexyl cinnamic aldehyde [HCA] and potassium dichromate) were also tested using different vehicles (one study for HCA and two studies for potassium dichromate). Due to the small number of repeated studies (5% of total studies), all were treated independently for accuracy evaluation. When the studies for the substances repeated in the same vehicle were considered together to yield an overall skin sensitization classification, there were 465 studies with unique substance–vehicle combinations.

3.3 Detailed Description of Substances Included in the Evaluation

Annex II provides information on the physicochemical properties (e.g., physical form tested), Chemical Abstracts Service Registry Number (CASRN), and chemical class for each substance tested. This information was obtained from the published reports, submitted data, or literature searches.

When available, chemical classes for each substance were retrieved from the National Library of Medicine's ChemIDplus® database. If chemical class information was not located, chemical classes were assigned for each test substance using a standard classification scheme based on the National Library of Medicine Medical Subject Headings.[42] A substance could be assigned to more than one chemical class; however, no substance was assigned to more than three classes. Certain complex pharmaceuticals and pharmaceutical intermediates were simply identified as pharmaceutical substances. Chemical class information is presented only to indicate the variety of structural elements present in the substances evaluated in this analysis; it is not intended to evaluate the impact of structure on skin sensitization activity or potency.

[40] May 17, 2007, available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf

[41] Some substances were tested in more than one vehicle. In such instances, each substance–vehicle combination was considered separately, thus a total of 465 unique substance–vehicle combinations were evaluated.

[42] Available at http://www.nlm.nih.gov/mesh/meshhome.html

Table D-1 Summary of Traditional LLNA Data Sources and Rationale for Substance Selection

Data Source	Number of Studies	Primary Data Source and Substance Selection Rationale
Gerberick et al. (2005)[1]	210	Compiled from previously conducted studies (published literature and unpublished sources) on substances with varying skin sensitization potential
M.J. Olson/GlaxoSmithKline	124	Pharmaceuticals, pharmaceutical intermediates
Basketter, Gerberick, and Kimber[2]	31	Compiled from previously conducted studies (published literature and unpublished sources) on substances with varying skin sensitization potential
K. Skirda/CESIO (TNO Report V7217)	18	Data were provided by CESIO member companies for use in a paper titled "Limitations of the Local Lymph Node Assay (LLNA) as preferred test for skin sensitisation: concerns about false positive and false negative test results" (TNO report V7217)
Lalko and Api (2006)	17	Original research conducted on essential oils, which were representative of the oils commonly used in perfumery. Each contains significant amounts of one or more known skin sensitizers.
H.W. Vohr/BGIA	16	Original research with epoxy resin components as part of a validation effort for non-radioactive versions of the local lymph node assay
Ryan et al. (2002)	15	Original research with known water-soluble haptens and known skin sensitizers to assess the usefulness of a novel vehicle
D. Germolec/NIEHS	15	Substances evaluated by the National Toxicology Program for skin sensitization potential
E. Debruyne/Bayer CropScience SA	10	Original research on different pesticide types and formulations
P. Ungeheur/EFfCI	9	Data for selected unsaturated chemicals were provided in the report entitled "Comparative Experimental Study on the Skin Sensitising Potential of Selected Unsaturated Chemicals as Assessed by the Murine Local Lymph Node Assay (LLNA) and the Guinea Pig Maximisation Test (GPMT)"
P. Botham/ECPA	6	Plant protection products (i.e., pesticides) were evaluated in the local lymph node assay with a novel vehicle to assess its usefulness
Basketter et al., 2007	1	Original research that re-evaluated resorcinol in the local lymph node assay, which identified resorcinol as a sensitizer.
Total	**471**[3]	

Abbreviations: BGIA = Berufsgenossenschaftliches Institut fur Arbeitsschutz; CESIO = Comite Europeen des Agents de Surface et de Leurs Intermediaires Organiques; ECPA = European Crop Protection Association; EFfCI = European Federation for Cosmetic Ingredients; NIEHS = National Institute for Environmental Health Sciences: TNO = TNO Nutrition and Food Research

[1] These data were submitted to ICCVAM in 1998 for the original evaluation of the validation status of the LLNA (ICCVAM 1999) and were evaluated by the European Centre for the Validation of Alternative Methods (ECVAM) Scientific Advisory Committee in its evaluation of the rLLNA (Gerberick et al. 2005).

[2] Data were included in a submission to ECVAM for the validation of the traditional LLNA as a stand-alone assay for potency determination.

[3] The total number of studies does not take into account the fact that some substances were tested more than once (see **Section 3.2**)

Table D-2 provides chemical class information for the test substances in this rLLNA evaluation. The table distinguishes the chemical classifications of the 211 substances in the original evaluation of the rLLNA (Kimber et al. 2006; ESAC 2007) and the chemical classifications of the additional substances received in response to 72 FR 27815.[43] Of the 211 substances initially evaluated by Kimber et al. (2006), the known chemical classes with the greatest number of substances were carboxylic acids (29) and halogenated hydrocarbons (27). Of the additional 246 substances in this evaluation, the known chemical classes with the greatest number of substances tested were pharmaceutical chemicals (125), carboxylic acids (15), and lipids (14). Ten of the substances included in this evaluation were formulations. Seventy substances could not be assigned to a specific chemical class due to incomplete information (e.g., the lack of a CASRN or structure).

3.4 Coding Procedures

Neither the previous evaluation of these 211 substances (ICCVAM 1999) nor any additional studies used in this evaluation describe coding of substances to avoid potential scoring bias.

Table D-2 Chemical Classes[1] Represented in the Current Traditional LLNA Database

Chemical Class	Number of Substances - Original[2]	Number of Substances - Additional[2]	Chemical Class	Number of Substances - Original	Number of Substances - Additional
Alcohols	9	4	Inorganic Chemicals	0	2
Aldehydes	21	4	Isocyanates	1	0
Amides	4	0	Ketones	5	0
Amidines	1	0	Lactones	2	2
Amines	14	7	Lipids	7	14
Anhydrides	1	0	Macromolecular Substances[3]	0	5
Carbohydrates	3	2	Nitriles	1	1
Carboxylic Acids	29	15	Nitro Compounds	2	0
Esters	3	0	Nitroso Compounds	3	0
Ethers	14	2	Onium Compounds	1	0
Formulations[3]	0	10	Pharmaceutical chemicals[4]	0	125
Heterocyclic Compounds	18	4	Phenols	18	2
Hydrocarbons, Acyclic	2	1	Polycyclic Compounds	5	3
Hydrocarbons, Cyclic	14	7	Quinones	1	1
Hydrocarbons, Halogenated	27	1	Sulfur Compounds	20	2
Hydrocarbons, Other	7	8	Urea	3	0
Imines	0	1	Unknown	28	42

[1] Total number of substances assigned to chemical classes does not equal the total number of substances evaluated because some substances were assigned to more than one class and some substances were not assigned to a specific chemical class.
[2] Number of substances - original represents the substances evaluated in Kimber et al. (2006).
Number of substances - additional represents the substances received in response to 72 FR 27815 (May 17, 2007) (see below)
[3] No chemical class could be assigned. The terms "formulation" or "macromolecular substance" was used to identify these substances.
[4] The chemical classification of "pharmaceutical chemicals" for the GlaxoSmithKline (GSK) substances was suggested by Dr. Michael Olson of GSK to capture three types of pharmaceutical substances (actives, intermediates, and starting materials).

[43] May 17, 2007, available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf

[This Page Intentionally Left Blank]

4.0 Comparative *In Vivo* Reference Data – the Traditional LLNA

4.1 The Traditional LLNA Protocol Used to Generate Comparative *In Vivo* Reference Data

As described in **Section 2.1**, the traditional LLNA protocol was consistent with the original ICCVAM-recommended protocol (ICCVAM 1999). That original LLNA test method protocol was accepted by U.S. regulatory agencies (e.g., 2003 EPA Health Effects Test Guidelines) and is itself consistent with procedures described in OECD TG 429, having served as the basis for development of the test guideline. Still, TG 429 allows for more procedural variation than the ICCVAM-recommended protocol (ICCVAM 1999).

4.2 Comparative Traditional LLNA Reference Data Used

The traditional LLNA data used to evaluate the rLLNA were obtained from 12 sources (**Table D-1)**. In addition to calculated SI values for each of the tested dose levels, the vehicle tested and values for the estimated concentration needed to produce an SI of 3 (EC3) for substances classified as sensitizers were provided in Gerberick et al. (2005). The data received in response to 72 FR 27815 (May 17, 2007[44]) included calculated SI values for each of the dose levels tested and the vehicle used. If EC3 values were not included in the data source, they were calculated, where possible, using either interpolation or extrapolation (Dearman et al. 2007). This information and the database (by each source) follow in **Annex III**.

4.3 Availability of Original Records for Comparative Traditional LLNA Reference Data

An attempt was made to obtain the original records for the traditional LLNA data through the *FR* notice (72 FR 27815, May 17, 2007[44]) and requests to specific stakeholders. Although the original study records were not obtained for any of the studies, compiled *in vivo* reports and/or transcribed results were obtained and/or are available for all studies included in this evaluation.

4.4 Quality of Comparative Traditional LLNA Reference Data

Good Laboratory Practice (GLP) guidelines are internationally recognized rules designed to produce high-quality laboratory records (OECD 1998; EPA 2006a, 2006b; U.S. Food and Drug Administration [FDA] 2007a). They provide an internationally standardized procedure for the conduct of studies, reporting requirements, archiving of study data and records, and information about the test protocol to ensure the integrity, reliability, and accountability of a study.

Ideally, all data supporting the validity of a test method should be obtained from studies reported and conducted in accordance with GLP guidelines. The extent to which the traditional LLNA studies complied with GLP guidelines is based on the information provided in published and submitted reports. Based on the available information, the following papers and data submissions were identified as originating from studies that followed GLP guidelines or used data obtained according to GLP guidelines:

- H.W. Vohr/Berufsgenossenschaftliches Institut fur Arbeitsschutz (BGIA)

[44] Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf

- P. Ungeheuer/European Federation for Cosmetic Ingredients (EFfCI)
- E. Debruyne/Bayer CropScience SA
- P. Botham/European Crop Protection Association (ECPA)
- M.J. Olson/GlaxoSmithKline (GSK)
- D. Germolec/National Institute for Environmental Health Sciences (NIEHS)

The publication by Gerberick et al. (2005) does not address the GLP compliance of any of the studies discussed. Several of the substances listed in Gerberick et al. (2005) were included in the original LLNA submission to ICCVAM (ICCVAM 1999). According to the submission, "Much of the data used here to support this submission and much of the data contained within the publications cited in this document have been derived from audited Good Laboratory Practice (GLP) compliant studies. Where this is not the case all investigations have been conducted to the spirit of GLP or Good Research Practice in GLP compliant facilities" (reproduced in ICCVAM 1999). Furthermore, in response to requests from ICCVAM, records were provided indicating compliance with GLP guidelines for some of the studies.

4.5 Accuracy and Reliability of the Traditional LLNA

4.5.1 Accuracy

ICCVAM (1999) reviewed the performance of the traditional LLNA with comparisons to (1) the Guinea Pig Maximization Test and the Buehler Test (EPA 2003) and (2) human results obtained from the human maximization test[45] and human patch test allergen[46] panels. The evaluation concluded that the LLNA demonstrated adequate accuracy (ICCVAM 1999).

4.5.2 Reliability

ICCVAM (1999) also reviewed the reliability of the traditional LLNA as assessed by intra- and interlaboratory reproducibility. The evaluation concluded that the LLNA demonstrated adequate intra- and interlaboratory repeatability and reproducibility (ICCVAM 1999).

[45] The human maximization test involves application of occluded patches on the same skin site with a rest period between each reapplication. Two weeks after the last induction patch, sensitization is evaluated using a 48-hour occluded patch test. The site is scored 24 and 48 hours after patch removal.

[46] Allergen patch tests are diagnostic tests applied to the surface of the skin to identify the cause of contact dermatitis. Chemicals and substances included in these tests (e.g., nickel, rubber, and fragrance mixes) are known to cause contact dermatitis (i.e., skin sensitization) (http://www.fda.gov/cber/allergenics.htm).

5.0 rLLNA Test Method Data and Results

5.1 Description of the rLLNA Test Method Protocol Used to Generate Data

No specific rLLNA studies were conducted for this evaluation; rather, data from traditional LLNA studies were evaluated retrospectively. The only difference in the test method protocols between the traditional LLNA and the rLLNA is the number of dose levels tested. In the traditional LLNA, at least three test-substance dose levels are tested, with the highest dose level based on maximum solubility and the avoidance of systemic toxicity and/or excessive local irritation. In contrast, only the highest dose level of a substance is tested in the rLLNA (Kimber et al. 2006). This retrospective evaluation assumes that the top dose level tested in the traditional LLNA studies was in fact the maximum soluble concentration that did not cause overt systemic toxicity and/or excessive local irritation. Because the criteria for choosing the top dose in the traditional LLNA and in the rLLNA are the same, the maximum dose level tested should be the same for both. However, it is important to consider that the highest possible dose level selected in a prospective validation study may differ between the two versions of the LLNA. Thus, the accuracy analysis of these same substances in a prospective rLLNA study may differ from the accuracy analysis obtained in this retrospective rLLNA analysis.

5.2 Availability of Original rLLNA Data Used to Evaluate Accuracy and Reliability

While original study records were not obtained for any of the previously conducted studies, compiled *in vivo* reports and/or transcribed results were obtained and/or available for all studies included in this evaluation.[47]

5.3 Description of the Statistical Procedure Used to Evaluate rLLNA Data

The performance analysis in this BRD focuses on the ability of the rLLNA to identify potential skin sensitizers as determined by the calculated SI for each test substance (see **Section 2.1**).

5.4 Summary of Results

The data evaluated here were obtained from 12 sources (**Table D-1**). Where available, the specific information extracted for each substance includes its name, CASRN, physicochemical properties (e.g., form tested, Log K_{ow}), and chemical class[48] (**Annex II**). Dose levels tested, along with calculated SI and/or EC3 values, sensitizing hazard classification, and the data source are provided in **Annex III**. If EC3 values were not included in the data source, they were calculated, where possible, using either interpolation or extrapolation (Dearman et al. 2007). Other than the information provided in the submitted data, no additional attempt was made to identify the source or purity of the test substance.

[47] The LLNA data for several of the substances evaluated for this report were included in the database that was submitted to ICCVAM in 1998 for the initial evaluation of LLNA (ICCVAM 1999). Therefore, some of the original data for these substances were available for review.

[48] Chemical classes were assigned by NICEATM based on the classification of the National Library of Medicine's Medical Subject Heading (available at http://www.nlm.nih.gov/mesh/meshhome.html).

5.5 Use of Coded Substances

Neither the previous evaluation of these 211 substances (ICCVAM 1999) nor any additional studies used in this evaluation describe coding of substances to avoid potential scoring bias.

5.6 Lot-to-Lot Consistency of Test Substances

Ideally, a single lot of each substance is used during the validation of a test method. In situations where multiple lots of a chemical must be used, the lot-to-lot consistency of a test substance must be evaluated to ensure the consistency of the substance evaluated over the course of the study. The procedures used to evaluate lot-to-lot consistency are described in the published reports. No attempt was made to review original records to assess the procedures used to evaluate different batches.

Data submitted by P. Botham/ECPA, P. Ungheuer/EFfCI, and D. Germolec/NIEHS included the source and the batch number of each tested substance.

5.7 Availability of Original Data for External Audit

The LLNA data included in the ICCVAM (1999) database were reviewed during the original evaluation. The original data for the other studies included in this evaluation were not available.

6.0 Accuracy of the rLLNA

6.1 Performance Statistics

A critical component of a formal evaluation of the validation status of a test method is an assessment of the accuracy of the proposed tested method when compared to the current reference test method (ICCVAM 2003). This aspect of assay performance is typically evaluated by calculating:

- *Accuracy* (concordance): the proportion of correct outcomes (positive and negative) of a test method
- *Sensitivity*: the proportion of all positive substances that are classified as positive
- *Specificity*: the proportion of all negative substances that are classified as negative
- *Positive predictivity*: the proportion of correct positive responses among substances testing positive
- *Negative predictivity*: the proportion of correct negative responses among substances testing negative
- *False positive rate*: the proportion of all negative substances that are falsely identified as positive
- *False negative rate*: the proportion of all positive substances that are falsely identified as negative

The ability of the rLLNA to correctly identify potential skin sensitizers was compared to that of the traditional LLNA for 471 studies.[49] Of the 471 studies, 318 detected skin sensitizers and 153 detected non-sensitizers.[50] Classification of substances and complete data for each substance are located in **Annex III**. When studies for the substances tested more than once in the same vehicle were considered together to yield an overall skin sensitization classification, 465 unique substance–vehicle combination studies resulted. Of these, 315 detected sensitizers and 150 detected non-sensitizers.

Based on the available study data, the rLLNA has an accuracy of 98.7% (465/471), a sensitivity of 98.1% (312/318), a specificity of 100% (153/153), a false positive rate of 0% (0/153), and a false negative rate of 1.9% (6/318) when compared to the traditional LLNA. When substances tested more than once in the same vehicle were considered together, the resulting 465 studies give an accuracy of 98.7% (459/465), a sensitivity of 98.1% (309/315), a specificity of 100% (150/150), a false positive rate of 0% (0/150), and a false negative rate of 1.9% (6/315). The performance characteristics of the rLLNA as discussed in Kimber et al. (2006) are presented in **Table D-3**.

[49] Due to the small number of repeated studies (5%), all studies were treated independently for this accuracy evaluation. When the studies for the substances repeated in the same vehicle were considered together to yield an overall skin sensitization classification, there were 465 studies with unique substance–vehicle combinations.

[50] For two of the repeated studies (HCA and linalool alcohol), the LLNA obtained discordant results. In both cases, one study classified the substance as a non-sensitizer and the other classified it as a sensitizer. Review of the studies indicates differences in the highest dose levels tested. For each of the studies, the traditional LLNA and the rLLNA both classified the substance as a sensitizer or as a non-sensitizer.

Table D-3 Performance of the rLLNA in Predicting Skin Sensitizers Compared to the Traditional LLNA

Data	N	Accuracy		Sensitivity		Specificity		Positive Predictivity		Negative Predictivity		False Positive		False Negative	
		%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.
Kimber et al. (2006)	211	98.6	208/211	98.2	166/169	100	42/42	100	166/166	93.3	42/45	0	0/42	1.8	3/169
rLLNA	471	98.7	465/471	98.1	312/318	100	153/153	100	312/312	96.2	153/159	0	0/153	1.9	6/318
rLLNA - substances repeated in the same vehicle were considered together	465	98.7	459/465	98.1	309/315	100	150/150	100	309/309	96.2	150/156	0	0/150	1.9	6/315

Abbreviations: N = number of studies; No. = numbers used to calculate percentage

Kimber et al. (2006) proposed that a minimum testing concentration be considered for the purpose of judging the appropriateness of a non-sensitizing classification for a test substance. In their evaluation, Kimber et al. proposed testing a minimum concentration of 10% in a dose solution (2006). However, lack of sensitizing potential at 10% does not necessarily indicate that a substance will not elicit skin sensitization when tested at a higher concentration. In fact, 51 substances (16% [51/315]) within the current database were non-sensitizers at concentrations of 10%[51] but were sensitizers at higher concentrations (see **Annex IV**).

According to the 1999 ICCVAM-recommended LLNA protocol, the maximum concentration tested should be "the highest achievable level while avoiding overt systemic toxicity and/or excessive local irritation." Similar text is included in OECD TG 429 (2002). Thus, setting a minimum testing concentration is not advised because the maximum soluble concentration that avoids systemic toxicity and/or excessive local irritation may be less than 10% with a non-sensitizing result.

6.2 Discordant Results

In the current analysis, six substances yielded false negative results in the rLLNA. The discordant substances were 2-methyl-2H-isothiazol-3-one, C19-azlactone, azithromycin, camphorquinone, nickel sulfate, and a substance designated as non-ionic surfactant 2. A review of the data for the false negatives indicates that the traditional LLNA classification of the substances as skin sensitizers was based on a low- or mid-dose level that produced an SI ≥ 3, while the highest dose level tested produced an SI < 3 (see **Table D-4**). Because the rLLNA evaluates only the highest dose level tested, all six substances were identified as non-sensitizers (i.e., false negatives). Four of the six substances that resulted in false negatives using the rLLNA compared to the traditional LLNA came from LLNA studies that used pooled data. Graphs of the dose-response curves for these six substances are provided in **Figure D-1**.

Table D-4 Traditional LLNA Data for Substances Identified as False Negatives by the rLLNA

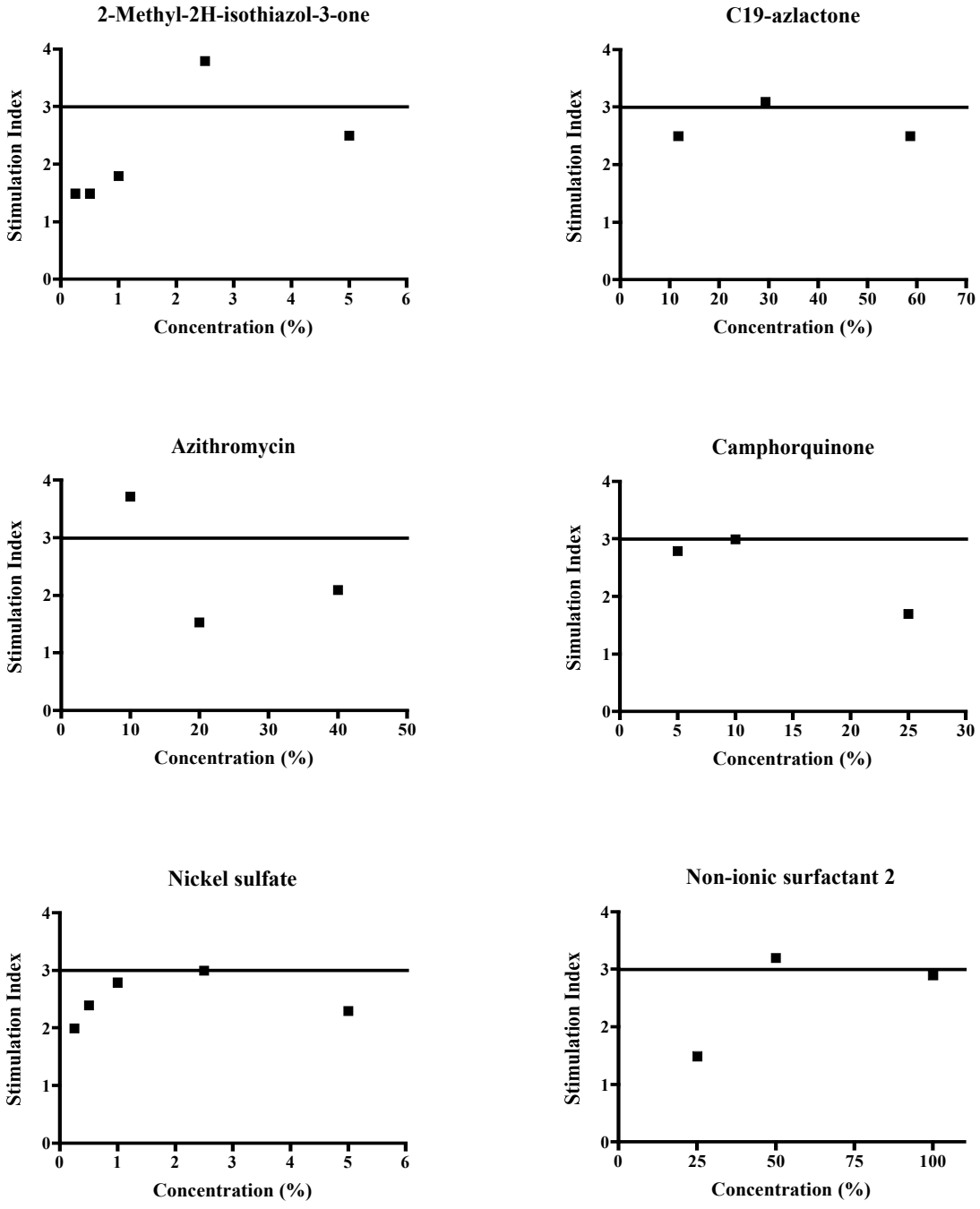
Substance	Vehicle	EC3	Traditional LLNA Data (Low- or Mid-Dose Level)		Traditional LLNA Data (Highest Dose Level)	
			Dose (%)	SI	Dose (%)	SI
2-methyl-2H-isothiazol-3-one	AOO	1.9	2.5	3.8	5	2.5
C19-azlactone	AOO	26	29.33	3.1	58.67	2.5
Azithromycin	Acetone	NC[1]	10	3.7	40	2.1
Camphorquinone	AOO	10	10	3.0	25	1.7
Nickel sulfate	Pluronic L92 (1%)	2.5	2.5	3.0	5	2.3
Non-ionic surfactant 2	AOO	47.1	50	3.2	100	2.9

Abbreviations: AOO = acetone: olive oil (4:1 by volume); EC3 = estimated concentration needed to produce a stimulation index of 3; NC = not calculated; SI = stimulation index

[1] Data was not calculated because extrapolation between points that bracket an SI of 3 could not be done.

[51] An initial dose was tested at a concentration of 10% or greater and resulted in an SI < 3, while a subsequent higher concentration resulted in an SI ≥ 3.

Figure D-1 Dose-Response Curves for Substances Identified as Sensitizers by the Traditional LLNA but as Non-Sensitizers by the rLLNA



Note: The horizontal line in each figure indicates a stimulation index of 3, which is the threshold for a positive response in the LLNA. Points on or above this line would indicate a positive (sensitizer) response, while points below this line would indicate a negative (non-sensitizer) response.

Table D-5 provides a summary of the available physicochemical properties of these substances and the vehicle used.

Table D-5 Summary of Available Physicochemical Properties for False Negatives, as Identified by the rLLNA

Substance	CASRN	Vehicle	Molecular Weight (g/mol)	K_{OW}[1]
2-Methyl-2H-isothiazol-3-one	2682-20-4	AOO	115.15	0.68[2]
C19-azlactone	—	AOO	379.63	5.21[2]
Azithromycin	83905-01-5	Acetone	748.99	3.24[3]
Camphorquinone	465-29-2	AOO	166.22	2.15[2]
Nickel sulfate	7786-81-4	Pluronic L92 (1%)	154.76	-0.17[3]
Non-ionic surfactant 2	—	AOO	—	—

Abbreviations: AOO = acetone: olive oil (4:1 by volume); CASRN = Chemical Abstracts Service Registry Number

[1] K_{OW} represents the octanol-water partition coefficient (expressed on log scale).

[2] K_{OW} calculated by the method of Moriguchi et al. (1994) and provided in Gerberick et al. (2005).

[3] K_{OW} calculated by the method of Meylan and Howard (1995) and obtained from the web site http://www.srcinc.com/what-we-do/databaseforms.aspx?id=385

[This Page Intentionally Left Blank]

7.0 Reliability of the rLLNA

An assessment of test method reliability (intralaboratory repeatability and intra- and interlaboratory reproducibility) is essential to evaluate the performance of an alternative test method (ICCVAM 2003). *Repeatability* refers to the closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period (ICCVAM 1997, 2003). *Intralaboratory reproducibility* refers to the determination of the extent to which qualified personnel within the same laboratory can replicate results using a specific test protocol at different times. *Interlaboratory reproducibility* refers to the determination of the extent to which different laboratories can replicate results using the same protocol and test substances, and indicates the extent to which a test method can be transferred successfully among laboratories.

In the data review, interlaboratory reproducibility of the rLLNA could be assessed with traditional LLNA data available for only five substances that had been tested in the same vehicle at multiple labs (**Annex III**). These are dinitrochlorobenzene (DNCB), HCA, linalool alcohol, methyl salicylate, and potassium dichromate. **Table D-6** provides a summary of the responses obtained by the rLLNA. Among these five substances, tested independently in two to three laboratories, DNCB, methyl salicylate, and potassium dichromate (3/5 = 60%) were classified as sensitizers or non-sensitizers in all studies (i.e., 100% concordance). For the other two substances, HCA and linalool alcohol, tested independently in two laboratories, one traditional LLNA study indicated each substance as a sensitizer and the other traditional LLNA study indicated each substance as a non-sensitizer.

Review of the studies indicates that the discordant results were due to differences in the highest dose levels tested. However, because the rLLNA and traditional LLNA use identical protocols and use similar data sets to evaluate the accuracy of the rLLNA and traditional LLNA, the reliability of the two methods would be expected to be similar. That is, the intra- and interlaboratory reliability of the rLLNA would be expected to be similar to that of the traditional LLNA (see ICCVAM 1999 for these statistics).

Table D-6 rLLNA Responses for Repeated Studies

Substance	Data Source	Vehicle	Traditional LLNA Response in Multiple Studies						rLLNA Classification [1]
			Dose (%)/SI	Dose (%)/SI	Dose (%)/SI	Dose (%)/SI	Dose (%)/SI	Dose (%)/SI	
1-Chloro-2-dinitrobenzene	Gerberick et al. (2005)	AOO	0.01/1.50	0.025/1.80	0.05/2.40	0.1/8.90	0.25/38.00	NA	+
	Data submitted by D. Germolec		0.01/1.17	0.025/1.12	0.05/1.93	0.1/1.95	0.25/7.10	NA	+
Hexyl cinnamic aldehyde	Gerberick et al. (2005)	AOO	2.5/1.30	5/1.10	10/2.50	25/10.00	50/17.00	NA	+
	Data Submitted by H.W. Vohr		2.5/1.10	5/1.20	10/2.84	NA	NA	NA	-
Linalool alcohol	Gerberick et al. (2005)	AOO	NA	NA	NA	25/2.50	50/4.80	100/8.30	+
	Data Submitted by D. Basketter, I. Kimber, and F. Gerberick		1/1.00	10/1.30	30/1.30	NA	NA	NA	-
Methyl salicylate	Gerberick et al. (2005)	AOO	1/1.00	2.5/1.10	5/1.60	10/1.40	20/0.90	NA	-
	Data submitted by D. Germolec		1/0.86	2.5/1.19	5/1.16	10/1.41	20/1.72	NA	-
Potassium dichromate	Gerberick et al. (2005)	DMSO	0.025/1.60	0.05/1.40	0.1/3.80	0.25/5.30	0.5/16.10	NA	+
	Data submitted by D. Germolec		0.025/1.21	0.05/1.84	0.1/2.22	0.25/3.39	NA	NA	+
	Ryan et al. (2002)		0.025/1.40	0.05/2.50	0.1/9.50	0.25/25.90	0.5/10.10	NA	+

Abbreviations: AOO = acetone: olive oil; DMSO = dimethyl sulfoxide; NA = not applicable because dose level was not tested; SI = stimulation index

[1] - = non-sensitizer, + = sensitizer

8.0 rLLNA Data Quality

8.1 Adherence to National and International GLP Guidelines

The extent to which the LLNA studies complied with GLP guidelines is based on the information provided in published and submitted reports. Based on the available information, the following papers and data submissions were identified as originating from studies that followed GLP guidelines or used data obtained according to GLP guidelines: H.W. Vohr/BGIA, P. Ungeheuer/EFfCI, E. Debruyne/Bayer CropScience SA, P. Botham/ECPA, M.J. Olson/GSK, and D. Germolec/NIEHS.

8.2 Data Quality Audits

Formal assessments of data quality, such as quality assurance audits, generally involve a systematic and critical comparison of the data provided in a study report to the laboratory records generated for a study.

Much of the data published by Gerberick et al. (2005) was conducted following GLP guidelines or were conducted in GLP-compliant facilities. Therefore, it was previously inferred that data audits were conducted on the data (ICCVAM 1999).

A formal assessment of the quality of the remainder of the LLNA data included in this BRD was not feasible. The published data on the LLNA were limited to tested concentrations and calculated SI and EC3 values. Auditing the reported values would require obtaining the original individual animal data for each LLNA experiment, which were not obtained. However, the conduct of many of the studies according to GLP guidelines implies that an independent quality assurance audit was conducted.

8.3 Impact of Deviations from GLP Guidelines

The impact of deviations from GLP guidelines cannot be evaluated for the data reviewed in this BRD, because no information on data quality audits was obtained.

8.4 Availability of Laboratory Notebooks or Other Records

The original records were not obtained for the studies included in this evaluation. Data were available for several of the substances included in the ICCVAM 1999 evaluation, thus some of the raw data for these substances were available for review.

[This Page Intentionally Left Blank]

9.0 Other rLLNA Scientific Reports and Reviews

9.1 Reports in the Peer-Reviewed Literature

A search of the terms "reduced LLNA," "cut-down LLNA," "limit dose LLNA," and "limit test LLNA" in the MEDLINE®, TOXLINE®, and Web of Science® search engines through December 2007 produced one relevant published report in addition to that of Kimber et al. (2006). Three related presentations (two posters and one platform) were included in the 2007 SOT Annual Meeting held in Charlotte, NC, from March 25-29. One of the posters (Basketter et al. 2007) and the platform presentation (Basketter 2007) detailed the evaluation that resulted in the Kimber et al. (2006) publication and are therefore not discussed below. The information in the second poster, Chaney et al. (2007), described the impact of reducing the number of animals per dose group on the performance of the rLLNA and is summarized below from the subsequent publication (Ryan et al. 2008; published online ahead of print as Ryan et al. 2007).

9.1.1 Ryan et al. (2008)

Ryan et al. (2008) evaluated the impact of reducing the number of mice (from five animals to two) on the performance characteristics using the rLLNA. Nineteen sensitizing and five non-sensitizing sustances were evaluated with 33 sensitizer datasets and eight non-sensitizer data sets.

SI values were determined for all possible two-animal combinations for the control- and high-dose groups. With 10 possible data combinations per experimental group, there were 100 possible sets of four values (two control animals and two high-dose animals) for each data set. The 100 possible SI values, each based on a unique set of four values, were plotted for each data set, and the percentage of combinations that resulted in an SI $\geq$ 3 was calculated. Of the sensitizers evaluated, at least 96% of the combinations yielded an SI $\geq$ 3 for 76% (25/33) of the data sets. Thirteen or fewer percent ($\leq$13%) of the possible combinations of non-sensitizers (excluding three data sets for sodium lauryl sulfate) had an SI $\geq$ 3. For the data sets with threshold SI values (2–4.9), however, 90% or more of the combinations resulted in SI $\geq$ 3 for only 20% (4/20) of the sensitizers. Thirteen of the 20 (65%) sensitizer data sets had less than 75% of the combinations producing SI $\geq$ 3. The authors concluded that the decreased sensitivity produced by using two mice per group was inappropriate for using the rLLNA to identify skin sensitization hazard.

[This Page Intentionally Left Blank]

10.0 Animal Welfare Considerations

10.1 How the rLLNA will Refine, Reduce, or Replace Animal Use

Compared to the traditional LLNA, the rLLNA will reduce the number of animals used to assess skin sensitization. In addition to a concurrent vehicle-control group and a positive-control group, the traditional LLNA requires testing four to five mice with each of at least three test-substance dose levels (ICCVAM 1999). Because the rLLNA tests only the highest dose level of the test substance being evaluated, in addition to the concurrent control groups, the number of animals tested would be decreased by at least 40% for each test.

10.2 Requirements for the Use of Animals

The rationale for the use of animals and the basis for determining the number of animals used in the rLLNA are the same as those for the traditional LLNA (ICCVAM 1999).

[This Page Intentionally Left Blank]

11.0 Practical Considerations

Several issues in addition to performance evaluations must be taken into account when assessing the practicality of an alternative test method in comparison to the existing test method:

- Assessments of the laboratory equipment and supplies needed to conduct the alternative test method
- Level of personnel training
- Labor costs
- Time required to complete the test method

The time, personnel cost, and effort required to conduct the proposed test method(s) must be considered reasonable in comparison to those of the test method it is intended to replace.

11.1 Transferability of the rLLNA

Test method transferability addresses the ability of a method to be performed accurately and reliably by multiple laboratories (ICCVAM 2003), including those experienced in the particular type of procedure as well as laboratories with less or no experience in the particular procedure. The degree of transferability of a test method can be evaluated by its interlaboratory reproducibility. **Section 7.0** discusses the minimum variability expected. The transferability of the rLLNA is equal to that of the traditional LLNA (ICCVAM 1999), which includes considerations for the required facilities, major fixed equipment, and any other necessary supplies.

11.2 rLLNA Training Considerations

The level of training and expertise needed to conduct the rLLNA, and the training requirements needed to demonstrate proficiency, are identical to that for the traditional LLNA (ICCVAM 1999).

11.3 Cost Considerations

The rLLNA uses the same basic protocol as the traditional LLNA. However, because fewer animals are tested, the related test costs (e.g., animal care, radioactivity, scintillation fluid, etc.) would be expected to be proportionally lower than the traditional LLNA.

11.4 Time Considerations

Because at least 40% fewer animals are tested in the rLLNA than in the traditional LLNA, the overall time required to conduct the method (e.g., dosing mice, removing the auricular lymph nodes from the animals) would be expected to decrease proportionally.

[This Page Intentionally Left Blank]

12.0 References

Basketter D, Patlewicz G, Gerberick F, Ryan C, Kern P, Betts C, et al. 2007. Identification of skin sensitizing chemicals in a reduced LLNA. [Abstract]. Toxicologist 96:236.

Basketter DA. 2007. Integrated systems and a modified local lymph node assay [Abstract]. Toxicologist 96:124.

Basketter DA, Sanders D, Jowsey IR. 2007. The skin sensitization potential of resorcinol: Experience with the local lymph node assay. Contact Dermatitis 56:196–200.

Basketter DA, Smith Pease CK, Patlewicz GY. 2003. Contact allergy: The local lymph node assay for the prediction of hazard and risk. Clin Exp Dermatol 28:218–221.

Chaney J, Ryan C, Kern P, Patlewicz G, Basketter D, Betts C, et al. 2007. The impact of reducing animal numbers in the local lymph node assay. [Abstract]. Toxicologist 96:236.

Dearman RJ, Betts CJ, Farr C, McLaughlin J, Berdasco N, Wiench K, et al. 2007. Comparative analysis of skin sensitization potency of acrylates (methyl acrylate, ethyl acrylate, butyl acrylate, and ethylhexyl acrylate) using the local lymph node assay. Contact Dermatitis 57:242–247.

EPA. 2003. Health Effects Test Guideline, OPPTS 870.2600. Skin Sensitization EPA 712–C–03–197. Washington, DC: U.S. Environmental Protection Agency. Available: http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Revised/870r-2600.pdf.

EPA. 2006a. Good Laboratory Practice Standards. Toxic Substances Control Act. 40 CFR 792. Available: http://www.access.gpo.gov/nara/cfr/waisidx_06/40cfr792_06.html.

EPA. 2006b. Good Laboratory Practice Standards. Federal Insecticide, Fungicide, and Rodenticide Act. 40 CFR 160. http://www.access.gpo.gov/nara/cfr/waisidx_06/40cfr160_06.html.

ESAC. 2007. Statement on the Reduced Local Lymph Node Assay (rLLNA). European Commission Directorate General, Joint Research Centre, Institute for Health and Consumer Protection, European Centre for the Validation of Alternative Methods, April 2007. Available: http://ecvam.jrc.it/ft_doc/ESAC26_statement_rLLNA_20070525-1.pdf

FDA. 2007a. Good laboratory practice for nonclinical laboratory studies. 21 CFR 58.

Gerberick GF, Ryan CA, Kern PS, Schlatter H, Dearman RJ, Kimber I, et al. 2005. Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. Dermatitis 16:157–202.

ICCVAM. 1997. Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publication No.: 97-3981. Research Triangle Park, NC: National Institute of Environmental Sciences. Available: http://iccvam.niehs.nih.gov/docs/about_docs/validate.pdf.

ICCVAM 1999. The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemical/Compounds. NIH Publication No. 99-4494. Research Triangle Park, NC: National Institute of Environmental Sciences. Available: http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf.

ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No: 03-4508. Research Triangle Park, NC: National Institute of Environmental Sciences. Available: http://iccvam.niehs.nih.gov/SuppDocs/SubGuidelines/SD_subg034508.pdf.

ICCVAM Authorization Act. 2000. Public Law 106-545, 42 U.S.C. 285*l*-3. Available: http://iccvam.niehs.nih.gov/docs/about_docs/PL106545.pdf.

ISO. 2002. Biological evaluation of medical devices -- Part 10: Tests for irritation and delayed-type hypersensitivity. Available for purchase at: http://www.iso.org/iso/home.htm.

Jowsey IR, Basketter DA, Westmoreland C, Kimber I. 2006. A future approach to measuring relative skin sensitising potency: a proposal. J Appl Toxicol 26:341–350.

Kimber I, Dearman RJ. 1991. Investigation of lymph node cell proliferation as a possible immunological correlate of contact sensitizing potential. Food Chem Toxicol 29:125–129.

Kimber I, Dearman RJ. 1996. Contact hypersensitivity: immunological mechanisms. In: Toxicology of Contact Hypersensitivity (Kimber I, Maurer T, eds). London: Taylor and Francis, 4–25.

Kimber I, Dearman RJ, Betts CJ, Gerberick GF, Ryan CA, Kern PS, et al. 2006. The local lymph node assay and skin sensitization: a cut-down screen to reduce animal requirements? Contact Dermatitis 54:181–185.

Lalko J, Api AM. 2006. Investigation of the dermal sensitization potential of various essential oils in the local lymph node assay. Food Chem Toxicol 44:739–746.

Meylan WM, Howard PH. 1995. Atom/fragment contribution method for estimating octanol-water partition coefficients. J Pharm Sci 84:83–92.

Moriguchi I, Hirono S, Nakagome I, Hirano H. 1994. Comparison of reliability of log P values for drugs calculated by several methods. Chem Pharm Bull 42:976–978.

OECD. 1998. OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 1: OECD Principles on Good Laboratory Practice (as revised in 1997). ENV/MC/CHEM (98)17. Paris: OECD.

OECD. 2002. Test No. 429. Skin Sensitisation: Local Lymph Node Assay. Adopted April 24, 2002. In: OECD Guidelines for Testing of Chemicals. Paris: OECD. Available: http://oberon.sourceoecd.org/vl=2675708/cl=14/nw=1/rpsv/cgi-bin/fulltextew.pl?prpsv=/ij/oecdjournals/1607310x/v1n4/s30/p1.idx.

Ryan CA, Chaney JG, Kern PS, Patlewicz GY, Basketter DA, Betts CJ, et al. 2008. The reduced local lymph node assay: the impact of group size. J Appl Toxicol 28:518–523.

Ryan CA, Cruse LW, Skinner RA, Dearman RJ, Kimber I, Gerberick GF. 2002. Examination of a vehicle for use with water soluble materials in the murine local lymph node assay. Food Chem Toxicol 40:1719–1725.

U.N. 2005. Globally Harmonised System of Classification and Labelling of Chemicals (GHS). Geneva: United Nations Publications.

[This Page Intentionally Left Blank]

13.0 Glossary

Accuracy: (a) The closeness of agreement between a test method result and an accepted reference value. (b) The proportion of correct outcomes of a test method. It is a measure of test method performance and one aspect of *relevance*. The term is often used interchangeably with *concordance* (see also *two-by-two table*). Accuracy is highly dependent on the prevalence of positives in the population being examined.*

Allergic contact dermatitis (ACD): A Type IV allergic reaction of the skin that results from repeated skin contact with a skin sensitizer. Clinical signs include the development of erythema (redness) and edema (swelling), blistering, and itching. Also referred to as *skin sensitization*.

Assay: The experimental system used. Often used interchangeably with *test* and *test method*.*

Coded substances: Substances labeled by code rather than name so that they can be tested and evaluated without knowledge of their identity or anticipation of test results. Coded substances are used to avoid intentional or unintentional bias when evaluating laboratory or test method performance.

Concordance: The proportion of all substances tested that is correctly classified as positive or negative. It is a measure of test method performance and one aspect of *relevance*. The term is often used interchangeably with *accuracy* (see also *two-by-two table*). Concordance is highly dependent on the prevalence of positives in the population being examined.*

EC3: The estimated concentration needed to produce a stimulation index of 3, as compared to the concurrent vehicle control.

Essential test method component: Structural, functional, and procedural elements of a test method that are used to develop the test method protocol. These components include unique characteristics of the test method, critical procedural details, and quality control measures. Adherence to essential test method components is necessary when the acceptability of a proposed test method is being evaluated based on performance standards derived from mechanistically and functionally similar validated test method. [Note: Previously referred to as *minimum procedural standards.*]*

False negative: A substance incorrectly identified as negative by a test method.*

False negative rate: The proportion of all positive substances falsely identified by a test method as negative (see *two-by-two table*). It is one indicator of test method accuracy.*

False positive: A substance incorrectly identified as positive by a test method.*

False positive rate: The proportion of all negative substances that are falsely identified by a test method as positive (see *two-by-two table*). It is one indicator of test method accuracy.*

Good Laboratory Practices (GLP): Regulations promulgated by the U.S. Food and Drug Administration and the U.S. Environmental Protection Agency, and principles and procedures adopted by the OECD and Japanese authorities, which describe record keeping and quality assurance procedures for laboratory records that will be the basis for data submissions to national regulatory agencies.*

The definitions in this glossary are restricted to their uses with respect to the rLLNA and the traditional LLNA.

* Definition used by ICCVAM (ICCVAM 2003).

Hazard: The potential for an adverse health or ecological effect. Hazard potential results only if an exposure occurs that leads to the possibility of an adverse effect being manifested.*

Interlaboratory reproducibility: A measure of whether different qualified laboratories using the same protocol and test substances can produce qualitatively and quantitatively similar results. Interlaboratory reproducibility is determined during the prevalidation and validation processes and indicates the extent to which a test method can be transferred successfully among laboratories.*

Intralaboratory repeatability: The closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period.*

Intralaboratory reproducibility: The first stage of validation; a determination of whether qualified people within the same laboratory can successfully replicate results using a specific test protocol at different times.*

Immunological: Relating to the immune system and immune responses.

In vivo: In the living organism. Refers to assays performed in multicellular organisms.

Local lymph node assay (LLNA): An *in vivo* test method used to assess the skin sensitization potential of a substance by measuring the proliferation of lymphocytes in the lymph nodes draining the ears (i.e., auricular lymph nodes) of mice, subsequent to topical exposure on the ear to the substance. The traditional LLNA measures lymphocyte proliferation by quantifying the amount of tritiated thymidine (^{3}H) incorporated into the cells of the draining lymph nodes.

Lymphocyte: A white blood cell found in the blood, lymph, and lymphoid tissues, which regulates and plays a role in acquired immunity.

Negative predictivity: The proportion of correct negative responses among substances testing negative by a test method (see *two-by-two table*). It is one indicator of test method accuracy. Negative predictivity is a function of the sensitivity of the test method and the prevalence of negatives among the substances tested.*

Non-sensitizer: A substance that does not cause skin sensitization after repeated skin contact.

Performance: The accuracy and reliability characteristics of a test method (see *accuracy*, *reliability*).*

Positive control: A substance known to induce a positive response used to demonstrate the sensitivity of the test method and to allow for an assessment of variability in the conduct of the assay over time. For most test methods, the positive-control substance is tested concurrently with the test substance and the vehicle/solvent control. However, for some *in vivo* test methods, periodic studies using a positive-control substance is considered adequate by the OECD.

Positive predictivity: The proportion of correct positive responses among substances testing positive by a test method (see *two-by-two table*). It is one indicator of test method accuracy.

The definitions in this glossary are restricted to their use with respect to the rLLNA and the traditional LLNA.

* Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM 2003).

Positive predictivity is a function of the sensitivity of the test method and the prevalence of positives among the substances tested.*

Prevalence: The proportion of positives in the population of substances tested (see *two-by-two table*).*

Protocol: The precise, step-by-step description of a test, including the listing of all necessary reagents, criteria, and procedures for the evaluation of the test data.*

Quality assurance: A management process by which adherence to laboratory testing standards, requirements, and record keeping procedures is assessed independently by individuals other than those performing the testing.*

Reduction alternative: A new or modified test method that reduces the number of animals required.*

Reference test method: The accepted *in vivo* test method used for regulatory purposes to evaluate the potential of a test substance to be hazardous to the species of interest.*

Refinement alternative: A new or modified test method that refines procedures to lessen or eliminate pain or distress in animals or enhances animal well-being.*

Relevance: The extent to which a test method correctly predicts or measures the biological effect of interest in humans or another species of interest. Relevance incorporates consideration of the *accuracy* or *concordance* of a test method.*

Reliability: A measure of the degree to which a test method can be performed reproducibly within and among laboratories over time. It is assessed by calculating intra- and interlaboratory reproducibility and intralaboratory repeatability.*

Replacement alternative: A new or modified test method that replaces animals with non-animal systems or one animal species with a phylogenetically lower one (e.g., a mammal with an invertebrate).*

Reproducibility: The consistency of individual test results obtained in a single laboratory (*intralaboratory reproducibility*) or in different laboratories (*interlaboratory reproducibility*) using the same protocol and test substances (see *intra-* and *interlaboratory reproducibility*).*

rLLNA (reduced LLNA): Also called the *cut-down LLNA*, *limit test LLNA*, or *LLNA limit dose procedure*. A variant of the traditional LLNA that employs a single high dose level of the test substance rather than multiple dose levels to determine its skin sensitization potential.

Sensitivity: The proportion of all positive substances that are classified correctly as positive in a test method. It is a measure of test method accuracy (see *two-by-two table*).*

Skin sensitizer: A substance that induces an allergic response following skin contact (U.N. 2005).

Specificity: The proportion of all negative substances that are classified correctly as negative in a test method. It is a measure of test method accuracy (see *two-by-two table*).*

Stimulation index (SI): A value calculated for the local lymph node assay to assess the skin sensitization potential of a test substance. The value is calculated as the ratio of radioactivity

---

The definitions in this glossary are restricted to their uses with respect to the rLLNA and the traditional LLNA.

* Definition used by ICCVAM (ICCVAM 2003).

incorporated into the auricular lymph nodes of a group of treated mice to the radioactivity incorporated into the corresponding lymph nodes of a group of vehicle-control mice. For the traditional LLNA and the rLLNA, an SI ≥ 3 classifies a substance as a skin sensitizer.

Test: The experimental system used; used interchangeably with *test method* and *assay*.*

Test method: A process or procedure used to obtain information on the characteristics of a substance or agent. Toxicological test methods generate information regarding the ability of a substance or agent to produce a specified biological effect under specified conditions. Used interchangeably with *test* and *assay*. See also *validated test method* and *reference test*.*

Transferability: The ability of a test method or procedure to be accurately and reliably performed in different, competent laboratories.*

Two-by-two table: The two-by-two table can be used for calculating accuracy (concordance) ([c+d]/[a+b+c+d]), negative predictivity (d/[c+d]), positive predictivity (a/[a+b]), prevalence ([a+c]/[a+b+c+d]), sensitivity (a/[a+c]), specificity (d/[b+d]), false positive rate (b/[b+d]), and false negative rate (c/[a+c]).*

		New Test Outcome		
		Positive	**Negative**	**Total**
Reference Test Outcome	Positive	a	c	a + c
	Negative	b	d	b + d
	Total	**a + b**	**c + d**	**a + b + c + d**

Validated test method: An accepted test method for which validation studies have been completed to determine the relevance and reliability of this method for a specific proposed use.*

Validation: The process by which the reliability and relevance of a procedure are established for a specific purpose.*

Vehicle control: An untreated sample containing all components of a test system, including the vehicle that is processed with the test substance-treated and other control samples to establish the baseline response for the samples treated with the test substance dissolved in the same vehicle.

Weight-of-evidence (process): The strengths and weaknesses of a collection of information are used as the basis for a conclusion that may not be evident from the individual data.

The definitions in this glossary are restricted to their use with respect to the rLLNA and the traditional LLNA.

* Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM 2003).

Annex I

ECVAM Scientific Advisory Committee Statement on the Validity of the rLLNA

[This Page Intentionally Left Blank]

EUROPEAN COMMISSION
DIRECTORATE GENERAL JRC
JOINT RESEARCH CENTRE
Institute for Health and Consumer Protection
European Centre for the Validation of Alternative Methods (ECVAM)

ESAC Statement on the Reduced Local Lymph Node Assay (rLLNA)

At its 26th Meeting, held on 26-27 April 2007 at the European Centre for the Validation of Alternative Methods (ECVAM), Ispra, Italy, the non-Commission members of the ECVAM Scientific Advisory Committee (ESAC)[1] unanimously endorsed the following statement:

Skin sensitisation is an important toxicological endpoint with respect to human safety.

Having reviewed the final report of the independent peer review evaluation co-ordinated by ICCVAM and NICEATM[2], the report by the EMEA[3], the pre-report of the SCCNFP[4], and evidence made available since the original submissions to ICCVAM, in March 2000 the 14th meeting of ESAC stated:

> "Following a review of the scientific report and publications on the local lymph node assay (LLNA) it is concluded that the LLNA is a scientifically validated test which can be used to assess the skin sensitisation potential of chemicals. The LLNA should be the preferred method, as it uses fewer animals and causes less pain and distress than the conventional guinea-pig methods. In some instances and for scientific reasons, the conventional methods can be used."

Since its acceptance for regulatory purposes, the LLNA has proved suitable for the purposes of satisfying a range of EU and other regulatory requirements[5].

The developers of the LLNA have now undertaken a retrospective analysis of published data obtained with the LLNA[6].

They conclude that within a tiered testing strategy in the context of REACH a "reduced" version of the LLNA (rLLNA), using only a negative control group and the equivalent of the high-dose group from the full LLNA, can be used as a screening test to distinguish between sensitisers and non-sensitisers.

ESAC established a peer review panel to evaluate if there was the potential to minimise animal use by employing the rLLNA as a screening test as part of a tiered-testing strategy for chemicals.

Mindful that with the rLLNA:

- When compared with the full LLNA the rLLNA cannot and will not result in additional false positives.
- When compared with the full LLNA the rLLNA may produce a few false negatives (3:169 in the reference document, reducing to 2:169 when negative results obtained with concentrations of <10% are considered invalid)
- The test results provided by the rLLNA do not allow the determination of the potency of a sensitising chemical.

ESAC states that the peer reviewed and published information is of a quality and nature to support the use of the rLLNA within tiered-testing strategies to reliably distinguish between chemicals that are skin sensitisers and non-sensitisers, and that animal use can be minimised providing:

EUROPEAN COMMISSION
DIRECTORATE GENERAL JRC
JOINT RESEARCH CENTRE
Institute for Health and Consumer Protection
European Centre for the Validation of Alternative Methods (ECVAM)

- The concentration used to evaluate sensitisation potential is the maximum consistent with solubility and the need to avoid local and other systemic adverse effects, and that this principle rather than strict adherence to the specific recommended absolute concentrations as in OECD TG 429 should be used.
- Negative test results associated with testing using concentrations of less than 10%, should undergo further evaluation.
- Positive and negative (vehicle) control groups are used, as appropriate, per OECD TG 429.
- The full LLNA should be performed when it is known that an assessment of sensitisation potency is required.

ESAC recommends that further work should be undertaken to determine if the 10% concentration threshold referenced above is optimal.

Thomas Hartung
Head of Unit
ECVAM
Institute for Health & Consumer Protection
Joint Research Centre
European Commission
Ispra

27 April 2007

1. The ESAC was established by the European Commission, and is composed of nominees from the EU Members States, industry, academia and animal welfare, together with representatives of the relevant Commission services.

   This statement was endorsed by the following members of the ESAC:

   Ms Sonja Beken (Belgium)
   Ms Dagmar Jírová (Czech Republic)
   Mr Tõnu Püssa (Estonia)
   Mr Lionel Larue (France)
   Mr Manfred Liebsch (Germany)
   Ms Annalaura Stammati (Italy)
   Mr Jan van der Valk (The Netherlands)
   Mr Constantin Mircioiu (Romania)
   Mr Albert Breier (Slovakia)
   Ms Argelia Castaño (Spain)
   Mr Patric Amcoff (Sweden)
   Mr Jon Richmond (UK)
   Mr Carl Westmoreland (COLIPA)
   Ms Vera Rogiers (ECOPA)
   Ms Nathalie Alépée (EFPIA)
   Mr Robert Combes (ESTIV)
   Mr Hasso Seibert (European Science Foundation)

   The following Commission Services and Observer Organisations were involved in the consultation process, but not in the endorsement process itself.

   Mr Thomas Hartung (ECVAM; chairman)
   Mr Jens Linge (ECVAM; ESAC secretary)
   Ms Elke Anklam (Director of IHCP)
   Ms Susanna Louhimies (DG Environment)
   Ms Barbara Mentré (DG ENTR)
   Ms Grace Patlewicz (ECB, DG JRC)
   Mr Christian Wimmer (DG Research)
   Mr Hajime Kojima (JACVAM)
   Ms Laurence Musset (OECD)
   Mr Barry Philips (Eurogroup for Animal Welfare)
   Mr William Stokes (NICEATM, USA)

2. NIH (1999). The murine local lymph node assay. The results of an independent peer review evaluation coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication n.99-4494.
   (http://iccvam.niehs.nih.gov/methods/immunotox/immunotox.htm)

3. EMEA (2000). Report from the ad-hoc expert meeting on testing for immunohypersensitivity (11/01/2000). European Agency for the Evaluation of Medicinal Products.

4. SCCNFP (2000). Opinion adopted by the SCCNFP during the 11 th plenary meeting, 17 February 2000. (http://ec.europa.eu/health/ph_risk/committees/sccp/docshtml/sccp_out114_en.htm)

5. A Cockshott, P Evans, CA Ryan, GF Gerberick, CJ Betts, RJ Dearman, I Kimber and DA Basketter (2006). The local lymph node assay in practice: a current regulatory perspective. Human & Experimental Toxicology **25**, 387-394.

6. I Kimber, RJ Dearman, CJ Betts, GF Gerberick, CA Ryan, PS Kern, GY Patlewicz and DA Basketter (2006.) The local lymph node assay and skin sensitisation: a cut-down screen to reduce animal requirements? Contact Dermatitis **54**, 181-185.

Annex II

Physicochemical Properties of Substances Evaluated in the rLLNA

[This Page Intentionally Left Blank]

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
(16-beta)-21-(Acetyloxy)-17-hydroxy-16-methylpregna-1,4,9(11)-triene-3,20-dione	17,21-Dihydroxy-16beta-methylpregna-1,4,9(11)-triene-3,20-dione 21-acetate	910-99-6	398.50	3.56	Solid	Pharmaceutical chemicals	GSK
(1r)-1,2,3,4-Tetrahydro-6,7-dimethoxy-2-methyl-1-[(3,4,5-trimethoxyphenyl)methyl]isoquino-line [r-(r*,r*)]-2,3-bis(benzoyloxy)-butanedioate (1:1)	–	104832-01-1	745.79	3.16	Solid	Pharmaceutical chemicals	GSK
(1R,4R)-4-Isopropenyl-1-methyl-2-methylenecyclohexane	–	–	–	–	–	–	LLNA/EC3 Validation Study
(2-Bromo-5-propoxyphenyl)(2-hydroxy-4-methoxyphenyl)-methadone	–	190965-45-8	365.23	5.26	Solid	Pharmaceutical chemicals	GSK
(2e)-2-[(2-Formyl-4-hydroxyphenyl)-methylidene]-butanedioic acid	–	773059-57-7	250.21	0.83	Solid	Pharmaceutical chemicals	GSK
(2-Oxo-1-phenyl-pyrrolidin-3-yl)(triphenyl)-phosphonium bromide	–	148776-18-5	502.40	7.51	Solid	Pharmaceutical chemicals	GSK
(2R,4S)-4-(4-Acetyl-1-piperazinyl)-n-{(1r)-1-[3,5-bis(trifluoro-methyl)phenyl]-ethyl}-2-(4-fluoro-2-methylphenyl)-n-methyl-1-piperidine-carboxamide monomethane-sulfonate	–	414910-30-8	712.73	5.63	Solid	Pharmaceutical chemicals	GSK
(2S,4S)-1-[(2s)-2-Amino-3,3-bis(4-fluorophenyl)-1-oxopropyl]-4-fluoro-2-pyrroli-dine carbonitrile	Denagliptin	483369-58-0	373.38	2.31	Solid	Pharmaceutical chemicals	GSK

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
(3as,4r,5s,6s,8r,9r,9ar,10r)-6-Ethenyldeca-hydro-5-hydroxy-4,6,9,10-tetramethyl-1-oxo-3a,9-propano-3ah-cyclopentacyclo-octen-8-yl [[(3-exo)-8-methyl-8-azabicyclo-[3.2.1]oct-3-yl]thio]-acetate	Retapamulin	224452-66-8	517.78	5.21	Solid	Pharmaceutical chemicals	GSK
(3as,4r,5s,6s,8r,9r,9ar,10r)-6-Ethenyldeca-hydro-5-hydroxy-4,6,9,10-tetra-methyl-1-oxo-3a,9-propano-3ah-cyclopentacycloocten-8-yl hydroxyacetate	Pleuromulin	125-65-5	378.51	3.98	Solid	Pharmaceutical chemicals	GSK
(3-Endo)-8-methyl-8-azabicyclo[3.2.1]-octan-3-ol	Tropine	120-29-6	141.22	-0.39	Solid	Pharmaceutical chemicals	GSK
(3r,3as,6ar)-Hexahydrofuro-[2,3-b]furan-3-ol	–	156928-09-5	130.14	-1.19	Solid	Pharmaceutical chemicals	GSK
(3r,3as,6ar)-Hexahydrofuro-[2,3-b]furan-3-yl [(1s,2r)-3-[(1,3-benzodioxol-5-ylsulfonyl)(2-methylpropyl)-amino]-2-hydroxy-1-[[4-[(2-methyl-4-thiazolyl)methoxy]phenyl]methyl]-propyl]carbamate	Brecanavir	313682-08-5	703.84	4.32	Solid	Pharmaceutical chemicals	GSK
(3R6R)-3-(2,3-Dihydro-1h-inden-2-yl)-1-[(1r)-1-(2-methyl-1,3-oxazol-4-yl)-2-(4-morpholinyl)-2-oxoethyl]-6-[(1s)-1-methylpropyl]-2,5-piperazinedione	–	820957-38-8	494.60	2.89	Solid	Pharmaceutical chemicals	GSK

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
(3S,6R)-3-isopropyl-6-methylcyclohexene	(+)-trans-p-Menth-2-ene	5113-93-9	138.25	4.70	Liquid	–	LLNA/EC3 Validation Study
(4r,5s)-(-)-1,5-Dimethyl-4-phenyl-2-imidazolidinone	–	92841-65-1	190.25	1.38	Solid	Pharmaceutical chemicals	GSK
(4r,5s)-1,5-Dimethyl-3-(1-oxo-2-propenyl)-4-phenyl-2-imidazo-lidinone	–	139109-23-2	244.30	3.33	Solid	Pharmaceutical chemicals	GSK
(4S)-1-(tert-Butoxycarbonyl)-4-fluoro-l-prolinamide	–	426844-22-6	232.26	0.98	Solid	Pharmaceutical chemicals	GSK
(4S)-1-(tert-Butoxycarbonyl)-4-fluoro-l-proline	–	203866-13-1	233.24	1.75	Solid	Pharmaceutical chemicals	GSK
(4S,5R)-1-[(1R,2R,3S)-3-(1,3-Benzodioxol-5-yl)-1-(2-benzyloxy-4-methoxyphenyl)-1-hydroxy-6-propoxy-2-indanoyl]-3,4-dimethyl-5-phenyl-2-imidazolidinone	–	190965-47-0	740.86	9.58	Solid	Pharmaceutical chemicals	GSK
(4Z)-2-Methyl-6-methyleneoct-4-ene	–	–	–	–	–	–	LLNA/EC3 Validation Study
(5R)-5-Isopropenyl-2-methyl-1-methylene-2-cyclohexene	–	–	–	–	Liquid	–	LLNA/EC3 Validation Study
(Alpha-r)-n-alpha-dimethyl-3,5-bis(trifluoro-methyl	–	334477-60-0	409.30	3.59	Solid	Pharmaceutical chemicals	GSK
(R,S)-3-Amino-2,3,4,5-tetrahydro-n-(1-methylethyl)-2,4-dioxo-n,5-diphenyl-1h-1,5-benzodiazepine-1-acetamide	–	184944-86-3	442.52	3.21	Solid	Pharmaceutical chemicals	GSK
(s)-(-)-1-Phenylpropyl-amine	–	3789-59-1	135.21	1.93	Liquid	Pharmaceutical chemicals	GSK

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
(S)-2-(4-Fluoro-2-methylphenyl)4-piperidinone (s)-alpha-hydroxybenzene-acetic acid salt	–	414910-13-7	359.40	1.68	Solid	Pharmaceutical chemicals	GSK
[3aS-(3aAlpha,4beta,5alpha,6alpha,8beta,9alpha,9abeta,10S*)]-6-Ethenyldecahydro-5-hydroxy-4,6,9,10-tetramethyl-1-oxo-3a,9-propano-3aH-cyclopentacycloocten-8-yl [(methylsulfonyl)-oxy]acetate	–	60924-38-1	456.60	4.11	Solid	Pharmaceutical chemicals	GSK
[4-(Ethoxymethyl)-2,6-dimethoxyphenyl]-boronic acid	–	591249-50-2	240.07	1.79	Solid	Pharmaceutical chemicals	GSK
[4S-[1(E),4Alpha,5alpha]]-1-[3-[2-[4-Methoxy-2-(phenylmethoxy)-benzoyl]-4-propoxyphenyl]-1-oxo-2-propenyl]-3,4-dimethyl-5-phenyl-2-imidazoli-dinone	–	190965-46-9	618.74	9.34	Solid	Pharmaceutical chemicals	GSK
1-(2',3',4',5'-Tetramethylphenyl)-3-(4'-tetrabutylphenyl)-propane-1,3-dione	–	–	336.47	5.35	–	–	Gerberick
1-(2',3',4',5'-Tetramethylphenyl)butane-1,3-dione	–	167998-73-4	221.32	3.14	–	–	Gerberick
1-(2,3-epoxypropoxy)-2,2-bis [(2,3-epoxypropoxy)methylbutane	–	–	–	–	–	–	BGIA
1-(2',5' Dimethylphenyl)butane-1,3-dione	–	56290-55-2	193.27	2.65	–	–	Gerberick
1-(2',5'-diethylphenyl)butane-1,3,-dione	–	167998-76-7	221.32	3.14	–	–	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
1-(3',4',5'-Tetramethoxyphenyl)-4-dimethylpentane-1,3-dione	–	135099-98-8	297.37	2.47	–	–	Gerberick
1-(4-Ethoxy-phenyl)-2-[4-(methyl-sulfonyl)phenyl]-ethanone	–	346413-00-1	318.40	2.46	Solid	Pharmaceutical chemicals	GSK
1-(p-methoxyphenyl)-1-penten-3-one	Powdery ketone	104-27-8	190.24	2.65	Solid	–	Gerberick
1,1,3-Trimethyl-2-formyylcyclohexa-2,4-dione	Safranal	116-26-7	150.22	2.54	Liquid	Hydrocarbons, Cyclic	Gerberick
1,1-Dimethylethyl [(1s)-1-[bis(4-fluorophenyl)-methyl]-2-[(2s,4s)-2-cyano-4-fluoro-1-pyrrolidinyl]-2-oxoethyl]carba-mate	–	483368-24-7	473.50	4.14	Solid	Pharmaceutical chemicals	GSK
1,1-Dimethylethyl [(1s)-2-[4-[(2-methyl-4-thiazolyl)methoxy]phenyl]-1-(2s)-oxiranylethyl]-carbamate	–	313680-92-1	390.51	3.32	Solid	Pharmaceutical chemicals	GSK
1,1-Dimethylethyl 3-[[[[(3s)-2,3,4,5-tetrahydro-1-[2-[(1-methylethyl)phenylamino]-2-oxoethyl]-2,4-dioxo-5-phenyl-1h-1,5-benzodiazepin-3-yl]amino]carbonyl]amino] benzoate	–	305366-94-3	661.76	6.74	Solid	Pharmaceutical chemicals	GSK
1,2,3,5,6,7-Hexahydro-2-thioxo-4h-cyclopentapyrimi-din-4-one	–	35563-27-0	168.22	0.65	Solid	Pharmaceutical chemicals	GSK
1,2-Benzisothiazolin-3-one	Proxan; Proxel active	2634-33-5	151.19	1.42	Solid	Sulfur Compounds; Heterocyclic Compounds	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
1,2-Diaminocyclohexane	cis-1,2-Cyclohexanediamine	1436-59-5	114.19	0.09	Liquid	Amines	BGIA
1,2-Dibromo-2,4-dicyanobutane	–	35691-65-7	265.93	1.91	Solid	Nitriles	Gerberick
1,3-Benzodioxazole-5-sulphonyl chloride	–	115010-10-1	220.63	0.14	Solid	Pharmaceutical chemicals	GSK
1,4-dihydroquinone	–	123-31-9	110.11	1.17	Solid	Phenols	Gerberick
1,6-Bis(2,3-epoxypropoxy)hexane	Digylidyl hexanediol; 1,6-Hexanediol diglycidyl ether	16096-31-4	230.30	0.84	Liquid	Ethers	BGIA
1-[3-(Cyclopentyl-oxy)-4-methoxy-phenyl]-4-oxocyclohexane carbonitrile	–	152630-47-2	313.40	2.23	Solid	Pharmaceutical chemicals	GSK
1-[5-[(4-Fluorophenyl)methyl]-2-furanyl]ethanone	–	280571-34-8	218.23	2.97	Solid	Pharmaceutical chemicals	GSK
12-Bromo-1-dodecanol	12-Bromolauryl alcohol	3344-77-2	265.23	3.40	Solid	Alcohols	Gerberick
12-Bromododecanoic acid	12-Bromolauric acid	73367-80-3	279.21	3.02	Solid	Lipids	Gerberick
14-Hydroxynor-morphinone	–	84116-46-1	285.30	NA	Solid	Pharmaceutical chemicals	GSK
1-Bromobutane	–	109-65-9	137.02	1.82	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Bromodocosane	–	6938-66-5	389.51	6.25	Solid	Hydrocarbons, Halogenated	Gerberick
1-Bromododecane	Lauryl bromide	143-15-7	249.23	3.79	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Bromoeicosane	–	4276-49-7	361.45	5.76	Solid	Hydrocarbons, Halogenated	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
1-Bromoheptadecane	–	3508-00-7	319.36	5.02	Solid	Hydrocarbons, Halogenated	Gerberick
1-Bromohexadecane	n-Hexadecyl bromide; Palmityl bromide; Cetyl bromide	112-82-3	305.34	4.77	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Bromohexane	n-Hexyl bromide	111-25-1	165.07	2.31	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Bromononane	–	693-58-3	207.15	3.05	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Bromooctadecane	–	112-89-0	333.39	5.26	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Bromopentadecane	n-Pentadecyl bromide	629-72-1	291.31	4.53	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Bromotetradecane	–	112-71-0	277.28	4.28	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Bromotridecane	–	765-09-3	263.26	4.03	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Bromoundecane	–	693-67-4	235.20	3.54	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Butanol	–	71-36-3	74.12	1.06	Liquid	Alcohols; Lipids	Gerberick
1-Chloro-2-dinitrobenzene	Dinitrochlorobenzene	97-00-7	202.55	-0.06	Solid	Hydrocarbon, Halogenated; Nitro Compounds; Hydrocarbons, Cyclic	NTP, Gerberick
1-Chlorohexadecane	–	4860-03-1	260.89	4.65	Liquid	Hydrocarbons, Halogenated	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
1-Chloromethylpyrene	–	1086-00-6	250.72	4.89	Solid	Hydrocarbons, Cyclic; Polycyclic Compounds	Gerberick
1-Chlorononane	n-Nonyl chloride	2473-01-0	162.70	2.93	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Chlorooctadecane	Stearyl chloride	3386-33-2	288.94	5.14	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Chlorotetradecane	Myristyl chloride	2425-54-9	232.83	4.16	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Iododecane	–	4292-19-7	296.24	3.91	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Iodohexadecane	Palmityl iodide; Hexadecyl iodide	544-77-4	352.34	4.89	Liquid/ Solid	Hydrocarbons, Halogenated	Gerberick
1-Iodohexane	–	638-45-9	212.07	2.43	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Iodononane	n-Nonyl iodide	4282-42-2	254.15	3.17	Liquid	Hydrocarbons, Halogenated	Gerberick
1-Iodooctadecane	–	629-93-6	380.39	5.39	Solid	Hydrocarbons, Halogenated	Gerberick
1-Iodotetradecane	Myristyl iodide, n-Tetradecyl iodide	19218-94-1	324.29	4.40	–	Hydrocarbons, Halogenated	Gerberick
1-Methyl-3-nitronitrosoguanidine	MNNG	70-25-7	147.09	-2.13	Solid	Amidines; Nitroso Compounds	Gerberick
1-Napthol	–	90-15-3	144.17	2.54	Solid	Hydrocarbons, Cyclic	Gerberick
1-Phenyl-1,2-propanedione	–	579-07-7	148.16	1.91	Liquid	Ketones; Heterocyclic Compounds	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
1-Phenyl-2-methylbutane-1,3-dione	–	6668-24-2	179.24	2.40	–	–	Gerberick
1-Phenyloctane-1,3-dione	–	55846-68-1	221.32	3.14	–	–	Gerberick
2-(3,4-Dimethyl-phenyl)-5-methyl-2,4-dihydropyrazol-3-one	–	18048-64-1	202.26	2.28	Solid	Pharmaceutical chemicals	GSK
2-(4-Amino-2nitro-phenylamino)-ethanol	HC Red No. 3	2871-01-4	197.19	0.12	Solid	Amines	Gerberick
2-(4-Ethoxyphenyl)-3-[4-(methyl-sulfonyl)phenyl-]pyrazolo[1,5-b]-pyridazine	GW 406381	221148-46-5	393.47	3.86	Solid	Pharmaceutical chemicals	GSK
2-(4-Oxopentyl)-1h-isoindole-1,3(2h)-dione	–	3197-25-9	231.25	1.57	Solid	Pharmaceutical chemicals	GSK
2-(4-tert-Amylcyclohexyl) acetaldehyde	QRM 2113	620159-84-4	196.33	3.28	–	–	Gerberick
2-(Benzyl)tert-butyl)amino)-1-(alpha,4-dihydroxy-m-tolyl)ethane	alpha-((Benzyl-tert-butylamino)methyl)-m-xylene-4,alpha,alpha'-triol	24085-03-8	329.44	2.51	Solid	Pharmaceutical chemicals	GSK
2,2,6,6-Tetramethyl-heptane-3,5-dione	–	1118-71-4	186.30	2.40	Liquid	Ketones	Gerberick
2,2-bis-[4-(2-hydroxy-3 methacryloxypropoxy)phenyl)]-propane	Bis-GMA	1565-94-2	512.59	4.94	Liquid	Carboxylic Acids; Phenols; Macromolecular Substances	LLNA/EC3 Validation Study
2,3,4,5-Tetrahydro-n-(1-methylethyl)-2,4-dioxo-n,5-diphenyl-3-[(phenylmethoxy)-imino]-1h-1,5-benzodiazepine-1-acetamide	–	305366-97-6	546.63	7.64	Solid	Pharmaceutical chemicals	GSK
2,3-Butanedione	Erythritol anhydride; Butadiene diepoxide	431-03-8	86.09	0.68	Liquid	Ketones	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
2,3-Dimethyl-2h-indazol-6-amine	–	444731-72-0	161.21	1.01	Solid	Pharmaceutical chemicals	GSK
2,4,6-Trichloro-1,3,5-triazine	Cyanuric chloride	108-77-0	184.41	0.78	Solid	Heterocyclic Compounds	Gerberick
2,4-Diaminophenoxyethanol HCl	–	66422-95-5	168.19	-1.28	–	Amines	LLNA/EC3 Validation Study
2,4-Dichloro-pyrimidine	–	3934-20-1	148.98	1.17	Solid	Pharmaceutical chemicals	GSK
2,4-Heptadienal	–	5910-85-0	110.16	1.80	–	Aldehydes; Hydrocarbons, Acyclic	Gerberick
2,4-Hexadienal	–	142-83-6	96.13	1.37	Liquid	Aldehydes; Hydrocarbons, Acyclic	LLNA/EC3 Validation Study
2,5-Diaminotoluene	–	95-70-5	122.08 (sulfate 156.25)	1.42	Solid (sulfate)	Amines	Gerberick
2,6-Dimethoxy-4-methyl-5-[3-(trifluoromethyl)-phenoxy]-8-quinolinamine	–	106635-86-3	378.35	5.73	Solid	Pharmaceutical chemicals	GSK
2,6-Dimethoxy-4-methyl-8-nitro-5-[3-(trifluoromethyl)-phenoxy]quinoline	–	189746-15-4	408.34	6.09	Solid	Pharmaceutical chemicals	GSK
2.4-Dinitrobenzene sulfonic acid	DNBS	89-02-1	248.17	-1.53	Liquid	Hydrocarbons, Cyclic; Sulfur Compounds	Ryan
2-[(Benzyloxy)-imino]malonic acid	–	305366-96-5	223.19	1.36	Solid	Pharmaceutical chemicals	GSK

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
2-[1-(4-Bromophenyl)-1-phenylethoxy]-n,n-dimethylethanamine hydrochloride	Bromadryl	13977-28-1	384.75	4.71	Solid	Pharmaceutical chemicals	GSK
2-Acetylcyclohexanone	–	874-23-7	143.21	1.66	Solid	Hydrocarbons, Cyclic; Ketones	Gerberick
2-Amino-6-chloro-4-nitrophenol	–	6358-09-4	188.57	0.26	Solid	Amines	Gerberick
2-Amino-di-phenylamine	–	534-85-0	184.24	2.39	Solid	Pharmaceutical chemicals	GSK
2-Aminoethyl-methylsulfone	–	49773-20-8	159.63	-1.64	Solid	Pharmaceutical chemicals	GSK
2-Aminophenol	o-Aminophenol; 2-Hydroxyaniline	95-55-6	109.13	1.17	Solid	Amines; Phenols	Gerberick
2-Benzyl-tert-butylamino-3'-hydroxymethyl-4'-hydroxyaceto-phenone hydrochloride	Ethanone, 2-((1,1-dimethylethyl) (phenymethyl)amino)-1-(4-hydroxy-3-(hydroxymethyl) phenyl)-, hydrochloride	24085-08-3	363.89	3.34	Solid	Pharmaceutical chemicals	GSK
2-Bromo-5-hydroxy-benzaldehyde	–	2973-80-0	201.02	2.45	Solid	Pharmaceutical chemicals	GSK
2-Bromo-5-propoxybenzoic acid	–	190965-43-6	259.10	3.33	Solid	Pharmaceutical chemicals	GSK
2-Bromotetradecanoic acid	2-Bromomyristic acid	10520-81-7	307.27	3.51	Solid	Carboxylic Acids	Gerberick
2-Chloro-6-methoxy-4-methylquinoline	–	6340-55-2	207.66	3.57	Solid	Pharmaceutical chemicals	GSK
2-Hydroxyethyl acrylate	HEA	818-61-1	116.12	0.54	Liquid	Carboxylic Acids	Gerberick
2-Hydroxypropyl methacrylate	2-HPMA	923-26-2	144.17	1.03	Solid	Carboxylic Acids	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
2-Mercaptobenzothiazole	Captax	149-30-4	167.25	1.80	Solid	Heterocyclic Compounds	Gerberick
2-Methoxy-4-methylphenol	Cresol	93-51-6	138.16	1.66	Liquid	Phenols	Gerberick
2-Methyl-2H-isothiazol-3-one	–	2682-20-4	115.15	0.68	Solid	Sulfur Compounds; Heterocyclic Compounds	Gerberick
2-Methyl-4H,3,1-benzoxazin-4-one	Product 240	525-76-8	161.16	1.52	Solid	–	Gerberick
2-Methyl-5-hydroxyethylaminophenol	–	55302-96-0	167.21	1.32	–	–	Gerberick
2-Methylundecanal	–	110-41-8	184.32	3.03	Liquid	Aldehydes	Gerberick
2-Morpholinoethyl isocyanide	–	443882-99-3	281.67	4.55	Solid	Pharmaceutical chemicals	GSK
2-Nitro-4-(propylthio)benzen-amine	–	54393-89-4	212.27	3.45	Liquid	Pharmaceutical chemicals	GSK
2-Nitro-p-phenylenediamine	–	5307-14-2	153.13	0.01	Solid	Amines	Gerberick
3 and 4-(4-Hydroxy-4-methylpentyl)-3-cyclohexane-1-carboxaldehyde	Lyral	31906-04-4	210.32	2.89	Liquid	Aldehydes; Hydrocarbons, Cyclic	Gerberick
3, 3', 4', 5-Tetrachlorosalicylanilide	3,5-Dichloro-N-(3,4-dichlorophenyl)-2-hydroxybenzamide; TCS	1154-59-2	351.01	3.49	Solid	Amides; Amines	Gerberick
3,4-Dichloroaniline hydrochloride	–	95-76-1	162.02	2.60	Solid	Pharmaceutical chemicals	GSK
3,4-Dihydrocoumarin	Hydroxydihydro-cinnamic acid lactone	119-84-6	148.16	1.91	Liquid	Heterocyclic Compounds	Gerberick
3,4-epoxyclohexylethyl-cyclopolymethylsiloxane	Tet-sil	–	–	–	–	–	LLNA/EC3 Validation Study

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
3,5,5-Trimethylhexanoyl chloride	–	36727-29-4	176.68	2.54	Liquid	Carboxylic Acids	Gerberick
3-[(2r)-3-[[2-(2,3-Dihydro-1h-inden-2-yl)-1,1-dimethyl-ethyl]amino]-2-hydroxypropoxy]-4,5-difluoro-benzene propanoic acid	–	753449-67-1	447.53	2.13	Solid	Pharmaceutical chemicals	GSK
3'-[(2z)-[1-(3,4-Dimethylphenyl)-1,5-dihydro-3-methyl)-5-oxo-4h-pyrazol-4-ylidene]hydrazino]-2'-hydroxy-[1,1'-biphenyl]-3-carboxylic acid, compound with 2-aminoethanol (2:1)	Eltrombopag Olamine	496775-62-3	564.65	5.25	Solid	Pharmaceutical chemicals	GSK
3-[4-[(6-Bromohexyl)oxy]-butyl]benzene-sulfonamide	–	452342-04-0	392.36	3.48	Solid	Pharmaceutical chemicals	GSK
3-Aminomethyl-3,5,5-trimethylcyclohexylamine	5-Amino-1,3,3-trimethyl-cyclohexane-methylamine; IPDA; Isophorone diamine	2855-13-2	170.30	1.90	Liquid	Amines	BGIA
3-Aminophenol	m-Aminophenol; 3-Hydroxyaniline	591-27-5	109.13	1.17	Solid	Amines; Phenols	Gerberick
3-Bromomethyl-5, 5'-dimethyl-dihydro-2(3H)-furanone	–	154750-20-6	207.07	1.79	–	–	Gerberick
3-Chloro-4-fluorobenzoyl chloride	–	65055-17-6	193.01	2.42	Solid	Pharmaceutical chemicals	GSK
3-Dimethylaminopropylamine	N,N-Dimethyl-1,3-propanediamine; DMAPA	109-55-7	102.18	0.92	Liquid	Amines	Gerberick
3-Ethoxy-1-(2',3',4',5'-tetramethylphenyl)propane-1,3-dione	–	170928-69-5	248.32	3.00	–	–	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
3-Fluoro-5-(3-pyridinyl)benzen-amine	–	181633-36-3	188.21	1.80	Solid	Pharmaceutical chemicals	GSK
3-Hydroxy-2-phenyl-4-quinolinecarboxylic acid	Oxycinchophen	485-89-2	265.27	4.95	Solid	Pharmaceutical chemicals	GSK
3-Hydroxy-4-methoxybenzaldehyde	Isovanillin	621-59-0	152.15	1.28	Solid	Pharmaceutical chemicals	GSK
3-Methyl-4-phenyl-1,2,5-thiadiazole-1,1-dioxide	MPT	3775-21-1	208.24	1.14	–	–	Gerberick
3-Methyleugenol	–	186743-26-0	178.23	2.40	–	Ethers; Phenols	Gerberick
3-Methylisoeugenol	–	186743-29-3	178.23	2.40	–	Carboxylic Acids	Gerberick
3-Phenylenediamine	m-Phenylenediamine	108-45-2	108.14	1.17	Solid	Amines	Gerberick
3-Propoxybenzoic acid	–	190965-42-5	180.21	3.08	Solid	Pharmaceutical chemicals	GSK
3-Propylidenphthalide	–	17369-59-4	174.20	2.40	Liquid	–	Gerberick
4-(Bromomethyl)-benzoic acid ethyl ester	–	26496-94-6	243.10	3.42	–	Pharmaceutical chemicals	GSK
4-(N-Ethyl-N-2-methan-sulfamido-ethyl)-2-methyl-1,4,-phenylenediamine	CD-4 developer	25646-71-3	836.97	-2.12	Solid	Amides; Sulfur Compounds	Gerberick
4'-(Trifluoro-methyl)-[1,1'-biphenyl]-4-carboxaldehyde	–	90035-34-0	250.22	4.31	Solid	Pharmaceutical chemicals	GSK
4,4,4-Trifluoro-1-phenylbutane-1,3-dione	BFA	362-06-7	219.18	2.52	–	–	Gerberick
4,4-Dibromobenzil	–	35578-47-3	368.02	5.34	Solid	–	LLNA/EC3 Validation Study

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
4-[4-[[(3R)-1-Butyl-3-[(r)-cyclohexyl-hydroxymethyl]-2,5-dioxo-1,4,9-triazaspiro[5.5]-undec-9-yl]methyl]phenoxy]benzoic acid	Aplaviroc	461443-59-4	577.73	3.92	Solid	Pharmaceutical chemicals	GSK
4-Allylanisole	Estragole	140-67-0	148.20	2.54	Liquid	Ethers; Phenols	Gerberick
4-Amino-3-nitrophenyl thiocyanate	–	54029-45-7	195.20	2.21	Solid	Pharmaceutical chemicals	GSK
4-Bromo-1-phthalimidopentane	–	59353-62-7	296.17	3.48	Liquid	Pharmaceutical chemicals	GSK
4-Chloro-6-iodoquinazoline	–	98556-31-1	290.49	2.96	Solid	Pharmaceutical chemicals	GSK
4-Fluoro-2-pyrrolidine-carboxamide	–	748165-40-4	132.14	-1.01	Solid	Pharmaceutical chemicals	GSK
4-Hydroxybenzoic acid	–	99-96-7	138.12	1.03	Solid	Phenols; Carboxylic Acids	Gerberick
4-Iodo-1-phthalimido-pentane	–	63460-47-9	343.17	3.87	Solid	Pharmaceutical chemicals	GSK
4-Isopropyl-1-methylenecyclohexane	–	–	–	–	–	–	LLNA/EC3 Validation Study
4'-Methoxyacetophenone	–	100-06-1	150.18	1.91	Solid	Ethers	Gerberick
4-Methylaminophenol sulfate	Metol; Paramethyl-aminophenol sulfate	55-55-0	344.38	-0.13	Solid	Amines; Phenols	Gerberick
4-Nitrobenzyl bromide	1-(Bromomethyl)-4-nitrobenzene	100-11-8	216.03	1.40	Solid	Hydrocarbons, Cyclic; Nitro Compounds	Gerberick
4-Phenylenediamine	p-PDA, p-Phenylenediamine	106-50-3	108.14	1.17	Solid	Amines	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
5,5-Dimethyl-3-methylenedihydro-2(3H)-furanone	–	29043-97-8	126.16	1.42	–	Heterocyclic Compounds; Sulfur Compounds; Lactones	Gerberick
5-[[4-[(2,3-Dimethyl-2h-indazol-6-yl)-methylamino]-2-pyrimidinyl]amino]-2-methylbenzene-sulfonamide	Pazopanib	444731-52-6	437.53	3.65	Solid	Pharmaceutical chemicals	GSK
5-Amino-2-methylbenzene-sulfonamide	–	69733-09-7	186.23	-0.07	Solid	Pharmaceutical chemicals	GSK
5-Amino-O-Cresol	2-Hydroxy-p-toluidine	2835-95-2	123.15	0.79	Solid	–	NTP
5-Chloro-2,6-dimethoxy-4-methyl-8-nitroquinoline	–	189746-21-2	282.69	3.95	Solid	Pharmaceutical chemicals	GSK
5-Chloro-2,6-dimethoxy-4-methylquinoline	–	189746-19-8	237.69	4.13	Solid	Pharmaceutical chemicals	GSK
5'-Chloro-2'-hydroxy-3'-nitro-[1,1'-biphenyl]-3-carboxylic acid	–	376592-58-4	293.67	4.03	Solid	Pharmaceutical chemicals	GSK
5-Chloro-2-methyl-4-isothiazolin-3-one	–	26172-55-4	149.60	0.92	Liquid	Sulfur Compounds; Heterocyclic Compounds	Gerberick
5-Chloro-6-methoxy-4-methyl-8-nitro-2(1h)quinolinone	–	189746-23-4	268.66	1.99	Solid	Pharmaceutical chemicals	GSK
5-Methoxy-2-nitro-4-(trifluoromethyl)benzene acetonitrile	–	178896-77-0	260.17	2.42	Solid	Pharmaceutical chemicals	GSK
5-Methoxy-6-(trifluoromethyl)-2,3-dihydro-1h-indole	–	178896-79-2	217.19	3.25	Solid	Pharmaceutical chemicals	GSK
5-Methyl-2,3-hexanedione	Acetyl isovaleryl	13706-86-0	128.17	1.42	Liquid	–	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
5-Methyl-2-phenyl-2-hexenal	–	21834-92-4	188.27	3.77	Liquid	–	LLNA/EC3 Validation Study
5-Methyleugenol	–	186743-25-9	178.23	2.40	–	Ethers; Phenols	Gerberick
6-(Diethylamino)-1-hexanol	–	06947-12-2	173.30	1.73	Liquid	Pharmaceutical chemicals	GSK
6-(Trifluoro-methyl)-2,3-dihydro-5-methyl-1h-indole, hydrochloride	–	280121-24-6	237.65	3.69	Solid	Pharmaceutical chemicals	GSK
6-[(2-Methyl-3-pyridinyl)oxy]-3-pyridinamine	–	181633-42-1	201.23	1.42	Solid	Pharmaceutical chemicals	GSK
6-Chloro-1-hexanol	–	2009-83-8	136.62	1.59	Liquid	Pharmaceutical chemicals	GSK
6-Diethylaminohexyl bromide hydrobromide	–	64993-14-2	317.11	3.57	Solid	Pharmaceutical chemicals	GSK
6-Iodo-quinazolin-4-ol	–	16064-08-7	272.05	1.49	Solid	Pharmaceutical chemicals	GSK
6-Methoxy-4-methyl-2(1H)-quinolinone	–	5342-23-4	189.22	1.51	Solid	Pharmaceutical chemicals	GSK
6-Methylcoumarin	6-MC	92-48-8	160.17	2.15	Solid	Heterocyclic Compounds	Gerberick
6-Methyleugenol	–	186743-24-8	178.23	2.40	–	Ethers; Phenols	Gerberick
6-Methylisoeugenol	–	13041-12-8	178.23	2.40	–	Carboxylic Acids	Gerberick
7,12-Dimethylbenz[a]anthracene	DMBA; 9,10-Dimethyl-1,2-benzanthracene	57-97-6	256.34	5.39	Solid	Hydrocarbons, Cyclic; Polycyclic Compounds	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow} [1,2]	Physical Form	Chemical Class[3]	Data Source
7-[(4z)-3-(Aminomethyl)-4-(methoxyimino)-1-pyrrolidinyl]-1-cyclopropyl-6-fluoro-1,4-dihydro-4-oxo-1,8-naphthyridine-3-carboxylic acid, monomethane-sulfonate	Gemifloxacin mesylate	210353-53-0	485.50	-1.25	Solid	Pharmaceutical chemicals	GSK
7-Bromotetradecane	7-Tetradecyl bromide; 7-Myristyl bromide	74036-97-8	277.29	4.28	–	Hydrocarbons, Halogenated	Gerberick
8-[(4-Phthalimido-1-methylbutyl)amino]-2,6-dimethoxy-4-methyl-5-(3-trifluoromethylphenoxy) quinoline	–	106635-87-4	593.61	8.70	Solid	Pharmaceutical chemicals	GSK
8-Amino-6-methoxy-4-methylquinoline	–	57514-21-3	188.23	2.30	Solid	Pharmaceutical chemicals	GSK
8-Chloro-3-pentyl-3,7-dihydro-1h-purine-2,6-dione	–	862892-90-8	256.69	2.27	Solid	Pharmaceutical chemicals	GSK
8-Hydroxy-5-[(1r)-1-hydroxy-2-[[2-[4-[(6-methoxy[1,1'-biphenyl]-3-yl)amino]phenyl]-ethyl]amino]ethyl]-2(1h)-quinolinone	–	530084-87-8	521.62	3.98	Solid	Pharmaceutical chemicals	GSK
A SC600	–	–	–	–	–	Formulation	Bayer
Abietic acid	Sylvic acid	514-10-3	302.46	4.61	Solid	Hydrocarbons, Cyclic; Polycyclic Compounds	Gerberick
Adipic acid	1,4-Butanedicarboxylic acid	124-04-9	146.14	-0.02	Solid	Pharmaceutical chemicals	GSK
AE F016382 00 TK71 A101	–	–	–	–	–	Formulation	Bayer

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Alpha-(p-toluenesulfonyl)-4-fluorobenzyliso-nitrile	–	165806-95-1	289.33	2.04	Solid	Pharmaceutical chemicals	GSK
alpha-Amyl cinnamic aldehyde	–	122-40-7	202.30	3.52	Solid	Aldehydes	Gerberick
alpha-Butyl cinnamic aldehyde	–	7492-44-6	188.27	3.28	Liquid	Aldehydes	Gerberick
alpha-Methyl cinnamic aldehyde	–	101-39-3	146.19	2.54	Liquid	Aldehydes	Gerberick
alpha-Methylphenylacetaldehyde	2-Phenyl propionaldehyde	93-53-8	134.18	2.29	Liquid	Aldehydes	Gerberick
alpha-Phellandrene	Menthadiene; Dihydro-p-cymene; p-Mentha-1,5-diene; 2-methyl-5-(1-methylethyl)- 1,3-cyclohexadiene	99-83-2	136.23	4.62	Solid	Hydrocarbons, Cyclic; Hydrocarbons, Other	LLNA/EC3 Validation Study
alpha-Terpinene	1-Isopropyl-4-methyl-1,3-cyclohexadiene; p-Mentha-1,3-diene	99-86-5	136.23	4.75	Solid	Hydrocarbons, Other	LLNA/EC3 Validation Study
Aniline	Benzenamine	62-53-3	93.13	1.56	Liquid	Amines	Gerberick
Anthranilic acid	–	118-92-3	131.14	1.21	Solid	Pharmaceutical chemicals	GSK
Atrazine SC	–	1912-24-9	215.68	2.82	Solid	Heterocyclic Compounds	ECPA, NTP
Azithromycin	–	83905-01-5	748.99	3.24	Solid	Polycyclic Compounds; Carbohydrates, Lactones	NTP
Bakelite EPR 161	–	9012-45-7	–	–	Solid	Macromolecular substances	BGIA
Bakelite EPR 162	–	9012-45-7	–	–	Solid	Macromolecular substances	BGIA

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Bakelite EPR 164	–	9012-45-7	–	–	Solid	Macromolecular substances	BGIA
Bandrowski's base	1,4-Cyclohexadiene-1,4-diamine; 1,4-Benzenediamine; N,N''-(2,5-diamino-2,5-cyclohexadiene-1,4-diylidene)bis-(9CI)	20048-27-5	318.38	0.74	Solid	Amines	LLNA/EC3 Validation Study
Basil oil	Ocimum basilicum herb oil	8015-73-4	–	–	Liquid	Lipids	Lalko & Api
Benzaldehyde	–	100-52-7	106.12	1.80	Liquid	Aldehydes	Gerberick
Benzalkonium chloride	–	8001-54-5	–	–	–	Onium Compounds	CESIO
Benzene-1,3,4-tricarboxylic anhydride	Trimellitic anhydride	552-30-7	192.13	0.75	Solid	Anhydrides; Carboxylic Acids	Gerberick
Benzo[a]pyrene	–	50-32-8	252.31	5.39	Solid	Hydrocarbons, Cyclic; Polycyclic Compounds	Gerberick
Benzocaine	–	94-09-7	165.19	1.52	Solid	Carboxylic Acids	Gerberick
Benzoquinone	p-Quinone; 1,4-Cyclohexadienedione	106-51-4	108.10	1.17	Solid	Quinones	Gerberick
Benzyl benzoate	–	120-51-4	212.25	3.14	Liquid	Carboxylic Acids	Gerberick
Benzyl bromide	alpha-Bromotoluene	100-39-0	171.03	2.56	Liquid	Hydrocarbons, Cyclic	Gerberick
Benzylidene acetone	4-phenyl-3-buten-2-one	122-57-6	146.19	2.54	Solid	Ketones	Gerberick
beta-Phellandrene	3-methylene-6-(1-methylethyl) cyclohexene; p-Mentha-1(7),2-diene	555-10-2	136.23	4.70	–	Hydrocarbons, Cyclic; Hydrocarbons, Other	LLNA/EC3 Validation Study

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
beta-Phenylcinnamaldehyde	–	1210-39-5	208.23	2.78	Liquid	–	LLNA/EC3 Validation Study
beta-Propiolactone	–	57-57-8	72.06	0.43	Liquid	Lactones	Gerberick
beta-Terpinene	p-Mentha-1(7),3-diene; (1-methylethyl)-4-methylene-1-cyclohexene	99-84-3	136.23	4.83	–	Hydrocarbons, Other	LLNA/EC3 Validation Study
bis-1,3-(2',5'-dimethylphenyl)-propane-1,3-dione	–	–	282.38	4.37	–	–	Gerberick
Bis-3,4-epoxycyclohexyl-ethyl-phenyl-methylsilane	Ph-Sil	–	–	–	–	–	LLNA/EC3 Validation Study
Bisphenol A-diglycidyl ether	–	1675-54-3	340.42	4.09	Liquid	Ethers	Gerberick
Butyl acrylate	n-butyl acrylate; n-Butyl propenoate; 2-Propenoic acid; Butyl ester	141-32-2	128.17	2.20	Liquid	Carboxylic Acids	NTP, LLNA/EC3 Validation Study
Butyl glycidyl ether	–	2426-08-6	130.19	1.42	Liquid	Ethers	Gerberick
C11-azlactone	–	176665-06-8	267.41	3.24	–	–	Gerberick
C15-azlactone	–	176665-09-1	323.52	4.23	–	–	Gerberick
C17-azlactone	–	176665-11-5	351.58	4.72	–	–	Gerberick
C19-azlactone	–	–	379.63	5.21	–	–	Gerberick
C4-azlactone	–	176664-99-6	169.22	1.52	–	–	Gerberick
C6-azlactone	–	176665-02-4	197.28	2.01	–	–	Gerberick
C9-azlactone	–	176665-04-6	239.36	2.75	–	–	Gerberick
Camphorquinone	Camphoroquinone	465-29-2	166.22	2.15	Solid	Hydrocarbons, Other	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Chlorobenzene	–	108-90-7	112.56	2.19	Liquid	Hydrocarbons, Cyclic; Hydrocarbons, Halogenated	Gerberick
Chlorothalonil	Tetrachloroiso-phthalodinitrile	1897-45-6	265.91	3.66	Solid	Nitriles	LLNA/EC3 Validation Study
Cinnamic alcohol	–	104-54-1	134.18	2.29	Solid	Alcohols	Gerberick
Cinnamic aldehyde	Cinnamaldehyde	104-55-2	132.16	2.29	Liquid	Aldehydes	Gerberick
cis-4-Cyano-4-[3-(cyclopentyloxy)-4-methoxyphenyl]cyclo-hexanecarboxylic acid	Cilomilast	153259-65-5	343.23	3.20	Solid	Pharmaceutical chemicals	GSK
cis-6-Nonenal	–	2277-19-2	140.23	2.29	Liquid	Aldehydes	Gerberick
Citral	3,7-Dimethyl-2,6-octadienal; Geranial-Neral mixture	5392-40-5	152.23	2.54/ 3.45	Liquid	Hydrocarbons, Other	Lalko & Api, Gerberick
Citronella oil	–	8000-29-1	–	–	Liquid	Lipids	Lalko & Api
Clarithromycin	–	81103-11-9	747.95	3.18	Solid	Polycyclic Compounds; Carbohydrates; Lactones	NTP
Clotrimazole	–	23593-75-1	344.84	5.35	Solid	Heterocyclic Compounds	Gerberick
Clove bud oil	Clove oil; Oil of cloves	8000-34-8	–	–	Liquid	Lipids	Lalko & Api
Clove leaf oil	–	8015-97-2	–	–	Liquid	Lipids	Lalko & Api
Clove stem oil	–	8015-98-3	–	–	Liquid	Lipids	Lalko & Api
Coumarin	–	91-64-5	146.15	1.91	Solid	Heterocyclic Compounds	Gerberick
Cyclamen aldehyde	–	103-95-7	190.29	3.28	Liquid	Carboxylic Acids	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Cytosine	4-Amino-2(1H)-pyrimidinone	71-30-7	120.11	-1.85	Solid	Pharmaceutical chemicals	GSK
D EC25®	–	–	–	–	–	Formulation	Bayer
D EW 15	–	–	–	–	–	Formulation	Bayer
Dicyclohexylcarbodiimide	–	538-75-0	206.33	6.83	Solid	Imines	NTP
Diethyl maleate	–	141-05-9	172.18	0.89	Liquid	Carboxylic Acids	Gerberick
Diethyl sulfate	–	64-67-5	154.19	-0.09	Liquid	Sulfur Compounds	Gerberick
Diethylacetaldehyde	–	97-96-1	100.16	1.56	Liquid	Aldehydes	Gerberick
Diethylenetriamine	–	111-40-0	103.17	0.29	Liquid	Amines	Gerberick
Diethylpthalate	–	84-66-2	222.24	1.87	Liquid	Carboxylic Acids	Gerberick
Dihydroeugenol	2-Methoxy-4-propylphenol; 4-Propylguaicol	2785-87-7	166.22	2.15	Liquid	Ethers; Phenols	Gerberick
Dimethyl 4-cyano-4-(3-cyclopentyloxy-4-methoxyphenyl)-pimelate	–	152630-48-3	403.48	3.31	Solid	Pharmaceutical chemicals	GSK
Dimethyl carbonate	–	616-38-6	90.08	0.10	Liquid	Pharmaceutical chemicals	GSK
Dimethyl sulfate	–	77-78-1	126.13	-0.59	Liquid	Sulfur Compounds	Gerberick
Dimethylsulfoxide	–	67-68-5	78.13	0.57	Liquid	Sulfur Compounds	Gerberick
Dinocap EC	–	39300-45-3	364.39	5.76	Liquid	Hydrocarbons, Cyclic	ECPA
Dipropylene triamine	Bis(3-aminopropyl)amine	56-18-8	131.22	-1.15	Liquid	Amines	BGIA
Dodecyl methanesulfonate	Lauryl methanesulfonate	51323-71-8	264.43	2.51	–	Esters; Sulfur Compounds	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Endo-tropine-3-mesylate	–	35130-97-3	219.31	-0.11	Solid	Pharmaceutical chemicals	GSK
Ethyl (3-endo)-8-methyl-8-azabicyclo[3.2.1]-octane-3-acetate	–	56880-11-6	211.31	1.53	Liquid	Pharmaceutical chemicals	GSK
Ethyl (z)-alpha-[[2-(1,1-dimethylethoxy)-1,1-dimethyl-2-oxoethoxy]imino]-2-[(triphenylmethyl)amino]- 4-thiazoleacetate	–	68672-65-1	599.76	7.75	Solid	Pharmaceutical chemicals	GSK
Ethyl 1h-1,2,4-triazole-3-carboxylate	–	64922-04-9	141.13	-0.02	Solid	Pharmaceutical chemicals	GSK
Ethyl 2,6-dichloro-5-fluoro-beta-oxo-3-pyridinepropanoate	–	96568-04-6	280.09	2.15	Solid	Pharmaceutical chemicals	GSK
Ethyl 4-iodobenzoate	–	51934-41-9	276.08	3.76	Liquid	Pharmaceutical chemicals	GSK
Ethyl benzoylacetate	–	94-02-0	192.21	2.01	Liquid	Esters; Ethers	Gerberick
Ethyl vanillin	–	121-32-4	166.18	1.52	Solid	Aldehydes	Gerberick
Ethyl-2-(Hydroxymethyl)-1,3-Propanediol Triacrylate	–	–	–	–	–	–	NTP
Ethylacrylate	–	140-88-5	100.12	0.92/ 1.22	Liquid	Carboxylic Acids	NTP, Gerberick
Ethylene glycol dimethacrylate	EGDMA	97-90-5	198.22	1.38	Liquid	Carboxylic Acids	Gerberick
Ethylenediamine free base	–	107-15-3	60.10	0.19	Liquid	Amines	Gerberick
Ethylhexyl acrylate	Octyl acrylate; 2-Ethylhexyl 2-propenoate; Acrylic acid; 2-ethylhexyl ester	103-11-7	184.28	4.09	Liquid	Carboxylic Acids	LLNA/EC3 Validation Study

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Eugenol	2-Methoxy-4-(2-propenyl)phenol; 4-Allyl-2-methoxyphenol; 4-Allylguaiacol	97-53-0	164.20	2.15/ 2.73	Liquid	Carboxylic Acids	Lalko & Api, Gerberick
EXP 10810 A	–	–	–	–	–	Formulation	Bayer
EXP 11120 A	–	–	–	–	–	Formulation	Bayer
F & Fo WG 50 + 25	–	–	–	–	–	Formulation	Bayer
FAR01042-00	–	–	–	–	–	Formulation	Bayer
FAR01060-00	–	–	–	–	–	Formulation	Bayer
Farnesal	–	502-67-0	220.36	3.77	Liquid	Alcohols; Hydrocarbons, other; Lipids	Gerberick
Fatty acid glutamate	–	–	–	–	–	–	CESIO
Fluorescein isothiocyanate	FITC	27072-45-3	389.38	3.32	Solid	Polycyclic Compounds; Isocyanates; Sulfur Compounds	Gerberick
Formaldehyde	–	50-00-0	30.03	0.33/ 0.35	Liquid	Aldehydes	Ryan, Gerberick
Fumaric acid	2-Butenedioic acid; Butenedioic acid; Fumarate	110-17-8	116.07	0.05	Solid	Carboxylic Acids	EFfCI
Furil	–	492-94-4	190.15	1.38	Solid	Heterocyclic Compounds	Gerberick
Fx + Me EW 69	–	–	–	–	–	Formulation	Bayer
Geraniol	Rhodinol	106-24-1	154.25	2.54/ 3.47	Liquid	Hydrocarbons, Other	Lalko & Api, Gerberick
Geranium oil	Pelargonium oil	8000-46-2	–	–	Liquid	–	Lalko & Api
Glutaraldehyde	–	111-30-8	100.12	0.92	Liquid	Aldehydes	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Glycerol	–	56-81-5	92.09	0.05	Liquid	Alcohols; Carbohydrates	Gerberick
Glyceryl thioglycolate	Acetic acid, mercapto-, monoester with 1,2,3-propanetriol	30618-84-9	166.19	-1.29	–	Lipids	CESIO
Glyoxal	Oxaldehyde; Ethanedial; Biformyl	107-22-2	58.04	0.19	Liquid	Aldehydes	Gerberick
Hexane	–	110-54-3	86.18	1.94	Liquid	Hydrocarbons, Acyclic	Gerberick
Hexyl cinnamic aldehyde	HCA; alpha-Hexyl-cinnamaldehyde; 2-(Phenylmethylene) octanal	101-86-0	216.32	3.77/ 4.82	Liquid	Aldehydes	BGIA, Gerberick
Hydroxycitronellal	–	107-75-5	172.26	2.15	Liquid	Hydrocarbons, Other	Gerberick
Hydroxyethylethylenediamine	N-(2-Hydroxyethyl) ethylenediamine	111-41-1	104.15	-2.13	Liquid	Alcohols; Amines	BGIA
Imidazolidinyl urea	Germall 115, Imidurea	39236-46-9	388.29	-3.00	Solid	Urea	Gerberick
Iodopropynyl butylcarbamate	3-iodo-2-propynylbutyl-carbamate	87977-28-4	281.09	2.45	Solid	Carboxylic Acids	LLNA/EC3 Validation Study
Isoeugenol	2-Methoxy-4-propenylphenol; 4-Propenylguaiacol	97-54-1	164.20	2.15	Liquid	Carboxylic Acids	Gerberick
Isononanoyl chloride	–	57077-36-8	176.69	2.54	–	Carboxylic Acids	Gerberick
Isopropanol	Isopropyl alcohol, 2-Propanol	67-63-0	60.10	0.82	Liquid	Alcohols	Gerberick
Isopropyl dicyandiamide	–	35695-36-4	126.16	0.51	Solid	Pharmaceutical chemicals	GSK
Isopropyl myristate	–	110-27-0	270.46	3.88	Liquid	Lipids	Gerberick
Isopropyleugenol	–	51474-90-9	206.29	2.89	–	Ethers; Phenols	Gerberick
Isopropylisoeugenol	–	2953-00-7	206.29	2.89	–	Ethers	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Jasmine absolute (Grandiflorum)	Jasmine oil	8022-96-6	–	–	Liquid	Lipids	Lalko & Api
Jasmine absolute (Sambac)	Jasmine oil	8022-96-6	–	–	Liquid	Lipids	Lalko & Api
Kanamycin	–	59-01-8; 8063-07-8	484.50	-0.90	Solid	Carbohydrates	Gerberick
Lactic acid	2-Hydroxypropanoic acid	598-82-3	90.08	0.05	Liquid	Carboxylic Acids	Gerberick
Lauryl gallate	–	1166-52-5	338.44	3.21	Solid	Carboxylic Acids	Gerberick
Laurylglycerin derivitive	–	–	–	–	–	–	CESIO
Lemongrass oil	Citral terpenes; Indian melissa oil; Indian oil of verbena; Cymbopogon citratus oil	8007-02-1	–	–	Liquid	Lipids; Hydrocarbons, other	Lalko & Api
Linalool alcohol	Linalool; Linalol; Linalyl alcohol	78-70-6	154.25	2.54/ 3.38	Liquid	Hydrocarbons	Gerberick, LLNA/EC3 Validation Study
Linalool aldehyde	–	–	–	–	–	–	LLNA/EC3 Validation Study
Linoleic acid	Grape seed oil	60-33-3	280.45	7.51	Liquid	Lipids	EFfCI
Linolenic acid	9,12,15-Octadecatrienoic acid	463-40-1	278.43	7.30	Liquid	Lipids	EFfCI
Litsea cubeba oil	–	68855-99-2	–	–	Liquid	–	Lalko & Api
Maleic acid	cis-Butenedioic acid; Toxilic acid	110-16-7	116.07	0.05	Solid	Carboxylic Acids	EFfCI
m-Chloropropio-phenone	3'-Chloropropiophenone	34841-35-5	168.62	2.90	Solid	Pharmaceutical chemicals	GSK
Methyl 4-(bromomethyl)benzoate	–	2417-72-3	229.08	2.89	Solid	Pharmaceutical chemicals	GSK

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Methyl 4-(bromomethyl)benzoate	–	2417-72-3	229.08	2.89	Solid	Pharmaceutical chemicals	GSK
Methyl acrylate	Methyl propenoate; Acrylic acid methyl ester; Methoxy-carbonylethylene	96-33-3	86.09	0.73	Liquid	Carboxylic Acids	LLNA/EC3 Validation Study
Methyl dodecanesulfonate	–	2374-65-4	264.43	2.51	–	Esters; Sulfur Compounds	Gerberick
Methyl hexacecyl sulfonate	–	4230-15-3	320.53	3.49	–	Hydrocarbons, Acyclic; Sulfur Compounds	Gerberick
Methyl hexadecenesulfonate	–	26452-48-2	318.52	3.49	–	Ethers; Sulfur Compounds	Gerberick
Methyl methanesulfonate	–	66-27-3	110.13	-0.20	Liquid	Hydrocarbons, Acyclic; Sulfur Compounds	Gerberick
Methyl pyruvate	–	600-22-6	102.09	-0.96	Liquid	Carboxylic Acids	LLNA/EC3 Validation Study
Methyl salicylate	Oil of wintergreen, 2-Hydroxybenzoic acid methyl ester	119-36-8	152.15	1.28/2.60	Liquid	Phenols; Carboxylic Acids	NTP, Gerberick
Methyl(2-sulfomethyl) octadecanoate	–	–	454.67	4.89	–	Ethers; Sulfur Compounds	Gerberick
Methyl-2-nonynoate	–	111-80-8	168.24	2.15	Liquid	Lipids	Gerberick
Methyl-4-hydroxybenzoate	Methylparaben	99-76-3	152.15	1.28	Solid	Carboxylic Acids	Gerberick
Methylmethacrylate	Pegalan	80-62-6	100.12	1.28	Liquid	Carboxylic Acids; Macromolecular Substances	LLNA/EC3 Validation Study

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
m-Phenylenebis(methylamine)	1,3-xylenediamine; m-Xylylenediamine; 1,3-Bis(aminomethyl)-benzene	1477-55-0	136.19	0.15	Liquid	Hydrocarbons, Cyclic	BGIA
n-(2-Chloro-4-pyrimidinyl)-2,3-drimethyl-2h-indazol-6-amine	–	444731-74-2	273.73	3.02	Solid	Pharmaceutical chemicals	GSK
n-(2-Chloro-4-pyrimidinyl)-n,2,3-trimethyl-2h-indazol-6-amine	–	444731-75-3	287.75	2.88	Solid	Pharmaceutical chemicals	GSK
n-(3,4-Dichlorophenyl)-n'-(1-methylethyl)-imidodicarbonimidic diamide monohydrochloride	Chlorproguanil hydrochloride	15537-76-5	324.64	3.22	Solid	Pharmaceutical chemicals	GSK
n-(4-Methoxyphenyl)-3-oxobutanamide	–	5437-98-9	207.23	0.88	Solid	Pharmaceutical chemicals	GSK
n-[(1,1-Dimethylethoxy)-carbonyl]-l-tyrosine, ethyl ester	–	72594-77-5	309.37	2.66	Solid	Pharmaceutical chemicals	GSK
n-[(1-Butyl-4-piperidinyl)methyl]-3,4-dihydro-2h-[1,3]oxazino[3,2-a]indole-10-carboxamide	Piboserod	152811-62-6	369.51	4.01	Solid	Pharmaceutical chemicals	GSK
n-[2-(Diethylamino)ethyl]-2-[[(4-fluorophenyl)-methyl]thio]-4,5,6,7-tetrahydro-4-oxo-n-[[4'-(trifluoromethyl)-[1,1'-biphenyl]-4-yl]methyl]-1h-cyclopentapyrim-idine-1-acetamide	–	356057-34-6	666.79	8.33	Solid	Pharmaceutical chemicals	GSK
n-[2-Benzyloxy-5-(2-bromo-1-hydroxy-ethyl)-phenyl]-formamide	–	201677-59-0	350.22	2.51	Solid	Pharmaceutical chemicals	GSK

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
n-{[(1,1-Dimethylethyl)oxy]carbonyl}-4-fluoro-beta-(4-fluorophenyl)-l-phenylalanine	–	481055-29-2	377.39	4.31	Solid	Pharmaceutical chemicals	GSK
n-Amino-pyridinium	–	35073-04-2	223.02	0.35	Solid	Pharmaceutical chemicals	GSK
N-Ethyl-N-nitrosourea	ENU	759-73-9	117.11	-0.73	Solid	Nitroso Compounds; Urea	Gerberick
Nickel Sulfate	–	7786-81-4	154.76	-0.17	–	Inorganic Chemical, Metals; Inorganic Chemical, Elements	Ryan
n-Isopropyl-n-phenyl-2-(2-phenylamino-phenylamino)-acetamide	–	161455-90-9	359.48	4.91	Solid	Pharmaceutical chemicals	GSK
N-Methyl-N-nitrosourea	MNU	684-93-5	103.08	-0.97	Solid	Nitroso Compounds; Urea	Gerberick
Nonanoyl chloride	Pelargonoyl chloride	764-85-2	176.68	2.54	Liquid	Carboxylic Acids	Gerberick
Non-ionic surfactant 1	–	–	–	–	–	–	CESIO
Non-ionic surfactant 2	–	–	–	–	–	–	CESIO
Non-ionic surfactant 3	–	–	–	–	–	–	CESIO
Non-ionic surfactant 4	–	–	–	–	–	–	CESIO
Non-ionic surfactant 5	–	–	–	–	–	–	CESIO
Non-ionic surfactant 6	–	–	–	–	–	–	CESIO
Non-ionic surfactant 7	–	–	–	–	–	–	CESIO
Non-ionic surfactant 8	–	–	–	–	–	–	CESIO
Non-ionic surfactant 9	–	–	–	–	–	–	CESIO

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Norbornene fluoroalcohol	2[(bicyclo[2.2.1]hept-5-ene-2-yloxy)methyl]-1,1,1,3,3,3-hexafluoro-2-propanol	305815-63-8	290.20		Liquid	–	LLNA/EC3 Validation Study
Octanoic acid	–	124-07-2	144.21	1.66	Liquid	Carboxylic Acids; Lipids	Gerberick
Octinol	Capryl alcohol; Octyl alcohol	111-87-5	130.23	2.81	Liquid	Alcohols; Lipids	EFfCI
Oleic acid	cis-9-Octadecenoic acid; Elainic acid	112-80-1	282.46	7.73	Liquid	Lipids	EFfCI
Oleyl methane sulfonate	–	35709-09-2	346.57	3.98	Liquid	Hydrocarbons, Acyclic; Sulfur Compounds	Gerberick
Oripavine	Oripavine	467-04-9	297.36	1.21	Solid	Pharmaceutical chemicals	GSK
Oxalic acid	–	144-62-7	90.03	-0.59	Solid	Carboxylic Acids	Gerberick
Oxazolone	4-Ethoxymethylene-2-phenyloxazol-5-one	15646-46-5	217.22	1.87	Solid	Heterocyclic Compounds	Gerberick
Oxirane, mono((C12-14-alkyloxy)methyl) derivs	–	68609-97-2	–	–	Liquid	–	BGIA
Oxyfluorfen EC	–	42874-03-3	361.70	5.21	Solid	Ethers	ECPA
Palmarosa oil	Cymbopogon martini oil; Geranium oil, east indian	8014-19-5	–	–	Liquid	–	Lalko & Api
Palmitoyl chloride	–	112-67-4	274.88	4.26	Liquid	Lipids	Gerberick
Penicillin G	–	61-33-6	334.39	2.09	Solid	Amides; Sulfur Compounds; Heterocyclic Compounds	Gerberick
Pentachlorophenol	Penta; PCP	87-86-5	266.34	2.79	Solid	Phenols	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Pentaerythritol Triacrylate	–	3524-68-3	298.29	0.91	–	Carboxylic Acids, Alcohols	NTP
Perillaldehyde	–	2111-75-3	150.22	2.54	Liquid	Hydrocarbons, Other	Gerberick
Phenyl benzoate	–	93-99-2	198.22	2.89	Solid	Carboxylic Acids	Gerberick
Phenylacetaldehyde	–	127-78-1	120.15	2.05	Liquid	Aldehydes	Gerberick
Phenylmethyl 2-(4-fluoro-2-methylphenyl)-4-oxo-3,4-dihydro-1(2h)-pyridine-carboxylate	–	414909-98-1	339.37	3.94	Solid	Pharmaceutical chemicals	GSK
Pluronic L92®	–	–	–	–	–	–	Ryan
p-Methylhydrocinnamic aldehyde	p-Cresyl propionaldehyde	5406-12-2	148.21	2.54	Liquid	–	Gerberick
Potassium dichromate	PDC	7778-50-9	294.18	0.62	Solid	Inorganic Chemical, Chromium Compounds; Inorganic Chemical, Potassium Compounds	NTP, Ryan, Gerberick
Precursor surfactant 1	–	–	–	–	–	–	CESIO
Precursor surfactant 2	–	–	–	–	–	–	CESIO
Propylene glycol	1,2-Dihydroxypropane; 1,2-Propanediol	57-55-6	76.09	0.43	Liquid	Alcohols	Gerberick
Propylparaben	Propyl 4-hydroxybenzoate	94-13-3	180.20	1.77	Solid	Phenols; Carboxylic Acids	Gerberick
p-tert-Butyl-a-ethyl-hydrocinnamal	Lilial	80-54-6	204.31	3.52	Liquid	Aldehydes	Gerberick
p-tert-Butylphenylglycidylether	4-tert-Butylphenyl 2,3-epoxypropyl ether	3101-60-8	206.28	3.52	Liquid	–	BGIA

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Pyridine	–	110-86-1	79.10	1.31	Liquid	Heterocyclic Compounds	Gerberick
Quinoxyfen SC	–	124495-18-7	308.13	5.69	Liquid	Heterocyclic Compounds	ECPA
Quinoxyfen/cyproconazole	–	124495-18-7/ 113096-99-4	308.134/291.776	5.69/ 3.25	Liquid	Heterocyclic Compounds	ECPA
R(+)-Limonene	–	5989-27-5	136.24	2.93	Liquid	Hydrocarbons; Hydrocarbons, Cyclic	Gerberick
R-Carvone	–	2244-16-8	150.22	3.07	Liquid	Hydrocarbons, Other	LLNA/EC3 Validation Study
R-Carvoxime	–	2051-55-0	165.23	3.57	Solid	–	LLNA/EC3 Validation Study
rel-(3r,3as,6ar)-Hexahydrofuro[2,3-b]furan-3-yl 4-nitrophenyl carbonate	–	252873-35-1	295.25	0.83	Solid	Pharmaceutical chemicals	GSK
Resorcinol	1,3-Dihydroxybenzene	108-46-3	110.11	1.17	Solid	Phenols	Basketter
Rifamycin SV	–	14897-39-3	697.77	5.04	Solid	Heterocyclic Compounds, Polycyclic Compounds	NTP
Saccharin	–	81-07-2	183.18	0.64	Solid	Sulfur Compounds; Heterocyclic Compounds	Gerberick
Salicylic acid	2-Hydroxybenzoic acid	69-72-7	138.12	1.03	Solid	Phenols; Carboxylic Acids	Gerberick
Saturated diglycerin	–	–	–	–	–	–	CESIO
Sodium ethyl xanthate	Carbonodithioic acid, O-ethyl ester, sodium salt	140-90-9	144.19	2.11	Solid	Pharmaceutical chemicals	GSK

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Sodium lauroyl lactylate	Pationic 138C	13557-75-0	366.43	2.58	–	–	Gerberick
Sodium lauryl sulfate	Sodium dodecyl sulfate, SLS, SDS, Irium	151-21-3	288.38	1.87/ 1.69	Solid	Alcohols; Sulfur Compounds; Lipids	BGIA, Gerberick
Sodium metasilicate	–	6834-92-0	122.06	-5.65	–	Minerals, Silicon Compounds	NTP
Sodium-3,3,5-trimethylhexanoyloxy-benzenesulfonate	–	94612-91-6	336.38	2.23	–	–	Gerberick
Spearmint oil	–	68917-46-4	–	–	Liquid	–	Lalko & Api
Squalene	2,6,10,15,19,23-Hexamethyl-2,6,10,14,18,22-tetracosahexaene	111-02-4	410.72	14.12	Liquid	Hydrocarbons, Acyclic	EFfCI
Streptomycin sulfate	–	3810-74-0	1457.39	-8.50	Solid	Carbohydrates	Gerberick
Succinic acid	Butanedioic acid	110-15-6	118.09	-0.75	Solid	Carboxylic Acids	EFfCI
Sulfanilamide	4-Aminobenzene-sulfonamide; p-Anilinesulfonamide; p-Sulfamidoaniline	63-74-1	172.21	0.40	Solid	Amides; Sulfur Compounds; Amines	Gerberick
Sulfanilic acid	p-Aminobenzene-sulfonic acid; p-Anilinesulfonic acid	121-57-3	173.19	0.40	Solid	Hydrocarbons, Cyclic; Sulfur Compounds	Gerberick
Tartaric acid	[R-(R*,R*)]-2,3-Dihydroxybutanedioic acid; L-Tartaric acid	87-69-4	150.09	0.87	Solid	Alcohols; Carboxylic Acids	Gerberick
tert-Butyl-3-aminobenzoate	–	92146-82-2	193.25	2.63	Solid	Pharmaceutical chemicals	GSK
Tetramethyl thiuram disulfide	Thiram; Bis (dimethylthio-carbamoyl) disulfide	137-26-8	240.44	1.17	Solid	Carboxylic Acids; Sulfur Compounds	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
trans-2-Decenal	–	3913-71-1	154.25	2.54	Liquid	Aldehydes; Hydrocarbons, Other	Gerberick
trans-2-Hexenal	–	6728-26-3	98.15	1.56	Liquid	Heterocyclic Compounds	Gerberick
Trans-2-methyl-2-butenal	–	497-03-0	84.12	1.15	Liquid	Aldehydes	LLNA/EC3 Validation Study
trans-Anethol	–	104-46-1	148.21	2.54	Liquid	Ethers; Phenols	Gerberick
Trienol	17,21-Dihydroxy-16beta-methylpregna-1,4,9(11)-triene-3,20-dione	13504-15-9	356.47	3.02	Solid	Pharmaceutical chemicals	GSK
Trifluralin EC	–	1582-09-8	335.28	5.31	–	Hydrocarbons, Cyclic; Amine	ECPA
Trimethylhexamine diamine	–	–	–	–	–	–	BGIA
Trimethylolpropane Triacrylate	–	15625-89-5	296.32	2.86	Liquid	Carboxylic Acids	NTP
Undec-10-enal	–	112-45-8	168.28	2.79	Liquid	Aldehydes	Gerberick
Undecylenic acid	10-Undecenoic acid	112-38-9	184.28	4.37	Liquid	Lipids	EFfCI
Unsaturated fatty acid	–	–	–	–	–	–	CESIO
Unsaturated fatty acid ester	–	–	–	–	–	–	CESIO
Vanillin	–	121-33-5	152.15	1.28	Solid	Aldehydes	Gerberick
Veratraldehyde	–	120-14-9	166.18	1.45	Solid	Pharmaceutical chemicals	GSK
Vinylidene dichloride	–	75-35-4	96.94	1.45	Liquid	Hydrocarbons, Acyclic; Hydrocarbons, Halogenated	Gerberick
Vinylpyridine	Ethylenepyridine	1337-81-1	105.14	1.80	Liquid	Heterocyclic Compounds	Gerberick

Substance Name	Synonyms	CASRN	Mol. Weight (g/mol)	Log K_{ow}[1,2]	Physical Form	Chemical Class[3]	Data Source
Ylang Ylang (Extra)	Cananga oil; Canangium odoratum genuina oil	8006-81-3	–	–	Liquid	–	Lalko & Api
Ylang Ylang (III)	Cananga oil; Canangium odoratum genuina oil	8006-81-3	–	–	Liquid	–	Lalko & Api

Abbreviations: CASRN = Chemical Abstracts Service Registry Number; g/mol = grams per mole.

[1] K_{ow} represents the octanol-water partition coefficient (expressed on log scale).

[2] When two numbers are shown for K_{ow}, the first number is the value calculated by the method of Moriguchi et al. (1994) and provided in Gerberick et al. (2005). The second number was calculated by the method of Meylan and Howard (1995) and obtained from the website: http://www.srcinc.com/what-we-do/databaseforms.aspx?id=385. LogP (log K_{ow}) values for GSK chemicals were calculated using the method provided by Daylight Chemical Information Systems (see: http://www.daylight.com/dayhtml/doc/clogp/index.html).

[3] Chemical classifications based on the Medical Subject Headings classification for chemicals and drugs, as developed by the National Library of Medicine (available at http://www.nlm.nih.gov/mesh/meshhome.html). Chemical classification of "pharmaceutical chemicals" for the GSK chemicals was suggested by Dr. Michael Olson of GSK, which in spirit captures three types of pharmaceutical active substances (actives, intermediates, and starting materials).

[4] Basketter = Basketter et al. 2007; Bayer = Bayer CropScience SA Studies, submitted by E. Debruyne; BGIA = Berufsgenossenschaftliches Institut fur Arbeitsschutz (German Institute for Occupational Safety and Health) Study Report, submitted by H.W. Vohr; CESIO = Comite Europeen des Agents de Surface et de Leurs Intermediaires Organiques (European Committee of Surfactants and Their Organic Intermediates) Report, submitted by K. Skirda; ECPA = European Crop Protection Association LLNA Project Report, submitted by P. Botham; EFfCI = European Federation for Cosmetic Ingredients study, submitted by P. Ungeheuer; Gerberick = Gerberick et al. 2005; GSK = Glaxo SmithKline, submitted by M.J. Olson; Lalko & Api = Lalko & Api (2006), submitted by A. Api (Research Institute for Fragrance Materials [RIFM]); LLNA/EC3 Validation Study, submitted by D. Basketter, I. Kimber, and F. Gerberick; NTP = NTP Study, submitted by D. Germolec; Ryan = Ryan et al. (2002).

Annex III

Traditional LLNA Data Used for the Performance Analysis of the rLLNA

[This Page Intentionally Left Blank]

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Resorcinol	108-46-3	AOO	+	+	6	1	1.80	2.5	2.30	5.0	2.60	10	6.30	25	10.10	50	12.50	Basketter et al. (2007)
A SC600		PA/H_2O	-	-	NC	10	1.40	25	1.80	50	2.30	100	1.60					Bayer CropScience SA Studies, Submitted by E. Debruyne
AE F016382 00 TK71 A101		PA/H_2O	-	-	NC	3.6	1.00	7.1	0.80	17.9	1.00	35.7	1.10					Bayer CropScience SA Studies, Submitted by E. Debruyne
D EC25®		PA/H_2O	-	-	NC	0.5	0.56	1	0.63	2.5	0.59							Bayer CropScience SA Studies, Submitted by E. Debruyne
D EW 15		PA/H_2O	-	-	NC	2.5	1.90	5	1.50	10	2.50	25	2.50					Bayer CropScience SA Studies, Submitted by E. Debruyne
EXP 10810 A		PA/H_2O	+	+	2.1	10	6.40	25	8.40	50	9.20							Bayer CropScience SA Studies, Submitted by E. Debruyne
EXP 11120 A		PA/H_2O	+	+	64.9	10	0.96	25	0.66	50	1.60	100	6.30					Bayer CropScience SA Studies, Submitted by E. Debruyne
F & Fo WG 50 + 25		PA/H_2O	+	+	0.0031	2.5	11.70	5	12.60	10	14.10	25	15.20					Bayer CropScience SA Studies, Submitted by E. Debruyne
FAR01042-00		PA/H_2O	-	-	NC	10	1.40	25	2.10	50	1.40	100	2.50					Bayer CropScience SA Studies, Submitted by E. Debruyne
FAR01060-00		PA/H_2O	+	+	88.5	10	0.40	25	0.80	50	1.00	100	3.60					Bayer CropScience SA Studies, Submitted by E. Debruyne
Fx + Me EW 69		PA/H_2O	+	+	25.2	5	0.83	10	1.55	25	2.95	50	8.55					Bayer CropScience SA Studies, Submitted by E. Debruyne
1-(2,3-epoxypropoxy)-2,2-bis [(2,3-epoxypropoxy)-methylbutane		ACE	+	+	1.4	1	2.06	3	6.52	10	12.00							BGIA Study Report, Submitted by H.W. Vohr
1,2-Diaminocyclo-hexane	1436-59-5	ACE	+	+	0.4	0.1	1.19	0.3	1.81	1	8.39							BGIA Study Report, Submitted by H.W. Vohr

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
1,6-Bis(2,3-epoxypropoxy)-hexane	16096-31-4	ACE	+	+	1.9	0.3	0.94	1	1.67	3	4.65							BGIA Study Report, Submitted by H.W. Vohr
3-Aminomethyl-3,5,5-trimethylcyclohexyl amine	2855-13-2	ACE	+	+	1.0	0.3	1.17	1	2.68	3	20.16							BGIA Study Report, Submitted by H.W. Vohr
Bakelite EPR 161	9012-45-7	ACE	+	+	0.7	0.1	1.02	0.3	2.37	1	3.49							BGIA Study Report, Submitted by H.W. Vohr
Bakelite EPR 162	9012-45-7	ACE	+	+	0.1	0.3	10.53	1	19.94	3	39.89							BGIA Study Report, Submitted by H.W. Vohr
Bakelite EPR 164	9012-45-7	ACE	+	+	0.2	0.3	5.58	1	16.11	3	28.13							BGIA Study Report, Submitted by H.W. Vohr
Dipropylene triamine	56-18-8	ACE	+	+	0.9	0.3	2.16	1	3.17	3	12.45							BGIA Study Report, Submitted by H.W. Vohr
Hexyl cinnamic aldehyde	101-86-0	AOO	-	-	NC	2.5	1.12	5	1.19	10	2.84							BGIA Study Report, Submitted by H.W. Vohr
Hexyl cinnamic aldehyde	101-86-0	ACE	+	+	1.2	3	4.56	10	6.63	30	9.86							BGIA Study Report, Submitted by H.W. Vohr
Hydroxyethyl-ethylenediamine	111-41-1	ACE	+	+	IDR[3]	3	2.00	10	1.72	30	6.60							BGIA Study Report, Submitted by H.W. Vohr
m-Phenylenebis-(methylamine)	1477-55-0	ACE	+	+	0.4	0.3	1.92	1	9.09	3	44.20							BGIA Study Report, Submitted by H.W. Vohr
Oxirane, mono((C12-14-alkyloxy)methyl) derivs	68609-97-2	ACE	+	+	0.6	0.3	2.35	1	4.16	3	22.74							BGIA Study Report, Submitted by H.W. Vohr
p-tert-Butylphenyl-glycidylether	3101-60-8	ACE	+	+	0.4	0.1	1.36	0.3	1.68	1	14.22							BGIA Study Report, Submitted by H.W. Vohr
Sodium lauryl sulfate	151-21-3	Pluronic L92	+	+	4.9	5	3.05	10	4.78	25	8.46							BGIA Study Report, Submitted by H.W. Vohr
Trimethylhexamine diamine		ACE	+	+	1.9	1	2.15	3	4.00	10	8.86							BGIA Study Report, Submitted by H.W. Vohr
Benzalkonium chloride	8001-54-5	ACE	+	+	0.1	0.5	9.00	1	11.10	2	7.60							CESIO Report, Submitted by K. Skirda

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Fatty acid glutamate			+	+	IDR[3]	5	1.50	25	1.80	50	1.20	100	4.80					CESIO Report, Submitted by K. Skirda
Glyceryl thioglycolate	30618-84-9	AOO	+	+	4.7	10	8.00	25	14.00	50	31.00							CESIO Report, Submitted by K. Skirda
Laurylglycerin derivative		DMF	+	+	24.3	5	1.62	10	2.36	25	3.03							CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 1		AOO	+	+	27.5	25	2.80	50	4.80	100	6.50							CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 2		AOO	-	+	47.1	25	1.50	50	3.20	100	2.90							CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 3		AOO	+	+	19.8	25	4.70	50	9.80	100	13.30							CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 4		AOO	+	+	0.012	25	36.00	50	39.00	100	162.00							CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 5		AOO	+	+	37.5	25	2.70	50	3.30	100	3.20							CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 6		AOO	+	+	34.4	25	2.70	50	3.50	100	6.50							CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 7		AOO	+	+	IDR[3]	25	6.30	50	50.80	100	7.40							CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 8		AOO	+	+	IDR[3]	25	4.20	50	3.30	100	5.60							CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 9		AOO	+	+	10.5	25	3.50	50	3.90	100	7.70							CESIO Report, Submitted by K. Skirda
Precursor surfactant 1		AOO	+	+	60.7	25	2.20	50	2.70	100	4.10							CESIO Report, Submitted by K. Skirda
Precursor surfactant 2		AOO	+	+	24.0	25	3.10	50	4.80	100	4.40							CESIO Report, Submitted by K. Skirda
Saturated diglycerin		EtOH/ H_2O	-	-	NC	25	1.40	50	2.10	100	1.90							CESIO Report, Submitted by K. Skirda
Unsaturated fatty acid		AOO	+	+	22.2	25	3.40	50	5.70	100	6.50							CESIO Report, Submitted by K. Skirda

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Unsaturated fatty acid ester		AOO	+	+	27.1	25	2.80	50	5.20	100	4.70							CESIO Report, Submitted by K. Skirda
Atrazine SC	1912-24-9	Pluronic L92 (1%)	+	+	31.3	12.5	1.80	25	2.80	50	3.60	75	7.10	100	7.30			ECPA LLNA Project Report, Submitted by P. Botham
Dinocap EC	39300-45-3	Pluronic L92 (1%)	+	+	1.1	0.8	2.00	4	14.20	21	26.70							ECPA LLNA Project Report, Submitted by P. Botham
Oxyfluorfen EC	42874-03-3	Pluronic L92 (1%)	-	-	NC	1	0.30	7	0.90	33	2.30							ECPA LLNA Project Report, Submitted by P. Botham
Quinoxyfen SC	124495-18-7	Pluronic L92 (1%)	-	-	NC	7	1.10	33	1.70	100	0.80							ECPA LLNA Project Report, Submitted by P. Botham
Quinoxyfen/ cyproconazole	124495-18-7/ 113096-99-4	Pluronic L92 (1%)	+	+	27.8	12.5	2.00	25	2.30	50	8.60	75	15.80	100	30.10			ECPA LLNA Project Report, Submitted by P. Botham
Trifluralin EC	1582-09-8	Pluronic L92 (1%)	+	+	7.0	7	3.10	33	26.30	100	61.50							ECPA LLNA Project Report, Submitted by P. Botham
Fumaric acid	110-17-8	DMSO	-	-	NC	5	1.30	10	2.30	25	1.40							EFfCI study, Submitted by P. Ungeheuer
Linoleic acid	60-33-3	AOO	+	+	14.1	10	1.50	25	7.00	50	9.10							EFfCI study, Submitted by P. Ungeheuer
Linolenic acid	463-40-1	AOO	+	+	9.9	10	3.10	25	9.30	50	10.30							EFfCI study, Submitted by P. Ungeheuer
Maleic acid	110-16-7	DMSO	+	+	7.0	10	6.70	25	16.10	50	16.10							EFfCI study, Submitted by P. Ungeheuer
Octinol	111-87-5	AOO	+	+	4.7	10	5.60	25	8.80	50	11.20							EFfCI study, Submitted by P. Ungeheuer
Oleic acid	112-80-1	AOO	+	+	10.5	10	2.60	25	14.90	50	6.90							EFfCI study, Submitted by P. Ungeheuer
Squalene	111-02-4	AOO	+	+	7.9	10	3.80	25	6.90	50	8.20							EFfCI study, Submitted by P. Ungeheuer

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Succinic acid	110-15-6	DMSO	-	-	NC	5	1.20	10	1.20	25	1.30							EFfCI study, Submitted by P. Ungeheuer
Undecylenic acid	112-38-9	AOO	+	+	19.4	10	2.50	25	3.30	50	4.40							EFfCI study, Submitted by P. Ungeheuer
1-(2',3',4',5'-Tetramethylphenyl) -3-(4'-tetrabutylphenyl)-propane-1,3-dione		ACE	-	-	NC	10	1.60	20	1.20	40	1.60							Gerberick et al. (2005)
1-(2',3',4',5'-Tetramethylphenyl) butane-1,3-dione	167998-73-4	ACE	+	+	8.3	10	7.00	20	22.10	40	22.40							Gerberick et al. (2005)
1-(2',5' Dimethylphenyl) butane-1,3-dione	56290-55-2	ACE	+	+	12.5	10	2.3	20	5.1	40	9.5							Gerberick et al. (2005)
1-(2',5'-diethylphenyl) butane-1,3,-dione	167998-76-7	ACE	+	+	9.6	10	3.9	20	19.2	40	18.7							Gerberick et al. (2005)
1-(3',4',5'-Tetramethoxy-phenyl)-4-dimethylpentane-1,3-dione	135099-98-8	ACE	-	-	NC	10	2.80	20	1.10	40	0.70							Gerberick et al. (2005)
1-(p-methoxyphenyl)-1-penten-3-one	104-27-8	AOO	+	+	9.3	10	3.5	25	10	50	26.1							Gerberick et al. (2005)
1,1,3-Trimethyl-2-formyylcyclohexa-2,4-dione	116-26-7	AOO	+	+	7.5	0.5	0.70	1	1.10	2.5	1.10	5	2.70	10	3.30			Gerberick et al. (2005)
1,2-Benzisothiazolin-3-one	2634-33-5	DMF	+	+	2.3	10	3.80	30	4.40	50	4.90							Gerberick et al. (2005)
1,2-Dibromo-2,4-dicyanobutane	35691-65-7	AOO	+	+	0.9	0.5	1.40	1	3.40	2.5	3.50	5	5.40					Gerberick et al. (2005)
1,4-dihydroquinone	123-31-9	AOO	+	+	0.1	0.1	2.80	0.25	5.80	0.5	13.70	1	15.20	2.5	13.10			Gerberick et al. (2005)
12-Bromo-1-dodecanol	3344-77-2	AOO	+	+	6.9	5	2.20	10	4.3	25	9.8							Gerberick et al. (2005)
12-Bromododecanoic acid	73367-80-3	AOO	+	+	17.9	5	1.30	10	2.00	25	3.9							Gerberick et al. (2005)
1-Bromobutane	109-65-9	AOO	-	-	NC	5	1.1	10	1.2	25	1							Gerberick et al. (2005)
1-Bromodocosane	6938-66-5	AOO	+	+	8.3	2.5	1.2	5	1.6	10	3.7							Gerberick et al. (2005)
1-Bromododecane	143-15-7	AOO	+	+	17.7	5	1.1	10	1.4	25	4.5							Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
1-Bromoeicosane	4276-49-7	AOO	+	+	6.1	5	2.1	10	6.2	25	8.4							Gerberick et al. (2005)
1-Bromoheptadecane	3508-00-7	AOO	+	+	4.8	5	3.2	10	6	25	9.6							Gerberick et al. (2005)
1-Bromohexadecane	112-82-3	AOO	+	+	2.3	1	1.1	2.5	3.3	5	7.9	10	11.1	25	13.5	50	16.8	Gerberick et al. (2005)
1-Bromohexane	111-25-1	AOO	+	+	10.3	1	1.7	10	2.9	50	18.6							Gerberick et al. (2005)
1-Bromononane	693-58-3	AOO	-	-	NC	5	1.2	10	1.4	25	2.8							Gerberick et al. (2005)
1-Bromooctadecane	112-89-0	AOO	+	+	15.2	5	1.8	10	2.2	25	4.5							Gerberick et al. (2005)
1-Bromopentadecane	629-72-1	AOO	+	+	5.1	5	2.9	10	7.8	25	19.6							Gerberick et al. (2005)
1-Bromotetradecane	112-71-0	AOO	+	+	9.2	5	1.5	10	3.3	25	11.3							Gerberick et al. (2005)
1-Bromotridecane	765-09-3	AOO	+	+	10.2	5	1.6	10	2.9	25	10.4							Gerberick et al. (2005)
1-Bromoundecane	693-67-4	AOO	+	+	19.6	5	1.3	10	1.4	25	3.9							Gerberick et al. (2005)
1-Butanol	71-36-3	dH_2O	-	-	NC	5	1.6	10	1.2	20	1.4							Gerberick et al. (2005)
1-Chloro-2,4-dinitrobenzene	97-00-7	AOO	+	+	0.05	0.01	1.50	0.025	1.8	0.05	2.4	0.1	8.9	0.25	38			Gerberick et al. (2005)
1-Chlorohexadecane	4860-03-1	AOO	+	+	9.1	5	1.6	10	3.3	25	5.7							Gerberick et al. (2005)
1-Chloromethyl-pyrene	1086-00-6	AOO	+	+	0.005	0.025	11.6	0.05	15.4	0.1	18.6							Gerberick et al. (2005)
1-Chlorononane	2473-01-0	AOO	-	-	NC	10	1	25	1.6	50	2.3							Gerberick et al. (2005)
1-Chlorooctadecane	3386-33-2	AOO	+	+	16.3	10	1.7	25	4.8	50	7.3							Gerberick et al. (2005)
1-Chlorotetradecane	2425-54-9	AOO	+	+	20.2	10	1.1	25	3.9	50	6.3							Gerberick et al. (2005)
1-Iododecane	4292-19-7	AOO	+	+	13.1	5	1.70	10	2.30	25	5.70							Gerberick et al. (2005)
1-Iodohexadecane	544-77-4	AOO	+	+	19.1	10	1.60	25	3.90	50	6.40							Gerberick et al. (2005)
1-Iodohexane	638-45-9	AOO	-	-	NC	10	0.90	25	1.20	50	2.50							Gerberick et al. (2005)
1-Iodononane	4282-42-2	AOO	+	+	24.2	10	1.30	25	3.10	50	4.60							Gerberick et al. (2005)
1-Iodooctadecane	629-93-6	AOO	-	-	NC	5	1	10	1.4	25	1.9							Gerberick et al. (2005)
1-Iodotetradecane	19218-94-1	AOO	+	+	13.8	10	1.70	25	6.90	50	9.70							Gerberick et al. (2005)
1-Methyl-3-nitro-nitrosoguanidine	70-25-7	AOO	+	+	0.03	0.05	27.5	0.1	60.4	0.25	78.3							Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
1-Napthol	90-15-3	AOO	+	+	1.3	0.1	1.40	0.25	1.00	0.5	1.20	1	1.50	2.5	8.50			Gerberick et al. (2005)
1-Phenyl-1,2-propanedione	579-07-7	AOO	+	+	1.3	5	12.80	10	17.70	25	20.10							Gerberick et al. (2005)
1-Phenyl-2-methylbutane-1,3-dione	6668-24-2	ACE	+	+	29.1	10	1.70	20	2.00	40	4.20							Gerberick et al. (2005)
1-Phenyloctane--1,3-dione	55846-68-1	ACE	+	+	10.4	10	2.80	20	6.60	40	8.70							Gerberick et al. (2005)
2-(4-Amino-2nitro-phenylamino)-ethanol	2871-01-4	AOO	+	+	2.2	0.1	0.50	0.25	1.20	0.5	1.90	1	1.8	2.5	3.3			Gerberick et al. (2005)
2-(4-tert-Amylcyclohexyl) acetaldehyde	620159-84-4	AOO	+	+	36.8	25	2.1	50	4.00	100	9.10							Gerberick et al. (2005)
2,2,6,6-Tetramethyl-heptane-3,5-dione	1118-71-4	ACE	+	+	26.7	10	2.10	20	2.80	40	3.40							Gerberick et al. (2005)
2,3-Butanedione	431-03-8	AOO	+	+	11.3	5	1.4	10	2.8	25	5.2							Gerberick et al. (2005)
2,4,6-Trichloro-1,3,5-triazine	108-77-0	AOO	+	+	0.09	1	21.80	2.5	28.90	5	34.00							Gerberick et al. (2005)
2,4-Heptadienal	5910-85-0	AOO	+	+	4.0	0.5	1.1	1	1.4	2.5	1.9	5	3.7	10	8.10			Gerberick et al. (2005)
2,5-Diaminotoluene	95-70-5	DMSO	+	+	0.17	0.125	2.6	0.25	3.5	0.5	4.1	1	5.5					Gerberick et al. (2005)
2-Acetylcyclo-hexanone	874-23-7	ACE	-	-	NC	10	0.8	20	0.7	40	0.8							Gerberick et al. (2005)
2-Amino-6-chloro-4-nitrophenol	6358-09-4	AOO	+	+	2.2	0.1	1.7	0.25	1.4	0.5	2.1	1	1.5	2.5	3.4			Gerberick et al. (2005)
2-Aminophenol	95-55-6	AOO	+	+	0.4	0.5	3.5	1	5	2.5	7.4							Gerberick et al. (2005)
2-Bromotetra-decanoic acid	10520-81-7	AOO	+	+	3.4	5	4.7	10	7.7	25	10.1							Gerberick et al. (2005)
2-Hydroxyethyl acrylate	818-61-1	AOO	+	+	1.4	5	10.70	10	14.80	25	18.10							Gerberick et al. (2005)
2-Hydroxypropyl methacrylate	923-26-2	AOO	-	-	NC	10	1.1	25	1.2	50	1.3							Gerberick et al. (2005)
2-Mercapto-benzothiazole	149-30-4	DMF	+	+	1.7	1	2.3	3	4.4	10	8.6							Gerberick et al. (2005)
2-Methoxy-4-methylphenol	93-51-6	AOO	+	+	5.8	4.2	1.80	8.4	5.00	21	8.50							Gerberick et al. (2005)
2-Methyl-2H-isothiazol-3-one	2682-20-4	AOO	-	+	1.9	0.25	1.50	0.5	1.50	1	1.8	2.5	3.8	5	2.5			Gerberick et al. (2005)
2-Methyl-4H,3,1-benzoxazin-4-one	525-76-8	DMSO	+	+	0.7	5	7.60	10	9.20	25	10.80							Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
2-Methyl-5-hydroxy-ethylaminophenol	55302-96-0	AOO	+	+	0.4	0.1	1.20	0.25	0.80	0.5	3.60	1	2.6	2.5	7.4			Gerberick et al. (2005)
2-Methylundecanal	110-41-8	AOO	+	+	10.0	0.5	1.40	1	1.30	2.5	1.30	5	2.40	10	3.00			Gerberick et al. (2005)
2-Nitro-p-phenylenediamine	5307-14-2	AOO	+	+	0.4	0.1	1.80	0.25	2.20	0.5	3.30	1	7.90	2.5	11.90			Gerberick et al. (2005)
3 and 4-(4-Hydroxy-4-methylpentyl)-3-cyclohexane-1-carboxaldehyde	31906-04-4	AOO	+	+	17.1	1	0.60	2.5	0.70	5	0.60	10	1.30	25	4.90			Gerberick et al. (2005)
3, 3', 4', 5-Tetrachloro-salicylanilide	1154-59-2	ACE	+	+	0.04	0.25	11.20	0.5	14.40	1	18.00							Gerberick et al. (2005)
3,4-Dihydrocoumarin	119-84-6	AOO	+	+	5.6	2.5	1.6	5	2.5	10	6.6							Gerberick et al. (2005)
3,5,5-Trimethyl-hexanoyl chloride	36727-29-4	AOO	+	+	2.7	5	7.20	10	12.00	25	19.00							Gerberick et al. (2005)
3-Aminophenol	591-27-5	AOO	+	+	3.2	2.5	2.8	5	3.5	10	5.7							Gerberick et al. (2005)
3-Bromomethyl-5, 5'-dimethyl-dihydro-2(3H)-furanone	154750-20-6	AOO	+	+	**3.5**	3.19	2.7	6.37	5.1	12.74	7.1							Gerberick et al. (2005)
3-Dimethylamino-propylamine	109-55-7	AOO	+	+	2.2	0.5	1.30	1	1.10	2.5	3.50	5	7.00	10	13.90			Gerberick et al. (2005)
3-Ethoxy-1-(2',3',4',5'-tetramethylphenyl)propane-1,3-dione	170928-69-5	ACE	+	+	33	10	1.1	20	1.7	40	3.7							Gerberick et al. (2005)
3-Methyl-4-phenyl-1,2,5-thiadiazole-1,1-dioxide	3775-21-1	AOO	+	+	1.4	0.1	1.3	0.25	1.1	0.5	2.1	1	1.9	2.5	5.6			Gerberick et al. (2005)
3-Methyleugenol	186743-26-0	AOO	+	+	32	11	1.5	27	2.3	54	6.4							Gerberick et al. (2005)
3-Methylisoeugenol	186743-29-3	AOO	+	+	3.6	2.5	2.20	5.5	4.30	11	6.00							Gerberick et al. (2005)
3-Phenylene-diamine	108-45-2	AOO	+	+	0.5	2.5	11.70	5	15.50	10	19.20							Gerberick et al. (2005)
3-Propyliden-phthalide	17369-59-4	AOO	+	+	3.7	5	4.90	10	9.10	25	15.10							Gerberick et al. (2005)
4-(N-Ethyl-N-2-methan-sulfamido-ethyl)-2-methyl-1,4,-phenylene-diamine	25646-71-3	DMSO	+	+	0.6	0.1	1.2	1	4.5	5	5.9	10	6.3					Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
4,4,4-Trifluoro-1-phenylbutane-1,3-dione	362-06-7	ACE	+	+	20	10	2.10	20	3.00	40	4.60							Gerberick et al. (2005)
4-Allylanisole	140-67-0	AOO	+	+	18	10	1.20	25	4.7	50	4.5	100	8					Gerberick et al. (2005)
4-Hydroxybenzoic acid	99-96-7	DMSO	-	-	NC	5	1.4	10	1.5	25	1.3							Gerberick et al. (2005)
4'-Methoxy-acetophenone	100-06-1	AOO	-	-	NC	10	1.3	25	1	50	1							Gerberick et al. (2005)
4-Methylamino-phenol sulfate	55-55-0	DMF	+	+	0.8	0.5	2.50	1	3.40	2.5	6.70							Gerberick et al. (2005)
4-Nitrobenzyl bromide	100-11-8	AOO	+	+	0.05	0.01	0.90	0.03	1.30	0.05	3.50	0.1	11.50					Gerberick et al. (2005)
4-Phenylene-diamine	106-50-3	AOO	+	+	0.16	0.05	1.90	0.1	2.30	0.25	4.00	0.5	5.70	1.0	6.60			Gerberick et al. (2005)
5,5-Dimethyl-3-methylenedihydro-2(3H)-furanone	29043-97-8	AOO	+	+	**2.0**	2	3	4	7.4	8	9.2							Gerberick et al. (2005)
5-Chloro-2-methyl-4-isothiazolin-3-one	26172-55-4	DMF	+	+	0.009	0.01	3.50	0.03	12.30	0.1	22.70							Gerberick et al. (2005)
5-Methyl-2,3-hexanedione	13706-86-0	AOO	+	+	26	25	2.90	50	6.00	100	14.30							Gerberick et al. (2005)
5-Methyleugenol	186743-25-9	AOO	+	+	13	11	2.7	27	4.9	54	4.3							Gerberick et al. (2005)
6-Methylcoumarin	92-48-8	ACE	-	-	NC	5	1	10	1	25	1.1							Gerberick et al. (2005)
6-Methyleugenol	186743-24-8	AOO	+	+	17	11	1.9	27	4.9	54	8.3							Gerberick et al. (2005)
6-Methylisoeugenol	13041-12-8	AOO	+	+	1.6	2.5	5.90	5.5	11.10	11	15.7							Gerberick et al. (2005)
7,12-Dimethyl-benz[a]anthracene	57-97-6	DMF	+	+	0.006	0.025	7.60	0.5	17.70	1	15.60							Gerberick et al. (2005)
7-Bromotetradecane	74036-97-8	AOO	+	+	21	5	0.9	10	1.2	25	3.6							Gerberick et al. (2005)
Abietic acid	514-10-3	AOO	+	+	15	5	1.5	10	2	25	5.2							Gerberick et al. (2005)
alpha-Amyl cinnamic aldehyde	122-40-7	AOO	+	+	10.6	1	1.5	2.5	1.7	5	2.2	10	2.8	25	8.2			Gerberick et al. (2005)
alpha-Butyl cinnamic aldehyde	7492-44-6	AOO	+	+	11.2	1	1.4	2.5	1.7	5	1.7	10	2.1	25	13			Gerberick et al. (2005)
alpha-Methyl cinnamic aldehyde	101-39-3	AOO	+	+	4.5	1	1.80	2.5	1.50	5	3.40	10	3.3	25	15.3			Gerberick et al. (2005)
alpha-Methylphenyl-acetaldehyde	93-53-8	AOO	+	+	6.3	0.5	2	1	2.2	2.5	1	5	2.2	10	5.2			Gerberick et al. (2005)
Aniline	62-53-3	AOO	+	+	89	5	1.1	10	0.9	25	2	50	1.9	100	3.3			Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Benzaldehyde	100-52-7	AOO	-	-	NC	1	2.1	2.5	1.7	5	2.2	10	1.8	25	2			Gerberick et al. (2005)
Benzene-1,3,4-tricarboxylic anhydride	552-30-7	AOO	+	+	9.2	1	1.10	2.5	2.00	5	2.00	10	3.20	25	4.60			Gerberick et al. (2005)
Benzo[a]pyrene	50-32-8	AOO	+	+	0.0009	0.5	17.6	1	19.2	2.5	27							Gerberick et al. (2005)
Benzocaine	94-09-7	AOO	-	-	NC	2.5	2.1	5	1.8	10	2.7	25	1.8	50	1.2			Gerberick et al. (2005)
Benzoquinone	106-51-4	AOO	+	+	0.0099	0.5	36.4	1	42.3	2.5	52.3							Gerberick et al. (2005)
Benzyl benzoate	120-51-4	AOO	+	+	17	5	2.3	25	3.5									Gerberick et al. (2005)
Benzyl bromide	100-39-0	AOO	+	+	0.2	0.25	3.5	0.5	11.5	1	16.1	2.5	16.4	5	25.1			Gerberick et al. (2005)
Benzylidene acetone	122-57-6	AOO	+	+	3.7	10	8.5	25	13.6	50	12.8							Gerberick et al. (2005)
beta-Propiolactone	57-57-8	AOO	+	+	0.15	0.025	1.50	1.0	13.00	2.5	19.90							Gerberick et al. (2005)
bis-1,3-(2',5'-dimethylphenyl)-propane-1,3-dione		ACE	-	-	NC	10	1.8	20	1.6	40	2.1							Gerberick et al. (2005)
Bisphenol A-diglycidyl ether	1675-54-3	AOO	+	+	1.5	1	2	3	6	10	17.4							Gerberick et al. (2005)
Butyl glycidyl ether	2426-08-6	AOO	+	+	30.9	10	1.40	25	2.20	50	5.60							Gerberick et al. (2005)
C11-azlactone	176665-06-8	AOO	+	+	16	8.3	1.30	20.7	4.00	41.3	8.50							Gerberick et al. (2005)
C15-azlactone	176665-09-1	AOO	+	+	18	10	1.80	25	4.10	50	7.50							Gerberick et al. (2005)
C17-azlactone	176665-11-5	AOO	+	+	19	10.87	1.70	27.17	4.30	54.33	4.60							Gerberick et al. (2005)
C19-azlactone		AOO	-	+	26	11.73	2.50	29.33	3.10	58.67	2.50							Gerberick et al. (2005)
C4-azlactone	176664-99-6	AOO	+	+	1.4	0.52	1.10	1.31	2.30	2.62	4.10	5.23	11.70					Gerberick et al. (2005)
C6-azlactone	176665-02-4	AOO	+	+	1.3	0.61	1.20	1.52	3.50	3.05	7.60							Gerberick et al. (2005)
C9-azlactone	176665-04-6	AOO	+	+	2.8	1.85	1.40	3.7	4.60	7.4	10.10							Gerberick et al. (2005)
Camphorquinone	465-29-2	AOO	-	+	10	5	2.8	10	3	25	1.7							Gerberick et al. (2005)
Chlorobenzene	108-90-7	AOO	-	-	NC	5	1.1	10	1.7	25	1.6							Gerberick et al. (2005)
Cinnamic alcohol	104-54-1	AOO	+	+	21	10	1.8	25	3.5	50	3.9	90	5.7					Gerberick et al. (2005)
Cinnamic aldehyde	104-55-2	AOO	+	+	3.0	0.5	1.40	1.0	0.90	2.5	1.90	5.0	7.10	10.0	15.80			Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
cis-6-Nonenal	2277-19-2	AOO	+	+	22	10	1.60	25.0	3.30	50.0	4.50	100.0	13.70					Gerberick et al. (2005)
Citral	5392-40-5	AOO	+	+	13	5	1.20	10	2.10	25	6.30							Gerberick et al. (2005)
Clotrimazole	23593-75-1	AOO	+	+	4.8	2.5	1.6	5	3.1	10	3							Gerberick et al. (2005)
Coumarin	91-64-5	AOO	-	-	NC	5	2.70	10	2.90	25	2.30							Gerberick et al. (2005)
Cyclamen aldehyde	103-95-7	AOO	+	+	22.0	1	1.40	2.5	1.30	10	1.80	25	3.3	50	5.2			Gerberick et al. (2005)
Diethyl maleate	141-05-9	AOO	+	+	5.8	25	16.30	50	22.60	100	13.10							Gerberick et al. (2005)
Diethyl sulfate	64-67-5	AOO	+	+	3.3	1	0.8	2.5	1.9	10	12							Gerberick et al. (2005)
Diethyl-acetaldehyde	97-96-1	AOO	+	+	76	25	1.2	50	0.8	75	2.4	100	16.3					Gerberick et al. (2005)
Diethylenetriamine	111-40-0	AOO	+	+	5.8	10	6.40	25	12.10									Gerberick et al. (2005)
Diethylpthalate	84-66-2	AOO	-	-	NC	25	1.00	50	1.30	100	1.50							Gerberick et al. (2005)
Dihydroeugenol	2785-87-7	AOO	+	+	6.8	5.1	2.70	10.1	3.60	25.3	7.80							Gerberick et al. (2005)
Dimethyl sulfate	77-78-1	AOO	+	+	0.19	0.25	3.80	0.5	6.00	1	5.70							Gerberick et al. (2005)
Dimethyl sulfoxide	67-68-5	AOO	+	+	72	25	2.70	50	2.30	100	3.90							Gerberick et al. (2005)
Dodecyl methanesulfonate	51323-71-8	AOO	+	+	8.8	5	2.10	10	3.30	25	9.00							Gerberick et al. (2005)
Ethyl benzoylacetate	94-02-0	ACE	-	-	NC	10	0.9	20	0.9	40	1.2							Gerberick et al. (2005)
Ethyl vanillin	121-32-4	AOO	-	-	NC	2.5	0.65	5	1.05	10	0.74	25	0.36	50	0.29			Gerberick et al. (2005)
Ethyl acrylate	140-88-5	AOO	+	+	28	10	1.2	25	2.7	50	5							Gerberick et al. (2005)
Ethylene glycol dimethacrylate	97-90-5	MEK	+	+	28	10	1.20	25	2.40	50	7.00							Gerberick et al. (2005)
Ethylenediamine free base	107-15-3	AOO	+	+	2.2	0.1	1.10	0.25	1.20	0.5	1.60	1	1.90	2.5	3.30	5	6.10	Gerberick et al. (2005)
Eugenol	97-53-0	AOO	+	+	13	2.5	1.60	5	1.50	10	2.40	25	5.50					Gerberick et al. (2005)
Farnesal	502-67-0	AOO	+	+	12	1	0.60	2.5	1.10	5	1.70	10	2.50	25	7.00			Gerberick et al. (2005)
Fluorescein isothiocyanate	27072-45-3	ACE/ DBP (50:50)	+	+	0.143	0.5	8.60	1	11.70	2.5	16.60							Gerberick et al. (2005)
Formaldehyde	50-00-0	ACE	+	+	0.61	0.093	1.10	0.185	2.30	0.37	2.30	0.925	3.90	1.85	4.00			Gerberick et al. (2005)
Furil	492-94-4	AOO	-	-	NC	5	1.20	10	1.70	25	2.20							Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Geraniol	106-24-1	EtOH/ DEP (75:25)	+	+	26	1	1.00	3	1.00	10	1.30	30	3.40	50	3.90			Gerberick et al. (2005)
Glutaraldehyde	111-30-8	ACE	+	+	0.1	0.05	1.30	0.125	4.30	0.25	7.60	0.5	11.60	1.25	17.70	2.5	18.00	Gerberick et al. (2005)
Glycerol	56-81-5	DMF	-	-	NC	25	1.10	50	0.70	100	0.50							Gerberick et al. (2005)
Glyoxal	107-22-2	AOO	+	+	1.4	1	2.50	2.5	4.20	5	5.20	10	10.30	25	15.80			Gerberick et al. (2005)
Hexane	110-54-3	AOO	-	-	NC	25	0.8	50	0.8	100	2.2							Gerberick et al. (2005)
Hexyl cinnamic aldehyde	101-86-0	AOO	+	+	11	2.5	1.30	5	1.10	10	2.50	25	10.00	50	17.00			Gerberick et al. (2005)
Hydroxycitronellal	107-75-5	AOO	+	+	33	2.5	2.20	5	1.00	10	0.80	25	1.10	50	7.10			Gerberick et al. (2005)
Imidazolidinyl urea	39236-46-9	DMF	+	+	24	10	1.70	25	3.10	50	5.50							Gerberick et al. (2005)
Isoeugenol	97-54-1	AOO	+	+	1.7	0.5	1.00	1	1.10	5	12.40							Gerberick et al. (2005)
Isononanoyl chloride	57077-36-8	AOO	+	+	2.7	5	6.60	10	10.60	25	12.60							Gerberick et al. (2005)
Isopropanol	67-63-0	AOO	-	-	NC	10	1.7	25	1.1	50	1							Gerberick et al. (2005)
Isopropyl myristate	110-27-0	AOO	+	+	44	25	2.10	50	3.30	100	3.40							Gerberick et al. (2005)
Isopropyleugenol	51474-90-9	AOO	-	-	NC	12	1.8	29.0	1.8	59.0	2.2							Gerberick et al. (2005)
Isopropyl-isoeugenol	2953-00-7	AOO	+	+	0.6	0.6	3.00	1.2	5.70	3.0	10.70							Gerberick et al. (2005)
Kanamycin	59-01-8; 8063-07-8	AOO	-	-	NC	5	2.2	10	0.8	25	1							Gerberick et al. (2005)
Lactic acid	598-82-3	DMSO	-	-	NC	5	1	10	1.4	25	2.2							Gerberick et al. (2005)
Lauryl gallate	1166-52-5	DMSO	+	+	0.3	1	12.10	10	29.70	25	29.30	50	36					Gerberick et al. (2005)
Linalool alcohol	78-70-6	AOO	+	+	30	25	2.5	50	4.8	100	8.3							Gerberick et al. (2005)
Methyl dodecanesulfonate	2374-65-4	AOO	+	+	0.4	1	21.6	2.5	39.9	5	48.6							Gerberick et al. (2005)
Methyl hexacecyl sulfonate	4230-15-3	AOO	-	-	NC	5	1	10	1.3	25	1.5							Gerberick et al. (2005)
Methyl hexadecane-sulfonate	26452-48-2	AOO	+	+	0.8	5	26.7	10	35.4	25	32.9							Gerberick et al. (2005)
Methyl methanesulfonate	66-27-3	AOO	+	+	2.7	0.25	0.7	1	0.7	10	3.6							Gerberick et al. (2005)
Methyl salicylate	119-36-8	AOO	-	-	NC	1.0	1	2.5	1.1	5.0	1.6	10	1.4	20	0.9			Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Methyl(2-sulfomethyl) octadecanoate		AOO	+	+	2.0	2.5	5.10	5.0	11.60	10.0	25.60							Gerberick et al. (2005)
Methyl-2-nonynoate	111-80-8	EtOH (80%)	+	+	2.5	5	10.4	10	17.7	20	24.4							Gerberick et al. (2005)
Methyl-4-hydroxybenzoate	99-76-3	DMF	-	-	NC	10	0.80	25	0.90	50	0.80							Gerberick et al. (2005)
N-Ethyl-N-nitrosourea	759-73-9	AOO	+	+	1.1	0.25	1.00	1	2.70	10	22.30							Gerberick et al. (2005)
N-Methyl-N-nitrosourea	684-93-5	AOO	+	+	0.05	0.05	2.7	0.1	7.1	0.25	15.4							Gerberick et al. (2005)
Nonanoyl chloride	764-85-2	AOO	+	+	1.8	5	12.70	10	19.40	25	20.90							Gerberick et al. (2005)
Octanoic acid	124-07-2	AOO	-	-	NC	10	0.70	25.0	1.00	50.0	1.60							Gerberick et al. (2005)
Oleyl methane sulfonate	35709-09-2	AOO	+	+	25	5	1.00	10.0	1.30	25.0	3.00							Gerberick et al. (2005)
Oxalic acid	144-62-7	DMF	+	+	15	5	2.40	10	2.80	25	3.40							Gerberick et al. (2005)
Oxazolone	15646-46-5	AOO	+	+	0.003	0.003	2.90	0.005	4.90	0.01	12.00	0.025	22.00	0.05	33.00			Gerberick et al. (2005)
Palmitoyl chloride	112-67-4	AOO	+	+	8.8	5	2.10	10	3.30	25	4.50							Gerberick et al. (2005)
Penicillin G	61-33-6	DMSO	+	+	30	2.5	1.00	5.0	1.00	10	1.40	25.0	2.10	50.0	6.60			Gerberick et al. (2005)
Pentachlorophenol	87-86-5	DMSO	+	+	20	10	2.10	25.0	3.50	50.0	5.40							Gerberick et al. (2005)
Perillaldehyde	2111-75-3	AOO	+	+	4.0	0.5	1.20	1.0	1.10	2.5	0.90	5.0	4.30					Gerberick et al. (2005)
Phenyl benzoate	93-99-2	AOO	+	+	20	5	2.30	10	2.10	25	3.50							Gerberick et al. (2005)
Phenylacetaldehyde	127-78-1	AOO	+	+	3.0	1	0.70	2.5	1.80	5.0	7.80	10	8.80	25.0	19.00			Gerberick et al. (2005)
p-Methylhydro-cinnamic aldehyde	5406-12-2	AOO	+	+	14	2.5	1.20	5	1.40	10	2.60	25	4.2	50	10.7			Gerberick et al. (2005)
Potassium dichromate	7778-50-9	DMSO	+	+	0.08	0.025	1.60	0.05	1.40	0.1	3.80	0.25	5.30	0.5	16.10			Gerberick et al. (2005)
Propylene glycol	57-55-6	dH_2O	-	-	NC	50	1.20	100.0	1.60									Gerberick et al. (2005)
Propylparaben	94-13-3	AOO	-	-	NC	5	1.40	10	1.00	25.0	1.30							Gerberick et al. (2005)
p-tert-Butyl-a-ethyl-hydrocinnamal	80-54-6	AOO	+	+	19	1	1.3	2.5	2.5	10	2	25	3.7	50	9.3			Gerberick et al. (2005)
Pyridine	110-86-1	AOO	+	+	72	25	1.10	50.0	2.30	100.0	3.90							Gerberick et al. (2005)
R(+)-Limonene	5989-27-5	AOO	+	+	69	25	1.8	50	2.4	100	4							Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Saccharin	81-07-2	DMSO	-	-	NC	25	1.30	50.0	1.30	75.0	1.50							Gerberick et al. (2005)
Salicylic acid	69-72-7	AOO	-	-	NC	5	0.80	10	1.50	25.0	2.50							Gerberick et al. (2005)
Sodium lauroyl lactylate	13557-75-0	AOO	+	+	15	5	1.40	10.0	2.50	25.0	3.90							Gerberick et al. (2005)
Sodium lauryl sulfate	151-21-3	DMF	+	+	14	1	0.90	2.5	1.10	5.0	1.70	10.0	2.60	20.0	3.50			Gerberick et al. (2005)
Sodium-3,3,5-trimethyl-hexanoyloxy-benzenesulfonate	94612-91-6	DMSO	+	+	6.4	5	2.30	10.0	4.80	25.0	7.80							Gerberick et al. (2005)
Streptomycin sulfate	3810-74-0	DMF	-	-	NC	2.5	1.20	5.0	1.40	10	1.30	25.0	2.00	50.0	1.90			Gerberick et al. (2005)
Sulfanilamide	63-74-1	DMF	-	-	NC	10.0	1.00	25.0	1.00	50.0	0.90							Gerberick et al. (2005)
Sulfanilic acid	121-57-3	DMF	-	-	NC	5.0	1.50	10	1.90	25.0	2.20							Gerberick et al. (2005)
Tartaric acid	87-69-4	DMF	-	-	NC	5	1.00	10	0.90	25	1.50							Gerberick et al. (2005)
Tetramethyl thiuram disulfide	137-26-8	AOO	+	+	5.2	2.5	2.40	5.0	2.90	10	5.10							Gerberick et al. (2005)
trans-2-Decenal	3913-71-1	AOO	+	+	2.5	0.5	1.30	1	1.10	2.5	3.00	5	6	10	9.5			Gerberick et al. (2005)
trans-2-Hexenal	6728-26-3	AOO	+	+	5.5	0.5	1.2	1	1.2	2.5	2.3	5	2.6	10	6.4			Gerberick et al. (2005)
trans-Anethol	104-46-1	AOO	+	+	2.3	4.5	13.50	9	24.7	22.6	37.3							Gerberick et al. (2005)
Undec-10-enal	112-45-8	AOO	+	+	6.8	5.0	1.70	10	5.30	25.0	7.50	50.0	8.70	75.0	8.80			Gerberick et al. (2005)
Vanillin	121-33-5	AOO	-	-	NC	2.5	0.90	5.0	1.40	10	1.50	25.0	1.20	50.0	1.40			Gerberick et al. (2005)
Vinylidene dichloride	75-35-4	AOO	-	-	NC	10	0.80	25.0	0.80	50.0	0.90							Gerberick et al. (2005)
Vinylpyridine	1337-81-1	AOO	+	+	1.6	2.5	7.40	5.0	14.20	10	14.80							Gerberick et al. (2005)
(16-beta)-21-(Acetyloxy)-17-hydroxy-16-methylpregna-1,4,9(11)-triene-3,20-dione	910-99-6	DMF	-	-	NC	2.5	1.30	5	1.27	10	0.89							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
(1r)-1,2,3,4-Tetrahydro-6,7-dimethoxy-2-methyl-1-[(3,4,5-trimethoxyphenyl) methyl]isoquino-line [r-(r*,r*)]-2,3-bis(benzoyloxy)-butanedioate (1:1)	104832-01-1	DMF	-	-	NC	10	0.92	25	1.12	50	1.27							Glaxo SmithKline, Submitted by M.J. Olson
(2-Bromo-5-propoxyphenyl)(2-hydroxy-4-methoxyphenyl)-methadone	190965-45-8	ACE	-	-	NC	0.5	1.10	5	0.90	50	1.70							Glaxo SmithKline, Submitted by M.J. Olson
(2e)-2-[(2-Formyl-4-hydroxyphenyl)-methylidene]-butanedioic acid	773059-57-7	DMF	+	+	48	0.5	0.80	5	1.45	50	3.08							Glaxo SmithKline, Submitted by M.J. Olson
(2-Oxo-1-phenyl-pyrrolidin-3-yl)(triphenyl)-phosphonium bromide	148776-18-5	DMSO	-	-	NC	2.5	1.64	5	2.45	10	1.40							Glaxo SmithKline, Submitted by M.J. Olson
(2R,4S)-4-(4-Acetyl-1-piperazinyl)-n-{(1r)-1-[3,5-bis(trifluoro-methyl)phenyl]-ethyl}-2-(4-fluoro-2-methylphenyl)-n-methyl-1-piperidine-carboxamide monomethane-sulfonate	414910-30-8	DMF	-	-	NC	2.5	1.07	5	0.90	10	1.47							Glaxo SmithKline, Submitted by M.J. Olson
(2S,4S)-1-[(2s)-2-Amino-3,3-bis(4-fluorophenyl)-1-oxopropyl]-4-fluoro-2-pyrroli-dine carbonitrile	483369-58-0	DMSO	-	-	NC	5	0.99	10	1.60	25	2.44							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
(3as,4r,5s,6s,8r,9r,9ar,10r)-6-Ethenyldeca-hydro-5-hydroxy-4,6,9,10-tetramethyl-1-oxo-3a,9-propano-3ah-cyclopentacyclo-octen-8-yl [[(3-exo)-8-methyl-8-azabicyclo-[3.2.1]oct-3-yl]thio]-acetate	224452-66-8	AOO	-	-	NC	0.5	1.00	2	1.40	5	1.20							Glaxo SmithKline, Submitted by M.J. Olson
(3as,4r,5s,6s,8r,9r,9ar,10r)-6-Ethenyldeca-hydro-5-hydroxy-4,6,9,10-tetra-methyl-1-oxo-3a,9-propano-3ah-cyclopentacyclooct en-8-yl hydroxyacetate	125-65-5	DMF	-	-	NC	10	1.38	25	1.42	50	1.56							Glaxo SmithKline, Submitted by M.J. Olson
(3-Endo)-8-methyl-8-azabicyclo[3.2.1]-octan-3-ol	120-29-6	DMF	+	+	8.9	10	4.11	25	12.85	50	26.14							Glaxo SmithKline, Submitted by M.J. Olson
(3r,3as,6ar)-Hexahydrofuro-[2,3-b]furan-3-ol	156928-09-5	DMF	-	-	NC	1	0.78	3	0.91	10	0.95							Glaxo SmithKline, Submitted by M.J. Olson
(3r,3as,6ar)-Hexahydrofuro-[2,3-b]furan-3-yl [(1s,2r)-3-[(1,3-benzodioxol-5-ylsulfonyl)(2-methylpropyl)-amino]-2-hydroxy-1-[[4-[(2-methyl-4-thiazolyl)methoxy] phenyl]methyl]-propyl]carbamate	313682-08-5	DMF	-	-	NC	2.5	1.46	7.5	0.92	25	1.04							Glaxo SmithKline, Submitted by M.J. Olson
(3R6R)-3-(2,3-Dihydro-1h-inden-2-yl)-1-[(1r)-1-(2-methyl-1,3-oxazol-4-yl)-2-(4-morpholinyl)-2-oxoethyl]-6-[(1s)-1-methylpropyl]-2,5-piperazinedione	820957-38-8	DMF	-	-	NC	5	0.74	10	1.41	25	1.62							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
(4r,5s)-(-)-1,5-Dimethyl-4-phenyl-2-imidazolidinone	92841-65-1	DMSO	-	-	NC	10	1.80	25	2.10	50	1.90							Glaxo SmithKline, Submitted by M.J. Olson
(4r,5s)-1,5-Dimethyl-3-(1-oxo-2-propenyl)-4-phenyl-2-imidazo-lidinone	139109-23-2	ACE	+	+	0.004	0.5	14.10	5	19.50	50	16.80							Glaxo SmithKline, Submitted by M.J. Olson
(4S)-1-(tert-Butoxycarbonyl)-4-fluoro-l-prolinamide	426844-22-6	DMF	-	-	NC	2.5	1.12	5	0.99	10	1.60							Glaxo SmithKline, Submitted by M.J. Olson
(4S)-1-(tert-Butoxycarbonyl)-4-fluoro-l-proline	203866-13-1	DMF	-	-	NC	5	0.88	10	0.82	25	1.29							Glaxo SmithKline, Submitted by M.J. Olson
(4S,5R)-1-[(1R,2R,3S)-3-(1,3-Benzodioxol-5-yl)-1-(2-benzyloxy-4-methoxyphenyl)-1-hydroxy-6-propoxy-2-indanoyl]-3,4-dimethyl-5-phenyl-2-imidazolidinone	190965-47-0	DMF	-	-	NC	0.5	0.78	5	1.14	10	1.39							Glaxo SmithKline, Submitted by M.J. Olson
(Alpha-r)-n-alpha-dimethyl-3,5-bis(trifluoro-methyl	334477-60-0	DMF	-	-	NC	10	1.32	25	1.41	50	1.63							Glaxo SmithKline, Submitted by M.J. Olson
(R,S)-3-Amino-2,3,4,5-tetrahydro-n-(1-methylethyl)-2,4-dioxo-n,5-diphenyl-1h-1,5-benzodiazepine-1-acetamide	184944-86-3	PG	-	-	NC	5	1.06	10	1.00	25	1.19							Glaxo SmithKline, Submitted by M.J. Olson
(s)-(-)-1-Phenylpropyl-amine	3789-59-1	AOO	-	-	NC	0.5	0.81	5	0.69	50	0.99							Glaxo SmithKline, Submitted by M.J. Olson
(S)-2-(4-Fluoro-2-methylphenyl)4-piperidinone (s)-alpha-hydroxybenzene-acetic acid salt	414910-13-7	DMF	-	-	NC	10	1.79	25	1.85	50	2.10							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
[3aS-(3aAlpha, 4beta,5alpha, 6alpha,8beta, 9alpha,9abeta, 10S*)]-6-Ethenyldecahydro-5-hydroxy-4,6,9,10-tetramethyl-1-oxo-3a,9-propano-3aH-cyclopentacyclooct en-8-yl [(methylsulfonyl)-oxy]acetate	60924-38-1	DMF	-	-	NC	10	2.46	25	2.40	50	1.54							Glaxo SmithKline, Submitted by M.J. Olson
[4-(Ethoxymethyl)-2,6-dimethoxyphenyl]-boronic acid	591249-50-2	DMF	-	-	NC	10	0.87	25	0.58	50	1.00							Glaxo SmithKline, Submitted by M.J. Olson
[4S-[1(E),4Alpha, 5alpha]]-1-[3-[2-[4-Methoxy-2-(phenylmethoxy)-benzoyl]-4-propoxyphenyl]-1-oxo-2-propenyl]-3,4-dimethyl-5-phenyl-2-imidazoli-dinone	190965-46-9	DMF	-	-	NC	0.05	0.88	0.5	0.59	5	0.56							Glaxo SmithKline, Submitted by M.J. Olson
1-(4-Ethoxy-phenyl)-2-[4-(methyl-sulfonyl)phenyl]-ethanone	346413-00-1	DMSO	-	-	NC	1	1.82	2.5	2.34	5	1.89							Glaxo SmithKline, Submitted by M.J. Olson
1,1-Dimethylethyl [(1s)-1-[bis(4-fluorophenyl)-methyl]-2-[(2s,4s)-2-cyano-4-fluoro-1-pyrrolidinyl]-2-oxoethyl]carba-mate	483368-24-7	AOO	-	-	NC	10	0.97	25	0.81	50	0.99							Glaxo SmithKline, Submitted by M.J. Olson
1,1-Dimethylethyl [(1s)-2-[4-[(2-methyl-4-thiazolyl)methoxy] phenyl]-1-(2s)-oxiranylethyl]-carbamate	313680-92-1	DMF	-	-	NC	5	1.04	10	0.84	25	1.16							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
1,1-Dimethylethyl 3-[[[[(3s)-2,3,4,5-tetrahydro-1-[2-[(1-methylethyl)phenyl amino]-2-oxoethyl]-2,4-dioxo-5-phenyl-1h-1,5-benzodiazepin-3-yl]amino] carbonyl]amino] benzoate	305366-94-3	DMF	+	+	38	10	2.10	25	2.20	50	3.80							Glaxo SmithKline, Submitted by M.J. Olson
1,2,3,5,6,7-Hexahydro-2-thioxo-4h-cyclopentapyrimi-din-4-one	35563-27-0	PG	-	-	NC	5	0.63	10	1.71	25	1.37							Glaxo SmithKline, Submitted by M.J. Olson
1,3-Benzodioxazole-5-sulphonyl chloride	115010-10-1	AOO	+	+	0.4	10	13.54	25	16.56	50	16.76							Glaxo SmithKline, Submitted by M.J. Olson
1-[3-(Cyclopentyl-oxy)-4-methoxy-phenyl]-4-oxocyclohexane carbonitrile	152630-47-2	DMSO	-	-	NC	10	2.70	25	2.80	50	2.20							Glaxo SmithKline, Submitted by M.J. Olson
1-[5-[(4-Fluorophenyl)meth yl]-2-furanyl]ethanone	280571-34-8	AOO	-	-	NC	0.5	1.00	5	1.00	50	1.20							Glaxo SmithKline, Submitted by M.J. Olson
14-Hydroxynor-morphinone	84116-46-1	PG	+	+	8.4	5	1.20	10	3.88	25	6.24							Glaxo SmithKline, Submitted by M.J. Olson
2-(3,4-Dimethyl-phenyl)-5--methyl-2,4-dihydropyrazol-3-one	18048-64-1	DMF	+	+	IDR[3]	2.5	4.41	7.5	4.82	25	8.46							Glaxo SmithKline, Submitted by M.J. Olson
2-(4-Ethoxyphenyl)-3-[4-(methyl-sulfonyl)phenyl-]pyrazolo[1,5-b]-pyridazine	221148-46-5	DMF	-	-	NC	5	0.98	10	0.97	25	0.94							Glaxo SmithKline, Submitted by M.J. Olson
2-(4-Oxopentyl)-1h-isoindole-1,3(2h)-dione	3197-25-9	AOO	-	-	NC	0.25	0.58	2.5	1.54	25	0.67							Glaxo SmithKline, Submitted by M.J. Olson
2-(Benzyl)tert-butyl)amino)-1-(alpha,4-dihydroxy-m-tolyl)ethane	24085-03-8	DMF	-	-	NC	10	0.67	25	1.10	50	1.28							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
2,3,4,5-Tetrahydro-n-(1-methylethyl)-2,4-dioxo-n,5-diphenyl-3-[(phenylmethoxy)-imino]-1h-1,5-benzodiazepine-1-acetamide	305366-97-6	DMF	-	-	NC	5	1.18	10	1.76	25	1.67							Glaxo SmithKline, Submitted by M.J. Olson
2,3-Dimethyl-2h-indazol-6-amine	444731-72-0	DMF	-	-	NC	5	0.74	10	0.95	25	1.16							Glaxo SmithKline, Submitted by M.J. Olson
2,4-Dichloro-pyrimidine	3934-20-1	DMF	+	+	0.7	0.25	0.76	0.75	3.46	2.5	8.64							Glaxo SmithKline, Submitted by M.J. Olson
2,6-Dimethoxy-4-methyl-5-[3-(trifluoromethyl)-phenoxy]-8-quinolinamine	106635-86-3	AOO	-	-	NC	0.5	2.30	5	2.10	50	1.80							Glaxo SmithKline, Submitted by M.J. Olson
2,6-Dimethoxy-4-methyl-8-nitro-5-[3-(trifluoromethyl)-phenoxy]quinoline	189746-15-4	PG	+	+	3.5	3	2.90	10	4.29	30	5.34							Glaxo SmithKline, Submitted by M.J. Olson
2-[(Benzyloxy)-imino]malonic acid	305366-96-5	AOO	-	-	NC	10	1.17	25	1.88	50	2.40							Glaxo SmithKline, Submitted by M.J. Olson
2-[1-(4-Bromophenyl)-1-phenylethoxy]-n,n-dimethyl-ethanamine hydrochloride	13977-28-1	DMF	+	+	5.5	0.5	2.38	5	2.88	15	5.08							Glaxo SmithKline, Submitted by M.J. Olson
2-Amino-di-phenylamine	534-85-0	AOO	+	+	0.5	10	10.20	25	12.40	50	7.70							Glaxo SmithKline, Submitted by M.J. Olson
2-Aminoethyl-methylsulfone	49773-20-8	0.5% Tween 80 in H_2O	-	-	NC	10	0.40	25	0.30	50	0.30							Glaxo SmithKline, Submitted by M.J. Olson
2-Benzyl-tert-butylamino-3'-hydroxymethyl-4'-hydroxyaceto-phenone hydrochloride	24085-08-3	DMF	+	+	22	0.5	0.96	5	1.54	50	5.44							Glaxo SmithKline, Submitted by M.J. Olson
2-Bromo-5-hydroxy-benzaldehyde	2973-80-0	AOO	+	+	2.6	0.5	1.25	5	4.93	50	21.40							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
2-Bromo-5-propoxybenzoic acid	190965-43-6	ACE	-	-	NC	0.5	0.60	5	0.70	50	1.10							Glaxo SmithKline, Submitted by M.J. Olson
2-Chloro-6-methoxy-4-methylquinoline	6340-55-2	DMF	-	-	NC	0.1	1.07	1	0.98	10	1.11							Glaxo SmithKline, Submitted by M.J. Olson
2-chloro-1-[(3-fluorophenyl)-methoxy]-4-nitrobenzene	443882-99-3	AOO	+	+	IDR[3]	5	3.96	10	2.62	25	3.22							Glaxo SmithKline, Submitted by M.J. Olson
2-Nitro-4-(propylthio)benzen-amine	54393-89-4	AOO	Equiv	+	IDR[3]	0.5	1.97[4]	5	1.34[4]	15	8.00[5]							Glaxo SmithKline, Submitted by M.J. Olson
3,4-Dichloroaniline hydrochloride	95-76-1	DMF	+	+	18	0.25	1.02	2.5	1.75	25	3.53							Glaxo SmithKline, Submitted by M.J. Olson
3-[(2r)-3-[[2-(2,3-Dihydro-1h-inden-2-yl)-1,1-dimethyl-ethyl]amino]-2-hydroxypropoxy]-4,5-difluoro-benzene propanoic acid	753449-67-1	DMF	-	-	NC	5	0.71	15	1.02	50	1.28							Glaxo SmithKline, Submitted by M.J. Olson
3'-[(2z)-[1-(3,4-Dimethylphenyl)-1,5-dihydro-3-methyl)-5-oxo-4h-pyrazol-4-ylidene]hydrazino]-2'-hydroxy-[1,1'-biphenyl]-3-carboxylic acid, compound with 2-aminoethanol (2:1)	496775-62-3	AOO	-	-	NC	5	1.38	15	1.05	50	0.84							Glaxo SmithKline, Submitted by M.J. Olson
3-[4-[(6-Bromohexyl)oxy]-butyl]benzene-sulfonamide	452342-04-0	AOO	-	-	NC	10	1.02	25	0.82	50	0.68							Glaxo SmithKline, Submitted by M.J. Olson
3-Chloro-4-fluorobenzoyl chloride	65055-17-6	PG	+	+	7.8	3	2.24	10	3.36	30	8.99							Glaxo SmithKline, Submitted by M.J. Olson
3-Fluoro-5-(3-pyridinyl)benzen-amine	181633-36-3	DMSO	+	+	15	0.5	1.90	5	2.20	50	5.90							Glaxo SmithKline, Submitted by M.J. Olson
3-Hydroxy-2-phenyl-4-quinolinecarboxylic acid	485-89-2	DMSO	-	-	NC	0.05	0.56	0.5	0.79	5	1.04							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
3-Hydroxy-4-methoxy-benzaldehyde	621-59-0	DMF	-	-	NC	0.25	0.75	2.5	1.15	25	1.35							Glaxo SmithKline, Submitted by M.J. Olson
3-Propoxybenzoic acid	190965-42-5	ACE	-	-	NC	0.5	1.10	5	1.20	50	1.10							Glaxo SmithKline, Submitted by M.J. Olson
4-(Bromomethyl)-benzoic acid ethyl ester	26496-94-6	AOO	+	+	IDR[3]	0.5	11.73	5	12.87	50	ND[6]							Glaxo SmithKline, Submitted by M.J. Olson
4'-(Trifluoro-methyl)-[1,1'-biphenyl]-4-carboxaldehyde	90035-34-0	DMF	-	-	NC	1	1.36	3	1.55	10	2.58							Glaxo SmithKline, Submitted by M.J. Olson
4-[4-[[(3R)-1-Butyl-3-[(r)-cyclohexyl-hydroxymethyl]-2,5-dioxo-1,4,9-triazaspiro[5.5]-undec-9-yl]methyl]phenoxy] benzoic acid	461443-59-4	DMSO	-	-	NC	2.5	1.07	5	1.27	10	1.63							Glaxo SmithKline, Submitted by M.J. Olson
4-Amino-3-nitrophenyl thiocyanate	54029-45-7	DMSO	+	+	0.8	0.5	2.30	1	3.32	5	3.55							Glaxo SmithKline, Submitted by M.J. Olson
4-Bromo-1-phthalimidopentane	59353-62-7	ACE	+	+	27	0.5	1.00	5	1.10	50	4.20							Glaxo SmithKline, Submitted by M.J. Olson
4-Chloro-6-iodoquinazoline	98556-31-1	AOO	+	+	IDR[3]	5	11.30	10	9.30	25	17.30							Glaxo SmithKline, Submitted by M.J. Olson
4-Fluoro-2-pyrrolidine-carboxamide	748165-40-4	DMF	-	-	NC	10	1.22	25	1.15	50	1.03							Glaxo SmithKline, Submitted by M.J. Olson
4-Iodo-1-phthalimido-pentane	63460-47-9	ACE	+	+	5.0	0.5	1.70	5	3.00	50	9.50							Glaxo SmithKline, Submitted by M.J. Olson
5-[[4-[(2,3-Dimethyl-2h-indazol-6-yl)-methylamino]-2-pyrimidinyl]amino] -2-methylbenzene-sulfonamide	444731-52-6	AOO	-	-	NC	5	1.13	10	0.91	25	0.91							Glaxo SmithKline, Submitted by M.J. Olson
5-Amino-2-methylbenzene-sulfonamide	6973-09-7	DMF	-	-	NC	5	1.36	10	1.12	25	1.42							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
5-Chloro-2,6-dimethoxy-4-methyl-8-nitroquinoline	189746-21-2	DMSO	+	+	IDR[3]	0.5	8.00	5	7.00	10	7.50							Glaxo SmithKline, Submitted by M.J. Olson
5-Chloro-2,6-dimethoxy-4-methylquinoline	189746-19-8	DMSO	-	-	NC	0.5	0.90	5	0.70	25	1.00							Glaxo SmithKline, Submitted by M.J. Olson
5'-Chloro-2'-hydroxy-3'-nitro-[1,1'-biphenyl]-3-carboxylic acid	376592-58-4	DMF	+	+	20	5	1.37	15	2.83	50	3.96							Glaxo SmithKline, Submitted by M.J. Olson
5-Chloro-6-methoxy-4-methyl-8-nitro-2(1h)quinolinone	189746-23-4	PG	+	+	IDR[3]	2.5	10.82	5	9.86	10	10.72							Glaxo SmithKline, Submitted by M.J. Olson
5-Methoxy-2-nitro-4-(trifluoromethyl) benzene acetonitrile	178896-77-0	DMSO	-	-	NC	0.5	1.30	5	1.50	50	1.60							Glaxo SmithKline, Submitted by M.J. Olson
5-Methoxy-6-(trifluoromethyl)-2,3-dihydro-1h-indole	178896-79-2	DMSO	+	+	37	0.5	1.10	5	1.30	50	3.70							Glaxo SmithKline, Submitted by M.J. Olson
6-(Diethylamino)-1-hexanol	06947-12-2	PG	+	+	10	3	0.79	10	2.92	30	25.50							Glaxo SmithKline, Submitted by M.J. Olson
6-(Trifluoro-methyl)-2,3-dihydro-5-methyl-1h-indole, hydrochloride	280121-24-6	ETOH (100%)	-	-	NC	0.5	1.10	5	1.00	50	1.20							Glaxo SmithKline, Submitted by M.J. Olson
6-[(2-Methyl-3-pyridinyl)oxy]-3-pyridinamine	181633-42-1	DMSO	+	+	45	0.5	1.00	5	1.40	50	3.20							Glaxo SmithKline, Submitted by M.J. Olson
6-Chloro-1-hexanol	2009-83-8	AOO	-	-	NC	5	2.35	15	1.66	50	1.92							Glaxo SmithKline, Submitted by M.J. Olson
6-Diethylaminohexyl bromide hydrobromide	64993-14-2	PG	+	+	5.3	3	1.76	10	5.46	30	14.69							Glaxo SmithKline, Submitted by M.J. Olson
6-Iodo-quinazolin-4-ol	16064-08-7	DMF	-	-	NC	1	0.72	2.5	1.16	5	0.93							Glaxo SmithKline, Submitted by M.J. Olson
6-Methoxy-4-methyl-2(1H)-quinolinone	5342-23-4	PG	-	-	NC	3	1.21	10	1.49	30	1.32							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
7-[(4z)-3-(Aminomethyl)-4-(methoxyimino)-1-pyrrolidinyl]-1-cyclopropyl-6-fluoro-1,4-dihydro-4-oxo-1,8-naphthyridine-3-carboxylic acid, monomethane-sulfonate	210353-53-0	DMSO	+	+	8.6	0.1	0.75	1	1.38	10	3.30							Glaxo SmithKline, Submitted by M.J. Olson
8-[(4-Phthalimido-1-methylbutyl) amino]-2,6-dimethoxy-4-methyl-5-(3-trifluoromethylphenoxy)quinoline	106635-87-4	PG	-	-	NC	3	2.22	10	1.50	30	1.49							Glaxo SmithKline, Submitted by M.J. Olson
8-Amino-6-methoxy-4-methylquinoline	57514-21-3	PG	-	-	NC	3	1.23	10	2.83	30	2.50							Glaxo SmithKline, Submitted by M.J. Olson
8-Chloro-3-pentyl-3,7-dihydro-1h-purine-2,6-dione	862892-90-8	DMF	+	+	32	5	2.00	15	0.85	50	5.29							Glaxo SmithKline, Submitted by M.J. Olson
8-Hydroxy-5-[(1r)-1-hydroxy-2-[[2-[4-[(6-methoxy[1,1'-biphenyl]-3-yl)amino]phenyl]-ethyl]amino]ethyl]-2(1h)-quinolinone	530084-87-8	DMF	-	-	NC	5	0.96	10	2.34	25	1.58							Glaxo SmithKline, Submitted by M.J. Olson
Adipic acid	124-04-9	DMSO	-	-	NC	10	1.01	25	0.93	50	0.79							Glaxo SmithKline, Submitted by M.J. Olson
Alpha-(p-toluenesulfonyl)-4-fluorobenzyliso-nitrile	165806-95-1	DMF	+	+	45	0.5	4.72	5	2.78	50	3.03							Glaxo SmithKline, Submitted by M.J. Olson
Anthranilic acid	118-92-3	AOO	-	-	NC	10	0.90	25	1.10	50	1.40							Glaxo SmithKline, Submitted by M.J. Olson
cis-4-Cyano-4-[3-(cyclopentyloxy)-4-methoxyphenyl] cyclo-hexanecarboxylic acid	153259-65-5	DMSO	-	-	NC	0.5	1.16	5	1.27	10	1.28							Glaxo SmithKline, Submitted by M.J. Olson
Cytosine hemihydrate	71-30-7	PG	-	-	NC	5	0.40	10	0.90	25	0.70							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Dimethyl 4-cyano-4-(3-cyclopentyloxy-4-methoxyphenyl)-pimelate	152630-48-3	AOO	-	-	NC	0.5	1.66	5	1.59	50	1.76							Glaxo SmithKline, Submitted by M.J. Olson
Dimethyl carbonate	616-38-6	AOO	-	-	NC	0.5	0.64	5	0.69	50	1.71							Glaxo SmithKline, Submitted by M.J. Olson
Endo-tropine-3-mesylate	35130-97-3	DMF	+	+	4.4	5	3.45	10	5.98	25	25.06							Glaxo SmithKline, Submitted by M.J. Olson
Ethyl (3-endo)-8-methyl-8-azabicyclo[3.2.1]-octane-3-acetate	56880-11-6	DMF	+	+	5.5	10	6.77	25	12.58	50	ND[6]							Glaxo SmithKline, Submitted by M.J. Olson
Ethyl (z)-alpha-[[2-(1,1-dimethylethoxy)-1,1-dimethyl-2-oxoethoxy]imino]-2-[(triphenylmethyl) amino]- 4-thiazoleacetate	68672-65-1	Buta-none	-	-	NC	2.5	0.80	5	1.32	10	0.92							Glaxo SmithKline, Submitted by M.J. Olson
Ethyl 1h-1,2,4-triazole-3-carboxylate	64922-04-9	DMF	-	-	NC	0.25	1.00	2.5	1.20	25	1.00							Glaxo SmithKline, Submitted by M.J. Olson
Ethyl 2,6-dichloro-5-fluoro-beta-oxo-3-pyridinepropanoate	96568-04-6	AOO	-	-	NC	0.25	0.59	2.5	1.33	25	1.97							Glaxo SmithKline, Submitted by M.J. Olson
Ethyl 4-iodobenzoate	51934-41-9	AOO	+	+	8.0	0.5	1.10	5	2.20	50	14.30							Glaxo SmithKline, Submitted by M.J. Olson
Isopropyl dicyandiamide	35695-36-4	DMF	-	-	NC	0.25	1.36	2.5	1.35	25	1.24							Glaxo SmithKline, Submitted by M.J. Olson
m-Chloropropio-phenone	34841-35-5	AOO	-	-	NC	10	0.86	25	0.73	50	1.25							Glaxo SmithKline, Submitted by M.J. Olson
Methyl 4-(bromomethyl) benzoate	2417-72-3	AOO	+	+	IDR[3]	0.5	26.83	5	18.47	50	ND[6]							Glaxo SmithKline, Submitted by M.J. Olson
n-(2-Chloro-4-pyrimidinyl)-2,3-drimethyl-2h-indazol-6-amine	444731-74-2	DMF	-	-	NC	1	0.80	2.5	0.74	5	1.07							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
n-(2-Chloro-4-pyrimidinyl)-n,2,3-trimethyl-2h-indazol-6-amine	444731-75-3	DMF	-	-	NC	1	1.28	2.5	1.47	5	2.15							Glaxo SmithKline, Submitted by M.J. Olson
n-(3,4-Dichlorophenyl)-n'-(1-methylethyl)-imidodicarbon-imidic diamide monohydrochloride	15537-76-5	DMF	+	+	1.3	0.25	1.64	2.5	4.61	25	ND[6]							Glaxo SmithKline, Submitted by M.J. Olson
n-(4-Methoxyphenyl)-3-oxobutanamide	5437-98-9	PG	+	+	2.2	3	3.10	10	3.49	30	10.33							Glaxo SmithKline, Submitted by M.J. Olson
n-[(1,1-Dimethylethoxy)-carbonyl]-l-tyrosine, ethyl ester	72594-77-5	AOO	-	-	NC	10	1.18	25	1.09	50	0.64							Glaxo SmithKline, Submitted by M.J. Olson
n-[(1-Butyl-4-piperidinyl)methyl]-3,4-dihydro-2h-[1,3]oxazino[3,2-a]indole-10-carboxamide	152811-62-6	AOO	-	-	NC	0.5	0.80	5	1.20	25	1.30							Glaxo SmithKline, Submitted by M.J. Olson
n-[2-(Diethylamino)ethyl]-2-[[(4-fluorophenyl)-methyl]thio]-4,5,6,7-tetrahydro-4-oxo-n-[[4'-(trifluoromethyl)-[1,1'-biphenyl]-4-yl]methyl]-1h-cyclopentapyrim-idine-1-acetamide	356057-34-6	EtOH/ dH_2O (4:1)	+	+	11	5	1.11	10	2.43	25	12.71							Glaxo SmithKline, Submitted by M.J. Olson
n-[2-Benzyloxy-5-(2-bromo-1-hydroxy-ethyl)-phenyl]-formamide	201677-59-0	DMF	-	-	NC	10	0.98	25	0.68	50	0.97							Glaxo SmithKline, Submitted by M.J. Olson
n-{[(1,1-Dimethylethyl)oxy]carbonyl}-4-fluoro-beta-(4-fluorophenyl)-l-phenylalanine	481055-29-2	DMF	-	-	NC	10	1.47	25	2.41	50	2.28							Glaxo SmithKline, Submitted by M.J. Olson
1-Aminopyridazin-ium iodide	35073-04-2	DMF	-	-	NC	10	1.17	25	1.43	50	1.25							Glaxo SmithKline, Submitted by M.J. Olson

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
n-Isopropyl-n-phenyl-2-(2-phenylamino-phenylamino)-acetamide	161455-90-9	DMF	+	+	2.1	2.5	3.80	5	7.30	10	6.60							Glaxo SmithKline, Submitted by M.J. Olson
Oripavine	467-04-9	DMF	+	+	8.6	2.5	1.90	5	2.50	10	3.20							Glaxo SmithKline, Submitted by M.J. Olson
Phenylmethyl 2-(4-fluoro-2-methylphenyl)-4-oxo-3,4-dihydro-1(2h)-pyridine-carboxylate	414909-98-1	DMF	-	-	NC	10	1.67	25	1.85	50	1.47							Glaxo SmithKline, Submitted by M.J. Olson
rel-(3r,3as,6ar)-Hexahydrofuro[2,3-b]furan-3-yl 4-nitrophenyl carbonate	252873-35-1	DMF	+	+	IDR[3]	2.5	3.48	7.5	3.40	25	3.54							Glaxo SmithKline, Submitted by M.J. Olson
Sodium ethyl xanthate	140-90-9	PG	+	+	7.3	5	1.57	10	4.70	25	9.42							Glaxo SmithKline, Submitted by M.J. Olson
tert-Butyl-3-aminobenzoate	92146-82-2	DMF	-	-	NC	10	1.24	25	1.02	50	1.08							Glaxo SmithKline, Submitted by M.J. Olson
Trienol	13504-15-9	DMF	-	-	NC	5	1.40	10	1.10	25	1.00							Glaxo SmithKline, Submitted by M.J. Olson
Veratraldehyde	120-14-9	AOO	+	+	3.2	0.5	2.63	5	3.24	50	3.47							Glaxo SmithKline, Submitted by M.J. Olson
Basil oil	8015-73-4	EtOH/ DEP (1:3)	+	+	IDR[3]	2.5	3.00	5.0	3.00	10.0	8.00	25.0	17.60	50.0	25.20			Lalko & Api (2006), Submitted by A. Api (RIFM)
Citral	5392-40-5	EtOH/ DEP (1:3)	+	+	6.3	2.5	2.80	5.0	2.30	10.0	5.10	25.0	11.40	50.0	22.10			Lalko & Api (2006), Submitted by A. Api (RIFM)
Citronella oil	8000-29-1	EtOH/ DEP (1:3)	-	-	NC	2.5	1.40	5.0	0.90	10.0	1.20	25.0	1.20	50.0	2.70			Lalko & Api (2006), Submitted by A. Api (RIFM)
Clove bud oil	8000-34-8	EtOH/ DEP (1:3)	+	+	7.1	1.0	1.10	2.5	1.80	5.0	2.50	10.0	3.70	25.0	5.90			Lalko & Api (2006), Submitted by A. Api (RIFM)
Clove leaf oil	8015-97-2	EtOH/ DEP (1:3)	+	+	8.0	2.5	1.60	5.0	1.50	10.0	4.00	25.0	9.50	50.0	11.40			Lalko & Api (2006), Submitted by A. Api (RIFM)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Clove stem oil	8015-98-3	EtOH/ DEP (1:3)	+	+	7.0	1.0	1.60	2.5	1.70	5.0	2.20	10.0	4.20	25.0	8.90			Lalko & Api (2006), Submitted by A. Api (RIFM)
Eugenol	97-53-0	EtOH/ DEP (1:3)	+	+	5.5	2.5	1.20	5.0	2.70	10.0	6.00	25.0	14.30	50.0	19.40			Lalko & Api (2006), Submitted by A. Api (RIFM)
Geraniol	106-24-1	EtOH/ DEP (1:3)	+	+	12	2.5	1.70	5.0	2.40	10.0	2.80	25.0	4.80	50.0	6.00			Lalko & Api (2006), Submitted by A. Api (RIFM)
Geranium oil	8000-46-2	EtOH/ DEP (1:3)	-	-	NC	2.5	1.20	5.0	0.70	10.0	1.70	25.0	1.80	50.0	2.80			Lalko & Api (2006), Submitted by A. Api (RIFM)
Jasmine absolute (Grandiflorum)	8022-96-6	EtOH/ DEP (1:3)	+	+	5.9	1.0	1.20	2.5	1.80	5.0	2.00	10.0	7.40	25.0	11.80			Lalko & Api (2006), Submitted by A. Api (RIFM)
Jasmine absolute (Sambac)	8022-96-6	EtOH/ DEP (1:3)	+	+	36	10	1.70	25.0	2.50	50.0	3.60	75.0	10.80	100.0	16.20			Lalko & Api (2006), Submitted by A. Api (RIFM)
Lemongrass oil	8007-02-1	EtOH/ DEP (1:3)	+	+	6.5	2.5	0.90	5.0	2.10	10.0	5.10	25.0	10.30	50.0	13.10			Lalko & Api (2006), Submitted by A. Api (RIFM)
Litsea cubeba oil	68855-99-2	EtOH/ DEP (1:3)	+	+	8.5	2.5	2.00	5.0	2.30	10.0	3.30	25.0	7.90	50.0	16.00			Lalko & Api (2006), Submitted by A. Api (RIFM)
Palmarosa oil	8014-19-5	EtOH/ DEP (1:3)	+	+	9.5	2.5	1.10	5.0	2.10	10.0	3.10	25.0	3.60	50.0	5.00			Lalko & Api (2006), Submitted by A. Api (RIFM)
Spearmint oil	68917-46-4	EtOH/ DEP (1:3)	+	+	8.2	0.5	1.20	1.0	1.10	2.5	1.20	5.0	1.90	10.0	3.60			Lalko & Api (2006), Submitted by A. Api (RIFM)
Ylang Ylang (Extra)	8006-81-3	EtOH/ DEP (1:3)	+	+	6.8	0.5	1.50	1.0	1.40	2.5	2.10	5.0	2.50	10.0	3.90			Lalko & Api (2006), Submitted by A. Api (RIFM)
Ylang Ylang (III)	8006-81-3	EtOH/ DEP (1:3)	-	-	NC	0.5	1.30	1.0	1.70	2.5	2.10	5.0	2.60	10.0	2.60			Lalko & Api (2006), Submitted by A. Api (RIFM)
(1R,4R)-4-Isopropenyl-1-methyl-2-methylene-cyclohexane		AOO	-	-	NC	1	1.30	5	1.80	10	1.20	15	2.30	25	2.90			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
(3S,6R)-3-isopropyl-6-methylcyclohexene	5113-93-9	AOO	-	-	NC	1	0.84	10	1.00	25	2.90							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
(4Z)-2-Methyl-6-methyleneoct-4-ene			-	-	NC	1	1.10	5	0.87	10	0.78	15	0.89	25	2.10			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
(5R)-5-Isopropenyl-2-methyl-1-methylene-2-cyclohexene		AOO	+	+	7.3	0.5	0.94	5	1.90	15	6.60							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
2,2-bis-[4-(2-hydroxy-3 methacryloxypropoxy)phenyl)]-propane	1565-94-2	AOO	+	+	45	35	2.00	75	5.90									LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
2,4-Diamino-phenoxyethanol HCl	66422-95-5	AOO	+	+	5.5	1	1.60	2,5	1.60	5	2.70	10	5.70	25	8.30			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
2,4-Hexadienal	142-83-6	AOO	+	+	3.5	0.5	0.90	1	1.50	2.5	2.20	5	4.20	10	14.80			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
3,4-epoxyclohexylethyl-cyclopoly-methylsiloxane		AOO	-	-	NC	50	1.20	100	1.20									LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
4,4-Dibromobenzil	35578-47-3	AOO	+	+	21	5	1.50	10	1.60	25	3.60	50	5.70					LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
4-Isopropyl-1-methylene-cyclohexane		AOO	-	-	NC	1	1.20	10	0.71	25	1.40							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
5-Methyl-2-phenyl-2-hexenal	21834-92-4	AOO	+	+	4.4	0.5	1.00	1.0	1.30	2.5	0.50	5	3.80	10	17.70			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
a-Phellandrene	99-83-2	AOO	+	+	5.4	1.0	1.10	10	5.00	25	28.00							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
a-Terpinene	99-86-5	AOO	+	+	8.9	1	1.1	5	1.5	10	3.40	15	8.90	25	23.00			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Bandrowski's base	20048-27-5	AOO	+	+	0.02	0.01	1.10	0.025	3.10	0.05	5.70	0.1	6.50	0.25	5.60			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
beta-Phenyl-cinnamaldehyde	1210-39-5	AOO	+	+	0.6	0.1	2.00	0.25	2.30	0.5	1.90	1	5.90	2.5	10.60			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Bis-3,4-epoxycyclohexyl-ethyl-phenyl-methylsilane (Ph-Sil)		AOO	+	+	16	25	3.70	35	4.20	50	7.90							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
b-Phellandrene	555-10-2	AOO	+	+	NC	1.0	1.10	10	4.80	20	23.00							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
b-Terpinene	99-84-3	AOO	-	-	NC	1	1.4	10	1.30	25	2.1							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Butyl acrylate	141-32-2	AOO	+	+	11	1	0.70	2.5	1.30	5	1.50	10	2.50	25	8.70			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Chlorothalonil	1897-45-6	DMF	+	+	0.004	0.003	2.10	0.01	9.40	0.03	13.80	0.1	18.40	0.3	27.20			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Ethyl hexyl acrylate	103-11-7	AOO	+	+	9.7	0.5	1.10	1	1.20	2.5	0.90	5	1.20	10	3.10			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Iodopropynyl butylcarbamate	87977-28-4	AOO	+	+	0.9	0.1	0.70	1	3.40	5	4.20	10	12.00					LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Linalool alcohol	78-70-6	AOO	-	-	NC	1	1.00	10	1.30	30	1.30							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Linalool aldehyde		AOO	+	+	9.5	1	1.20	5	2.00	15	4.20							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Methyl acrylate	96-33-3	AOO	+	+	20	1	0.80	2.5	0.80	5	1.30	10	1.60	25	3.80			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Methyl pyruvate	600-22-6	AOO	+	+	2.4	1.0	1.20	2.5	3.10	5.0	4.70	10	8.00					LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Methyl methacrylate	80-62-6	AOO	+	+	90	10	1.40	30	1.50	50	1.50	75	2.10	100	3.60			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Norbornene fluoroalcohol	305815-63-8	AOO	+	+	46	5.0	0.70	10	0.80	25.0	1.90	50	3.20	100	3.70			LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
R-Carvone	2244-16-8	AOO	+	+	13	6	1.30	12	2.60	20	6.20							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
R-Carvoxime	2051-55-0	AOO	+	+	0.6	0.1	2.10	1	3.70	5	8.10							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Trans-2-methyl-2-butenal	497-03-0	AOO	-	-	NC	10	1.50	25	1.00	50	1.80							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
1-Chloro-2-dinitrobenzene	97-00-7	AOO	+	+	0.1	0.01	1.17	0.025	1.12	0.05	1.93	0.10	1.95	0.25	7.10			NTP Study, Submitted by D. Germolec
5-Amino-O-Cresol	2835-95-2	AOO	+	+	7.7	2.5	1.45	5.00	2.77	10.00	3.19							NTP Study, Submitted by D. Germolec
Atrazine	1912-24-9	ACE	-	-	NC	10	1.29	20.00	1.38	30.00	0.76							NTP Study, Submitted by D. Germolec
Azithromycin	83905-01-5	ACE	-	+	IDR[3]	10	3.72	20.00	1.54	40.00	2.10							NTP Study, Submitted by D. Germolec

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Butyl acrylate	141-32-2	ACE	+	+	24	10	1.00	20.00	2.18	30.00	4.07							NTP Study, Submitted by D. Germolec
Clarithromycin	81103-11-9	ACE	-	-	NC	1	1.78	2.00	1.03	4.00	1.18							NTP Study, Submitted by D. Germolec
Dicyclohexylcarbodiimide	538-75-0	ACE	+	+	0.057	0.006	1.03	0.03	1.71	0.06	3.16							NTP Study, Submitted by D. Germolec
Ethyl acrylate	140-88-5	ACE	-	-	NC	10	0.89	20.00	1.19	30.00	0.91							NTP Study, Submitted by D. Germolec
Ethyl-2-(Hydroxymethyl)-1,3- Propanediol Triacrylate		ACE	+	+	0.13	0.3	1.00	0.10	1.52	0.15	4.13	0.30	4.59					NTP Study, Submitted by D. Germolec
Methyl salicylate	119-36-8	AOO	-	-	NC	1	0.86	2.50	1.19	5.00	1.16	10.00	1.41	20.00	1.72			NTP Study, Submitted by D. Germolec
Pentaerythritol Triacrylate	3524-68-3	ACE	-	-	NC	0.005	1.19	0.01	0.92	0.05	1.68	0.10	2.43					NTP Study, Submitted by D. Germolec
Potassium dichromate	7778-50-9	DMSO	+	+	0.2	0.025	1.21	0.05	1.84	0.10	2.22	0.25	3.39					NTP Study, Submitted by D. Germolec
Rifamycin SV	14897-39-3	AOO	-	-	NC	3	0.94	10.00	1.02	30.00	1.33							NTP Study, Submitted by D. Germolec
Sodium metasilicate	6834-92-0	EtOH (15%)	-	-	NC	2	0.87	4.00	1.40	6.00	1.29							NTP Study, Submitted by D. Germolec
Trimethylolpropane Triacrylate	15625-89-5	ACE	-	-	NC	0.05	0.96	0.10	0.87	0.25	1.62							NTP Study, Submitted by D. Germolec
2.4-Dinitrobenzene sulfonic acid	89-02-1	H_2O	+	+	15	1	1.70	10	1.50	20	4.40							Ryan et al. (2002)
2.4-Dinitrobenzene sulfonic acid	89-02-1	Pluronic L92	+	+	6.4	1	0.90	10	4.40	20	11.60							Ryan et al. (2002)
2.4-Dinitrobenzene sulfonic acid	89-02-1	DMF	+	+	0.8	1	4.00	10	16.30	20	18.50							Ryan et al. (2002)
2.4-Dinitrobenzene sulfonic acid	89-02-1	DMSO	+	+	2.0	1	1.70	10	13.70	20	16.10							Ryan et al. (2002)
Formaldehyde	50-00-0	H_2O	+	+	15	1	1.20	10	2.50	20	3.60							Ryan et al. (2002)
Formaldehyde	50-00-0	Pluronic L92	+	+	4.2	1	2.00	10	4.80	20	8.80							Ryan et al. (2002)
Formaldehyde	50-00-0	DMF	+	+	0.3	1	6.70	10	13.20	20	17.70							Ryan et al. (2002)
Formaldehyde	50-00-0	DMSO	+	+	0.3	1	7.50	10	16.00	20	17.60							Ryan et al. (2002)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	6 Conc. (%)	6 SI	Data Source
Nickel sulfate	7786-81-4	Pluronic L92 (1%)	-	+	2.5	0.25	2.00	0.5	2.40	1	2.80	2.5	3.00	5	2.30			Ryan et al. (2002)
Nickel sulfate	7786-81-4	DMF	-	-	NC	0.25	0.90	0.5	1.10	1	1.60	2.5	1.60	5	2.20			Ryan et al. (2002)
Nickel sulfate	7786-81-4	DMSO	+	+	4.8	0.25	1.30	0.5	1.40	1	1.40	2.5	1.80	5	3.10			Ryan et al. (2002)
Pluronic L92®		H_2O	-	-	NC	1	1.30	2.5	1.00	5	1.00	10	0.80	25	0.80	50	2	Ryan et al. (2002)
Potassium dichromate	7778-50-9	Pluronic L92 (1%)	+	+	0.2	0.025	1.10	0.05	1.10	0.1	1.40	0.25	4.90	0.5	5.40			Ryan et al. (2002)
Potassium dichromate	7778-50-9	DMF	+	+	0.03	0.025	2.90	0.05	4.30	0.1	9.10	0.25	15.10	0.5	22.60			Ryan et al. (2002)
Potassium dichromate	7778-50-9	DMSO	+	+	0.05	0.025	1.40	0.05	2.50	0.1	9.50	0.25	25.90	0.5	10.10			Ryan et al. (2002)

Abbreviations: ACE = acetone; AOO = acetone: olive oil (4:1 by volume); BGIA = Berufsgenossenschaftliches Institut fur Arbeitsschutz (German Institute for Occupational Safety and Health); CASRN = Chemical Abstract Service Registry Number; CESIO = Comite Europeen des Agents de Surface et de Leurs Intermediaires Organiques (European Committee of Surfactants and their Organic Intermediates; Conc. = concentration; DBP = dibutyl phosphate; DEP = diethyl phthalate; dH2O = distilled water; DMF = dimethylformamide; DMSO = dimethyl sulfoxide; ECPA = European Crop Protection Association; EFfCI = European Federation for Cosmetic Ingredients; ECPA = European Crop Protection Association; EtOH = ethanol; H_2O = water; LLNA = Local Lymph Node Assay; MEK = methyl ethyl ketone; NC = not calculated because no SI value was ≥ 3 (i.e., substance was a nonsensitizer); ND = no data; NTP = National Toxicology Program; PA/ H_2O = pluronic acid/ H_2O (1%); PG = propylene glycol; rLLNA = Reduced Local Lymph Node Assay; RIFM = Research Institute for Fragrance Materials; SI = stimulation index; Trad. = traditional

[1] "+" = Sensitizer; "-" = Non-sensitizer

[2] EC3 represents the estimated concentration needed to produce a stimulation index of three (i.e., a three fold increase in lymphocyte proliferation is observed for the test substance versus the vehicle control substance) and was calculated using the methods described in Ryan et al. (2007).

[3] IDR indicates an insufficient dose response to calculate an EC3 value using the methods in Ryan et al. (2007)

[4] Result of initial study

[5] Result of second study

[6] Data not obtained due to toxicity

Annex IV

Substances in the NICEATM LLNA Database for which an Initial Dose of 10% or Greater Elicited a Negative Result but a Subsequent Higher Dose Elicited a Positive Response

[This Page Intentionally Left Blank]

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	Data Source
1-(2',5' Dimethylphenyl)butane-1,3-dione	56290-55-2	ACE	+	+	12.5	10	2.30	20	5.10	40	9.50					Gerberick et al. (2005)
1,1-Dimethylethyl 3-[[[[(3s)-2,3,4,5-tetrahydro-1-[2-[(1-methylethyl)phenylamino]-2-oxoethyl]-2,4-dioxo-5-phenyl-1h-1,5-benzodiazepin-3-yl]amino]carbonyl]amino]-benzoate	305366-94-3	DMF	+	+	37.5	10	2.10	25	2.20	50	3.80					Glaxo SmithKline, Submitted by M.J. Olson
1-Chlorooctadecane	3386-33-2	AOO	+	+	16.3	10	1.70	25	4.80	50	7.30					Gerberick et al. (2005)
1-Chlorotetradecane	2425-54-9	AOO	+	+	20.2	10	1.10	25	3.90	50	6.30					Gerberick et al. (2005)
1-Iodohexadecane	544-77-4	AOO	+	+	19.1	10	1.60	25	3.90	50	6.40					Gerberick et al. (2005)
1-Iodononane	4282-42-2	AOO	+	+	24.2	10	1.30	25	3.10	50	4.60					Gerberick et al. (2005)
1-Iodotetradecane	19218-94-1	AOO	+	+	13.8	10	1.70	25	6.90	50	9.70					Gerberick et al. (2005)
1-Phenyl-2-methylbutane-1,3-dione	6668-24-2	ACE	+	+	29.1	10	1.70	20	2.00	40	4.20					Gerberick et al. (2005)
1-Phenyloctane--1,3-dione	55846-68-1	ACE	+	+	10.5	10	2.80	20	6.60	40	8.70					Gerberick et al. (2005)
2-(4-tert-Amylcyclohexyl)acetaldehyde	620159-84-4	AOO	+	+	36.8	25	2.10	50	4.00	100	9.10					Gerberick et al. (2005)
2,2,6,6-Tetramethyl-heptane-3,5-dione	1118-71-4	ACE	+	+	26.7	10	2.10	20	2.80	40	3.40					Gerberick et al. (2005)
2,2-bis-[4-(2-hydroxy-3 methacryloxypropoxy)phenyl)] -propane	1565-94-2	AOO	+	+	45.3	35	2.00	75	5.90							LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
3-Ethoxy-1-(2',3',4',5'-tetramethylphenyl)propane-1,3-dione	170928-69-5	ACE	+	+	33.0	10	1.10	20	1.70	40	3.70					Gerberick et al. (2005)
3-Methyleugenol	186743-26-0	AOO	+	+	31.6	11	1.50	27	2.30	54	6.40					Gerberick et al. (2005)
4,4,4-Trifluoro-1-phenylbutane-1,3-dione	362-06-7	ACE	+	+	20.0	10	2.10	20	3.00	40	4.60					Gerberick et al. (2005)
4-Allylanisole	140-67-0	AOO	+	+	17.7	10	1.20	25	4.70	50	4.50	100	8.00			Gerberick et al. (2005)
5-Methyl-2,3-hexanedione	13706-86-0	AOO	+	+	25.8	25	2.90	50	6.00	100	14.30					Gerberick et al. (2005)
5-Methyleugenol	186743-25-9	AOO	+	+	13.2	11	2.70	27	4.90	54	4.30					Gerberick et al. (2005)

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	Data Source
6-Methyleugenol	186743-24-8	AOO	+	+	16.9	11	1.90	27	4.90	54	8.30					Gerberick et al. (2005)
Atrazine SC	1912-24-9	Pluronic L92 (1%)	+	+	31.3	13	1.80	25	2.80	50	3.60	75	7.10	100	7.30	ECPA LLNA Project Report, Submitted by P. Botham
Butyl acrylate	141-32-2	ACE	+	+	24.4	10	1.00	20	2.18	30	4.07					NTP Study, Submitted by D. Germolec
Butyl glycidyl ether	2426-08-6	AOO	+	+	30.9	10	1.40	25	2.20	50	5.60					Gerberick et al. (2005)
C15-azlactone	176665-09-1	AOO	+	+	17.8	10	1.80	25	4.10	50	7.50					Gerberick et al. (2005)
C17-azlactone	176665-11-5	AOO	+	+	19.0	11	1.70	27	4.30	54	4.60					Gerberick et al. (2005)
C19-azlactone		AOO	-	+	26.4	12	2.50	29	3.10	59	2.50					Gerberick et al. (2005)
Cinnamic alcohol	104-54-1	AOO	+	+	20.6	10	1.80	25	3.50	50	3.90	90	5.70			Gerberick et al. (2005)
cis-6-Nonenal	2277-19-2	AOO	+	+	22.4	10	1.60	25	3.30	50	4.50	100	13.70			Gerberick et al. (2005)
Diethylacetaldehyde	97-96-1	AOO	+	+	76.1	25	1.20	50	0.80	75	2.40	100	16.30			Gerberick et al. (2005)
Dimethyl sulfoxide	67-68-5	AOO	+	+	71.9	25	2.70	50	2.30	100	3.90					Gerberick et al. (2005)
Ethyl acrylate	140-88-5	AOO	+	+	28.3	10	1.20	25	2.70	50	5.00					Gerberick et al. (2005)
Ethylene glycol dimethacrylate	97-90-5	MEK	+	+	28.3	10	1.20	25	2.40	50	7.00					Gerberick et al. (2005)
EXP 11120 A		Pluronic acid/H_2O (1%)	+	+	64.9	10	0.96	25	0.66	50	1.60	100	6.30			Bayer CropScience SA Studies, Submitted by E. Debruyne
FAR01060-00		Pluronic acid/H_2O (1%)	+	+	88.5	10	0.40	25	0.80	50	1.00	100	3.60			Bayer CropScience SA Studies, Submitted by E. Debruyne
Imidazolidinyl urea	39236-46-9	DMF	+	+	23.9	10	1.70	25	3.10	50	5.50					Gerberick et al. (2005)
Isopropyl myristate	110-27-0	AOO	+	+	43.8	25	2.10	50	3.30	100	3.40					Gerberick et al. (2005)
Jasmine absolute (Sambac)	8022-96-6	EtOH/DEP (1:3)	+	+	36.4	10	1.70	25	2.50	50	3.60	75	10.80	100	16.20	Lalko & Api (2006), Submitted by A. Api (RIFM)
Linalool alcohol	78-70-6	AOO	+	+	30.4	25	2.50	50	4.80	100	8.30					Gerberick et al. (2005)
Linoleic acid	60-33-3	AOO	+	+	14.1	10	1.50	25	7.00	50	9.10					EFfCI study, Submitted by P. Ungeheuer

Substance Name	CASRN	Vehicle	rLLNA[1]	Trad. LLNA[1]	EC3[2]	1 Conc. (%)	1 SI	2 Conc. (%)	2 SI	3 Conc. (%)	3 SI	4 Conc. (%)	4 SI	5 Conc. (%)	5 SI	Data Source
Methyl methacrylate	80-62-6	AOO	+	+	90.0	10	1.40	30	1.50	50	1.50	75	2.10	100	3.60	LLNA/EC3 Validation Study, Submitted by D. Basketter, I. Kimber, and F. Gerberick
Non-ionic surfactant 1		AOO	+	+	27.5	25	2.80	50	4.80	100	6.50					CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 2		AOO	-	+	47.1	25	1.50	50	3.20	100	2.90					CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 5		AOO	+	+	37.5	25	2.70	50	3.30	100	3.20					CESIO Report, Submitted by K. Skirda
Non-ionic surfactant 6		AOO	+	+	34.4	25	2.70	50	3.50	100	6.50					CESIO Report, Submitted by K. Skirda
Oleic acid	112-80-1	AOO	+	+	10.5	10	2.60	25	14.90	50	6.90					EFfCI study, Submitted by P. Ungeheuer
Pentachlorophenol	87-86-5	DMSO	+	+	19.6	10	2.10	25	3.50	50	5.40					Gerberick et al. (2005)
Precursor surfactant 1		AOO	+	+	60.7	25	2.20	50	2.70	100	4.10					CESIO Report, Submitted by K. Skirda
Pyridine	110-86-1	AOO	+	+	71.9	25	1.10	50	2.30	100	3.90					Gerberick et al. (2005)
Quinoxyfen/cyproconazole	124495-18-7/ 113096-99-4	Pluronic L92 (1%)	+	+	27.8	13	2.00	25	2.30	50	8.60	75	15.80	100	30.10	ECPA LLNA Project Report, Submitted by P. Botham
R(+)-Limonene	5989-27-5	AOO	+	+	68.8	25	1.80	50	2.40	100	4.00					Gerberick et al. (2005)
Undecylenic acid	112-38-9	AOO	+	+	19.4	10	2.50	25	3.30	50	4.40					EFfCI study, Submitted by P. Ungeheuer
Unsaturated fatty acid ester		AOO	+	+	27.1	25	2.80	50	5.20	100	4.70					CESIO Report, Submitted by K. Skirda

Abbreviations: ACE = acetone; AOO = Acetone: olive oil (4:1 by volume); CASRN = Chemical Abstract Services Registry Number; CESIO = Comite Europeen des Agents de Surface et de Leurs Intermediaires Organiques (European Committee of Surfactants and their Organic Intermediates; Conc. = concentration; DEP = diethyl phthalate; DMF = dimethylformamide; DMSO = dimethyl sulfoxide; EFfCI = European Federation for Cosmetic Ingredients; ECPA = European Crop Protection Association; EtOH = ethanol; H2O = water; LLNA = Local Lymph Node Assay; MEK = methyl ethyl ketone; NTP = National Toxicology Program; rLLNA = Reduced Local Lymph Node Assay; RIFM = Research Institute for Fragrance Materials; SI = stimulation Index; Trad. = traditional

[1] "+" = Sensitizer; "-" = Non-sensitizer

[2] EC3 represents the estimated concentration needed to produce a stimulation index of three (i.e., a three-fold increase in lymphocyte proliferation is observed for the test substance versus the vehicle control substance).

[This Page Intentionally Left Blank]

Appendix E

ECVAM Scientific Advisory Committee Statement on the Validity of the rLLNA

[This Page Intentionally Left Blank]

EUROPEAN COMMISSION
DIRECTORATE GENERAL JRC
JOINT RESEARCH CENTRE
Institute for Health and Consumer Protection
European Centre for the Validation of Alternative Methods (ECVAM)

ESAC Statement on the Reduced Local Lymph Node Assay (rLLNA)

At its 26th Meeting, held on 26-27 April 2007 at the European Centre for the Validation of Alternative Methods (ECVAM), Ispra, Italy, the non-Commission members of the ECVAM Scientific Advisory Committee (ESAC)[1] unanimously endorsed the following statement:

Skin sensitisation is an important toxicological endpoint with respect to human safety.

Having reviewed the final report of the independent peer review evaluation co-ordinated by ICCVAM and NICEATM[2], the report by the EMEA[3], the pre-report of the SCCNFP[4], and evidence made available since the original submissions to ICCVAM, in March 2000 the 14th meeting of ESAC stated:

> "Following a review of the scientific report and publications on the local lymph node assay (LLNA) it is concluded that the LLNA is a scientifically validated test which can be used to assess the skin sensitisation potential of chemicals. The LLNA should be the preferred method, as it uses fewer animals and causes less pain and distress than the conventional guinea-pig methods. In some instances and for scientific reasons, the conventional methods can be used."

Since its acceptance for regulatory purposes, the LLNA has proved suitable for the purposes of satisfying a range of EU and other regulatory requirements[5].

The developers of the LLNA have now undertaken a retrospective analysis of published data obtained with the LLNA[6].

They conclude that within a tiered testing strategy in the context of REACH a "reduced" version of the LLNA (rLLNA), using only a negative control group and the equivalent of the high-dose group from the full LLNA, can be used as a screening test to distinguish between sensitisers and non-sensitisers.

ESAC established a peer review panel to evaluate if there was the potential to minimise animal use by employing the rLLNA as a screening test as part of a tiered-testing strategy for chemicals.

Mindful that with the rLLNA:

- When compared with the full LLNA the rLLNA cannot and will not result in additional false positives.
- When compared with the full LLNA the rLLNA may produce a few false negatives (3:169 in the reference document, reducing to 2:169 when negative results obtained with concentrations of <10% are considered invalid)
- The test results provided by the rLLNA do not allow the determination of the potency of a sensitising chemical.

ESAC states that the peer reviewed and published information is of a quality and nature to support the use of the rLLNA within tiered-testing strategies to reliably distinguish between chemicals that are skin sensitisers and non-sensitisers, and that animal use can be minimised providing:

EUROPEAN COMMISSION
DIRECTORATE GENERAL JRC
JOINT RESEARCH CENTRE
Institute for Health and Consumer Protection
European Centre for the Validation of Alternative Methods (ECVAM)

- The concentration used to evaluate sensitisation potential is the maximum consistent with solubility and the need to avoid local and other systemic adverse effects, and that this principle rather than strict adherence to the specific recommended absolute concentrations as in OECD TG 429 should be used.
- Negative test results associated with testing using concentrations of less than 10%, should undergo further evaluation.
- Positive and negative (vehicle) control groups are used, as appropriate, per OECD TG 429.
- The full LLNA should be performed when it is known that an assessment of sensitisation potency is required.

ESAC recommends that further work should be undertaken to determine if the 10% concentration threshold referenced above is optimal.

Thomas Hartung
Head of Unit
ECVAM
Institute for Health & Consumer Protection
Joint Research Centre
European Commission
Ispra

27 April 2007

1. The ESAC was established by the European Commission, and is composed of nominees from the EU Members States, industry, academia and animal welfare, together with representatives of the relevant Commission services.

   This statement was endorsed by the following members of the ESAC:

   Ms Sonja Beken (Belgium)
   Ms Dagmar Jírová (Czech Republic)
   Mr Tõnu Püssa (Estonia)
   Mr Lionel Larue (France)
   Mr Manfred Liebsch (Germany)
   Ms Annalaura Stammati (Italy)
   Mr Jan van der Valk (The Netherlands)
   Mr Constantin Mircioiu (Romania)
   Mr Albert Breier (Slovakia)
   Ms Argelia Castaño (Spain)
   Mr Patric Amcoff (Sweden)
   Mr Jon Richmond (UK)
   Mr Carl Westmoreland (COLIPA)
   Ms Vera Rogiers (ECOPA)
   Ms Nathalie Alépée (EFPIA)
   Mr Robert Combes (ESTIV)
   Mr Hasso Seibert (European Science Foundation)

   The following Commission Services and Observer Organisations were involved in the consultation process, but not in the endorsement process itself.

   Mr Thomas Hartung (ECVAM; chairman)
   Mr Jens Linge (ECVAM; ESAC secretary)
   Ms Elke Anklam (Director of IHCP)
   Ms Susanna Louhimies (DG Environment)
   Ms Barbara Mentré (DG ENTR)
   Ms Grace Patlewicz (ECB, DG JRC)
   Mr Christian Wimmer (DG Research)
   Mr Hajime Kojima (JACVAM)
   Ms Laurence Musset (OECD)
   Mr Barry Philips (Eurogroup for Animal Welfare)
   Mr William Stokes (NICEATM, USA)

2. NIH (1999). The murine local lymph node assay. The results of an independent peer review evaluation coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication n.99-4494.
   (http://iccvam.niehs.nih.gov/methods/immunotox/immunotox.htm)

3. EMEA (2000). Report from the ad-hoc expert meeting on testing for immunohypersensitivity (11/01/2000). European Agency for the Evaluation of Medicinal Products.

4. SCCNFP (2000). Opinion adopted by the SCCNFP during the 11 th plenary meeting, 17 February 2000. (http://ec.europa.eu/health/ph_risk/committees/sccp/docshtml/sccp_out114_en.htm)

5. A Cockshott, P Evans, CA Ryan, GF Gerberick, CJ Betts, RJ Dearman, I Kimber and DA Basketter (2006). The local lymph node assay in practice: a current regulatory perspective. Human & Experimental Toxicology **25**, 387-394.

6. I Kimber, RJ Dearman, CJ Betts, GF Gerberick, CA Ryan, PS Kern, GY Patlewicz and DA Basketter (2006.) The local lymph node assay and skin sensitisation: a cut-down screen to reduce animal requirements? Contact Dermatitis **54**, 181-185.

Appendix F

Independent Scientific Peer Review Panel Assessment

[This Page Intentionally Left Blank]

Appendix F1

Independent Scientific Peer Review Panel Report: Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products

[This Page Intentionally Left Blank]

Independent Scientific Peer Review Panel Report: Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products

May 2008

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Institute of Environmental Health Sciences (NIEHS)
National Institutes of Health
U.S. Public Health Service
Department of Health and Human Services

National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709

This document is available electronically at http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

The findings and conclusions of this report are those of the Independent Scientific Peer Review Panel and should not be construed to represent the official views of ICCVAM or its member agencies.

Table of Contents

List of Tables

List of Abbreviations and Acronyms

ACD	Allergic contact dermatitis
ANOVA	Analysis of variance
AOO	Acetone: olive oil (4:1)
BRD	Background review document
BrdU	Bromodeoxyuridine
BT	Buehler Test
CD4	Cluster of differentiation 4
CPSC	U.S. Consumer Product Safety Commission
CRO	Clinical research organization
CV	Coefficient of variation
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNCB	Dinitrochlorobenzene
EC3	Estimated concentration needed to produce a stimulation index of 3
ECt	Estimated concentration needed to produce a stimulation index that is indicative of a positive response
ECVAM	European Centre for the Validation of Alternative Methods
ELISA	Enzyme Linked Immunosorbent Assay
eLLNA: BrdU-FC	Enhanced LLNA with BrdU detected by flow cytometry
EPA	U.S. Environmental Protection Agency
FC	Flow cytometry
FR	*Federal Register*
GLP	Good Laboratory Practice
GPMT	Guinea Pig Maximization Test
GSK	GlaxoSmithKline
HCA	Hexyl cinnamic aldehyde
HMT	Human Maximization Test
HRIPT	Human Repeat Insult Patch Test
HTdR	^{3}H-Methyl Thymidine
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ISO	International Organization for Standardization
IWG	Immunotoxicity Working Group
JaCVAM	Japanese Center for Validation of Alternative Methods
LLNA	Murine local lymph node assay
LLNA: BrdU-ELISA	LLNA with BrdU detected by ELISA
LLNA: BrdU-FC	LLNA with BrdU detected by FC
LLNA: DA	LLNA: Daicel adenosine triphosphate
LNC	Lymph node cells
LOEL	Lowest observed effect level
MEK	Methyl ethyl ketone
NICEATM	NTP Interagency Center for the Evaluation of Alternative Toxicological Methods

NIEHS	National Institute of Environmental Health Sciences
NIH	National Institutes of Health
NOEL	No observed effect level
NTP	National Toxicology Program
OD	Optical density
OECD	Organisation for Economic Co-operation and Development
QSAR	Quantitative structure-activity relationship
REACH	Registration, Evaluation, and Authorisation of Chemicals
rLLNA	Reduced local lymph node assay
SAR	Structure-activity relationship
SD	Standard deviation
SI	Stimulation index
SDS	Sodium dodecyl sulfate
SLS	Sodium lauryl sulfate
TG	Test guideline
Th	T-helper
vs.	Versus

Members of the Independent Scientific Peer Review Panel

Michael Luster, Ph.D. (Panel Chair), Senior Consultant to the National Institute of Occupational Safety and Health (NIOSH) Health Effects Laboratory, Morgantown, WV

Nathalie Alépée, Ph.D., Associate Research Fellow, Pfizer PDRD MCT Laboratory, France

Anne Marie Api, Ph.D., Vice President, Human Health Sciences, Research Institute for Fragrance Materials, Woodcliff Lake, NJ

Nancy Flournoy, M.S., Ph.D. (Biostatistician), Professor and Chair, Dept. of Mathematics & Statistics, University of Missouri-Columbia, Columbia, MO

Thomas Gebel, Ph.D., Regulatory Toxicologist, Federal Institute for Occupational Safety & Health, Dortmund, Germany

Sidney Green, Ph.D.[1], Graduate Professor, Howard University, Washington, DC

Kim Headrick, B. Admin., B.Sc., International Harmonization Senior Policy Advisor, Health Canada, Ottawa, Ontario, Canada

Dagmar Jírová, M.D., Ph.D., Toxicologist, Research Manager, Head of Reference Center for Cosmetics, Head of Reference Laboratory for Experimental Immunotoxicology, National Institute of Public Health, Czech Republic

David Lovell, Ph.D. (Biostatistician), Reader in Medical Statistics, Postgraduate Medical School, University of Surrey, Guildford, Surrey, U.K.

Howard Maibach, M.D., Professor, Department of Dermatology, University of California-San Francisco, San Francisco, CA

James McDougal, Ph.D., Professor and Director of Toxicology Research, Department of Pharmacology and Toxicology, Boonshoft School of Medicine, Wright State University, Dayton, OH

Michael Olson, Ph.D., Director of Occupational Toxicology, Corporate Environment Health and Safety, GlaxoSmithKline, Research Triangle Park, NC

Raymond Pieters, Ph.D., Associate Professor, Immunotoxicology Group Leader, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands

Jean Regal, Ph.D., Professor, Department of Pharmacology, University of Minnesota Medical School, Duluth, MN

Jonathan Richmond, MB ChB, FRCSEd[1], Head of Division, Animals Scientific Procedures Division, Home Office, London, U.K.

[1] Drs. Green and Richmond were unable to attend the public meeting on March 4-6, 2008. However, they were involved in the review of the background review documents and concur with the conclusions and recommendations included in this report.

Peter Theran, V.M.D., Massachusetts Society for the Prevention of Cruelty to Animals, Novato, CA

Stephen Ullrich, Ph.D., Dallas/Ft. Worth Living Legends Professor & Professor of Immunology, Post-graduate School of Biomedical Science, University of Texas M.D. Anderson Cancer Center, Houston, TX

Michael Woolhiser, Ph.D., Technical Leader - Immunotoxicology, Toxicology & Environmental Research & Consulting Immunology, Dow Chemical, Midland, MI

Takahiko Yoshida, M.D., Ph.D., Professor, Department of Health Science, Asahikawa Medical College, Hokkaido, Japan

[This Page Intentionally Left Blank]

Preface

In 1999, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) recommended the murine local lymph node assay (LLNA) to U.S. Federal agencies as a valid substitute for currently accepted guinea pig test methods to assess the allergic contact dermatitis potential of many, but not all, types of substances. The recommendation was based on a comprehensive evaluation of the validation status of the LLNA that included an assessment by an international independent scientific peer review panel (Panel). The Panel report and the ICCVAM LLNA test method recommendations (ICCVAM 1999) are available at the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)-ICCVAM website.[2] The LLNA was subsequently incorporated into national and international test guidelines for the assessment of skin sensitization (OECD 2002; ISO 2002; EPA 2003). For this Panel report, this LLNA will be referred to as the "traditional" LLNA.

On January 10, 2007, the U.S. Consumer Product Safety Commission (CPSC) formally requested through NICEATM that ICCVAM assess the validation status of:[3]

- The traditional LLNA as a stand-alone assay for potency determinations (including severity) for the purpose of hazard classification
- Three modifications of the traditional LLNA not requiring the use of radioactive materials
- The LLNA limit dose procedure (also referred to as the "reduced" LLNA)
- The ability of the traditional LLNA to test mixtures, metals, and aqueous solutions (i.e., to re-evaluate the applicability domain for the traditional LLNA)

NICEATM, in coordination with ICCVAM and the ICCVAM Immunotoxicity Working Group, prepared a comprehensive draft background review document (BRD) for each modified version of the traditional LLNA test method being evaluated, as well as a draft applicability domain addendum to the final BRD published previously on the traditional LLNA. Each draft BRD and the draft addendum detailed the available data and information from the published literature and submissions received in response to a 2007 *Federal Register (FR)* notice that had requested data related to CPSC's nomination (*FR* notice Vol. 72, No. 95, p. 27815-27817, May 17, 2007). In addition, ICCVAM developed draft LLNA Performance Standards intended for use in validating alternative test methods that are functionally and mechanistically similar to the traditional LLNA. Finally, ICCVAM, based on the information contained in each of the draft BRDs and the draft addendum, developed draft test method recommendations.

The various supporting documents and the draft ICCVAM recommendations were provided to a new international Panel for an independent scientific review. In addition, NICEATM

[2] The 1999 ICCVAM Panel report and recommendations can be obtained at: http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf

[3] The CPSC nomination can be obtained at: http://iccvam.niehs.nih.gov/methods/immunotox/llnadocs/CPSC_LLNA_nom.pdf

announced the availability of these documents on the NICEATM-ICCVAM website (http://iccvam.niehs.gov) for public comment in a *FR* notice (Vol. 73, No. 5, p. 1360-1362, January 8, 2008) and via the ICCVAM listserv. The *FR* notice also announced the public Panel meeting, to be convened at the CPSC Headquarters in Bethesda, MD on March 4–6, 2008.

The Panel was charged with:

- Reviewing each ICCVAM draft BRD and the draft addendum for completeness and identifying any errors or omissions of existing relevant data or information
- Evaluating the information in each draft BRD and the draft addendum to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003) had been appropriately addressed for the recommended use of the new versions and applications of the traditional LLNA
- Considering the ICCVAM draft test method recommendations for the following and commenting on the extent to which they are supported by the information provided in the draft BRDs and the draft addendum:
  - proposed test method uses
  - proposed recommended standardized protocols
  - proposed test method performance standards
  - proposed additional studies
- Evaluating the draft ICCVAM LLNA Performance Standards and considering whether they were adequate for assessing the accuracy and reliability of alternative test methods that are functionally and mechanistically similar to the traditional LLNA

During our public meeting in March 2008, the Panel discussed each charge, listened to public comments, and developed conclusions and recommendations for ICCVAM on each of the nominated activities. The Panel wished to emphasize that they were to consider two overall questions. They were to consider: (1) whether the validation status of the each of the above proposed modifications or alternative uses of the LLNA had been adequately characterized for its intended purpose according to established ICCVAM validation criteria (available on the NICEATM-ICCVAM website, http://iccvam.niehs.gov), and (2) whether proposed modifications or alternative uses of the LLNA are sufficiently accurate and reliable to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure.

This report details the Panel's independent conclusions and recommendations. ICCVAM will consider this report, along with all relevant public comments, as it develops final test method recommendations. The final ICCVAM test method recommendations will be forwarded to U.S. Federal agencies for their consideration in accordance with the ICCVAM Authorization Act of 2000 (Public Law 106-545).

The Panel gratefully acknowledges the efforts of NICEATM staff in coordinating the logistics of the peer review Panel meeting and in preparing materials for their review. The Panel also thanks each of the test method developers, Drs. George DeGeorge (LLNA: BrdU-FC), Kenji Idehara (LLNA: DA), and Masahiro Takeyoshi, (LLNA: BrdU-ELISA) for providing summaries and additional clarifications of the non-radioactive test methods under review. Finally, as Panel Chair, I want to thank each Panel member for her or his thoughtful and objective review of these LLNA-related activities.

Michael Luster, Ph.D.
Chair, LLNA Peer Review Panel
May 2008

[This Page Intentionally Left Blank]

Executive Summary

This report describes the conclusions and recommendations of an international independent scientific peer review panel (Panel). This Panel was charged by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) with evaluating the validation status of new versions and applications of the murine local lymph node assay (LLNA) for assessing the allergic contact dermatitis (ACD) potential of chemicals and products. The LLNA, which was first evaluated in 1999 by ICCVAM, is hereafter referred to as the "traditional LLNA" to distinguish it from other versions considered by the Panel. The new versions and applications considered include:

- The LLNA limit dose procedure (also referred to as the "reduced" LLNA[4])
- The ability of the traditional LLNA to test mixtures, metals, and aqueous solutions (i.e., a re-evaluation of the applicability domain for the traditional LLNA)
- Three modifications of the traditional LLNA not requiring the use of radioactive materials:
  - LLNA: DA (Local Lymph Node Assay: Daicel Adenosine Triphosphate)
  - LLNA: BrdU-FC (Local Lymph Node Assay: Bromodeoxyuridine detected by flow cytometry)
  - LLNA: BrdU-ELISA (Local Lymph Node Assay: Bromodeoxyuridine detected by ELISA)
- The traditional LLNA as a stand-alone assay for potency determinations (including severity) for the purpose of hazard classification

The Panel also evaluated the draft ICCVAM LLNA Performance Standards and considered whether they were adequate for assessing the accuracy and reliability of alternative test methods that are functionally and mechanistically similar to the traditional LLNA.

LLNA Limit Dose Procedure

The Panel agreed that the LLNA limit dose procedure, which normally allows for testing at one dose level, should be routinely recommended for hazard identification when used for testing purposes which do not require dose response information, because it would offer time, cost, throughput and logistical benefits as well as using fewer animals. In instances when a necessity to measure relative skin-sensitization potency for the purpose of risk assessment was present, then the traditional LLNA should be used in order to generate dose response information. Still,

[4] As described in this report, the Panel agreed that consideration should be given to applying the same term to the LLNA limit dose procedure since in various places throughout the draft BRD it was referred to differently as either the "cut-down," the "limit dose," or the "reduced LLNA" (i.e., "rLLNA"). Since the European Centre for the Validation of Alternative Methods (ECVAM) has already established a naming convention of "rLLNA," the Panel recommended adopting the ECVAM terminology to harmonize the terminology used among the international validation agencies. However, because the ICCVAM documents that were reviewed use "LLNA limit dose procedure" that term is retained in this report.

the Panel recommended use of the LLNA limit dose procedure as the initial testing procedure to identify sensitizers and non-sensitizers before conducting the traditional LLNA even when dose response information *is* required since if the test substance were negative in the limit dose procedure, it would not be necessary to conduct a multiple-dose LLNA test.

The draft background review document (BRD) for the LLNA limit dose procedure provides a comprehensive review of available data and information for assessing the usefulness and limitations of this modified version of the LLNA for the purpose of skin-sensitization hazard classification. The Panel evaluated the draft BRD for completeness, errors, and omissions and recommended that it be updated to reflect their suggestions/corrections relating to general, statistical, and specific editorial issues. In particular, the Panel noted that the differences in terminology used for this procedure caused confusion and recommended that an internationally harmonized term be adopted. They suggested referring to the procedure as the "reduced LLNA" (i.e., "rLLNA") since that is being used by the European Centre for the Validation of Alternative Methods (ECVAM).

The Panel concluded that the stimulation index (SI) based on the ratio of 3.0 as the cutoff value was indicative of a response that was sufficiently greater than the control and would be considered an immunologically relevant response, but recommended that statistical analyses be used to definitively establish that a response induced by a test substance is significantly different from the vehicle control. The Panel agreed that the LLNA protocol recommended by ICCVAM (ICCVAM 1999; Dean et al. 2001) should be the standard protocol for all future LLNA limit dose studies using the traditional LLNA protocol. Specifically, prospective LLNA limit dose procedure studies should require that lymph nodes be collected from individual animals instead of pooling them with other animals in a treatment group, which is also currently permitted by the Organisation for Economic Co-operation and Development (OECD) Test Guideline 429 (OECD 2002). Individual animal response data are necessary in order to statistically analyze for differences between treated and control data. In addition, having data from individual animals also allows for identification of technical problems and outlier animals within a dose group. Based on power calculations provided as supplemental information, the Panel agreed that five animals per dose group is an appropriate number to recommend for LLNA limit dose studies following the traditional LLNA protocol. It should be noted that the Panel's analysis of the LLNA limit dose dataset was not restricted to studies with confirmed individual animal data, and that the Panel considered data known to have been generated using pooled group data. The Panel stated that, internationally, both individual and pooled animal data have likely been used both for regulatory decisions and for in-house decisions relating to product development and risk management. In addition, the fact that the retrospective data analysis set out in the draft LLNA limit dose procedure BRD did not distinguish between individual or pooled animal data suggested that both met the quality standards for inclusion in the draft BRD.

Although they did not reach consensus, the Panel suggested that for laboratories in which the LLNA is "routinely" performed and have demonstrated the ability to consistently obtain positive results, hexyl cinnamic aldehyde (HCA) or another positive control (e.g., a substance

that matches the chemical class of the test substances) could be run at intervals for quality control purposes rather than concurrent with each experiment. The Panel cited Kimber et al. (2006), which describes "routine" use of the "rLLNA" utilizing only a vehicle and a high-dose group, as a rationale for this suggestion. However, the Panel does not recommend omitting the concurrent positive control in laboratories that perform the LLNA only "occasionally."

Based on the analyses presented in the draft BRD, the Panel considered the accuracy of the LLNA limit dose procedure to have been adequately evaluated and compared to the traditional LLNA, mindful of the limitations associated with a retrospective evaluation. For instance, it cannot be assumed that the compounds tested in the retrospective studies were always tested at the highest possible dose unless such information was explicitly indicated. In this regard, the Panel recommended that a more detailed description of what is considered "*avoidance of excessive irritation*" and "*evidence of systemic toxicity*" be included in any LLNA protocol in order to aid in choosing the most appropriate high (i.e., limit) dose, although specific indicators of "*systemic toxicity or excessive irritation*" were not formally discussed.

The Panel agreed that it was appropriate to assume that the intra- and inter-laboratory reproducibility of the LLNA limit dose procedure and the traditional LLNA would be similar, because reproducibility is more dependent on the method than on the number of dose groups. However, reducing the number of test substances dose groups from three to one might reduce the sensitivity of the assay. The traditional LLNA may have a greater chance of correctly identifying a sensitizer even in the presence of one or more technical errors since data from three dose groups are being considered and an SI $\geq$3.0 at any dose group would result in the substance being classified as a sensitizer. However, for the purpose of adopting an assay that uses fewer animals and provides increased throughput for testing purposes, these hypothetical considerations are not a sufficient reason to argue against use of the limit dose LLNA procedure.

LLNA for Testing Aqueous Solutions, Metals, and Mixtures

The draft ICCVAM recommendations state that, although more data are needed to assess the use of the LLNA for testing for mixtures and aqueous solutions before a recommendation can be made, the traditional LLNA appears to be useful for the testing of metal compounds, with the exception of nickel. The Panel agreed with these draft ICCVAM recommendations. Regarding the use of the LLNA for testing mixtures, the Panel acknowledged that the ability of ICCVAM to develop draft test method recommendations was limited not only by the amount of data available, but the relatively poor concordance of traditional LLNA outcomes in comparison to those obtained in guinea pig tests, and recommended that this be noted in the final ICCVAM recommendations. The term "mixtures" can represent an infinite number of materials and it would be more beneficial to specify types or formulations of mixtures that are being examined.

Regarding metals, the Panel concluded that the accuracy statistics for the traditional LLNA when compared to results obtained from evaluation in humans supported use of the traditional LLNA as a hazard identification tool for metals, with the exception of nickel, which produces variable responses. One minority opinion stated that the results for nickel compounds were not entirely questionable and that the traditional LLNA might also be suitable for testing nickel compounds.

Thus, the Panel recommended further evaluation of the variable results obtained for nickel in the context of the available literature on allergic contact dermatitis to nickel in humans.

Regarding substances tested in aqueous solutions, the Panel suggested expanding the brief section of the draft test method recommendations discussing the test method protocol for the traditional LLNA to specifically point out how the conclusions of the applicability domain evaluation may affect the standard traditional LLNA protocol. For instance, it could be suggested that aqueous test solutions be avoided due to problems associated with skin application. It would be preferable for a hierarchy of organic solvents to be considered as dosing vehicles, with emphasis on using a vehicle to which humans may actually be exposed in circumstances linked to occupational sensitization.

The Panel agreed with the draft ICCVAM recommendation for continued accrual of information from traditional LLNA evaluations of mixtures, metals, and aqueous solutions with comparative data for guinea pig (i.e., guinea pig maximization test [GPMT] or Buehler test [BT]) and human (i.e., human maximization test [HMT] or human repeat insult patch test [HRIPT]) tests. However, the Panel suggested that, given resource limitations, it would be important to organize the recommendations based on relative priority.

The draft Addendum to the original validation report for the traditional LLNA (ICCVAM 1999) provided a comprehensive review of currently available data and information for evaluating the usefulness and limitations of the traditional LLNA for assessing the skin-sensitization potential of mixtures, metal compounds, and substances tested in aqueous solutions. The Panel evaluated the draft Addendum for completeness, errors, and omissions and concluded that there were no apparent errors or omissions, although they did state that the term “mixtures” was used too broadly (i.e., can represent an infinite number of materials) and it would be more beneficial to specify types or formulations of mixtures that are being examined.

The Panel did not identify any classes of chemicals missing from the dataset used to review the utility of the traditional LLNA for testing aqueous solutions. However, while they did not propose an alternative, the Panel expressed concern over the most appropriate definition for an aqueous solution (defined in the draft Addendum as any solution containing ≥20% water). For the mixtures included in the analysis, the Panel noted that quantitative compositions had not been provided and therefore they could not comment on whether these mixtures were representative of the types of mixtures typically tested in the traditional LLNA. With respect to metals (none of which are mixtures), there was a paucity of important representatives of commercially useful metals such as platinum, palladium, iron, zinc, manganese and silver in the data set. The Panel suggested that to enlarge the group of metal non-sensitizers, substances used as cosmetic ingredients (e.g., titanium dioxide) and aluminum compounds currently used in antiperspirants might be considered.

The Panel agreed that, although it was important to identify data obtained according to GLP guidelines, data obtained from non-GLP studies should not be excluded automatically from this retrospective analysis. The Panel concluded that other factors could be used to identify high quality data. Examples include data published in peer-reviewed journals or obtained from a study conducted in a laboratory that has GLP capabilities.

The Panel concluded that, considering the limited comparative data that were available, particularly for mixtures and aqueous solutions, the accuracy assessment of the traditional LLNA for testing mixtures, metals, and aqueous solutions when compared to available human and/or guinea pig test results was as comprehensive as was possible. The limited amount of comparative data made it unfeasible to draw definitive conclusions for mixtures and aqueous solutions from the available accuracy statistics.

Non-Radioactive LLNA Protocol - The LLNA: DA Test Method

The Panel concluded that the available data and test method performance support the ICCVAM draft recommendations for the LLNA: Daicel Adenosine Triphosphate test method (LLNA: DA), and that the test method may be useful for identifying substances as potential skin sensitizers and non-sensitizers, but that this recommendation is contingent upon receipt, review, and analyses of additional existing data and information from the test method developer. Therefore, this non-radioactive version of the traditional LLNA cannot currently be recommended for the hazard identification of skin sensitizing substances, regardless of whether or not there are restrictions on the use of radioactive materials, until such time as this existing data has been received and confirmed.

The draft LLNA: DA BRD was compiled to provide a comprehensive review of available data and information evaluating the usefulness and limitations of the LLNA: DA test method to assess the allergic contact dermatitis potential of chemicals and other substances. The Panel evaluated the draft BRD for completeness, errors, and omissions and recommended that their suggestions/corrections relating to general, statistical, and specific editorial issues be incorporated into future revisions.

The Panel agreed that five animals per dose group should be recommended for validation of modified LLNA test methods. The Panel, however, noted that supplemental power calculations for the LLNA: DA test method indicated that the power for detecting a three-fold increase in the treatment group was estimated to be 95% for a sample size of three mice per dose group. Thus, the Panel identified the use of three animals per dose group as a potential opportunity to reduce animal number when using modified assays in the future, assuming all essential validation requirements can be successfully met. A minority opinion expressed by five Panel members was that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals per dose group could be considered.

Generally, the Panel viewed the difference in treatment schedule between the LLNA: DA and the traditional LLNA to be potentially significant if the LLNA: DA induced the elicitation phase of skin sensitization. The Panel was concerned that the 1% sodium lauryl sulfate (SLS) pretreatment step in the LLNA: DA might modify the inherent sensitivity of the LLNA. They recommended that the test method developer (Daicel Chemical Industries, Ltd.) justify the use of 1% SLS or consider an alternative decision criterion (i.e., an SI threshold other than 3.0) such that the 1% SLS pretreatment is no longer necessary.

The Panel considered the database of substances tested in the LLNA: DA to be representative of a sufficient range of chemicals expected to be tested for skin-sensitization potential, and concluded that the accuracy analysis had made appropriate comparisons to the traditional LLNA, guinea pig tests, and human data/experience. The Panel could not identify specific characteristics associated with the one false negative (i.e., 2-mercaptobenzothiazole) or the one false positive (i.e., benzalkonium chloride), but reemphasized that the potential impact of pretreatment with 1% SLS in this context needed to be considered.

With regard to test method reliability, the Panel concluded the intralaboratory reproducibility of the LLNA: DA had not been adequately evaluated. They noted that the two sensitizers tested had similar chemical structures (i.e., eugenol and isoeugenol) and that it was unclear if the tests were truly independent. The Panel also noted that the interlaboratory reproducibility of the assay could not be adequately evaluated given the lack of original laboratory data and limitations in the study design. In particular, they cited the use of pooled lymph nodes from the mice in each treatment group and the testing of each substance at predetermined dose levels established by the lead laboratory as study design limitations. Still, a Panel minority considered pooled data acceptable and the setting of dose levels for all laboratories based on results from the lead laboratory to be reasonable.

The Panel also commented that ideally, test substances should be coded during the validation of a new assay, although they did not feel that a lack of coding constituted a reason for rejecting the current LLNA: DA dataset. The Panel also commented that although GLP compliance is highly recommended for validation studies, the current studies should not be rejected solely on the basis of a lack of GLP compliance. However, the Panel considered it important to obtain the original records for all validation studies (which have been requested by NICEATM) in order to confirm that the reported data were the same as the data recorded in the laboratory notebooks.

With regard to the 5% (1/19) false negative and 10% (1/10) false positive rates obtained with the LLNA: DA, the Panel commented that it was important to identify reasons why the substances gave “false” results, taking into consideration factors such as intended use of the substances and the target population. They agreed that it might be useful to follow a suspected inaccuracy with an investigation of the mechanistic basis for the discordance since it may help to establish a biologically-based rationale for the discordance.

The Panel noted that the available LLNA: DA data did not support all of the ICCVAM draft recommendations in the proposed test method standardized LLNA: DA protocol. First, although the Panel agreed with the ICCVAM protocol that recommends five animals per dose group, they noted that supplemental statistical information provided for the LLNA: DA test method implied that using less than five animals per dose group was acceptable (e.g., a 3.0-fold increase in the SI value would likely be detected with 99% confidence when using four animals per dose group). In addition, the Panel considered it important to adequately characterize the effect of the 1% SLS pretreatment step in the LLNA: DA, and it should be demonstrated that the day 8 applications do not induce a skin reaction that could be indicative of the onset of the elicitation phase of skin sensitization. Keeping these points in mind, the Panel agreed that if the limit dose

procedure was applicable to the traditional LLNA, then it would also be applicable to the LLNA: DA in order to further reduce the number of animals used.

The Panel also stated that the available data supported the ICCVAM draft recommendations for the LLNA: DA in terms of future studies, which included performing a more comprehensive evaluation using more non-sensitizers within and across laboratories. A minority opinion stated by one Panel member was that although testing more sensitizers might be warranted for interlaboratory validation studies, a sufficient number of non-sensitizers (n=11) had already been tested within the same laboratory.

The Panel also commented that the protocol differences between the LLNA: DA and the traditional LLNA could not clearly be constituted as "major" or "minor" changes. However, they considered this issue largely irrelevant if a test method was able to correctly predict the dermal sensitization potential of a test substance. Consequently, the Panel concluded that the current draft ICCVAM Performance Standards could be applicable to the LLNA: DA as a mechanistically and functionally similar test method.

Non-Radioactive LLNA Protocol - The LLNA: BrdU-FC Test Method

Overall, the Panel concluded that the available data and test method performance of the LLNA with bromodeoxyuridine (BrdU) detected by flow cytometry (LLNA: BrdU-FC) supported the draft ICCVAM recommendations that it may be useful for identifying substances as potential skin sensitizers and non-sensitizers, but that more information and existing data must be made available before the LLNA: BrdU-FC can be recommended for routine use. The Panel concluded that the test method usefulness and limitations identified in the draft ICCVAM recommendations accurately summarized the limits of the information supplied and the additional information that would need to be generated or provided for further consideration of the test method. As a result, the Panel concluded that the LLNA: BrdU-FC could not currently be considered as a scientifically valid replacement alternative to the traditional LLNA. Still, the Panel suggested that the test method recommendation should clearly state that the test method was not "invalid," but simply that there was currently not sufficient evidence and information to state that it had been adequately validated.

The draft LLNA: BrdU-FC BRD was compiled to provide a comprehensive review of available data and information evaluating the usefulness and limitations of the LLNA: BrdU-FC test method to assess the ACD-inducing potential of substances. The Panel evaluated the draft BRD for completeness, errors, and omissions and recommended that their recommendations/corrections relating to general, statistical, and specific editorial issues be incorporated into future revisions.

The LLNA: BrdU-FC included routine measurements of ear swelling as an indicator of excessive skin irritation. The Panel viewed that this, or any other quantitative measurement of skin irritation, should be carefully considered for inclusion in all LLNA protocols. The Panel considered inclusion of optional quantification of immunophenotypic markers as an additional mechanism for distinguishing irritants from sensitizers to be useful, as it might reduce the frequency of false positives (i.e., substances which are actually skin irritants) and improve

comparisons with human data. However, they considered application of immunological markers too detailed and costly for routine LLNA use (i.e., for hazard classification purposes) and more suited for research purposes.

The Panel noted that the substances tested in the LLNA: BrdU-FC seemed representative of a sufficient range of chemical classes and physical chemical properties, and thus that the test method appeared applicable to many of the types of chemicals and products that are typically tested for skin-sensitization potential. However, the Panel considered the total database available for evaluation of the validation status of the LLNA: BrdU-FC to be relatively small compared to the large number of substances assessed in the traditional LLNA. Therefore, the Panel recommended caution when making conclusions related to its concordance with the traditional LLNA. Still, the accuracy of the LLNA: BrdU-FC was considered adequately evaluated and comparable to the traditional LLNA.

The Panel concluded that intralaboratory reproducibility was not adequately assessed and it should be better evaluated in order to support the validation of this test method. The Panel suggested that although the studies evaluated in the draft BRD were not GLP-compliant, this should not affect acceptance of the data for an evaluation of the validation status of this test method. However, some sources of variability in the intralaboratory data, such as failure to appreciate differences in composition of dosing solutions between experiments caused by test article instability or other phenomena, might be obscured if not in complete compliance with GLP guidelines. Thus, the Panel suggested that any additional studies undertaken to validate the test method should ideally be GLP-compliant.

The Panel agreed that the available data supported the ICCVAM draft recommendations for the LLNA: BrdU-FC procedure in terms of the proposed test method standardized protocol. They suggested that the utility of ear swelling or other methods to detect inflammation appeared warranted in every variation of the LLNA (including the traditional LLNA), but should be further investigated before routine inclusion in the protocol is recommended. The Panel also concluded that the traditional LLNA limit dose procedure could be applied to the LLNA: BrdU-FC, keeping in mind the limitations associated with a "limit dose" procedure.

The Panel further agreed that the ICCVAM draft recommendations for future studies highlighted the unanswered questions raised by the available data set. Specifically, conducting interlaboratory studies as a part of the validation process is important. The Panel considered the immunological markers suggested for the LLNA: BrdU-FC to be acceptable, but that additional immunological markers for discrimination of irritant versus sensitization phenomena were also possible. In general, for any future work, efforts should be made to decrease the variability and to thereby increase the power of the test in order to ensure that more animals were not needed relative to the traditional LLNA or other alternative LLNA protocols.

The Panel considered the protocol differences between the LLNA: BrdU-FC and the traditional LLNA to be "minor" changes, and therefore concluded that assessment of the validity of this test method could be based on the draft ICCVAM LLNA Performance Standards. The Panel also cautioned, however, that a clear definition of what constituted a "major" versus a "minor"

change, or a different protocol altogether could be better addressed once the recommendations for the current draft ICCVAM LLNA Performance Standards were finalized.

Non-Radioactive LLNA Protocol - The LLNA: BrdU-ELISA Test Method

The Panel concluded that the available data and test method performance for the LLNA with BrdU detected by enzyme-linked immunosorbent assay (LLNA: BrdU-ELISA) support the ICCVAM draft recommendations that it may be useful for identifying substances as potential skin sensitizers and non-sensitizers, but that more information and existing data must be made available before the LLNA: BrdU-ELISA can be recommended for use. The Panel also stated that a detailed protocol was needed, in addition to sufficient quantitative data for broader analysis on a larger set of balanced reference substances that take into account physicochemical properties and sensitization potency, as well as an appropriate evaluation of interlaboratory reproducibility.

The draft LLNA: BrdU-ELISA BRD was compiled to provide a comprehensive review of available data and information evaluating the usefulness and limitations of the LLNA: BrdU-ELISA test method to assess the ACD-inducing potential of chemicals and other substances. The Panel evaluated the draft BRD for completeness, errors, and omissions and recommended that their suggestions/corrections relating to general, statistical and specific editorial issues be incorporated into the final document.

The Panel's main concern with the test method was that the accuracy of the LLNA: BrdU-ELISA at SI ≥3.0 was inadequate and not equivalent to the traditional LLNA. Furthermore, although using a decision criterion of SI ≥1.3 improved the test's performance in identifying sensitizers from non-sensitizers, it did not resolve concerns about the test method. Based on a power analysis for the LLNA: BrdU-ELISA, which was provided to the Panel as supplemental information, the Panel concluded that it was difficult to justify using a SI ≥1.3 as the cutoff value, given the much larger number of animals that would be required to detect a 1.3-fold increase above vehicle controls with similar power to the traditional LLNA when five animals per dose group are used. For a three-fold increase, the supplemental statistical analyses indicated that a sample size of four was sufficient. Still, the Panel agreed with the ICCVAM recommendation to use five animals per dose group and to collect individual animal data. They concluded that this would allow for more robust calculations in the event that an outlier prevented some of the data from being included in the analysis. A minority opinion by five Panel members was stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from a least four animals could be considered.

The Panel noted that in organizations where the use or disposal of radioactive materials was restricted, the potential to use the LLNA: BrdU-ELISA could reduce the number of animals needed per test compared to the traditional LLNA and would result in less pain and suffering compared to using traditional guinea pig test methods. However, if the SI ≥1.3 was chosen as the decision criterion because of its improved accuracy compared to SI ≥3.0, the Panel stated that the number of mice needed to perform the LLNA: BrdU-ELISA test should be compared to

the number of guinea pigs that would be needed for skin-sensitization tests in order to assess if the LLNA: BrdU-ELISA actually reduced overall animal use for skin-sensitization testing.

In general, the Panel considered the number of substances tested in the LLNA: BrdU-ELISA too few, and that data from more substances tested using the traditional LLNA, guinea pig tests, and human tests should have been included. The Panel also did not consider the available data from the LLNA: BrdU-ELISA to be representative of a sufficient range of chemical classes and physical chemical properties. The limited dataset prevents an evaluation of whether the test method would be considered applicable to any of the types of chemicals and products typically tested for skin-sensitization potential.

However, the Panel concluded that the appropriate comparisons between the traditional LLNA, guinea pig test and human data had been made. The Panel agreed that the false negative rate for hazard identification using the SI ≥3.0 in the LLNA: BrdU-ELISA was excessive (i.e., using this SI threshold value, the LLNA: BrdU-ELISA misclassified 29% and 39% of the substances classified as sensitizers in the traditional LLNA or in humans, respectively).

The Panel also considered that the intralaboratory reproducibility of the LLNA: BrdU-ELISA was not adequately evaluated and compared to the traditional LLNA. The Panel indicated that the number of substances was too few, and in some cases there was a wide variation in repeat tests of the same substance. The Panel recommended a more comprehensive evaluation of the intralaboratory reproducibility of the test method, using different SI values, and that the analysis of the variability of the estimated concentration needed to produce a positive SI value (ECt values) be conducted on a log scale.

The Panel also noted that interlaboratory reproducibility for the LLNA: BrdU-ELISA could not be evaluated because neither the design of the study sponsored by the Japanese Center for Validation of Alternative Methods nor any of the resulting data had been provided in advance of their evaluation. The Panel agreed that a multi-laboratory validation study using a balanced set of chemicals would adequately characterize the interlaboratory reproducibility of the LLNA: BrdU-ELISA.

In general, the Panel agreed that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA procedure in terms of the proposed test method standardized protocols. However, as noted above, a minority opinion by five Panel members was that there could be circumstances in which pooled data from at least four animals could also be acceptable. The Panel also stated that if the LLNA: BrdU-ELISA was found to be equivalent to the traditional LLNA in the future that it would be appropriate to apply the LLNA limit dose procedure to the test. The Panel also agreed with ICCVAM's test method recommendations for future studies and emphasized that more data were needed in order to determine the appropriate threshold value for the decision criterion. The Panel concluded that it might be more appropriate to use a statistically-based decision criterion rather than a stimulation index to classify substances as sensitizers, and that this should be further investigated.

The Panel agreed that the LLNA: BrdU-ELISA protocol differed from the traditional LLNA only in the method used to assess lymphocyte proliferation and as such concluded that this represented a "minor" change (as defined in the current draft ICCVAM LLNA Performance Standards) and separate performance standards for the LLNA: BrdU-ELISA were not needed.

Draft ICCVAM LLNA Performance Standards

The draft ICCVAM LLNA Performance Standards are intended to evaluate the acceptability of proposed test methods that are mechanistically and functionally similar to the traditional LLNA. ICCVAM proposed that the applicability of the draft ICCVAM LLNA Performance Standards be restricted to protocols that incorporate "minor" modifications to the traditional LLNA procedure, defined as changes only to the method for measuring lymphocyte proliferation. The Panel agreed that different methods of measuring lymphocyte proliferation represent "minor" modifications, but recommended that, instead of trying to define "minor" modifications, a better strategy might be to define criteria that would need to be satisfied in order to ensure that the alternative test method was mechanistically and functionally similar to the traditional LLNA (e.g., only measure cell proliferation associated with the induction phase of a skin-sensitization reaction). The Panel considered that the draft performance standards were also appropriate for evaluating other modifications. Examples of acceptable modifications included test animal sex, strain, the use of rats rather than mice, the number of animals per group, and timing of test article treatment. One minority opinion considered the potential impact of changes to protocol components other than the method of measuring lymphocyte proliferation to be significant and therefore would require more extensive validation, which was not defined.

The Panel indicated that alternative LLNA protocols that are undergoing validation should contain essential test method components that follow the ICCVAM-recommended protocol (ICCVAM 1999; Dean et al. 2001), unless adequate scientific rationale for deviating from this protocol was provided.

The Panel also identified aspects of the LLNA that should be required as part of the test method validation process: (1) application of the test substance to the skin with sampling of the lymph nodes draining that site, (2) measurement of cell proliferation in the draining lymph node, (3) absence of a skin reaction that could be indicative of the onset of the elicitation phase of skin sensitization, (4) data collected at the level of the individual animal to allow for an estimate of the variance within control and treatment groups (using this variance, a power analysis needs to be conducted to demonstrate that the modified method is utilizing a sufficient number of animals per treatment group to permit hazard identification with at least 95% power), and (5) if dose response information is needed, there are an adequate number of dose groups ($n \geq 3$) with which to accurately characterize the dose response for a given test substance.

The Panel noted that the list of substances included in the draft ICCVAM LLNA Performance Standards was sufficiently representative of the types of materials that are likely to be tested for skin sensitization. However, among the 13 sensitizers in the list of "required" substances, only five were considered to have robust data (i.e., traditional LLNA data based on at least three independent studies).

To evaluate performance for use in hazard identification, the Panel concluded that all 22 substances in the draft ICCVAM-recommended list should be tested and accuracy statistics calculated (Note: this list of substances includes "required" substances as well as "optional" false negative and false positive substances, of which only 8/22 have "robust" datasets [$n \geq 3$ as defined by the Panel]). To the extent possible, a rationale for any discordant results should be provided. However, the most potent sensitizers (e.g., dinitrochlorobenzene [DNCB]) should always be identifiable. Also, considerable weight should be given to the balance between animal welfare and human safety when considering the adequacy of test method accuracy. Based on the limited data available for the sensitizers on the list and the lack of standardization of test methods from which the results were obtained, the current database does not support inclusion of ECt values as a component of the accuracy evaluation.

The Panel agreed with the draft ICCVAM recommendations for evaluating test method reliability. These recommendations included obtaining ECt values that are generally within 0.5x to 2.0x of the mean historical EC3 (i.e., estimated concentrations needed to produce an SI of 3) values for hexyl cinnamic aldehyde (HCA) (intralaboratory, n=4 experiments in one laboratory), or HCA and DNCB (interlaboratory, n=1 experiment in three laboratories). However, the Panel recommended that the criteria for independent tests should be specified (e.g., different animal shipment, different reagents, different operator). The Panel concluded that the proposed criteria for acceptability appeared to be appropriate in this case, because only one or two substances were being evaluated (i.e., a statistical multiple-comparisons[5] problem does not exist). The Panel also suggested that historical control data using HCA and DNCB in the same vehicle could be used to demonstrate adequate intra- and/or inter-laboratory reproducibility.

The Panel also recommended that statistical tests to analyze the data might allow for a more accurate interpretation. They recommended that a suitable variance-stabilizing transformation (e.g., log transformation, square root transformation) be applied in all statistical analyses and in reporting summary standard deviations. The Panel also recommended that a more rigorous evaluation be conducted of what would be considered an appropriate range of ECt values to include as a requirement. This would be a statistical evaluation that considers the variability of ECt values generated among the sensitizers included on the performance standards reference substances list and the statistical multiple comparisons problem.

Use of the LLNA for Potency Determinations

The Panel agreed with the draft ICCVAM recommendation that the LLNA should not be used as a stand-alone assay for categorizing skin sensitizers as strong vs. weak, but that it could be used as part of a weight-of-evidence evaluation (e.g., along with quantitative structure-activity

[5] When multiple experiments are conducted and multiple observations, comparisons, or hypothesis tests are conducted, the chance of observing rare events increases. Suppose, for example, that an interval is established such that 5% of observations from a particular population of data are outside that interval. Then if *k* independent experiments generate data from this population (e.g., a standard normal distribution), the chances that all 20 results will lie inside the interval is $(1.0 - 0.05)^k$ (N. Flournoy, personal communication).

relationships, peptide reactivity, human evidence, historical data from other experimental animal studies) for this purpose. The Panel also agreed with the draft ICCVAM recommendation that any LLNA studies conducted for the purpose of evaluating skin-sensitization potency should use the ICCVAM recommended LLNA protocol. In addition, the Panel viewed that the relevant testing guidelines for the traditional LLNA should be revised to include the procedure for calculating an EC3 value.

A draft BRD was compiled by ICCVAM that provided a comprehensive review of available data and information and an evaluation of the usefulness and limitations of the traditional LLNA for the categorization of substances with regard to skin-sensitization potency. The Panel evaluated the draft BRD for completeness, errors, and omissions and noted alternative analyses that would allow for a more complete evaluation of the use of the traditional LLNA for skin-sensitization potency categorizations (see below).

The Panel agreed that the database of substances evaluated for potency determinations was sufficient and represented a range of chemical classes and physicochemical properties applicable to products typically tested for skin-sensitization potential. The Panel also concluded that since the database was compiled from existing data, the lack of substance coding likely had no impact on the retrospective evaluation presented in the draft BRD. Still, the Panel recommended the coding of test substances in any future validation studies. The Panel generally agreed that potency determinations based on traditional LLNA results should ideally be limited to data from studies that evaluated lymph node proliferation in individual animals so that outliers and technical errors could be identified. However, they also agreed that pooled animal data should not be excluded automatically from a retrospective analysis.

The Panel indicated that the relevance of the LLNA for potency determinations had been adequately compared and evaluated to human (i.e., HMT or HRIPT) and guinea pig (i.e., GPMT or BT) data. A minority opinion stated by one Panel member was that the relevance of the traditional LLNA to human clinical observations had not been sufficiently determined.

In general, the Panel agreed that the proposed two-level categorization scheme (weak vs. strong sensitizers) for both human and guinea pig data was appropriate. However, a minority opinion stated by two Panel members was that a moderate category should be included since certain compounds might be on the border between weak and strong sensitizers. Thus, they suggested that the five-category scheme proposed by Kimber et al. (2003), which includes non-sensitizers, might be recommended.

The Panel concluded that the decision criteria providing the best overall performance was the use of <250 $\mu g/cm^2$ to distinguish between strong and weak sensitizers in humans and the use of an LLNA EC3 $\leq 9.4\%$ to distinguish between strong and weak sensitizers in the LLNA. The Panel stated that more data would be needed to determine if values different from these two would be more appropriate. The Panel also recommended that safety factors other than 10 for the lowest observed effect level (LOEL) be evaluated to determine if improved results could be obtained. The Panel also suggested an analysis that directly compares the LOEL values without using a safety factor (i.e., using LOEL data only) and an analysis that only uses no observed effect level data. The Panel further stated that traditional LLNA tests based on pooled or

individual lymph nodes for a dose group should be evaluated independently to assess the impact of using pooled data on the accuracy analysis for skin-sensitization potency. Finally, the Panel stated that the effect of different vehicles should be recognized as a limitation in the current data analysis and a likely contributor to the variability observed within and across laboratories.

The Panel stated that data from studies that could not be confirmed as being GLP-compliant, but that were from peer-reviewed literature or sources with high-quality laboratory management practices, were still appropriate to include in the accuracy analysis. However, the Panel stated that, ideally, GLP compliance should be the standard, as it is clearly the only objective way to judge the credibility of the data.

The Panel recommended that more data should be collected to determine the optimal threshold in humans for distinguishing between strong and weak sensitizers. In addition, the Panel discouraged conducting additional animal studies unless such studies would be expected to lead to an overall reduction in animal use. The Panel recommended that the LOELs from Akkan et al. (2003) be used instead of the DSA_{05} (i.e., the dose per skin area leading to a sensitization incidence of 5%) values from Schneider and Akkan (2004) in all of the potency analyses. A minority opinion by one Panel member stated that it was acceptable to use the DSA_{05} values from Akkan et al. (2003) as LOEL values in the evaluation. This panelist mentioned that the DSA_{05} value is a LOEL value adjusted to 5% incidence of induction in order to correct for human studies leading to different inductions. Furthermore, the panelist stated that because the DSA_{05} is corrected for an induction rate of 5%, it would be better to compare with the traditional LLNA EC3 than to use the default uncorrected LOEL.

1.0 Murine Local Lymph Node Assay (LLNA) Limit Dose Procedure[6]

1.1 Comments on the Draft Background Review Document (BRD) for Completeness, Errors and Omissions

1.1.1 General Comments

The international independent scientific peer review panel (hereafter, Panel) was asked if there were errors in the draft LLNA limit dose procedure BRD that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. The Panel agreed that consideration should be given to applying the same term to the LLNA limit dose procedure since in various places throughout the draft BRD it was referred to differently as either the "cut-down," the "limit dose," or the "reduced LLNA" (rLLNA). Since the European Centre for the Validation of Alternative Methods (ECVAM) has already established a naming convention of "rLLNA," the Panel recommended adopting the ECVAM terminology to harmonize the terminology used among the international validation agencies.

The Panel recommended that since the validation of the LLNA limit dose procedure encompassed data that was analyzed retrospectively, a discussion of the limitations of a retrospective evaluation of previously published LLNA results should be included in the final BRD. In particular, the assumption that the highest dose in the retrospective dose-response study would be equivalent to the highest possible dose tested in the limit dose procedure should be addressed. Discussing such a limitation would be important since it bears directly on the validation of the limit dose procedure.

Further, since determination of the appropriate "limit dose" is critical to the LLNA limit dose procedure, the Panel suggested that a discussion of how to arrive at the maximal concentration for test substance dosing should be included in the final BRD. The final BRD should also specifically define what is meant by the terms "avoidance of excessive irritation" and "systemic toxicity" to aid in choosing the most appropriate maximum dose. In this regard, the Panel suggested that a systematic and quantitative measurement of ear thickness and systemic toxicity be considered or evaluated for routine inclusion in the LLNA protocol.

The Panel discussed modifying the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) LLNA protocol requirement for testing concurrent positive controls (ICCVAM 1999; Dean et al. 2001) as a means of further streamlining the LLNA limit dose procedure (i.e., reducing animal number, cost, etc.). Although the Panel did not reach consensus, a suggestion was made that for laboratories in which the LLNA is "routinely" performed and which had demonstrated the ability to consistently obtain positive results, hexyl cinnamic aldehyde (HCA) or another positive control (e.g., a substance that matches the chemical class of the test substances) could be run at intervals for quality control purposes rather than concurrent with each experiment. The Panel noted that Kimber et al. (2006) have described the "routine" use of the "rLLNA" utilizing only a vehicle and a high-dose group. The Panel also recommended that for laboratories that perform the LLNA only "occasionally," a concurrent positive control should be used. However, in their discussions, the Panel was not able to conclude what would constitute "routine" or "occasional" LLNA use or what would be

[6] Also referred to as the "reduced" LLNA

an appropriate interval between positive control testing when a concurrent positive control is not used.

The Panel also noted that including the following additional information in the final BRD might prove informative if included:

- An indication of any procedural problems reported for the tests
- An indication of the range of historical values obtained with the negative and positive controls (the positive control historical range might give insight into the need for a concurrent positive control)
- Any discussion of global harmonization should expand on why the draft BRD did not place more reliance upon Organisation for Economic Co-operation and Development (OECD) Test Guideline (TG) 429 (OECD 2002) as a normative reference
- For two of the substances tested multiple times (HCA and linalool alcohol), different doses were used and discordant results were obtained. It should be noted for which (if either) of these tests, the highest dose tested was consistent with the dose selection principles set out in the draft BRD

1.1.2 General Statistical Comments

The Panel also had some statistical comments related to the LLNA limit dose procedure. First, the Panel viewed that a reference to the supplemental statistical information in which Dr. Joseph Haseman performed power calculations on the traditional LLNA would be useful in determining if the sample size used in the LLNA limit dose procedure was adequate for evaluating skin-sensitization potential. Also, the Panel concluded that although a stimulation index (SI) based on the ratio of 3.0 as the decision criterion for a sensitizer is informative, statistical analysis determining if the test substance is significantly different from the control substance should be recommended.

1.1.3 Comments with Specific References to the Text

The Panel also identified the following minor formatting and grammatical errors, and information gaps in the draft BRD:

- The manner of notating numerical data in the draft BRD tables was not consistent (e.g., in some places the value “one” was shown as 1, elsewhere as 1.0; in a few places the SI was shown over the percent concentration used).
- Lines 291-294: The statement was made that “chemical class information is included to provide an indication of the variety of structural elements present in the substances that were evaluated in this analysis, but it is not intended to suggest an impact of structure on sensitization potential.” The latter concept is not entirely correct; the portion of the sentence stating “but it is not intended to suggest an impact of structure on sensitization potential” should be omitted. Consideration should be given to using the large database of chemicals to selectively modify structure-activity relationship (SAR) software for improved predictivity. This

could likely be accomplished by communication with software developers to point out availability of the newly expanded ICCVAM dataset developed for evaluation of the LLNA limit dose procedure.

- Lines 299-300: The sentence is incomplete; "non-sensitizers" should be inserted at the end of the sentence.
- Line 358: The citation to Sailstad et al. (2001) was not listed in Section 12.0 (References) and should be included.
- Lines 365-384 (Section 1.1.2): Consideration should be given to expanding the background on the mechanism and natural history of allergic contact dermatitis (ACD). Some additional detail regarding the biochemistry and cell biology of immune induction and elicitation would be useful as an orientation to how the LLNA functions as an integrated method of detection for ACD.
- Lines 366–368: The introductory sentence on the prevalence of ACD as an occupational health issue would benefit from amplification to also indicate that ACD is of variable severity with some potentially severe ramifications, and that ACD is recurrent upon rechallenge possibly requiring workplace accommodation or change of employment.
- Lines 366-368: There is no reference provided for the statistic from the U.S. Department of Labor Bureau and Labor Statistics cited in Section 1.1.2 of the draft BRD.
- Line 433: The reason for further evaluation of negative results with concentrations less than 10% should be clarified (Kimber et al. 2006).
- Lines 435-436: This bullet point re-plays conclusions made in the summary of the "rLLNA" issued by the scientific advisory committee of the European Centre for the Validation of Alternative Methods (ECVAM). However, in the draft BRD the statement has been altered and should instead read "..., as appropriate, per OECD TG 429 (OECD 2002)" rather than citing ICCVAM (1999) and Dean et al. (2001).
- Lines 452-453: The intent of the sentence would be clarified by modifying to read "...to identify potential human skin sensitizers through quantification of lymphocyte proliferation in the test method."
- Lines 496-500: Reading of the Kimber et al. (2006) citation does not indicate a recommendation for a concurrent positive control group. Thus, the sentence in the draft BRD that reports use of vehicle and positive control groups in the limit dose procedure based on the Kimber et al. (2006) paper is incorrect.
- Lines 509-510: This sentence should also mention that the LLNA limit dose procedure, as published by Kimber et al. (2006), is rationalized not only as a means of bioresource economy but as a valid means of streamlining the LLNA for regulatory screening assessment purposes under regulation such as Registration, Evaluation and Authorisation of Chemicals (REACH).

- Lines 520-523: A footnote might be needed to explain why OECD TG 429 (OECD 2002) is not referenced here.
- Lines 629-636: Data donated by GlaxoSmithKline (GSK) to the National Toxicology Program Interagency Center for the Validation of Alternative Toxicological Methods (NICEATM)-ICCVAM LLNA review were generated under GLP conditions at a clinical research organization (CRO) repeatedly audited for GLP compliance by GSK. This information should be added to the text.
- Section 6.1: The presentation of data and associated discussion regarding limitations in accepting a 10% dose concentration cut-off should be repositioned for emphasis. This information is important in developing suggestions for a standard LLNA limit dose procedure. Data from Appendix D could be reduced to a small table or figure and be integrated into the body of the final BRD.
- Lines 722-723: This appears to be an incomplete sentence.
- Lines 815-822: Data donated by GSK to the NICEATM-ICCVAM LLNA review were generated under GLP conditions at a CRO repeatedly audited for GLP compliance by GSK. This information should be added to the text.
- Table following line 1126: There is a typographical error in the 2 x 2 table. It appears that the cell for Negative (New Test) x Total (Reference Test) should read "c + d" rather than "a + d."

1.2 Comments on the Validation Status of the LLNA Limit Dose Procedure

1.2.1 Test Method Protocol

For the proposed LLNA limit dose procedure, ICCVAM recommended that the number of animals used in each group should be the same as that recommended by ICCVAM for the traditional LLNA based on its 1998 evaluation (i.e., at least five animals per group), and that individual animal data should be collected and reported (ICCVAM 1999; Dean et al. 2001). The Panel was asked whether they agreed that these are appropriate protocol requirements for the limit dose procedure. The Panel agreed that, based on the supplemental power calculations for the traditional LLNA performed by Dr. Joseph Haseman (see **Table F-1**), a minimum of five animals per treatment group should be recommended for all future studies employing the limit dose procedure. In addition, the collection of individual animal data, as recommended by ICCVAM for the traditional LLNA (ICCVAM 1999; Dean et al. 2001), should also apply to all future studies following the LLNA limit dose procedure. Similarly, application of the LLNA limit dose procedure to a modified LLNA protocol would require adherence to a validated protocol with the exception of omitting the middle and low dose groups. Respective power calculations would indicate if application of the LLNA limit dose procedure to a validated modified LLNA protocol would allow fewer animals per dose group.

Table F-1 Power Calculations for the Traditional LLNA[1]

Parameter	3.0-fold increase[2]	2.5-fold increase	2.0-fold increase	1.5-fold increase	1.3-fold increase
Mean Rx response	1034.4	862.0	689.6	517.2	448.24
Log (Mean Rx response)	6.942	6.759	6.536	6.248	6.105
Difference (log scale)	1.099	0.916	0.693	0.405	0.262
Difference/SD	2.40	2.00	1.51	0.88	0.57
Power for N=5	95%	80-90%	50-80%	<50%	<50%
Power for N=4	90%	80%	50%	<50%	<50%
Power for N=3	50-80%	50-80%	<50%	<50%	<50%
Other power	–	–	95% (N=11)	95% (N=29)	95% (N=68)
Other power	–	–	90% (N=9)	90% (N=23)	90% (N=54)

Abbreviations: N = number of animals; Rx = Treatment; SD = standard deviation.

[1] The power calculations above are based on a one-sided $p<0.05$ Student's t test applied to log-transformed data from vehicle control LLNA tests.

[2] Fold-increase = Required increase above the vehicle control for a positive response (i.e., the stimulation index)

The primary rationale for both provisions is to underpin robust statistical analysis of LLNA results. Furthermore, the use of individual animal data would allow for the evaluation of dosing errors or other anomalies that might be masked by the use of pooled animal data.

1.2.2 Substances Used for the Validation Studies

The Panel was asked whether they considered the traditional LLNA database representative of a sufficient range of chemical classes and physical chemical properties such that it would be applicable to any of the types of chemicals and products that are typically tested for skin-sensitization potential. If not, the Panel was asked which relevant chemical classes/properties (other than those identified as limitations in the traditional LLNA) should be tested with caution, or not evaluated using the limit dose procedure, and which chemicals or products should be evaluated to fill this data gap. The Panel agreed that, in general, the traditional LLNA database included in the LLNA limit dose procedure evaluation was representative of a sufficient range of chemical classes and physical chemical properties and that it should be applicable to any of the types of chemicals and products that are typically tested in the traditional LLNA for skin-sensitization potential. It was notable that the substances included in this evaluation provided a diverse chemical database. Since much is known about the mechanism of sensitization, the LLNA should theoretically identify any chemical that works by migration of haptens to the lymph node. However, the Panel noted that substance classes that are sometimes problematic in the LLNA (i.e., metals) would also likely be problematic in the LLNA limit dose procedure. There were also some substance classes that had limited or no representation in the draft BRD (i.e., mixtures/formulations, higher molecular weight

biopharmaceuticals, and medical device materials). Thus, in general, the LLNA (and the LLNA limit dose procedure) is best used as part of a weight-of-evidence appraisal in which attributes such as physical chemical parameters, SAR evaluation, and indications of other biological activity involving potential chemical-to-biological macromolecule interactions, are carefully considered along with LLNA results to evaluate dermal sensitizing potential.

Because the LLNA limit dose procedure was based on a retrospective evaluation of existing data, most of which was not generated using coded chemicals, the Panel was asked whether a lack of coding of test substances adversely impacted or biased the current evaluation. The Panel considered that although coding of chemicals should be recommended for prospective validation studies, this evaluation was likely not adversely impacted or biased because of a retrospective evaluation of existing data. This is supported by the fact that many of the chemical structures included in the analyses do not appear to contain known structural motifs associated with ACD or chemical hypersensitivity and therefore there was no a priori expectation that the chemical tested would be a sensitizer. The Panel viewed it important to consider the issue of coding or bias in prospective validation studies.

For some substances submitted using the traditional LLNA test method, it was not possible to confirm whether the data were generated using pooled animal data for each dose group (as allowed in OECD TG 429 [OECD 2002]). ICCVAM (1999), Dean et al. (2001), and EPA (2003) recommend the use of statistical analyses to help interpret LLNA study results, which necessitates data collected at the level of the individual animal. Furthermore, Cockshott et al. (2006) reported that using individual animal data allowed for technical problems or other outliers during an experiment to be identified. The Panel was asked what impact the inclusion of pooled animal data might have on the accuracy analysis of the LLNA limit dose procedure. The Panel concluded that, although it would be important to note whether individual or pooled animal data were reported, the retrospective analysis of the LLNA limit dose procedure versus the traditional LLNA should not be limited to studies with confirmed individual animal data. The Panel stated that internationally, both individual and pooled animal data have likely been used both for regulatory decisions and for in-house decisions relating to product development and risk management. Also, the fact that the retrospective data analysis presented in the draft LLNA limit dose procedure BRD did not distinguish between individual or pooled animal data suggested that both met the quality standards for inclusion in the draft BRD.

1.2.3 Test Method Accuracy

The Panel was asked whether the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA limit dose procedure had been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft ICCVAM BRD). The Panel concluded that the relevance of the LLNA limit dose procedure had been adequately evaluated and compared to the traditional LLNA. Comparisons resulting in an accuracy of 98.9% (461/466), a sensitivity of 98.4% (308/313) and a specificity of 100% (153/153) for the LLNA limit dose procedure when compared to the traditional LLNA were sufficient to consider it adequately validated for use in the evaluation of skin sensitization, mindful of its known limitations that are described elsewhere. Still, the Panel noted that it was

important to keep in mind that a prospective analysis may not have the same accuracy as this retrospective analysis.

Furthermore, there were five substances for which the highest concentration tested produced an SI < 3.0, while lower concentrations of these substances produced an SI >3.0 (see Table 6-2 of the draft ICCVAM BRD). These substances were classified as false negatives compared to what was obtained in the traditional LLNA. The Panel was asked to identify any characteristics associated with these or other substances that might signal that this type of abnormal dose response might occur, and therefore, that using the LLNA limit dose procedure would not be appropriate. The Panel could not identify any common characteristics associated with the five false negative substances that would explain the non-linear dose response obtained. It was not known if any procedural problems were reported with these studies or what values were returned by the negative/positive control groups (in relation to other historical positive control values).

Thus, the Panel suggested that it might be worthwhile to examine whether LLNA results with these five false negative substances should be repeated. If the difference turned out to be repeatable, there could be hypothetical reasons to explain why the higher doses did not pass the SI threshold of 3.0. For example, under certain experimental conditions, the target lymphocytes may be selectively induced to a highly sensitive state by some chemicals at higher doses and may undergo either induction of apoptosis or inhibition of cell proliferation. Still, there was no evidence that these substances were immunomodulators that might have differentially stimulated or depressed the immune response depending on the dose and exposure. In any case, understanding false negatives is encouraged to ensure adequate protection of public health.

The Panel was asked whether the draft BRD adequately characterized the usefulness and limitations of the LLNA limit dose procedure based on the accuracy analyses. Overall, the Panel agreed that the draft BRD adequately characterized the usefulness and limitations of the LLNA limit dose procedure based on the accuracy analyses. Since the LLNA limit dose procedure and the traditional LLNA have close concordance, there was no need for detailed discussion in the draft BRD. However, it was not explicitly stated in the draft BRD that compared to a fully conducted traditional LLNA, a false positive result in the LLNA limit dose procedure is not possible (i.e., if the single dose used in the proposed limit dose procedure gives an SI ≥3.0, then so would the top dose in the traditional LLNA). Furthermore, prospective testing with the LLNA limit dose procedure to predict the sensitization potential of an unknown chemical was not discussed.

1.2.4 Test Method Reliability

The Panel was asked if it was appropriate to assume that the intra- and inter-laboratory reproducibility of the LLNA limit dose procedure and the traditional LLNA would be similar, based on the fact that they use identical protocols with the exception of the number of doses used (i.e., would reducing the number of test substance dose groups from three to one reduce the reliability of the assay?). The Panel agreed that it was appropriate to assume that the intra- and inter-laboratory reproducibility of the LLNA limit dose procedure and the traditional LLNA would be similar, because reproducibility is more dependent on the method than on the

number of dose groups. However, reducing the number of test substances dose groups from three to one could reduce the sensitivity of the assay (i.e., the ability to correctly identify sensitizers). The traditional LLNA may have a greater chance of correctly identifying a sensitizer even in the presence of one or more technical errors since there are data from three dose groups for consideration and an SI ≥3.0 at any dose group would result in the substance being classified as a sensitizer. However, for the purpose of adopting an assay that uses fewer animals and provides increased throughput for screening purposes, these hypothetical considerations are not a sufficient reason to argue against use of the limit dose LLNA procedure.

1.2.5 Data Quality

For some studies included in the draft BRD, it was not possible to determine whether or not they had been conducted in accordance with GLP guidelines. Furthermore, original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data was the same as the data recorded in laboratory notebooks. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. The Panel was asked whether the results of such studies (all of which are currently included) should be excluded from the performance analyses. The Panel concluded that it was important to note if the data were obtained from studies conducted according to GLP guidelines, as ideally this should be the case. However, the Panel concluded that the data resulting from the retrospective studies that could not be confirmed as GLP-compliant should not be excluded from the performance analysis. Since there was not an indication that the reliability of the data presented for consideration may have been compromised, omitting any data would likely lessen the impact of the analysis. Furthermore, data obtained from peer-reviewed literature or final reports were likely of sufficient quality.

1.2.6 Consideration of All Available Data and Relevant Information

The Panel was asked if all the relevant data identified in published or unpublished studies conducted using the traditional LLNA had been adequately considered in the draft BRD. If not, the Panel was asked what other traditional LLNA data needed to be considered and how such data could be obtained. The Panel considered that the draft BRD had taken into account a large majority of the relevant data identified in published and unpublished traditional LLNA studies. The data received as a result of the *Federal Register (FR)* notices and the key literature citations seemed to be inclusive of the relevant data for this analysis. Although additional data that could have been included might exist, it was deemed unlikely that the current outcome (which is based on 466 substances) would be altered given the very small change in accuracy statistics relative to Kimber et al. (2006), which was based on 211 substances.

1.3 Comments on the Draft ICCVAM Test Method Recommendations on the LLNA Limit Dose Procedure

1.3.1 Test Method Usefulness and Limitations

The Panel was asked to comment on whether the available data supported the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed test method usefulness and limitations (i.e., that the LLNA limit dose procedure should be routinely recommended for hazard identification when dose response information is not required). The Panel considered that, based on the available information, the draft recommendations appeared valid, but made the following suggestions:

- Further emphasis should be given to using the LLNA limit dose procedure as a part of a comprehensive weight-of-evidence evaluation of dermal sensitizing potential (e.g., including physical chemical evaluation, SAR information, including likelihood of dermal penetration, ability of materials to adduct biomacromolecules).
- Such information in addition to LLNA results might also be useful in confirming or questioning LLNA outcomes terms of in human hazard identification, since it should be emphasized that a major application of the method is to prospectively detect harmful chemicals.
- Solubility or thermodynamic activity data, beyond visual assessment (e.g., use of chemically-specific methods to document solubility), should be used to confirm the appropriateness of the maximum dose tested.
- Vehicle selection for the LLNA can affect the results and may not allow accurate comparisons between chemicals applied in different vehicles. In choosing the best vehicle, consider measured solubility information for the potential vehicle. Then, it would be important to take into account how the vehicle affects the amount of the chemical that can be applied to the ear. More importantly, the impact that vehicle selection has on the amount of applied chemical that actually gets into the mouse to induce the sensitization response should be evaluated. Some of the recommended LLNA vehicles (e.g., 4:1 acetone:olive oil (AOO), dimethylsulfoxide (DMSO), methyl ethyl ketone) could be expected to disrupt the barrier properties of the skin. Additionally, although propylene glycol might allow an increased amount of chemical to be applied, it might also inhibit the penetration of a chemical by enhancing partitioning in the vehicle relative to the skin.

The Panel was asked whether the LLNA limit dose procedure should be routinely recommended for the hazard identification of skin sensitizing chemicals when dose response information *is not* required. With the points noted above in mind, the Panel agreed that it should be routinely recommended since the LLNA limit dose procedure offers time, cost, throughput, and logistical benefits as well as using fewer animals. Still, the investigator should keep in mind what is known of the chemical regarding general toxicity and note scenarios where abnormal dose-response relationships in the traditional LLNA might result in false negatives in the limit dose procedure (see Table 6-2 of the draft ICCVAM BRD).

The Panel was then asked whether the LLNA limit dose procedure should be routinely recommended as the initial test to identify sensitizers before conducting the traditional LLNA, as a way to further reduce animal use, even if dose response information *is* required, since negative results would not require further testing. The Panel agreed that use of the LLNA limit dose procedure, as the initial testing procedure to identify sensitizers and non-sensitizers before conducting the traditional LLNA, is justifiable even when dose response information is required. This is applicable in the occupational and public health setting where obtaining hazard information is of critical importance. There is a benefit since dose-response information generated in subsequent testing in the traditional LLNA for substances that were positive in the limit dose procedure then gives further assurance of detecting hazardous substances and allowing a potency estimate. The benefits of screening out the negatives (which do not require dose response information) is clear; however the animal welfare gains will depend on the proportion of test substances in any class that turn out to be non-sensitizers and there might be possible consequences of the delays resulting from a further round of testing for those materials that are identified as sensitizers.

Based on the existing database, there is a false negative rate of 1.6% (5/313 positive compounds) for the LLNA limit dose approach compared to the results obtained in the traditional LLNA. The Panel was asked whether they considered that this is adequately addressed by the proposed cautionary language and weight-of-evidence consideration for negative substances. The Panel agreed that the small rate of false negatives was adequately addressed in the draft test method recommendations by giving cautionary and weight-of-evidence consideration to the negative substances (and any possible false positive results). Furthermore, given that the dose responses for these five materials were rather unusual, it was not known whether these studies were repeatable, whether any procedural problems were reported with these studies, or what values were returned by the negative/positive control groups (in relation to other historical positive control values). In general, the Panel viewed that the false negative rate of 1.6% would likely be unimportant when the larger differences between the animal model and humans are considered.

1.3.2 Test Method Protocol

The Panel was asked whether they agreed that the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed standardized test method protocol. The Panel agreed and recommended adherence to the ICCVAM (1999) LLNA protocol for future studies of the LLNA limit dose procedure with the exception of omitting the middle and low dose groups. Similarly, application of the LLNA limit dose procedure to a modified LLNA protocol would require adherence to the modified LLNA protocol with the exception of omitting the middle and low dose groups. Adhering to the use of individual animals for future studies was specifically stressed because it would allow for an estimate of inter-animal variability.

The recommended ICCVAM protocol (ICCVAM 1999; Dean et al. 2001; EPA 2003), as well as OECD TG 429 (OECD 2002), specifies that the highest dose tested should be the highest soluble concentration that does not induce systemic toxicity and/or excessive skin irritation.

However, Kimber et al. (2006) concluded that negative results obtained from studies where the highest concentration tested was below 10% should be considered invalid, and adopted a 10% application concentration as a threshold of confidence for categorization of a chemical as being negative while noting that the figure should not be considered as inviolable. The Panel was asked whether the data presented in the draft BRD (i.e., 51/313 positive substances in the NICEATM database were negative at concentrations equal to or above 10%, but were positive at even higher concentrations) were adequate to conclude that this threshold concentration is not appropriate. The Panel viewed that this point should be clarified. ICCVAM recommended that no threshold should be used to determine the validity of conduct of the LLNA limit dose procedure. Instead, formal attempts to maximize dose delivery including documentation of solubility of the test substance in the vehicle used should be undertaken.

The Panel was asked whether additional testing should be required if a negative result was obtained for a test substance in a study where the highest concentration that could be tested (based on systemic toxicity or excessive local irritation, as described in ICCVAM [1999], Dean et al. [2001], and EPA [2003]) was <10%. The Panel considered that, if a negative result was obtained for a test substance under these conditions, additional testing should not be required, because at that point it would likely be a toxic effect and not sensitization. In contrast, the imperative should be to minimize the number of false negatives. For this purpose, rigorous examination of maximum solubility or other parameters to ensure testing at maximum concentration should be employed. In addition, weight-of-evidence considerations such as SAR and physicochemical characteristics should be documented. More animal testing to verify negative results should only be undertaken if the weight-of-evidence suggests that it would be appropriate.

The Panel was asked if the current approach for selecting the "limit dose" was appropriate or whether there is a threshold concentration for the LLNA at which a negative result could always be considered as an acceptable result. The Panel agreed that the current recommendation to select a maximum applied dose in the LLNA limit dose procedure is appropriate. However, the data presented in the draft BRD implied that at present it is not possible to establish a uniform concentration threshold for the "limit dose." Thus, it seemed justifiable that preliminary experimentation (as would be typically performed during a dose range finding study) should be conducted on vehicle selection, test substance solubility, and stability in the vehicle.

1.3.3 Future Studies

The Panel was asked if they agreed that the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed future studies. Although limited in scope, the Panel considered that the available data supported the ICCVAM draft recommendations for additional studies. The Panel agreed that attempts be made to investigate if maximum solubility was achieved (e.g., use of chemically-specific methods to document solubility). For hazard assessment, it was considered troublesome that there were so many vehicle choices, because the vehicle could have a significant effect on whether (and how much) a test substance penetrated the skin barrier. Observed vehicle effects may relate to dermal penetration as well as to immunomodulation. The Panel considered it desirable to follow

the hierarchy of vehicles recommended in the ICCVAM (ICCVAM 1999; Dean et al. 2001) protocol. In addition, dedicated attempts must be made to investigate solubility in AOO mixtures before using other vehicles. Regardless of the vehicle used, it is important to ensure that a vehicle does not promote lymph node cell proliferation. The Panel also suggested that it might be informative to test both known mild and severe sensitizers concurrently in all recommended vehicles to evaluate if a specific vehicle choice(s) might influence the results.

Although the false negative rate in the current analysis was small, a need exists to better understand factors that could lead to false negative results with future use of the LLNA limit dose procedure. Thus, consideration should also be given to formal statistical assessments to verify group size and use of individual animal data in routine performance of the LLNA limit dose procedure. Criteria should be established to verify proficiency with the LLNA limit dose procedure. Such criteria could be used to answer questions about the necessity to perform concurrent positive controls.

1.3.4 Comments with Specific References to the Text

The Panel also identified the following comments and/or corrections to the draft ICCVAM test method recommendations document on the LLNA limit dose procedure that should be considered by ICCVAM:

- Lines 26-28: Conclusions given here regarding the relative potency ratings of the five materials classified as false negative in the analysis in Section 6.2 of the draft BRD were newly introduced. This assessment should also have been considered for inclusion in Section 6.2 of the draft BRD.
- Line 28: The citation of Gerberick et al. (2004) was not accompanied by a reference.
- Lines 62-70: The listing of substances not amenable to test in the LLNA could have been expanded to also include agents with anticipated pharmacodynamic action as immune suppressants.
- Line 69: The citation of Gaspari et al. (2007) was not accompanied by a reference.
- Lines 75-79: Dependent upon other considerations, this portion of the text could have been modified to (1) clarify recommendations regarding routine use of concurrent positive control (i.e., possible exception for laboratories conducting a high volume of LLNA work in which periodic positive control for quality control purposes might suffice), and (2) expand on the details regarding indications of excessive irritation and/or systemic toxicity to aid in choice of maximal test dose.

2.0 LLNA for Testing Aqueous Solutions, Metals, And Mixtures

2.1 Comments on the Draft Addendum for Completeness, Errors, and Omissions

In regard to the draft Addendum to the traditional LLNA BRD, the Panel was asked to comment on any errors that should be corrected or omissions of relevant data/information that should have been included. The Panel concluded that there were no apparent errors or omissions to the draft Addendum.

2.2 Comments on the Validation Status of the Traditional LLNA for Testing Aqueous Solutions, Metals, and Mixtures

2.2.1 Substances Used for the Validation Studies

The Panel was asked whether the database of substances evaluated was representative of a sufficient range of mixtures, metal compounds, and substances in aqueous solutions that are typically tested for skin-sensitization potential. While there were limited data available on the effects of mixtures, metals, and aqueous solution on skin-sensitization potential, the Panel considered the database to be generally representative. The Panel indicated that there did not seem to be obvious classes of chemicals missing from the data set used to evaluate the utility of the traditional LLNA for testing aqueous solutions. However, quantitative compositions for the mixtures included in the analysis had not been provided. Thus, it was difficult to determine if those mixtures were representative of the types of mixtures typically tested in the traditional LLNA. With respect to metals, there was a paucity of commercially useful metals such as platinum, palladium, iron, zinc, manganese, and silver compounds. To enlarge the group of metal non-sensitizers, substances used as cosmetic ingredients (e.g., titanium dioxide) and aluminum compounds currently used in antiperspirants might be considered. However, the Panel considered that the inclusion of an array of other metals and at least one zinc and manganese salt likely weighted the data set appropriately and it appeared sufficiently broad to support conclusions about the utility of the traditional LLNA for testing the skin-sensitization potential of metals.

Substances or mixtures that were tested in an aqueous or an organic:aqueous vehicle were labeled as aqueous solutions. For the purpose of this evaluation, a substance or mixture containing at least 20% water was defined as an aqueous solution. The Panel was asked whether this criterion was appropriate for defining an aqueous solution. The Panel was uncertain about the appropriateness of this definition of an aqueous solution, but did not offer an alternative definition. However, the Panel indicated that an organic:aqueous solution that is not miscible would likely produce varying results because of partitioning of the chemical into either phase.

The Panel was asked whether the lack of coding of test substances might adversely impact or bias the current evaluation. While coding of chemicals is recommended for prospective validation studies, the retrospective evaluations in the draft Addendum were based on existing data, most of which were not generated using coded chemicals. However, the Panel agreed that the lack of chemical coding was not likely to bias the evaluation since this study was retrospective. This is supported by the fact that many of the chemical structures included in the

analyses did not contain known structural motifs associated with allergic contact dermatitis/chemical hypersensitivity and therefore there was no a priori expectation that the chemical tested would be a sensitizer. Furthermore, many of the substances tested were apparently evaluated for hazard assessment purposes rather than to test the predictive ability of the traditional LLNA. Thus, there does not appear to be any bias in chemical selection for the expanded dataset considered in the study of applicability domain for the traditional LLNA.

For some substances submitted using the traditional LLNA test method, it was not possible to confirm whether the data were generated based on pooled lymph nodes among animals within a dose group, as allowed in OECD TG 429 (OECD 2002), or individual animal responses, as recommended by ICCVAM (1999) and required by EPA (2003). ICCVAM (1999) and EPA (2003) both recommend the use of statistical analyses to aid in the interpretation of traditional LLNA study results; such analyses necessitate data collected from individual animals. Additionally, Cockshott et al. (2006) reported that using individual animal data allowed for outlier animal results within a dose group to be identified. The Panel was asked whether the analysis of the performance of the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions should be limited to data from studies that collected individual animal data, and then to comment on the potential impact on the accuracy analysis of including results from studies in which pooled animal data were collected. The Panel concluded that, although individual animal data were preferred, pooled animal data should not be excluded automatically from this retrospective analysis.

2.2.2 Test Method Accuracy

The Panel was asked whether the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions had been adequately evaluated and compared to the human and guinea pig test results. The Panel agreed that the comparative assessment of the relevance of the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions appeared to be as comprehensive as was feasible. However, because of the limited number of comparisons available, the accuracy statistics probably do not give a complete picture of the usefulness and limitations of the traditional LLNA for identifying skin sensitizers among these types of substances or when using an aqueous vehicle.

When multiple traditional LLNA studies were available for the same substance, the "majority call" (among studies using the same vehicle and generally tested over the same concentration range) was used by ICCVAM to assign an overall classification for the purposes of the accuracy analysis. For example, if chemical X was tested five times and was positive in three studies and negative in two, the overall classification was positive. The Panel was asked whether they agreed with this approach. They expressed their concern about the approach in the following way; if all nickel-containing compounds in the analysis were viewed as a group, there were four positive calls and four negative calls (see Appendix C2 of the draft Addendum). Using the "majority call" approach, the overall call would be determined by the next available study, which may not provide the correct call. More data would be needed to confirm whether the classification was appropriate. For this dataset, most of the "negative calls" had SI values that

approached 3.0. Thus, a more suitable method might be to base the overall call on the SI data, while giving greater positive call consideration/weight to SI values just below 3.0. It may also be useful to perform a meta-analysis. It is important for the Addendum to mention the potential impact of using the "majority call" decision, rather than relying on a weight-of-evidence approach, on the accuracy analyses.

2.2.3 Data Quality

For some studies included in the draft Addendum, it was not possible to determine whether or not they had been conducted in accordance with GLP guidelines. Original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data was the same as the data recorded in laboratory notebooks. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. The Panel was asked to discuss what impact this lack of information might have on the evaluation of the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions and whether such studies should be excluded from an analysis of test method accuracy. The Panel considered it important to note if the data were obtained from studies conducted according to international GLP guidelines, since ideally this should be the process followed. However, the Panel viewed that data from studies that could not be confirmed as being GLP-compliant were still appropriate to include in the accuracy analysis, provided that the data were from the peer-reviewed literature or from sources with high quality laboratory management practices. Much of the value for this draft Addendum was the potential to supplement the data available at the time of the ICCVAM (1999) analysis. Additional information on test substance identification would clearly be useful in the continued evaluation of the applicability domain of the traditional LLNA, but omitting data on mixtures, metals, or use of aqueous solutions based solely on the lack of GLP compliance would lessen the impact of the current retrospective analysis and did not seem warranted. However, if the original data were not available, it would be appropriate to note this in the final version of the Addendum.

2.2.4 Consideration of All Available Data and Relevant Information

The Panel was asked whether the draft Addendum included all of the relevant data for studies conducted using the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions. The Panel considered that, although it was possible that there might be a few studies in the literature to augment the analysis, it seemed that the relevant data had been identified and the response to the *FR* notice and the literature citations examined had included the most relevant studies.

2.3 Comments on the Draft ICCVAM Test Method Recommendations on the Traditional LLNA for Testing Aqueous Solutions, Metals, and Mixtures

2.3.1 Test Method Usefulness and Limitations

ICCVAM stated that more data would be needed before a recommendation on the usefulness and limitations of the traditional LLNA for testing mixtures could be made, due to the limitations associated with the available mixtures database (i.e., unknown formulae, lack of

human data). The Panel was asked whether they agreed that the available data supported the ICCVAM draft recommendations for the traditional LLNA with regard to testing mixtures in terms of the proposed test method usefulness and limitations. The Panel agreed that ICCVAM's draft recommendation with respect to the traditional LLNA testing of mixtures appeared valid based on the limitations inherent in the available data set. Still, the Panel urged that the ICCVAM recommendation indicate that the approach may be viable. The Panel further recommended that the test method recommendations summary should indicate that the limitations include relatively poor concordance of traditional LLNA outcomes for mixtures with to those obtained in guinea pig tests. Routine comparisons of accuracy according to classification criteria may not be sufficient to evaluate the concordance for mixtures, and furthermore, the guinea pig tests are not necessarily valid for mixtures. The Panel also indicated that the term "mixtures" was used too broadly (i.e., can represent an infinite number of materials) and it would be more beneficial to specify types or formulations of mixtures that are being examined.

ICCVAM recommended that, based on the available data for metals, the traditional LLNA was useful for the testing of metal compounds, with the exception of nickel. The Panel was asked whether they agreed that the available data supported the ICCVAM draft recommendations for the traditional LLNA with regard to testing metals in terms of the proposed test method usefulness and limitations. Based on the available information, the Panel agreed that the draft recommendations with regard to testing metals appeared to be valid. In particular, the evidence for most metals (e.g., accuracy of 86% (12/14), sensitivity of 100% (9/9), specificity of 60% (3/5), 0% (0/9) false negatives) when comparing traditional LLNA results to those obtained from evaluations in humans supported the use of the traditional LLNA as a hazard identification tool for metals, excluding nickel. However, the Panel recommended that it would be worthwhile to study further the variable results obtained for nickel since there is a wealth of literature on allergic contact dermatitis of nickel in humans.

In a minority opinion, Dr. Dagmar Jírová stated that it should not be concluded that the traditional LLNA was not suitable for testing nickel compounds, because the different vehicles used may have had a significant impact on the ability of nickel to penetrate the skin and be bioavailable. She noted that nickel chloride and nickel sulfate were both positive in aqueous solutions, and negative only when non-aqueous vehicles were used. In human exposures, nickel compounds were applied in aqueous solutions. Thus, this may serve as sufficient justification to use aqueous vehicles when nickel, and perhaps also other substances, are tested and evaluated in the traditional LLNA. When DMSO was used as the vehicle, the SI value increased with increasing nickel concentration. Unfortunately, no data were available for concentrations over 5% for either nickel compound in DMSO. Nickel chloride as 10% in aqueous solution reached an SI of 6.6. Inconsistent test results due to the vehicle have also occurred in other *in vitro* studies (e.g., phototoxicity). Thus, Dr. Jírová concluded that the traditional LLNA could be used even for testing nickel compounds when other vehicles (in particular aqueous) are used.

Due to the limited number of substances tested in aqueous solutions, ICCVAM recommended that more data would be needed before a recommendation on the usefulness and limitations of

the traditional LLNA for testing substances in aqueous solutions could be made. The Panel was asked whether they agreed that the available data supported this ICCVAM draft recommendation for the traditional LLNA with regard to the testing of substances in aqueous solutions. The Panel agreed that the draft ICCVAM recommendation was appropriate and that more data were required before an adequate evaluation of the use of the traditional LLNA with aqueous solutions could be conducted.

2.3.2 Test Method Protocol

The Panel was asked whether they agreed that the available data supported the ICCVAM draft recommendations for the traditional LLNA in terms of the proposed test method standardized protocol. The Panel agreed that, in general, the results of the assessment in the draft Addendum supported the proposals for standardized conduct of the traditional LLNA. However, this conclusion depended on a side-by-side reading of the draft Addendum and the ICCVAM (1999) protocol. The Panel suggested expanding the brief section of the draft test method recommendations dealing with test method protocol for the traditional LLNA (Section 2.0) to specifically point out how the conclusions of the applicability domain evaluation may affect the standard traditional LLNA protocol. For example, the evaluation of aqueous solutions apparently resulted in the methodological recommendation that aqueous test solutions be avoided and the further recommendation of a hierarchy of organic solvents to be considered as dosing vehicles. The emphasis might be on using a vehicle to which humans may actually be exposed.

2.3.3 Future Studies

The Panel was asked whether they agreed that the available data support the ICCVAM draft recommendations for the traditional LLNA in terms of the proposed future studies. The Panel agreed that the ICCVAM recommendation for continued accrual of information from traditional LLNA evaluations of mixtures, metals, and aqueous solutions with comparative data for guinea pig and human tests was appropriate. The traditional LLNA accuracy for metals of 86% and sensitivity of 100% (0% false negative) was excellent; a specificity of 60% (40% false positive) was considered acceptable as over-classification maintains safe human use. The Panel encouraged the use of the traditional LLNA to acquire further information on mixtures, metals, and aqueous solutions. However, the Panel suggested that, given resource limitations, it would be important to prioritize the recommendations in order to focus on what is most important.

[This Page Intentionally Left Blank]

3.0 Non-Radioactive LLNA Protocol - The LLNA: Daicel Adenosine Triphosphate (LLNA: DA) Test Method

3.1 Comments on the Draft BRD for Completeness, Errors, and Omissions

3.1.1 General Comments

The Panel was asked if there were any errors in the draft LLNA: DA BRD that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. As a general comment, the Panel noted that the draft BRD clearly and succinctly provided an overview of the LLNA: DA test method and the relevant validation study data. The draft BRD indicated that the LLNA: DA differs from the traditional LLNA in the method of measuring proliferation (measures levels of adenosine triphosphate [ATP] instead of radioactivity), substance treatment (pretreating the test site with 1% sodium lauryl sulfate [SLS] prior to test substance application and an additional treatment on day 7), and sampling time (draining auricular lymph nodes are collected on day 8 rather than on day 6). Because the traditional LLNA evaluates the induction phase only, the relevance of results with the LLNA: DA (and any other LLNA protocol) should always be considered in the context of human experimental sensitization data, human epidemiologic data, and elicitation in the clinical setting.

3.1.2 Comments with Specific References to the Text

The Panel noted the following text that should be clarified or corrected in the final version of the LLNA: DA BRD:

- Line 428: The text should read “1% SLS,” not “1% SDS”; the same terminology should be used throughout rather than going back and forth between SLS and sodium dodecyl sulfate (SDS).
- Line 449: The text and formula in lines 448-450 appear misplaced and instead seem to belong in Section 7.0. Additionally, X and Y should be defined and the “Var (ln SI)” formula should be clarified.
- Table 3-1: There were 33 substances in the table, yet the discussion of the table in the text referred to 31 substances. Although the reason for this apparent discrepancy becomes more evident later on, this should be discussed up front.
- Table 3-2 and 3-3: The interlaboratory distribution and testing of the sensitizers versus non-sensitizers should be indicated here.
- Table 6-1: Although the table clearly provided a comparison of the different methods, it would be useful if the footer for this table also indicated the basis for the differences in substances included in each analysis (i.e., n=25, 26, or 29) as stated in the text.
- Table 7-1 was only moderately helpful because the standard deviations (SD) were not calculated on a log scale. Given the skewness in the data, the ranges given were misleading indicators of increases due to the vehicles. Without the samples sizes, an

analysis of variance was impossible to calculate, and that would have been the appropriate measure of differences between the experiments. In order to make recommendations regarding needed reproducibility experiments, it would have been helpful to have a power analysis for this situation. That is, for each vehicle with each chemical, using two, three, four and five animals per dose group per experiment, how many experiments need to be run to detect significant differences between the experiments?

3.2 Comments on the Validation Status of the LLNA: DA

3.2.1 Test Method Protocol

Based on its 1998 evaluation of the traditional LLNA procedure, ICCVAM recommended that at least five animals be used in each test group (ICCVAM 1999). The LLNA: DA validation studies presented in the draft BRD were performed using four animals per dose group. Thus, the Panel was asked to comment on the potential impact of using fewer than five mice per dose group. The Panel noted that supplemental statistical information they were provided with indicated that the power for detecting a three-fold increase in the SI value in the treatment group for the LLNA: DA dataset evaluated in the draft BRD was estimated to be 95% for a sample size of three mice per dose group (see **Table F-2**). Since an increase of false negatives may not be an issue, the potential opportunity exists for utilizing this smaller group size. The Panel cautioned, however, that using less than five animals per group might result in a less precise estimate of the mean response, which, in turn, will impact accuracy. Also, if technical errors further reduce the sample size, accuracy is further reduced. Thus, the Panel recommended that all initial validation studies adhere to the ICCVAM-recommended protocol (ICCVAM 1999; Dean et al. 2001) of five animals per dose group until sufficient information is generated to indicate that the use of fewer animals per dose group is statistically valid.

Table F-2 Power Calculations for the LLNA: DA[1]

Parameter	3.0-fold Increase[2]	2.5-fold Increase	2.0-fold Increase	1.5-fold Increase	1.3-fold Increase
Mean Rx response	8835	7362.5	5890	4417.5	3828.5
Log (mean Rx response)	9.086	8.904	8.681	8.393	8.250
Difference from control (log scale)	1.098	0.916	0.693	0.405	0.262
Difference/SD	3.95	3.29	2.49	1.46	0.94
Power for N=5	99%	99%	95%	50-80%	<50%
Power for N=4	99%	95-99%	90%	50%	<50%
Power for N=3	95%	90-95%	80%	<50%	<50%
Other power	–	–	–	95% (N=11)	95% (N=25)
Other power	–	–	–	90% (N=9)	90% (N=20)

Abbreviations: N = number of animals; Rx = Treatment; SD = standard deviation.

[1] The power calculations above are based on a one-sided $p<0.05$ Student's t test applied to log-transformed data from vehicle control LLNA:DA tests.

[2] Fold-increase = Required increase above the vehicle control for a positive response

The data generated for the substances analyzed in the LLNA: DA interlaboratory validation studies came from auricular lymph nodes that were pooled across animals in each treatment group. The Panel was asked to comment on the potential impact of including pooled animal data on the accuracy analysis of the LLNA: DA. The Panel noted that a statistical analysis of differences between treatment-related and vehicle control ATP levels could not be determined without measures of variability. Individual animal data highlights technical issues and allows for consideration of dose-response information and statistical analyses.

The LLNA: DA differs from the traditional LLNA in the treatment schedule and by including a pretreatment step with 1% SLS just prior to application of the test substance. The Panel was asked to comment on the appropriateness of these protocol differences. The Panel did not consider these differences to be significant, as long as it could be demonstrated that the 1% SLS pretreatment step and the additional test substance treatment on day 7 did not induce a skin reaction indicative of the elicitation phase of skin sensitization. Although it was being used at a lower concentration than the estimated concentration needed to produce a stimulation index of 3 (EC3), the Panel expressed concern about pretreating the mouse ear with 1% SLS since SLS is an irritant and positive in the traditional LLNA. Consequently, the inherent sensitivity of the LLNA may be modified by the 1% SLS pretreatment step. To demonstrate that these concerns are not justified, the Panel concluded that weak irritants and weak sensitizers needed to be tested in the LLNA: DA assay with and without pretreatment with 1% SLS. The test method developer might also consider using decision criteria other than SI ≥3.0 such that 1% SLS pretreatment is no longer necessary.

3.2.2 Substances Used for the Validation Studies

The Panel was asked if they considered the substances tested in the LLNA: DA to be representative of a sufficient range of chemical classes and physicochemical properties that the test method would be applicable to any of the types of chemicals and products that are typically tested for skin-sensitization potential. The Panel considered the database of test substances tested in the LLNA: DA representative of a sufficient range of chemicals. The selected substances included solids and liquids and a range of solvents/vehicles. The database also represented a range of sensitizing potency, a variety of different chemical classes and substances with differing requirements for metabolic activation. However, it might have been useful to have also included substances with clearly different protein reaction mechanisms (protein binding), as well as dyes, natural extracts, and mixtures.

3.2.3 Test Method Accuracy

The accuracy analysis in the draft LLNA: DA BRD was based on overall concordance with the traditional LLNA. Accuracy statistics compared to the guinea pig tests and human data/experience were also provided. The Panel, when asked if they considered these comparisons appropriate for assessing the accuracy of the LLNA: DA, agreed that the comparisons to the traditional LLNA performance and also to the guinea pig and human sensitization data were important. The Panel also stressed that, because the traditional LLNA only evaluates the induction phase, the relevance of the LLNA: DA results should always be considered in the context of human experimental sensitization data, human epidemiologic data, and elicitation in the clinical setting.

The Panel was asked if they considered the evaluation of the relevance of the LLNA: DA and the comparison to the traditional LLNA to be adequate. The Panel noted that Table 6-1 of the draft LLNA: DA BRD clearly provided a comparison of the different reference methods (i.e., traditional LLNA, human tests, and guinea pig tests). Thus, the Panel concluded that the relevance of the LLNA: DA had been adequately evaluated. However, including data on more substances is likely to further strengthen confidence in the concordance data.

One substance, 2-mercaptobenzothiazole, produced a false negative response compared to the traditional LLNA when tested using the LLNA: DA. The Panel was asked if they could identify any characteristics associated with this or similar substances, compared to the correctly identified sensitizers, that might signal that this type of discordant response would occur, and therefore using the LLNA: DA to test such substances would not be appropriate (or that negative results for substances with such properties may warrant additional testing). The Panel could not identify specific characteristics that might explain the false negative response using the LLNA: DA. Although understanding the solubility and stability of the test substance in different vehicles is important, the differences in response did not seem to be explained by the vehicle differences (AOO and dimethylformamide [DMF]) between the two tests. In addition, the impact of 1% SLS pretreatment on the negative response in the LLNA: DA is not known but should be considered.

One substance, benzalkonium chloride, produced a false positive response compared to the traditional LLNA and guinea pig test when tested using the LLNA: DA. The Panel was asked if they could identify any characteristics associated with this or similar substances, compared to the correctly identified non-sensitizers that might signal that this type of discordant response would occur, and therefore using the LLNA: DA to test such substances would not be appropriate (or that positive results for substances with such properties may warrant additional testing). The Panel could not identify specific characteristics that might explain the false positive response for this substance in the LLNA: DA. The Panel viewed that it was important to note, however, that this chemical is a well-known skin irritant, and on occasion it had also been considered a human sensitizer, typically on the basis of positive diagnostic patch test data. Thus, the Panel reiterated that the relevance of LLNA: DA results should always be considered in the context of human experimental sensitization data, human epidemiologic data, and elicitation in the clinical setting. The actual impact of the 1% SLS pretreatment step on the LLNA: DA has not been well established, although van Och et al. (2000) and De Jong et al. (2002) have reported that 1% SLS pretreatment enhances the response in the traditional LLNA.

3.2.4 Test Method Reliability

The Panel was asked if they considered the intralaboratory reproducibility of the LLNA: DA to have been adequately evaluated and compared to the traditional LLNA (refer to Table 7-1 of the draft LLNA: DA BRD). The Panel noted that only eugenol and isoeugenol, two sensitizers with similar chemical structures, were tested. The Panel recommended testing a positive control commonly used in the traditional LLNA (e.g., HCA) for a more complete evaluation of intralaboratory reproducibility. In addition, it was unclear if the tests were truly independent. Factors that might indicate independence should have been documented (e.g., time interval between experiments, different animal shipment, different reagents, different operator).

The Panel was also asked if they considered the interlaboratory reproducibility of the LLNA: DA to have been adequately evaluated and compared to the traditional LLNA. The Panel noted that the interlaboratory reproducibility of the assay could not be adequately evaluated given the lack of original study data and limitations in the study design. Study design limitations included:

- Pooled lymph nodes were used from mice within a dose group. This precluded an analysis of variation between laboratories.
- The lead laboratory established the dose levels to be tested by the other laboratories participating in the interlaboratory validation effort. In a minority opinion, Drs. Nathalie Alépée and Michael Woolhiser asserted that for an effective and efficient interlaboratory evaluation, it seemed reasonable to set dose levels for all laboratories based on results from the lead laboratory.

In addition, the Panel considered that the interlaboratory studies could benefit by performing more than one test on two commonly used positive controls (i.e., HCA and DNCB).

The draft LLNA: DA BRD contained an analysis of data from two interlaboratory reproducibility validation studies that used coded substances, as well as an intralaboratory accuracy validation study with 31 substances that were not coded. The Panel was asked if they

considered the lack of coding of the test substances to have adversely impacted or biased the intralaboratory accuracy evaluation. The Panel commented that, in the validation of a new assay, it is better to avoid the potential for bias by testing coded substances. However, the Panel concluded the data already generated for the LLNA: DA test method should be considered and not be rejected in the current validation evaluation.

The lead laboratory established the dose levels tested by the participating laboratories in the two interlaboratory validation studies. The Panel was asked if this adversely impacted or biased the evaluation. The Panel considered that the choice of the maximum test substance concentration is crucial for the proper performance of the traditional LLNA as well as any modified LLNA. Thus, predetermining the dose levels to be tested for each substance might have reduced variability between the two interlaboratory studies. In a minority opinion, Drs. Nathalie Alépée and Michael Woolhiser asserted that for an effective and efficient interlaboratory evaluation, it seemed reasonable to set dose levels for all laboratories based on results from the lead laboratory.

3.2.5 Data Quality

The studies evaluated in the draft BRD for the LLNA: DA were not conducted in accordance with GLP guidelines although they were reportedly done in laboratories that conduct GLP studies, and were conducted "in the spirit" of GLP (K. Idehara, personal communication). Furthermore, the original records for the interlaboratory studies were requested but have not yet been obtained. As a result, an independent audit could not be conducted to confirm that the reported data was the same as the data recorded in laboratory notebooks. The Panel was asked to comment on the potential impact this might have had on the evaluation of the LLNA: DA. The Panel commented that, ideally, GLP compliance is recommended for validation studies, but the current studies should not be rejected based on the lack of GLP compliance alone. However, all the raw data obtained through the validation process should be made available and audited for accuracy. The Panel further commented that since the original records for the interlaboratory studies have not yet been provided, recommendations from ICCVAM should be contingent upon receiving these data. Obtaining original laboratory records is a necessary step to confirm that all data generated during the validation studies have been provided, and that the reported data are the same as the data recorded in laboratory notebooks.

3.2.6 Consideration of All Available Data and Relevant Information

The Panel was asked to comment on whether all of the relevant data identified in published or unpublished studies that employed the LLNA: DA had been adequately compared. The Panel viewed that, generally, it seemed that all of the relevant results had been adequately identified and considered. However, as mentioned above, all of the original data supporting these results have not been provided. The Panel again expressed concern related to the effect of pretreating the mouse ear with 1% SLS and the Panel therefore recommended that the results from van Och et al. (2000) and De Jong et al. (2002) should be considered.

3.3 Comments on the Draft ICCVAM Test Method Recommendations on the LLNA: DA

3.3.1 Test Method Usefulness and Limitations

The Panel was asked to comment on whether the available data supported the ICCVAM draft recommendations for the LLNA: DA procedure in terms of the proposed test method usefulness and limitations. The Panel agreed with ICCVAM's recommendation, which stated that the LLNA: DA might be useful for identifying substances as potential skin sensitizers and non-sensitizers, but this recommendation was contingent upon the receipt of additional data and information. The Panel further added that information on the possibility of skin reactions suggestive of the onset of the elicitation phase and the impact of the 1% SLS pretreatment step on the performance of the LLNA: DA should be evaluated. The Panel also considered that the ICCVAM proposed limitations needed to be more clearly defined, as it was not clear from the draft recommendations what points were considered as limitations. For instance, limitations that are known for the traditional LLNA would likely apply to this modified protocol as well and these should be noted.

The Panel was asked whether restrictions on using radioactive materials would warrant that the LLNA: DA be routinely recommended for hazard identification of skin sensitizing substances in lieu of having to possibly use guinea pig tests. The Panel noted that, based on gaps in the currently available dataset and information described in this report, the LLNA: DA could not yet be recommended for the routine use for hazard identification of skin sensitizing substances, regardless of whether restrictions on using radioactive materials were present or not. Generally, non-radioactive LLNA test methods are preferred in lieu of using guinea pig tests because fewer animals are used and animal pain and distress is reduced. However, policy issues regarding restrictions on radioactivity should have no impact on this science-based conclusion.

The Panel was asked if, from a public health perspective, the recommended guidance for evaluating negatives were sufficient to address concerns associated with the false negative rate of 5% (1/19 substances) calculated for the LLNA: DA. The Panel noted that this was not a scientific question, rather a risk characterization issue, and could not be answered without considering other factors such as intended use, target population, etc. The Panel was also asked if, from a testing strategy perspective, the ICCVAM guidance addressed concerns associated with the false positive rate of 10% (1/10 substances) calculated for the LLNA: DA and/or if they had other suggestions for additional guidance or limitations. The Panel again commented that this was not a scientific question but a risk characterization issue and could not be answered without considering other factors such as intended use, target population, etc. Furthermore, the Panel noted that it would be difficult to generalize the finding of one test substance being a "false" result. Instead, they considered it better to identify reasons why a substance was a "false" result. Certainly, if a "false" result is suspected, confirmatory testing with another mouse LLNA method was not recommended. It might be important to follow a suspected inaccuracy with an investigation of the mechanistic basis for the discordance.

3.3.2 Test Method Protocol

The Panel was asked if they agreed that the available data supported the ICCVAM draft recommendations for the LLNA: DA procedure in terms of the proposed test method standardized protocols or what recommendations they would make. The Panel noted that available data did not support all of the ICCVAM draft recommendations in the LLNA: DA standardized protocol. First, the ICCVAM protocol (ICCVAM 1999; Dean et al. 2001) for the traditional LLNA recommends using at least five animals per dose group. Although the Panel agreed that five animals per dose group should be recommended for validation studies, they suggested that power calculations would be useful in determining if subsequent use of the modified test method could use fewer animals per dose group. For the LLNA: DA test method, the Panel noted that based on statistical power calculations that were provided as supplemental information, using four animals per group instead of five did not appear to be a limitation (i.e., detecting a 3.0-fold increase in the SI with four animals per group was estimated to have a 99% confidence level). In addition, the Panel generally agreed with the recommendation in the ICCVAM protocol (ICCVAM 1999; Dean et al. 2001) that individual animal data should be collected. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable.

Of greater importance, the Panel concluded that pretreatment with 1% SLS should not be accepted until its impact on the performance of the LLNA: DA has been adequately characterized. Although used at a concentration below its EC3, the Panel was concerned about pretreating the mouse ear with an irritant reported as positive in the traditional LLNA. To demonstrate that these concerns are not justified, the Panel recommended that substances that are weak irritants and weak sensitizers be tested in the LLNA: DA with and without pretreatment with 1% SLS. It also needed to be demonstrated that the 1% SLS pretreatments, as well as the additional test substance treatment on day 7, did not induce a skin reaction that could be indicative of the onset of the elicitation phase of skin sensitization.

The Panel was asked to comment on whether the limit dose procedure could be applied to the LLNA: DA. The Panel concluded that if the limit dose procedure is considered applicable to the traditional LLNA, then it should also be applicable to the LLNA: DA, in order to reduce the number of animals used. This would require adherence to the LLNA: DA test method protocol, with the exception that the middle and low dose groups would be omitted in the limit dose version.

3.3.3 Future Studies

The Panel was asked if they agreed that the available data support the ICCVAM draft recommendations for the LLNA: DA in terms of the proposed future studies or, if not, what recommendations they would make. The Panel stated that the available data supported the ICCVAM draft recommendations for the LLNA: DA in terms of a more comprehensive evaluation using more non-sensitizers within and across laboratories. In a minority opinion, Dr.

Thomas Gebel stated that although testing of more non-sensitizers might be warranted for interlaboratory validation studies, a sufficient number of non-sensitizers had been tested within the same laboratory (Table 6.3 in the draft BRD).

However, the Panel viewed that there were additional studies that ICCVAM might consider. As previously mentioned, the Panel recommended that the 1% SLS pretreatment step should not be accepted until its impact on the performance of the LLNA: DA had been adequately characterized. Furthermore, it should be demonstrated that such pretreatments did not induce a skin reaction that could be indicative of the onset of the elicitation phase of skin sensitization. It might also be of interest to evaluate ATP as a marker of lymph node proliferation using the traditional LLNA dosing scheme and lymph node collection schedule. Lastly, the Panel considered that studies on the reliability of outlier analysis in small sample sizes and the effects of reduced sample size on the power of the LLNA: DA test method should be proposed.

3.3.4 Performance Standards

The draft LLNA: DA BRD indicated that the LLNA: DA protocol differed from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. In addition, there are differences between the two protocols that relate to how and when the test substance is applied and when the lymph nodes are collected (Table 2-1 and Appendix A in the draft LLNA: DA BRD). According to the proposed draft ICCVAM LLNA Performance Standards for the traditional LLNA, any change to the LLNA protocol other than the method used to assess lymphocyte proliferation would be considered a "major" change. The Panel was asked if they agreed that these should be considered "major" changes and therefore the usefulness and limitations of the LLNA: DA should not be assessed using the draft ICCVAM LLNA Performance Standards. The Panel commented that answering this question depended on having a clear definition of what constitutes a "major" versus a "minor" change, and what may constitute a different protocol altogether. Depending on the goal of the assay, whether a change is "major" versus "minor" may not be relevant. Ultimately, if a test method is able to make the correct prediction with regard to the dermal sensitization potential of a test substance, then the issue of "major" versus "minor" modifications might not apply. Considering the robust nature of the current draft ICCVAM LLNA Performance Standards, it is difficult to identify the need for additional requirements for methods like the LLNA: DA. Thus, the draft ICCVAM LLNA Performance Standards could be used to evaluate the LLNA: DA as a mechanistically and functionally similar test method.

The Panel was asked, even if the draft ICCVAM LLNA Performance Standards were not found applicable to the LLNA: DA, whether an analysis based on 13 of the 18 proposed required reference substances in the performance standards would impact the overall evaluation of the test method accuracy. The Panel commented that the accuracy analysis based on 13 of the 18 proposed required reference substances in the performance standards (with one false negative substance) should have no impact on the overall evaluation of test method accuracy as 31 substances have been tested. However, given the concern regarding pretreatment with 1% SLS, the Panel stated that testing of substances with and without 1% SLS was needed to characterize

the effect of this pretreatment on the performance of the assay. The Panel concluded that as described above, the idea of "major" versus "minor" changes might be reconsidered, thus the current draft ICCVAM LLNA Performance Standards could be applicable to the LLNA: DA as a mechanistically and functionally similar test method.

4.0 Non-Radioactive LLNA Protocol - LLNA: Bromodeoxyuridine Detected by Flow Cytometry (BrdU-FC) Test Method

4.1 Comments on the Draft BRD for Completeness, Errors, and Omissions

4.1.1 General Comments

The Panel was asked if there were any errors in the draft LLNA: BrdU-FC BRD that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. The Panel noted that overall, errors and omissions in the draft LLNA: BrdU-FC BRD were few. The majority of omissions relating to the data records were identified in the text, and all reasonable efforts to obtain additional information from MB Research Labs, the developer of the LLNA: BrdU-FC, appear to have been made.

The following describes the identified errors, omissions, and/or information gaps in the draft LLNA: BrdU-FC BRD that should be addressed:

- Data are available in the peer-reviewed literature on the application of BrdU in the LLNA with histochemical or enzyme-linked immunosorbent assay (ELISA) detection. This could be briefly mentioned in the final BRD for the LLNA: BrdU-FC method, simply as a means of indicating the utility of non-radiolabeled tracer methods in the LLNA.
- It should be noted that a potential reason why nickel chloride was negative in the LLNA: BrdU-FC may be due to oral tolerance in the mice that was induced by nickel-containing nipples of drinking bottles and nickel cages (Van Hoogstraten et al. 1993).
- The vehicle(s) used with the test substances should be stated.
- Information on experience of the inter-laboratory transferability of other technologies that depend upon flow cytometry technology as the key data read-out should be included.
- All raw data for the LLNA: BrdU-FC and the enhanced LLNA (eLLNA): BrdU-FC should be made available.

4.1.2 Comments with Specific References to the Text

The Panel stated that the following comments and/or suggested corrections relevant to specific parts of the draft LLNA: BrdU-FC BRD text should be addressed:

- Line 226: Citation was made to a reference dated 2001 by MB Research Labs which established their development of the LLNA: BrdU-FC; however, no reference was included in Section 12.0 (References) of the draft BRD.
- Lines 232-233: For a sensitizer, the SI should be greater than *or equal to* 3.
- Line 246: “i.e., positive” should be explained.
- Line 254: 11% should be 17%.

- Line 263: For purposes of completeness, it may be worthwhile to add a brief description of the comparative accuracy of the available traditional LLNA, LLNA: BrdU-FC, and eLLNA: BrdU-FC results versus human maximization/patch test data. This information is of importance and displayed in Tables 6-1 and 6-2.
- Lines 286-288: The issue of the refinement/reduction in animal use that might follow the availability of a scientifically validated non-radioactive variant of the LLNA was mentioned in the draft BRD as a benefit but it was not quantified, and no authoritative reference was cited in support.
- Line 288: The final LLNA: BrdU-FC BRD should explain why the BrdU method would result in less pain and distress to the animals (i.e., does the route of injection of BrdU vs. ^{3}H-methyl thymidine produce less discomfort?).
- Line 335: Reference was made to a citation dated 2001 by MB Research Labs which established their development of the LLNA: BrdU-FC; however, no citation was included in Section 12.0 (References) of the draft BRD.
- Line 356: The sentence starting "To evaluate excessive skin…" implies that evaluation of excessive skin irritation by measuring ear thickness is recommended by the ICCVAM LLNA protocol although it is only recommended in the LLNA: BrdU-FC protocol.
- Lines 365-366: Consider supplementing the list of abbreviations for Figure 2-1 with B220+, B:T, CD69+, and IAk+. Also, the figure shows I-Ak+ while all other text uses IAk+. The MB Research Labs protocol shows I-Ak+.
- Figure 2-1: Should be redrawn to show the SI decision point lines coming off of the "Analyze Proliferating LNC (lymph node cells)" box rather than the "Inject BrdU and Excise…" box.
- Lines 500-503: Classification of "equivocal results" was unclear without data comparison (i.e., benzocaine produced divergent results in both tests). Were these results unlike what was expected from human data? What were the data for salicylic acid and mercaptobenzothiazole?
- Lines 552-558: It may be useful for comparative purposes to add summary accuracy data for the traditional LLNA versus human maximization/patch test data from the larger data set reported in the 1999 ICCVAM LLNA report to the section of the final BRD which discusses performance of the LLNA: BrdU-FC method.
- Table 6-3: Benzocaine was missing and it seems that salicylic acid was the same in both traditional LLNA and LLNA: BrdU-FC. Mercaptobenzothiazole was not reported for the LLNA: BrdU-FC. This needs to be corrected or explained.
- Table 6-5: The human outcome for benzalkonium chloride and ethylene glycol methacrylate should be negative. See also Table 6-6.

4.2 Comments on the Validation Status of the LLNA: BrdU-FC

4.2.1 Test Method Protocol

The LLNA: BrdU-FC protocol includes routine measurements of ear swelling as an indicator of excessive dermal irritation. The Panel was asked if they considered this procedure to be an appropriate approach and if this measurement should be recommended for routine inclusion into all LLNA protocols. The Panel stated that, as a quantitative parameter associated with inflammatory cell influx and fluid retention near the site of test substance application, ear swelling (or other quantitative measurements) should be carefully considered for inclusion into all LLNA protocols. This might assist in differentiating between sensitizers and irritants, assist in the interpretation of equivocal results, and possibly detect other procedure-related problems that might require further exploration/consideration.

The LLNA: BrdU-FC protocol also includes optional quantification of immunophenotypic markers as an additional mechanism for distinguishing irritants from sensitizers. The Panel was asked if they considered this to be an appropriate approach to reduce false positives, and if the correct markers were being considered. The Panel was also asked if these measurements should be recommended for routine inclusion in the LLNA: BrdU-FC. The Panel agreed that the use of immunological markers would be appropriate for detailed studies, as it might reduce the frequency of false positives (irritants) and improve comparisons with human data. However, since the primary use of the LLNA is for discrimination of human hazard from direct chemical contact, it could be argued that some false positives are acceptable (especially for methods which have relatively lower rates of false negatives). Given this dominant use, application of immunological markers would likely be too detailed and costly for routine LLNA use. Thus, the Panel suggested that results of ear swelling measurements be compared with the more technically complex flow cytometry markers to determine if similar results might be obtained. Furthermore, alternative immunological markers for discriminating between irritants and sensitizers may be available, although the draft LLNA: BrdU-FC BRD did not contain information allowing any informed decision on whether other markers might be more predictive. Thus, based on current knowledge, the current markers suggested in the draft LLNA: BrdU-FC BRD seemed acceptable to the Panel. Two other possibilities suggested were a surface marker relating to CD4 T-helper cells (Th) or Th1 cells (interferon-γ).

The Panel was also asked to comment on the appropriateness of the "sequential strategy" used in the eLLNA: BrdU-FC (see Figure 2-1 of the draft LLNA: BrdU-FC BRD). Generally, the Panel viewed that incorporation of immune parameters improved the value of a predictive assay and may also help explain mechanisms, which is important. Still, the "sequential strategy" used in the eLLNA: BrdU-FC for discriminating irritation from sensitization might be more sensible for research studies because of resource and cost considerations, and may not be appropriate for routine use of the LLNA in hazard identification. For human hazard detection, more simplified methods should be available for discrimination of irritants.

4.2.2 Substances Used for the Validation Studies

The Panel was asked to consider if the substances tested in the LLNA: BrdU-FC were representative of a sufficient range of chemical classes and physicochemical properties such that the test method would be applicable to any of the types of chemicals and products that are typically tested for skin-sensitization potential. The Panel agreed that if the proviso that the applicability domain limitations published for the traditional LLNA remained in force, the substances tested in the LLNA: BrdU-FC seemed representative of a sufficient range of chemical classes and physical chemical properties and it would likely be applicable to many of the types of chemicals and products that are typically tested for skin-sensitization potential. However, the available LLNA: BrdU-FC database was relatively small compared to the large number of substances assessed in the traditional LLNA and this implied some caution in assuming that assay performance was concordant with the traditional LLNA.

4.2.3 Test Method Accuracy

The accuracy analysis presented in the draft LLNA: BrdU-FC BRD was based on overall concordance with the traditional LLNA. Accuracy statistics compared to the guinea pig tests and human data/experience were also provided. The Panel was asked if these comparisons were appropriate for assessing the accuracy of the LLNA: BrdU-FC. The Panel viewed that since the traditional LLNA is used to provide human hazard identification and information relevant to human health, the accuracy statistics compared to human data/experience were important. Since the LLNA: BrdU-FC is fairly similar to the traditional LLNA, guinea pig comparisons might not have been necessary. However, taken together, the availability of both human data/experience and guinea pig data allowed additional insights that might have expanded the applicability domain of the LLNA: BrdU-FC, or indicated improved performance with respect to LLNA false negatives and positives.

The Panel was then asked if the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: BrdU-FC had been adequately evaluated and compared to the traditional LLNA. The Panel agreed that the relevance of the LLNA: BrdU-FC was adequately evaluated and compared to the traditional LLNA, and supported the inclusion of accuracy analyses with and without equivocal materials.

Three substances (benzalkonium chloride, resorcinol, and Tween 80) produced a false positive response compared to the traditional LLNA and guinea pig test when tested using the LLNA: BrdU-FC (based on immunophenotyping, benzalkonium chloride was subsequently classified as an irritant rather than a sensitizer). The Panel was asked if they could identify any characteristics associated with these or similar substances that might suggest that using the LLNA: BrdU-FC to test such substances would not be appropriate or that positive results for substances with such properties may warrant additional testing. Overall, the Panel stated that there were not any patterns or unifying concepts that explained the three false positive results in the available data set. They noted that only a single laboratory is using the LLNA: BrdU-FC method and recommended that the raw data on which the reports were prepared be made available in order to allow further investigation. The Panel also suggested that additional studies

be conducted to determine whether LLNA: BrdU-FC results with these three substances are repeatable.

Dr. Raymond Pieters stated that benzalkonium chloride and Tween 80 are considered aggressive irritants, but both published data (Manetz and Meade. 1999; Varani et al. 2008) and unpublished data from his laboratory has shown that benzalkonium chloride (5%) is more potent than SLS in the stimulation of lymph node cell proliferation and may therefore may actually be considered a sensitizer. However, in the traditional LLNA these compounds did not increase the SI above the threshold for a positive response (i.e., SI ≥3), so they were identified as non-sensitizers.

4.2.4 Test Method Reliability

The Panel was asked if the intralaboratory reproducibility of the LLNA: BrdU-FC had been adequately evaluated and compared to the traditional LLNA and if any limitations were apparent based on this assessment. The draft LLNA: BrdU-FC BRD analyzed data from repeat testing of HCA in six different vehicles and intralaboratory reproducibility was assessed by a coefficient of variation (CV) evaluation. The calculated CVs ranged from 30% to 53%. The Panel agreed that the relatively large SD and associated CV values raised questions about the extent of experiment-to-experiment variability. There was less concern about vehicle choice and effects on the range of group means than about the CVs greater than 50% for the group means of HCA tested in DMSO and AOO. The large number of repeated experiments for these tests would have been expected to dramatically reduce variability. The Panel concluded that the results suggested that key elements of assay standardization were not yet developed. Further evaluation using other positive control substances would have been valuable to more adequately characterize reproducibility.

4.2.5 Data Quality

The studies evaluated in the draft BRD for the LLNA: BrdU-FC were not all conducted in accordance with GLP guidelines although they were done in a laboratory that routinely conducts GLP studies (G. DeGeorge, personal communication). The Panel was asked to discuss what impact this might have on the evaluation of the LLNA: BrdU-FC. The Panel considered that, even without formal GLP compliance, the current LLNA: BrdU-FC results appeared to reflect a sincere attempt to perform work of high quality. The only area in which a lack of full GLP compliance may have been a source of assay variability was in the quantitative analysis of dosing solutions. For instance, failure to appreciate differences in composition of dosing solutions between experiments caused by test article instability or other phenomena may account for the relative large variability in intralaboratory data and possibly of some of the discordant results (i.e., false negatives and differences in LLNA: BrdU-FC results between repeat studies for the same substance). Thus, the Panel viewed that any additional studies undertaken to validate the test method should ideally be GLP-compliant.

Furthermore, the original records for these studies were requested but had not yet been obtained at the time of the Panel review. As a result, an independent audit could not be conducted to confirm that the reported data was the same as the data recorded in laboratory notebooks. The

Panel was asked if they agreed that any recommendations from ICCVAM should be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription. The Panel agreed that, although a request for original data had been made, it was good practice to hold final recommendations until an independent audit could be performed. While it would be expected that no serious errors would be uncovered which would alter the current findings, an audit would confirm assay performance to date and position ICCVAM for further consideration of the LLNA: BrdU-FC.

4.2.6 Consideration of All Available Data and Relevant Information

The Panel was also asked if, based on the draft LLNA: BrdU-FC BRD, all the relevant data identified in published or unpublished studies that employ this test method had been adequately considered. Furthermore, they were asked that if there were other comparative test method data that were not considered in the draft BRD, how such data might be obtained. Overall, the Panel considered that all the relevant data identified in published or unpublished studies that employed this test method had been adequately considered in the draft LLNA: BrdU-FC BRD. However, some additional information was available in the peer-reviewed literature on application of BrdU in the LLNA with other methods of detection (e.g., histochemistry, ELISA). The Panel felt that these could have been briefly mentioned in the draft BRD for the LLNA: BrdU-FC method, simply as a means of indicating the utility of non-radiolabeled tracer methods in the LLNA. Furthermore, if an analysis of the CV for the traditional LLNA was undertaken, a more direct comparison with the LLNA: BrdU-FC could have been performed.

4.3 Comments on the Draft ICCVAM Test Method Recommendations on the LLNA: BrdU-FC

4.3.1 Test Method Usefulness and Limitations

The Panel was asked if they agreed that the available data and test method performance support the ICCVAM draft recommendations for the LLNA: BrdU-FC in terms of the proposed test method usefulness and limitations (i.e., that it may be useful for identifying substances as potential skin sensitizers and non-sensitizers, but that more information and data are needed before a recommended use of the LLNA: BrdU-FC can be made). The Panel agreed that the available data and test method performance of the LLNA: BrdU-FC support the draft ICCVAM recommendations. They considered the proposed test method usefulness and limitations to have well summarized the limits of the information supplied and the additional information that would need to be generated or provided for further consideration of this test method. As a result, the LLNA: BrdU-FC could not at this stage be considered scientifically validated as a replacement alternative to the traditional LLNA. Still, the test method recommendation should clearly state that the test method was not "invalid" but simply that there was currently not sufficient evidence and information to affirm that it had been adequately validated by ICCVAM. Instead, the Panel considered that the LLNA: BrdU-FC could be recommended in instances where mechanistic information about a sensitizer is required.

The Panel was asked if restrictions on using radioactive materials were or were not present, whether or not the LLNA: BrdU-FC should be routinely recommended for hazard identification

of skin sensitizing substances in lieu of having to possibly use guinea pig tests. The Panel agreed that it is preferable to use alternative methods for the LLNA (i.e., ELISA detection of BrdU or histochemical detection of BrdU-labeled cells), as opposed to application of guinea pig test methods, if a limitation on radioisotope use exists (e.g., the lack of a radioactivity use license). This rationale is based on avoidance of the less quantitative guinea pig test methods, which may employ adjuvant treatment with associated animal stress and harm. Still, at this time, the Panel considered that data gaps in the LLNA: BrdU-FC method precluded recommending it for routine hazard identification of skin sensitizing substances in lieu of the traditional LLNA, whether or not limitations on using radioactive materials exist. Policy issues regarding restrictions on radioactivity should have no impact on this science-based conclusion.

The Panel was asked if the ICCVAM recommendations adequately addressed concerns associated with the false positive rate of 17% (3/18 substances) calculated for the LLNA: BrdU-FC and if there were other suggestions for additional guidance or limitations that should be considered. The Panel agreed that the relatively high false positive rate was adequately identified and discussed, and that no mechanistic reason could be identified for these results based on available information. The Panel noted that it might be worthwhile to point out in the final BRD the impact on human health of false positive results versus false negative results in the context of hazard screening and identification. Consideration of factors such as intended use and target population of the false positive substances would further aid in characterizing human risk for these substances.

4.3.2 Test Method Protocol

The Panel was asked whether or not they agreed that the available data supported the ICCVAM draft recommendations for the LLNA: BrdU-FC procedure in terms of the proposed test method standardized protocol and if not, what recommendations would they make. The Panel noted that the draft ICCVAM recommendations for conduct of a standardized method for the LLNA: BrdU-FC variant were relatively brief and stated only that all applicable portions of the 1999 ICCVAM procedure be carefully followed. The Panel agreed that the available data supported the ICCVAM draft recommendations for the LLNA: BrdU-FC procedure in terms of adhering to the ICCVAM LLNA protocol (ICCVAM 1999; Dean et al. 2001). In particular, the Panel agreed that at least five animals per dose group should be used, particularly in light of Dr. Haseman's power analysis (see **Table F-3**). The Panel did note however that power calculations could be undertaken to determine if fewer animals per dose group might be adequate for post-validation studies, though Dr. Haseman's power analyses suggest that this is unlikely. The majority of the Panel also agreed with the ICCVAM-recommended protocol to use individual animal data although a minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD TG 429 guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals per dose group could be considered acceptable.

Further, the Panel considered the methodological description of the LLNA: BrdU-FC procedure supplied by MB Research Labs (Appendix A to the draft BRD) to be comprehensive. The

utility of ear swelling or other methods to detect inflammation/excessive local irritation appear to be warranted in every variation of the LLNA (including the traditional LLNA), but should be further investigated before routine inclusion in any protocol is recommended.

The Panel was asked whether the LLNA limit dose procedure could be applied to the LLNA: BrdU-FC. The Panel agreed that the LLNA limit dose procedure could be applied to the LLNA: BrdU-FC as long as the limitations associated with the limit dose procedure were appreciated. Furthermore, application of the limit dose procedure to the LLNA: BrdU-FC would require adherence to a validated LLNA: BrdU-FC test method protocol with the exception that the middle and low dose groups would be excluded. Furthermore, it would need to be confirmed that the number/pattern of sensitizers that would have been identified/missed from the "high dose" group would mirror that of the traditional LLNA.

Table F-3 Power Calculations for the LLNA: BrdU-FC[1]

Parameter	3.0-fold Increase[2]	2.5-fold Increase	2.0-fold Increase	1.5-fold Increase	1.3-fold Increase
Mean Rx response	30279	25232.5	20186	15139.5	13120.9
Log (Mean Rx response)	10.318	10.136	9.913	9.625	9.482
Difference from control (log scale)	1.098	0.916	0.693	0.405	0.262
Difference/SD	1.75	1.46	1.10	0.65	0.42
Power for N=5	80%	50-80%	<50%	<50%	<50%
Power for N=4	50-80%	50%	<50%	<50%	<50%
Power for N=3	50%	<50%	<50%	<50%	<50%
Other Power	95% (N=9)	95% (N=12)	95% (N=19)	95% (N=52)	95% (N>100)
Other Power	90% (N=7)	90% (N=10)	90% (N=15)	90% (N=42)	90% (N>100)

Abbreviations: N = number of animals; Rx = Treatment; SD = standard deviation.

[1] The power calculations above are based on a one-sided $p<0.05$ Student's t test applied to log-transformed data from vehicle control LLNA BrdU-FC tests.

[2] Fold-increase = Required increase above the vehicle control for a positive response

4.3.3 Future Studies

Finally, the Panel was asked whether they agreed that the available data supported the ICCVAM draft recommendations for the LLNA: BrdU-FC in terms of the proposed future studies. The Panel agreed that the ICCVAM draft recommendations for future studies highlighted the unanswered questions raised by the available database. Specifically, conducting interlaboratory studies as a part of the validation process was considered important. As mentioned previously, the Panel viewed that the immunological markers suggested for the LLNA: BrdU-FC in the draft BRD were acceptable but that additional immunological markers for discrimination of irritant versus sensitization phenomena might also be identified. A suggestion for a future study was to use the surface marker relating to CD4 Th cells or internal marker relating to Th1 cells (interferon-γ).

In general, for any future work, the Panel considered that efforts should be made to decrease the variability and thereby increase the power of the test in order to ensure that more animals were not needed relative to the traditional LLNA or other alternative LLNA protocols. For instance, further optimization of the LLNA: BrdU-FC method should include kinetic studies to demonstrate that the optimal protocol was being used.

4.3.4 Performance Standards

The draft BRD indicated that the LLNA: BrdU-FC protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. According to the proposed draft ICCVAM LLNA Performance Standards for the traditional LLNA, any change to the LLNA protocol other than the method used to assess lymphocyte proliferation was considered a "major" change. According to this criterion, the Panel considered the protocol differences between the LLNA: BrdU-FC and the traditional LLNA to be "minor" changes, and therefore considered that the validity of this test method could be based on the draft ICCVAM LLNA Performance Standards. However, the Panel also recognized that this depended on a clear definition of what constituted a "major" versus a "minor" change, or a different protocol altogether. Thus, further consideration of this topic could be addressed once the recommendations for the current draft ICCVAM LLNA Performance Standards were finalized. The Panel found it difficult to identify any additional requirements for methods like the LLNA: BrdU-FC.

Even if the draft ICCVAM LLNA Performance Standards were not found to apply to the LLNA: BrdU-FC, the Panel considered that the impact of the LLNA: BrdU-FC accuracy analysis based on 13 of the 18 proposed required reference substances in the draft ICCVAM LLNA Performance Standards should not have a major impact on the overall evaluation of test method accuracy, as 45 substances, representative of an appropriate range, were tested. However, based on consideration for development of LLNA performance standards, it would be desirable for validation purposes that the substances missing from the range of 18 standard materials be assessed in the LLNA: BrdU-FC protocol.

The draft LLNA: BrdU-FC BRD also indicated that three out of six sensitizers for which EC3 data were available had EC3 values that were outside of the proposed 0.5x to 2.0x EC3 acceptability range, which was developed based on the traditional LLNA. The Panel viewed that the primary concern seemed to have less to do with the variation in the response than with a concern that the range of response would skew the interpretation of any LLNA: BrdU-FC results used for sensitization potency estimates. Furthermore, it was not known if the same vehicle was used to derive both EC3 values/ranges. The proposed 0.5 x to 2.0 x range seemed to be based upon empirical/goodness of fit rather than any biological constant. The appropriateness of this range should be considered further when the finalized ICCVAM LLNA Performance Standards document is considered. In general, if the vehicles were different the question is irrelevant.

4.3.5 Comments with Specific References to the Text

The Panel made the following comment with specific reference to the text in the draft ICCVAM test method recommendations on the LLNA: BrdU-FC and suggested that it be addressed:

- Lines 39-41: For parallel construction of this sentence with the preceding sentence, suggest substituting the following "One of the other equivocal substances, salicylic acid, is one of the recommended reference standard materials used as a non-sensitizer in the draft ICCVAM LLNA Performance Standards…." This is based on the assumption that salicylic acid was the substance intended for discussion and that it was used in the draft ICCVAM LLNA Performance Standards as a model non-sensitizer.

5.0 Non-Radioactive LLNA Protocol - The LLNA: Bromodeoxyuridine Detected by ELISA (BrdU-ELISA) Test Method

5.1 Comments on the Draft BRD for Completeness, Errors, and Omissions

5.1.1 General Comments

The Panel was asked if there were any errors in the draft LLNA: BrdU-ELISA BRD that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. The Panel noted that, in general, all of the data included were relevant, and that it was apparent that considerable effort had been involved in carefully developing the comprehensive database. The Panel noted that they would have preferred to have the original papers by Dr. Takeyoshi included in the review materials, but they were easily retrieved from the journal websites. The Panel indicated that raw data (i.e., the actual optical density at 370 nm [OD_{370}] readings for the triplicates and the SD of the triplicates) are necessary for a thorough evaluation. Additionally, the Panel noted that only a relatively small number of substances had been tested in the LLNA: BrdU-ELISA.

When considering the animal welfare impact of implementing the LLNA: BrdU-ELISA, the Panel agreed that it would be less painful than guinea pig tests in those circumstances where the use of radioactive materials are restricted. Thus, the Panel agreed that the test represents a potential refinement. The Panel further stated that, if there is not an option to replace the guinea pig test with a non-animal test, decreasing the extent of pain and distress should be the first animal welfare priority. The Panel cautioned that at some point, however, the numbers of animals being utilized must be considered. An eventual recommendation that the LLNA: BrdU-ELISA be routinely used instead of guinea pig test methods where the use of radioactive substances are restricted would apparently require a significant increase in the number of mice killed per test (to increase the statistical power of the test method - see **Section 5.1.2** below) if an SI ≥1.3 is deemed the appropriate criterion to use for determining a positive response. The Panel stated that it would be helpful to know how many guinea pigs are currently being used nationally and internationally for skin-sensitization tests, and how many mice would be used in the LLNA: BrdU-ELISA with the SI ≥1.3 criterion. Even an order of magnitude estimate would help the Panel judge whether the increase in numbers of mice needed is justified as the quest to relieve pain in guinea pigs is pursued.

5.1.2 General Statistical Comments

The Panel was concerned about using an SI of ≥1.3 to optimize the performance of the LLNA: BrdU-ELISA method. One Panel member's extensive experience with ELISA protocols was cited as evidence that the difference between the OD_{370} of the vehicle and the positive test at 1.3 would not likely be statistically significant. The Panel recommended that the raw data must be reviewed to evaluate this. In addition, based on Dr. Joseph Haseman's power analysis (see **Table F-4**), the Panel stated that it was difficult to justify using a SI ≥1.3 as the decision criterion since it would result in a significant increase in the number of animals needed to obtain an acceptable confidence level. In this regard, the Panel recommended that power

calculations should be routinely recommended to ensure that the appropriate number of animals per dose group is being analyzed.

Table F-4 Power Calculations for the LLNA: BrdU-ELISA[1]

Parameter	3.0-fold Increase[2]	2.0-fold Increase	1.3-fold Increase
Mean Rx response	0.399	0.266	0.173
Log (mean Rx response)	-0.92	-1.32	-1.75
Difference from control (log scale)	1.10	0.70	0.27
Difference/SD	3.64	2.32	0.89
Power for N=4	99%	80-90%	<50%
Other power	95% (N=3)	95% (N=5)	50% (N=8)
Other power	–	50-80% (N=3)	80% (N=16)
Other power	–	–	90% (N=22)

Abbreviations: N = number of animals; Rx = Treatment; SD = standard deviation.

[1] The power calculations above are based on a one-sided $p<0.05$ Student's *t* test applied to log-transformed data from vehicle control LLNA:BrdU-ELISA tests.

[2] Fold-increase = Required increase above the vehicle control for a positive response

5.1.3 Comments with Specific References to the Text

The Panel also identified the following minor formatting and grammatical errors, as well as information gaps, in the draft BRD:

- The Panel noted a discrepancy between the draft LLNA: BrdU-ELISA BRD and the draft ICCVAM LLNA Performance Standards in the vehicle used for testing 2-mercaptobenzothiazole. Table 6-2 of the draft LLNA: BrdU-ELISA BRD indicated that the vehicle was AOO but the revised draft ICCVAM LLNA Performance Standards indicated that the vehicle was DMF (see page C15, C22 of September 7, 2007, draft and page B-6 of January 7, 2008, revised draft). Additionally, Table 1 on page C-7 of the revised draft ICCVAM LLNA Performance Standards listed AOO as the vehicle for 2- mercaptobenzothiazole. For both vehicles, the revised draft ICCVAM LLNA Performance Standards indicated that the EC3 value is 2.5%, although the text on page 10 of the draft LLNA: BrdU-ELISA BRD stated "the NICEATM database of traditional LLNA studies indicates that 2-mercaptobenzothiazole has a higher EC3 value when tested in AOO (mean EC3=9.8%) compared with DMF (mean EC3=2.5%)...."
- The Panel noted that Table 6-1 of the draft LLNA: BrdU-ELISA BRD indicated that, when compared to the guinea pig and human test data, the sensitivity and specificity of the LLNA: BrdU-ELISA is lower than that of the traditional LLNA. In

fact, depending on the SI threshold value used, the sensitivity and specificity of the LLNA: BrdU-ELISA can be higher than that of the traditional LLNA.

The Panel recommended that the draft LLNA: BrdU-ELISA BRD be updated to rectify these errors and omissions.

5.2 Comments on the Validation Status of the LLNA: BrdU-ELISA

5.2.1 Test Method Protocol

The data generated for the substances analyzed in the LLNA: BrdU-ELISA test method came from auricular lymph nodes from four individual mice in each dose group. The ICCVAM-recommended LLNA protocol (ICCVAM 1999; Dean et al. 2001) and OECD TG 429 (OECD 2002) recommend a minimum of five animals per dose group when collecting individual animal data. The Panel was asked what impact might the use of four animals per dose group have on the accuracy and reliability of the LLNA: BrdU-ELISA, and if the Panel agreed with the ICCVAM recommendation that future use of this test method protocol should include five animals per dose group. The Panel majority agreed with the ICCVAM recommendation that future use of this test method should use five animals per dose group and collect individual animal data, as per the ICCVAM-recommended protocol. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable. Based on the supplemental data provided by Dr. Haseman, the power to detect a three-fold increase with a sample size of four was determined to be 99%. These calculations, however, assume that a sample size of four is always obtained. If a sample size of four was planned and fewer usable data values were obtained, then the experiment might be compromised. Furthermore, the Panel concluded that testing for and eliminating "outliers" from experiments with small sample sizes is questionable. A reduction in sample size from five to four was not recommended unless data was provided on the frequency with which "outliers" occurred and an analysis is performed that establishes that a reduction in the nominal sample size from five to four would not compromise the performance of the test method. The Panel stated that the handling of suspected "outliers" and the use of robust statistics are issues that need to be addressed in such an analysis. For example, robust procedures may compensate for apparent "outliers" and eliminate the impulse to discard data. An example is calculating the mean values used in the SI on a log scale and then exponentiating the result to construct the SI.

The Panel also indicated that it was important to routinely include a positive control group in test method validation experiments (e.g., HCA), which was likely not the case for most of the LLNA: BrdU-ELISA validation experiments. Although the Panel did not reach consensus, they did consider the suggestion that for laboratories in which the LLNA is "routinely" performed, positive controls (e.g., HCA or a substance that matches the chemical class of the test substances) could be run at intervals for quality control purposes rather than concurrent with each experiment in which substances are tested. The Panel also discussed that omitting the concurrent positive control should not be recommended for laboratories that perform the LLNA only "occasionally." In their discussions, the Panel was not able to conclude what should

constitute "routine" or "occasional" LLNA use or what would be an appropriate interval between positive control testing when a concurrent positive control is not used.

5.2.2 Substances Used for the Validation Studies

The Panel was asked whether the LLNA: BrdU-ELISA database was representative of a sufficient range of chemical classes and physicochemical properties such that the test method would be applicable to any of the types of chemicals and products typically tested for skin-sensitization potential. The Panel indicated that the ratio of solids to liquids was not comparable; more solids should be included. The Panel further indicated that more substances for which traditional LLNA data are available should be tested, and that compounds including metals (e.g., nickel, cobalt), mixtures, and substances in aqueous solutions should be included.

5.2.3 Test Method Accuracy

The current accuracy analysis using an SI ≥3.0 or SI ≥1.3 to identify sensitizers is based on overall concordance with the traditional LLNA. Accuracy statistics compared to the guinea pig tests and human data/experience were also provided to the Panel. The Panel was asked whether these comparisons were appropriate for assessing the accuracy of the LLNA: BrdU-ELISA. The Panel indicated that comparing the LLNA: BrdU-ELISA performance to the traditional LLNA and the guinea pig tests were appropriate. Comparisons between the LLNA: BrdU-ELISA and human data were considered particularly valuable because the traditional LLNA doesn't match human data with 100% accuracy. For this reason, the Panel considered comparing the performance of the LLNA: BrdU-ELISA with that of the traditional LLNA with respect to predicting the human outcomes to be the best method of comparing these two LLNA protocols. The Panel concluded that in moving forward with any test method recommendation, key importance should be placed on interpreting the test results and making them clinically applicable to humans.

Takeyoshi et al. (2007) performed an accuracy analysis using decision criteria other than an SI ≥3.0 to classify substances as sensitizers. Maximal accuracy for the LLNA: BrdU-ELISA occurred when an SI ≥1.3 was used to distinguish between sensitizers and non-sensitizers. Using this decision criterion, the LLNA: BrdU-ELISA achieved an accuracy of 91% (21/23), with a sensitivity of 100% (16/16) and a specificity of 71% (5/7) (i.e., there were no false negatives and two false positives). The Panel was asked whether this analysis supported a recommendation that the decision criteria be based on an SI ≥1.3, and if there were concerns with using such a small increase (i.e., 1.3-fold) above the vehicle control response as the basis for identifying a positive response. The Panel did not support using an SI ≥1.3 as the criterion for positive results. An SI=1.0 means there was no difference between the vehicle control and the test substance. An SI=1.3 represents a 30% increase from the vehicle control. The difference between the OD_{370} of the vehicle and the positive test at 1.3 may not be statistically significant. An SI=3.0, which represents a three-fold difference between the vehicle and a positive test, would be a more believable positive difference. If the positive test criteria must be reduced to 1.3, then the Panel questioned whether the protocol is useful in its current state. The supplemental information that provided power calculations indicated that it would not be

realistic to expect to detect a 1.3 fold increase in the control response without a significant addition of animals. Although using SI ≥1.3 increases the accuracy of the test, it comes at an increased cost to animals, which merits consideration. Furthermore, the ICCVAM (1999) report stated that an irritating chemical might induce proliferation, but that the response seldom exceeds an SI ≥3.0 (page 6). The Panel concluded that this might provide further justification against using a low SI (e.g., 1.3) as a threshold for a positive response.

The Panel was asked if the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: BrdU-ELISA, using the SI ≥3.0 criterion, had been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft LLNA: BrdU-ELISA BRD). If not, the Panel was asked what other analyses should be performed. The Panel agreed that the relevance of the LLNA: BrdU-ELISA, using the SI ≥3.0 criterion, had been adequately evaluated. The Panel further stated that a better evaluation could be performed, however, if the database for the LLNA: BrdU-ELISA included more substances with traditional LLNA, guinea pig, and human data. The Panel considered the false negative rate of the test method to be excessive when results are compared with that obtained in the traditional LLNA (29/33/27% for the various datasets) or with human data (39%) – the results should be at least comparable with the traditional LLNA.

Using the SI ≥3.0 criterion, there were four substances (aniline, 4-chloroaniline, 2-mercaptothiazole, and hydroxycitronellal) that produced false negative responses when tested using the LLNA: BrdU-ELISA compared to the traditional LLNA. The Panel was asked whether it could identify any characteristics associated with these or similar substances, compared to the correctly identified sensitizers, that might indicate that such substances should not be tested in the LLNA: BrdU-ELISA or that negative results for such substances should indicate a need for confirmatory testing. The Panel could not identify any characteristics associated with these substances that might allow the identification of these substances as false negatives prior to testing. The Panel stated that the LLNA: BrdU-ELISA test, using the standard SI ≥3.0 to indicate positive results, simply does not perform well for identifying sensitizers.

5.2.4 Test Method Reliability

The Panel was asked whether the intralaboratory reproducibility of the LLNA: BrdU-ELISA had been adequately evaluated and compared to the traditional LLNA, and whether any limitations were apparent based on this intralaboratory reproducibility assessment. The Panel indicated that the number of substances evaluated for intralaboratory reproducibility was too few and, in some cases, there was a wide variation in repeat test results for the same substance. Only six substances (five sensitizers and only one non-sensitizer) were tested multiple times. The non-sensitizer, propylene glycol, was tested only twice and opposite results were obtained. The Panel considered the results of an intralaboratory reproducibility evaluation that was based on two discordant results only to be unacceptable. The numbers calculated in Table 7-1 of the draft LLNA: BrdU-ELISA BRD are correct, but the Panel questioned the dependability of the data since only two to three values were available for calculating the mean and CV. The Panel considered the CV values (over 30%) high, compared to the traditional LLNA (draft LLNA: BrdU-ELISA BRD Tables 7-1 to 7-3). The Panel stated that at least four independent tests with

three concentrations tested represent a solid basis for calculation. The Panel considered the number of tests for intralaboratory concordance analysis to be insufficient, and stated that more intralaboratory testing is needed. The Panel recommended an evaluation of the intralaboratory reproducibility of the EC ≥1.3 and that the analysis of the variability of the ECt be conducted on a log scale.

The substances evaluated for intralaboratory reproducibility of the LLNA: BrdU-ELISA study were not coded. The Panel was asked whether the lack of coding of test substances adversely impacts or biases the current evaluation. The Panel stated that, although coding of substances is preferred for independent testing and evaluation of test results, the current data should not be rejected from consideration because the substances tested were not coded.

The Japanese Center for Validation of Alternative Methods (JaCVAM) has implemented a multi-laboratory validation study of the LLNA: BrdU-ELISA. The Panel was asked whether the study design was appropriate to adequately determine the extent of interlaboratory reproducibility for the LLNA: BrdU-ELISA. If not, the Panel was asked what other studies should be performed. The Panel stated that they had insufficient time to evaluate the study design and that they could not evaluate interlaboratory reproducibility because the study data were not available at the time of their evaluation.

5.2.5 Data Quality

The studies evaluated in the draft BRD for the LLNA: BrdU-ELISA were not conducted in strict accordance with GLP guidelines, although there were reportedly performed in laboratories that conduct GLP studies (M. Takeyoshi, personal communication). In other words, an audit report was not available. Also, the raw data were unavailable for an independent audit. The Panel was asked to discuss what impact this might have on the evaluation of the LLNA: BrdU-ELISA. The Panel concluded that ideally, validation studies should be performed in accordance with GLP guidelines. Although the systems employed for tests (i.e., test facilities, staff, reagents, etc.) were identical to those for GLP-compliant studies, the data quality may be questioned and therefore should at least be available for a retrospective independent audit. However, in this case, the Panel concluded that the lack of GLP compliance was not likely the reason for the poor results obtained with the LLNA: BrdU-ELISA.

The original records for these studies were requested but had not been received by the time the Panel convened. As a result, an independent audit could not be conducted to confirm that the reported data in peer reviewed publications and a poster presentation is the same as the raw data. The Panel was asked whether any recommendations from ICCVAM should be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription. The Panel concluded that, to have confidence in data quality, ICCVAM recommendations should be contingent upon the completion of an independent audit. Moreover, if an SI ≥1.3 is used as the criterion for positive results, review of the raw data is necessary to confirm statistically significant differences. The Panel concluded that this test, as described, had poor accuracy, poor sensitivity, and poor specificity. The Panel stated that changing the SI decision criterion from 3.0 to improve test performance, especially to such a drastic change as SI ≥1.3, is a mistake and sets a dangerous precedent.

5.2.6 Consideration of All Available Data and Relevant Information

Based on the draft LLNA: BrdU-ELISA BRD, the Panel was asked whether all the relevant data identified in published or unpublished studies that employ this test method had been adequately considered, and if other comparative test method data that were not considered were available. If yes, the Panel was asked to suggest how to obtain such data. The Panel believed that all of the relevant data, with the exception of the interlaboratory reproducibility study, were presented and that they were not aware of any omissions.

5.3 Comments on the Draft ICCVAM Test Method Recommendations on the LLNA: BrdU-ELISA

5.3.1 Test Method Usefulness and Limitations

The Panel was asked whether the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA in terms of the proposed test method usefulness and limitations. The Panel agreed with the ICCVAM recommendation that the LLNA: BrdU-ELISA may be useful for identifying substances as potential skin sensitizers and non-sensitizers but that, at this time, more information and data are needed before a recommended use of the LLNA: BrdU-ELISA can be made. The Panel also stated that a detailed protocol is needed, in addition to sufficient quantitative data for a more comprehensive analysis based on a larger set of balanced reference substances with regard to physicochemical properties and sensitization potency, as well as an evaluation of interlaboratory reproducibility.

The Panel was asked whether the LLNA: BrdU-ELISA should be routinely recommended for hazard identification of skin sensitizing substances in lieu of using guinea pig tests if restrictions on using radioactive materials *are* present, due to the fact that fewer animals might be used and because pain and distress would be avoided. The Panel stated that if the accuracy of the test method was at least similar to the traditional LLNA, the LLNA: BrdU-ELISA might be routinely recommended for hazard identification of skin-sensitizing substances in terms of reduction of animals and refinement of the pain and distress associated with guinea pig tests. Clearly, using the LLNA: BrdU-ELISA instead of the traditional LLNA or guinea pig test methods would also offer advantages for the environment due to the use of a non-radioactive probe chemical. However, the Panel stated that the accuracy of the current LLNA: BrdU-ELISA dataset at SI ≥3.0 was inadequate and not equivalent to the traditional LLNA. The Panel also noted that if an SI ≥1.3 was used because of its apparent increased accuracy, additional mice (over and above the number needed in the standard LLNA test) would apparently be needed (see **Table F-4**). Thus, the Panel stated that reduction of animals would not be achieved. In this regard, the Panel noted that some quantification of the total animal use numbers would be useful as it is not clear whether the increased number of mice used would outweigh the avoidance of pain and distress in guinea pigs.

The Panel was asked whether the LLNA: BrdU-ELISA procedure or other valid and accepted non-radioactive method could be routinely recommended for hazard identification of skin sensitizing substances instead of the traditional LLNA if limitations in using radioactive materials *are not* present. The Panel stated that the LLNA: BrdU-ELISA procedure could not be routinely recommended for hazard identification of skin sensitizing substances instead of the

traditional LLNA, because the accuracy of this test at SI ≥3.0 was inadequate. In other words, the current dataset available for the LLNA: BrdU-ELISA did not predict the guinea pig or human outcomes as accurately as the traditional LLNA. Thus, the Panel acknowledged that there is the possibility that additional data might impact on the accuracy statistics and eliminate this concern. The Panel stated that factors that weigh on a decision of replacement of the LLNA with a non-radioactive method would include:

- Are more animals needed?
- Is the replacement test safer and less complex?
- Is the replacement test more efficient?
- Is the replacement test less costly?

The Panel stated that additional factors to consider might exist, but overall recommended that whether or not restrictions on radioactivity exist, a test that causes the least pain and uses the fewest number animals should be preferred, as long as adequate test method performance is maintained. Clearly, policy issues regarding restrictions on radioactivity should have no impact on this science-based conclusion.

The Panel was asked whether using a decision criterion of SI ≥1.3 instead of SI ≥3.0 resolved any concerns with respect to potential false positives or false negatives that may occur in this test method. The Panel was also asked for other suggestions for additional guidance or limitations that should be considered. The Panel stated that using a decision criterion of SI ≥1.3 instead of SI ≥3.0 would not itself resolve any concerns; more raw data are needed for a broader set of reference positive and negative sensitizers, including metals, mixtures, and aqueous solutions. The Panel also stated that a detailed protocol is needed, as is an evaluation of interlaboratory reproducibility. The Panel considered the current database to be inadequate, but based on the limited database, concluded that it might be more appropriate to use a statistically based decision criteria than a stimulation index.

5.3.2 Test Method Protocol

The ICCVAM draft recommendations state that the LLNA: BrdU-ELISA protocol should adhere to the ICCVAM LLNA protocol (ICCVAM 1999; Dean et al. 2001), except for measurement of lymphocyte proliferation. The Panel was asked whether the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA procedure in terms of the proposed test method standardized protocols. In general, the Panel agreed that the available data support the ICCVAM draft recommendations for this test method in terms of the standardized protocol. As stated previously, the Panel majority agreed with the ICCVAM recommendation that future studies should use five animals per dose group and collect individual animal data, as per the ICCVAM-recommended protocol. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable. The Panel further noted that using an SI <3.0 would require more

animals to achieve adequate statistical power (**Table F-4**) and therefore any considerations of reducing the SI to improve test method accuracy should include this point.

The Panel was asked whether the traditional LLNA limit dose procedure could be applied to the LLNA: BrdU-ELISA. The Panel stated that, if the LLNA: BrdU-ELISA was considered equivalent to the traditional LLNA, then it would be appropriate to apply the LLNA limit dose procedure to this test method. The Panel explained that, as in the case of the traditional LLNA, the protocol would be the same except for testing the maximum dose only, so applying the limit dose procedure would appear to have the same opportunity to reduce the number of animals needed to perform the test. However, using an SI ≥3.0 would not be appropriate because of the associated low accuracy in identifying sensitizers.

5.3.3 Future Studies

The Panel was asked whether the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA in terms of the proposed future studies. The Panel stated that the proposed future studies were justified. The Panel concluded that it is important to consider non-radioactive methods because, in some laboratories, it is difficult or not permissible to use radioactivity. The Panel also stated that, if more data were available and there was less variability in this test method, it might warrant re-evaluation. The Panel concluded that more data are needed, especially for determination of the appropriate threshold value for the decision criterion, and that interlaboratory reproducibility should be also evaluated (which presumably will occur once the Japanese interlaboratory validation effort is complete).

5.3.4 Performance Standards

The LLNA: BrdU-ELISA protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. According to the proposed draft ICCVAM LLNA Performance Standards for the traditional LLNA, any change to the LLNA protocol other than the method used to assess lymphocyte proliferation is considered a "major" change. The Panel was asked whether protocol differences between the LLNA: BrdU-ELISA and the traditional LLNA should be considered only "minor" changes and therefore if the validity of this test method should be based on the draft ICCVAM LLNA Performance Standards. In general, the Panel agreed that the LLNA: BrdU-ELISA protocol differs only in the method used to assess lymphocyte proliferation. Thus, based on the current draft ICCVAM LLNA Performance Standards, it should be considered as having only "minor" changes and therefore the validity of this test method could be based only on the draft ICCVAM LLNA Performance Standards.

However, the Panel concluded also that the answer to this question might differ depending on what the draft ICCVAM LLNA Performance Standards ultimately constitutes as a "major" change, a "minor" change, or a different protocol altogether. The Panel further stated that, depending on the goal of the assay, these distinctions may not be relevant. Ultimately, if a test method is able to make the correct prediction with regard to the sensitization potential of a test substance, then the issue of "major" versus "minor" changes in the protocol should not apply.

The Panel was asked, even if the draft ICCVAM LLNA Performance Standards do not apply to the LLNA: BrdU-ELISA, what impact should the accuracy analysis based on eight of the 18 proposed required reference substances in the draft ICCVAM LLNA Performance Standards (only one false negative and no false positives) have on the overall evaluation of test method accuracy. The Panel concluded that the accuracy analysis based only on eight of the 18 proposed required reference substances had a significant impact on the evaluation of test method accuracy. The low number of experiments provided data that resulted in unacceptable test method performance.

The Panel was asked whether there were concerns that 4/4 sensitizers, for which EC3 data were available, had EC3 values that were outside of the proposed recommended 0.5x to 2.0x EC3 acceptability range developed based on the traditional LLNA. The Panel concluded that the EC3 values outside the recommended 0.5x to 2.0x EC3 acceptability range raised concerns related to test reproducibility and reliability.[7]

The Panel was asked whether separate performance standards should be developed for the LLNA: BrdU-ELISA. The Panel concluded that separate performance standards for the LLNA: BrdU-ELISA were not needed because the test principles are identical to the traditional LLNA.

[7] During their public meeting on March 4-6, 2008, the Panel's discussion, conclusions and recommendations on the LLNA: BrdU-ELISA took place prior to the discussion, conclusions and recommendations on the draft ICCVAM LLNA performance standards. Following their discussion of the draft LLNA performance standards, the Panel concluded that that an evaluation of test method accuracy should be based on overall accuracy statistics when compared to the traditional LLNA, and not on a chemical-by-chemical match that is based on obtaining an EC3 value within a specified range of EC3 values.

6.0 Draft ICCVAM LLNA Performance Standards

6.1 Comments on the Proposed Purpose and Applicability

ICCVAM proposed that these performance standards should only be applicable to versions of the LLNA that incorporate "minor" modifications to the traditional LLNA. Currently, *this is limited to the use of non-radioactive reagents to measure lymphocyte proliferation*. It is considered essential that the modified LLNA should otherwise adhere to all other aspects of the traditional LLNA protocol, as defined by ICCVAM (1999) and Dean et al. (2001). This includes aspects such as: the sex and strain of mouse used, the number of mice per dose group, the timing and site of test article treatment, the duration between the last treatment and lymph node collection, the inclusion of concurrent negative and positive control groups, the measured endpoint (i.e., lymphocyte proliferation in the draining auricular lymph node), and the collection of data at the level of the individual mouse. The Panel was asked if they agreed that the use of non-radioactive reagents for measuring cell proliferation in the lymph nodes constitutes a "minor" modification to the traditional LLNA protocol.

The Panel noted that the draft ICCVAM LLNA Performance Standards are proposed for evaluating the acceptability of test methods that are functionally and mechanistically similar to the traditional LLNA (i.e., measuring the same biological effect), and understood that ICCVAM proposed that these performance standards should only be applicable to protocols that incorporate "minor" modifications to the traditional LLNA, as defined above. The Panel unanimously agreed that based on ICCVAM's definition, the use of non-radioactive reagents for measuring cell proliferation is a "minor" modification of the traditional LLNA protocol. However, the Panel also agreed that other modifications may be considered "minor" and that a better strategy for the performance standards might be to define criteria that need to be satisfied to insure that the method is mechanistically and functionally similar (see criteria listed under essential test method components). Examples of potentially acceptable modifications identified by the Panel include sex, strain, the use of rats rather than mice, number of animals per group, and timing of test article treatment.

Regardless of the modification, the Panel stated that the modified test method should be designed to measure only the induction phase of the immune response. This is crucial, since the traditional LLNA is intended for hazard identification with the underlying principle that stimulation of cell proliferation in the draining lymph node suggests that sensitization (i.e., induction) is occurring. Using only the induction phase as the method to identify hazardous substances involves a short time frame, and reduces pain and distress in treated animals (i.e., no dermatitis response). Furthermore, the Panel stated that the performance standards should not imply that the traditional LLNA, or any alternative LLNA protocol, is capable of specifically distinguishing a type IV hypersensitivity reaction (as might be inferred from the text beginning with line 342 of the draft ICCVAM LLNA Performance Standards document). Therefore, reference to type IV hypersensitivity reaction should be removed from the document.

The Panel was asked if they considered it necessary that a modified LLNA keep the same decision criteria for distinguishing between sensitizers and non-sensitizers as the traditional

LLNA (i.e., an SI ≥3.0). The Panel considered it unnecessary for a modified LLNA to keep this same decision criteria as a different method for measuring cellular proliferation might have better concordance with the human data at a SI different than 3.0. Thus, with any modified LLNA, the SI threshold defining a sensitizer would need to be established (i.e., it is important to consider if the results are biologically relevant to humans).

The Panel was asked if other procedural modifications could be identified as "minor," based on the description in the draft ICCVAM LLNA Performance Standards document, and therefore could be evaluated for equivalence to the traditional LLNA using the proposed performance standards. The Panel reiterated that sex, strain, the use of rats rather than mice, animals per group, and timing of test article treatment are also potentially "minor" modifications. Furthermore, the proposed performance standards appear robust; therefore, regardless of the modification (i.e., "major" or "minor"), there is the same expectation for test method performance. Dr. James McDougal offered a minority opinion to express his concern about the potential impact that allowing alternative LLNA protocols with modifications other than the method by which lymphocyte proliferation was measured would have.

The Panel was asked if they considered the draft ICCVAM LLNA Performance Standards applicable to the LLNA limit dose procedure. The Panel noted that the current draft ICCVAM LLNA Performance Standards could be applicable to the LLNA limit dose procedure as long as it is recognized that this procedure can only be used for a yes/no hazard classification (i.e., an ECt estimate is not feasible).

6.2 Comments on the Essential Test Method Components

The essential test method components are based on the ICCVAM-recommended protocol (ICCVAM 1999; Dean et al. 2001), which is the basis for the current EPA (2003) test guideline. There are some notable differences between these protocols and OECD TG 429 (OECD 2002) for the LLNA. The Panel was asked to comment on, when evaluations of non-radioactive versions of the traditional LLNA are conducted using these performance standards, whether it is necessary that the validation studies follow the ICCVAM-recommended protocol. The Panel indicated that ideally, there would be one globally recognized set of performance standards (ICCVAM, ECVAM, JaCVAM). However, when validating versions of the traditional LLNA where the only difference is in the use of a non-radioactive method to measure cell proliferation, the ICCVAM-recommended protocol should be used. If more extensive changes to the protocol are being considered, the following requirements should be considered during modifications of the LLNA:

- Application of the test substance should be to the skin, with sampling of the lymph nodes draining that site.
- Cell proliferation should be measured in the draining lymph node.
- No skin reaction should be present, since presence of a skin reaction might indicate the onset of the elicitation phase of skin sensitization.

- Data should be collected at the level of the individual animal to allow for an estimate of the variance within control and treatment groups. Using this variance, a power analysis needs to be conducted to demonstrate that the modified method is utilizing a sufficient number of animals per treatment group to permit hazard identification with at least 95% power.
- If dose response information is needed, there should be an adequate number of dose groups ($n \geq 3$) with which to adequately characterize the dose response for a given test substance.

The Panel was asked to comment on whether validation studies should include a concurrent positive control with each test substance and if so, whether the concurrent testing of the positive control and test substance should be conducted in the same vehicle or if different vehicles were acceptable. The Panel noted that a concurrent positive control should be included in each validation study to ensure that the test system was operating as expected and technical errors were not occurring. A concurrent positive control would be especially useful when an unknown test material was being tested or when a laboratory was collecting a dataset to serve as historical control data. However, if a known sensitizer was being tested, a concurrent positive control might not be needed, thus reducing animal use. Finally, the Panel concluded that the positive control should be tested in the same vehicle as the test substance. Using a different vehicle for the positive control would require an additional set of vehicle control animals.

The Panel was also asked whether the validation studies should use a minimum of five animals per dose group and collect lymph node data from individual animals. The Panel commented that until sufficient data were collected to enable a reliable power calculation to be conducted to determine the optimal number of animals per dose group, at least five animals per dose group should be used. The Panel also agreed that when validating a modified LLNA protocol, lymph node proliferation should be evaluated at the level of the individual animal within each dose group. Variance is only measurable if lymph nodes from individual animals are assessed. If the variability within a dose group of a modified LLNA protocol was substantially less than the traditional LLNA, reducing the number of animals per dose group might yield similar results. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable.

6.3 Comments on the Proposed Reference Substances

The Panel was asked if they agreed with the selection and prioritization criteria used to select the performance standards reference substances. The Panel noted that the rationale for selection of the reference substances included in the draft ICCVAM LLNA Performance Standards was well documented (taking into account the physicochemical characteristics, the purity, the stability, the quality of the *in vivo* data, and the chemical classes covered). The substances also appeared to be distributed over a wide range of EC3 values. However, the available database for some of the substances was insufficient. Among the 13 sensitizers in the "required" list, only five appear to have a robust database (i.e., have been tested in at least three independent

studies). Thus, consideration should be given to revising the list of substances and/or making the data for the substances on the current list more robust. Ideally, the reference list should be based only on substances with robust data for LLNA, human, and guinea pig tests.

The rationale for the number of substances included on the "required" list of substances (n=18) was provided in the draft ICCVAM LLNA Performance Standards. In addition, there were four additional substances that were described as problematic in the traditional LLNA (i.e., false negatives and false positives). The Panel was asked if they considered 18 required" substances to be an adequate number upon which to evaluate the performance of non-radioactive LLNA test methods, where the only protocol modification is the method for assessing cell proliferation in the auricular lymph nodes, and if not, how many reference chemicals should be tested. The Panel commented that ideally, one would like to be able to demonstrate that an assay is equivalent to the traditional LLNA. However, with the small number of reference substances available, establishing equivalence will be extremely difficult. Therefore, the Panel recommended that, for use in hazard identification, a modified method should be evaluated with all 22 substances (including false negatives and false positives) and accuracy statistics calculated. To the extent possible, rationale for any discordant results should be provided, but the most potent sensitizers (e.g., DNCB) should always be identifiable. There also should be considerable weight given to the balance between animal welfare and human safety when considering the adequacy of test method accuracy.

The panel considered it noteworthy that 19 of the 22 substances on the draft ICCVAM list are in common with the ECVAM performance standards list. The Panel also considered it important that substances be coded during validation studies.

It is also relevant to note that the Panel discussed the value of GLP procedures on several occasions during the meeting. In each instance, the Panel agreed that data collected under GLP conditions would be greatly preferred, particularly for reasons of data quality and the associated reliability of any interpretations. However, they noted that GLP compliance would not be considered a requirement that would automatically exclude data from consideration. The Panel concluded that other factors could be used to identify high quality data. Examples would include published in a peer-reviewed journal or obtained from a study conducted in a laboratory that routinely conducts GLP studies. Data generated under non-GLP conditions would be subject to a critical quality review, and as such the Panel considered it important to obtain the original records in order to confirm the reported data.

The Panel was asked if they considered the types of substances included in the reference substance list, with regard to relative sensitization potency, physicochemical characteristics, and vehicles, to be representative of the overall diversity of substances that are likely to be tested for skin sensitization. The Panel concluded that although the list should not be considered all-inclusive, it was sufficiently representative.

The Panel was also asked if there were other types of information relevant to skin sensitization that should be considered in order to demonstrate an adequately diverse reference list. The Panel commented that identifying concentrations of each of the substances that are known to cause excessive local irritation or overt systemic toxicity would be useful.

The Panel was asked if there were other substances that they considered to be more appropriate for assessing the sensitivity (i.e., ability of the test method to correctly identify sensitizing substances) and specificity (i.e., ability of the test method to correctly identify non-sensitizing substances) of non-radioactive LLNA test methods, and for which there is available LLNA, guinea pig, and human data. The Panel could not identify such substances given the time frame for consideration but reiterated that substances in the reference list should have robust data.

A subset of "discordant chemicals" (i.e., false negative or two false positive compared to guinea pig tests or human data) were included as "optional" substances that could be studied to evaluate if the proposed modifications might provide improved performance relative to the traditional LLNA. The Panel was asked to comment on the appropriateness of including these specific substances in the reference list, whether they should be required, whether different substances should be included, and if more false negative/positive substances should be tested. As mentioned previously, the Panel commented that it was appropriate to include such substances in the reference list and that they should be required and evaluated during the validation of alternative LLNA assays that are functionally and mechanistically similar to the traditional LLNA assay. The Panel noted that the substances that were considered discordant depended on the species to which comparisons are made (i.e., LLNA vs. guinea pig or LLNA vs. human). Still, since the "discordant compounds" were false negatives or positives in the traditional LLNA, they would provide an opportunity to determine if modifications to the traditional LLNA may even have increased accuracy.

Finally, the Panel was asked if "correct" results with these discordant chemicals would be sufficient to consider the alternative test method to be more predictive of skin sensitization than the traditional LLNA. The Panel concluded that correct results with the "discordant chemicals" would not be sufficient to consider the alternative test method to be more predictive of skin sensitization, but it could provide supporting evidence to indicate further testing with additional compounds would be of value.

6.4 Comments on the Test Method Accuracy Standards

The draft ICCVAM LLNA Performance Standards state that the non-radioactive proposed LLNA test method should exactly match the accuracy of the traditional LLNA when evaluated with the minimum set of 18 reference substances. The Panel was asked if they agreed that test method accuracy should be based on a chemical-by-chemical match with regard to identifying the chemicals as sensitizers or non-sensitizers. The Panel commented that although an assay that is able to predict the same hazard classification for the reference substances as the traditional LLNA is desired, with the small number of reference substances available (n=18), clearly establishing equivalence will be extremely difficult. Furthermore, even with this small number, there is a statistical multiple-comparisons[8] problem because more than one chemical is

[8] When multiple experiments are conducted and multiple observations, comparisons, or hypothesis tests are conducted, the chance of observing rare events increases. Suppose, for example, that an interval is established such that 5% of observations from a particular population of data are outside that interval. Then if k independent experiments generate data from this population (e.g., a standard normal distribution), the chances that all 20 results will lie inside the interval is $(1.0 - 0.05)k$ (N. Flournoy, personal communication).

being tested. The likelihood that a modified LLNA will fail to demonstrate equivalence to the traditional LLNA will increase with the number of chemicals that must be identified correctly. A statistical measure of concordance should be calculated so that accuracies can be compared between methods.

The Panel reiterated their recommendation that, for use in hazard identification, a modified test method should be evaluated with all 22 substances (including false negatives and false positives) and accuracy statistics calculated. A statistical measure of concordance should be calculated so that accuracies can be compared between methods. To the extent possible, rationale for any discordant results should be provided. However, the most potent sensitizers (e.g., DNCB) should always be identifiable. Considerable weight should be given to the balance between animal welfare and human safety when considering the adequacy of test method accuracy.

The draft ICCVAM LLNA Performance Standards recommend that, for each sensitizer, the threshold concentration that induces a positive SI response should be within 0.5x to 2.0x of the concentration obtained for the EC3 in the traditional LLNA. As described in the draft ICCVAM LLNA Performance Standards, statistical approaches have been used in an attempt to identify an appropriate range, but these calculated ranges do not appear to be the most practical. The Panel was asked to comment on the appropriateness of using this criterion to judge the equivalency of a non-radioactive version of the traditional LLNA and, if this approach was not acceptable, to suggest an alternative along with the basis for this approach. The Panel commented that the usefulness and limitations of the traditional LLNA for determining relative potency have not been definitively established, and therefore equivalence should not be based strictly on potency. Furthermore, the current database does not support the inclusion of EC3 values as a component of the accuracy evaluation. The range of 0.5x to 2x EC3 value suggested in the draft ICCVAM LLNA Performance Standards document are based on the experience with a range of known skin sensitizers tested in the standard LLNA. However, based on the available data provided, the 0.5x to 2.0x EC3 range may be too restrictive *if a strict interpretation of equivalence is applied*. The chances of a failure to achieve an EC3 within this range would vary from chemical to chemical depending upon the inherent underlying variability and robustness in the estimation of the EC3. For those chemicals for which the EC3 can be accurately estimated, the failure rate may be close to zero. For other, more variable chemicals, perhaps with fewer data points, the failure rate for a single chemical will be much higher.

The Panel reiterated their concern with regard to EC3 values (i.e., the statistical multiple comparisons problem). The likelihood that a modified LLNA will fail to demonstrate equivalence to the traditional LLNA will increase with the number of chemicals tested, the extent to which the new test must obtain the same EC3 value, and how independent the results are for different chemicals in the same lab.

For five of the 13 sensitizers on the draft ICCVAM reference substances list, the reference EC3 value was based on a single LLNA study (Table C1 of the draft ICCVAM LLNA Performance Standards). The Panel was asked to comment on the appropriateness of including such

chemicals in the list of recommended reference substances and whether or not the 0.5x to 2.0x criteria should be applied to such substances. The Panel concluded that the appropriateness of the 0.5x to 2.0x ECt range had not been adequately justified. It was inappropriate to include chemicals represented by only one LLNA study on a list of recommended reference substances, as there was insufficient data by which to calculate a robust mean ECt value. Thus, those compounds should either (1) be exchanged for compounds with sufficient EC3 data (i.e., have been tested in at least three independent studies using the same solvent), or (2) retained but not considered to be part of the ECt criterion until such data has been collected.

6.5 Comments on the Test Method Reliability Standards

The draft ICCVAM LLNA Performance Standards state that acceptable intralaboratory reproducibility will be indicated by a laboratory obtaining, in each of four independent experiments conducted with at least one week between each experiment, ECt values (the estimated concentration needed to produce an SI of a defined threshold [e.g., EC3]) for HCA that are generally within 0.5x to 2.0x (i.e., 5% to 20%) of the historical mean EC3 concentration (10%) for this substance, based on existing available traditional LLNA data. The Panel was asked if they considered four repeat experiments to be adequate. The Panel concluded that four experiments would be adequate, as requiring four independent experiments is similar to the original LLNA submission, as is a one-week interval between experiments. Therefore, these requirements were appropriate for a comparison of modified methods to the traditional LLNA. However, it would be useful to have this number evaluated statistically (see **Section 6.7**).

The Panel was asked if they considered testing HCA adequate for demonstrating intralaboratory reproducibility and if not, which substance(s) should be tested. The Panel concluded that HCA testing would be adequate for demonstrating intralaboratory reproducibility and would allow an effective comparison to the traditional LLNA.

The Panel was asked to comment on whether the required one-week interval between independent tests was adequate and/or appropriate. The Panel concluded that the minimum one-week interval seemed logical and that the more important clarifying information might be the elements that define independent tests (e.g., different animal shipment, different reagents, different operator, blind testing).

The Panel was asked to comment on the appropriateness of the criteria for acceptability (generally within 0.5x to 2.0x EC3 for HCA), or to describe another criteria and explain the basis for their recommendation. The Panel concluded that the criteria for acceptability appeared to be appropriate because the statistical multiple comparisons issue does not exist. However, given that there is so much data and experience with HCA and the fact that only one compound is being tested (not 18), it is reasonable to evaluate reproducibility using the mean ± 3 standard deviations rather than the 0.5x to 2.0x EC3 range to account for a single comparison (see **Section 6.7** regarding data transformation recommendations). The Panel noted that historical control data using HCA in the same vehicle could be used to demonstrate adequate intralaboratory reproducibility.

The draft ICCVAM LLNA Performance Standards state that acceptable interlaboratory reproducibility will be indicated by each of three laboratories obtaining ECt values for HCA and DNCB from a single experiment that are generally within 0.5x to 2.0x (5% to 20% and 0.025 to 0.1%, respectively) of the mean historical EC3 concentration (10% and 0.05%, respectively) obtained for these two substances in the traditional LLNA. The Panel was asked if they considered the single experiment per substance in each laboratory to be adequate. The Panel concluded that, considering the overall validation plan for a given laboratory, multiple experiments (n=3) within each laboratory should be conducted.

The Panel was asked if they considered testing HCA and DNCB to be adequate for demonstrating interlaboratory reproducibility and if not, which substance(s) should be tested. The Panel concluded that, since there is a great deal of data and experience with HCA and DNCB, and many laboratories have successfully worked with them in the traditional LLNA, they should be considered adequate for this purpose.

The Panel was asked if they considered the criteria for acceptability to be appropriate. The Panel concluded that the criteria for acceptability (i.e., generally within 0.5x to 2.0x ECt for HCA and DNCB) appeared to be appropriate because the statistical multiple comparisons problem was relatively minor given that only two substances are being tested. However, given that there is so much data and experience with HCA and DNCB and the fact that two compounds are being tested (not 18), it is reasonable to evaluate reproducibility using the mean ± 4.5 standard deviations to account for statistical multiple comparisons (see **Section 6.7** regarding data transformation recommendations). The Panel also noted that historical control data using HCA and DNCB in the same respective vehicle could be used to demonstrate adequate interlaboratory reproducibility.

6.6 Summary

The Panel was asked what criteria should be used to evaluate the equivalence of a radioactive or non-radioactive LLNA method to the traditional LLNA, if one were proposed with a "major" change, as defined in the draft ICCVAM LLNA Performance Standards (e.g., different mouse strain or use of male mice, change in the schedule for test article administration, change in schedule for lymph node excision, etc.). The Panel commented that the idea of what is a "major" and a "minor" change should be re-considered (refer to Question 2 regarding essential test components). The final version of the performance standards should be adequate to evaluate any protocol modifications.

The Panel was asked if a new set of performance standards would be required for a modified version of the LLNA that incorporated one or more "major" protocol changes. Based on the above response, the Panel concluded that a new set would not be required.

The Panel was asked to comment on how many reference substances might be considered adequate for evaluating the validity of a modified version of the LLNA with a "major" protocol change; specifically, if the 18 minimum reference substances in the draft ICCVAM LLNA Performance Standards would be sufficient. The Panel concluded that additional substances should not be considered necessary. However, since eight of the proposed sensitizers had

limited data (i.e., EC3 values based on ≤2 LLNA studies), other substances with more robust data should be considered as replacements. Furthermore, if the goal is to evaluate a specific applicability domain, additional test substances might be needed.

The Panel was asked to comment, regardless of the number of reference substances, whether the alternative LLNA with a "major" change should be required to obtain the same "call" (and potency for sensitizers) as the traditional LLNA for the 18 minimum reference substances in the draft ICCVAM LLNA Performance Standards. The Panel reiterated that an assay that is equivalent to the traditional LLNA is desired, but with the small number of reference substances available, clearly establishing equivalence will be extremely difficult. They also reiterated their concern regarding the statistical multiple comparisons problem.

For use in hazard identification, a proposed modified LLNA should be evaluated with all 22 substances (including false negatives and false positives) and accuracy statistics calculated so that accuracies can be compared between the modified test method and the traditional LLNA. To the extent possible, rationale for any discordant results should be provided. However, the most potent sensitizers (e.g., DNCB) should always be identifiable. Considerable weight should be given to the balance between animal welfare and human safety when considering the adequacy of test method accuracy.

The Panel was asked to identify any additional specific substances that should be used. The Panel concluded that while additional substances should not be needed, it would be useful to identify replacements for the eight proposed sensitizers with limited test data. If the goal is to evaluate a specific applicability domain, additional test substances might be needed.

6.7 Additional Statistical Comments

During the evaluation of the draft ICCVAM LLNA Performance Standards, the Panel noted a number of statistical issues that should be addressed. They suggested that in order to achieve a normal distribution of the data and to reduce differences between groups, a suitable variance stabilizing transformation (e.g., log transformation, square root transformation) should be applied in all statistical analyses and in reporting summary standard deviations. The Panel also suggested that there should be a more rigorous evaluation of what would be considered an appropriate range of ECt values to include as a requirement. This would be a statistical evaluation that takes into consideration the variability of ECt values generated among the sensitizers included on the performance standards reference substances list and the statistical multiple comparisons problem and the fact that sample sizes that are less than 30 invalidate statistics based on the normal distribution (Young 2007).

Furthermore, bioequivalence models have been developed (Berger and Hsu 1996) and should be applied to the LLNA. Probability values can be used as descriptive statistics and as such provide a summary measure of weight-of-evidence that would be useful for comparison of performance standards across test methods. In this context, it would be informative to have statistical tests of data generated for these purposes. A test of concordance for measuring the accuracy of classification should be done.

Intralaboratory tests should include analysis of variance (ANOVA)-like tests with a test for no trend, with the null hypothesis being that there is a difference and the alternative being that the difference is bioequivalent. Interlaboratory tests should include ANOVA-like tests with the null hypothesis being that there is a difference and the alternative hypothesis being that the difference is bioequivalent. The reliability tests require "bioequivalence" to be defined (i.e., what is acceptable to be considered equivalent).

It is not known whether these specific statistical tests can be identified in the literature or if they need to be developed. If they do need to be developed, this should be given a priority. Prior to running reliability studies, these statistical methods should be used to determine the appropriate number of substances and the number of times each substance needs to be tested within and among laboratories in the study design (see also ISO 5725 [ISO 1994] and ASTM Standard E691 [ASTM 2005]). The power for the traditional LLNA should be established for comparison purposes.

7.0 Use of the LLNA for Potency Determinations

7.1 Comments on the Draft BRD for Completeness, Errors, and Omissions

The Panel was asked if there were any errors in the draft BRD on the use of the LLNA for potency determinations that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. The Panel noted alternative analyses that would better help evaluate the use of the traditional LLNA for skin-sensitization potency (see the discussion of the categorization scheme in **Section 7.2** and the discussion of future studies in **Section 7.3**).

7.2 Comments on the Validation Status of the Traditional LLNA to Determine Skin-Sensitization Potency

7.2.1 Substances Used for the Validation Studies

The Panel was asked to consider whether the validation status of the traditional LLNA for potency categorization (i.e., "strong" vs. "weak" sensitizers) has been adequately characterized, and if the traditional LLNA is sufficiently accurate and reliable to be used as a stand-alone assay for characterizing the potency of sensitizing substances, based on the comparison to human and guinea pig responses. The Panel agreed that the LLNA database of 170 substances with comparative guinea pig (i.e., Guinea Pig Maximization Test or Buehler Test) and/or human data (i.e., Human Maximization Test [HMT] and/or Human Repeat Insult Patch Test [HRIPT], but not human clinical observations) is sufficient in number and well balanced for this evaluation. The database included 112 substances (97 sensitizers, 15 non-sensitizers) with comparative human data and 105 substances (52 sensitizers, 53 non-sensitizers) with comparative guinea pig data. Known contact sensitizers of public health concern from various chemical groups are included. The Panel further agreed that these substances were representative of a sufficient range of chemical classes and physical chemical properties so that it would be applicable to the types of chemicals and products typically tested for skin-sensitization potential.

While coding of chemicals to reduce bias is recommended for validation studies, this evaluation was based on a retrospective evaluation of existing data, most of which were generated using chemicals that were not coded. The Panel was asked whether the lack of coding of test substances adversely impacted or biased the current evaluation. Given the nature of the studies (i.e., the testing was not conducted to demonstrate the ability of the LLNA to be used for potency characterization), the Panel stated that the lack of coding likely had no impact on the current evaluation.

For some substances tested for sensitization using the traditional LLNA, it was not possible to determine whether the data were generated using pooled or individual animal lymph node samples within a dose group (the former allowed in OECD TG 429 [OECD 2002]; the latter as recommended in the ICCVAM 2001 protocol and required in the EPA 2003 skin sensitization test guideline). Cockshott et al. (2006) reported that using individual animal data allowed for technical problems during an experiment and outlier animals within a dose group to be

identified. Considering this, the Panel was asked whether the analysis of the performance of the traditional LLNA for potency determinations should be limited to data from studies that can be confirmed as using individual animal data collection procedures.

A majority of the Panel agreed that, ideally, future traditional LLNA potency determinations should be based on data from studies that use individual data collection procedures, as this would allow for the identification of outliers that might skew the average group stimulation index. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable.

7.2.2 Test Method Accuracy

The Panel was further asked what impact the inclusion of pooled animal data might have on the accuracy analysis included in Section 6.0 of the draft ICCVAM LLNA potency BRD. With regard to this retrospective dataset, the Panel agreed that pooled data should not be excluded from the current analysis to assess potency determinations for the traditional LLNA. The Panel stated that it is impossible to assess the impact of using pooled data without a separate analysis of the ability of the traditional LLNA to be used for characterizing skin-sensitization potency using pooled vs. individual data, which the Panel recommended be done (see the discussion of future studies in **Section 7.3**).

A minority opinion from Dr. Dagmar Jírová stated that, since OECD TG 429 (OECD 2002) allows the use of both pooled and individual animal data, the analysis that includes both types of data is appropriate. Even with the diversity of data sources (the vehicle is not known for 43% of substances tested in the traditional LLNA; human data were obtained by different, even undefined methods, etc.), the outcome of the evaluation was good, which documents the strength and robustness of the traditional LLNA.

The Panel was asked whether the correct classification, as well as the over- and under-classification, rates of the traditional LLNA for sensitization potency determinations had been adequately compared and appropriately evaluated based on the corresponding human and guinea pig data (refer also to Section 6.0 of the draft ICCVAM LLNA potency BRD). The Panel agreed that the two approaches used in the draft BRD for analyzing the ability of the traditional LLNA to discriminate between strong and weak skin sensitizers were appropriate and correct. In these two approaches, the traditional LLNA was evaluated, after identifying the optimal EC3 value, for its ability to correctly classify strong and weak sensitizers as defined by human or guinea pig threshold values based on: (1) sensitizers only, and (2) sensitizers combined with false positives, false negatives, and non-sensitizers.

A minority opinion from Dr. Howard Maibach stated that the relevance of the traditional LLNA to human clinical observations has not been sufficiently determined and should be.

The accuracy analysis (see Section 6.0 of draft ICCVAM BRD) focuses on a proposed two-level categorization scheme (weak sensitizers vs. strong sensitizers) for both human and guinea pig data. The Panel was asked whether this was an appropriate categorization scheme, or if

other categorization schemes should be considered. The Panel agreed that the two-level categorization scheme was appropriate, especially considering the fact that, for human situations, risk assessment should be performed, and therefore more categories are not needed. Even a weak sensitizer under heavy exposure and individual circumstances may reach a comparable risk level as a strong sensitizer under conditions of low exposure.

A minority opinion from Drs. Raymond Pieters and Michael Woolhiser recommended the addition of at least a moderate category since certain compounds will always be on the border between weak and strong. Dr. Pieters specifically recommended the categorization scheme of Kimber et al. (2003), which is based on five categories if non-sensitizers are included.

Of the two human threshold concentrations that are proposed in this two-category categorization scheme (i.e., <250 μg/cm^2 or <500 μg/cm^2), the Panel was asked which threshold was the most appropriate for categorizing sensitizing substances as strong vs. weak for humans, or if another threshold was more appropriate for this purpose. The Panel noted that this validation was based on comparison to guinea pig and HMT/HRIPT information. These data relate only to induction and do not permit an assessment of risk in humans for elicitation.

For the data provided, the Panel concluded that the best results were obtained using the decision criterion of 250 μg/cm^2 and the corresponding optimal traditional LLNA EC3 value of 9.4%. Using this cut-off when traditional LLNA false negative and false positive substances are included in the analysis, in addition to sensitizers in both the traditional LLNA and in humans using the HRIPT and/or HMT, correct classification of strong sensitizers was 79% and underclassification was 21%. Underclassification of substances in this context means classification as weak instead of strong sensitizers (i.e., they are not missed as sensitizers regarding the labeling and safety of consumers). The Panel stated that more data are needed to determine if another threshold is more appropriate.

When the potency categorization analysis was based on sensitizers only, the guinea pig tests predicted weak sensitizers with higher accuracy than did the LLNA (89% vs. 75% for the 250 μg/cm^2 cutoff and 83% vs. 60% for the 500 μg/cm^2 cutoff), which is logical because the guinea pig test methodology involves all phases of the sensitization process and usually involves adjuvants. However, the guinea pig tests were less accurate for the prediction of strong sensitizers compared to LLNA (48% vs. 71% for the 250 μg/cm^2 cutoff and 42% vs. 63% for the 500 μg/cm^2 cutoff), which represents a higher risk for consumers. For the protection of public health, it is more important to correctly identify strong sensitizers than weak sensitizers.

The Panel was asked whether the draft BRD adequately characterized the usefulness and limitations of the LLNA for potency categorizations. If not, the Panel was asked what additions or changes should be made to the description of usefulness and limitations in the draft BRD. The Panel stated that additional evaluations should be conducted to determine the impact on potency categorization if the human threshold data are evaluated differently (e.g., alternative lowest observed effect level [LOEL] safety factors other than 10, using LOEL data only, using no observed effect level [NOEL] data only), and if this might improve the correlation between the LLNA and the human results. According to the Panel, the approach of directly comparing the LOEL values without using a safety factor compares values of similar significance in

humans and in the LLNA. In other words, the LOEL in humans describes the threshold induction area dose in humans and the EC3 value in the traditional LLNA is the threshold induction area dose and thus could be the analogous value to the human LOEL. The Panel further stated that traditional LLNA tests based on pooled or individual lymph nodes for a dose group should be evaluated independently to assess the impact of using pooled data on the accuracy analysis for skin-sensitization potency. Finally, the Panel stated that the effect of different vehicles should be recognized as a limitation in the data analysis given the demonstrated variability of results.

7.2.3 Test Method Reliability

The Panel was asked whether the reliability (e.g., intralaboratory repeatability, intra- and inter-laboratory reproducibility) of the traditional LLNA for potency determinations had been adequately evaluated. If not, the Panel was asked what other analyses should be performed. Similar to their recommendations for test method accuracy, the Panel stated that additional evaluations of reliability should be conducted based on using different approaches for human threshold data (e.g., using alternative LOEL safety factors other than 10, using LOEL data only, using NOEL data only). The Panel further stated that the reliability of LLNA based on using pooled or individual animal data should be evaluated independently. Finally, the Panel stated that the effect of different vehicles should be recognized as a limitation in the data analysis, as a source of increased variability.

7.2.4 Data Quality

It was not possible to determine whether or not all studies included in the draft LLNA potency BRD had been conducted in accordance with GLP guidelines, nor was it possible to obtain the results of GLP audits for all studies determined to be GLP-compliant. The Panel was asked to discuss what impact this might have on the evaluation of the LLNA for potency determinations and whether any of the non-GLP studies should be excluded from the analyses. The Panel concluded that it was important to note if the data were obtained from studies conducted according to international GLP guidelines, as ideally this should be the case. However, the Panel concluded that data from studies that could not be confirmed as being GLP-compliant but that were from peer-reviewed literature or other sources with high quality laboratory management practices were still appropriate to include in this retrospective analysis.

As described in the draft BRD, original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Considering this, the Panel was asked whether the results of these studies (all of which are currently included) be excluded from any of the performance analyses. The Panel considered the data to have been generated by repeatedly published and reliable laboratories and therefore did not question the adequacy/quality of the retrospective data analysis. Thus, although data should be checked when available, exclusion of data was not deemed necessary, in this case.

7.2.5 Consideration of All Available Data and Relevant Information

Based on the draft BRD, the Panel was asked whether all the relevant data identified in published or unpublished studies conducted using the traditional LLNA had been adequately considered. If not, the Panel was asked what other studies should be considered. The Panel recommended that the LOELs from Akkan et al. (2003) be used instead of the DSA_{05} values from Schneider and Akkan (2004) in all of the potency analyses. A minority opinion by Dr. Thomas Gebel stated that it was acceptable to use the DSA_{05} values from Akkan et al. (2003) as LOEL values in the evaluation. Dr. Gebel mentioned that the DSA_{05} value is a LOEL value adjusted to 5% incidence of induction. Akkan et al. (2003) used the DSA_{05} value to correct for different human studies leading to different inductions. Dr. Gebel further stated that as the DSA_{05} is corrected for an induction rate of 5%, it would be better to compare with the traditional LLNA EC3 than to use the default uncorrected LOEL.

7.3 Comments on the Draft ICCVAM Test Method Recommendations for the Use of the LLNA for Potency Determination

7.3.1 Test Method Usefulness and Limitations

With regard to the use of the LLNA for potency categorization (i.e., strong vs. weak sensitizers), the ICCVAM draft recommendation is that the traditional LLNA should not be considered as a stand-alone test method for predicting sensitization potency, but must instead be used as part of a weight-of-evidence evaluation to discriminate between strong and weak sensitizers. This is based on the fact that, although there is a significant positive correlation between traditional LLNA EC3 values and human sensitization threshold doses, this correlation is not strong [see detailed discussion in the draft ICCVAM recommendations]. The Panel agreed that the traditional LLNA should not be considered a stand-alone assay for categorization of skin-sensitization potency, but it could be used in a weight-of-evidence evaluation (e.g., along with quantitative structure-activity relationship [QSAR], peptide reactivity, human evidence) to discriminate between strong and weak sensitizers. The Panel further stated that there are additional studies proposed that may provide a better correlation and improve prediction of potency categorization (see the discussion of future studies below).

A minority opinion from Drs. Thomas Gebel and Dagmar Jírová stated that there is a significant positive correlation between EC3 values and human threshold values. It is likely that limitations in estimating human threshold values and the inclusion of human NOEL values in the current evaluation contributed negatively to the resulting R^2 value of 0.405 (when LLNA EC3 data vs. human threshold data were compared, see Table 6-2 of the draft ICCVAM BRD). Thus, the R^2 value may improve when the additional analyses that have been suggested by the Panel are conducted.

The Panel stated that the effect of different vehicles should be recognized as a limitation in the data analysis and a likely source of within and between laboratory variability.

7.3.2 Test Method Protocol

The Panel was asked whether the ICCVAM-recommended LLNA protocol (ICCVAM 1999; EPA 2003) should be used when generating data that will or might be considered for sensitization potency categorization decisions. The Panel agreed that this protocol should be used. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable.

The Panel was asked whether the relevant testing guidelines for the traditional LLNA should be updated to include the calculation of an EC3 value. The Panel agreed with this recommendation. The calculation of an EC3 value is briefly described in the draft ICCVAM LLNA Performance Standards for specific situations with references to Basketter et al. (2000) and Ryan et al. (2007).

7.3.3 Future Studies

The Panel was asked whether the available data support the ICCVAM draft recommendations for the traditional LLNA in terms of the proposed future studies. The Panel agreed and concluded that more data are needed to determine the optimal threshold in humans for distinguishing between strong and weak sensitizers. However, the Panel discouraged conducting new animal studies unless it was likely that results from such studies would lead to an overall reduction in animal use. The Panel stated further that the traditional LLNA appears to be a robust rodent assay for the quantification of the induction of cell-mediated immunity. Thus, use of the traditional LLNA for potency determination can be used in conjunction with QSAR information, guinea pig assays, HRIPT/HMT, and the quantitative data of elicitation and frequency of positive response in humans in a weight-of-evidence approach. The Panel further stated that additional evaluations should be conducted to determine the impact on potency categorization if the human threshold data are evaluated differently (e.g., alternative LOEL safety factors other than 10, using LOEL data only, using NOEL data only). This might improve the correlation between LLNA and human data. The Panel further stated that LLNA tests based on pooled or individual animal data should be evaluated independently to assess the impact of using pooled data on the accuracy for determining skin-sensitization potency.

The Panel recommended a statistical analysis to determine where an appropriate cutoff value between weak or strong sensitizers might be best defined for traditional LLNA data. For example, receiver operating characteristic curves could be used to identify the optimum cut-off for determining the difference between weak and strong sensitizers.

Finally, the Panel stated that the effect of different vehicles should be recognized as a limitation in the current data analysis, that this was a source of variability within and between laboratories, and that its impact should be considered in future analyses.

8.0 References

Akkan Z, Kalberlah F, Oltmanns J, Schneider K. 2004. Beurteilung der Wirkst€arke hautsensibilisierender Chemikalien anhand des Local Lymph Node Assay (LLNA) [in German]. Bremerhaven, Germany: Federal Institute for Occupational Safety and Health.

ASTM. 2005. ASTM Standard E691-05 Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method In: Annual Book of ASTM Standards. West Conshohocken: ASTM International. www.astm.org.

Basketter DA, Blaikie L, Dearman RJ, Kimber I, Ryan CA, Gerberick GF, et al. 2000. Use of the local lymph node assay for the estimation of relative contact allergenic potency. Contact Dermatitis 42:344-348.

Berger RL, Hsu JC. 1996. Bioequivalence trials, intersection-union tests and equivalence confidence sets. Stat Sci 11:283-319.

Cockshott A, Evans P, Ryan CA, Gerberick GF, Betts CJ, Dearman RJ, et al. 2006. The local lymph node assay in practice: a current regulatory perspective. Hum Exp Toxicol 25:387-394.

Dean JH, Twerdok LE, Tice RR, Sailstad DM, Hattan DG, Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay: II. Conclusions and recommendations of an independent scientific peer review panel. Regul Toxicol Pharmacol 34:258-273.

De Jong WH, van Och FM, Den Hartog Jager CF, Spiekstra SW, Slob W, Vandebriel RJ, et al. 2002. Ranking of allergenic potency of rubber chemicals in a modified local lymph node assay. Toxicol Sci 66:226–232.

EPA. 2003. Health Effects Test Guideline OPPTS 870.2600. Skin Sensitization. Research Triangle Park, NC: Environmental Protection Agency.

Gaspari AA. 2007. Mechanism of action and other potential roles of an immune response modifier. Cutis 79(4 SUPPL.):36-45.

Gerberick GF, Ryan CA, Kern PS, Dearman RJ, Kimber I, Patlewicz GY, et al. 2004. A chemical dataset for evaluation of alternative approaches to skin-sensitization testing. Contact Dermatitis 50:274-288.

Gerberick GF, Vassallo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin JP. 2007. Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. Toxicol Sci 97:417-27.

ICCVAM. 1999. The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals/Compounds. The Results of an Independent Peer Review Evaluation Coordinated by the Interagency Coordinating Committee on the Validation of Alternative Toxicological Methods (ICCVAM) and the National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication No. 99-4494. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No: 03-4508. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

ISO. 1994. ISO 5725. Accuracy (trueness and precision) of measurement methods and results. Geneva: International Organization for Standardization.

ISO. 2002. ISO 10993-10: Tests for Irritation and Sensitization. Geneva: International Organization for Standardization.

Jowsey IR, Clapp CJ, Safford R, Gibbons BT, Basketter D. 2008. The impact of vehicle on the relative potency of skin sensitising chemicals in the local lymph node assay. In Press.

Kimber I, Basketter DA, Butler M, Gamer A, Garrigue JL, Newsome C, et al. 2003. Classification of allergens according to potency: proposals. Food Chem Toxicol 41:1799-1809.

Kimber I, Dearman RJ, Betts CJ, Gerberick GF, Ryan CA, Kern PS, et al. 2006. The local lymph node assay and skin sensitization: a cut-down screen to reduce animal requirements? Contact Dermatitis 54:181-185.

Manetz TS, Meade BJ. 1999. Development of a flow cytometry assay for the identification and differentiation of chemicals with the potential to elicit irritation, IgE-mediated, or T cell-mediated hypersensitivity responses. Toxicol Sci. 48:206-17.

OECD. 2002. Guideline for the Testing of Chemicals 429 - Skin Sensitisation: Local Lymph Node Assay. Paris: OECD.

Ryan CA, Chaney JG, Gerberick GF, Kern PS, Dearman RJ, Kimber I, Basketter DA. 2007. Extrapolating local lymph node assay EC3 values to estimate relative sensitizing potency. Cutan Ocul Toxicol 26:135-145.

Schneider K, Akkan Z. 2004. Quantitative relationship between the local lymph node assay and human skin sensitization assays. Regul Toxicol Pharmacol 39:245-55.

Sailstad DM, Hattan D, Hill RN, Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay: I. The ICCVAM review process. Regul Toxicol Pharmacol 34:249-257.

Takeyoshi M. 2007. Promising endpoint and assay performance of non-radioisotopic local lymph node assay based on BrdU incorporation. Sixth World Congress on Alternatives and Animal Use in the Life Sciences. August 21-25, 2007. Tokyo, Japan.

Van Hoogstraten IM, Boos C, Boden D, Von Blomberg ME, Scheper RJ, Kraal G. 1993. Oral induction of tolerance to nickel sensitization in mice. J Invest Dermatol 101:26-31.

Van Och FMM, Slob W, De Jong WH, Vandebriel RJ, Van Loveren H. 2000. A quantitative method for assessing the sensitizing potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of the uncertainty margins. Toxicology 146:49-59.

Varani J, Perone P, Spahlinger DM, Singer LM, Diegel KL, Bobrowski WF, et al. 2008. Human skin in organ culture and human skin cells (keratinocytes and fibroblasts) in monolayer culture for assessment of chemically induced skin damage. Toxicol Pathol 35:693-701.

Young SS. 2007. Gaming the system: chaos from multiple testing. Institute of Mathematical Statistics Bulletin 36:13

[This Page Intentionally Left Blank]

.

Annex I

Peer Review Panel Member Biosketches

[This Page Intentionally Left Blank]

Panel Member Biosketches

Nathalie Alépée, Ph.D.

Dr. Alépée performed research leading to a Ph.D. in Medical Virology and Microbiology at the Centre National de la Recherché Scientifique research institute, Gif sur Yvette, France. She is currently the Global Pfizer Leader for photosafety, including the global portfolio support and Associate Research Fellow in Investigative Toxicology, at Pfizer Global Research and Development, Amboise, France. As a laboratory manager in the Molecular and Cellular Toxicology Group with Pfizer, she implemented the local lymph node assay (LLNA) in the laboratory. She serves on the European Centre for the Validation of Alternative Methods (ECVAM) Scientific Advisory Committee (ESAC), representing the European Federation of Pharmaceutical Industries Associations (EFPIA). She is also the Pfizer representative to the European Partnership on Alternative to Animal Testing (EPAA), in two working groups; Identification of Opportunities, Including R&D (working group 2), and Validation and Acceptance (working group 5). She served as a peer reviewer of the reduced LLNA test protocol and prediction model for ESAC in 2007 and has been designated as an ESAC peer reviewer for ECVAM's performance standards for the standard LLNA.

Anne Marie Api, Ph.D.

Dr. Api received a Ph.D. from Aston University in Birmingham, England and is currently Vice President of Human Health Sciences at the Research Institute for Fragrance Materials (RIFM), as well as the Scientific Director. She is responsible for the human health scientific program, and the investigation and initiation of new research and testing projects for RIFM. She is also Adjunct Assistant Professor at the University of Medicine and Dentistry of New Jersey. She is a member of 10 professional organizations, including the American Contact Dermatitis Society, the European Society of Contact Dermatitis, and the Society of Investigative Dermatology. She participated in the World Health Organization (WHO) International Workshop in Skin Sensitization in Chemical Risk Assessment held in Berlin, Germany in 2006. She is author of over 100 publications and presentations relevant to dermatology and dermatotoxicology.

Nancy Flournoy, M.S., Ph.D.

Dr. Flournoy received a M.S. degree in Biostatistics from the University of California at Los Angeles, and a Ph.D. in Biomathematics from the University of Washington. She is Professor and Chair of the Department of Statistics at the University of Missouri-Columbia. Her research interests include adaptive designs, bioinformatics, chemometrics, clinical trials, and environmetrics. She has an extensive list of edited volumes and papers on statistical theory, statistical genetics and immunology, epidemiology in immune suppressed subjects, clinical trials for prevention and treatment of viral infection, transplantation biology and its effects on digestion, lungs, eyes, mouth, and central nervous system, optimization of statistical processing, and additional papers, interviews, and technical reports. She has editorial responsibilities for numerous statistical journals, serves on numerous advisory boards, and nominating committees. She is a member and past Chair of the Council of Sections of the American Statistical Association, and served in various other statistical, medical and toxicological societies or programs as Chair or as a member of the Board of Directors. She is a former member of the Scientific Advisory Committee on Alternative Toxicological

Methods. She also served on the Expert Panels for the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) that evaluated the Revised Up-and-Down Procedure; the Current Validation Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants; and Five *In Vitro* Pyrogen Test Methods.

Thomas Gebel, Ph.D.

Dr. Gebel received a Ph.D. in Toxicology from the University of Mainz and is certified as a toxicologist by the German Society of Toxicology. His scientific interests are in biomonitoring, genetic toxicology, environmental hygiene, and occupational toxicology. He has published over 40 papers in peer-reviewed scientific journals. He is employed by the German Federal Institute for Occupational Safety and Health, and is an Associate Professor at the University of Goettingen. Dr. Gebel is currently a member of the Organisation for Economic Co-operation and Development (OECD) Globally Harmonized System of Classification and Labeling of Chemicals (GHS) expert group on sensitization and head of the German advisory committee on classification and labeling of existing substances and biocides. Dr. Gebel also is head of the German Delegations to the United Nations Economic and Social Council Sub-Committee of Experts on the GHS, and to the OECD Task Force on Harmonisation of Classification and Labeling. He participated in the WHO International Workshop in Skin Sensitization in Chemical Risk Assessment held in Berlin, Germany in 2006.

Sidney Green Ph.D., F.A.T.S.

Dr. Green received a Ph.D. in Biochemical Pharmacology from Howard University. His research interests include toxicology, mutagenic assay systems, and alternatives to animals in toxicology. He is currently Graduate Professor of Pharmacology at Howard University and a faculty member at the Centers for Alternatives to Animal Testing at the Johns Hopkins University School of Public Health. Previously, he has been Director of the Department of Toxicology at Covance Laboratories Inc. and the Director of the Division of Toxicological Research at the U.S. Food and Drug Administration (FDA). Dr. Green is a Fellow of the Academy of Toxicological Sciences (F.A.T.S.). He has served on numerous expert panels and committees. He was a participant in an International Workshop organized by ICCVAM and NICEATM on *In Vitro* Methods for Assessing Acute Systemic Toxicity in 2000. He served on the ICCVAM/NICEATM Expert Panels that evaluated the Corrositex® Test Method for Assessing Dermal Corrosivity Potential of Chemicals, and *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants. He is a former member of the ICCVAM Advisory Committee on Alternative Toxicological Methods (ACATM) and of SACATM. He has authored over 60 publications for peer-reviewed journals.

Kim Headrick, B.Admin., B.Sc.

Kim Headrick received Bachelor of Administration and B.Sc. degrees from the University of Ottawa, Canada. She is currently International Harmonization and Senior Policy Advisor for Health Canada, and Chair of the UN Sub-Committee of Experts on GHS. She manages the overall strategy for the implementation of the GHS in Canada. She was awarded the Queen Elizabeth Commemorative Golden Jubilee Medal in 2002, which focuses on the achievements of people who, over the past 50 years, have created the Canada of today. She is

a member of the OECD Task Force on Harmonization of Classification and Labelling and the OECD Expert Group Meeting on Sensitization Hazards.

Dagmar Jírová, M.D., Ph.D.

Dr. Jírová received a Ph.D. from the Medical Faculty of Hygiene at Charles University in Prague. She is currently the Head of the Reference Center for Cosmetics, and Head of National Reference Laboratory for Experimental Immunotoxicology at the National Institute of Public Health in the Czech Republic. Her main responsibilities include safety assessment of consumer products, particularly cosmetics and their ingredients, performance of toxicological methods *in vivo* in animals, human patch testing for local toxicity assessment, and introduction of *in vitro* techniques for screening of toxicological endpoints using cell and tissue cultures. She represents the Czech Republic in the Standing Committee on Cosmetics of the European Commission. She is an ESAC-ECVAM member and was involved in Peer Review Panel for Skin Irritation Validation Study and LLNA test protocol and prediction model. She is author of more than 100 publications and presentations relevant to dermatotoxicology including recent presentation at the 6th World Congress on Alternatives & Animal Use in the Life Sciences, held in Tokyo, 2007, titled "Comparison of Human Skin Irritation and Photoirritation Patch Test Data with Cellular *in vitro* Assays and Animal *in vivo* data."

David Lovell, Ph.D., B.Sc. (Hons), F.S.S., FIBiol, CStat, CBiol

Dr. Lovell received a Ph.D. from the Department of Human Genetics and Biometry, University College, London. He is currently Reader in Medical Statistics at the Postgraduate Medical School at the University of Surrey. Previously, he was Associate Director and Head of Biostatistics support to Clinical Pharmacogenomics at Pfizer Global Research and Development in Sandwich, Kent providing data management and statistical support to pharmacogenetics and genomics. He joined Pfizer in 1999 as the Biometrics Head of Clinical Pharmacogenetics. Before joining Pfizer, Dr. Lovell was the Head of the Science Division at BIBRA International, Carshalton, which included Molecular Biology, Genetic Toxicology, Biostatistics and Computer Services. At BIBRA, Dr. Lovell managed the statistical and computing group providing specialized statistical support to BIBRA's Clinical Unit and contract research work. He conducted and managed research programs on genetics, statistics and quantitative risk assessment for the European Union (EU) and U.K. Government Departments. His research interests at BIBRA were in the use of mathematical and statistical methods together with genetic models in the understanding of toxicological mechanisms and risk assessment problems. Dr. Lovell had previously been a Senior Research Officer with the U.K. Medical Research Council (MRC) Experimental Embryology and Teratology Unit, a visiting Postdoctoral Fellow at the National Institute of Environmental Health Sciences (NIEHS) in North Carolina, U.S., a Geneticist at the MRC Laboratories, Carshalton, and a Research Assistant in Cytogenetics at Birmingham University. He has acted as a consultant to a number of organizations, has considerable experience of working with Regulatory Authorities, has many publications related to his work and has wide experience of making presentations to a wide range of audiences. He is a member of the U.K. Government's advisory Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment (COM) and the Independent Scientific Advisory Committee for Medicines and Healthcare Products Regulatory Agency database research. He served on the NICEATM-ICCVAM Expert Panels

that evaluated the Frog Embryo Teratogenesis Assay - Xenopus, *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants, and Five *In Vitro* Pyrogen Test Methods.

Michael Luster, Ph.D.

Dr. Luster received a Ph.D. in Immunology from Loyola University of Chicago. He was formerly Chief, Toxicology and Molecular Biology Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health (NIOSH), and currently serves as a senior advisor to the Director of the Health Effects Laboratories and the staff of Toxicology and Molecular Biology Branch at NIOSH. Program areas include neuroscience, dermatology, molecular carcinogenesis, molecular epidemiology, molecular toxicology, molecular epidemiology, and inflammation/immunotoxicology. In addition, Dr. Luster conducts basic and applied research in immunotoxicology including its application in risk assessment. Current research activities include molecular epidemiology studies of genetic polymorphism involved in workplace-related diseases and experimental studies involving occupational allergic rhinitis. Dr. Luster is also working with various staff at the U.S. Environmental Protection Agency (EPA) through the Risk Assessment Forum to develop immunotoxicity testing guidelines. He also directed two studies for the NTP on the Toxicology and the Carcinogenesis of Promethazine and Ortho-phenylphenol, in 1990 and 1986, respectively. He is a co-author of over 300 publications in peer-reviewed journals.

Howard Maibach, M.D.

Dr. Maibach received an M.D. from Tulane University. He is currently a professor in the Department of Dermatology at the University of California, San Francisco (USCF), where he is also Chief of the Occupational Dermatology Clinic. In his 35 years at UCSF, Dr. Maibach has written and lectured extensively on dermatotoxicology and dermatopharmacology. His current research programs include defining the chemical-biologic faces of irritant dermatitis and the study of percutaneous penetration. Dr. Maibach served on the 1998 ICCVAM Peer Panel that evaluated the Murine LLNA. Dr. Maibach has been on the editorial boards of over 30 scientific journals and is a member of 19 professional societies including the American Academy of Dermatology, San Francisco Dermatological Society, and the International Commission on Occupational Health. He has co-authored over 1500 publications related to dermatology.

James McDougal, Ph.D., F.A.T.S.

Dr. McDougal earned a Ph.D. in Pharmacology/Toxicology at the University of Arizona. He is currently Professor and Director of Toxicology Research in the Department of Pharmacology and Toxicology at Wright State University's Boonshoft School of Medicine. Prior to his appointment at Wright State, he worked in the Air Force toxicology research organization for about 17 years. He has active skin research programs related to dermal pharmacokinetics, molecular biology of skin irritation, dermal risk assessment, and biologically-based mathematical modeling. He has served on many national committees, published more than 75 manuscripts, and consults for a wide variety of government and industry organizations. Dr. McDougal is a member of the National Academy of Sciences (National Research Council) Committee on Toxicology and the American Congress of Governmental Industrial Hygienists Threshold Limit Value Committee for Chemical substances. Dr. McDougal is also past president of the Dermal Toxicology Specialty Section of the Society of Toxicology.

Michael Olson, Ph.D., A.T.S.

Dr. Olson received a Ph.D. in Toxicology from the University of Arkansas for Medical Sciences, with dissertation research conducted at the FDA National Center for Toxicological Research. Following graduate training, he served as NIEHS National Research Service Award Post-doctoral Fellow in the Department of Pharmacology, School of Medicine - University of North Carolina. Currently he is Director, Occupational Toxicology, Corporate Environment Health and Safety for GlaxoSmithKline. Dr. Olson is a Fellow of the Academy of Toxicological Sciences (A.T.S.). His research interests include mechanisms of chemically-induced toxicity; genetic toxicity; xenobiotic metabolism; alternative methods in toxicology; hazard evaluation, risk assessment, and communication. Dr. Olson has authored a number of peer-reviewed manuscripts and book chapters in these areas as well as preparing many occupational health effects reviews for pharmaceutical active ingredients, isolated intermediates, and associated chemicals. He has served as an editorial board member and *ad hoc* referee for numerous toxicology and biosciences journals. In addition, he has worked as a Visiting Scientist, EPA, as well as advisor to EPA Risk Assessment Forum, U.S. National Institutes of Health (NIH) (Toxicology Study Section I), U.S. Air Force, Transportation Research Board, and the National Research Council - National Academy of Sciences. A member of several biomedical professional societies, Dr. Olson has served in elective and appointed positions in the Society of Toxicology, including Chairman of the Society of Toxicology (SOT) Occupational Health Specialty Section.

Raymond Pieters, Ph.D.

Dr. Pieters received a Ph.D. at Utrecht University and is currently an Associate Professor at the Institute for Risk Assessment Sciences, and Group Leader for Immunotoxicology at that institution. In 2007, he presented a paper on Development of Strategies to Assess Drug Hypersensitivity at the Congress of the European Societies of Toxicology. He was involved in the development of the Reporter Antigen Popliteal Lymph Node Assay, an assay to assess the immunomodulating potential of chemicals, which enables differentiation between immunosensitizing chemicals (sensitizers), immunostimulating chemicals (irritants), and chemicals that have no apparent immunological effects. He has published over 70 papers on sensitization and other subjects in immunotoxicology in peer-reviewed journals, including a review article, *Murine Models of Drug Hypersensitivity*, in 2005.

Jean Regal, Ph.D.

Dr. Regal received a Ph.D. in Pharmacology from the University of Minnesota. She is currently a Professor in the Department of Pharmacology, Department of Biochemistry & Molecular Biology and Associate Dean of Faculty Affairs, Medical School Duluth, University of Minnesota. Her current research is focused on respiratory allergy, especially asthma. She has served on multiple NIH review panels regarding asthma, as an immunotoxicologist in 2000 for an Institute of Medicine Committee on Health Effects Associated with Exposures Experienced during the Persian Gulf War, as well as on the 1998 ICCVAM Peer Panel that evaluated the Murine LLNA. In 2007 she served as an ad hoc reviewer for the NTP Board of Scientific Counselors for two nominations: Artificial Butter Flavoring Mixture & O-phthalaldehyde, at NIEHS. Also in 2007, she served on an NIEHS Center in Environmental Toxicology pilot project program for the University of Texas Medical Branch at Galveston. She is currently Vice-President-elect of the Immunotoxicology

Specialty Section of SOT and Associate Editor of the Journal of Immunotoxicology. Dr. Regal has authored over 50 research articles and reviews in peer-reviewed journals and holds two patents on pulmonary administration of sCR1 and other complement inhibitory proteins.

Jonathan Richmond, B.Sc. (Hons) Med.Sci., MB ChB, FRCSEd, FRMS

Dr. Richmond received a Bachelor of Science in Medical Science with Honors (B.Sc. [Hons] Med.Sci.) and Bachelor of Medicine and Bachelor of Surgery (MB ChB) degrees with Distinction in Medicine and Therapeutics from Edinburgh University. Presently, he is head of the Animals Scientific Procedures Division at the Home Office. He is a Fellow of the Royal College of Surgeons of Edinburgh (FRCSEd) and a Fellow of the Royal Society of Medicine (FRMS). Other appointments include convener of the U.K. interdepartmental group on the 3Rs, board member U.K. National Centre for the 3Rs, convener of the International Standards Organization Technical Corrigendum 194/Working Group 3 (*Biocompatibility of Medical Device Materials*), and member of related expert working groups. He is a former member of the EU Committee on Scientific and Technical Progress and past Chairman of the European Commission Technical Expert Working Group on ethical review. He served as chair of the peer review panel for the reduced LLNA test protocol and prediction model for ESAC in 2007 and has been designated as an ESAC peer reviewer for ECVAM's performance standards for the standard LLNA. He served on the ICCVAM/NICEATM Expert Panel that evaluated Five *In Vitro* Pyrogen Test Methods. He has a variety of publications in peer-reviewed journals and national and international meetings, on the principles and practice of surgery, regulation of biomedical research, principles of humane research, bioethics, and public policy.

Peter Theran, V.M.D.

Dr. Theran holds a Doctor of Veterinary Medicine degree from the University of Pennsylvania. He has had many years of experience both as a veterinary internal medicine specialist at the Massachusetts Society for Prevention of Cruelty to Animals' Angell Memorial Animal Hospital in Boston, and as the director of Boston University Medical Center's Laboratory Animal Science Center. He presently serves on a number of government committees as an animal welfare member, and is a member of the Board of Directors of the Institute for *In Vitro* Sciences in Gaithersburg, MD and Chimp Haven in Shreveport, Louisiana. He served on the NICEATM-ICCVAM Expert Panels that evaluated the *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants, and Five *In Vitro* Pyrogen Test Methods. He is a former member of ACATM and SACATM. He is presently working as a consultant.

Stephen Ullrich, Ph.D.

Dr. Ullrich received a Ph.D. in Microbiology from Georgetown University. He is currently the Dallas/Fort Worth Living Legends Professor, and Professor of Immunology at the University of Texas, M.D. Anderson Cancer Center, where he is also Associate Director, The Center for Cancer Immunology Research. He is also a member of the Animal Research Strategic Advisory Committee. He has served numerous national review committees and panels, including the 1998 ICCVAM Peer Panel that evaluated the Murine LLNA. Dr. Ullrich has authored over 75 peer-reviewed publications, over 30 invited articles, and he holds four patents in the U.S., E.U., and Australia for a UV-induced Immunosuppressive Substance. He is the past President of the American Society for Photobiology.

Michael Woolhiser, Ph.D.

Dr. Woolhiser received a Ph.D. in Pharmacology and Toxicology from the Medical College of Virginia at Virginia Commonwealth University. He is a specialist in immunotoxicology and is currently a toxicologist for the Dow Chemical Company where he serves as a Technical Leader for Immunotoxicology, and Polyurethane Business Toxicology Consultant. Dr. Woolhiser is also an Adjunct Professor at the Center for Integrative Toxicology, Michigan State University. He is a member of the Program Committee of the Society of Toxicology's Immunotoxicology Specialty Section. He has served on numerous working groups, including an LLNA Expert Working Group under the European Crop Protection Agency's Toxicology Expert Group, a European Centre for Ecotoxicology and Toxicology of Chemicals LLNA Task Force. He has authored 29 peer-reviewed publications.

Takahiko Yoshida, M.D., Ph.D.

Dr. Yoshida earned his M.D. and a Ph.D. in Medical Science from Tokai University. He is currently Professor in the Department of Health Science at Asahikawa Medical College. Prior to this appointment, he held the posts of Instructor, Assistant Professor and Associate Professor at the Tokai University School of Medicine. He has also been a Guest Researcher at NIEHS. He has also worked as an occupational physician for major Japanese corporations, including Toyota and Sony. Dr. Yoshida’s research interests include occupational health, public health, environmental health and preventative medicine. He is a member of the International Congress of Occupational Health, the Japanese Society of Hygiene, the Japanese Society of Immunotoxicology, the Japanese Society of Clinical Ecology, and the SOT.

[This Page Intentionally Left Blank]

Annex II

Questions for the Peer Review Panel

[This Page Intentionally Left Blank]

Annex F1

Questions for the Peer Review Panel: LLNA Limit Dose Procedure

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: LLNA Limit Dose Procedure

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the murine local lymph node assay (LLNA) limit dose procedure as a substitute for the traditional LLNA for assessing the allergic contact dermatitis potential of chemicals and other substances.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA limit dose procedure. You are first asked to review the information in the draft ICCVAM LLNA limit dose procedure background review document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in this BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[9]) have been appropriately addressed for the proposed use of the LLNA limit dose procedure. Adequate validation[10] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. This validation process characterizes the usefulness and limitations of a test method for its intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA limit dose procedure (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft LLNA limit dose procedure BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA limit dose procedure.

These include questions prepared by the ICCVAM IWG to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method, and/or adequate guidance for organizations that may be involved in conducting or supporting further development, standardization, and/or validation.

The overall question to consider is whether the validation status of the LLNA limit dose procedure has been adequately characterized for its intended purpose, and is it sufficiently accurate and reliable to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure when there is not a need for dose response information, in order to reduce the number of animals required for such testing.

---

[9] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[10] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

I. Questions to the Panel: Review for Errors and Omissions

1. Are there any errors or omissions of existing relevant data or information that should be included in the draft BRD?

II. Questions to the Panel: LLNA Limit Dose Procedure Draft BRD

1. For the proposed LLNA limit dose procedure, ICCVAM recommends that the number of animals used in each group should be the same as that recommended by ICCVAM for the traditional LLNA based on its 1998 evaluation of the LLNA, and that individual animal data should be collected and reported (ICCVAM, 1999). Do you agree that these are appropriate provisions for the limit dose procedure? Please explain your answer.
2. Do you consider the traditional LLNA database representative of a sufficient range of chemical classes and physical chemical properties that it would be applicable to any of the types of chemicals and products that are typically tested for skin-sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the limit dose procedure? What chemicals or products should be evaluated to fill this data gap? Please explain your answer.
3. While coding of chemicals is recommended for prospective studies, this evaluation is based on a retrospective evaluation of existing data, most of which was not generated using coded chemicals to reduce the potential for bias. Does the lack of coding of test substances adversely impact or bias the current evaluation? Please explain your answer.
4. For some substances submitted using the traditional LLNA test method, it was not possible to confirm whether the data were generated using pooled animal data for each dose group (as allowed in Test Guideline [TG] 429 of the Organisation for Economic Co-operation and Development [OECD]). ICCVAM (1999), Dean et al. (2001), and EPA (2003) recommend the use of statistical analyses to help interpret LLNA study results, which necessitates data collected at the level of the individual animal, while Cockshott et al. (2006) reported that using individual animal data allowed for technical problems during an experiment to be identified. Considering this, should the analysis of the performance of the LLNA limit dose procedure against the traditional LLNA be limited to data from studies that can be confirmed as using individual animal data collection procedures? What impact might the inclusion of pooled animal data have on the accuracy analysis of the LLNA limit procedure? Please explain your answer.
5. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA limit dose procedure been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft ICCVAM BRD)? If not, what other analyses should be performed?
6. There were five substances for which the highest concentration tested produced an SI of less than 3.0, while lower concentrations of these substances produced an SI of greater than 3.0 (see Table 6-2 of the draft ICCVAM BRD). These

substances are classified as "false negatives" compared to what was obtained in the traditional LLNA. Can you identify any characteristics associated with these or other substances that might signal that this type of abnormal dose response might occur, and therefore using the LLNA limit dose procedure would not be appropriate? Please explain your answer.

7. Does the BRD adequately characterize the usefulness and limitations of the LLNA limit dose procedure based on the accuracy analyses? If not, what additions or changes should be made to the current usefulness and limitations? Please explain your answer.
8. Is it appropriate to assume that the intra- and inter-laboratory reproducibility of the LLNA limit dose procedure and the traditional LLNA will be similar, based on the fact that they use identical protocols with the exception of the number of doses used? Do you agree? Does reducing the number of test substances dose groups from three to one reduce the reliability of the assay? Please explain your answer.
9. For some studies included in the draft BRD, it was not possible to determine whether or not they had been conducted in accordance with Good Laboratory Practice (GLP) guidelines. Original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. Considering this, should the results of these studies (all of which are currently included) be excluded from any of the performance analyses? Please explain your answer.
10. Based on the draft BRD, have all the relevant data identified in published or unpublished studies conducted using the traditional LLNA been adequately considered? If not, what other traditional LLNA data needs to be considered and how can it be obtained?

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA Limit Dose Procedure

1. Do the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed test method usefulness and limitations? If not, what recommendations would you make? Please explain your answer.

   - Should the LLNA limit dose procedure be routinely recommended for the hazard identification of skin sensitizing chemicals when potency information is not required? Please explain your answer.
   - If potency information is required, should the LLNA limit dose procedure be routinely recommended as the initial test to identify sensitizers before conducting the traditional LLNA as a way to further reduce animal use, since negative results would not require further testing? Please explain your answer.
   - Based on the existing database, there is a false negative rate of 1.6% (5/313 positive compounds) for the LLNA limit dose approach compared to the results obtained in the traditional LLNA. Do you consider that this is adequately addressed by the proposed cautionary language and weight of evidence consideration for negative substances? Please explain your answer.

2. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed test method standardized protocol? If not, then what recommendations would you make? Please explain your answer.

   - The recommended ICCVAM protocol (ICCVAM 1999; Dean et al. 2001; EPA 2003), as well as OECD TG 429, specifies that the highest dose tested should be the highest soluble concentration that does not induce systemic toxicity and/or excessive skin irritation. However, Kimber et al. (2006) concluded that negative results obtained from studies where the highest concentration tested was below 10% should be considered invalid, and adopted a 10% application concentration as a threshold of confidence for categorization of a chemical as being negative while noting that the figure should not be considered as inviolable. Are the data presented in the draft BRD (i.e., 5/313 positive substances in the NICEATM database were negative at concentrations ≤10%, but were positive at higher concentrations) adequate to conclude that this threshold concentration is not appropriate? If a negative result was obtained for a test substance in a study where the highest concentration that could be tested (based on systemic toxicity or excessive local irritation, as described in ICCVAM [1999], Dean et al. [2001], and EPA [2003]) was <10%, should additional testing be required? Do you agree that the current approach for selecting the "limit" dose is appropriate or do you conclude that there is a threshold concentration for the LLNA at which a negative result could always be considered as an acceptable result? If so, what is that concentration? Please explain your answer.

3. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed future studies? If not, then what recommendations would you make? Please explain your answer.

Annex F2

Questions for the Peer Review Panel:
LLNA for Testing Aqueous Solutions, Metals, and Mixtures

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: LLNA for Testing Aqueous Solutions, Metals, and Mixtures

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is currently updating the original validation report of the LLNA (ICCVAM 1999) based on a comprehensive review of available data and information regarding the current validity of the LLNA for assessing the skin sensitizing potential of mixtures, metal compounds, and substances tested in aqueous solutions. The information is based on a retrospective review of LLNA data derived from a database of over 500 substances (including mixtures) tested in the LLNA and builds on the previous ICCVAM evaluation of the LLNA, which was based on 209 substances (ICCVAM 1999). In the original ICCVAM report, the performance of the LLNA was compared to 1) the results from guinea pig tests and 2) information about sensitizers in humans (e.g., human maximization test [HMT] results, substances used in human repeat insult patch test [HRIPT], clinical data), where available. This addendum updates the LLNA performance analyses for mixtures, metal compounds, and substances tested in aqueous solutions when compared to human and guinea pig results.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA. You are first asked to review the information in the draft Addendum to the ICCVAM (1999) report for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in this Addendum to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[11]) have been appropriately addressed for the proposed use of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions. Adequate validation[12] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. This validation process characterizes the usefulness and limitations of a test method for its intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft Addendum.

The questions relating to the draft Addendum that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions.

---

[11] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[12] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

These include questions prepared by the ICCVAM IWG to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method, and/or adequate guidance for organizations that may be involved in conducting or supporting further development, standardization, and/or validation.

The overall question to consider is whether the validation status of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions has been adequately characterized, and is it sufficiently accurate and reliable to be used for the identification of sensitizing substances based on a comparison to either human or guinea pig responses.

I. Questions to the Panel: Review for Errors and Omissions

1. In the draft Addendum, are there any errors that need to be corrected or omissions of existing relevant data or information that should be included?

II. Questions to the Panel: Updated LLNA Applicability Domain Addendum

1. Do you consider the database of substances evaluated representative of a sufficient range of mixtures, metal compounds, and substances in aqueous solutions that are typically tested for skin-sensitization potential? Please explain your answer.
2. For the purpose of this evaluation, aqueous solutions were defined by the proportion of water (at least 20%) (i.e., substances or mixtures that were tested in an aqueous or an organic:aqueous vehicle were labeled as aqueous solutions). Do you consider this to be an appropriate criterion for defining aqueous solutions? If not, what would be more appropriate? Please explain your answer.
3. While coding of chemicals is recommended for prospective studies, this evaluation is based on a retrospective evaluation of existing data, most of which was not generated using coded chemicals to reduce bias. Does the lack of coding of test substances adversely impact or bias the current evaluation? Please provide a rationale for your answer.
4. For some substances submitted using the LLNA, it was not possible to confirm whether the data were generated using pooled animal data for each dose group (as allowed in Test Guideline [TG] 429 of the Organisation for Economic Co-operation and Development [OECD]) rather than individual animal data (as recommended in the ICCVAM 2001 protocol)? Cockshott et al. (2006) reported that using individual animal data allowed for technical problems during an experiment to be identified. Considering this, should the analysis of the performance of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions be limited to data from studies that can be confirmed as using individual animal data collection procedures? What impact might the inclusion of pooled animal data have on the accuracy analysis included in Section 5.0 of the draft Addendum? Please explain your answer.
5. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions been adequately evaluated and compared to the human and guinea pig (refer also to Section 5.0 of the draft

Addendum)? If not, what other analyses should be performed? Please explain your answer.

6. When multiple LLNA studies were available for the same substance, the majority call (where all studies used the same vehicle and the same concentration range) was used to assign an overall classification for the purposes of the accuracy analysis. For example, if chemical X was tested 5 times and was positive in 3 studies and negative in two, the overall classification was positive. Do you agree with the approach to assigning overall classifications? If not, how would you propose that this be accomplished? Please explain your answer.
7. Does the Addendum adequately characterize the usefulness and limitations of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions based on the accuracy analyses? If not, what additions or changes should be made to the current usefulness and limitations? Please explain your answer.
8. For some studies included in the draft Addendum, it was not possible to determine whether or not they had been conducted in accordance with Good Laboratory Practice (GLP) guidelines. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. Please discuss what impact this lack might have on the evaluation of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions and whether such studies should be excluded from any analysis.
9. As described in the draft Addendum, original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Considering this, should the results of these studies (all of which are currently included) be excluded from any of the performance analyses? Please explain your answer.
10. Based on the draft Addendum, have all the relevant data identified in published or unpublished studies conducted using the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions been adequately considered? If not, what other studies should to be considered?

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA for Testing Aqueous Solutions, Metals, and Mixtures

1. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA with regard to testing mixtures, metal compounds, and substances in aqueous solutions in terms of the proposed test method usefulness and limitations? Please explain your answer. If not, what recommendations would you make?
2. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA in terms of the proposed test method standardized protocol? Please explain your answer. If not, then what recommendations would the Panel make?

3. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA in terms of the proposed future studies? Please explain your answer. If not, then what recommendations would you make?

Annex F3

Questions for the Peer Review Panel:

Non-Radioactive LLNA Protocol - LLNA: DA Test Method

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Non-Radioactive LLNA Protocol: LLNA: DA Test Method

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the LLNA: DA (Local Lymph Node Assay-Daicel adenosine triphosphate [ATP]) for assessing the allergic contact dermatitis (ACD) potential of chemicals and other substances. This test method, developed by Daicel Chemical Industries, Ltd. (Tokyo, Japan), is a non-radiolabeled version of the traditional LLNA, and is based on measuring levels of ATP in the auricular lymph nodes as an indicator of increased cell proliferation.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA: DA. You are first asked to review the information in the draft ICCVAM LLNA: DA background review document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in the BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[13]) have been appropriately addressed for the proposed use of the LLNA: DA. Adequate validation[14] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. The validation process characterizes the usefulness and limitations of a test method for a specific intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA: DA (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft LLNA: DA BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA: DA.

These include questions prepared by the ICCVAM Immunotoxicity Working Group (IWG) to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method. The questions are also intended to obtain guidance that will be helpful to federal agencies and other organizations that may be involved in conducting or supporting further development, standardization, and/or validation studies.

The overall question to consider is whether the validation status of the LLNA: DA has been adequately characterized for its intended purpose, and is it sufficiently accurate and reliable to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure.

---

[13] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[14] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

I. Questions to the Panel: Review of the Draft LLNA: DA BRD for Errors and Omissions

1. In the draft LLNA: DA BRD, are there any errors in the BRD that should be corrected, or omissions of existing relevant data or information that should be included?

II. Questions to the Panel: Draft LLNA: DA BRD

1. Test Method Protocol
    i. The traditional LLNA protocol (ICCVAM 1999; Dean et al. 2001) recommends a minimum of five successfully treated animals per dose group. Current validation of the LLNA: DA was performed using four animals per dose group. What impact might using fewer mice have on the accuracy analysis of the LLNA: DA? Please explain your answer.
    ii. The data generated for the substances analyzed in the LLNA: DA test method come from auricular lymph nodes that were pooled across mice in each dose group rather than being analyzed on an individual animal data. What impact might the inclusion of pooled animal data have on the accuracy analysis of the LLNA: DA? Please explain your answer.
    iii. The LLNA: DA differs from the traditional LLNA in the treatment schedule and by including a pre-treatment with 1% SLS prior to application of the test substance. Do you consider these changes to be appropriate? Please explain your answer.
2. Substances Used for the Validation Studies
    i. Do you consider the LLNA: DA database representative of a sufficient range of chemical classes and physicochemical properties that it would be applicable to any of the types of chemicals and products that are typically tested for skin-sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the LLNA: DA? What chemicals or products should be evaluated to fill this data gap? Please explain your answer.
3. Test Method Accuracy
    i. The current accuracy analysis is based on overall concordance with the traditional LLNA. Accuracy statistics compared to the guinea pig tests and human data/experience have also been provided. Are these comparisons appropriate for assessing the accuracy of the LLNA: DA? Please explain your answer.
    ii. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: DA been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft ICCVAM BRD)? If not, what other analyses should be performed? Please explain your answer.

iii. There was one substance (2-mercaptobenzothiazole) that produced a "false negative" response compared to the traditional LLNA when tested using the LLNA: DA. The mean EC3 in the traditional LLNA for this substance is 2.5 (n=2), it is positive in both the guinea pig and human, and has "high" peptide reactivity as per Gerberick et al. (2007). Can you identify any characteristics associated with this or similar substances, compared to the correctly identified sensitizers, that might signal that this type of discordant response might occur, and therefore using the LLNA: DA to test such substances would not be appropriate or that negative results for such substances should indicate a need for confirmatory testing? Please explain your answer.

iv. There was one substance (benzalkonium chloride) that produced a "false positive" response compared to the traditional LLNA and guinea pig test when tested using the LLNA: DA. Can you identify any characteristics associated with this or similar substances, compared to the correctly identified non-sensitizers that might signal that this type of discordant response might occur, and therefore using the LLNA: DA to test such substances would not be appropriate, or that positive results for substances with such properties may warrant additional testing? Please explain your answer.

4. Test Method Reliability (Intra- and Inter-laboratory reproducibility)

   i. Has the intralaboratory reproducibility of the LLNA: DA been adequately evaluated and compared to the traditional LLNA (refer also to Table 7-1 of the draft LLNA: DA BRD)? If not, what other analyses should be performed? Are any limitations apparent based on this intralaboratory reproducibility assessment? Please explain your answer.

   ii. Has the interlaboratory reproducibility of the LLNA: DA been adequately evaluated and compared to the traditional LLNA (refer also to Tables 7-2 and 7-3 of the draft LLNA: DA BRD)? If not, what other analyses should be performed? Are any limitations apparent based on this interlaboratory reproducibility assessment? Please explain your answer.

   iii. The draft LLNA: DA BRD analyzes data from two interlaboratory validation studies that used coded substances, as well as an intralaboratory validation study with 31 substances that were not coded. Does the lack of coding of test substances adversely impact or bias the current evaluation? In addition, it appears that the lead laboratory established the dose levels tested in the two interlaboratory validation studies and the participating laboratories did not determine their own dose levels for testing. Does this adversely impact or bias the current evaluation? Please explain your answer.

5. Data Quality

   i. The studies evaluated in the draft BRD for the LLNA: DA were not conducted in accordance with Good Laboratory Practices (GLP) guidelines although there were reportedly done in laboratories that conduct GLP studies and were conducted "in the spirit" of GLP (K. Idehara, personal communication).

Please discuss what impact this might have on the evaluation of the LLNA: DA.

ii. The original records for these studies were requested but have not yet been obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Should any recommendations from ICCVAM be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription? Please explain your answer.

6. Consideration of all available data and relevant information

i. Based on the draft LLNA: DA BRD, have all the relevant data identified in published or unpublished studies that employ this test method been adequately considered? Are there other comparative test method data that were not considered in the draft BRD, but are available for consideration? If yes, please explain how to obtain such data.

IV. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA: DA

1. Test Method Usefulness and Limitations

i. Do you agree that the available data and test method performance (accuracy and reliability) support the ICCVAM draft recommendations for the LLNA: DA in terms of the proposed test method usefulness and limitations? Please explain your answer.

ii. If restrictions on using radioactive materials are present, should the LLNA: DA be routinely recommended for hazard identification of skin sensitizing substances in lieu of having to possibly use guinea pig tests? Please explain your answer.

iii. Even if limitations in using radioactive materials are not present, should the LLNA: DA procedure or other valid and accepted non-radioactive method be routinely recommended for hazard identification of skin sensitizing substances instead of the traditional LLNA? Please explain your answer.

iv. From a public health perspective, is the recommended guidance for evaluating negatives sufficient to address concerns associated with the false negative rate of 5% (1/19 substances) calculated for the LLNA: DA? Do you have suggestions for additional guidance or limitations? Please explain your answer.

v. From a testing strategy perspective, does the ICCVAM guidance address concerns associated with the false positive rate of 10% (1/10 substances) calculated for the LLNA: DA? Are there other suggestions for additional such guidance or limitations? Please explain your answer.

2. Test method protocol

i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: DA procedure in terms of the proposed test

method standardized protocols? If not, what recommendations would you make? Please explain your answer.

ii. Can the traditional LLNA limit dose procedure be applied to the LLNA: DA? Please explain your answer.

3. Future Studies

   i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: DA in terms of the proposed future studies? If not, then what recommendations would you make? Please explain your answer.

4. Performance Standards

   i. The LLNA: DA protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001; EPA 2003) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. In addition, there are differences between the two protocols that relate to how and when the test substance is applied and when the lymph nodes are collected (Table 2-1 and Appendix A in the draft LLNA: DA BRD). According to the proposed draft ICCVAM Performance Standards for the traditional LLNA (http://iccvam.niehs.nih.gov/methods/immunotox/llna_PerfStds.htm), any change to the LLNA protocol other than the method used to assess lymphocyte proliferation is considered a major change. Do you agree that these should be considered major changes and therefore the usefulness and limitations of the LLNA: DA should not be assessed using the draft ICCVAM Performance Standards? Please explain your answer.

   ii. Even if the draft ICCVAM Performance Standards do not apply to the LLNA: DA, what impact should the accuracy analysis based on 13 of the 18 required performance standards substances (only one false negative and no false positives) have on the overall evaluation of test method accuracy? Please explain your answer.

   iii. Should separate performance standards be developed for the LLNA: DA? Please explain your answer.

[This Page Intentionally Left Blank]

Annex F4

Questions for the Peer Review Panel:

Non-Radioactive LLNA Protocol - LLNA: BrdU-FC Test Method

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Non-Radioactive LLNA Protocol: LLNA: BrdU-FC Test Method

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the LLNA: BrdU-FC (Local Lymph Node Assay: Bromodeoxyuridine detected by flow cytometry) for assessing the allergic contact dermatitis potential of chemicals and other substances. This test method, developed by MB Research Labs (Spinnerstown, PA), is a non-radiolabeled version of the traditional LLNA, and is based on measuring the incorporation of bromodeoxyuridine (BrdU) into the DNA of dividing lymphocytes using flow cytometry as in indicator of cell proliferation.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA: BrdU-FC. You are first asked to review the information in the draft ICCVAM LLNA: BrdU-FC Background Review Document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in the BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[15]) have been appropriately addressed for the proposed use of the LLNA: BrdU-FC. Adequate validation[16] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. The validation process characterizes the usefulness and limitations of a test method for a specific intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA: BrdU-FC (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft LLNA: BrdU-FC BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA: BrdU-FC.

These include questions prepared by the ICCVAM Immunotoxicity Working Group (IWG) to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method. The questions are also intended to obtain guidance that will be helpful to federal agencies and other organizations that may be involved in conducting or supporting further development, standardization, and/or validation studies.

The overall question to consider is whether the validation status of the LLNA: BrdU-FC has been adequately characterized for its intended purpose, and is it sufficiently accurate and

---

[15] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[16] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

reliable to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure.

I. Questions to the Panel: Review of the Draft LLNA: BrdU-FC BRD for Errors and Omissions

1. In the draft LLNA: BrdU-FC BRD, are there any errors in the BRD that should be corrected, or omissions of existing relevant data or information that should be included?

II. Questions to the Panel: Draft LLNA: BrdU-FC BRD

1. Test Method Protocol
   i. The LLNA: BrdU-FC includes routine measurements of ear swelling as an indicator of excessive dermal irritation. Do you consider this procedure to be an appropriate approach? Do you think that this measurement should be recommended for routine inclusion into all LLNA protocols? Please explain your answers.
   ii. The LLNA: BrdU-FC also includes optional quantification of immunophenotypic markers as an additional mechanism for distinguishing irritants from sensitizers. Do you consider this to be an appropriate approach to reduce false positives? Are the correct markers being considered or do you recommend other/additional markers? Should these measurements be recommended for routine inclusion in the LLNA: BrdU-FC? Please explain your answers.
   iii. Please comment on the appropriateness of the "sequential strategy" used in the eLLNA: BrdU-FC (see Figure 2-1 of the draft BRD).
2. Substances Used for the Validation Studies
   i. Do you consider the LLNA: BrdU-FC database representative of a sufficient range of chemical classes and physicochemical properties that it would be applicable to any of the types of chemicals and products that are typically tested for skin-sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the LLNA: BrdU-FC? What chemicals or products should be evaluated to fill this data gap? Please explain your answers.
3. Test Method Accuracy
   i. The current accuracy analysis is based on overall concordance with the traditional LLNA. Accuracy statistics compared to the Guinea Pig tests and human data/experience have also been provided. Are these comparisons appropriate for assessing the accuracy of the LLNA: BrdU-FC? Please explain your answer.
   ii. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: BrdU-FC been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of

the draft ICCVAM BRD)? If not, what other analyses should be performed? Please explain your answer.

iii. Three substances (benzalkonium chloride, resorcinol, and Tween 80) produced a "false positive" response compared to the traditional LLNA and guinea pig test when tested using the LLNA: BRDU-FC (Based on immunophenotyping, benzalkonium chloride was subsequently classified as an irritant rather than a sensitizer). Can you identify any characteristics associated with these or similar substances, compared to the correctly identified non-sensitizers that might signal that this type of discordant response might occur, and therefore using the LLNA: BrdU-FC to test such substances would not be appropriate, or that positive results for substances with such properties may warrant additional testing? Please explain your answer.

4. Test Method Reliability (Intra- and Inter-laboratory reproducibility)

i. Has the intralaboratory reproducibility of the LLNA: BrdU-FC been adequately evaluated and compared to the traditional LLNA (refer also to Table 7-1 of the draft LLNA: BrdU-FC BRD)? If not, what other analyses should be performed? Are any limitations apparent based on this intra-laboratory reproducibility assessment? Please explain your answer.

ii. The draft LLNA: BrdU-FC BRD analyzes data from repeat testing of hexyl cinnamic aldehyde (HCA) in six different vehicles and intralaboratory reproducibility is assessed by coefficient of variation (CV). The calculated CVs ranged from 30% to 53%. Based on these data, are there concerns with the intralaboratory reproducibility of the LLNA: BrdU-FC? Please explain your answer.

5. Data Quality

i. The studies evaluated in the draft BRD for the LLNA: BrdU-FC were not all conducted in accordance with Good Laboratory Practices (GLP) guidelines although there were done in a laboratory that routinely conducts GLP studies (G. DeGeorge, personal communication). Please discuss what impact this might have on the evaluation of the LLNA: BrdU-FC.

ii. The original records for these studies were requested but have not yet been obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Do you agree that any recommendations from ICCVAM should be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription? Please explain your answer.
Consideration of all available data and relevant information

iii. Based on the draft LLNA: BrdU-FC BRD, have all the relevant data identified in published or unpublished studies that employ this test method been adequately considered? Are there other comparative test method data that were not considered in the draft BRD, but are available for consideration? If yes, please explain how to obtain such data.

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA: BrdU-FC

1. Test Method Usefulness and Limitations

   i. Do you agree that the available data and test method performance (accuracy and reliability) support the ICCVAM draft recommendations for the LLNA: BrdU-FC in terms of the proposed test method usefulness and limitations? Please explain your answer.

   ii. If restrictions on using radioactive materials are present, should the LLNA: BrdU-FC be routinely recommended for hazard identification of skin sensitizing substances in lieu of having to possibly use guinea pig tests? Please explain your answer.

   iii. Even if limitations in using radioactive materials are not present, should the LLNA: BrdU-FC procedure be routinely recommended for hazard identification of skin sensitizing substances instead of the traditional LLNA? If not, then why? Please explain your answer.

   iv. Do the ICCVAM recommendations adequately address concerns associated with the false positive rate of 17% (3/18 substances) calculated for the LLNA: BrdU-FC? Are there other suggestions for additional such guidance or limitations that should be considered? Please explain your answer.

2. Test Method Protocol

   i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-FC procedure in terms of the proposed test method standardized protocol? If not, then what recommendations would you make? Please explain your answer.

   ii. Can the traditional LLNA limit dose procedure be applied to the LLNA: BrdU-FC? Please explain your answer.

3. Future Studies

   i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-FC in terms of the proposed future studies? What other recommendations would you make? Please explain your answer.

4. Performance Standards

   i. The LLNA: BrdU-FC protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. According to the proposed draft ICCVAM Performance Standards for the traditional LLNA (http://iccvam.niehs.nih.gov/methods/immunotox/llna_PerfStds.htm), any change to the LLNA protocol other than the method used to assess lymphocyte proliferation is considered a major change. Do you agree that protocol differences between the LLNA: BrdU-FC and the traditional LLNA

should be considered only minor changes and therefore the validity of this test method should be based only on the draft ICCVAM Performance Standards? Please explain your answer.

ii. Even if the draft ICCVAM Performance Standards do not apply to the LLNA: BrdU-FC, what impact should the accuracy analysis based on 13 of the 18 required performance standards substances have on the overall evaluation of test method accuracy? Please explain your answer.

iii. Are there concerns that 3/6 sensitizers, for which EC3 data were available, had EC3 values that were outside of the proposed 0.5x to 2.0x EC3 acceptability range developed based on the traditional LLNA? Please explain your answer.

[This Page Intentionally Left Blank]

Annex F5

Questions for the Peer Review Panel:

Non-Radioactive LLNA Protocol - LLNA: BrdU-ELISA Test Method

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Non-Radioactive LLNA Protocol: LLNA: BrdU-ELISA Test Method

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the LLNA: BrdU-ELISA (local lymph node assay with bromodeoxyuridine [BrdU] detected by ELISA) for assessing the allergic contact dermatitis potential of chemicals and other substances. This test method, developed by Dr. Masahiro Takeyoshi (Tokyo, Japan), is a non-radiolabeled version of the traditional LLNA based on measuring levels of incorporated BrdU in the auricular lymph nodes as an indicator of increased cell proliferation.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA: BrdU-ELISA. You are first asked to review the information in the draft ICCVAM LLNA: BrdU-ELISA Background Review Document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in the BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[17]) have been appropriately addressed for the proposed use of the LLNA: BrdU-ELISA. Adequate validation[18] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. The validation process characterizes the usefulness and limitations of a test method for a specific intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA: BrdU-ELISA (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft LLNA: BrdU-ELISA BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA: BrdU-ELISA.

These include questions prepared by the ICCVAM Immunotoxicity Working Group (IWG) to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method. The questions are also intended to obtain guidance that will be helpful to federal agencies and other organizations that may be involved in conducting or supporting further development, standardization, and/or validation studies.

The overall question to consider is whether the validation status of the LLNA: BrdU-ELISA has been adequately characterized for its intended purpose, and is it sufficiently accurate and

---

[17] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[18] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

reliable to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure.

I. Questions to the Panel: Comments on the Draft LLNA: BRDU-ELISA BRD for Errors and Omissions

1. In the draft LLNA: BrdU-ELISA BRD, are there any errors in the BRD that should be corrected, or omissions of existing relevant data or information that should be included?

II. Questions to the Panel: Draft LLNA: BrdU-ELISA BRD

1. Test Method Protocol
   i. The data generated for the substances analyzed in the LLNA: BrdU-ELISA test method come from auricular lymph nodes from four individual mice in each dose group. The recommended ICCVAM LLNA protocol and OECD Test Guideline 429 recommend a minimum of five animals per dose group for collecting individual animal data. What impact might the use of four animals per dose group have on the accuracy and reliability of the LLNA: BrdU-ELISA? Do you agree with the ICCVAM recommendation that future use of this test method protocol should include five animals per dose group? Please explain your answer.
2. Substances Used for the Validation Studies
   i. Do you consider the LLNA: BrdU-ELISA database representative of a sufficient range of chemical classes and physicochemical properties that the test method would be applicable to any of the types of chemicals and products that are typically tested for skin-sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the LLNA: BrdU-ELISA? What chemicals or products should be evaluated to fill this data gap? Please explain your answers.
3. Test Method Accuracy
   i. The current accuracy analysis using a stimulation index (SI) ≥3 or SI ≥1.3 to identify sensitizers is based on overall concordance with the traditional LLNA. Accuracy statistics compared to the Guinea Pig tests and human data/experience have also been provided. Are these comparisons appropriate for assessing the accuracy of the LLNA: BrdU-ELISA? Please explain your answer.
   ii. Takeyoshi et al. (2007) performed an accuracy analysis using decision criteria other than SI ≥3 to classify substances as sensitizers. Maximal accuracy for the LLNA: BrdU-ELISA occurred when an SI ≥1.3 was used to distinguish between sensitizers and non-sensitizers. Using this decision criteria, they achieved an accuracy of 91% (21/23), with a sensitivity of 100% (16/16) and a specificity of 71% (5/7) (i.e., there were no false negatives and two false positives). Does this analysis support a recommendation that the decision criteria be based on an SI ≥1.3? Are there concerns with using a small

increase (i.e., 1.3-fold) above the vehicle control response as the basis for identifying a positive response? Please explain your answers.

iii. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: BrdU-ELISA, using the SI ≥3 criterion, been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft ICCVAM BRD)? If not, what other analyses should be performed? Please explain your answer.

iv. Using the SI ≥3 criterion, there were four substances (aniline, 4-chloroaniline, 2-mercaptothiazole, and hydroxycitronellal) when tested using the LLNA: BrdU-ELISA that produced "false negative" responses compared to the traditional LLNA. 4-Chloroanline and aniline are amines. 2-Mercaptobenzothiazole is a heterocylic compound and hydroxycitronellal is a hydrocarbon. 2-Mercaptobenzothiazole is a liquid, but the other three substances are solids. Can you identify any characteristics associated with these or similar substances, compared to the correctly identified sensitizers, that might signal that this type of discordant response might occur, and therefore using the LLNA: BrdU-ELISA to test such substances would not be appropriate or that negative results for such substances should indicate a need for confirmatory testing? Please explain your answer.

4. Test Method Reliability (Intra- and Inter-laboratory reproducibility)

   i. Has the intralaboratory reproducibility of the LLNA: BrdU-ELISA been adequately evaluated and compared to the traditional LLNA (refer also to Tables 7-1 through 7-3 of the draft LLNA: BrdU-ELISA BRD)? If not, what other analyses should be performed? Are any limitations apparent based on this intra-laboratory reproducibility assessment? Please explain your answers.

   ii. The substances evaluated for intralaboratory reproducibility of the LLNA: BrdU-ELISA study were not coded. Does the lack of coding of test substances adversely impact or bias the current evaluation? Please explain your answer.

   iii. The Japanese Center for the Validation of Alternative Methods (JaCVAM) has implemented a multi-laboratory validation study of the LLNA: BrdU-ELISA. Although the results from this study have yet to be reported, we are hoping to obtain information on the study design (i.e., with regard to number and types of chemicals tested and the number of laboratories involved). If we do, do you consider the design appropriate to adequately determine the extent of interlaboratory reproducibility for the LLNA: BrdU-ELISA? If not, what other analyses should be performed? Are any limitations apparent based on this study design? Please explain your answer.

5. Data Quality

   i. The studies evaluated in the draft BRD for the LLNA: BrdU-ELISA were not conducted in strict accordance with all provisions of the Good Laboratory Practice (GLP) guidelines, although there were reportedly performed in laboratories that conduct GLP studies (M. Takeyoshi, personal

communication). Please discuss what impact this might have on the evaluation of the LLNA: BrdU-ELISA.

ii. The original records for these studies were requested but were not available. As a result, an independent audit could not be conducted to confirm that the reported data in peer reviewed publications and a poster presentation is the same as the data recorded in laboratory notebooks. Should any recommendations from ICCVAM be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription? Please explain your answer.

6. Consideration of all available data and relevant information

i. Based on the draft LLNA: BrdU-ELISA BRD, have all the relevant data identified in published or unpublished studies that employ this test method been adequately considered? Are there other comparative test method data that were not considered in the draft BRD, but are available for consideration? If yes, please explain how to obtain such data.

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA: BrdU-ELISA

1. Test Method Usefulness and Limitations

   i. Do you agree that the available data and test method performance (accuracy and reliability) support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA in terms of the proposed test method usefulness and limitations? Please explain your answer.

   ii. If restrictions on using radioactive materials are present, should the LLNA: BrdU-ELISA be routinely recommended for hazard identification of skin sensitizing substances in lieu of using guinea pig tests due to the advantages of fewer animals and the avoidance of pain and distress? Please explain your answer.

   iii. Even if limitations in using radioactive materials are not present, should the LLNA: BrdU-ELISA procedure or other valid and accepted non-radioactive method be routinely recommended for hazard identification of skin sensitizing substances instead of the traditional LLNA? Please explain your answer.

   iv. Does using a decision criterion of SI $\geq$1.3 instead of SI $\geq$3.0 resolve any concerns with respect to potential false positives or false negatives that may occur in this test method? Are there other suggestions for additional such guidance or limitations that should be considered? Please explain your answer.

2. Test Method Protocol

   i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA procedure in terms of the proposed test method standardized protocols? If not, then what recommendations would you make? Please explain your answer.

   ii. Can the traditional LLNA limit dose procedure be applied to the LLNA: BrdU-ELISA? Please explain your answer.

3. Future Studies

   a. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA in terms of the proposed future studies? If not, then what recommendations would you make? Please explain your answer.

i. Performance Standards

   i. The LLNA: BrdU-ELISA protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. According to the proposed draft ICCVAM Performance Standards for the traditional LLNA (http://iccvam.niehs.nih.gov/methods/immunotox/llna_PerfStds.htm), any change to the LLNA protocol other than the method used to assess

lymphocyte proliferation is considered a major change. Do you agree that protocol differences between the LLNA: BrdU-ELISA and the traditional LLNA should be considered only minor changes and therefore the validity of this test method should be based only on the draft ICCVAM Performance Standards? Please explain your answer.

ii. Even if the draft ICCVAM Performance Standards do not apply to the LLNA: BrdU-ELISA, what impact should the accuracy analysis based on eight of the 18 required performance standards substances (only one false negative and no false positives) have on the overall evaluation of test method accuracy? Please explain your answer.

iii. Are there concerns that 4/4 sensitizers, for which EC3 data were available, had EC3 values that were outside of the recommended 0.5x to 2.0x EC3 acceptability range developed based on the traditional LLNA? Please explain your answer.

iv. Should separate performance standards be developed for the LLNA: BrdU-ELISA? Please explain your answer.

Annex F6

Questions for the Peer Review Panel:
Draft ICCVAM Performance Standards for the LLNA

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Draft ICCVAM Performance Standards for the LLNA

ICCVAM has developed draft LLNA performance standards consisting of essential test method components, a minimum list of reference substances, and expected accuracy and reliability values. These are proposed for evaluating the acceptability of proposed test methods that are mechanistically and functionally similar to the traditional LLNA test method protocol previously recommended by ICCVAM.

The overall question for the Panel is whether these performance standards are considered adequate for assessing the accuracy and reliability of test method protocols that are based on similar scientific principles and that measure the same biological effect as the traditional LLNA?

1) Purpose and Applicability

a) ICCVAM proposes that these performance standards should only be applicable to versions of the LLNA that incorporate minor modifications to the traditional LLNA. Currently, this is limited to the use of non-radioactive reagents to measure lymphocyte proliferation. It is considered essential that the modified LLNA should otherwise adhere to all other aspects of the traditional LLNA protocol, as defined by ICCVAM (1999) and Dean et al. (2001). This includes aspects such as: the sex and strain of mouse used, the number of mice per dose group, the timing and site of test article treatment, the duration between the last treatment and lymph node collection, the inclusion of concurrent negative and positive control groups, the measured endpoint (i.e., lymphocyte proliferation in the draining auricular lymph node), and the collection of data at the level of the individual mouse. Do you agree that the use of non-radioactive reagents for measuring cell proliferation in the lymph nodes, if that is the only difference, constitutes a minor modification to the traditional LLNA protocol? Is it necessary to keep the same decision criteria for distinguishing between sensitizers and non-sensitizers (i.e., an SI of 3)? Please explain your answer.

b) Are there other procedural modifications that you consider minor and therefore could be evaluated for equivalence to the traditional LLNA using the proposed performance standards? If yes, please explain what they are and why.

c) Do you consider these performance standards to also be applicable to the LLNA limit dose procedure? Please explain your answer.

2) Essential Test Method Components

a) The essential test method components are based on the ICCVAM recommended protocol (ICCVAM 1999; Dean et al. 2001), which is the basis for the current U.S. Environmental Protection Agency (EPA 2003) test guideline (TG). There are some notable differences between these protocols and the Organisation for Economic Co-operation and Development TG 429 for the LLNA (OECD 2002). When evaluations of non-radioactive versions of the traditional LLNA are conducted using these performance standards, is it necessary that the validation studies follow the ICCVAM recommended protocol? Specifically, should the studies include: 1) a concurrent positive control with each test substance; 2) using a minimum of five animals per dose group; and 3) measuring proliferation in lymph nodes from individual animals

rather than pooling lymph nodes across all animals in a dose group? Please explain your answers.

b) Should the concurrent testing of the positive control and test substance be conducted in the same vehicle or can different vehicles be used? Please explain your answer.

3) Proposed Reference Substances

a) Do you agree with the selection and prioritization criteria used to select the performance standards reference substances? Please explain your answer.

b) The rationale for the number of substances included on the "required" list of substances (n=18) is provided in Appendix C of the draft ICCVAM Performance Standards. Do you consider this to be an adequate number upon which to evaluate the performance of non-radioactive LLNA test methods, where the only protocol modification is the method for assessing cell proliferation in the auricular lymph nodes? If not, how many reference chemicals should be tested? Please explain your answer.

c) Do you consider the types of substances included in the reference substance list, with regard to relative sensitization potency, physicochemical characteristics, and vehicles, to be representative of the overall diversity of substances that are likely to be tested for skin sensitization? Please explain your answer.

d) Are there other types of information relevant to skin sensitization that should be considered in order to demonstrate an adequately diverse reference list? If yes, please explain what additional information should be included.

e) Are there other substances that you consider to be more appropriate for assessing the sensitivity (ability of the test method to correctly identify sensitizing substances) and specificity (ability of the test method to correctly identify non-sensitizing substances) of non-radioactive LLNA test methods, and for which there is available LLNA, guinea pig, and human data? If yes, please name the substances and explain why.

f) Four "discordant chemicals" (i.e., two LLNA false negatives and two false positives compared to guinea pig tests or human data) are included as optional substances that could be studied to evaluate if the proposed modifications might provide improved performance relative to the traditional LLNA.

- Please comment on the appropriateness of including these specific substances in the reference list. Should different substances be included? Should more false negative/positive substances be tested? If so, what are they? Please explain your answers.
- Do you consider their "optional" status appropriate, or should testing these substances be required? Please explain your answer.
- Would "correct" results with these four discordant chemicals be sufficient to consider the alternative test method to be more predictive of skin sensitization than the traditional LLNA? Please explain your answer.

4) Test Method Accuracy Standards

a) The draft ICCVAM Performance Standards state that the non-radioactive proposed LLNA test method should exactly match the accuracy of the traditional LLNA when evaluated with the minimum set of 18 reference substances. Do you agree that test method accuracy should be based on a chemical-by-chemical match with regard to identifying the chemicals as sensitizers or non-sensitizers? Please explain your answer.

b) The draft ICCVAM LLNA Performance Standards recommend that, for each sensitizer, the threshold concentration that induces a positive SI response should be within 0.5x to 2.0x of the concentration obtained for the EC3 in the traditional LLNA. As described in Appendix D of the draft ICCVAM Performance Standards, statistical approaches have been used in an attempt to identify an appropriate range, but these calculated ranges do not appear to be the most practical. In contrast, the NICEATM LLNA database demonstrates that EC3 values from replicate tests for a sensitizing chemical when tested using the same solvent are rarely outside of this proposed 0.5x to 2.0x acceptability EC3 range. Please comment on the appropriateness of using this criterion to judge the equivalency of a non-radioactive version of the traditional LLNA. If this approach is not acceptable, please explain why, and present an alternative approach along with the basis for this approach.

c) For five of the 13 sensitizers on the draft ICCVAM reference substances list, the reference EC3 value is based on a single LLNA study (see Table C1 of the draft ICCVAM Performance Standards). Please comment on the appropriateness of including such chemicals in the list of recommended reference substances and whether or not the 0.5x to 2.0x criteria should be applied to such substances. Please explain your answer.

5) Test Method Reliability Standards

a) The draft ICCVAM Performance Standards state that acceptable intralaboratory reproducibility will be indicated by a laboratory obtaining, in each of four independent experiments conducted with at least one week between each experiment, ECt values (the estimated concentration needed to produce an SI of a defined threshold [e.g., EC3]) for hexyl cinnamic aldehyde (HCA) that are generally within 0.5x to 2.0x (i.e., 5% to 20%) of the historical mean EC3 concentration (10%) for this substance, based on existing available traditional LLNA data.

- Do you consider the number of repeat experiments (n=4) to be adequate? Please explain your answer.
- Do you consider testing HCA adequate for demonstrating intralaboratory reproducibility? If not, which substance(s) should be tested? Please explain your answer.
- Is the required one-week interval between independent tests adequate and/or appropriate? If not, please provide an alternative schedule and explain the basis for your recommendation.
- Do you consider the criteria for acceptability to be appropriate? If not, please describe another criteria and explain the basis for your recommendation.

b) The draft ICCVAM Performance Standards state that acceptable interlaboratory reproducibility will be indicated by each of three laboratories obtaining ECt values for HCA and 2,4-dinitrochlorobenzene (DNCB) from a single experiment that are generally within 0.5x to 2.0x (5% to 20% and 0.025 to 0.1%, respectively) of the mean historical EC3 concentration (10% and 0.05%, respectively) obtained for these two substances in the traditional LLNA.

- Do you consider the single experiment per substance in each laboratory to be adequate? If not, please provide an alternative approach and explain why.
- Do you consider testing HCA and DNCB to be adequate for demonstrating interlaboratory reproducibility? If not, which substance(s) should be tested? Please explain your answer.
- Do you consider the criteria for acceptability to be appropriate? If not, please describe another criteria and explain the basis for your recommendation.

6) Summary Question

a) If a radioactive or non-radioactive LLNA method were proposed with a "major change" (e.g., different mouse strain or use of male mice, change in the schedule for test article administration, change in schedule for lymph node excision, etc.), what criteria should be used to evaluate the equivalence of this method to the traditional LLNA?

- Would a new set of performance standards be required for this method? Please explain your answer.
- How many reference substances might be considered adequate? Would the 18 minimum reference substances in the draft ICCVAM LLNA Performance Standards be sufficient? If more substances are considered necessary, how many should there be tested and what should their characteristics be? Please explain your answer.
- Regardless of the number of reference substances, should the alternative LLNA be required to obtain the same call (and potency for sensitizers) as the traditional LLNA for the 18 minimum reference substances in the draft ICCVAM LLNA Performance Standards? Please explain your answer.
- Are there additional specific substances that should be used? If yes, what are they? Please explain your answer.
- What, if any, additional information would be considered necessary and why?

Annex F7

Questions for the Peer Review Panel:

Use of the LLNA for Potency Determinations

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Use of the LLNA for Potency Determinations

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the murine local lymph node assay (LLNA) for assessing the potential potency of substances to cause allergic contact dermatitis.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA for potency determinations. You are first asked to review the information in the draft ICCVAM LLNA for potency determinations Background Review Document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in this BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[19]) have been appropriately addressed for the proposed use of the LLNA for potency determinations. Adequate validation[20] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. This validation process characterizes the usefulness and limitations of a test method for its intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA for potency determinations (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA for potency determinations.

These include questions prepared by the ICCVAM IWG to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method, and/or adequate guidance for organizations that may be involved in conducting or supporting further development, standardization, and/or validation.

The overall question to consider is whether the validation status of the LLNA for potency determinations has been adequately characterized, and is it sufficiently accurate and reliable to be used for the identification of sensitizing substances according to their relative potency classification based on a comparison to either human or guinea pig responses.

I. Questions to the Panel: Review for Errors and Omissions

1. Are there any errors or omissions of existing relevant data or information that should be included in the draft BRD?

---

[19] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[20] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

II. Questions to the Panel: LLNA for Potency Determinations Draft BRD

1. Do you consider the database of substances evaluated representative of a sufficient range of chemical classes and physical chemical properties that it would be applicable to any of the types of chemicals and products that are typically tested for skin-sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the for potency determinations? What chemicals or products should be evaluated to fill this data gap? Please explain your recommendation.
2. While coding of chemicals is recommended for prospective studies, this evaluation is based on a retrospective evaluation of existing data, most of which was not generated using coded chemicals to reduce bias. Does the lack of coding of test substances adversely impact or bias the current evaluation? Please provide a rationale for your answer.
3. For some substances submitted using the LLNA, it was not possible to confirm whether the data were generated using pooled animal data for each dose group (as allowed in Test Guideline [TG] 429 of the Organisation for Economic Co-operation and Development [OECD]) rather than individual animal data (as recommended in the ICCVAM 2001 protocol)? Cockshott et al. (2006) reported that using individual animal data allowed for technical problems during an experiment to be identified. Considering this, should the analysis of the performance of the LLNA for potency determinations be limited to data from studies that can be confirmed as using individual animal data collection procedures? What impact might the inclusion of pooled animal data have on the accuracy analysis included in Section 6.0 of the draft ICCVAM BRD? Please explain your answer.
4. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA for potency determinations been adequately evaluated and compared to the human and guinea pig (refer also to Section 6.0 of the draft ICCVAM BRD)? If not, what other analyses should be performed?
5. The accuracy analysis (see Section 6.0 of the draft ICCVAM BRD) focuses on the two-level categorization scheme proposed by the United Nations Globally Harmonized System for Classification and Labelling for both human and guinea data. Should other categorization schemes be considered?
6. Does the BRD adequately characterize the usefulness and limitations of the LLNA for potency determinations based on the accuracy analyses? If not, what additions or changes should be made to the current usefulness and limitations?
7. Has the reliability (e.g., intralaboratory repeatability, intra- and inter-laboratory reproducibility) of the LLNA for potency determinations been adequately evaluated (refer also to Section 7.0 of the draft ICCVAM BRD)? If not, what other analyses should be performed?

8. For some studies included in the draft BRD, it was not possible to determine whether or not they had been conducted in accordance with Good Laboratory Practice (GLP) guidelines. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. Please discuss what impact this lack might have on the evaluation of the LLNA for potency determinations and whether such studies should be excluded from any analysis.
9. As described in the draft BRD, original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Considering this, should the results of these studies (all of which are currently included) be excluded from any of the performance analyses? If yes, please explain.
10. Based on the draft BRD, have all the relevant data identified in published or unpublished studies conducted using the LLNA for potency determinations been adequately considered? If not, what other studies should to be considered?

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA for Potency Determinations

1. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA for potency determinations in terms of the proposed test method usefulness and limitations? Please explain your answer. If not, why recommendations would you make?
    - Should the LLNA be routinely recommended for the hazard classification of the skin-sensitization potency of chemicals?
2. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA in terms of the proposed test method standardized protocol? Please explain your answer. If not, then what recommendations would you make?
3. Should the relevant testing guidelines for the LLNA be updated to include the calculation of an EC3 value?
4. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA in terms of the proposed future studies? Please explain your answer. If not, then what recommendations would you make?

[This Page Intentionally Left Blank]

Appendix F2

Summary Minutes of Independent Scientific Review Panel Meeting

March 4–6, 2008

[This Page Intentionally Left Blank]

Summary Minutes

Independent Scientific Peer Review Panel Meeting

Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay (LLNA): A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products

Consumer Product Safety Commission (CPSC), Headquarters

Bethesda, MD

March 4 – 6, 2008

8:30 a.m. – 5:30 p.m.

Peer Review Panel Members:

Michael Luster, Ph.D. (Peer Review Panel Chair)	Senior Consultant to the NIOSH Health Effects Laboratory, Morgantown, WV, U.S.
Nathalie Alépée, Ph.D.	Associate Research Fellow, Pfizer PDRD MCT Laboratory, France
Anne Marie Api, Ph.D.	Vice President, Human Health Sciences, Research Institute for Fragrance Materials, Woodcliff Lake, NJ, U.S.
Nancy Flournoy, M.S., Ph.D.	Professor and Chair, Dept. of Mathematics and Statistics, University of Missouri-Columbia, Columbia, MO, U.S.
Thomas Gebel, Ph.D.	Regulatory Toxicologist, Federal Institute for Occupational Safety and Health, Dortmund, Germany
Kim Headrick, B. Admin., B.Sc.	International Harmonization Senior Policy Advisor, Health Canada, Ottawa, Ontario, Canada
Dagmar Jírová, M.D., Ph.D.	Toxicologist, Research Manager, Head of Reference Center for Cosmetics, Head of Reference Laboratory for Experimental Immunotoxicology, National Institute of Public Health, Czech Republic
David Lovell, Ph.D	Reader in Medical Statistics, Postgraduate Medical School, University of Surrey, Guildford, Surrey, U.K.
Howard Maibach, M.D.	Professor, Dept. of Dermatology, University of California-San Francisco, San Francisco, CA, U.S.
James McDougal, Ph.D.	Professor and Director of Toxicology Research, Dept. of Pharmacology and Toxicology, Boonshoft School of Medicine, Wright State University, Dayton, OH, U.S.

Peer Review Panel Members:

Michael Olson, Ph.D.	Director of Occupational Toxicology, Corporate Environment Health and Safety, GlaxoSmithKline, RTP, NC, U.S.
Raymond Pieters, Ph.D.	Associate Professor, Immunotoxicology Group Leader, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands
Jean Regal, Ph.D.	Professor, Dept. of Pharmacology, University of Minnesota Medical School, Duluth, MN, U.S.
Peter Theran, V.M.D.	Massachusetts Society for the Prevention of Cruelty to Animals, Novato, CA, U.S.
Stephen Ullrich, Ph.D.	Dallas/Ft. Worth Living Legends Professor & Professor of Immunology, Post-graduate School of Biomedical Science, University of Texas M.D. Anderson Cancer Center, Houston, TX, U.S.
Michael Woolhiser, Ph.D.	Technical Leader - Immunotoxicology, Toxicology and Environmental Research and Consulting Immunology, Dow Chemical, Midland, MI, U.S.
Takahiko Yoshida, M.D., Ph.D.	Professor, Dept. of Health Science, Asahikawa Medical College, Hokkaido, Japan

ICCVAM and ICCVAM IWG Members:

Paul Brown, Ph.D.	FDA, Silver Spring, MD, U.S.
Ruth Barratt, Ph.D., D.V.M.	FDA, Rockville, MD, U.S.
Karen Hamernik, Ph.D.	EPA, Washington, DC, U.S.
Masih Hashim, Ph.D.	EPA, Washington, DC, U.S.
Abigail Jacobs, Ph.D. (IWG Co-Chair)	FDA, Silver Spring, MD, U.S.
Kristina Hatlelid, Ph.D.	CPSC, Bethesda, MD, U.S.
Joanna Matheson, Ph.D. (IWG Co-Chair)	CPSC, Bethesda, MD, U.S.
Tim McMahon, Ph.D.	EPA, Washington, DC, U.S.
Amy Rispin, Ph.D.	EPA, Washington, DC, U.S.
William Stokes, D.V.M., D.A.C.L.A.M.	NIEHS, RTP, NC, U.S.
Raymond Tice, Ph.D.	NIEHS, RTP, NC, U.S.
Ron Ward, Ph.D.	EPA, Washington, DC, U.S.
Marilyn Wind, Ph.D. (ICCVAM Chair)	CPSC, Bethesda, MD, U.S.
Jiaqin Yao, Ph.D.	FDA, Silver Spring, MD, U.S.

ECVAM Observer:

David Basketter, Ph.D.	DABMEB Consultancy Ltd., Bedfordshire, U.K.

Invited Experts:

George DeGeorge, Ph.D., D.A.B.T.	MB Research Laboratories, Spinnerstown, PA, U.S.
Kenji Idehara, Ph.D.	Daicel Chemical Industries, Hyogo, Japan
Masahiro Takeyoshi, Ph.D.	Chemicals Evaluation and Research Institute, Saitama, Japan

Public Attendees:

Odette Alexander	Syngenta Crop Protection, Inc., Greensboro, NC, U.S.
Nancy Beck, Ph.D.	PCRM, Washington, DC, U.S.
Ann Blacker, Ph.D.	Bayer CropScience, RTP, NC, U.S.
Stuart Cagan, Ph.D.	Shell Oil Company, Houston, TX, U.S.
Joan Chapdelaine, Ph.D.	Calvert Laboratories, Inc., Olyphant, PA, U.S.
Adriana Doi, Ph.D.	BASF Corporation, RTP, NC, U.S.
Carol Eisenmann, Ph.D.	Personal Care Products Council, Washington, DC, U.S.
Charles Hastings, Ph.D.	BASF Corporation, RTP, NC, U.S.
Kailash Gupta, D.V.M., Ph.D.	Retired CPSC, Bethesda, MD, U.S.
John Lyssikatos	Hill Top Research, Miamiville, OH, U.S.
Laurence Musset. Ph.D.	OECD, Paris, France
Carol O'Neil	NuPathe, Conshohocken, PA, U.S.
Kui Lea Park, Ph.D.	National Institute of Toxicological Research, KFDA, Seoul, Korea
Rafael Rivas	AFRRI/USHUS, Bethesda, MD, U.S.
Terri Sebree	NuPathe, Conshohocken, PA, U.S.
Libby Sommer	EPA, Washington, DC, U.S.
Merrill Tisdel	Syngenta Crop Protection Inc., Greensboro, NC, U.S.
Jeffrey Toy, Ph.D.	FDA, Rockville, MD, U.S.

NICEATM:

William Stokes, D.V.M., D.A.C.L.A.M.	Director
Raymond Tice, Ph.D.	Deputy Director
Debbie McCarley	Special Assistant to the Director
Support Contract Staff— Integrated Laboratory Systems, Inc. (ILS)	
David Allen, Ph.D.	Michael Paris
Thomas Burns, M.S.	Eleni Salicru, Ph.D.
Linda Litchfield	Judy Strickland, Ph.D., D.A.B.T.
Douglas Winters, M.S.	ILS, Inc.

Abbreviations: AFFRI = Armed Forces Radiobiology Research Institute; CPSC = U.S. Consumer Product Safety Commission; ECVAM = European Centre for the Validation of Alternative Methods; EPA = U.S. Environmental Protection Agency; FDA = U.S. Food and Drug Administration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; ILS = Integrated Laboratory Systems; IWG = Immunotoxicology Working Group; KFDA = Korea Food and Drug Administration; NICEATM = National Toxicology Program Interagency Center for Evaluation of Alternative Toxicological Methods; NIEHS = National Institute of Environmental Health Sciences; NIOSH = National Institute of Occupational Safety and Health; OECD = Organisation for Economic Co-operation and Development; PCRM = Physicians Committee for Responsible Medicine; USDA = U.S. Department of Agriculture; USHUS = Uniformed Services University of the Health Sciences.

Tuesday, March 4, 2008

Call to Order and Introductions

Dr. Michael Luster (Peer Review Panel Chair) called the meeting to order at 8:30 a.m. and introduced himself. He then asked all Peer Review Panel (hereafter Panel) members to introduce themselves and to state their name and affiliation for the record. He then asked all the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) staff, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) members, the ICCVAM Immunotoxicity Working Group (IWG) members, the European Centre for the Validation of Alternative Methods (ECVAM) observer, and members of the public to also introduce themselves. Dr. Luster stated that there would be opportunity for public comments during each of the seven local lymph node assay (LLNA)-related topics. He asked that all those interested in making a comment, register at the registration table and provide a written copy of their comments, if available, to NICEATM staff. Dr. Luster emphasized that the comments would be limited to seven minutes per individual and that, while an individual would be welcome to make comments during each commenting period, repeating the same comments at each comment period would be inappropriate. He further stated that the meeting was being recorded and that Panel members should speak directly their microphone. Finally, Dr. Luster noted that if the Panel finished early with the assigned topics on the agenda for that day, they would proceed to the next day's topics if time permitted.

Welcome from the ICCVAM Chair

Dr. Marilyn Wind, U.S. Consumer Product Safety Commission (CPSC) and Chair of ICCVAM, welcomed everyone to CPSC and to the Panel meeting. Dr. Wind stressed the importance of this Panel's efforts especially considering recent reports that allergies and asthma have increased markedly over the past number of years and that contact dermatitis is the most common occupational illness in the United States. Dr. Wind thanked the Panel members for giving their expertise, time, and effort and acknowledged their important role to the ICCVAM test method evaluation process. Dr. Wind also emphasized the important role of the public and their comments in this process.

Welcome from the Director of NICEATM, and Conflict of Interest Statements

Dr. William Stokes, Director of NICEATM, stated the Panel meeting was being convened as a National Institutes of Health (NIH) special emphasis panel and was being held in accordance with the Federal Advisory Committee Act regulations. As such, Dr. Stokes indicated that he would serve as the Designated Federal Official for this public meeting. He reminded the Panel that they had signed a conflict-of-interest statement when they were selected for the Panel, in which they identified any potential conflicts of interest. He then read this statement to provide another opportunity for members of the Panel to identify any conflicts not previously declared. Dr. Luster asked the Panel members to declare any direct or indirect conflicts based on Dr. Stokes statements and to recuse themselves from discussion and voting on any aspect of the

meeting where there might be a conflict. None of the Panel members declared a conflict of interest.

Overview of the ICCVAM Test Method Evaluation Process

Dr. Stokes provided an overview of the ICCVAM test method evaluation process. He stated that the Panel was made up of 19 different scientists from eight different countries (Canada, Czech Republic, France, Germany, Japan, The Netherlands, United Kingdom, and the United States). Dr. Stokes thanked the Panel members for the significant amount of time and effort that they had devoted to prepare for and attend the meeting. He explained that the purpose of the Panel was to assist ICCVAM by carrying out an independent scientific peer review of the information provided on a series of proposed new versions of the LLNA and some expanded applications of the assay. Dr. Stokes mentioned that the original LLNA peer review panel in 1998 considered the LLNA a valid substitute for the guinea pig-based test in most testing situations, but not all. He mentioned that three Panel members from the 1998 review are also on the current Panel (i.e., Drs. Howard Maibach, Jean Regal, and Stephen Ullrich). Dr. Stokes also reviewed the nomination that was received from CPSC in January 2007[21], which provides the basis for the current evaluation.

Dr. Stokes then identified the 15 Federal agencies that comprise ICCVAM and summarized ICCVAM's mission. He noted that ICCVAM, as an interagency committee, does not carry out research and development or validation studies. Instead, ICCVAM, in conjunction with NICEATM, carries out the critical scientific evaluation of proposed test methods with regard to their usefulness and limitations for regulatory testing and then makes formal recommendations to ICCVAM agencies.

Dr. Stokes provided a brief review of ICCVAM's history and summarized the ICCVAM Authorization Act of 2000[22], detailing the purpose and duties of ICCVAM. He noted that one of ICCVAM's duties is to review and evaluate new, revised, and alternative test methods applicable to regulatory testing. He stated that all of the reports produced by NICEATM are available on the NICEATM-ICCVAM website or can be obtained upon request from NICEATM. He also mentioned that ICCVAM provides guidance on test method development, validation criteria, and processes, and helps to facilitate not only the acceptance of scientifically valid alternative methods, but also encourages international harmonization.

Dr. Stokes then described the ICCVAM test method evaluation process, which begins with a test method nomination or submission. NICEATM conducts a prescreen evaluation to summarize the extent to which the proposed submission or nomination addresses the ICCVAM prioritization criteria. A report of this evaluation is then provided to ICCVAM, which in turn develops recommendations regarding the priority for evaluation. ICCVAM then seeks input on their recommendations from the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) and the public. Given sufficient regulatory applicability, sufficient data, resources, and priority, a test method will move forward into a formal evaluation. A draft

[21] http://iccvam.niehs.nih.gov/methods/immunotox/llnadocs/CPSC_LLNA_nom.pdf
[22] http://iccvam.niehs.nih.gov/docs/about_docs/PL106545.pdf

background review document (BRD), which provides a comprehensive review of all available data and information, is prepared by NICEATM, in conjunction with an ICCVAM working group designated for the relevant toxicity testing area (e.g., the IWG). In addition, ICCVAM considers all of the available information and makes draft test method recommendations on the proposed usefulness and limitations of the test methods, test method protocol, performance standards, and future studies. The BRD and the draft ICCVAM test method recommendations are made available to the Panel and the public for review and comment. The Panel peer reviews the BRD and evaluates the extent to which it supports the draft ICCVAM test method recommendations. A Panel report is published, which is then considered, along with public and SACATM comments by ICCVAM in making final recommendations. These final recommendations are forwarded to the ICCVAM member agencies for their consideration and possible incorporation into relevant testing guidelines.

Dr. Stokes reviewed the ICCVAM criteria for adequate validation. He stated that validation is defined by ICCVAM as the process by which the reliability and relevance of a procedure are established for a specific purpose, and that adequate validation is a prerequisite for consideration of a test method by U.S. Federal regulatory agencies. Dr. Stokes listed the ICCVAM acceptance criteria for test method validation and acceptance. He concluded by summarizing the timeline of the review activities beginning with CPSC's nomination in January 2007 and ending with the present Panel meeting.

ICCVAM Charge to the Panel

Dr. Stokes reviewed the charge to the Panel, which was to: (1) review the draft BRDs, the draft Addendum to the traditional[23] LLNA, and the draft performance standards for completeness and identify any errors or omissions; (2) determine the extent to which each of the applicable criteria for validation and regulatory acceptance had been addressed for the proposed revised or modified versions of the LLNA; and (3) consider and provide comment on the extent to which the ICCVAM draft test method recommendations including the proposed use, standardized protocols, performance standards, and additional studies are supported by the information provided in the draft BRDs and draft Addendum.

Dr. Stokes thanked the IWG and ICCVAM for their contributions to this project, and acknowledged the contributions from the participating liaisons from ECVAM and JaCVAM (Japanese Center for the Validation of Alternative Methods). He also acknowledged the NICEATM staff for their support and assistance in organizing the Panel meeting and preparing the materials being reviewed.

[23] For the purposes of this document, the radioactive LLNA test method, which was first evaluated by ICCVAM in 1999, and subsequently recommended to U.S. Federal agencies as a valid substitute for currently accepted guinea pig test methods to assess the allergic contact dermatitis potential of many, but not all, types of substances, is referred to as the traditional LLNA.

Current Regulatory Testing Requirements and Hazard Classification Schemes for Allergic Contact Dermatitis and the Traditional LLNA Procedure

Dr. Joanna Matheson, Chair of the IWG, briefly reviewed the regulatory testing requirements of U.S. Federal agencies for skin-sensitization hazard identification and provided a brief description of the LLNA protocol.

Overview of the Agenda

Dr. Luster provided a brief synopsis of the agenda. He stated that there were six test methods and applications, along with the draft LLNA performance standards for review and that the same agenda would be followed for each: (1) introductory summary of the draft ICCVAM recommendations from one of the NICEATM staff members; in addition, test method developers would provide a brief description of the methodology for each of the three non-radioactive tests, (2) presentation of the Evaluation Group draft comments by the Evaluation Group leader, (3) Panel discussion, (4) public comments, (5) recommendations and conclusions by the Panel.

Overview of the Draft LLNA Limit Dose Procedure[24] BRD and Draft ICCVAM Test Method Recommendations

Dr. David Allen, Integrated Laboratory Systems, Inc., the NICEATM support contractor, presented an overview of the draft ICCVAM BRD for the LLNA limit dose procedure. He mentioned that the draft ICCVAM BRD provided a comprehensive review of the available data and information regarding the usefulness and limitations of the LLNA limit dose procedure. The method was reviewed for its accuracy in correctly identifying sensitizers and non-sensitizers, when compared to the traditional LLNA.

NICEATM published a series of *Federal Register (FR)* notices, including an *FR* notice (72FR27815, May 17, 2007) requesting original data from the LLNA. This *FR* notice was also sent to over 100 potentially interested stakeholders for their input and comment. As a result, data on 255 substances tested in the LLNA were received. The resulting LLNA database consisted of 471 studies of 466 unique substances, 211 of which were included in the original ICCVAM 1999 evaluation. Dr. Allen briefly summarized the performance characteristics of the LLNA limit dose procedure test method, which is detailed in the draft ICCVAM BRD[25], and briefly summarized the draft ICCVAM test method recommendations for the LLNA limit dose procedure[26]

Panel Evaluation:

Dr. Michael Olson led the Panel discussion on the LLNA limit dose procedure and specifically thanked the members of his Evaluation Group (i.e., Drs. James McDougal, Raymond Pieters, Jonathan Richmond [not present], and Takahiko Yoshida) for their collegial review of the information presented in the draft ICCVAM LLNA Limit Dose Procedure BRD. Dr. Olson also

[24] Also know as the reduced LLNA (rLLNA).
[25] http://iccvam.niehs.nih.gov/methods/immunotox/LLNA-LD/LLNAldBRD07Jan08FD.pdf
[26] http://iccvam.niehs.nih.gov/methods/immunotox/LLNA-LD/IWGrecLLNA-LD07Jan08FD.pdf

thanked the NICEATM staff for their technical support during the BRD review process. He then presented the draft responses to ICCVAM's questions to the Panel for consideration by the entire Panel. The focus was on review of the BRD for errors and omissions, assessment of the validation status of the test method, and review of draft ICCVAM test method recommendations. The Panel discussion and their recommended revisions to each section of the draft ICCVAM BRD and recommendations are reflected in the *Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products*, published in May 2008 (hereafter, the Panel report[27]).

During the Panel's evaluation, discussion arose regarding what might have resulted in the inverted-U-shaped dose response that was seen with the false-negative substances in the LLNA limit dose procedure. Dr. Olson responded that although it was difficult to understand what the cause might have been, he speculated that the top dose was either toxic at a systemic-effect level or that those substances were immunosuppressive at the highest dose level. He also stated that there did not seem to be any structural features of the substances that could be attributed for the false negative response in the LLNA limit dose procedure.

The Panel also discussed the use of concurrent versus intermittent positive controls in the LLNA limit dose procedure. Dr. Olson indicated that the Evaluation Group had discussed the possibility to allow intermittent positive controls for laboratories that exhibited repeatable and adequate performance with the LLNA but he indicated that it would be important to describe a set of performance criteria that would determine when this practice would be acceptable. Clearly, if the laboratory was not performing the assay routinely or if there were other reasons to suspect variability in response with any substance, the positive control would be necessary. Dr. Stokes indicated that this discussion was pertinent and indicated that the Panel's suggestions for what the performance criteria might be for intermittent positive control testing would be of interest to the IWG. Dr. Stokes also wanted to clarify that the OECD TG, is consistent with the EPA TG and the ICCVAM-recommended test method protocol for the LLNA although the OECD TG allows additional latitude in how tests are run (i.e., four animals per dose group, use of pooled data, and the option to not run a positive concurrent positive).

Public Comments:

Dr. Amy Rispin, EPA

Dr. Rispin stated that the ICCVAM LLNA report (1999[28]) and standardized protocol (2001[29]) recommends the use of a concurrent positive control in addition to the concurrent negative control required for each study. Subsequently, the OECD (Organisation for Economic Co-operation and Development) Test Guideline (TG) 429 (Skin Sensitisation: Local Lymph Node Assay) was finalized (2002). She said that originally, OECD TG 429 was drafted without a concurrent positive control but that language was added to include the recommended use of a concurrent positive control until laboratories demonstrate competence. Subsequent to that, EPA

[27] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf
[28] http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf
[29] http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/LLNAProt.pdf

put forth its LLNA guideline for sensitization[30], which states that concurrent positive and negative controls are to be included in each study. Dr. Rispin then added that U.S. Federal regulatory agencies, most notably the EPA and FDA, received LLNA data from studies in which the positive control did not achieve the appropriate limits of performance (i.e., the control values were not in the appropriate range) and therefore the studies were deemed unacceptable, underscoring the importance of a concurrent positive control for regulatory acceptance in the United States.

In response to Dr. Rispin's public comment, Drs. Ullrich and Theran asked how competence is determined and if laboratories have difficulties reaching a level of competence, respectively. Dr. Abby Jacobs responded by stating that the FDA has seen large data variations in laboratories that conduct the LLNA. It is often difficult to determine what the variations might be due to (e.g., new technicians, tail vein injection, lymph node removal) and these variations have been seen both in laboratories that are established and those that are not.

Dr. David Basketter, ECVAM Observer

Dr. Basketter said that the main point he wanted to address is that efforts should be made to harmonize the LLNA protocol with that described in OECD TG 429. He stated that although there is referral to the "ICCVAM protocol" throughout the BRDs under consideration, OECD TG 429 is more globally recognized for regulatory use of the LLNA and therefore should be the referenced protocol. Dr. Basketter further stated that if the LLNA limit dose procedure followed the ICCVAM protocol using five animals per group instead of following OECD TG 429, which allows using four animals per group, there would only be a savings of one animal for substances that were negative. He stated that the goal of ECVAM was actually to halve the number of animals by omitting the mid- and low-dose groups and that this would achieve significant animal savings since the likely prevalence of non-sensitizers is approximately two-thirds of chemicals tested and non-sensitizers would not require further testing even if dose response information for sensitizers was needed.

Dr. Basketter also mentioned that the retrospective evaluation of the LLNA being presented to the Panel analyzed whether the top dose could identify a substance as a sensitizer and how that compares to the traditional LLNA's performance. Since the traditional LLNA assay was determined to be positive or negative based on a stimulation index (SI) of three, it is problematic if the focus is on statistics when using the five-animal model as this would require also going back and re-evaluating all the preceding data using the statistical approach.

Dr. McDougal responded to Dr. Basketter's comment by stating that one wouldn't have to go back and retrospectively re-evaluate previous data but that new data generated could be analyzed statistically. This approach would include determining if the treatment group was statistically different from the vehicle control group and then determining the biological relevance. This might help to eliminate irritants.

[30] http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Revised/870r-2600.pdf

Panel Conclusions and Recommendations:

Dr. Luster asked the Panel to review the conclusions and recommendations for the LLNA limit dose procedure they had discussed earlier and to make any revisions, if necessary. One particular question that was asked during the Panel's conclusions and recommendations was whether an OECD TG existed for the LLNA limit dose procedure. Dr. Stokes indicated that the OECD TG would need to be updated to allow for the provision of a limit dose procedure and that's why the Panel's conclusions and recommendations are even more relevant. Dr. Stokes indicated that ICCVAM has already submitted a proposal to update the OECD TG based on the outcome of these deliberations and recommendations from the IWG.

The Panel agreed to use the term weight-of-evidence to refer to existing information that would aid the LLNA limit dose procedure in identifying a substance as a sensitizer or a non-sensitizer. The Panel also discussed the use of concurrent positive controls and recommended that a laboratory that is proficient at conducting the limit dose procedure can test a positive control at routine intervals rather than concurrently (although the Panel did not identify what constituted routine intervals). The Panel also discussed the use of individual versus pooled data and agreed with the ICCVAM-recommended protocol that individual animal data should always be collected. The Panel concluded that individual animal response data are necessary in order to allow for statistical analyses of any differences between treated and control data. In addition, having data from individual animals also allows for identification of technical problems and outlier animals within a dose group. Dr. Luster asked the Panel if they agreed with the changes and revisions made at this point and with the Panel conclusions and recommendations as presented and revised. The Panel unanimously agreed. The Panel's detailed recommendations and conclusions on the LLNA limit dose procedure are included in their final Panel report[31].

Overview of the Draft Addendum for the Applicability Domain of the LLNA and Draft ICCVAM Test Method Recommendations

Dr. Eleni Salicru, Integrated Laboratory Systems, Inc. (the NICEATM support contractor), summarized the information provided in the draft ICCVAM Addendum to the ICCVAM LLNA report (1999). This Addendum provided an updated assessment of the validity of the LLNA for testing the sensitizing potential of mixtures, metals, and aqueous solutions. The database used for this evaluation contained traditional LLNA data submitted as part of the original LLNA evaluation (ICCVAM 1999), data extracted from peer-reviewed articles published after the original evaluation, and data submitted to NICEATM in response to the *FR* notice (72 FR 27815, May 17, 2007) requesting such data. Dr. Salicru then summarized the performance characteristics of the LLNA when used to test mixtures, metals, and aqueous solutions[32], as well as the draft ICCVAM test method recommendations for each of the three categories of test substances[33].

[31] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf
[32] http://iccvam.niehs.nih.gov/methods/immunotox/LLNA-app/LLNAappADD19Jan08FD.pdf
[33] http://iccvam.niehs.nih.gov/methods/immunotox/LLNA-app/LLNAappRecs19Jan08FD.pdf

Panel Evaluation:

Dr. McDougal, on behalf of his Evaluation Group, presented for consideration by the entire Panel the draft responses to the questions asked of the Panel by ICCVAM. The Panel then discussed the completeness of the draft ICCVAM Addendum, identified any errors and omissions, and reviewed the draft ICCVAM test method recommendations with regard to the ability of the LLNA to be used to test the sensitizing potential of mixtures, metals, and aqueous solutions. The Panel discussion and their recommended revisions to each section of the draft ICCVAM Addendum are reflected in the Panel report, published in May 2008[34]. During the Panel's evaluation of the LLNA's applicability domain, the difficulty of testing metals in the LLNA was discussed and Dr. Woolhiser asked if testing metals was also problematic in the guinea pig. Dr. Api indicated that with the metals, most of the data has come from the clinical experience because animal studies are not predicting accurately what is happening in the clinic. Dr. Maibach indicated that metals have been tested in the guinea pig and that they are sensitized easily. Dr. Maibach further commented that metals in man need to be patch-tested for clinical relevance at a level close to the irritant dose and that a thoughtful series of algorithms is necessary to determine this. He also pointed out that patch test results to some metals (e.g., nickel, palladium) may indicate that a cell mediated reaction is occurring (i.e., contact allergy) but it needs to be sorted out if this cell mediated reaction actually results in a disease (i.e., allergic contact dermatitis) and this is where the LLNA could prove useful.

With regard to mixtures, Dr Api commented that based on her experience, when the mixture tested in the LLNA contains a predominant material (loosely defined that as greater than 70 percent) then the LLNA for the mixture mirrors what occurs for that one material. When evidence indicates that the substance is a true mixture, some times the LLNA does what is expected and other times the results are unexpected. In those cases, a weight-of-evidence approach (e.g., structure-activity relationships, clinical evidence) is employed.

Public Comments:

Dr. Charles Hastings, BASF Corporation

Dr. Hastings, representing CropLife America (an industry association of companies in the crop protection business), provided an overview of current activities in industry related to the use of the LLNA to detect dermal sensitizers and the global issues that are of importance. Dr. Hastings mentioned that CropLife America's primary concern is the testing of pesticide mixtures and formulations. He stated that they support the use of the LLNA for testing the dermal sensitization of mixtures and formulations as well as single ingredients.

Dr. Hastings mentioned that in the United States, EPA OPPTS (Office of Prevention, Pesticides and Toxic Substances) Guideline 870.2600[35] allows for the use of the LLNA as the preferred alternative to the standard guinea pig test. Based on this recommendation, member companies of CropLife America conducted a large number of LLNA studies for both active ingredients

[34] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

[35] http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Revised/870r-2600.pdf

and formulations in the European Union (E.U.) and were at the point of submitting data in the United States, as well. Then, in early 2007, they were informed that EPA had concerns about the validity of using the LLNA to test mixtures and formulations, and were advised to discontinue using this test method for that purpose until it had been adequately validated. Dr. Hastings stated that, in contrast to the EPA, E.U. regulators consider the LLNA acceptable for testing pesticide formulations and actually prefer it to a guinea pig test.

Dr. Pieters asked if the E.U. has conducted any evaluations of the validity of the LLNA for testing mixtures and formulations. Dr. Hastings replied that he was not certain if they had performed an extensive evaluation or not but that the E.U. considered the LLNA a validated method and therefore likely considered it appropriate to test not only the active ingredient but also the formulation or mixture.

Dr. Hastings mentioned that one concern in terms of using the LLNA for testing mixtures or formulations, particularly in the E.U., is the testing of aqueous substances. Many of the industry formulations are aqueous-based and may be incompatible with traditional LLNA vehicles. The European Crop Protection Association sponsored a study that evaluated the use of an aqueous vehicle known as Pluronic L92, which helps adhere the test material to the mouse ear. In the study, they tested three aqueous pesticide formulations that contained known sensitizers, using Pluronic L92 as the vehicle. As expected, the test results demonstrated sensitizing activity. Regarding global considerations, Dr. Hastings mentioned that if the LLNA is not accepted for mixture/formulation testing in the United States, industry will have no choice but to conduct both the LLNA, with 18 to 24 animals, and a guinea pig test, with 20 to 30 animals, for each formulation they may develop for global distribution. This scenario counters the ICCVAM goal of "reducing, refining, and replacing" animal use in regulatory safety testing.

Dr. Hastings ended with the following conclusions:

- CropLife America believes the LLNA test can be used for pesticide formulations.
- CropLife America supports the efforts of EPA and ICCVAM to confirm the validity of the LLNA for testing mixtures/formulations and encourages a quick evaluation.
- CropLife America is willing to help, as needed.
- If and, when, it is determined that the LLNA is acceptable, CropLife America requests that EPA notify them so they can then begin conducting the LLNA again for the United States.

Dr. Api asked if CropLife America has data comparing pesticides that have been evaluated in the LLNA and in guinea pigs and/or humans. Dr. Hastings replied that they do and that generally there is not much discrepancy with guinea pig test results. Occasionally they might see a false positive compared to a guinea pig test, but he did not recall ever seeing a false negative. In most cases, they would feel comfortable accepting an occasional false positive because human health is still protected.

Dr. David Basketter, ECVAM Observer

Dr. Basketter stated that he had personal reservations about testing complex mixtures and formulations in assays that were designed for testing substances (e.g., the LLNA) since no single test has ever been validated for testing mixtures. On another point, he stated that most of the metals of importance have been tested in both the guinea pig and the LLNA and the "right" answers have been generated. Thus, it does not seem worthwhile to produce new tests with revised protocols for hazard and potency categorization for testing metals.

Panel Conclusions and Recommendations:

Dr. Luster asked the Panel if they agreed with the comments and recommendations that were made earlier during the Panel discussion. The Panel agreed with the draft ICCVAM recommendation for continued collection of information from traditional LLNA evaluations of mixtures, metals, and aqueous solutions with comparative data for guinea pig (i.e., guinea pig maximization test [GPMT] or Buehler test [BT]) and human (i.e., human maximization test [HMT] or human repeat insult patch test [HRIPT]) tests. However, the Panel suggested that, given resource limitations, it would be important to organize the recommendations based on relative priority. Dr. Luster asked the Panel if they agreed with this suggestion about prioritization of activities; all members of the Panel agreed with one abstention. Dr. Howard Maibach abstained from voting stating that he hoped this public meeting and the subsequent Panel report would emphasize to industry the need for them to submit more data on mixtures, metals, and aqueous substances in order to provide a clearer evidence of the validity of the LLNA in testing these types of substances. The Panel's detailed recommendations and conclusions on the applicability domain of the LLNA are included in their final Panel report[36].

Method Description and Overview of the LLNA: Daicel Adenosine Triphosphate (LLNA: DA) Test Method

Dr. Kenji Idehara, Daicel Chemical Industries, Ltd. (private limited company), summarized the technical aspects of the LLNA: DA test method. He described the LLNA: DA as a non-radioisotopic version of the LLNA method in which lymph node adenosine triphosphate (ATP) content is used as a measure of cell proliferation instead of radiolabeled thymidine incorporation. Dr. Idehara indicated that the LLNA: DA was developed six years ago at Daicel Chemical Industries, Ltd., and that they use the test method regularly for in-house assessments of the skin-sensitization potential of chemical materials, intermediates, or products. He summarized the protocol differences between the LLNA: DA and the traditional LLNA. In the LLNA: DA, the application site is treated with 1% sodium lauryl sulfate (SLS) one hour before each test substance (or vehicle control) application, and the test substance is applied to the test site on day 7 as well as on days 1, 2, and 3. The auricular lymph nodes are excised from individual animals on day 8 rather than on day 6 and the amount of ATP in the lymph nodes is measured with a luciferin-luciferase assay. Dr. Idehara mentioned that these modifications (i.e., 1% SLS pretreatment and additional application on day 7) enhance lymph node cell

[36] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

proliferation in order to achieve an SI = 3 in the LLNA: DA, which allows for a more direct comparison to the traditional LLNA.

Dr. Idehara mentioned that after excision, ATP content gradually decreased with time. Therefore, the overall assay time for measuring ATP content needs to be similar (i.e., within approximately 30 minutes) among all test animals. He noted that this was an important point for this method and recommended that the LLNA: DA be conducted by at least two persons. Dr. Idehara mentioned that ATP content assays are conducted using commercially available kits, and his laboratory has experience with two different commercial sources in Japan, Kikkoman and Lonzar.

Overview of the Draft LLNA: DA BRD and Draft ICCVAM Test Method Recommendations

Dr. Allen then presented an overview of the draft ICCVAM BRD for the LLNA: DA test method. He mentioned that the draft ICCVAM BRD provided a comprehensive review of the available data and information regarding the usefulness and limitations of the LLNA: DA to distinguish between sensitizers and non-sensitizers, compared to the traditional LLNA. The objective of the BRD was to describe the current validation status of the LLNA: DA test method, including its relevance and reliability, scope of substances tested, and the availability of a standardized protocol.

Dr. Allen mentioned that the data analyzed in the BRD included data provided by Daicel Chemical Industries, Ltd., on 31 substances tested at their laboratories. In addition, data for 14 different coded substances were generated from a two-phased interlaboratory validation study that included 17 total labs. Taken together, the total database represented in the LLNA: DA BRD included 33 different substances. Dr. Allen briefly summarized the performance characteristics of the LLNA: DA test method, which is detailed in the draft ICCVAM BRD[37]. Dr. Allen concluded by briefly summarizing the draft ICCVAM test method recommendations for the LLNA: DA test method[38].

Panel Evaluation:

Dr. Michael Woolhiser thanked the Panel members of his Evaluation Group (i.e., Drs. Nathalie Alépeé, Thomas Gebel, Sidney Green [not present], and Jean Regal) for their tireless efforts in reviewing their Evaluation Group's assigned documents. He also thanked the NICEATM staff for their technical support during the review process. Dr. Woolhiser then presented the draft responses to ICCVAM's questions about this test method for consideration by the entire Panel. This included their review of the draft BRD for errors and omissions, their overall assessment of the validation status of the test method, and their comments on the draft ICCVAM test method recommendations. The Panel discussion and their recommended revisions to each section of the draft ICCVAM BRD are reflected in the Panel report, published in May 2008[39].

[37] http://iccvam.niehs.nih.gov/methods/immunotox/llna-DA/LLNA-DAbrd07Jan08FD.pdf
[38] http://iccvam.niehs.nih.gov/methods/immunotox/llna-DA/LLNA-DARecs07Jan08FD.pdf
[39] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

Adjournment

The meeting was adjourned for the day at 5:03 p.m., to reconvene at 8:30 a.m., Wednesday, March 5, 2008.

Wednesday, March 5, 2008

Reconvening of the Panel Meeting

Dr. Luster reconvened the Panel Meeting at 8:30 a.m. He introduced himself and then asked that all Panel members, followed by all others in attendance, introduce themselves as well.

Overview of the Draft LLNA: DA BRD and Draft ICCVAM Test Method Recommendations

Panel Evaluation:

Dr. Woolhiser continued his presentation from the previous day of the draft responses to ICCVAM's questions to the Panel, for consideration by the entire Panel. The Panel discussion and their recommended revisions to each section of the draft ICCVAM BRD are reflected in the Panel report, published in May 2008[40]. Dr. Woolhiser indicated that the Evaluation Group had two main concerns with the LLNA: DA test method. The first concern related to pretreatment with 1% SLS and understanding how this impacted the biology of the response. Second, the time course of the study was different than the traditional LLNA because it extended the study by one day and included an additional challenge. This brought forth a question about the immunology of the response as it relates to the potential for elicitation and whether or not that is a significant change from the traditional LLNA, which is purely an induction model.

Public Comments:

Dr. George DeGeorge, MB Research Laboratories

In response to a question raised during the Panel discussion, Dr. DeGeorge commented that using lymph node weight as the readout to differentiate between sensitizers and non-sensitizers in the LLNA is problematic because although there are more lymph node cells packed into a node, each cell has less cytoplasm. The lymph nodes swell to a point, and then excrete water and become smaller lymphocytes that are countable. He cited examples from his laboratory with several different sensitizers, which demonstrate that lymphocytes in the node are smaller when a large SI (e.g., SI = 25) is obtained relative to when a smaller SI (e.g., SI = 3) is obtained.

Dr. DeGeorge also commented that he agreed with a point made during the Panel discussion that the LLNA: DA method and the LLNA: Bromodeoxyuridine Detected by ELISA (LLNA: BrdU-ELISA) method should be considered separately, because they are so dissimilar.

In his final comment, Dr. DeGeorge stated that in the traditional LLNA, in the LLNA: Bromodeoxyuridine Detected by Flow Cytometry (LLNA: BrdU-FC), and probably also in the LLNA: DA, strong sensitizing substances do not need to be administered three times. For instance, if one administers with a single, moderately high dose of dinitrochlorobenzene

[40] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

(DNCB) (i.e., one that would induce an SI of 20 to 40) and then measure lymph node cell proliferation on day 1, 2, 3, or 4, an increase in the number of cells in the node and the number of cells that are positive for BrdU would likely be observed. Thus, administrations of additional applications have the potential to cause cumulative irritation. Dr. DeGeorge stated that the LLNA: DA method, which extends the assay to eight days instead of six days, should evaluate what happens to lymph node cell number at earlier sample times. In addition, if the animals receive just one application using a high dose, with or without the SLS, is there an increase in the SI? If so, that would lead to the possibility that the extra applications are not necessary and might lead to cumulative irritation.

Dr. David Basketter, ECVAM Observer

Dr. Basketter made a statement that from a clinical perspective, substances are typically described as significant sensitizers or not significant sensitizers, and within that latter group some of the substances may indeed be non-sensitizing. Thus, just because a substance has been shown in an isolated case report to be a human sensitizer does not mean that there is sufficient evidence to consider it as positive for comparison with outcomes of predictive assays. It has to be of sufficient importance (i.e., potency) to trigger a positive classification. Dr. Basketter mentioned SLS, methyl salicylate, and isopropanol, as substances which will always be positive in some human cases although they shouldn't be positive in a predictive assay.

Dr. Basketter also commented that caution should be given to making sensitization assumptions based on chemical class references. As an example, eugenol and isoeugenol are structurally similar and have similar physical properties, but they act by different chemical reaction mechanisms and could fit into distinctly different chemical classes.

Dr. Basketter's last comment acknowledged that much work has been done in terms of validating the traditional LLNA. If one makes minor changes to the LLNA in terms of a different readout for proliferation, then they benefit from all the experience generated in validating the traditional LLNA and less effort is needed to prove that the minor modification is valid. In contrast, if more significant modifications are made, one cannot rely on that same experience. Dr. Basketter cautioned that more importance should be placed on distinguishing whether something has changed substantially enough such that you can no longer rely on the traditional LLNA as a reference.

Dr. Masahiro Takeyoshi, Chemicals Evaluation and Research Institute

Dr. Takeyoshi made a short presentation about differences in LLNA sensitization responsiveness among different strains of mice. He mentioned that this was an important issue when evaluating the modified LLNA methods being developed in Japan. He showed differences in responsiveness among three different mouse strains commonly used in Japan (i.e., BALB/cAnN, CBA/JN, and CD-1) tested with parabenzoquinone in his group's non-radioactive LLNA (i.e., LLNA: BrdU-ELISA). The data indicated that the CBA/JN mouse strain exhibited a higher responsiveness, as indicated by an increased SI, to parabenzoquinone than the other two mouse strains tested. Based on these results, CBA/JN mice were chosen for testing substances in the LLNA: BrdU-ELISA test method. Dr. Takeyoshi also indicated that based on evaluating different SI cutoffs in the LLNA: BrdU-ELISA, 2-

mercaptobenzothiazole, 3-(4-isopropylphenyl)isobutyraldehyde, and hydroxycitronellal, had low responsiveness (i.e., SI values). He noted that 2-mercaptobenzothiazole is an OECD TG 429 recommended positive control for the LLNA however repeat tests could not detect this substance as positive when using an SI value of 1.7 or more. Dr. Takeyoshi suggested that a substance-specific lower response might exist in the test system. Dr. Takeyoshi also summarized LLNA data by Dr. Ullmann and coworkers with the contract lab RCC, Ltd. in which they investigated the responsiveness of six different mouse strains (CBA/CaOlaHsd, CBA/Ca (CruBR), CBA/Jlbm (SPF), CBA/JNCrj, BALB/c and NMRI) to 25% 2-mercaptobenzothiazole. The data indicated that CBA/JNCrj mice showed markedly lower responsiveness compared to the other strains tested. These studies indicate that strain related differences would not be negligible with regard to measuring different endpoints of cellular proliferation in the LLNA because depending on the chemicals tested, responsiveness might be potentially impacted. For instance, some of the discordance seen in the LLNA: DA test method (e.g., 2-mercaptobenzothiazole) could be a strain specific effect.

Panel Conclusions and Recommendations:

Dr. Luster asked the Panel to review their conclusions and recommendations and discuss any revisions, if necessary. The Panel viewed the difference in treatment schedule between the LLNA: DA and the traditional LLNA to potentially be significant if the treatment schedule for the LLNA: DA corresponds to entering the elicitation phase of skin sensitization. The Panel was concerned that the 1% SLS pretreatment step in the LLNA: DA might modify the inherent sensitivity of the LLNA. They recommended that the test method developer (Daicel Chemical Industries, Ltd.) justify the use of 1% SLS or consider an alternative decision criterion (i.e., an SI threshold other than three) such that the 1% SLS pretreatment is no longer necessary. Dr. Luster asked the Panel if they agreed with the recommendations and conclusions that the Panel made along with the revisions; unanimously, the Panel agreed. The Panel's detailed recommendations and conclusions on the LLNA: DA test method are included in their final Panel report.[41]

Method Description and Overview of the LLNA: BrdU-FC Test Method

Dr. George DeGeorge, MB Research Laboratories, presented an overview of the LLNA: BrdU-FC test method. He stated that mice are dosed topically on the ears once daily for three consecutive days (i.e., days 1, 2, and 3), just like the traditional LLNA protocol. On day 6, the mice receive an intraperitoneal injection with bromodeoxyuridine (BrdU), and five hours later, the auricular lymph nodes are removed. The lymph nodes from individual animals are processed and, using flow cytometry, the number of BrdU-positive cells are counted from treated animals and compared to control animals as a measure of lymph node cell proliferation.

Dr. DeGeorge described in detail how the cells are processed and gated for flow cytometric analysis. He mentioned that the cells are also permeabilized and treated with propidium iodide which allows gates to be drawn around the G_0, G_1, S, and G_2M phases of the cell cycle. Dr.

[41] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

DeGeorge projected specific examples of flow cytometry plots and histograms for DNCB, hexyl cinnamic aldehyde (HCA), and positive and negative control data.

Dr. DeGeorge also described the tiered protocol for the assessment of sensitization potential using the LLNA: BrdU-FC and how ear swelling measurements and additional immunophenotypic endpoints (i.e., the enhanced LLNA: BrdU-FC) aid in distinguishing skin irritants from an irritating sensitizer.

Overview of the Draft LLNA: BrdU-FC BRD and Draft ICCVAM Test Method Recommendations

Dr. Judy Strickland, Integrated Laboratory Systems, Inc. (the NICEATM support contractor), presented an overview of the draft ICCVAM BRD for the LLNA: BrdU-FC test method. She stated that the draft ICCVAM BRD provided a comprehensive review of the available data and information regarding the usefulness and limitations of the LLNA: BrdU-FC test method. Specifically, the test method was reviewed for its ability to distinguish between sensitizers and non-sensitizers compared with the traditional LLNA. The objective of the BRD was to describe the current validation status of the LLNA: BrdU-FC test method, including its relevance and reliability, scope of substances tested, and the availability of a standardized protocol.

Dr. Strickland indicated that MB Research Laboratories submitted data to NICEATM for the 48 substances analyzed in the BRD in response to an *FR* notice (72FR27815, May 17, 2007) that requested such data. Dr. Strickland briefly summarized the performance characteristics of the LLNA: BrdU-FC test method, which is detailed in the draft ICCVAM BRD[42], and the draft ICCVAM test method recommendations for the LLNA: BrdU-FC test method[43].

Panel Evaluation:

Dr. Raymond Pieters, on behalf of his Evaluation Group, presented the Evaluation Group's review of the draft BRD and the draft test method recommendations for the LLNA: BrdU-FC test method. Specifically, he presented the draft responses to ICCVAM's questions to the Panel for consideration by the entire Panel. This included their review of the draft BRD for errors and omissions, their overall assessment of the validation status of this test method, and their comments on the draft ICCVAM test method recommendations. The Panel discussion and their recommended revisions to each section of the draft ICCVAM BRD are reflected in the Panel report, published in May 2008[44]. The applicability of the draft ICCVAM-recommended LLNA performance standards to the LLNA: BrdU-FC test method was discussed, particularly with regard to the number of substances tested in the LLNA: BrdU-FC method and whether more data would be necessary for review before the validation status of the assay could be determined. Dr. Stokes reminded the Panel that the proposed LLNA performance standards didn't exist when the studies for the LLNA: BrdU-FC test method were performed. The

[42] http://iccvam.niehs.nih.gov/methods/immunotox/fcLLNA/FC-LLNAbrd07Jan08FD.pdf
[43] http://iccvam.niehs.nih.gov/methods/immunotox/fcLLNA/FCLLNARecs07Jan08FD.pdf

[44] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

questions should be whether the adequacy of the substances that have been tested is sufficient or if more studies need to be done to cover any gaps that might exist (e.g., range of potencies or activity, chemical classes).

Public Comments

Dr. David Basketter, ECVAM Observer

Dr. Basketter commented on the statement that Dr. DeGeorge made during his overview of the LLNA: BrdU-FC test method that HCA is irritating. He said that he is not convinced it is a significant irritant. Based on previous data, they had to use 50% HCA in a 48 hour occlusive application in the guinea pig in order to produce a mildly irritating response. Dr. Api added to Dr. Basketter's comment by stating that RIFM has also not found HCA to be an irritant when tested up to 20% in humans.

Dr. Basketter also commented that in the draft BRD for the LLNA: BrdU-FC, resorcinol was noted to be negative in the traditional LLNA and this is not correct. Dr. Basketter's group published results in 2007 in the journal Contact Dermatitis that resorcinol is clearly positive in the traditional LLNA when tested at higher concentrations and therefore this should be corrected for the record.

Dr. George DeGeorge, MB Research Laboratories

Dr. DeGeorge wanted to clarify that the LLNA: BrdU-FC test method was compared to the traditional LLNA to determine if the LLNA: BrdU-FC was more predictive of skin-sensitization potential. He stated that in some cases it was better while in others it wasn't, but overall, using human data as the gold standard reference, the LLNA: BrdU-FC exceeded the traditional LLNA predictivity values and accuracy. He also noted that the additional endpoints included in the LLNA: BrdU-FC allow for them to distinguish irritating substances that typically are considered false positives in the LLNA.

Dr. DeGeorge also noted that since the LLNA: BrdU-FC is so similar to the traditional LLNA the issue of refinement and reduction in animal use is not immediately apparent but if the assay is done in as few as four mice per group with a periodic positive control (e.g., every six months) this represents a significant decrease in animal numbers compared to guinea pig tests. Furthermore, there is a refinement since mice are phylogenitically lower than guinea pigs, and undergo less pain and distress during the assay than guinea pigs undergo.

With regard to the discussion of coefficients of variation (CVs) and the 0.5x to 2.0x EC3 (i.e., the estimated concentration needed to produce a stimulation index of three) range, Dr. DeGeorge suggested that a larger range might be more reasonable because the current range is likely too restrictive.

Dr. George also noted that ICCVAM requires interlaboratory validation if a test method is to be transferred to other laboratories. With regard to the LLNA: BrdU-FC, it is a "me-too" assay and only has "minor" changes from the traditional LLNA and is currently only used in one laboratory. Therefore, the current dataset should suffice for determining the validity of the LLNA: BrdU-FC. In response to Dr. DeGeorge's comment, Dr. Stokes stated that if a method is only proposed to be used by one laboratory, having only intralaboratory data certainly would

suffice but if it was proposed for broader use (e.g., adopted or endorsed by regulatory authorities), then other laboratories would have to demonstrate interlaboratory reproducibility. Dr. Luster asked if there was any mechanism available so that a company or small laboratory could apply for funding to help support an inter-laboratory validation. Dr. Stokes indicated that they could nominate the test method for additional validation studies to ICCVAM. It would go through a nomination review process and a prioritization would be given to that. The nomination would then be considered by the member agencies as to whether funding would be provided.

Panel Conclusions and Recommendations:

Dr. Luster asked the Panel to review their conclusions and recommendations and discuss any revisions, if necessary. The Panel suggested that the utility of ear swelling or other methods to detect inflammation appeared warranted for inclusion in every variation of the LLNA (including the traditional LLNA), but should be further investigated before routine inclusion in the protocol is recommended. The Panel further agreed that the draft ICCVAM test method recommendations for future studies highlighted the unanswered questions raised by the available data set. Specifically, conducting interlaboratory studies as a part of the validation process is important.

The Panel considered the immunological markers suggested for the LLNA: BrdU-FC to be appropriate, but noted that other immunological markers for discrimination of irritant versus sensitization phenomena were also available. In general, for any future work, efforts should be made to decrease the variability and to thereby increase the power of the test in order to ensure that more animals were not needed relative to the traditional LLNA or other modified LLNA protocols.

Dr. Luster asked the Panel to indicate if they agreed with the recommendations and conclusions that the Panel made along with the revisions; the Panel unanimously agreed. The Panel's detailed recommendations and conclusions on the LLNA: BrdU-FC test method are included in their final Panel report.[45]

Method Description and Overview of the LLNA: BrdU-ELISA Test Method

Dr. Masahiro Takeyoshi, Chemicals Evaluation and Research Institute, presented an overview of the LLNA: BrdU-ELISA test method. He stated that the LLNA: BrdU-ELISA test method is very similar to the traditional LLNA test method. Unique to the LLNA: BrdU-ELISA test method, after test substance applications on days 1, 2, and 3, BrdU is injected interperitoneally on day 5. Approximately 24 hours after the BrdU injection, lymph nodes are collected, and detection of the amount of BrdU incorporated into the DNA of lymph node cells is conducted with an ELISA.

In the development process of this method, experiments were conducted to detect the most efficient injection schedule of BrdU. Based on the various injection schedules tested, a single injection protocol on day four was identified as the optimal injection schedule for BrdU administration.

[45] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

Dr. Takeyoshi then showed a video of laboratory personnel preparing the lymph node cells for BrdU detection by ELISA. He went on to describe data for the LLNA: BrdU-ELISA compared to the traditional LLNA and how performance could be improved using alternative decision criteria (i.e., an SI other than three as the threshold for a positive response).

Overview of the Draft LLNA: BrdU-ELISA BRD and Draft ICCVAM Test Method Recommendations

Dr. Salicru presented an overview of the draft ICCVAM BRD for the LLNA: BrdU-ELISA test method. She noted that the draft ICCVAM BRD provided a comprehensive review of the available data and information regarding the usefulness and limitations of the LLNA: BrdU-ELISA test method. Specifically, the test method was reviewed for its ability to distinguish between sensitizers and non-sensitizers compared with the traditional LLNA and guinea pig test methods. The objective of the BRD was to describe the current validation status of the LLNA: BrdU-ELISA test method, including its relevance and reliability, scope of substances tested, and the availability of a standardized protocol.

Dr. Salicru stated that data from a total of 29 substances were considered in the accuracy analysis for the LLNA: BrdU-ELISA, and they were all tested in one laboratory. Dr. Salicru briefly summarized the performance characteristics of the LLNA: BrdU-ELISA test method, which are detailed in the draft ICCVAM BRD[46], and the draft ICCVAM test method recommendations for the LLNA: BrdU-ELISA test method[47].

Panel Evaluation:

Ms. Kim Headrick, presented her Evaluation Group's (Drs. Anne Marie Api, Howard Maibach, Peter Theran, and Stephen Ullrich), review of the draft BRD and draft ICCVAM test method recommendations for the LLNA: BrdU-ELISA test method. Specifically, she presented the draft responses to ICCVAM's questions to the Panel for consideration by the entire Panel. This included their review of the draft BRD for errors and omissions, their overall assessment of the validation status of the test method, and their comments on the draft ICCVAM test method recommendations. The Panel discussion and their recommended revisions to each section of the draft ICCVAM BRD are reflected in the Panel report, published in May 2008.[48]

Public Comments:

Dr. David Basketter, ECVAM Observer

Dr. Basketter noted that when the traditional LLNA was first suggested as an alternative to the guinea pig tests, it went through a comprehensive validation process, and one of the concerns was that it should perform reliably and distinctly better than the guinea pig assays. He emphasized that this point should be kept in mind when thinking about the modified LLNA protocols with alternative end points that are currently being reviewed. He stated that the current rigor of examination for the modified LLNA protocols being reviewed for validation is

[46] http://iccvam.niehs.nih.gov/methods/immunotox/llna-ELISA/BrdUELISAbrd07Jan08.pdf
[47] http://iccvam.niehs.nih.gov/methods/immunotox/llna-ELISA/BrdUELISARecs07Jan08FD.pdf
[48] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

higher than that for the traditional LLNA. He speculated that in the not-too-distant future, *in vitro* alternatives are likely to be going through a similar review process and it is going to become ever more difficult to put these alternatives in place, not because there is ill-will against the selections but because of the high standard of being good scientists. Thus, it is important that pragmatic decisions are made using the tools that are available.

Dr. George DeGeorge, MB Research Laboratories
Dr. DeGeorge commented that he agreed with Dr. Basketter's statements. He said that based on his experience in this peer review process, it is unlikely that he would bring any of the three *in vitro* test methods that MB Research Laboratories is developing for consideration by ICCVAM, given the many high hurdles that have to be negotiated.

In response to the comments by Drs. Basketter and DeGeorge, Dr. McDougal commented that it does not seem unreasonable to raise the bar for what is expected of new or modified tests. Dr. Luster added that understandably, the focus on animal refinement and reduction is paramount, but that as scientists we have to ensure that the bar is maintained sufficiently high so that as the years go by scientific quality is not compromised.

Panel Conclusions and Recommendations:

Dr. Luster asked the Panel to review their conclusions and recommendations and discuss any revisions, if necessary. The Panel concluded that the available data and test method performance for the LLNA: BrdU-ELISA support the draft ICCVAM test method recommendations that it may be useful for identifying substances as potential skin sensitizers and non-sensitizers, but that more information and existing data must be made available before the LLNA: BrdU-ELISA can be recommended for use. The Panel also stated that a detailed protocol was needed, in addition to sufficient quantitative data for broader analysis on a larger set of balanced reference substances that take into account physicochemical properties and sensitization potency, as well as an appropriate evaluation of interlaboratory reproducibility.

The Panel's main concern with this test method was that the accuracy of the LLNA: BrdU-ELISA at SI $\geq$ 3 was inadequate and not equivalent to the traditional LLNA. Furthermore, although using a decision criterion of SI $\geq$ 1.3 improved the test's performance in identifying sensitizers from non-sensitizers, it did not resolve concerns about the test method, particularly considering that power calculations suggest a much larger number of animals per group would be required to identify a positive response. Thus, the Panel also concluded that it might be more appropriate to use a statistically based decision criterion rather than a stimulation index to classify substances as sensitizers, and that this should be further investigated. Dr. Luster asked the Panel to indicate if they agreed with the recommendations and conclusions that the Panel made along with the revisions; unanimously, the Panel agreed. The Panel's detailed recommendations and conclusions on the LLNA: BrdU-ELISA test method are included in their final Panel report[49].

[49] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

Overview of the Draft ICCVAM Performance Standards for the LLNA

Dr. Allen presented an overview of the draft ICCVAM Performance Standards for the LLNA. He briefly summarized the overall purpose of performance standards (i.e., to provide a basis for evaluating the performance of a proposed test method that is mechanistically and functionally similar to the validated test method) and the three elements encompassed within such performance standards (i.e., essential test method components, a minimum list of reference substances, and accuracy/reliability values). He noted that the proposed applicability of these draft ICCVAM LLNA performance standards is for the evaluation of LLNA protocols that deviate from the ICCVAM-recommended LLNA protocol only with respect to the method for assessing lymphocyte proliferation (e.g., using non-radioactive instead of radioactive reagents). Dr. Allen then provided an overview of the essential test method components, the minimum list of reference substances, and the accuracy/reliability values as detailed in the draft ICCVAM LLNA Performance Standards[50].

Panel Evaluation:

Dr. Woolhiser, on behalf of his Evaluation Group, presented the Evaluation Group's responses to the ICCVAM questions asked about the draft ICCVAM LLNA Performance Standards for the entire Panel to consider. The overall question for the Panel was whether these performance standards were considered adequate for assessing the accuracy and reliability of test method protocols that were based on similar scientific principles and that measured the same biological effect as the traditional LLNA. The Panel discussion and their recommended revisions to the draft ICCVAM LLNA Performance Standards are reflected in the Panel report published in May 2008[51].

Adjournment

The meeting was adjourned at 5:42 p.m., to reconvene at 8:30 a.m., Thursday, March 6, 2008.

Thursday, March 6, 2008

Reconvening of the Panel Meeting

Dr. Luster reconvened the Panel Meeting at 8:30 a.m. He introduced himself and then asked that all Panel members and all others in attendance introduce themselves as well.

Overview of the Draft ICCVAM LLNA Performance Standards

Panel Evaluation:

Dr. Woolhiser reviewed some of the important points highlighted during the previous day's discussion on this topic, and then continued to summarize the remaining comments of his Evaluation Group on the questions asked by ICCVAM on the draft ICCVAM LLNA Performance Standards for consideration by the entire Panel. As mentioned above, the Panel

[50] http://iccvam.niehs.nih.gov/methods/immunotox/PerfStds/LLNAPerfStd07Jan08FD.pdf
[51] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

discussion and their recommended revisions to the draft ICCVAM LLNA Performance Standards are reflected in the Panel report published in May 2008[52].

Dr. Woolhiser noted was that there were general comments on the topic order for the Panel's review. He asked if Dr. Stokes would comment on the rationale for the topic order. Dr. Stokes indicated that as the IWG deliberated the order of topics for this review, consideration was given to the fact that the three non-radioactive methods had undergone validation studies prior to the creation of LLNA performance standards. Thus, the non-radioactive test methods were reviewed before the performance standards, so as to not bias the Panel's assessment of each test method's performance. The performance standards could then be considered for their application to future test methods.

Public Comments:

Dr. Amy Rispin, EPA

Dr. Rispin stated that her intent was to provide some additional regulatory perspective on some of the points that have been discussed. When Federal agencies evaluate the validation status of a test method under ICCVAM, they conduct a comprehensive analysis of overall performance (i.e., accuracy and reliability) in the context of making regulatory decisions with data from the test method. Thus, in a regulatory situation, equal or greater accuracy compared to the reference test method is the expectation. If the number of animals can be decreased only at the expense of accuracy, the acceptability of such a test method for the particular regulatory purpose would need to be carefully considered. Certain methods, instead of being complete replacements, might have to be relegated to the role of screens, where positives would be accepted, but negatives would require further testing - a less than ideal situation.

Dr. Rispin commented that performance standards are the regulating agencies basis for the acceptability of variations of accepted test methods. If an agency receives data from a modified LLNA method that has not been reviewed and validated in the ICCVAM process, there is unlikely to be a comprehensive peer review of it within the agency, given resource limitations. Therefore, the question of major versus minor departures from the functional criteria is important to ICCVAM and its member agencies. One cannot anticipate that there will be anything other than these performance standards to adequately evaluate the usefulness and limitations of a new method.

Dr. David Basketter, ECVAM Observer

Dr. Basketter first commented on a point that Dr. Thomas Gebel alluded to during the Panel's discussion of the draft ICCVAM LLNA Performance Standards ,which was that if a new laboratory performed the traditional LLNA to assess 18 or 22 chemicals, they probably wouldn't get a complete match. Dr. Basketter disagreed with Dr. Gebel's statement and viewed that a competent laboratory performing the LLNA would get it 100% correct.

Dr. Basketter then provided some comments that he stated were "from the ECVAM perspective." He stated that the ECVAM performance standards tried to address adhering to a standard

[52] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

protocol and that any change to the protocol other than the method for evaluating lymph node proliferation (e.g., strain, species, number of applications, time) was considered not to be minor, and therefore such a protocol would not be applied to these performance standards. By restricting the performance standards to minor changes, ECVAM was trying to minimize the number of chemicals required to evaluate sensitivity. Furthermore, the EC3 value could be used to see if the test method could classify substances in the appropriate range of sensitization potency.

ECVAM initially chose their reference substances in order to determine whether a modified method (differing only in the method for measuring cell proliferation) would give the same answer as the traditional LLNA. Thus, there was no intent to compare to the guinea pig or human data.

Dr. Basketter speculated that it is doubtful that data from multiple LLNA studies on the same substance are available and therefore it is unlikely that much larger sample sizes from which to calculate mean EC3 values and associated ranges will be obtained.

Dr. Basketter concluded by stating that ECVAM will not include more false positives and false negatives in its list. It has included one false positive and false negative in order to harmonize with ICCVAM but they don't see an added statistical value of just having one more false positive and false negative.

Karen Hamernik, EPA

Dr. Hamernik concurred with the comments that Dr. Rispin made previously, that performance standards, if developed such that they are too generalized with respect to minor versus major changes, would be problematic for regulatory agencies when they are reviewing submissions that include data from a modified LLNA protocol. Dr. Hamernik also asked for clarification from the Panel on a statement made during their discussions that a test for concordance for measuring the accuracy of classification (i.e., yes/no answer) should be done and that a chemical-for-chemical match is not necessary. Dr. Flournoy responded that concordance is not absolute but a continuum. Dr. Luster further clarified that the Panel discussion was based on the fact that the traditional LLNA is not a perfect match when compared to the guinea pig tests. Because there are false negatives and false positives compared to the guinea pig, there should be some flexibility so that an absolute chemical-by-chemical match is not required. In addition, a scientifically valid explanation can be provided for any discordance. Dr. Stokes emphasized that this was an important point and that additional clarity on the differences between a chemical-by-chemical match and overall accuracy need to be carefully considered before the final test method accuracy requirements are defined.

Panel Conclusions and Recommendations:

Dr. Luster asked the Panel to review the conclusions and recommendations for the ICCVAM LLNA performance standards they had discussed earlier and to make any revisions, if necessary. The Panel indicated that modified LLNA protocols that are undergoing validation should contain essential test method components that follow the ICCVAM-recommended

protocol[53], unless adequate scientific rationale for deviating from this protocol was provided. The Panel also identified aspects of the LLNA that should be required as part of the test method validation process, if more extensive changes to the protocol are being considered: (1) application of the test substance to the skin with sampling of the lymph nodes draining that site, (2) measurement of cell proliferation in the draining lymph node, (3) absence of a skin reaction that could be indicative of the onset of the elicitation phase of skin sensitization, (4) data collected at the level of the individual animal to allow for an estimate of the variance within control and treatment groups,[54] and (5) if dose response information is needed, there are an adequate number of dose groups ($n \geq 3$) with which to accurately characterize the dose response for a given test substance.

The Panel also recommended that statistical tests to analyze the data might allow for a more accurate interpretation. They recommended that a suitable variance-stabilizing transformation (e.g., log transformation, square root transformation) be applied in all statistical analyses and in reporting summary standard deviations. The Panel also recommended that a more rigorous evaluation be conducted of what would be considered an appropriate range of ECt values (i.e., estimated concentration needed to produce a stimulation index that is indicative of a positive response) to include as a requirement. This would be a statistical evaluation that considers the variability of ECt values generated among the sensitizers included on the performance standards reference substances list and the statistical multiple comparisons problem.

Dr. Luster asked the Panel if they agreed with the changes and revisions made at this point and with the Panel conclusions and recommendations as presented and revised. The members of the Panel agreed with one abstention; Dr. McDougal abstained from voting stating that he still had a concern about what constitutes a "major/minor" change. The Panel's detailed recommendations and conclusions on the ICCVAM LLNA performance standards are included in their final Panel report.[55]

Overview of the Draft LLNA Potency Determinations BRD and Draft ICCVAM Test Method Recommendations

Dr. Strickland presented an overview of the draft ICCVAM BRD for the use of the LLNA to determine skin-sensitization potency. She mentioned that the draft ICCVAM BRD provided a comprehensive review of the available data and information regarding the usefulness and limitations of the LLNA as a stand-alone assay for hazard categorization of skin-sensitization potency. In the BRD, the LLNA was evaluated for its ability to categorize substances for skin-sensitization potency using EC3 values.

Dr. Strickland noted that the analyses conducted in the BRD were based on LLNA studies obtained from ICCVAM (1999), the published literature, and data received in response to an *FR* notice (72 FR 27815, May 17, 2007) requesting original data from the LLNA. As a result, the analyzed data included 170 substances with LLNA, human, and/or guinea pig data. Dr.

[53] http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/LLNAProt.pdf

[54] Individual animal data will allow the application of a formal statistical test, if deemed necessary, and will also allow power calculations associated with the modified LLNA test.

[55] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

Strickland noted that three sets of data were analyzed and briefly summarized the results which are detailed in the draft ICCVAM BRD.[56] Dr. Strickland also briefly summarized the draft ICCVAM test method recommendations for potency determinations.[57]

Panel Evaluation:

Ms. Headrick presented her Evaluation Group's draft responses to ICCVAM's questions to the Panel for consideration by the entire Panel. This included their review of the draft BRD for errors and omissions, their overall assessment of the validation status of the test method, and their comments on the draft ICCVAM test method recommendations. The Panel discussion and their recommended revisions to each section of the draft ICCVAM BRD and recommendations are reflected in the Panel report published in May 2008.[58]

During the course of the discussion on the potency applicability of the LLNA, Dr. Woolhiser asked what the basis for the human threshold concentration cutoff values of 250 and 500 $\mu g/cm^2$ were. Dr. Wind replied that a number of experts and clinicians from throughout the world went back and looked at what, in their countries, they demarcated as strong sensitizers. The proposed Globally Harmonized System of Classification and Labeling of Chemicals (GHS) subcategory guidance values for the LLNA, guinea pig tests (GPMT, BT) and human data (HMT and HRIPT) were made on the basis of an impact analysis of 175 chemicals. In addition, the two proposed cut-offs were evaluated by the GHS Expert Group on Sensitization based upon chemicals already regulated as strong sensitizers to ensure their inclusion within the GHS categorization scheme. Clinical members of the Expert Group also confirmed relevance of the cut-off values such that clinically important skin sensitizers fell into the appropriate subcategory. The proposed guidance values were also in line with the European Commission's Expert Working Group recommendations.

Public Comments:

Dr. David Basketter, ECVAM Observer

Dr. Basketter commented that reviewing the potency data by splitting it into pooled and unpooled groups could be interesting but might be difficult since the majority of available data likely comes from pooled groups. Furthermore, much of the deliberation concluding that individual animal data must be used, was derived from analyses based only or largely on pooled data from four animals.

Dr. Basketter further stated that he viewed the analyses, which make the assumption that the human threshold data is the gold standard, as fundamentally flawed. Human data comes from studies conducted at different times, with different protocols, according to varying quality standards, and by different people. Therefore, there is no definitive knowledge of the reproducibility of the data. However, he considers the analyses adequate for recommending the LLNA as a part of a weight-of-evidence decision on human sensitization potency categorizations.

[56] http://iccvam.niehs.nih.gov/methods/immunotox/LLNA-pot/LLNApotency18Jan08FD.pdf
[57] http://iccvam.niehs.nih.gov/methods/immunotox/LLNA-pot/LLNAPotencyRecs18Jan08FD.pdf
[58] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

Dr. Amy Rispin, EPA

Dr. Rispin noted that there has been much discussion about various ways of handling the potency data. The OECD expert task force on skin sensitization needs to see an analytical comparison of what is considered to be the most appropriate approach for evaluating the data. The question for categorization purposes is, *What is the ideal testing modality for separating strong versus weak sensitizers for potency categorization*? A regulator who must assign a categorization is going to be confronted with all available test data and must know which data should be given the greatest weight in their evaluation.

Dr. Rispin noted that the OECD task force also reviewed the draft BRD on potency determinations and sent a list of several questions to the Panel, some of which have been answered, many of which have not been. One of the questions is, can the LLNA protocols be refined (e.g., by selection of solvents or choice of other test parameters) to improve correlation? She concluded by noting that she hopes that the additional analyses that the Panel has suggested will bring some clarity to the matter.

Panel Conclusions and Recommendations:

Dr. Luster asked the Panel to review the conclusions and recommendations for the LLNA potency determinations they had discussed earlier and to make any revisions, if necessary. The Panel agreed with the draft ICCVAM recommendation that the LLNA should not be used as a stand-alone assay for categorizing skin sensitizers as strong versus weak, but that it could be used as part of a weight-of-evidence evaluation (e.g., along with quantitative structure-activity relationships, peptide reactivity, human evidence, historical data from other experimental animal studies) for this purpose. The Panel also agreed with ICCVAM's recommendation that any LLNA studies conducted for the purpose of evaluating skin-sensitization potency should use the ICCVAM-recommended LLNA protocol. In addition, the Panel stated that the relevant testing guidelines for the traditional LLNA should be revised to include the procedure for calculating an EC3 value. Dr. Luster asked the Panel if they agreed with the changes and revisions made at this point and with the Panel conclusions and recommendations as presented and revised; the Panel unanimously agreed. The Panel's detailed recommendations and conclusions on the LLNA potency determinations are included in their final Panel report.[59]

Concluding Remarks

Dr. Luster, on behalf of the Panel, thanked the NICEATM-ICCVAM staff for their continued assistance during the review process and the Panel meeting. He also thanked Drs. Joanna Matheson and Abby Jacobs, the IWG co-chairs, and Dr. Marilyn Wind, ICCVAM Chair and IWG member, for the hard work they put into the project. Dr. Luster also thanked the Panel and the Panel Chairs for their involvement in the huge task of reviewing seven topics. He commented that, for future reference for ICCVAM, the Panel in their individual groups were able to do a good job in reviewing the materials, but because they were so focused on their particular topics due to serious time constraints, there may not have been the full benefit of their expertise for other topics in all cases.

[59] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

Drs. Wind and Stokes thanked the Panel again for their hard work, thoughtful and objective deliberations, and advice. Dr. Stokes further thanked the invited test method developers for their excellent summaries of their method for the benefit of the Panel, and CPSC for hosting the Panel meeting. He mentioned that there has been discussion about obtaining additional existing data (i.e., on mixtures, on one or more of the non-radiolabeled test methods), and that should these data become available in a timely manner and if NICEATM is able to assimilate and analyze the data, the Panel might be reconvened by teleconference to review the data. Dr. Stokes concluded by saying he looked forward to further working with the Panel members to complete their Panel report.

Adjournment

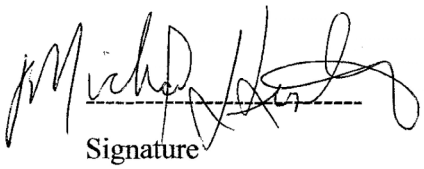
The meeting was adjourned and concluded at 3:20 p.m.

William S. Stokes, D.V.M.
NIEHS
P.O. Box 12233
MD-EC17
Research Triangle Park, NC 27709

Dear Dr. Stokes,

The Meeting Summary Minutes, Independent Scientific Peer Review Panel Meeting, Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay (LLNA): A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products, accurately summarizes the Peer Review Panel meeting of March 4-6, 2008, in Bethesda, MD.

Sincerely,

Signature	Printed Name	Date
	MICHAEL I LUSTER	11-18-08

Appendix G

Federal Register Notices and Public Comments

[This Page Intentionally Left Blank]

Appendix G1

Federal Register Notices

[This Page Intentionally Left Blank]

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); the Murine Local Lymph Node Assay: Request for Comments, Nominations of Scientific Experts, and Submission of Data

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Request for comments, submission of relevant data, and nominations of scientific experts.

SUMMARY: The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) received a nomination from the U.S. Consumer Product Safety Commission (CPSC) to evaluate the validation status of: (1) The murine local lymph node assay (LLNA) as a stand-alone assay for determining potency (including severity) for the purpose of hazard classification; (2) the ''cut-down'' or ''limit dose'' LLNA approach; (3) non-radiolabeled LLNA methods; (4) the use of the LLNA for testing mixtures, aqueous solutions, and metals; and (5) the current applicability domain (i.e., the types of chemicals and substances for which the LLNA has been validated). ICCVAM reviewed the nomination, assigned it a high priority, and proposed that NICEATM and ICCVAM carry out the following activities in its evaluation: (1) Initiate a review of the current literature and available data, including the preparation of a comprehensive background review document, and (2) convene a peer review panel to review the various proposed LLNA uses and procedures for which sufficient data and information are available to adequately assess their validation status. ICCVAM also recommends development of performance standards for the LLNA. At this time, NICEATM requests: (1) Public comments on the appropriateness and relative priority of these activities, (2) nominations of expert scientists to consider as members of a possible peer review panel, and (3) submission of data for the LLNA and/or modified versions of the LLNA.

DATES: Submit comments, data, and nominations by June 15, 2007. Relevant data will also be accepted after this date and considered when feasible.

ADDRESSES: Dr. William S. Stokes, NICEATM Director, NIEHS, P.O. Box 12233, MD EC–17, Research Triangle Park, NC 27709, (fax) 919–541–0947, (e-mail) *niceatm@niehs.nih.gov.* Courier address: NICEATM, 79 T.W. Alexander Drive, Building 4401, Room 3128, Research Triangle Park, NC 27709. Responses can be submitted electronically at the ICCVAM–NICEATM Web site: *http://iccvam.niehs.nih.gov/contact/FR_pubcomment.htm* or by e-mail, mail, or fax.

FOR FURTHER INFORMATION CONTACT: Other correspondence should be

directed to Dr. William S. Stokes (919–541–2384 or *niceatm@niehs.nih.gov*).

SUPPLEMENTARY INFORMATION:

Background

ICCVAM previously evaluated the validation status of the LLNA as a stand-alone alternative method to the Guinea Pig Maximization Test (GPMT) and the Buehler Assay (NIH publication No. 99–4494; available at *http://iccvam.niehs.nih.gov/methods/immunotox/llna.htm*). Based on this evaluation, ICCVAM recommended the LLNA as a valid substitute for the guinea pig methods for most testing situations. The Environmental Protection Agency, Food and Drug Administration, and the CPSC subsequently accepted the method as a valid substitute. The OECD also adopted the LLNA as OECD Test Guideline 429.

In January 2007, the CPSC submitted a nomination to NICEATM (*http://iccvam.niehs.nih.gov/SuppDocs/submission.htm*) requesting that ICCVAM assess the validation status of:

- The LLNA as a stand-alone test for potency determinations (including severity) for the purpose of hazard classification.
- LLNA protocols that do not require the use of radioactive materials.
- The LLNA ''cut-down'' or ''limit dose'' procedure.
- The ability of the LLNA to test mixtures, aqueous solutions, and metals.
- The current applicability domain (i.e., the types of chemicals and substances for which the LLNA has been determined to be useful).

Since 2003, ICCVAM has routinely developed performance standards for test methods; however, they were not developed for the LLNA, which was reviewed in 1999. Accordingly, ICCVAM proposes to now develop performance standards for the LLNA. Performance standards communicate the basis by which new proprietary and nonproprietary test methods have been determined to have sufficient relevance and reliability for specific testing purposes. Performance standards based on test methods accepted by regulatory agencies can be used to evaluate the reliability and relevance of other test methods that are based on similar scientific principles and measure or predict the same biological or toxic effect. On January 24, 2007, ICCVAM unanimously endorsed with a high priority: (1) Developing performance standards for the LLNA and (2) initiating a review of the available data and information associated with the CPSC nominated activities. A determination of which (if any) of the nominated activities will move forward will be made subsequent to this review and after consideration of comments by the public and the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM). If a decision is made to proceed with evaluation of these test methods, ICCVAM and NICEATM propose convening a peer review panel to review the usefulness and limitations of each of the LLNA methods listed above. The panel would also formulate conclusions on the adequacy of draft ICCVAM performance standards, any proposed future validation studies, and draft ICCVAM-proposed standardized test method protocols.

Request for Public Comments and Nominations of Scientific Experts

NICEATM requests public comments on the appropriateness and relative priority of the nominated activities. NICEATM also requests the nominations of scientists with relevant knowledge and experience to serve on the panel if a panel meeting occurs. Areas of relevant expertise include, but are not limited to: physiology, pharmacology, immunology, skin sensitization testing in animals, development and use of in vitro methodologies, biostatistics, knowledge about the use of chemical datasets for validation of toxicity studies, and hazard classification of chemicals and products. Each nomination should include the person's name, affiliation, contact information (i.e., mailing address, e-mail address, telephone and fax numbers), curriculum vitae, and a brief summary of relevant experience and qualifications.

Request for Data

NICEATM invites the submission of data from standard LLNA testing (i.e., OECD TG 429) with mixtures, aqueous solutions, and/or metals, as well as corresponding data from human and other animal studies. In addition, NICEATM invites the submission of data supporting the use of (1) the LLNA as a stand-alone test for determining potency (including severity) for the purpose of hazard classification, (2) the LLNA ''cut-down'' or ''limit dose'' procedure, and (3) LLNA protocols that do not require the use of radioactivity. Although data can be accepted at any time, data submitted by June 15, 2007, will be considered during the ICCVAM evaluation process. Submitted data will be used to further evaluate the usefulness and limitations of the LLNA and may be incorporated into future NICEATM and ICCVAM reports and publications as appropriate. The data will also be included in a database to support the investigation of other test methods for assessing skin sensitization.

When submitting chemical and protocol information/test data, please reference this **Federal Register** notice and provide appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, as applicable).

NICEATM prefers data to be submitted as copies of pages from study notebooks and/or study reports, if available. Raw data and analyses available in electronic format may also be submitted. Each submission for a chemical should preferably include the following information, as appropriate:

- Common and trade name.
- Chemical Abstracts Service Registry Number (CASRN).
- Chemical class.
- Product class.
- Commercial source.
- LLNA protocol used.
- Individual animal responses.
- The extent to which the study complied with national or international Good Laboratory Practice (GLP) guidelines.
- Date and testing organization.
- Sensitization data from other test methods.

Consideration by SACATM

On June 12, 2007, SACATM will meet at the Marriott Bethesda North Hotel and Conference Center in Bethesda, Maryland. The agenda includes consideration of the nominated LLNA activities, priorities, and proposed activities *http://ntp.niehs.nih.gov/go/7441*) and an opportunity for oral public comments. The SACATM meeting was announced in a separate **Federal Register** notice (**Federal Register** Vol. 72, No. 83, pp. 23831–32, May 1, 2007).

Background Information on ICCVAM and NICEATM

ICCVAM is an interagency committee composed of representatives from 15 federal regulatory and research agencies that use or generate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability and promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety and hazards of chemicals and products and that refine, reduce, or replace animal use. The ICCVAM Authorization Act of 2000 (42 U.S.C. 285l–3, available at *http://iccvam.niehs.nih.gov/about/PL106545.htm*) establishes ICCVAM as a permanent interagency committee of the

NIEHS under NICEATM. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of federal agencies. Additional information about ICCVAM and NICEATM is available on the following Web site: *http://iccvam.niehs.nih.gov*.

Dated: May 8, 2007.

David A. Schwartz,

Director, National Institute of Environmental Health Sciences and National Toxicology Program.

[FR Doc. E7–9544 Filed 5–16–07; 8:45 am]

BILLING CODE 4140–01–P

[This Page Intentionally Left Blank]

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Draft Performance Standards for the Murine Local Lymph Node Assay: Request for Comments

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Request for comments.

SUMMARY: The Murine Local Lymph Node Assay (LLNA) is the first alternative test method evaluated and recommended by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). It was subsequently accepted by regulatory authorities to determine the allergic contact dermatitis potential of chemicals and products. In January 2007, the U.S. Consumer Product Safety Commission (CSPC) submitted a nomination requesting that NICEATM and ICCVAM assess the validation status of (1) The LLNA as a stand-alone assay for potency determination for hazard classification purposes; (2) modified LLNA protocols; (3) the LLNA limit test; (4) the use of LLNA to test mixtures, aqueous solutions, and metals; and (5) the applicability domain for LLNA. In order to facilitate the review of the modified LLNA protocols, ICCVAM proposed developing performance standards for the LLNA. In May 2007, a **Federal Register** notice was published (Vol. 72, No. 95, pages 27815–27817, May 17, 2007) requesting comments and data relevant to these nominated activities. In June 2007, the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) endorsed the nominated activities as high priorities for ICCVAM. In response to SACATM comments, along with those provided by the public in response to the previous **Federal Register** notice, ICCVAM also endorsed these activities as high priorities. ICCVAM subsequently prepared draft performance standards for the LLNA and now requests public comments on this draft document, which is available on the NICEATM/ICCVAM Web site at: (*http://iccvam.niehs.nih.gov/methods/immunotox/immunotox.htm*) or by contacting NICEATM (see **FOR FURTHER INFORMATION CONTACT** below).

DATES: Submit comments on or before October 29, 2007.

ADDRESSES: Dr. William S. Stokes, NICEATM Director, NIEHS, P.O. Box 12233, MD EC–17, Research Triangle Park, NC 27709, (fax) 919–541–0947, (e-mail) *niceatm@niehs.nih.gov.* Courier address: NICEATM, 79 T.W. Alexander Drive, Building 4401, Room 3128, Research Triangle Park, NC 27709. Responses can be submitted electronically at the ICCVAM–NICEATM Web site: *http://iccvam.niehs.nih.gov/contact/FR_pubcomment.htm* or by e-mail, mail, or fax.

FOR FURTHER INFORMATION CONTACT: Other correspondence should be directed to Dr. William S. Stokes (919–541–2384 or *niceatm@niehs.nih.gov*).

SUPPLEMENTARY INFORMATION:

Background

The LLNA is an alternative test method used for skin sensitization testing that reduces the number of animals needed, reduces the time required for testing, and can substantially reduce or avoid pain and distress associated with traditional guinea pig testing methods. The LLNA was the first alternative test method evaluated and recommended by ICCVAM and based on the recommendations of ICCVAM and an independent scientific peer review panel, the LLNA has been accepted by U.S. and international regulatory authorities as an alternative to the guinea pig maximization test and Buehler test for assessing allergic contact dermatitis (EPA 2003; ISO 2002; OECD 2002). Since 2003, ICCVAM has routinely developed performance standards for test methods; however, because the concept of performance standards was not developed by ICCVAM until 2003, they were not developed during the ICCVAM evaluation of the LLNA in 1998 (NIH Publication No. 99–4494, available: (*http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf*).

In January 2007, CSPC submitted a nomination requesting that NICEATM and ICCVAM assess the validation status of (1) The LLNA as a stand-alone assay for potency determination for classification purposes; (2) modified LLNA protocols; (3) the LLNA limit test; (4) the use of LLNA to test mixtures, aqueous solutions, and metals; and (5) the applicability domain for LLNA. ICCVAM endorsed the nomination and also decided to develop performance standards to facilitate evaluation of modified LLNA protocols to the traditional LLNA. In May 2007, a **Federal Register** notice was published requesting comments and data relevant to these activities (Vol. 72, No. 95, pages 27815–27817, May 17, 2007; available,

[This Page Intentionally Left Blank]

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

National Toxicology Program (NTP); NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Announcement of an Independent Scientific Peer Review Panel Meeting on the Murine Local Lymph Node Assay; Availability of Draft Background Review Documents; Request for Comments

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Meeting announcement and request for comments.

SUMMARY: NICEATM in collaboration with the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) announces an independent scientific peer review panel meeting to evaluate modifications and new applications for the Murine Local Lymph Node Assay (LLNA). The LLNA is an alternative test method that can be used to determine the allergic contact dermatitis potential of chemicals and products. The panel will review the following:

- The validation status of three modified LLNA test method protocols that use non-radioactive probe chemicals.
- The validation status of a LLNA limit dose procedure.
- The use of the LLNA to test mixtures, aqueous solutions, and metals (applicability domain for the LLNA).
- The use of the LLNA to determine potency (potential for causing allergic contact dermatitis).
- Revised draft recommended performance standards for the LLNA.

At this meeting, the panel will peer review the draft background review documents and revised draft LLNA performance standards for each topic and evaluate the extent that established validation and acceptance criteria have been appropriately addressed. The panel will also comment on the extent

that the review documents support draft ICCVAM recommendations on proposed test method protocols, proposed uses of the LLNA, and the revised draft LLNA performance standards.

NICEATM invites public comments on the draft background review documents, draft ICCVAM test recommendations, draft test method protocols, and revised draft LLNA performance standards. All documents will be available on the NICEATM–ICCVAM Web site at *http://iccvam.niehs.nih.gov/methods/immunotox/immunotox.htm* by January 8, 2008.

DATES: The meeting is scheduled for March 4–6, 2008, from 8:30 a.m. to 5 p.m. each day. The meeting is open to the public free of charge, with attendance limited only by the space available. In order to facilitate planning for this meeting, persons wishing to attend are asked to register by February 20, 2008, via the NICEATM–ICCVAM Web site (*http://iccvam.niehs.nih.gov/contact/reg_LLNAPanel.htm*). The deadline for written comments is February 22, 2008.

ADDRESSES: The meeting will be held at the U.S. Consumer Product Safety Commission (CPSC) Headquarters, Bethesda Towers Bldg., 4330 East West Highway, Bethesda, MD.

FOR FURTHER INFORMATION CONTACT: Comments may also be submitted via the NICEATM–ICCVAM Web site at *http://iccvam.niehs.nih.gov/contact/FR_pubcomment.htm*. Comments or other correspondence can be sent to Dr. William S. Stokes, NICEATM Director, NIEHS, P.O. Box 12233, MD EC–17, Research Triangle Park, NC, 27709, (phone) 919–541–2384, (fax) 919–541–0947, (e-mail) *niceatm@niehs.nih.gov.* Courier address: NICEATM, 79 T.W. Alexander Drive, Building 4401, Room 3128, Research Triangle Park, NC 27709.

SUPPLEMENTARY INFORMATION:

Background

The LLNA is a reduction and refinement alternative test method for skin sensitization testing because it reduces the number of animals needed and can substantially reduce or avoid pain and distress compared to traditional guinea pig testing methods for sensitization. The LLNA was the first alternative test method evaluated and recommended by ICCVAM (NIH Publication No. 99–4494, available at: *http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf*). Based on the recommendations of ICCVAM and an independent scientific peer review panel, U.S. and international regulatory authorities have accepted the LLNA as an alternative to the guinea pig maximization test and Buehler test for assessing allergic contact dermatitis (ISO 2002; OECD 2002; EPA 2003). This review will evaluate the potential for broader use of the LLNA for regulatory testing of chemicals and products for allergic contact dermatitis potential, enabling further reduction and refinement (less pain and suffering) of animal use for this purpose. In January 2007, the CPSC submitted a nomination requesting that NICEATM and ICCVAM assess the validation status of (1) the LLNA as a stand-alone assay for potency determination for hazard classification purposes; (2) modified LLNA protocols; (3) the LLNA limit test; (4) the use of the LLNA to test mixtures, aqueous solutions, and metals; and (5) the applicability domain for the LLNA. In June 2007, the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) endorsed these activities as high priorities for ICCVAM. NICEATM on behalf of ICCVAM also sought input from the public on these activities (**Federal Register**: Vol. 72, No. 95, pages 27815–27817, May 17, 2007). After considering these inputs, ICCVAM endorsed these activities as high priorities. ICCVAM is also developing performance standards to facilitate evaluation of modified LLNA protocols compared to the traditional LLNA. Although ICCVAM has routinely developed performance standards for test methods since 2003, they were not developed as part of the ICCVAM evaluation of the LLNA in 1998. These draft performance standards for the LLNA were made public and comments were requested via the **Federal Register** (Vol. 72, No. 176, pages 52130–52131, Sept. 12, 2007). The May 2007 **Federal Register** notice requested data from studies using the LLNA or modified versions of the LLNA.

Drawing on the submitted data and literature sources, ICCVAM and NICEATM drafted background review documents for each of the modifications and new applications of the LLNA. ICCVAM has also developed draft test method recommendations regarding the proposed usefulness, limitations, and validation status of these test methods. ICCVAM will convene an independent scientific panel to peer review the draft background review documents for the test methods and determine whether the data and analyses in the draft documents support the draft ICCVAM test method recommendations. The panel will also be asked to comment on the adequacy of the revised draft performance standards, proposed future studies, draft standardized test method protocols, and recommended reference substances. NICEATM will ask the panel to consider all available information, including the scientific studies cited in the draft review documents, public comments, and any new information identified during the peer review, for developing their conclusions and recommendations.

Peer Review Panel Meeting

The purpose of this meeting is to conduct a scientific peer review of the revised draft performance standards and an evaluation of modifications and new applications for the LLNA. The LLNA is an alternative test method that can be used to determine the allergic contact dermatitis potential of chemicals and products. The panel will review the following:

- The LLNA as a stand-alone assay for potency determination for hazard classification purposes
- Modified LLNA protocols
- The LLNA limit test
- The use of the LLNA to test mixtures, aqueous solutions, and metals (applicability domain for the LLNA)
- The use of the LLNA to determine potency (potential for causing allergic contact dermatitis).

The panel will consider the draft background review documents for each of these methods and evaluate the extent that established validation and acceptance criteria are appropriately addressed for each test method (as described in the ICCVAM document, *Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods,* NIH Publication No. 97–981, available at *http://iccvam.niehs.nih.gov/docs/about_docs/validate.pdf*). The panel will then comment on the extent to which the draft ICCVAM recommendations are supported by the information provided in the background review document for each topic. It is anticipated that the panel will address the topics in the following order:

1. The LLNA limit test.
2. The applicability domain of the LLNA including its suitability for mixtures, aqueous solutions, and metals.
3. The LLNA as a stand-alone assay for potency determination for hazard classification.
4. The revised draft performance standards for the LLNA.
5. The modified LLNA test method protocols using non-radioactive materials.

Additional information about the meeting, including a roster of the panel members and the draft agenda, will be made available two weeks prior to the meeting on the NICEATM-ICCVAM Web site (*http://iccvam.niehs.nih.gov*). This information will also be available after that date by contacting NICEATM (see **FOR FURTHER INFORMATION CONTACT** above).

Attendance and Registration

This public meeting will take place March 4–6, 2008, at the CPSC Headquarters, Bethesda Towers Bldg., 4330 East West Highway, Bethesda, MD (an area map, driving directions, and CPSC contact information are available at *http://www.cpsc.gov/about/contact.html*). The meeting will begin at 8:30 a.m. and is scheduled to conclude at approximately 5 p.m. each day, although adjournment on March 6 may occur earlier or later depending upon the time needed for the expert panel to complete its work. It is also possible that the panel may conclude its deliberations on March 5 and not need to meet on March 6. Persons needing special assistance in order to attend, such as sign language interpretation or other reasonable accommodation, should contact 919–541–2475 (voice), 919–541–4644 TTY (text telephone, through the Federal TTY Relay System at 800–877–8339), or e-mail *niehsoeeo@niehs.nih.gov.* Requests should be made at least seven days in advance of the event.

Availability of the Draft Background Review Documents and Draft ICCVAM Recommendations

NICEATM prepared draft background review documents on each of these modifications or applications of the LLNA that describe the current validation status of the modified test methods and applications and contain all of the data and analyses supporting this proposed validation status. The draft background review documents, draft ICCVAM test method recommendations, draft test method protocols, and revised draft test method performance standards are available from the NICEATM–ICCVAM Web site (*http://iccvam.niehs.nih.gov/methods/immunotox/immunotox.htm*) or by contacting NICEATM (see **FOR FURTHER INFORMATION CONTACT** above).

Request for Public Comments

NICEATM invites the submission of written comments on the draft background review documents, draft ICCVAM test method recommendations, draft test method protocols, and revised draft test method performance standards. Written comments should be submitted preferably electronically via the NICEATM–ICCVAM Web site or by e-mail (*niceatm@niehs.nih.gov*); the deadline for submission of written comments is February 22, 2008. When submitting written comments, please refer to this **Federal Register** notice and include appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, if applicable). Written comments may also be sent by mail, fax, or e-mail to Dr. William Stokes (see **FOR FURTHER INFORMATION CONTACT** above). All comments received will be placed on the NICEATM–ICCVAM Web site (*http://iccvam.niehs.nih.gov*) and identified by the individual's name and affiliation or sponsoring organization (if applicable). Comments will also be sent to the panel and ICCVAM agency representatives and made available at the meeting.

This meeting is open to the public, and time will be provided for the presentation of oral comments by the public at designated times during the peer review. Members of the public who wish to present oral statements at the meeting should contact NICEATM (see **FOR FURTHER INFORMATION CONTACT** above) no later than February 20, 2008, and provide contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, if applicable). Up to seven minutes will be allotted per speaker, one speaker per organization. Persons registering to make comments are asked to provide NICEATM a written copy of their statement by February 27, 2008, so that copies can be distributed to the panel prior to the meeting. If this is not possible, please bring 40 copies of your comments to the meeting for distribution and to supplement the record. Written statements can supplement and expand the oral presentation.

Summary minutes and the panel's final report will be available following the meeting on the NICEATM–ICCVAM Web site (*http://iccvam.niehs.nih.gov*). ICCVAM will consider the panel's conclusions and recommendations and any public comments received when finalizing their test method recommendations and performance standards for these methods.

Background Information on ICCVAM and NICEATM

ICCVAM is an interagency committee composed of representatives from 15 Federal regulatory and research agencies that use or generate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability, and promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety and hazards of chemicals and products and that refine, reduce, or replace animal use. The ICCVAM Authorization Act of 2000 (42 U.S.C. 285l–3, available at *http://iccvam.niehs.nih.gov/docs/about_docs/PL106545.pdf*) establishes ICCVAM as a permanent interagency committee of the NIEHS under NICEATM. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM is available on the NICEATM–ICCVAM Web site at *http://iccvam.niehs.nih.gov.*

References

EPA. 2003. EPA OPPTS 870.2600 Test Guideline—Skin Sensitization. Available: *http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Drafts/870–2600.pdf.*

ISO. 2002. ISO 10993–10 Biological evaluation of medical devices—Part 10: Tests for irritation and delayed-type hypersensitivity. Geneva: International Organization for Standardization.

OECD. 2002. OECD Guideline for the Testing of Chemicals—Test Guideline 429: Skin Sensitization: Local Lymph Node Assay (adopted 24 April 2002). Paris: Organisation for Economic Co-operation and Development.

Dated: December 19, 2007.

Samuel H. Wilson,

Acting Director, National Institute of Environmental Health Sciences and National Toxicology Program.

[FR Doc. E7–25553 Filed 1–7–08; 2:42 pm]

BILLING CODE 4140–01–P

[This Page Intentionally Left Blank]

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Toxicology Program (NTP); NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Peer Review Panel Report on the Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay (LLNA): A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products: Notice of Availability and Request for Public Comments

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Request for comments.

SUMMARY: NICEATM, in collaboration with the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), convened an independent international scientific peer review panel on March 4–6, 2008 to evaluate new versions and applications of the LLNA for assessing the allergic contact dermatitis potential of chemicals and products. The peer review panel (''the Panel'') report from this meeting is now available. The report contains (1) the Panel's evaluation of the validation status of the methods and (2) the Panel's comments and conclusions on draft ICCVAM test method recommendations. NICEATM invites public comment on the Panel's report. The report is available on the NICEATM–ICCVAM Web site at *http://iccvam.niehs.nih.gov/methods/immunotox/llna_PeerPanel.htm* or by contacting NICEATM at the address given below.

DATES: Written comments on the Panel report should be received by July 7, 2008.

ADDRESSES: Comments should be submitted preferably electronically via the NICEATM–ICCVAM Web site at *http://iccvam.niehs.nih.gov/contact/FR_pubcomment.htm*. Comments can also be submitted by e-mail to *niceatm@niehs.nih.gov*. Written comments can be sent by mail or fax to Dr. William S. Stokes, Director, NICEATM, NIH/NIEHS, P.O. Box 12233, MD EC–17, Research Triangle Park, NC 27709, (phone) 919–541–2384, (fax) 919–541–0947. Courier address: NICEATM, 79 T.W. Alexander Drive, Building 4401, Room 3128, Research Triangle Park, NC 27709.

FOR FURTHER INFORMATION CONTACT: Dr. William S. Stokes, Director, NICEATM (919–541–2384 or *niceatm@niehs.nih.gov*).

SUPPLEMENTARY INFORMATION:

Background

In January 2007, the Consumer Product Safety Commission submitted a nomination to NICEATM and ICCVAM to assess the validation status of (1) The use of the LLNA to determine potency for hazard classification purposes; (2) LLNA protocols using non-radioactive procedures; (3) the LLNA limit dose procedure; and (4) the use of the LLNA to test mixtures, aqueous solutions, and metals (*i.e.*, an updated assessment of the applicability domain of the LLNA). In June 2007, the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) endorsed these activities as high priorities for ICCVAM. NICEATM, on behalf of ICCVAM, also sought input from the public on these activities and requested data from studies using the LLNA or modified versions of the LLNA (**Federal Register** Vol. 72, No. 95, pages 27815–27817, May 17, 2007). After considering all comments received, ICCVAM endorsed carrying out these activities as high priorities. ICCVAM also developed draft LLNA performance standards to facilitate evaluation of modified LLNA protocols that are functionally and mechanistically similar to the traditional LLNA. These draft LLNA performance standards were made public and comments were requested via the **Federal Register** (Vol. 72, No. 176, pages 52130–52131, Sept. 12, 2007).

ICCVAM and NICEATM prepared draft background review documents (BRDs) that provided comprehensive reviews of available data and relevant information for each of the modifications and new applications of the LLNA. ICCVAM also developed draft test method recommendations regarding the proposed usefulness and limitations, standardized protocols, and future studies. Both the draft BRDs and draft recommendations were made available for public comment, and a public peer review meeting was announced in the **Federal Register** (Vol. 73, No. 5, pages 1360–1362, Jan. 8, 2008).

The Panel met in public session on March 4–6, 2008. The Panel reviewed the draft ICCVAM BRDs for completeness, errors, and omissions of any existing relevant data or information. The Panel evaluated the information in the BRDs to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM, 2003) had been appropriately addressed. The Panel then considered the ICCVAM draft test method

recommendations (*i.e.*, proposed test method uses, proposed recommended standardized protocol, proposed test method performance standards, and proposed additional studies) and commented on whether the recommendations were supported by the information provided in the draft BRDs.

The Panel's conclusions and recommendations are detailed in the *Peer Review Panel Final Report: Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay (LLNA): A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products* (available at *http://iccvam.niehs.nih.gov/methods/immunotox/llna_PeerPanel.htm*). The draft BRDs, draft ICCVAM test method recommendations, and the draft LLNA Performance Standards are available at *http://iccvam.niehs.nih.gov/methods/immunotox/immunotox.htm*.

Request for Comments

NICEATM invites the submission of written comments on the Panel's report. When submitting written comments, please refer to this **Federal Register** notice and include appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, if applicable). All comments received will be made publicly available on the NICEATM–ICCVAM Web site at *http://ntp-apps.niehs.nih.gov/iccvampb/searchPubCom.cfm*. In addition, there will be an opportunity for oral public comments on the Panel's report during an upcoming meeting of SACATM scheduled for June 18–19, 2008. Information concerning the SACATM meeting will be published in a separate **Federal Register** notice and available on the SACATM Web site at *http://ntp.niehs.nih.gov/go/7441*.

ICCVAM will consider the Panel report along with SACATM and public comments when finalizing test method recommendations. An ICCVAM test method evaluation report, which will include the final ICCVAM recommendations, will be forwarded to relevant Federal agencies for their consideration. The evaluation report will also be available to the public on the NICEATM–ICCVAM Web site and by request from NICEATM (see **ADDRESSES** above).

Background Information on ICCVAM, NICEATM, and SACATM

ICCVAM is an interagency committee composed of representatives from 15 Federal regulatory and research agencies that use, generate, or disseminate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability and promotes scientific validation, regulatory acceptance, and national and international harmonization of toxicological test methods that more accurately assess safety and hazards of chemicals and products and that refine, reduce, and replace animal use. The ICCVAM Authorization Act of 2000 (42 U.S.C. 285*l*-3, available at *http://iccvam.niehs.nih.gov/docs/about_docs/PL106545.pdf*) established ICCVAM as a permanent interagency committee of the NIEHS under NICEATM. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM can be found at the NICEATM–ICCVAM Web site (*http://iccvam.niehs.nih.gov*).

Additional information about SACATM, including the charter, roster, and records of past meetings, can be found at *http://ntp.niehs.nih.gov/go/167*.

References

ICCVAM, 2003, ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03–4508. Research Triangle Park, NC: NIEHS. Available at: *http://iccvam.niehs.nih.gov*.

Dated: May 8, 2008.

Samuel H. Wilson,

Acting Director, National Institute of Environmental Health Sciences and National Toxicology Program.

[FR Doc. E8–11195 Filed 5–19–08; 8:45 am]

BILLING CODE 4140–01–P

Appendix G2

Public Comments Received in Response to *Federal Register* Notices

Responses to 72 FR 27815 (May 17, 2007)—The Murine Local Lymph Node Assay: Request for Comments, Nominations of Scientific Experts, and Submission of Data

Responses to 72 FR 52130 (September 12, 2007)—Draft Performance Standards for the Murine Local Lymph Node Assay: Request for Comments

Responses to 73 FR 1360 (January 8, 2008)—Announcement of an Independent Scientific Peer Review Panel Meeting on the Murine Local Lymph Node Assay; Availability of Draft Background Review Documents; Request for Comments

Responses to 73 FR 29136 (May 20, 2008)—Peer Review Panel Report on the Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay (LLNA): A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products: Notice of Availability and Request for Public Comments

No comments received

Date: Thu, 7 Jun 2007 09:00:44 -0400
Subject: NTP NICEATM Nomination of experts and response to call for data - LLNA
Ref.: Federal Register vol. 72 no. 95, p. 27815, 17 May 2007

Dr. Stokes -
Responding to the request for comment on the US CPSC proposal to ICCVAM-NICEATM for evaluation of the validation status of the murine local lymph node assay, I am pleased to submit the following information for consideration. (The views expressed in item 1.) below are solely my own and do not necessarily reflect the corporate position of GSK.)

1.) Appropriateness and relative priority of items comprising the proposed review of the status of the LLNA: It seems entirely justified that the proposed review should be undertaken based on the large volume of high quality peer-reviewed information published on performance, data evaluation and proposed protocol modifications of the LLNA in the period since the original ICCVAM-sponsored LLNA validation exercise. As proposed by US CPSC, ICCVAM-NICEATM preparation of a comprehensive background review should precede activation of a study panel. Regarding the priority of items for the background review as presented in the Federal Register notice, I suggest that the priority sequence should be slightly rearranged to highlight items 1, 5, 4, 2 and 3 (as identified in the Fed. Reg. notice) in priority sequence. Thus, from most to least pressing: 1. development of data to allow the LLNA to be used as a stand-alone tool in determining potency / severity of sensitising potential of chemicals; 2. evaluation and extension of the domain of applicability of the LLNA; 3. use of the LLNA for testing mixtures, aqueous solutions, and metals; 4. development of an animal-sparing cut-down approach to the LLNA focused on use of untreated vs. single high-concentration test group; and 5. assessment of the status of LLNA methods using non-radiolabeled tracer for end-point analysis.

2.) Nomination of expert scientists to serve on a possible LLNA review panel: I am pleased to offer the name of my GSK colleague Frederick J. Guerriero as a possible panel member. Mr Guerriero is a key member of the GSK Occupational Toxicology working group and in this capacity has had the responsibility of protocol development, study contracting and evaluation of a large number of LLN assays over the past 7-8 years. In addition, Mr Guerriero has previously served on the NICEATM study panel which evaluated *in vitro* alternatives for evaluation of ocular irritant/corrosion effects of chemicals. As a secondary potential candidate for the study panel, I would also be pleased to volunteer my service which is based in similar experience to that of Mr. Guerriero.

3.) Submission of LLNA data: Over the past 5 years GSK has transitioned to sole use of the LLNA as a means for evaluating the sensitising potential of a wide variety of chemical materials used in the synthesis of pharmaceuticals. The spectrum of substances which have been evaluated includes commodity chemicals used as starting materials, proprietary synthetic intermediates of varying structural complexity, and active pharmaceutical entities. All of these

assessments have been conducted by the "traditional" control + 3 concentration protocol using 3H-thymidine label. A small proportion of materials also have companion data evolved with the M&K or Beuhler dermal sensitisation protocol. Although the composite data are not presently in a readily transmitted form, I believe that we could be in position to share results of assessment of *ca.*190 chemicals if materials from the pharmaceutical sector would be of interest in the assessment which NICEATM is planning.

I will send this letter in print form with mailing today. I look forward to your reply in due course.

Sincerely yours -
Michael J. Olson, Ph.D.
Director, Occupational Toxicology
Corporate Environment, Health and Safety
GlaxoSmithKline

Subject: FR Notice Comments - 72FR27815 - LLNA
Date: Friday, June 15, 2007 1:43 PM

Dear Dr Stokes,

Safepharm Laboratories Ltd., UK (SPL) has conducted Local lymph node assays on behalf of sponsoring companies since 1997. The assays have been conducted on a wide variety of chemicals and chemical preparations. Since August 2002 the use of other animal models for evaluation of skin sensitisation potential for regulatory purposes (e.g. methods that require the use of guinea pigs) has been permissible in the UK only if a valid scientific reason can be provided as to why a LLNA cannot be conducted. In effect, the LLNA is the only method that can be used in the UK for assessment of skin sensitisation potential forregulatory purposes. We therefore support the proposed activities of ICCVAM-NICEATM as detailed in the Federal Register vol. 72, No. 95, p.27815-27817, 17 May 2007 in response to the U.S. CPSC nomination of January 10, 2007.

We have witnessed concerns in some areas of the chemical industry, with regard to the applicability of the LLNA for testing of preparations, mixtures and irritant substances, and also with regard to the fact that the LLNA has not always provided results consistent with existing knowledge of the test substance or related test substances. We do not know if all of these concerns are justified, but they can only serve to reduce confidence in the predictive capability of the assay. This is not desirable when the assay offers significant scientific and animal welfare advantages over guinea pig models for many product types, and in a country where the assay is effectively the only available method for evaluation of skin sensitisation potential for regulatory purposes. An assessment of the applicability domain of the assay in its current form and the use of the assay for testing mixtures, preparations, aqueous solutions, irritant substances and metals is therefore very much welcomed. It seems very appropriate to initiate a review of the current peer-reviewed literature and available data, in order to prepare a comprehensive background review document, conduct a review of the validation status of the LLNA for its various uses and to develop relevant performance standards.

It is noted that at its 26th meeting held on 26-27th April 2007 at the European Centre for the Validation of Alternative Methods (ECVAM), the non-commission members of ECVAM Scientific and Advisory Committee (ESAC) considered the reduced version of the LLNA (rLLNA) to be scientifically validated, but only when used as a screening test to distinguish between sensitisers and non-sensitisers and with due regard to the conditions set forth in the official ESAC statement of 27th April 2007. This statement was based on the outcome of a review of LLNA data for 211 chemicals[1]. The review of existing and newly-provided LLNA data proposed by ICCVAM-NICEATM therefore presents an ideal opportunity to assess further the validity of the rLLNA for screening purposes.

As a contract research organisation, SPL is unable to provide data for review by ICCVAM-NICEATM without the permission of its Sponsors, although we

consider it may be possible to provide a summary of study outcomes, coupled to general product type, should this be of interest to ICCVAMNICEATM.

In conclusion, Safepharm Laboratories Ltd. welcomes the proposed activities of ICCVAM-NICEATM in response to the U.S. CPSC nomination of January 10, 2007, and will be pleased to explore ways in which our experience may be of use in the process.

Yours sincerely,
Robert L. Guest
Head of Alternative and Acute Toxicology
Safepharm Laboratories Ltd.

[1] I Kimber, RJ Dearman, CJ Betts, GF Gerberick, CA Ryan, PS Kern, GY Patlewicz, DA Basketter (2006). The local lymph node assay and skin sensitization: a cut-down screen to reduce animal requirements? Contact Dermatitis 2006: 54:181-185

June 15, 2007

Dr William S Stokes
Director, NICEATM
National Institute of Environmental Health Sciences
PO Box 12233, MD EC-17
Research Triangle Park, NC 27709

Re: **72 FR 27815; May 17, 2007; National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); the Murine Local Lymph Node Assay (LLNA): Request for Comments, Nominations of Scientific Experts, and Submission of Data**

Dear Dr. Stokes:

These comments are submitted on behalf of the Alternatives Research and Development Foundation, the American Anti-Vivisection Society, Humane Society Legislative Fund, The Humane Society of the United States, People for the Ethical Treatment of Animals and the Physicians Committee for Responsible Medicine. The parties to this submission are national animal protection, health, and scientific advocacy organizations with a combined constituency of more than 10 million Americans who share the common goal of promoting reliable and relevant regulatory testing methods and strategies that protect human health and the environment while reducing, and ultimately eliminating, the use of animals.

In January, 2007, (ICCVAM) received a nomination from the U.S. Consumer Product Safety Commission (CPSC) to evaluate the validation status of: (1) The murine local lymph node assay (LLNA) as a stand-alone assay for determining potency (including severity) for the purpose of hazard classification; (2) the ''cut-down'' or ''limit dose'' LLNA approach; (3) non-radiolabeled LLNA methods; (4) the use of the LLNA for testing mixtures, aqueous solutions, and metals; and (5) the current applicability domain (i.e., the types of chemicals and substances for which the LLNA has been validated).

ICCVAM reviewed the nomination, assigned it a high priority, and proposed that NICEATM and ICCVAM carry out the following activities in its evaluation: (1) Initiate a review of the current literature and available data, including the preparation of a comprehensive background review document, and (2) convene a peer review panel to review the various proposed LLNA uses and procedures for which sufficient data and information are available to adequately assess their validation status. ICCVAM also recommends development of performance standards for the LLNA. At this time, NICEATM requests: (1) Public comments on the appropriateness and relative priority of these activities, (2) nominations of expert scientists to consider as members of a possible peer review panel, and (3) submission of data for the LLNA and/or modified versions of the LLNA.

At the meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) on June 12, 2007, several comments were made that suggested ICCVAM was assuming a relatively rapid review of these issues. However, this is not borne out by the CSPC

nomination which does not mention an expedited process. In addition, ICCVAM has recommended the creation of a background review document (BRD) and review by an expert peer review panel, with no mention of an expedited process. The cost/benefit of this LLNA review has not been evaluated, and SACATM was asked to vote to accept or reject NICEATM/ICCVAM's decision to proceed without offering any alternatives. Doubts about the cost/benefit of this project caused one SACATM member to vote against proceeding.

Despite the fact that ICCVAM documents, including the Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods,[1] mention the possibility of an expedited review process, it would appear that this process has only been used in one case. Despite repeated critiques of ICCVAM for failing to act expeditiously, we are still unable to locate a description of the expedited review process in ICCVAM literature and the parameters for applying it.

In light of the fact that the LLNA has been used by regulatory agencies for classifying skin sensitizers for years and both research data and regulatory use of the LLNA have been extensively reviewed in the literature, yet another review of this widely accepted method is unwarranted. The only circumstance under which this proposal is acceptable is if ICCVAM quickly reviews the existing literature and makes an expedited evaluation regarding the relevance of this information to Agency regulatory needs. ICCVAM's limited resources should be spent validating and promoting for regulatory acceptance any of the number of non-animal methods for skin sensitization that are currently in development.

In March 1999, ICCVAM published a final peer review report concluding that the LLNA is a valid alternative to currently accepted guinea pig test methods.[2] The U.S. EPA, FDA, and OSHA announced their acceptance of the LLNA as an alternative to the guinea pig maximization test for assessing allergic contact dermatitis in October 1999. That same year, ESAC, the Scientific Advisory Committee of the European Centre for the Validation of alternative Methods (ECVAM), also endorsed the LLNA for regulatory use.

In September 2000, the European Centre for Ecotoxicology and Toxicity of Chemicals (ECETOC) published a comprehensive review of sensitization test methods with respect to hazard identification and labeling, (and?) to determine whether the various methods are appropriate for determining relative potency and risk assessment.[3] The conclusions from this review included: (1) the LLNA is a viable and complete alternative to traditional guinea pig test methods for the purposes of skin sensitization hazard identification, and (2) the LLNA is suitable for the determination of relative skin sensitizing potency and the adaptation of this method for derivation of comparative criteria such as EC3 values provides an effective and quantitative basis for such measurements. This report further recommends that "the LLNA is the recommended method for new assessments of relative potency and/or for the investigation of the influence of vehicle or formulation on skin sensitizing potency."

[1] http://iccvam.niehs.nih.gov/SuppDocs/SubGuidelines/SD_subg034508.pdf
[2] http://iccvam.niehs.nih.gov/methods/immunotox/immunotox.htm
[3] ECETOC. 2000. Skin Sensitization Testing for the Purpose of Hazard Identification and Risk Assessment.

More recent work has further verified the use of the LLNA as a stand-alone method for estimating potency for regulatory purposes, including a 2005 study that concludes that there is a "clear linear relationship between LLNA-derived EC3 values and historical human skin patch data."[4] A 2007 review concludes that "The LLNA, when conducted according to published guidelines, provides a robust method for skin sensitization testing that not only provides reliable hazard identification in formation but also data necessary for effective risk assessment and risk management." In addition, a retrospective analysis of the regulatory use of the LLNA in the EU was published in 2006 and concluded that "the LLNA is satisfactory for routine regulatory use."[5] We acknowledge that the LLNA must be validated for determining sensitization potency for regulatory use; however, we urge ICCVAM to take an abbreviated test validation approach, as was recommended by the recent International Programme on Chemical Safety Workshop on Skin Sensitization in Chemical Risk Assessment:[6] "An abbreviated test validation approach may be appropriate to assess the validity of potency assessment based on the LLNA and its appropriateness for predicting sensitizing induction potency in humans."

The ''cut-down'' or ''limit dose'' LLNA approach (reduced, or rLLNA) has recently been reviewed by an ECVAM peer review panel. In April, 2007, ESAC issued a statement supporting the use of the rLLNA "within tiered-testing strategies to reliably distinguish between chemicals that are skin sensitizers and non-sensitizers "thereby reducing animal use by as much as 50%."[7] The statement also notes the following limitations: that "the test results provided by the rLLNA do not allow the determination of the potency of a sensitising chemical," and that "negative test results associated with testing using concentrations of less than 10% should undergo further evaluation"

The applicability and limitations of this modification of the LLNA have been clearly established. Therefore, in lieu of a lengthy review of this method, ICCVAM should expeditiously review and endorse the ESAC peer review and circulate harmonized testing recommendations regarding this assay to US agencies before year's-end and NICEATM should collaborate with ECVAM to address the question of concentration threshold.

Other recent work has included the development of several applications of non-radioactive detection methods for the LLNA, including BrdU incorporation, methods measuring the release of various cytokines, and methods using fluorescent markers and quantification by flow cytometry. In many cases, these methods have been shown to be as sensitive as protocols involving radio-labeling.[8] In addition, in NIH-sponsored and contract work, MB Research has shown that "for a large range of chemicals, the FC-LLNA EC3 values were consistent with

[4] Basketter et al. Predictive identification of human skin sensitization thresholds. Contact Dermatitis. 2005; 53 (5): 260-267.
[5] Cockshott et al., The local lymph node assay in practice: a current regulatory perspective. Hum Exp Toxicol 2006; 25 (7): 387-394.
[6] http://www.who.int/ipcs/methods/harmonization/areas/sensitization_summary.pdf
[7] http://ecvam.jrc.it/publication/ESAC26_statement_rLLNA_20070525-1.pdf
[8] Takeyoshi et al. Advantage of using CBA/N strain mice in a non-radioisotopic modification of the local lymph node assay. J Appl Toxicol. 2006. 26:5-9. Takeyoshi et al. Novel approach for classifying chemicals according to skin sensitizing potency by non-radioisotopic modification of the local lymph node assay. J Appl Toxicol. 2005. 25:120-134. Suda et al. Local lymph node assay with non-radioisotope alternative endpoints. J Toxicol Sci. 2002. 27:205-218.

those reported in ICCVAM LLNA validation studies."[9] Both ECVAM and Japanese Center for the Validation of Alternative Methods (JaCVAM) are currently reviewing these methods and, rather than initiate a full independent review, ICCVAM must collaborate with these ongoing efforts.

With regard to the assessment of the LLNA for aqueous mixtures and metals, the information that is currently available should allow ICCVAM to make a rapid determination of the applicability and limitations of the LLNA for these classes of chemicals and, if it cannot, we do not endorse further validation efforts in this regard, but recommend the pursuit of *in vitro* methods for this purpose.

Several non-animal methods for estimating sensitivity are under development, including quantitative structure activity relationship (QSAR) modeling that shows a high concordance with both guinea pig and LLNA data,[10] quantification of peptide reactivity, which also shows a high concordance with LLNA data,[11] and human cell cultures.[12] We urge ICCVAM to secure an interagency grant from the CPSC to fund the validation of one or more of these non-animal methods. Clearly, ICCVAM and the CPSC both benefit from the sharing of resources, as the CPSC nominated the method and ICCVAM will be tasked with the final work product.

ICCVAM should consider taking an approach similar to the European Sens-it-iv project,[13] which involves the coordinated efforts of more than two dozen groups from industry, academia and other organizations, all working toward the common goal of developing *in vitro* methods to assess immunotoxicity. ICCVAM should consider facilitating the creation of such a goal-oriented task force.

To summarize, given the fact that the LLNA has been used by regulatory agencies for classifying skin sensitizers for years and both research data and regulatory use of the LLNA have been extensively reviewed in the literature and by other countries, yet another lengthy review of this widely accepted method is clearly unwarranted. Instead, we urge ICCVAM to perform an expedited review of the existing information regarding the LLNA's performance and limitations and to issue recommendations to US agencies with all due speed. In the interest of eventual replacement of animals in sensitization testing, ICCVAM must spend its time and resources promoting the development and regulatory use of non-animal methods, which it can do by engaging in integrated approaches to *in vitro* immunotoxicity.

Sincerely,

[9] http://www.mbresearch.com/TOXNOTE/TOXNOTE-LLNA.pdf
[10] Fedorowicz et al., Structure-activity models for contact sensitization. Chem Res Toxicol. 2005; 18(6): 954-969.
[11] Gerberick et al. Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. 2007; 97(2): 417-427.
[12] Schoeters et al. Microarray analyses in dendritic cells reveal potential biomarkers for chemical-induced shin sensitization. 2007; 44(12): 3222-3233.
[13] http://www.sens-it-iv.eu/

Catherine Willett, PhD
Science Policy Advisor
Regulatory Testing Division
People for the Ethical Treatment of Animals

Sara Amundson
Executive Director
Humane Society Legislative Fund

Dr. Martin Stephens
Vice President for Animal Research Issues
Humane Society of the United States

Kristie Stoick, MPH
Research Analyst
Physicians Committee for Responsible Medicine

Sue A. Leary
President
Alternatives Research & Development Foundation

Tracie Letterman, Esq.
Executive Director
American Anti-Vivisection Society

[This Page Intentionally Left Blank]

Data Submissions

In response to the *FR* notices, the following public commenters provided data for evaluation of the rLLNA. Data are available in Appendix D (see Annex III of the Final Background Review Document).

- Dr. Dori Germolec, National Institute for Environmental Health Sciences
- Dr. Hans Werner Vohr, Bayer HealthCare (Wuppertal-Elberfeld, Germany)
- Dr. Melissa Kirk and Daniel Cervins, MB Research Labs
- Dr. Michael Olson, GlaxoSmithKline
- Dr. Kirill Skirda, European Committee of Organic Surfactants and their Intermediates
- Dr. Phil Botham, Syngenta (Cheshire, United Kingdom)
- Dr. Peter Ungeheuer, The European Federation for Cosmetic Ingredients (Frankfurt, Germany)
- Dr. Eric Debruyne, Bayer CropScience (France)

[This Page Intentionally Left Blank]

Subject: FR Notice Comments - 72FR52130: LLNA Performance Standards
Date: Monday, October 29, 2007 4:31 PM

Dr William S Stokes
Director, NICEATM
National Institute of Environmental Health Sciences
PO Box 12233, MD EC-17
Research Triangle Park, NC 27709

Re: 72 FR 52130; September 12, 2007; National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Draft Performance Standards for the Murine Local Lymph Node Assay: Request for Comments.

Dear Dr. Stokes:

These comments are submitted on behalf of the Alternatives Research and Development Foundation, the American Anti- Vivisection Society, Humane Society Legislative Fund, The Humane Society of the United States, People for the Ethical Treatment of Animals, and the Physicians Committee for Responsible Medicine. The parties to this submission are national animal protection, health, and scientific advocacy organizations with a combined constituency of more than 10 million Americans who share the common goal of promoting reliable and relevant regulatory testing methods and strategies that protect human health and the environment while reducing, and ultimately eliminating, the use of animals.

In January, 2007, (ICCVAM) received a nomination from the U.S. Consumer Product Safety Commission (CPSC) to evaluate the validation status of: (1) The murine local lymph node assay (LLNA) as a stand-alone assay for determining potency (including severity) for the purpose of hazard classification; (2) the "cut-down"˙ or "limit dose" LLNA approach; (3) non-radiolabeled LLNA methods; (4) the use of the LLNA for testing mixtures, aqueous solutions, and metals; and (5) the current applicability domain (i.e., the types of chemicals and substances for which the LLNA has been validated). The development of these performance standards is an initial response to this nomination, and ICCVAM is requesting comment on these performance standards.

Although we fully support the development of performance standards that expedite the validation of new protocols that are similar to previously validated methods, we reiterate our disappointment that ICCVAM/ NICEATM has chosen to apply its limited resources to the lengthy process of developing performance standards for such a narrow scope of applicability. These performance standards apply only to modifications of the "standard LLNA" that involve incorporation of non-radioactive methods of detecting lymphocyte proliferation.

A major aspect of the ICCVAM Authorization Act of 2000 (Public Law 106-545, 42 U.S.C. 285l-3) is the charge to "reduce, refine, and/or replace the use of

animals in testing where feasible." The performance standards described in this FR notice apply to modifications of the standard LLNA that do not affect the number of animals used in this method. The only conceivable reduction could occur if the availability of accepted non-radioactive methods of detection would allow more laboratories to perform the LLNA, and if they then choose the LLNA over the Guinea Pig Maximization test or the Buehler Test. The issue of how this exercise (development of performance standards with this limited applicability) addresses ICCVAM's mandate of reducing, refining or replacing the use of animals is not currently mentioned in the draft document and needs to be adequately explained.

In addition, the draft performance standards require the use of a minimum of 20 reference compounds. The criteria by which the compounds were chosen and the characteristics of the compounds are described; however, there is no justification for the requirement of such a large number of compounds for this particular method modification. The methods to which these performance standards apply will differ from the "standard LLNA" only in the method of detection of lymphocyte proliferation; therefore the element of concern is sensitivity of the detection method. All other aspects of the methods to be evaluated will be identical to the standard LLNA, including delivery and biological response. It is therefore not necessary to test representatives for every chemical class or every solvent that has been tested in the standard LLNA. The important characteristic of the reference compound is the magnitude of proliferation response that is generated, and the list of reference compounds chosen should be limited to those that represent the range of response seen with the standard LLNA.

Finally, it is the belief of the parties to this submission that the limited resources available to ICCVAM/NICEATM would be better spent on activities that would have greater impact on the reduction, refinement or replacement of animal use, such as evaluating the use of human cell lines or one of the available in vitro skin models as a replacement for the LLNA.

Sincerely,

Catherine Willett, PhD
Science Policy Advisor
Regulatory Testing Division
People for the Ethical Treatment of Animals

Sara Amundson
Executive Director
Humane Society Legislative Fund

Dr. Martin Stephens
Vice President for Animal Research Issues
Humane Society of the United States

Kristie Stoick, MPH
Research Analyst
Physicians Committee for Responsible Medicine

Sue A. Leary
President
Alternatives Research & Development Foundation

Tracie Letterman, Esq.
Executive Director
American Anti-Vivisection Society

[This Page Intentionally Left Blank]

The Procter & Gamble Company
Central Product Safety
Miami Valley Laboratories
P.O. Box 538707
Cincinnati, Ohio 45253-8707
www.pg.com

February 22, 2008

William S. Stokes, D.V.M., DACLAM
RADM, U.S. Public Health Service
Director, National Toxicology Program Interagency Center for the
Evaluation of Alternative Toxicological Methods
Executive Director, Interagency Coordinating Committee on
the Validation of Alternative Toxicological Methods
National Institute of Environmental Health Sciences, NIH, DHHS
P.O. Box 12233
Research Triangle Park, NC 27709

Dear Dr. Stokes

Thank you for the opportunity to review and comment on the documents prepared by ICCVAM and NICEATM related to a number of the modifications/proposed uses for the traditional LLNA that will be considered by an independent international expert panel in early March.

The teams have done a great job summarizing the available data on the LLNA and for the most part we are in agreement with the conclusions and recommendations outlined in the documents. What makes the LLNA such a valuable tool for skin sensitization hazard identification and risk assessment is that the strengths and limitations of the assay are recognized so well. I am not sure there is another toxicological test that is more understood and evaluated than the LLNA. I am certain that most experts in the field of skin allergy would agree that the older guinea pig skin sensitization test methods are considerably less understood, specifically related to their lack of evaluation through a formal validation process. Our hope is that this peer review of the LLNA will lead to a better appreciation of the LLNA and more important help researchers develop non-animal test methods for evaluating potential skin sensitizing chemicals by using the robust and quantitative natureof the LLNA as a foundation to compare new alternative methods.

For your review and consideration our LLNA experts (Cindy Ryan, Pierre Aeby, Petra Kern and myself) have prepared comments on the LLNA documents posted on the website. I hope you will find them useful and please let us know if you need any additional information.

Sincerely,

G. Frank Gerberick, Ph.D.
Research Fellow Victor Mills Society

DRAFT ICCVAM Recommendations: LLNA Potency

Comparison of LLNAEC3 values to human data:

An evaluation of the ability of the LLNA to predict the relative sensitization potency of chemicals in humans necessitates the use of human sensitization data for comparative purposes. In order for such a comparison to provide meaningful information, one must be aware of and understand the limitations in each of the datasets. The human data used in the comparison are derived from either HRIPT or HMT studies in which single test concentrations, expressed as $\mu g/cm^2$, were used for the induction phase of the test protocol. Therefore, a test concentration could be defined as the NOEL, when in reality it may just be the highest concentration tested to date which did not induce sensitization and there is a probability that higher levels would also fail to induce. This certainly could be the case if a LOEL for the particular chemical has not been identified. Indeed, it is difficult to compare LLNA EC3 concentrations against a human NOEL or an arbitrary value of the LOEL/10 (which is intended to represent an estimation of a probable induction threshold value). On one side, the LLNA data were generated using a test protocol designed to produce quantitative values with dose response information which permit the calculation of the LLNA EC3 and on the other side, the human data were generated by a variety of different human repeated insult patch test and human maximization test protocols which, by design are more qualitative in nature, and unless a series of studies were conducted, provide limited if any information on an induction dose response.

It is concerning that in the evaluation of the LLNA to predict skin sensitization potency in humans key values for the comparison are "pragmatically determined", as is indicated in lines 335-337 of the background review document "Next, the optimal EC3 value that maximized obtaining the correct skin sensitization calls for strong and weak sensitizers (using one or the other proposed decision criterion) was pragmatically determined." Similar wording is used in lines 801-804. The method or rationale for this "pragmatic determination" are not clearly evident in the document. A sound statistical approach should have been used instead and would have provided a more scientifically robust comparison.

Comparison of LLNA EC3 values to guinea pig data:

To assess the ability of the LLNA to predict skin sensitization potency in Guinea Pigs is not relevant to the purpose of this review. Guinea pig tests such as the Buehler (BT) and Guinea Pig Maximization tests (GPMT) were designed for the purpose of hazard identification and are poorly suited for potency estimations. While the ECETOC Technical Report No. 87, Contact Sensitisation: Classification According to Potency proposes methods to categorize allergenic potency based on BT and GPMT data, it demands that the study was conducted in full accord with OECD TG 406 and advises judicious interpretation of the data as does a similar European Union commission expert review. While the BT and GPMT have served the toxicology community well for many years as predictive skin sensitization hazard methods, it is important to recognize that, unlike the LLNA, neither of these tests has been formally validated by a recognized organization nor has the inter-laboratory variability been adequately investigated.

In several sections of the background review document, for examples Lines 321-324 and lines 714-717, it is indicated that for each substance with comparative LLNA and guinea pig data, potency was evaluated by comparing the LLNA EC3 concentration against the percentage of responding guinea pigs in the BT or GPMT and the associated induction concentration. Comparing LLNA EC3 concentration against the percentage of responding guinea pigs is not appropriate in our opinion and resulting data are of very different natures; the LLNA measures events associated with the induction of skin sensitization and provides objective, quantitative dose response information whereas data derived from the guinea pig tests are based on a subjective evaluation of skin responses occurring at the elicitation phase of sensitization and provides no dose response information on the induction phase.

It appears that the authors understand the difficulty of comparing LLNA EC3 values with potency classifications based on guinea pig data. In line 395 of the background review document it states that "…for substances that had more than one EC3 or guinea pig response, the geometric mean EC3 value and the weight of evidence GP classification category was used. Although the data generated by the GPMT and the BT is categorical, using the weight of evidence categorization provided some measure of a mean response across multiple studies." Considering the admitted difficulties encountered in dealing with multiple sets of guinea pig-derived data, the authors should be consistent and not make any conclusion based on such comparison.

Proposed classification categories for sensitization:

While cut-off values for potency classification are proposed based on either Buehler test and GPMT responses (Table 1-1) we would caution the use of such data in the absence of any other supporting data due to the nature of the test design. In addition, the proposed scheme uses the intradermal induction dose of the GPMT along with the % responders as the basis for classification. We believe that the topical induction concentration should be considered as it is the more relevant route of exposure and the concentration used for intradermal injection is often limited by the addition of Freund's Complete Adjuvant.

The proposed classification (as well as the one proposed by ECETOC TR No. 87) considers only data from guinea pig tests which are defined as 'positive' by the accepted TG 406 definition of a sensitizing chemical (i.e. induces 30% or 15% positive responses in the GPMT or BT respectively). It is possible that a weakly sensitizing chemical tested in a guinea pig test could elicit positive responses in 20% or 25% of the test animals in a GPMT or 10% in the BT, and would be considered as a non-sensitizer and thus would not be classified according to the proposed scheme while a chemical with any LLNA EC3 value would be assigned to one of the 2 proposed categories. Data obtained through the LLNA allows for a continuous spectrum of EC3 values and thus provides a rank ordering of relative potencies which offer more opportunities for categorization beyond two categories. And on the other side, Human and GP tests which are designed to provide yes/no answers have various threshold values creatively proposed in order to force results in the same two categories.

In the proposed two level classification scheme for sensitization potency (Table 1-1), the criteria for classification for category 1 are given as "A high frequency of occurrence…." OR "A probability of occurrence of a high sensitization rate in humans…" and for category 2 are given

as "A low or moderate frequency …." OR "A probability of occurrence of a low to moderate sensitization rate in humans…". The frequency of sensitization or the sensitization rate within an exposed population concerns the **prevalence** of allergic contact sensitization to a particular chemical, which is entirely different from the inherent **potency** of the chemical. Therefore the use of such criteria to classify potency is not appropriate. The likelihood of a chemical inducing skin sensitization within an exposed population (i.e. the probable sensitization rate) depends on two key elements: the intrinsic allergenic potency of the chemical AND the conditions and extent of the allergen exposure (e.g. frequency, duration, exposure conditions, etc.). Clinically, the nature, extent and duration of exposure are commonly the predominant determinants of prevalence. The relative potency of a chemical concerns the amount of chemical required to induce sensitization. In general, the more potent the allergen, the lower the dose per unit area required to induce sensitization. Prevalence data are derived from diagnostic patch testing of patients with suspect allergic contact dermatitis, often presenting with clinical disease, in dermatology clinics. The diagnostic patch test itself is designed to detect the weakest degrees of allergy by using occluded exposure conditions for 48 hours and highest allergen concentrations possible to elicit a reaction. For example, the standard patch test concentration for nickel sulfate is 2.5%. Applied in a diagnostic patch test using an 8 mm Finn chamber delivers a dose per unit area of 750 $\mu g/cm^2$, well above the identified human induction threshold of 154 $\mu g/cm^2$ (see Table 2 of Appendix A of the LLNA potency background review documents). Many times the nature of the exposure conditions leading to the induction of allergy for these patients is not clearly defined. At best the published results of thousands of such diagnostic patch tests can be used to evaluate trends in patch test reactions.

One example often used to illustrate the difference between potency and prevalence is nickel. It is a very common contact allergen with a relatively high sensitization rate in the US and Europe. However, experimental evidence indicates that nickel is a relatively weak contact allergen, with LLNA EC3 of 140 $\mu g/cm^2$ and a human induction threshold of 154 $\mu g/cm^2$ for nickel sulfate. The high prevalence is due to the wide distribution, frequent exposure and the nature of exposure, often through 'compromised' skin such as body piercing.

Conversely, the preservative methylchloroisothiazolinone/methylisothiazolinone (MCI/MI) is a well known contact allergen considered to be of strong to extreme potency with LLNA EC3 of 2.25 $\mu g/cm^2$ and a human induction NOEL of 1.25 $\mu g/cm^2$. In Europe, the prevalence rate of allergy to MCI/MI is stable at 1-3% of patch-tested patients. Considering the number of MCI/MI-containing cosmetics and toiletries that are on the market, the opportunities for exposure and the allergenic potency of the preservative one would expect a much higher incidence rate. The prevalence rate for this potent allergen is kept low because of regulatory guidelines/limits on the level of MCI/MI permissible in certain products, thus limiting the dose per unit area of the exposure. Thus, the clinical prevalence of the strong allergen MCI/MI is low whereas for nickel, a known weak allergen, the prevalence is considerably higher which is opposite of what would be expected if only looking at potency and not considering exposure.

The proposed two level classification scheme for sensitization potency (Table 1-1) does not accurately reflect the range of allergenic potencies that have been demonstrated by both animal and human data. LLNA EC3 values and human induction thresholds clearly span several orders of magnitude as shown by the data in Table 2 of Appendix A of the LLNA potency background

review documents. Human threshold values range from 1.25 μg/cm^2 for MCI/MI, to 250 μg/cm^2 for isoeugenol, to 2755 μg/cm^2 for farnesol, to 20,690 μg/cm^2 for benzyl benzoate. Clinical experience with allergic contact dermatitis would also indicate that discrete classes of sensitizing potency exist (Contact Derm, 2000, 42:344-348).

DRAFT ICCVAM Recommendations: LLNA Applicability Domain

Draft Recommendations – Use of the LLNA to Test Mixtures:

A dataset of 18 mixtures was evaluated, 15 of which had guinea pig data and none had human data. As a result, the LLNA data were compared to the guinea pig data. Since the database is severely limited due to the lack of human data, there is no proof that the guinea pig data would be representative of the human response. Thus, using the guinea pig data as the standard to which the LLNA data should be compared is not appropriate.

In addition, the usefulness of these data is limited further by the fact that information on the ingredients is known for only one of the 15 mixtures and 11 were tested in the LLNA in an aqueous vehicle, the performance of which is also being assessed in this same report.

High quality LLNA mixture data is published in Lalko et al. (2006), cited in section 7.6 of Addendum No. 1 to the ICCVAM report. This publication concerns the evaluation of essential oils and includes analytical data on the composition of the oils as well as LLNA data on the identified major constituents. These data should have been included in the evaluation and not just mentioned as other available scientific reports.

Since the database is severely limited due to the lack of human data, we agree with the recommendation that an assessment of the suitability of the LLNA for testing mixtures should not be conducted until a sufficient quantity of quality data become available. A similar logic of course also applies to guinea pig test methods.

Draft Recommendations – Use of the LLNA to Test Metal Compounds:

The reference dataset contains human data for 17 metal compounds representing 13 different metals. Since the allergenic potential in humans of most all of the known metals has been established, one questions the importance of or need for an assessment of the LLNA's ability to detect metal allergens. However, we agree with the recommendation that the LLNA is useful for the testing of metal compounds. Whether or not the LLNA is useful for testing nickel compounds is of limited importance as nickel is a well known human contact allergen.

In addition, since only 1 of the 14 metal compounds with LLNA and human data was tested in both in an aqueous vehicle, the comparison does not add much value to the assessment, especially in light of the fact that the performance of the LLNA using aqueous vehicles is being assessed in this same report.

Draft Recommendations – Use of the LLNA to Test Substances in Aqueous Solutions:

A dataset of 21 substances tested in aqueous solutions was evaluated, 4 of which had had human data. Since the database is severely limited due to the lack of human data, we agree with the recommendation that an assessment of the suitability of the LLNA for testing substance in

aqueous solutions should not be conducted until a sufficient quantity of quality data become available.

DRAFT ICCVAM Recommendations: LLNA Limit Dose Procedure

Draft Recommendations – Limit Dose Procedure:

We agree with the recommendation that the LLNA limit dose procedure is appropriate for hazard identification purposes.

We must point out that a 10% concentration threshold for defining non-sensitizing chemicals is not, as suggested in line 44 of the recommendation, proposed by Kimber et al. (2006) as the absolute cut-off. In the discussion section of that same paper, Kimber et al. indicate that for the purposes of that article the 10% threshold was used and that that figure "should not be regarded as inviolable." They go on to say that a case could be made for using, for instance, either 15% or 20%. In the 2005 Gerberick et al. paper (Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. Dermatitis, 16(4):157-202), compounds that did not induce a positive response at any concentration tested, with the highest concentration being at least 20% or greater, were categorized as non-sensitizing.

In addition, the 10% threshold concentration at which all which all negative results would be considered valid did not originate in the cited Kimber et al 2006 publication. The original reference is Cockshott et al., 2006, Human and Experimental Toxicology, 25:387-394 in which the performance of the LLNA was evaluated in a regulatory context. In that paper, a negative result obtained with the highest concentration tested at 10% would be considered a valid result if the positive control, a mild to moderate sensitizer, gave a positive response. In other words, a chemical which is negative at a top concentration of 10% does not represent a significant human sensitization hazard. This is similar to the definition of a non-sensitizing chemical in the Guinea Pig Maximization Test (GPMT) or Buehler test as one which induces less than 30% or 15% positive responses respectively. Therefore, if a chemical elicits positive responses in 20% or 25% of the test animals in a GPMT, it would be considered as a non-sensitizer from a regulatory perspective.

Comments on DRAFT ICCVAM Recommendations: LLNA Non-Radioactive Methods

DRAFT ICCVAM Recommendations: LLNA BrdU ELISA Procedure

We agree with the recommendation that more information and data are needed on this method in order to conduct a meaningful assessment of the BrdU ELISA procedure's performance relative to the traditional LLNA. It is especially important to have information regarding the inter-laboratory performance of this assay.

We do have one suggestion for consideration. Table 6-2 of the Background Review Documents shows a comparison of standard LLNA EC3 values and 0.5x-2x range for the performance standard chemicals and EC3 values calculated from the BrdU ELISA LLNA. Since an alternative SI cutoff for the BrdU ELISA LLNA was identified that provides greater accuracy

than an SI = 3 cutoff i.e., SI = 1.3, a comparison of BrdU ELISA EC1.3 values to standard LLNA EC3 values would be helpful.

DRAFT ICCVAM Recommendations: LLNA BrdU FC Procedure

We agree with the recommendation that more information and data are needed on this method in order to conduct a meaningful assessment of the BrdU-FC procedure's performance relative to the traditional LLNA. While the total number of chemicals tested (45) is sufficient, it is especially important to have information regarding the inter-laboratory performance of this assay. The background review document speculates that the transferability of the LLNA: BrdU-FC and the eLLNA: BrdU-FC would be similar to the traditional LLNA. However, we do not think that will be the case. Flow cytometry is not a trivial technique. It is certainly more error prone than scintillation counting and often the quality of the results is very dependant on trained personnel and precise procedures.

Only 13 of the 18 minimum performance standard reference chemicals have been tested in the LLNA BrdU-FC procedure. This may not be sufficient to assess the test performance according to the ICCVAM Performance Standards for the LLNA. In addition, rather than focusing on the number of chemicals for which the BrdU-FC procedure produced equivocal results or did not obtain 100% concordance with the ICCVAN LLNA performance standard reference chemicals, we believe that it would be of greater value to investigate potential causes for those results. Such information would provide some understanding of the limitations of the methods.

Since the purpose of this evaluation of the LLNA BrdU-FC procedure is to assess its ability to be a non-radioactive alternative to the traditional LLNA, is a comparison with Guinea Pig data justified?

The provided test protocol indicates that at least 6 mice be employed for an irritation prescreen and a possible 12 more be used for the optional quantitative irritation test. Therefore, this method has the potential to use more mice than the traditional LLNA. This requirement for greater animal usage must be taken into consideration when evaluating the BrdU-FC Procedure and it must be determined that the quality or quantity of information provided by this method exceeds that which would be obtained with the traditional LLNA. In other words, are the additional mice required by the BrdU-FC worth any possible additional information that would be gained compared to conducting a traditional LLNA?

DRAFT ICCVAM Recommendations: LLNA DA Procedure

Beyond the method to assess lymph node cell proliferation, the test protocol for the LLNA DA contains several key deviations from the OECD Test Guideline 429 recommended protocol and the Essential Test Method Components as described in the Draft ICCVAM Performance Standards for the LLNA . As indicated in the recommendation document (lines 77-79), the LLNA DA has made major modification to the traditional LLNA in both the test substance treatment and sampling schedule. Therefore, this method is outside of the requirements of the draft ICCVAM Performance Standards for the LLNA and should not be consider for validation as an LLNA alternative at this time.

[This Page Intentionally Left Blank]

February 22, 2008

Dr William S Stokes
Director, NICEATM
National Institute of Environmental Health Sciences
PO Box 12233, MD EC-17
Research Triangle Park, NC 27709

Re: **73 FR 25553; January 8, 2008; National Toxicology Program (NTP); NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Announcement of an Independent Scientific Peer Review Panel Meeting on the Murine Local Lymph Node Assay; Availability of Draft Background Review Documents; Request for Comments**

Dear Dr. Stokes:

These comments are submitted on behalf of People for the Ethical Treatment of Animals and the Physicians Committee for Responsible Medicine. The parties to this submission are national animal protection, health, and scientific advocacy organizations with a combined constituency of more than two million Americans who share the common goal of promoting reliable and relevant regulatory testing methods and strategies that protect human health and the environment while reducing, and ultimately eliminating, the use of animals.

Please take note of the following thoughts and transmit them to the Peer Review Panel (PRP) accordingly.

In January, 2007, (ICCVAM) received a nomination from the U.S. Consumer Product Safety Commission (CPSC) to evaluate the validation status of: (1) The murine local lymph node assay (LLNA) as a stand-alone assay for determining potency (including severity) for the purpose of hazard classification; (2) the ''cut-down'' or ''limit dose'' LLNA approach; (3) non-radiolabeled LLNA methods; (4) the use of the LLNA for testing mixtures, aqueous solutions, and metals; and (5) the current applicability domain (i.e., the types of chemicals and substances for which the LLNA has been validated).

Now more than a year later, ICCVAM is preparing for a peer review meeting to evaluate its recommendations and findings on these four items. It is unclear when final recommendations will be transmitted to federal agencies, but if ICCVAM's review of *in vitro* pyrogenicity methods is any indication, it may be at least another year.

Since this review of the LLNA and the proposed recommendations contained therein will lead to little reduction or refinement of animal use in sensitization, the resources that ICCVAM devote to this exercise should be kept to a minimum, and any forthcoming recommendations should be transmitted to agencies immediately following the Peer Review.

We have divided our comments into sections following the FR Notice:

LLNA limit dose procedures (the reduced or rLLNA) —draft Background Review Document (BRD) and other related documents

In April, 2007, ESAC issued a statement supporting the use of the rLLNA "within tiered-testing strategies to reliably distinguish between chemicals that are skin sensitizers and non-sensitizers "thereby reducing animal use by as much as 50%."[1]

In spite of the ESAC recommendations, ICCVAM has conducted its own data call in and data review. The reviewed database is comprehensive and contains a broad cross-section of the chemical universe. The performance characteristics were all above 95% (false negative and positive rates are very low or zero). Even though this additional review was largely unnecessary, we are pleased that ICCVAM's draft recommendations concluded favorably for the rLLNA procedure and urge the Peer Review Panel to concur. ICCVAM should forward recommendations regarding the use of the rLLNA to federal agencies immediately following the Peer Review.

Mixtures, metals, and aqueous solutions—draft Updated Assessment of the Validity of the LLNA for Mixtures, Metals, and Aqueous Solutions and related documents

ICCVAM has evaluated available data with respect to the use of LLNA in predicting the skin sensitization potential of mixtures, metals, and aqueous solutions. In all cases, the limited availability of data prevented a conclusive recommendation for the use of the LLNA; for metals, the LLNA is recommended only as part of a weight-of-evidence approach, which does not significantly promote a reduction in the use of animals.

Clearly this approach to expanding the applicability domain of the LLNA has not proved terribly fruitful, and we do not endorse further validation efforts in this regard, but recommend all resources are directed towards the pursuit of *in vitro* methods for this purpose.

Potency—draft BRD and related documents

Once again, ICCVAM has reviewed all availed data and come to a conclusion that is in opposition to that of other experts in the field. For more than 10 years data has been accumulating indicating the potential for the LLNA to make a determination of the sensitization potency of a chemical.[2] Several publications by Basketter and others (many of which are referenced in the BRD) as well as the eloquent argument by Basketter et al. presented in Appendix A, conclude that LLNA is appropriate for determining potency. In September 2000, the European Centre for Ecotoxicology and Toxicity of Chemicals (ECETOC) published a comprehensive review of sensitization test methods with respect to hazard identification and labeling, to determine whether the various methods are appropriate for determining relative potency and risk assessment.[3] The conclusions from this review included: (1) the LLNA is a viable and complete alternative to traditional guinea pig test

[1] http://ecvam.jrc.it/publication/ESAC26_statement_rLLNA_20070525-1.pdf

[2] Kimber I, Basketter D A. Contact sensitization: A new approach to risk assessment. Human and Ecological Risk Assessment 1997: 3: 385 - 395.

[3] ECETOC. 2000. Skin Sensitization Testing for the Purpose of Hazard Identification and Risk Assessment.

methods for the purposes of skin sensitization hazard identification, and (2) the LLNA is suitable for the determination of relative skin sensitizing potency and the adaptation of this method for derivation of comparative criteria such as EC3 values provides an effective and quantitative basis for such measurements. This report further recommends that "the LLNA is the recommended method for new assessments of relative potency and/or for the investigation of the influence of vehicle or formulation on skin sensitizing potency."

More recent work has further verified the use of the LLNA as a stand-alone method for estimating potency for regulatory purposes, including a 2005 study that concludes that there is a "clear linear relationship between LLNA-derived EC3 values and historical human skin patch data."[4] A 2007 review concludes that "The LLNA, when conducted according to published guidelines, provides a robust method for skin sensitization testing that not only provides reliable hazard identification in formation but also data necessary for effective risk assessment and risk management." In addition, a retrospective analysis of the regulatory use of the LLNA in the EU was published in 2006 and concluded that "the LLNA is satisfactory for routine regulatory use."[5]

Despite all of this, ICCVAM's review of the LLNA for potency determination does not support such a finding, even though, according to the BRD, the LLNA was better overall at predicting sensitization potency than guinea pig data. It is clear from the BRD that different data treatments result in different R^2 values, and the BRD should more clearly discuss the reasons those analysis decisions were made. Further, the BRD should explain in detail why conclusions were drawn that are opposite to that of the evidence they reference.

We urge the PRP to take into account the submission in Appendix A of the draft LLNA-potency BRD, which details why the LLNA is a scientifically appropriate method of potency determination, and the subsequent submitted comment by Dr. David Basketter, a recognized expert in the field of skin sensitization, when making its final report to ICCVAM.

Non-radioactive methods—draft BRDs and related documents

Three new methods of measuring lymphocyte proliferation have been proposed. Unlike the traditional LLNA, these new methods do not use a radioactive indicator, which could increase the use of the LLNA in facilities that cannot use radioactive material. The new methods include two variants of a bromodioxyuridine system [BrdU: ELISA and BrdU: Flow Cytometry (FC)] and the LLNA: DA.

When compared to human data, the LLNA: BrdU-FC had a higher accuracy rate, higher sensitivity, the same specificity, the same false positive rate, and a lower false negative rate than the traditional LLNA. Despite this performance, the assay does not achieve complete concordance with the proposed LLNA Performance Standards the PRP will be evaluating. This is also the case with for the LLNA-DA method, which compares identically to human data, yet

[4] Basketter et al. Predictive identification of human skin sensitization thresholds. Contact Dermatitis. 2005; 53 (5): 260-267.
[5] Cockshott et al., The local lymph node assay in practice: a current regulatory perspective. Hum Exp Toxicol 2006; 25 (7): 387-394.

falls short when compared to the traditional LLNA. While reasons for this are not clear, it is worth an examination of whether we should compare new methods to the methods they are replacing or to the endpoint of actual interest.

The BrdU: ELISA has been recommended for use by ICCVAM pending receipt of additional information and using alternative decision criteria. We support this finding. Because of the incomplete concordance between these methods and the traditional LLNA, ICCVAM qualified their acceptance and recommends a "weight-of-evidence" approach. While it is usually good scientific practice to evaluate any test method results in weight-of-evidence manner, qualifications such as these undercut the recommendations and introduce undue confusion to the reader. In our view, this gives a company a clear incentive to conduct more testing, when in reality the methods evaluated have acceptable performance and should simply be recommended.

Performance Characteristics

Although we fully support the development of performance standards that expedite the validation of new protocols that are similar to previously validated methods, we reiterate our disappointment that ICCVAM/ NICETAM has chosen to apply its limited resources to the lengthy process of developing performance standards for such a narrow scope of applicability. These performance standards apply only to modifications of the "standard LLNA" that involve incorporation of non-radioactive methods of detecting lymphocyte proliferation.

In addition, the draft performance standards require the use of a minimum of 22 reference compounds. The criteria by which the compounds were chosen and the characteristics of the compounds are described; however, there is no justification for the requirement of such a large number of compounds for this particular method modification. The methods to which these performance standards apply will differ from the "standard LLNA" only in the method of detection of lymphocyte proliferation; therefore the element of concern is sensitivity of the detection method. All other aspects of the methods to be evaluated will be identical to the standard LLNA, including delivery and biological response. It is therefore not necessary to test representatives for every chemical class or every solvent that has been tested in the standard LLNA. The important characteristic of the reference compound is the magnitude of proliferation response that is generated, and the list of reference compounds chosen should be limited to those that represent the range of response seen with the standard LLNA.

In addition, a major criterion for the selection of the above compounds is that there are Guinea pig data available; more appropriately, chemicals should be chosen on the basis of available human data.

Conclusions and Future directions

This exercise is a good example of actions undertaken by ICCVAM which result in frustration in the animal protection community. In the future we hope that ICCVAM will take a more holistic approach to determine the ways in which it spends its limited time and resources so as to ensure maximum benefit for animals in laboratories.

Several non-animal methods for estimating sensitivity are under development, including quantitative structure activity relationship (QSAR) modeling that shows a high concordance with guinea pig and LLNA data,[6] quantification of peptide reactivity, which also shows a high concordance with LLNA data,[7,8] and human cell cultures.[9,10] We urge ICCVAM to secure an interagency grant from the CPSC to fund the validation of one or more of these non-animal methods. Clearly, ICCVAM and the CPSC both benefit from the sharing of resources, as the CPSC nominated the method and ICCVAM will be tasked with the final work product.

ICCVAM should consider taking a more pro-active approach similar to the European Sens-it-iv project,[11] which involves the coordinated efforts of more than two dozen groups from industry, academia and other organizations, all working toward the common goal of developing *in vitro* methods to assess immunotoxicity.

Sincerely,

/s/

Catherine Willett, PhD
Science Policy Advisor
Regulatory Testing Division
People for the Ethical Treatment of Animals

/s/

Kristie Stoick, MPH
Scientific and Policy Advisor
Physicians Committee for Responsible Medicine

[6] Fedorowicz et al. Structure-activity models for contact sensitization. Chem Res Toxicol. 2005; 18(6): 954-969.
[7] Gerberick et al. Quantification of chemical peptide reactivity for screening contact allergens: A classification tree model approach. Toxicol. Sci. 2007; 97(2): 417-427.
[8] Natsch and Emter. Skin sensitizers induce antioxidant response element dependent genes: Application to the in vitro testing of the sensitization potential of chemicals. Tox Sci. 2008; 102(1): 110-119.
[9] Sakaguchi, et al., Development of an in vitro skin sensitization test using human cell lines; huna Cell Line Activation Test (h-CLAT) II. An inter-laboratory study of the h-CLAT. Toxicol. In vitro. 2005; 20 (5): 774-784.
[10] Schoeters et al. Microarray analyses in dendritic cells reveal potential biomarkers for chemical-induced shin sensitization. Mol. Immunol. 2007; 44(12): 3222-3233.
[11] http://www.sens-it-iv.eu/

[This Page Intentionally Left Blank]

Appendix G3

Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) Comments SACATM Meeting on June 18–19, 2008

The following is excerpted from the final minutes and speaker presentations of the SACATM meeting convened on July 18-19, 2008. The full meeting minutes are available online at http://ntp-server.niehs.nih.gov/?objectid=AF6CC417-F1F6-975E-75B5F3FF7DF1CDDC.

[This Page Intentionally Left Blank]

Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay

Dr. Marilyn Wind presented the "Report on the Independent Scientific Peer Review Meeting: Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay (LLNA), a Test Method for Assessing the Contact Dermatitis Potential of Chemicals and Products – Introduction and Overview," authored by Dr. Joanna Matheson.

In 2007, the timeline for the ICCVAM evaluations included the nomination from the CPSC, endorsement by ICCVAM, and SACTAM's endorsement of the recommendation. In 2008, the LLNA Peer Review Panel met and a report was made available. The new/updated LLNA applications and protocols reviewed by the Peer Review Panel included:

- LLNA limit dose procedure
- LLNA for testing mixtures, metals, and aqueous solutions
- Non-radioactive LLNA: DA method
- Non-radioactive LLNA: BrdU-FC method
- Non-radioactive LLNA: BrdU-ELISA method
- Draft ICCVAM LLNA performance standards
- LLNA for potency determinations.

The documents prepared by the NICEATM and ICCVAM Immunotoxicity Working Group (IWG) for each new/updated LLNA applications included the draft background review document (BRD), the draft ICCVAM test method recommendations (usefulness and limitations, recommended protocol, future studies), and questions for the Peer Review Panel.

Dr. Wind presented an overview of the LLNA test method protocol.

- The LLNA protocol was initially described by Kimber et al. (1986).
- The purpose of the LLNA is to identify chemical sensitizers through quantification of lymphocyte proliferation.
- The LLNA uses a minimum of three dose levels. The highest dose level should be the maximum soluble concentration that does not cause systemic toxicity or excessive local irritation.
  - A stimulation index (SI) is calculated as the ratio of radioactivity incorporated into the cells of auricular lymph nodes of the treated animals to that of the vehicle control animals. The threshold for classifying a substance as a skin sensitizer is an $SI \geq 3$.
  - In order for an LLNA study to be considered acceptable, the concurrent positive control must yield an $SI \geq 3$.

A condensed version of the LLNA test method protocol was provided.

- Test substance is applied to mouse ears on Days 1, 2, and 3.
- On Day 6, mice are injected with radiolabeled thymidine (or an analogue of thymidine).
- Radiolabeled thymidine is incorporated into the DNA of proliferating cells and the auricular lymph nodes are removed.
- The amount of radiolabeled thymidine in the lymph nodes is determined as a measure of lymphocyte proliferation. The ratio of incorporated radioactivity in the auricular lymph nodes of treated vs. control mice (i.e., SI) is calculated, which leads to classification of a compound as negative or as a sensitizer.

The sole difference between the LLNA limit dose (rLLNA) test method protocol and that of the traditional LLNA protocol is that only a single dose, the highest dose that does not induce systemic toxicity or excessive local irritation, is used.

The LLNA Limit Dose Test Method Database contains information that is included in the BRD, which is based on a retrospective review of traditional LLNA data that were either submitted as part of the original LLNA evaluation (ICCVAM 1999), extracted from peer-reviewed publications, or submitted to NICEATM in response to an *FR* notice requesting available data and information. Data from 471 studies representing 466 unique substances are available (211 substances were included in the 1998 ICCVAM evaluation of the traditional LLNA). Results with the LLNA limit dose test procedure almost always agree with results from the traditional LLNA. Kimber et al. (2006) showed a 98.6% accuracy for 211 substances, and ICCVAM (2008) showed a 98.9% accuracy for 466 substances.

Dr. Wind provided the draft ICCVAM recommendations for the LLNA Limit Dose Test Method.

- The LLNA limit dose procedure should be used for the hazard identification of skin sensitizing substances if dose response information is not needed. Use all other LLNA protocol specifications recommended by ICCVAM (ICCVAM 1999, Dean et al. 2001).
- Users should be aware that the limit dose is the highest soluble concentration that does not induce overt systemic toxicity and/or excessive local irritation. A small possibility of a false negative result exists (1.6% [5/313]) when compared to the traditional LLNA.

Overview of the Peer Review Panel Report

Dr. Michael Luster presented the "Overview of the LLNA Independent Scientific Peer Review Panel Report," starting with the charge to the Panel to review the draft BRDs and to evaluate the extent to which applicable validation and acceptance criteria of toxicological test methods have been appropriately addressed. Further the Panel was to consider the ICCVAM draft test method recommendations for proposed method uses and limitations, recommended

standardized protocols, test method performance standards, and proposed future studies. The Panel was to then comment on the extent to which these items are supported by the information provided in the BRD. The Panel evaluated the LLNA modifications and applications as listed above.

Dr. Luster provided highlights of the final Independent Scientific Peer Review Panel report and recommended that the report should be consulted for a detailed description of the Panel's conclusions and recommendations. He provided evaluation highlights for all LLNA modifications and applications. The following points were made about the LLNA Limit Dose Procedure (rLLNA).

- The procedure follows the traditional ICCVAM LLNA protocol except for the number of doses tested. This procedure uses only the high dose group (requires 40% fewer animals).
- In general, the Panel concurred with the draft ICCVAM recommendation that the LLNA limit dose procedure should be routinely recommended for the hazard identification of skin sensitizing chemicals when dose response information is not required. The Panel also recommended that the limit dose procedure can be used as an initial test when dose-response information is required.
- The Panel also recommended that if dose-response information is required, as a way to further reduce animal use, the LLNA limit dose procedure should be routinely recommended as the initial test to identify sensitizers before conducting the traditional LLNA since negative results would not require further testing.
- Additionally, the Panel suggested that the test be referred to as the "reduced LLNA" (rLLNA) to be consistent with ECVAM terminology.

SACATM Discussion

There were no public comments specific to the rLLNA.

Dr. Grantley Charles concurred with the recommendation that the rLLNA protocol should include a discussion of how to determine the maximum dose if only a single dose is to be used in a screen process. An investigator must be able to define excessive irritation; otherwise, the testing may produce a bell-shaped response curve. The updated ICCVAM LLNA protocol has been revised to add specific guidance on determining the maximum concentration to be tested to avoid overt systemic toxicity and excessive local irritation.

Dr. Marion Ehrich suggested that the rLLNA appears favorable because 153/153 nonsensitizing agents and 308/318 sensitizing agents were predicted.

Dr. Luster said that the Panel made a very strong suggestion at the panel meeting that there be some histology associated with the rLLNA procedure. Histology was part of the plan, and it is embedded in the text.

Dr. Donald Fox stated that use of the reduced LLNA procedure could encourage the alternative use of animals.

[This Page Intentionally Left Blank]

Appendix H

Relevant Skin Sensitization Regulations and Testing Guidelines

[This Page Intentionally Left Blank]

Appendix H1
Relevant Skin Sensitization Test Regulations

United States

Agency/ Center or Office	Regulated Products	Legislation	Statutory Requirements
FDA/CDER	Pharmaceuticals	Federal Food, Drug, and Cosmetic Act (U.S.C. Title 21, Chapter 9) Public Health Service Act (U.S.C. Title 42, Chapter 6A)	21 CFR 312 21 CFR 314
EPA/OPPTS	Pesticides	Toxic Substances Control Act (U.S.C. Title 15, Chapter 53) Federal Insecticide, Fungicide, and Rodenticide Act (U.S.C. Title 7, Chapter 6)	40 CFR 158.50 40 CFR 158.100 40 CFR 158.340
CPSC	Consumer Products	Federal Hazardous Substances Act (U.S.C. Title 15, Chapters 1261-1278)	16 CFR 1500.3
OSHA	Chemicals	Occupational Safety and Health Act of 1970 (U.S.C. Title 29, Chapter 15)	29 CFR 1910.1200

Europe

Agency/ Center or Office	Regulated Products	Legislation	Statutory Requirements
EU	Dangerous Preparations (Chemicals and chemical mixtures)	Directive 1999/45/EC of the European Parliament and of the Councilof 31 May 1999 Annex V to Directive 67/548/EEC of 27 June 1967	
EU	Pesticides	Directive 91/414/EEC of the European Parliament and of the Council of 15 July 1991	

International

Agency/ Center or Office	Regulated Products	Legislation	Statutory Requirements
GHS	Chemicals		GHS Part 3, Chapter 3.4

]This Page Intentionally Left Blank]

Appendix H2

OECD Test Guideline 429: Skin Sensitisation – Local Lymph Node Assay

[This Page Intentionally Left Blank]

OECD/OCDE **429**

Adopted:
24th April 2002

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Skin Sensitisation: Local Lymph Node Assay

INTRODUCTION

1. The OECD Test Guideline Programme periodically reviews progress in test method development and refinement, both in terms of scientific advances and animal welfare, to determine whether existing Test Guidelines should be updated and whether new Guidelines should be developed. Toward that end, a new assay for the determination of skin sensitisation in the mouse, the Local Lymph Node Assay (LLNA) has been sufficiently validated and accepted to justify its adoption as a new Test Guideline (1)(2)(3). This is the second Guideline to be promulgated for assessing skin sensitisation potential of chemicals in animals. The other Guideline (406) utilises guinea pig tests, notably the guinea pig maximisation test and the Buehler test (4).

2. The LLNA provides certain advantages with regard to both scientific progress and animal welfare. It studies the induction phase of skin sensitisation and provides quantitative data suitable for dose response assessment. The details of the validation of the LLNA and a review of the associated work have been published (5)(6)(7)(8). In addition, it should be noted that the mild/moderate sensitisers, which are recommended as suitable positive control substances for guinea pig test methods, are also appropriate for use with the LLNA (6)(8)(9).

INITIAL CONSIDERATIONS

3. The LLNA provides an alternative method for identifying skin sensitising chemicals and for confirming that chemicals lack a significant potential to cause skin sensitisation. This does not necessarily imply that in all instances the LLNA should be used in place of guinea pig tests, but rather that the assay is of equal merit and may be employed as an alternative in which positive and negative results generally no longer require further confirmation.

4. The LLNA is an *in vivo* method and, as a consequence, will not eliminate the use of animals in the assessment of contact sensitising activity. It has, however, the potential to reduce the number of animals required for this purpose. Moreover, the LLNA offers a substantial refinement of the way in which animals are used for contact sensitisation testing. The LLNA is based upon consideration of immunological events stimulated by chemicals during the induction phase of sensitisation. Unlike guinea pig tests the LLNA does not require that challenged-induced dermal hypersensitivity reactions be elicited. Furthermore, the LLNA does not require the use of an adjuvant, as is the case for the guinea pig maximisation test. Thus, the LLNA reduces animal distress. Despite the advantages of the LLNA over traditional guinea pig tests, it should be recognised that there are certain limitations that may necessitate the use of traditional guinea pigs tests (e.g., false negative findings in the LLNA with certain metals, false positive findings with certain skin irritants)(10).

429 OECD/OCDE

PRINCIPLE OF THE TEST

5. The basic principle underlying the LLNA is that sensitisers induce a primary proliferation of lymphocytes in the lymph node draining the site of chemical application. This proliferation is proportional to the dose applied (and to the potency of the allergen) and provides a simple means of obtaining an objective, quantitative measurement of sensitisation. The LLNA assesses this proliferation as a dose-response in which the proliferation in test groups is compared to that in vehicle treated controls. The ratio of the proliferation in treated groups to that in vehicular controls, termed the Stimulation Index, is determined, and must be at least three before a test substance can be further evaluated as a potential skin sensitiser. The methods described here are based on the use of radioactive labelling to measure cell proliferation. However, other endpoints for assessment of proliferation may be employed provided there is justification and appropriate scientific support, including full citations and description of the methodology.

DESCRIPTION OF THE ASSAY

Selection of animal species

6. The mouse is the species of choice for this test. Young adult female mice of CBA/Ca or CBA/J strain, which are nulliparous and non-pregnant, are used. At the start of the study, animals should be between 8-12 weeks old, and the weight variation of the animals should be minimal and not exceed 20% of the mean weight. Other strains and males may be used when sufficient data are generated to demonstrate that significant strain and/or gender-specific differences in the LLNA response do not exist.

HOUSING AND FEEDING CONDITIONS

7. Animals should be individually housed. The temperature of the experimental animal room should be 22ºC (± 3ºC). Although the relative humidity should be at least 30% and preferably not exceed 70% other than during room cleaning, the aim should be 50-60%. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water.

PREPARATION OF ANIMALS

8. The animals are randomly selected, marked to permit individual identification (but not by any form of ear marking), and kept in their cages for at least 5 days prior to the start of dosing to allow for acclimatisation to the laboratory conditions. Prior to the start of treatment all animals are examined to ensure that they have no observable skin lesions.

Reliability check

9. Positive controls are used to demonstrate appropriate performance of the assay and competency of the laboratory to successfully conduct the assay. The positive control should produce a positive LLNA response at an exposure level expected to give an increase in the stimulation index (SI) >3 over the negative control group. The positive control dose should be chosen such that the induction is clear but not excessive. Preferred substances are hexyl cinnamic aldehyde (CAS No 101-86-0) and mercaptobenzothiazole (CAS No 149-30-4). There may be circumstances in which, given adequate justification, other control substances, meeting the above criteria, may be used. While ordinarily a positive control group may be required in each assay, there may be situations in which test laboratories will have available historic positive control data to show consistency of a satisfactory response over a six-month or

more extended period. In those situations, less frequent testing with positive controls may be appropriate at intervals of no greater than 6 months. Although the positive control substance should be tested in the vehicle that is known to elicit a consistent response (e.g., acetone:olive oil), there may be certain regulatory situations in which testing in a non-standard vehicle (clinically/chemically relevant formulation) will also be necessary. In such situations the possible interaction of a positive control with this unconventional vehicle should be tested.

TEST PROCEDURE

Number of animals and dose levels

10. A minimum of four animals is used per dose group, with a minimum of three concentrations of the test substance, plus a negative control group treated only with the vehicle for the test substance, and a positive control, as appropriate. In those cases in which individual animal data are to be collected, a minimum of five animals per dose group are used. Dose and vehicle selection should be based on the recommendations given in reference (2). Doses are selected from the concentration series 100%, 50%, 25%, 10%, 5%, 2.5%, 1%, 0.5% etc. Existing acute toxicity and dermal irritation data should be considered, where available, in selecting the three consecutive concentrations so that the highest concentration maximises exposure whilst avoiding systemic toxicity and excessive local skin irritation (2)(11). Except for absence of treatment with the test substance, animals in the control groups should be handled and treated in a manner identical to that of animals in the treatment groups.

11. The vehicle should be selected on the basis of maximising the test concentrations and solubility whilst producing a solution/suspension suitable for application of the test substance. In order of preference, recommended vehicles are acetone/olive oil (4:1 v/v), dimethylformamide, methyl ethyl ketone, propylene glycol and dimethyl sulphoxide (2)(10), but others may be used if sufficient scientific rationale is provided. In certain situations it may be necessary to use a clinically relevant solvent or the commercial formulation in which the test substance is marketed as an additional control. Particular care should be taken to ensure that hydrophilic materials are incorporated into a vehicle system, which wets the skin and does not immediately run off. Thus, wholly aqueous vehicles are to be avoided.

Experimental schedule

12. The experimental schedule of the assay is as follows:

- *Day 1:*
  Individually identify and record the weight of each animal. Open application of 25µL of the appropriate dilution of the test substance, the vehicle alone, or the positive control (as appropriate), to the dorsum of each ear.
- *Days 2 and 3:*
  Repeat the application procedure carried out on day 1.
- *Days 4 and 5* :
  No treatment.
- *Day 6* :
  Record the weight of each animal. Inject 250µL of phosphate-buffered saline (PBS) containing 20 µCi (7.4e+5 Bq) of ^{3}H-methyl thymidine into all test and control mice via the tail vein. Alternatively inject 250 µL PBS containing 2 µCi (7.4e + 4 Bq) of ^{125}I-iododeoxyuridine and 10^{-5}M fluorodeoxyuridine into all mice via the tail vein. Five hours (5 h) later, the animals are killed. The draining auricular lymph nodes from each ear are excised and pooled in PBS for each experimental group (pooled treatment group approach);

alternatively pairs of lymph nodes from individual animals may be excised and pooled in PBS for each animal (individual animal approach). Details and diagrams of the node identification and dissection can be found in Annex I of the ICCVAM Immunotoxicology Working Group LLNA Protocol (10).

Preparation of cell suspensions

13. A single cell suspension of lymph node cells (LNC) either from pooled treatment groups or bilaterally from individual animals is prepared by gentle mechanical disaggregation through 200 µm-mesh stainless steel gauze. Lymph node cells are washed twice with an excess of PBS and precipitated with 5% trichloroacetic acid (TCA) at 4°C for 18h(2). Pellets are either re-suspended in 1 mL TCA and transferred to scintillation vials containing 1.0 mL of scintillation fluid for ^{3}H-counting, or transferred directly to gamma counting tubes for ^{125}I-counting.

Determination of cellular proliferation (incorporated radioactivity)

14. Incorporation of ^{3}H-methyl thymidine is measured by β-scintillation counting as disintegrations per minute (DPM). Incorporation of ^{125}I-iododeoxyuridine is measured by ^{125}I-counting and also is expressed as DPM. Depending on the approach used, the incorporation will be expressed as DPM/treatment group (pooled approach) or DPM/animal (individual approach).

OBSERVATIONS

Clinical observations

15. Animals should be carefully observed once daily for any clinical signs, either of local irritation at the application site or of systemic toxicity. All observations are systematically recorded with individual records being maintained for each animal.

Body weights

16. As stated in paragraph 12, individual animal body weights should be measured at the start of the test and at the scheduled kill of the animals.

CALCULATION OF RESULTS

17. Results are expressed as the Stimulation Index (SI). When using the pooled approach, the SI is obtained by dividing the pooled radioactive incorporation for each treatment group by the incorporation of the pooled vehicle control group; this yields a mean SI. When using the individual approach, the SI is derived by dividing the mean DPM /mouse within each test substance group and the positive control group by the mean DPM/mouse for the solvent/vehicle control group. The average SI for vehicle treated controls is then 1.

18. Use of the individual approach to calculate the SI will enable the performance of a statistical analysis of the data. In choosing an appropriate method of statistical analysis, the investigator should maintain an awareness of possible inequalities of variances and other related problems that may necessitate a data transformation or a non-parametric statistical analysis. An adequate approach for interpreting the

data is to evaluate all individual data of treated and vehicle controls, and derive from these the best fitting dose response curve, taking confidence limits into account (10)(12)(13). However, the investigator should be alert to possible "outlier" responses for individual animals within a group that may necessitate the use of an alternative measure of response (e.g. median rather than mean) or elimination of the outlier.

19. The decision process with regard to a positive response includes a stimulation index ≥ 3, together with consideration of dose-response and, where appropriate, statistical significance (3)(6)(10)(13)(14).

20. If it is necessary to clarify the results obtained, consideration should be given to various properties of the test substance, including whether it has a structural relationship to known skin sensitisers, whether it causes excessive skin irritation, and the nature of the dose response seen. These and other considerations are discussed in detail elsewhere (7).

DATA AND REPORTING

Data

21. Data should be summarised in tabular form showing the mean and individual DPM values and stimulation indexes for each dose (including vehicle control) group.

Test report

22. The test report should contain the following information:

Test substance:

- identification data (e.g. CAS number, if available; source; purity; known impurities; lot number);
- physical nature and physicochemical properties (e.g. volatility, stability, solubility);
- if mixture, composition and relative percentages of components.

Vehicle:

- identification data (purity; concentration, where appropriate; volume used);
- justification for choice of vehicle.

Test animals:

- strain of mice used;
- microbiological status of the animals, when known;
- number, age and sex of animals;
- source of animals, housing conditions, diet, etc.

Test conditions:

- details of test substance preparation and application;
- justification for dose selection (including results from range finding study, if conducted);-vehicle and test substance concentrations used, and total amount of substance applied;
- details of food and water quality (including diet type/source, water source).

Reliability check:

- a summary of results of latest reliability check, including information on substance, concentration and vehicle used;
- concurrent and/or historical positive and negative control data for testing laboratory.

Results:

- individual weights of animals at start of dosing and at scheduled kill;
- a table of mean/median (pooled approach) and individual (individual approach) DPM values, as well as the range of values for both approaches, and stimulation indices for each dose (including vehicle control) group;
- statistical analysis, where appropriate;
- time course of onset and signs of toxicity, including dermal irritation at site of administration, if any, for each animal.

Discussion of results:

- A brief commentary on the results, the dose-response analysis, and statistical analyses, where appropriate, with a conclusion as to whether the test substance should be considered a skin sensitiser.

LITERATURE

(1) Kimber, I. and Basketter, D.A. (1992). The murine local lymph node assay; collaborative studies and new directions: A commentary. Food and Chemical Toxicology 30, 165-169.

(2) Kimber, I. Dearman, R.J. Scholes E.W, and Basketter, D.A (1994). The local lymph node assay: developments and applications. Toxicology, 93, 13-31.

(3) Kimber, I. Hilton, J. Dearman, R.J. Gerberick, G.F. Ryan, C.A. Basketter, D.A. Lea, L. House, R.V. Ladies, G.S. Loveless, S.E. Hastings, K.L. (1998). Assessment of the skin sensitisation potential of topical medicaments using the local lymph node assay: An interlaboratory exercise. Journal of Toxicology and Environmental Health, 53 563-79 (1998).

(4) OECD (1992). Guideline 406 Skin Sensitisation.

(5) Chamberlain, M. and Basketter, D.A. (1996). The local lymph node assay: status of validation. Food and Chemical Toxicology, 34, 999-1002.

(6) Basketter, D.A. Gerberick, G.F. Kimber, I. and Loveless, S.E. (1996). The local lymph node assay A viable alternative to currently accepted skin sensitisation tests. Food and Chemical Toxicology, 34, 985-997.

(7) Basketter, D.A. Gerberick, G.F. and Kimber, I. (1998). Strategies for identifying false positive responses in predictive sensitisation tests. Food and Chemical Toxicology. (in press) 36 327-33.

(8) Van Och, F.M.M, Slob, W., De Jong, W.H., Vandebriel, R.J., Van Loveren, H. (2000). A quantitative method for assessing the sensitising potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of uncertainty margins. Toxicology 146 49-59.

(9) Dearman, R.J. Hilton, J. Evans, P. Harvey, P. Basketter, D.A. and Kimber, I. (1998). Temporal stability of local lymph node assay responses to hexyl cinnamic aldehyde. Journal of Applied Toxicology 18 281-4.

(10) National Institute of Environmental Health Sciences (1999). The murine local lymph node Assay: A test method for assessing the allergic contact dermatitis potential of chemicals/compounds: The results of an independent peer review evaluation coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods (NICETAM). NIH Publication No: 99-4494, Research Triangle Park, N.C. (http://iccvam.niehs.nih.gov).

(11) OECD (2002). Guideline 404 Acute Dermal Irritation/Corrosion.

(12) Basketter, D.A. Selbie, E. Scholes, E.W. Lees, D. Kimber, I. and Botham, P.A. (1993). Results with OECD recommended positive control sensitisers in the maximisation. Buehler and local lymph node assays. Food and Chemical Toxicology, 31, 63-67.

(13) Basketter DA, Lea LJ, Dickens A, Briggs D, Pate I, Dearman RJ, Kimber I (1999). A comparison of statistical approaches to the derivation of EC3 values from local lymph node assay dose responses. J. Appl. Toxicology, 19, 261-266.

(14) Basketter DA, Blaikie L, Derman RJ, Kimber I, Ryan CA, Gerberick GF, Harvey P, Evans P, White IR, and Rycroft RTG (2000). Use of local lymph node assay for the estimation of relative contact allergenic potency. Contact Dermatitis, 42, 344-48.

[This Page Intentionally Left Blank]

Appendix H3

OECD Test Guideline 406: Skin Sensitisation

[This Page Intentionally Left Blank]

406

Adopted:
17.07.92

OECD GUIDELINE FOR TESTING OF CHEMICALS

Adopted by the Council on 17th July 1992

Skin Sensitisation

INTRODUCTION

1. OECD Guidelines for Testing of Chemicals are periodically reviewed in light of scientific progress. In such reviews, special attention is given to possible improvements in relation to animal welfare. This updated version of the original guideline 406, adopted in 1981, is the outcome of a meeting of OECD experts held in Paris in May 1991.

2. Currently, quantitative structure-activity relationships and *in vitro* models are not yet sufficiently developed to play a significant role in the assessment of the skin-sensitisation potential of substances which therefore must continue to be based on *in vivo* models.

3. The guinea pig has been the animal of choice for predictive sensitisation tests for several decades. Two types of tests have been developed: adjuvant tests in which sensitisation is potentiated by the injection of Freunds Complete Adjuvant (FCA), and non-adjuvant tests. In the original guideline 406, four adjuvant tests and three non-adjuvant tests were considered to be acceptable. In this updated version, the Guinea Pig Maximisation Test (GPMT) of Magnusson and Kligman which uses adjuvant (1)(2)(3)(4) and the non-adjuvant Buehler Test (5)(6) are given preference over other methods and the procedures are presented in detail. It is recognised, however, that there may be circumstances where other methods may be used to provide the necessary information on sensitisation potential.

4. The immune system of the mouse has been investigated more extensively than that of the guinea pig. Recently, mouse models for assessing sensitisation potential have been developed that offer the advantages of an endpoint which is measured objectively, short duration and minimal animal treatment. The mouse ear swelling test (MEST) and the local lymph node assay (LLNA) appear to be promising. Both assays have undergone validation in several laboratories (7)(8)(9)(10)(11) and it has been shown that they are able to detect reliably moderate to strong sensitisers. The LLNA or the MEST can be used as a first stage in the assessment of skin sensitisation potential. If a positive result is seen in either assay, a test substance may be designated as a potential sensitiser, and it may not be necessary to conduct a further guinea pig test. However, if a negative result is seen in the LLNA or MEST, a guinea pig test (preferably a GPMT or Buehler Test) must be conducted using the procedure described in this guideline.

5. Definitions used are set out in the Annex.

GENERAL PRINCIPLE OF SENSITISATION TESTS IN GUINEA PIGS

6. The test animals are initially exposed to the test substance by intradermal injection and/or epidermal application (induction exposure). Following a rest period of 10 to 14 days (induction

period), during which an immune response may develop, the animals are exposed to a challenge dose. The extent and degree of skin reaction to the challenge exposure in the test animals is compared with that demonstrated by control animals which undergo sham treatment during induction and receive the challenge exposure.

ELEMENTS COMMON TO SENSITISATION TESTS IN GUINEA PIGS

Sex of animals

7. Male and/or female healthy young adult animals can be used. If females are used they should be nulliparous and non-pregnant.

Housing and feeding conditions

8. The temperature of the experimental animal room should be 20ºC (± 3ºC) and the relative humidity 30-70 per cent. Where the lighting is artificial, the sequence should be 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. It is essential that guinea pigs receive an adequate amount of ascorbic acid.

Preparation of the animals

9. Animals are acclimatised to the laboratory conditions for at least 5 days prior to the test. Before the test, animals are randomised and assigned to the treatment groups. Removal of hair is by clipping, shaving or possibly by chemical depilation, depending on the test method used. Care should be taken to avoid abrading the skin. The animals are weighed before the test commences and at the end of the test.

Reliability check

10. The sensitivity and reliability of the experimental technique used should be assessed every six months by use of substances which are known to have mild-to-moderate skin sensitisation properties.

11. In a properly conducted test, a response of at least 30% in an adjuvant test and at least 15% in a non-adjuvant test should be expected for mild/moderate sensitisers. Preferred substances are hexyl cinnamic aldehyde (CAS No. 101-86-0), mercaptobenzothiazole (CAS No. 149-30-4) and benzocaine (CAS No. 94-09-7). There may be circumstances where, given adequate justification, other control substances meeting the above criteria may be used.

Removal of the test substance

12. If removal of the test substance is considered necessary, this should be achieved using water or an appropriate solvent without altering the existing response or the integrity of the epidermis.

DESCRIPTION OF THE GUINEA-PIG MAXIMISATION TEST METHOD

Number of animals

13. A minimum of 10 animals is used in the treatment group and at least 5 animals in the control group. When fewer than 20 test and 10 control guinea pigs have been used, and it is not possible to conclude that the test substance is a sensitiser, testing in additional animals to give a total of at least 20 test and 10 control animals is strongly recommended.

Dose levels

14. The concentration of test substance used for each induction exposure should be well-tolerated systemically and should be the highest to cause mild-to-moderate skin irritation. The concentration used for the challenge exposure should be the highest non-irritant dose. The appropriate concentrations can be determined from a pilot study using two or three animals. Consideration should be given to the use of FCA-treated animals for this purpose.

Induction: Intradermal Injections

Day 0 - treated group

15. Three pairs of intradermal injections of 0.1 ml volume are given in the shoulder region which is cleared of hair so that one of each pair lies on each side of the midline.

Injection 1: a 1:1 mixture (v/v) FCA/water or physiological saline

Injection 2: the test substance in an appropriate vehicle at the selected concentration

Injection 3: the test substance at the selected concentration formulated in a 1:1 mixture (v/v) FCA/water or physiological saline.

16. In injection 3, water soluble substances are dissolved in the aqueous phase prior to mixing with FCA. Liposoluble or insoluble substances are suspended in FCA prior to combining with the aqueous phase. The concentration of test substance shall be equal to that used in injection 2.

17. Injections 1 and 2 are given close to each other and nearest the head, while 3 is given towards the caudal part of the test area.

Day 0 - control group

18. Three pairs of intradermal injections of 0.1 ml volume are given in the same sites as in the treated animals.

Injection 1: a 1:1 mixture (v/v) FCA/water or physiological saline

Injection 2: the undiluted vehicle

Injection 3: a 50% w/v formulation of the vehicle in a 1:1 mixture (v/v) FCA/water or physiological saline.

Induction: Topical Application

Day 5-7 - treated and control groups

19. Approximately twenty-four hours before the topical induction application, if the substance is not a skin irritant, the test area, after close-clipping and/or shaving is painted with 0.5 ml of 10% sodium lauryl sulphate in vaseline, in order to create a local irritation.

Day 6-8 - treated group

20. The test area is again cleared of hair. A filter paper (2 x 4 cm) is fully-loaded with test substance in a suitable vehicle and applied to the test area and held in contact by an occlusive dressing for 48 hours. The choice of the vehicle should be justified. Solids are finely pulverised and

incorporated in a suitable vehicle. Liquids can be applied undiluted, if appropriate.

Day 6-8 - control group

21. The test area is again cleared of hair. The vehicle only is applied in a similar manner to the test area and held in contact by an occlusive dressing for 48 hours.

Challenge: Topical Application

Day 20-22 - treated and control groups

22. The flanks of treated and control animals are cleared of hair. A patch or chamber loaded with the test substance is applied to one flank of the animals and, when relevant, a patch or chamber loaded with the vehicle only may also be applied to the other flank. The patches are held in contact by an occlusive dressing for 24 hours.

Observations - treated and control groups

23. - approximately 21 hours after removing the patch the challenge area is cleaned and closely-clipped and/or shaved or depilated if necessary;

- approximately 3 hours later (approximately 48 hours from the start of the challenge application) the skin reaction is observed and recorded according to the grades shown below;

- approximately 24 hours after this observation a second observation (72 hours) is made and once again recorded.

Blind reading of test and control animals is encouraged.

TABLE: MAGNUSSON AND KLIGMAN GRADING SCALE FOR THE EVALUATION OF CHALLENGE PATCH TEST REACTIONS

0 = no visible change

1 = discrete or patchy erythema

2 = moderate and confluent erythema

3 = intense erythema and swelling

Rechallenge

24. If it is necessary to clarify the results obtained in the first challenge, a second challenge (i.e. a rechallenge), where appropriate with a new control group, should be considered approximately one week after the first one. A rechallenge may also be performed on the original control group.

Clinical observations

25. All skin reactions and any unusual findings, including systemic reactions, resulting from induction and challenge procedures should be observed and recorded. Other procedures, e.g.

histopathological examination, the measurement of skin fold thickness, may be carried out to clarify doubtful reactions.

DESCRIPTION OF THE BUEHLER TEST METHOD

Number of animals

26. A minimum of 20 animals is used in the treatment group and at least 10 animals in the control group.

Dose levels

27. The concentration of test substance used for each induction exposure should be the highest to cause mild irritation. The concentration used for the challenge exposure should be the highest non-irritating dose. The appropriate concentration can be determined from a pilot study using two or three animals.

28. For water soluble test materials, it is appropriate to use water or a dilute non-irritating solution of surfactant as the vehicle. For other test materials 80% ethanol/water is preferred for induction and acetone for challenge.

Induction: Topical application

Day 0 - treated group

29. One flank is cleared of hair (closely-clipped). The test patch system should be fully loaded with test substance in a suitable vehicle (the choice of the vehicle should be justified; liquid test substances can be applied undiluted, if appropriate). The test patch system is applied to the test area and held in contact with the skin by an occlusive patch or chamber and a suitable dressing for 6 hours.

30. The test patch system must be occlusive. A cotton pad is appropriate and can be circular or square, but should approximate 4-6 cm^2. Restraint using an appropriate restrainer is preferred to assure occlusion. If wrapping is used, additional exposures may be required.

Day 0 - control group

31. One flank is cleared of hair (closely-clipped). The vehicle only is applied in a similar manner to that used for the treated group. The test patch system is held in contact with the skin by an occlusive patch or chamber and a suitable dressing for 6 hours. If it can be demonstrated that a sham control group is not necessary, a naive control group may be used.

Days 6-8 and 13-15 - treated and control groups

32. The same application as on day 0 is carried out on the same test area (cleared of hair if necessary) of the same flank on day 6-8, and again on day 13-15.

Challenge

Day 27-29 - treated and control groups

33. The untreated flank of treated and control animals is cleared of hair (closely-clipped). An occlusive patch or chamber containing the appropriate amount of test substance is applied, at the maximum non-irritant concentration, to the posterior untreated flank of treated and control animals.

When relevant, an occlusive patch or chamber with vehicle only is also applied to the anterior untreated flank of both treated and control animals. The patches or chambers are held in contact by a suitable dressing for 6 hours.

Observations - treated and control groups

34. - approximately 21 hours after removing the patch the challenge area is cleared of hair;

- approximately three hours later (approximately 30 hours after application of the challenge patch) the skin reactions are observed and recorded according to the grades shown in the Guinea-Pig Maximisation Test (see paragraph 23);

- approximately 24 hours after the 30 hour observation (approximately 54 hours after application of the challenge patch) skin reactions are again observed and recorded.

Blind reading of test and control animals is encouraged.

Rechallenge

35. If it is necessary to clarify the results obtained in the first challenge, a second challenge (i.e. a rechallenge), where appropriate with a new control group, should be considered approximately one week after the first one. The rechallenge may also be performed on the original control group.

Clinical observations

36. All skin reactions and any unusual findings, including systemic reactions, resulting from induction and challenge procedures should be observed and recorded. Other procedures, e.g. histopathological examination, measurement of skin fold thickness, may be carried out to clarify doubtful reactions.

DATA AND REPORTING (GPMT and Buehler Test)

Data

37. Data should be summarised in tabular form, showing for each animal the skin reactions at each observation.

Test report

38. The test report must include the following information:

Test substance:

- physical nature and, where relevant, physicochemical properties;
- identification data.

Vehicle:

- justification of choice of vehicle.

Test animals:

- strain of guinea-pig used;

- number, age and sex of animals;
- source, housing conditions, diet, etc.;
- individual weights of animals at the start and at the conclusion of the test.

Test conditions:

- technique of patch site preparation;
- details of patch materials used and patching technique;
- result of pilot study with conclusion on induction and challenge concentrations to be used in the test;
- details of test substance preparation, application and removal;
- vehicle and test substance concentrations used for induction and challenge exposures and the total amount of substance applied for induction and challenge.

Reliability check:

- a summary of the results of the latest reliability check including information on substance, concentration and vehicle used.

Results:

- on each animal including grading system;
- narrative description of the nature and degree of effects observed;
- any histopathological findings.

Discussion of the results.

If a screening assay is performed before the guinea pig test the description or reference of the test, including details of the procedure, must be given together with results obtained with the test and reference substances.

LITERATURE

(1) Magnusson B. and Kligman A.M. (1969). The identification of contact allergens by animal assay. The guinea pig maximisation test. J. Invest. Dermatol., 52, 268.

(2) Magnusson B. and Kligman A.M. (1970). Allergic Contact Dermatitis in the Guinea Pig. Charles G. Thomas; Springfield, Illinois.

(3) Magnusson B. (1980). Identification of contact sensitizers by animal assay. Cont. Derm., 6, 46.

(4) Magnusson B., Fregert S. and Wahlberg J. (1979). Determination of skin sensitization potential of chemicals. Predictive testing in guinea pigs. Arbete och Hälsa, 26(E).

(5) Buehler E.V. (1965). Delayed contact hypersensitivity in the guinea pig. Arch. Dermatol., 91, 171.

(6) Ritz H.L. and Buehler E.V. (1980). Procedure for conducting the guinea pig assay. Current Concepts in Dermatology, Drill V.A. and Lazar P. (eds), Academic Press, New York, N.Y., 25-40.

(7) Kimber I., Hilton J. and Weisenberger C. (1989). The murine local lymph mode assay for identification of contact allergens: a preliminary evaluation of in situ measurement of lymphocyte proliferation Contact Dermatitis, 21, 215-220.

(8) Kimber I., Hilton J. and Botham P.A. (1990). Identification of contact allergens using the murine local lymph node assay: comparisons with the Buehler Occluded Patch Test in guinea pigs. Journal of Applied Toxicology, 10(3), 173-180.

(9) Kimber I., Hilton J., Botham P.A., Basketter D.A., Scholes E.W., Miller K., Robbins M.C., Harrison P.T.C., Gray T.J.B. and Waite S.J. (1991). The murine local lymph node assay: results of an inter-laboratory trial. Toxicology Letters, 55, 203-213.

(10) Basketter D.A., Scholes E.W., Kimber I., Botham P.A., Hilton J., Miller K., Robbins M.C., Harrison P.T.C. and Waite S.J. (1991). Interlaboratory evaluation of the local lymph node assay with 25 chemicals and comparison with guinea pig test data. Toxicology Methods, 1, 30-43.

(11) Gad S.C., Dunn B.J., Dobbs D.W., Reilly C. and Walsh R.D. (1986). Development and validation of an alternative dermal sensitisation test: the mouse ear swelling test (MEST). Toxicol. Appl. Pharmacol., 84, 93-114.

ANNEX

DEFINITIONS

Skin sensitisation (allergic contact dermatitis) is an immunologically mediated cutaneous reaction to a substance. In the human, the responses may be characterised by pruritis, erythema, oedema, papules, vesicles, bullae or a combination of these. In other species the reactions may differ and only erythema and oedema may be seen.

Induction exposure: an experimental exposure of a subject to a test substance with the intention of inducing a hypersensitive state.

Induction period: a period of at least one week following an induction exposure during which a hypersensitive state may develop.

Challenge exposure: an experimental exposure of a previously treated subject to a test substance following an induction period, to determine if the subject reacts in a hypersensitive manner.

[This Page Intentionally Left Blank]

Appendix H4

EPA Health Effects Test Guidelines
OPPTS 870.2600: Skin Sensitization

[This Page Intentionally Left Blank]

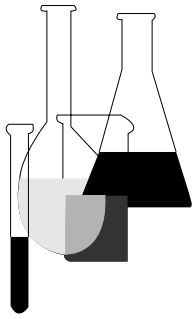
United States
Environmental Protection
Agency

Prevention, Pesticides
and Toxic Substances
(7101)

EPA 712–C–03–197
March 2003

Health Effects Test Guidelines

OPPTS 870.2600
Skin Sensitization



INTRODUCTION

This guideline is one of a series of test guidelines that have been developed by the Office of Prevention, Pesticides and Toxic Substances, United States Environmental Protection Agency for use in the testing of pesticides and toxic substances, and the development of test data that must be submitted to the Agency for review under Federal regulations.

The Office of Prevention, Pesticides and Toxic Substances (OPPTS) has developed this guideline through a process of harmonization that blended the testing guidance and requirements that existed in the Office of Pollution Prevention and Toxics (OPPT) and appeared in Title 40, Chapter I, Subchapter R of the Code of Federal Regulations (CFR), the Office of Pesticide Programs (OPP) which appeared in publications of the National Technical Information Service (NTIS) and the guidelines published by the Organization for Economic Cooperation and Development (OECD).

The purpose of harmonizing these guidelines into a single set of OPPTS guidelines is to minimize variations among the testing procedures that must be performed to meet the data requirements of the U. S. Environmental Protection Agency under the Toxic Substances Control Act (15 U.S.C. 2601) and the Federal Insecticide, Fungicide and Rodenticide Act (7 U.S.C. 136, *et seq.*).

Final Guideline Release: This guideline is available from the U.S. Government Printing Office, Washington, DC 20402 on disks or paper copies: call (202) 512–0132. This guideline is also available electronically in PDF (portable document format) from EPA's Internet Web site at http://www.epa.gov/opptsfrs/home/guidelin.htm.

OPPTS 870.2600 Skin sensitization.

(a) **Scope**—(1) **Applicability**. This guideline is intended to meet testing requirements of both the Federal lnsecticide, Fungicide, and Rodenticide Act (FIFRA) (7 U.S.C. l36, *et seq*.) and the Toxic Substances Control Act (TSCA) (15 U.S.C. 2601).

(2) **Background**. The source materials used in developing this harmonized OPPTS test guideline are OPPTS Harmonized Test Guidelines Series 870, Guideline 870.2600 Skin Sensitization, dated August 1998; 40 CFR 798.4100 Dermal Sensitization; OECD 406 Skin Sensitization (adopted July 1992); and OECD 429 Skin Sensitization: Local Lymph Node Assay (adopted April 2002).

(b) **Purpose**. The purpose of the selected test is to identify substances with skin sensitization potential. Determination of the potential to cause or elicit skin sensitization reactions (allergic contact dermatitis) is an important element in evaluating a substance's toxicity. Information derived from skin sensitization tests serves to identify possible hazards to a population exposed repeatedly to a test substance. Testing is not required if the test material is a known skin sensitizer. If it is suspected that the test material is a strong dermal irritant, see OPPTS 870.1000, paragraph (d)(2)(iii).

(c) **Definitions**. The following definitions apply to this test guideline. The definitions in Section 3 of TSCA and in 40 CFR Part 792—Good Laboratory Practice Standards (GLP) also apply to this test guideline.

Challenge exposure is an exposure of a previously treated subject to a test substance following an induction period to elicit a contact hypersensitivity response.

Induction exposure is the administration of a test substance to the test subject with the intention of inducing contact sensitization.

Induction period is a period of at least 1 week following an induction exposure during which sensitization may develop.

Skin sensitization (allergic contact dermatitis) is an immunologically mediated cutaneous reaction to a substance. In the human, the responses may be characterized by pruritis, erythema, edema, papules, vesicles, bullae, or a combination of these. In other mammalian species, the reactions may differ and only erythema and edema may be seen.

Stimulation index (SI) is the ratio of ^{3}H-methyl thymidine or ^{125}I-iododeoxyuridine (^{125}IU) incorporation into test group lymph nodes relative to that recorded for solvent/vehicle control group lymph nodes.

(d) **Test procedures**—(1) **Methods**. Any of the following test methods is considered to be acceptable:

(i) Local Lymph Node Assay (LLNA) test, or

(ii) Guinea-Pig Maximization Test (GPMT), or

(iii) Buehler test.

(2) **Choice of assays**. See OPPTS 870.1000 for a general discussion of factors to be considered prior to performing the test. In addition, the following considerations apply:

(i) The LLNA (see references in paragraphs (g)(1) through (g)(6) of this guideline) is a preferred alternative method, where applicable, to the traditional guinea pig test because it demonstrates an equivalent prediction of human allergic contact dermatitis as compared to the other sensitization tests, provides quantitative data and an assessment of dose-response, gives consideration to animal welfare concerns, and is suitable for testing colored substances. It should be recognized that there are certain testing situations that may necessitate the use of traditional guinea pig tests. The tester should note that the LLNA may not be appropriate for all types of test materials, such as certain metallic compounds, high molecular weight proteins, strong dermal irritants and materials that do not sufficiently adhere to the ear for an acceptable period of time during treatment. When using the LLNA, particular care should be taken to ensure that hydrophilic materials are incorporated into a vehicle system that wets the skin and does not immediately run off. Thus, wholly aqueous vehicles or test materials and runny liquids are to be avoided. In all instances, the tester must document that appropriate techniques were used to facilitate adherence to the mouse ear for an adequate exposure duration. It may be possible to use the LLNA to test some of these materials if appropriate techniques are used to facilitate adherence.

(ii) In situations for test materials where the LLNA is not applicable or may provide unreliable or problematic results, the GPMT or Buehler tests are recommended (see references in paragraphs (g)(7) through (g)(14) of this guideline).

(iii) Although the LLNA, GPMT, or Buehler tests are considered to be acceptable tests, it is recognized that other tests may give useful results. If other tests are used, the investigator must provide justification/reasoning for use of other procedures and methods and protocols must be provided. A positive and negative control group must be included in each test.

(e) **Test methods**—(1) **LLNA method**—(i) **Principle of the method**. The basic principle underlying the LLNA is that skin sensitizers induce proliferation of lymphocytes in the lymph nodes draining the site of chemical application. Generally, under appropriate test conditions, this proliferation is proportional to the dose applied, and provides a means of obtaining an objective, quantitative measurement of sensitization. The test measures cellular proliferation as a function of *in vivo* radioisotope incorporation

into the DNA of dividing lymphocytes. The LLNA assesses this proliferation in the draining auricular lymph nodes located in the cervical region at the bifurcation of the jugular vein. Lymphocyte proliferation in test groups is compared to that in concurrent solvent/vehicle-treated controls. A positive control is added to each assay to provide an indication of appropriate assay performance.

(ii) **Animal selection**—(A) **Sex and strain of animals**. Young adult female mice (nulliparous and non-pregnant) of the CBA/Ca or CBA/J strain should be used at age 8–12 weeks. All animals are to be age-matched (preferably within a one-week time frame). Females are used because the existing database is predominantly based on this gender. Males and other strains of mice should not be used until it is sufficiently demonstrated that significant strain-specific and/or gender-specific differences in the LLNA response do not exist.

(B) **Housing and feeding**. The temperature of the experimental animal room should be 21 ± 3 °C and the relative humidity 30–70%. When artificial lighting is used, the light cycle should be 12 hours light: 12 hours dark. For feeding, standard laboratory mouse diets are to be used with an unlimited supply of drinking water. The mice must be acclimatized for at least 5 days prior to the start of the test. Animals must be housed individually. Healthy animals are randomly assigned to control and treatment groups having statistically homogeneous body weights. The animals are uniquely identified prior to being placed on study. Although a variety of techniques exist to uniquely mark mice, any method that involves identification via ear marking (e.g., ear tags) must not be used.

(iii) **Test conditions**—(A) **Preparation of doses**. Solid test substances are to be dissolved in appropriate solvents or vehicles and diluted, if appropriate, prior to dosing of the animals. Stable suspensions might also be acceptable. Liquid test substances may be dosed directly or diluted prior to dosing. Fresh preparations of the test substance are to be prepared daily unless stability data demonstrate the acceptability of storage.

(B) **Solvent/vehicle**. The solvent/vehicle is to be selected on the basis of maximizing the test concentration while producing a solution/suspension suitable for application of the test substance. In order of preference, recommended solvents/vehicles are acetone/olive oil (4:1 v/v), *N*,*N*-dimethylformamide, methyl ethyl ketone, propylene glycol, and dimethyl sulfoxide, but others may be used if appropriately justified. The selected solvent/vehicle must not interfere with or bias the test result and should be selected to achieve the maximum concentration/skin exposure of the test substance. Ensure that hydrophilic materials are incorporated into a vehicle system that wets the skin and does not immediately run off. Thus, wholly aqueous vehicles are to be avoided.

(C) **Controls**. (*1*) Concurrent negative (solvent/vehicle) and positive controls are to be included in each test. In some circumstances, it may be useful to include a naive control. Except for treatment with the test substance, animals in the control groups are to be handled in an identical manner to animals of the treatment groups.

(*2*) Positive controls are used to ensure the appropriate performance of the assay. The positive control must produce a positive LLNA response at an exposure level expected to give an increase in the stimulation index (SI) of three or greater (SI ≥ 3) over the solvent or vehicle control group. The positive control dose is to be chosen such that the induction is clear but not excessive. Preferred positive control substances are hexyl cinnamic aldehyde (HCA) and mercaptobenzothiazole. There may be circumstances where, given adequate justification, other positive control substances may be used. However, benzocaine should not be used as a positive control in the LLNA.

(*3*) The positive control substance is tested in the vehicle that is known to elicit a consistent response (i.e., acetone/olive oil). If a non-standard vehicle (chemically relevant formulation) is used with a positive control, the non-standard vehicle (chemically relevant formulation) must be tested for a local lymph node response prior to the initiation of the study and the results reported.

(iv) **LLNA test procedure**—(A) **A minimum of five animals are used per dose group**. At least three consecutive doses of the test substance are to be used. A solvent/vehicle control group and a positive control group are also required. Doses are normally selected from within the concentration series 100%, 50%, 25%, 10%, 5%, 2.5%, 1%, 0.5%, 0.1%. In general, dose selection is based on factors such as toxicity, solubility, irritancy and any other available information such as the results of other testing and structure-activity relationships. To avoid false negatives, test as high a concentration as possible. Generally, the maximum concentration tested is the highest achievable level that avoids overt systemic toxicity and excessive local irritation. To identify the appropriate maximum test substance dose, an initial toxicity test, conducted under identical experimental conditions except for an assessment of lymph node proliferative activity, may be necessary. To support an ability to identify a dose-response relationship, data must be collected on at least three test substance treatment doses, in addition to the concurrent solvent/vehicle control group. Where the LLNA study results are negative, the concurrent positive control must induce a SI ≥ 3 relative to its solvent/vehicle-treated control.

(B) **LLNA experimental procedure**. The LLNA experimental procedure is to be performed by appropriately trained staff as follows:

(*1*) Day 1. Record the body weight of each mouse prior to dermal applications. Apply 25 μL/ear of the appropriate dilution of the test sub-

stance, or the positive control, or the solvent/vehicle control alone to the dorsum of both ears. A positive displacement pipettor may facilitate application of the test material.

(*2*) Days 2 and 3. Repeat the application procedure as carried out on day 1.

(*3*) Days 4 and 5. No treatment.

(*4*) Day 6. Record the body weight of each mouse. Inject 250 µL of sterile phosphate buffered saline (PBS) containing 20 µCi of ^{3}H-methyl thymidine or 250 µL PBS containing 2 µCi ^{125}IU and 10^{-5} M fluorodeoxyuridine into each experimental mouse via the tail vein. Five hours later, the draining (auricular) lymph node of each ear is excised and pooled in PBS for each animal. A single cell suspension of lymph node cells (LNC) is prepared for each mouse. The single cell suspension is prepared in PBS by either gentle mechanical separation through 200-mesh stainless steel gauze or another acceptable technique for generating a single cell suspension. The LNC are washed twice with an excess of PBS and the DNA precipitated with 5% trichloroacetic acid (TCA) at 4 ^{o}C for approximately 18h.

(*5*) For the ^{3}H-methyl thymidine method, pellets are resuspended in 1 mL TCA and transferred to 10 mL of scintillation fluid. Incorporation of ^{3}H-methyl thymidine is measured by B-scintillation counting as disintegrations per minute (dpm) for each mouse and expressed as dpm/mouse. For the ^{125}IU method, the 1 mL TCA pellet is transferred directly into gamma counting tubes. Incorporation of ^{125}IU is determined by gamma counting and also expressed as dpm/mouse.

(C) **Observations**. At a minimum, observe mice once daily for any clinical signs, either of local irritation at the application site or of systemic toxicity. Weighing mice prior to treatment and at the time of necropsy will aid in assessing systemic toxicity. All observations are systematically recorded, with records being maintained for each individual mouse.

(D) **Measurements and calculation of results**. (*1*) The proliferative response of lymph node cells from the pooled lymph nodes of each individual animal is expressed as the number of radioactive disintegrations per minute (dpm) per animal, subtracting out any background dpm. Then the group mean dpm, along with an appropriate measure of inter-animal variability (i.e., mean ± standard deviation), is calculated for each test group (i.e., positive, solvent/vehicle, and any other control groups) and the solvent/vehicle group. Final results are expressed as the SI which is calculated as a ratio (i.e., SI = mean dpm of test group divided by mean dpm of solvent/vehicle control group).

(*2*) In addition to an assessment of the magnitude of the ratio estimate, SI, conduct statistical analyses which include both an overall assess-

ment (e.g. ANOVA) of the dose-response relationships and pairwise comparisons of the SIs of the test groups, positive control group and any other control group versus that of the solvent/vehicle control group. In choosing an appropriate method of statistical analysis, the investigator should be aware of possible inequality of variances and other related problems that may necessitate a data transformation or a nonparametric statistical analysis.

(v) **Data interpretation and reporting for LLNA**—(A) **Data Interpretation**. (*1*) A substance is regarded as a skin sensitizer in the LLNA if at least one concentration of the test material results in a 3-fold or greater increase in ^{3}H-methyl thymidine or ^{125}IU incorporation in the lymph node cells of test group lymph nodes relative to that recorded for solvent/vehicle control lymph nodes, as indicated by the SI. However, the magnitude of the SI should not be the sole factor used in determining the biological significance of a skin sensitization response. A quantitative assessment must be performed by statistical analysis of individual animal data in order to provide a more complete evaluation of the test substance (see paragraph (e)(1)(iv)(D)(*2*) of this guideline). Factors to be considered in evaluating the biological significance of a response or outcome of the test include the results of the SI determinations, statistical analyses, the strength of the dose-response relationship, chemical toxicity, solubility, and the consistency of the solvent/vehicle and positive control responses.

(*2*) Strong irritants may yield false positive results in the LLNA due to the initiation of a significant lymphocyte proliferation. However, the dose-response information from the assay may help to uncover a strong irritant response since, for instance, it has been shown that the proliferation induced by irritation usually results in a shallow dose-response relationship. Concurrent evaluation of ear swelling may also provide helpful information on differentiating weak sensitizers from strong irritants.

(B) **Test report**. The test report for LLNA must contain the following specific information:

(*1*) Test substance. (*i*) Identification data and CAS number, if known, and EPA registration number, if applicable;

(*ii*) Physical nature and purity;

(*iii*) Physicochemical properties relevant to the conduct of the study;

(*iv*) Stability of the test substance, if known; and

(*v*) Lot number of the test substance.

(*2*) Solvent/vehicle. (*i*) Solvent/vehicle used and its purity;

(*ii*) Justification for choice of solvent/vehicle, if appropriate; and

(*iii*) Solubility and stability of the test substance in the solvent/vehicle.

(*3*) Test animals. (*i*) Strain of mice used;

(*ii*) Acclimation information;

(*iii*) Number, age, and sex of mice;

(*iv*) Source, housing conditions, diet, etc.;

(*v*) Individual body weight of the animals at the start and end of the test, including body weight range, mean, and associated error term for each group;

(*vi*) Health and microbiological/pathogen status of the mouse; and

(*vii*) Details of animal food and water quality;

(*4*) Test conditions. (*i*) Details of test substance preparation;

(*ii*) Details of the administration of the test substance;

(*iii*) Detailed description of treatment and sampling schedules; and

(*iv*) Methods for measurement of toxicity.

(*5*) Results. (*i*) Positive and negative (solvent/vehicle) control data in tabular form;

(*ii*) Data from range-finding study, if conducted;

(*iii*) Doses used;

(*iv*) Rationale for dose level selection;

(*v*) Signs of toxicity;

(*vi*) Dpm/mouse values for each mouse within each treatment group and control group;

(*vii*) Group mean dpm/mouse and associated error term for each treatment group and control group;

(*viii*) The SI calculated, compared to the concurrent solvent/vehicle control group, for each test substance treatment dose group, the concurrent positive control group, and any other concurrent control group;

(*ix*) Individual mouse dpm data must be presented in tabular form, along with the group mean dpm, its associated error term and the SI for each dose group;

(*x*) Criteria for considering studies as positive or negative (including information on any qualitative or quantitative measure of ear swelling);

(*xi*) Dose-response relationship;

(*xii*) Statistical analyses and method applied;

(*xiii*) Concurrent and negative control data as established in the tester's laboratory; and

(*xiv*) Concurrent positive control data.

(*6*) Discussion of the results.

(*7*) Conclusions.

(*8*) The reporting requirements specified under 40 CFR Part l58 (for pesticides) and 40 CFR Part 792, Subpart J (for toxic substances) should be followed.

(2) **GPMT and Buehler Methods**—(i) **Principle of the test methods**. Following initial exposure to a test substance, the animals are subjected, after a period of not less than 1 week, to a challenge exposure with the test substance to establish whether a hypersensitive state has been induced. Sensitization is determined by examining the reaction to the challenge exposure and comparing this reaction with that of the initial induction exposure. The test animals are initially exposed to the test substance by intradermal and/or epidermal application (induction exposure). Following a rest period of 10 to 14 days (the induction period), during which an immune response may develop, the animals are exposed to a challenge dose. The extent and degree of skin reaction to the challenge exposure is compared with that demonstrated by control animals that undergo sham treatment during induction and then receive the challenge exposure.

(ii) **Animal selection**—(A) **Species and strain**. The young adult guinea pig is preferred. Young adult commonly used laboratory strains must be employed.

(B) **Housing and feeding**. The temperature of the experimental animal room should be 20 ± 3 °C with the relative humidity 30–70 percent. Where the lighting is artificial, the sequence should be 12 h light/12 h dark. Conventional laboratory diets may be used with an unlimited supply of drinking water. It is essential that guinea pigs receive an adequate amount of ascorbic acid.

(C) **Number and sex**. The number and sex will depend on the method chosen. Either sex may be used in the Buehler test and the GPMT. If females are used, they must be nulliparous and not pregnant. The Buehler test recommends using a minimum of 20 animals in the treatment and at least 10 as controls. At least 10 animals in the treatment group and 5 in the control group must be used with the GPMT, with the stipulation that if it is not possible to conclude that the test substance is a sensitizer after using fewer than 20 test and 10 control guinea pigs, the testing of

additional animals to give a total of at least 20 test and 10 control animals is strongly recommended

(D) **Control animals**. (*2*) Every 6 months, assess the sensitivity and reliability of the experimental technique in naive animals by the use of positive control substances known to have mild-to-moderate skin-sensitizing properties. In a properly conducted test, a response of at least 30 percent in an adjuvant test and at least 15 percent in a nonadjuvant test is expected for mild-to-moderate sensitizers. Preferred substances are hexylcinnamic aldehyde (CAS No.101–86–0), mercaptobenzothiazole (CAS No. 149–30–4), benzocaine (CAS No. 94–09–7), dinitro-chloro-benzene (CAS No. 97–00–7), or DER 331 epoxy resin (CAS No. 25068–38–6). There may be circumstances where, given adequate justification, other control substances meeting the above criteria may be used.

(*2*) To ensure that the response to the challenge reaction in treated animals is truly of allergic origin and not due to skin irritancy, a sham-treated vehicle-only control is included in the test strategy. This sham-treated control group is treated in exactly the same manner as the test animals, except that during the induction phase the test article is omitted. The selected vehicle must not interfere or alter the test results.

(E) **Dose levels**. The dose level will depend on the test method selected. In the Buehler test, select the concentration of the induction dose such that it is high enough to cause mild irritation, and the challenge dose such that it is the highest non-irritating concentration. In the GPMT, the concentration of the induction dose must be well tolerated systemically, and must be high enough to cause mild-to-moderate skin irritation; the GPMT challenge dose must use the highest non-irritating concentration.

(F) **Observation of animals**. (*1*) Skin reactions are to be graded and recorded after the challenge exposures at the time specified by the methodology selected. This is usually at 24 and 48 hours. Additional notations are to be made as necessary to fully describe unusual responses.

(*2*) Regardless of the test method selected, initial and terminal body weights must be taken and recorded.

(G) **Procedures**. The procedures to be used are those described by the test method chosen. Brief summaries are given here, but the tester is referred to the original literature for more complete guidance on conducting the Buehler test (see references in paragraphs (g)(7) through (g)(10) of this guideline) or the GPMT (see references in paragraphs (g)(11) through (g)(14) of this guideline).

(*1*) The Buehler test uses topical administration via a closed patch on days 0, 6–8, and 13–15 for induction, with topical challenge of the untreated flank for 6 hours on day 27–28. Readings are made approximately 24 hours alter removing the challenge patch, and again 24 hours

after that. If the results are equivocal, the animals may be rechallenged one week later, using either the original control group or a new control group for comparison.

(*2*) The GPMT uses intradermal injection with and without Freund's complete adjuvant (FCA) for induction, followed on days 5–8 by topical irritation/induction, followed by topical challenge for 24 hours on day 20–22. Readings are made approximately 24 hours after removal of the challenge dose, and again after another 24 hours. As with the Buehler test, if the results are equivocal, the animals may be rechallenged 1 week later. If only 10 animals were used initially and gave equivocal results, the use of an additional 10 experimental and 5 control animals is strongly recommended.

(*3*) Blind reading of both test and control animals is recommended.

(*4*) Removal of the test material is accomplished with water or an appropriate solvent, without altering the existing response or the integrity of the epidermis.

(*5*) Hair is removed from the site of application by clipping, shaving, or possibly by depilation, depending on the test selected.

(iii) **Data and reporting for GPMT and Buehler Methods**. Data must be summarized in tabular form, showing for each individual animal the skin reaction, results of the induction exposure, and the challenge exposure at times indicated by the method chosen. As a minimum, the erythema and edema must be graded and any unusual finding must be recorded.

(A) **Evaluation of the results**. The evaluation of results will provide information on the proportion of each group that became sensitized and the extent (slight, moderate, severe) of the sensitization reaction in each individual animal.

(B) The following specific information is to be reported for the GPMT and Buehler Methods.

(*1*) A description of the method used and the commonly accepted name.

(*2*) Information on the positive control study, including the positive control substance used, the method used, and the time conducted.

(*3*) The number, species, strain, age, source, and sex of the test animals.

(*4*) Individual body weights of the animals at the start of the test and at the conclusion of the test.

(*5*) A brief description of the grading system.

(*6*) Each reading made on each individual animal.

(*7*) The chemical identification and relevant physicochemical properties of the test substance.

(*8*) Manufacturer, source, purity, and lot number of test substance.

(*9*) Physical nature, and, where appropriate, concentration and pH value for the test substance.

(*10*) The vehicles used for induction and challenge and justification for their use, if other than water or physiological saline. Any material that might reasonably be expected to react with or enhance or retard absorption of the test substance must be reported.

(*11*) The total amount of test substance applied for induction and challenge, and the technique of application in each case.

(*12*) Description of any pre-test conditioning, including diet, quarantine and treatment of disease.

(*13*) Description of caging conditions including number (and any change in number) of animals per cage, bedding material, ambient temperature and humidity, photoperiod, and identification of diet of test animals.

(*14*) Histopathological findings, if any.

(*15*) Discussion of results.

(*16*) A list of references cited in the body of the report, i.e., references to any published literature used in developing the test protocol, performing the testing, making and interpreting observations, and compiling and evaluating the results.

(*17*) The reporting requirements as specified under 40 CFR Part l58 (for pesticides) and 40 CFR Part 792, Subpart J (for toxic substances) should be followed

(f) **Screening tests**. The mouse ear swelling test (MEST) (see references in paragraphs (g)(15) through (g)(18) of this guideline) may be used as a screening test to detect moderate to strong sensitizers. If a positive result is seen in this assay, the test substance may be designated a potential sensitizer, and it may not be necessary to conduct a further test in guinea pigs. If the MEST does not indicate sensitization, the test substance should not be designated a nonsensitizer without confirmation in an accepted test using guinea pigs or LLNA if appropriate.

(g) **References**. The following references should be consulted for additional background information on this test guideline.

(1) *The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals/Compounds.* Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), National Institutes of Environmental Health Sciences, NIH Publication No. 99–4494.3 (1999). (Document available at http://iccvam.niehs.nih.gov/methods/llnadocs/llnarep.pdf.) Description and picture of auricular lymph node dissection available at at http://iccvam.niehs.nih.gov/methods/llnadocs/LLNAProt.pdf.

(2) Kimber, I. et al. The murine local lymph node assay for identification of contact allergens: a preliminary evaluation of in situ measurement of lymphocyte proliferation. *Contact Dermatitis* 21:215–220 (1989)

(3) Kimber, I. et al. Identification of contact allergens using the murine local lymph node assay: comparisons with the Buehler Occluded Patch Test in guinea pigs. *Journal of Applied Toxicology* 10:173–180 (1990).

(4) Kimber, I. et al. The murine local lymph node assay: results of an interlaboratory trial. *Toxicology Letters* 55:203–213 (1991).

(5) Basketter, D.A. et al. Interlaboratory evaluation of the local lymph node assay with 25 chemicals and comparison with guinea pig test data. *Toxicology Methods* 1:30–43 (1991).

(6) Basketter, D.A. et al. Identification of metal allergens in the local lymph node assay. *American Journal of Contact Dermatitis* 10 (4):207–212 (1999).

(7) Buehler, E.V. Delayed contact hypersensitivity in the guinea pig, *Archives of Dermatology* 91:171–177 (1965).

(8) Ritz, H.L. and Buehler, E.V. Planning, conduct and interpretation of guinea pig sensitization patch test in *Current Concepts in Cutaneous Toxicity,* ed. V. Drill and P. Lazar. Academic, New York, NY. pp. 25–42 (l980).

(9) Buehler, L.V. Occ1usive patch method for skin sensitization in guinea pigs: the Buehler method. *Food and Chemical Toxicology* 32:97–101 (1994).

(10) Buehler, E.V. Prospective testing for delayed contact hypersensitivity in guinea pigs: the Buehler method; in *Methods in Immunotoxico1ogy,* Vol. 2, ed. C. Burleson, A. Munson, and J. Dean. Wiley, NY, pp. 343–356(1995).

(11) Magnusson, B. and Kligman, A.M. The identification of contact allergens by animal assay. The guinea pig maximization test. *Journal of Investigative Dermatology* 52: 268–276 (1969).

(12) Magnusson, B. and Kligman, A.M. *Allergic contact dermatitis in the guinea pig.* Charles C. Thomas, Springfield, IL (1970).

(13) Magnusson, B. Identification of contact sensitizers by animal assay. *Contact Dermatology* 6:46 (1980).

(14) Magnusson, B. et al. Determination of skin sensitization potential of chemicals. Predictive testing in guinea pigs. *Arbete och Halsa:* 26(E) (1979).

(15) Gad, S.C. et al. Development and validation of an alternative dermal sensitization test: the mouse ear swelling test (MEST). *Toxicology and Applied Pharmacology 8* 4:93–114 (1986).

(16) Maisey, J. and Miller, K., Assessment of the ability of mice fed on Vitamin-A supplemented diet to respond to a variety of potential contact sensitizers. Contact Dermatitis 15:17–23 (l986).

(17) Thorne, P.S. et al., The noninvasive mouse ear swelling assay, I. Refinements for detecting weak contact sensitizers. *Fundamental and Applied Toxicology* 7:790–806 (1991).

(18) Thorne, P.S. et al. The noninvasive mouse ear swelling assay. II. Testing the contact sensitizing potency of fragrances. *Fundamental and Applied Toxicology*7:807–820 (1991).

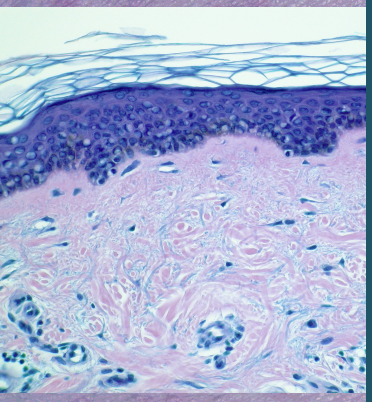
[This Page Intentionally Left Blank]

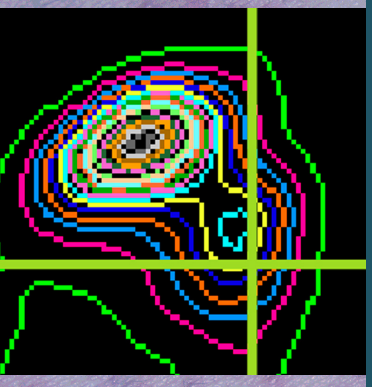
National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

NICEATM
National Toxicology Program
Interagency Center for the Evaluation
of Alternative Toxicological Methods

ICCVAM
Interagency Coordinating
Committee on the Validation
of Alternative Methods





INDEPENDENT SCIENTIFIC PEER REVIEW PANEL REPORT

Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products

May 2008

NIEHS
National Institute of
Environmental Health Sciences

NTP

Independent Scientific Peer Review Panel Report: Validation Status of New Versions and Applications of the Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals and Products

May 2008

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Institute of Environmental Health Sciences (NIEHS)
National Institutes of Health
U.S. Public Health Service
Department of Health and Human Services

National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709

This document is available electronically at:
http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRept2008.pdf

The findings and conclusions of this report are those of the Independent Scientific Peer Review Panel and should not be construed to represent the official views of ICCVAM or its member agencies.

Table of Contents

List of Tables

List of Abbreviations and Acronyms

ACD	Allergic contact dermatitis
ANOVA	Analysis of variance
AOO	Acetone: olive oil (4:1)
BRD	Background Review Document
BrdU	Bromodeoxyuridine
BT	Buehler Test
CD4	Cluster of differentiation 4
CPSC	U.S. Consumer Product Safety Commission
CRO	Clinical research organization
CV	Coefficient of variation
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNCB	Dinitrochlorobenzene
EC3	Estimated concentration needed to produce a stimulation index of three
ECt	Estimated concentration needed to produce a stimulation index that is indicative of a positive response
ECVAM	European Centre for the Validation of Alternative Methods
ELISA	Enzyme Linked Immunosorbent Assay
eLLNA: BrdU-FC	Enhanced LLNA with BrdU detected by flow cytometry
EPA	U.S. Environmental Protection Agency
FC	Flow cytometry
FR	*Federal Register*
GLP	Good Laboratory Practice
GPMT	Guinea Pig Maximization Test
GSK	GlaxoSmithKline
HCA	Hexyl cinnamic aldehyde
HMT	Human Maximization Test
HRIPT	Human Repeat Insult Patch Test
HTdR	^{3}H-Methyl Thymidine
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ISO	International Organization for Standardization
IWG	Immunotoxicity Working Group
JaCVAM	Japanese Center for Validation of Alternative Methods
LLNA	Local Lymph Node Assay
LLNA: BrdU-ELISA	LLNA with BrdU detected by ELISA
LLNA: BrdU-FC	LLNA with BrdU detected by FC
LLNA: DA	LLNA: Daicel Adenosine Triphosphate
LNC	Lymph node cells
LOEL	Lowest observed effect level
MEK	Methyl ethyl ketone
NICEATM	NTP Interagency Center for the Evaluation of Alternative Toxicological Methods

NIEHS	National Institute of Environmental Health Sciences
NIH	National Institutes of Health
NOEL	No observed effect level
NTP	National Toxicology Program
OD	Optical density
OECD	Organisation for Economic Co-operation and Development
QSAR	Quantitative structure-activity relationship
REACH	Registration, Evaluation, and Authorisation of Chemicals
rLLNA	Reduced LLNA
SAR	Structure-activity relationship
SD	Standard deviation
SI	Stimulation index
SDS	Sodium dodecyl sulfate
SLS	Sodium lauryl sulfate
TG	Test Guideline
Th	T-helper
vs.	Versus

Members of the Independent Scientific Peer Review Panel

Michael Luster, Ph.D. (Panel Chair), Senior Consultant to the National Institute of Occupational Safety and Health (NIOSH) Health Effects Laboratory, Morgantown, WV

Nathalie Alépée, Ph.D., Associate Research Fellow, Pfizer PDRD MCT Laboratory, France

Anne Marie Api, Ph.D., Vice President, Human Health Sciences, Research Institute for Fragrance Materials, Woodcliff Lake, NJ

Nancy Flournoy, M.S., Ph.D. (Biostatistician), Professor and Chair, Dept. of Mathematics & Statistics, University of Missouri-Columbia, Columbia, MO

Thomas Gebel, Ph.D., Regulatory Toxicologist, Federal Institute for Occupational Safety & Health, Dortmund, Germany

Sidney Green, Ph.D.,[1] Graduate Professor, Howard University, Washington, DC

Kim Headrick, B. Admin., B.Sc., International Harmonization Senior Policy Advisor, Health Canada, Ottawa, Ontario, Canada

Dagmar Jírová, M.D., Ph.D., Toxicologist, Research Manager, Head of Reference Center for Cosmetics, Head of Reference Laboratory for Experimental Immunotoxicology, National Institute of Public Health, Czech Republic

David Lovell, Ph.D. (Biostatistician), Reader in Medical Statistics, Postgraduate Medical School, University of Surrey, Guildford, Surrey, U.K.

Howard Maibach, M.D., Professor, Department of Dermatology, University of California-San Francisco, San Francisco, CA

James McDougal, Ph.D., Professor and Director of Toxicology Research, Department of Pharmacology and Toxicology, Boonshoft School of Medicine, Wright State University, Dayton, OH

Michael Olson, Ph.D., Director of Occupational Toxicology, Corporate Environment Health and Safety, GlaxoSmithKline, Research Triangle Park, NC

Raymond Pieters, Ph.D., Associate Professor, Immunotoxicology Group Leader, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands

Jean Regal, Ph.D., Professor, Department of Pharmacology, University of Minnesota Medical School, Duluth, MN

Jonathan Richmond, MB ChB, FRCSEd,[1] Head of Division, Animals Scientific Procedures Division, Home Office, London, U.K.

Peter Theran, V.M.D., Massachusetts Society for the Prevention of Cruelty to Animals, Novato, CA

---

[1] Drs. Green and Richmond were unable to attend the public meeting on March 4-6, 2008. However, they were involved in the review of the background review documents and concur with the conclusions and recommendations included in this report.

Stephen Ullrich, Ph.D., Dallas/Ft. Worth Living Legends Professor & Professor of Immunology, Post-graduate School of Biomedical Science, University of Texas M.D. Anderson Cancer Center, Houston, TX

Michael Woolhiser, Ph.D., Technical Leader - Immunotoxicology, Toxicology & Environmental Research & Consulting Immunology, Dow Chemical, Midland, MI

Takahiko Yoshida, M.D., Ph.D., Professor, Department of Health Science, Asahikawa Medical College, Hokkaido, Japan

[This Page Intentionally Left Blank]

Preface

In 1999, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) recommended the murine local lymph node assay (LLNA) to U.S. Federal agencies as a valid substitute for currently accepted guinea pig test methods to assess the allergic contact dermatitis potential of many, but not all, types of substances. The recommendation was based on a comprehensive evaluation of the validation status of the LLNA that included an assessment by an international independent scientific peer review panel (hereafter, Panel). The Panel report and the ICCVAM LLNA test method recommendations (ICCVAM 1999) are available at the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)-ICCVAM website.[1] The LLNA was subsequently incorporated into national and international test guidelines for the assessment of skin sensitization (OECD 2002; ISO 2002; EPA 2003). For this Panel report, this LLNA will be referred to as the "traditional" LLNA.

On January 10, 2007, the U.S. Consumer Product Safety Commission (CPSC) formally requested through NICEATM that ICCVAM assess the validation status of:[2]

- The traditional LLNA as a stand-alone assay for potency determinations (including severity) for the purpose of hazard classification
- Three modifications of the traditional LLNA not requiring the use of radioactive materials
- The LLNA limit dose procedure (also referred to as the "reduced" LLNA)
- The ability of the traditional LLNA to test mixtures, metals, and aqueous solutions (i.e., to re-evaluate the applicability domain for the traditional LLNA)

NICEATM, in coordination with ICCVAM and the ICCVAM Immunotoxicity Working Group, prepared a comprehensive draft background review document (BRD) for each modified version of the traditional LLNA test method being evaluated, as well as a draft applicability domain addendum to the final BRD published previously on the traditional LLNA. Each draft BRD and the draft addendum detailed the available data and information from the published literature and submissions received in response to a 2007 *Federal Register (FR)* notice that had requested data related to CPSC's nomination (*FR* notice Vol. 72, No. 95, p. 27815-27817, May 17, 2007). In addition, ICCVAM developed draft LLNA Performance Standards intended for use in validating alternative test methods that are functionally and mechanistically similar to the traditional LLNA. Finally, ICCVAM, based on the information contained in each of the draft BRDs and the draft addendum, developed draft test method recommendations.

The various supporting documents and the draft ICCVAM recommendations were provided to a new international Panel for an independent scientific review. In addition, NICEATM announced the availability of these documents on the NICEATM-ICCVAM website

---

[1] The 1999 ICCVAM Panel report and recommendations can be obtained at: http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf

[2] The CPSC nomination can be obtained at: http://iccvam.niehs.nih.gov/methods/immunotox/llnadocs/CPSC_LLNA_nom.pdf

(http://iccvam.niehs.gov) for public comment in a *FR* notice (Vol. 73, No. 5, p. 1360-1362, January 8, 2008) and via the ICCVAM listserv. The *FR* notice also announced the public Panel meeting, to be convened at the CPSC Headquarters in Bethesda, MD on March 4–6, 2008.

The Panel was charged with:

- Reviewing each ICCVAM draft BRD and the draft addendum for completeness and identifying any errors or omissions of existing relevant data or information
- Evaluating the information in each draft BRD and the draft addendum to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003) had been appropriately addressed for the recommended use of the new versions and applications of the traditional LLNA
- Considering the ICCVAM draft test method recommendations for the following and commenting on the extent to which they are supported by the information provided in the draft BRDs and the draft addendum:
  - proposed test method uses
  - proposed recommended standardized protocols
  - proposed test method performance standards
  - proposed additional studies
- Evaluating the draft ICCVAM LLNA Performance Standards and considering whether they were adequate for assessing the accuracy and reliability of alternative test methods that are functionally and mechanistically similar to the traditional LLNA

During our public meeting in March 2008, the Panel discussed each charge, listened to public comments, and developed conclusions and recommendations for ICCVAM on each of the nominated activities. The Panel wished to emphasize that they were to consider two overall questions. They were to consider: (1) whether the validation status of the each of the above proposed modifications or alternative uses of the LLNA had been adequately characterized for its intended purpose according to established ICCVAM validation criteria (available on the NICEATM-ICCVAM website, http://iccvam.niehs.gov), and (2) whether proposed modifications or alternative uses of the LLNA are sufficiently accurate and reliable to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure.

This report details the Panel's independent conclusions and recommendations. ICCVAM will consider this report, along with all relevant public comments, as it develops final test method recommendations. The final ICCVAM test method recommendations will be forwarded to U.S. Federal agencies for their consideration in accordance with the ICCVAM Authorization Act of 2000 (Public Law 106-545).

The Panel gratefully acknowledges the efforts of NICEATM staff in coordinating the logistics of the peer review Panel meeting and in preparing materials for their review. The

Panel also thanks each of the test method developers, Drs. George DeGeorge (LLNA: BrdU-FC), Kenji Idehara (LLNA: DA), and Masahiro Takeyoshi, (LLNA: BrdU-ELISA) for providing summaries and additional clarifications of the non-radioactive test methods under review. Finally, as Panel Chair, I want to thank each Panel member for her or his thoughtful and objective review of these LLNA-related activities.

Michael Luster, Ph.D.
Chair, LLNA Peer Review Panel
May 2008

[This Page Intentionally Left Blank]

Executive Summary

This report describes the conclusions and recommendations of an international independent scientific peer review panel (hereafter, Panel). This Panel was charged by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) with evaluating the validation status of new versions and applications of the murine local lymph node assay (LLNA) for assessing the allergic contact dermatitis (ACD) potential of chemicals and products. The LLNA, which was first evaluated in 1999 by ICCVAM, is hereafter referred to as the "traditional LLNA" to distinguish it from other versions considered by the Panel. The new versions and applications considered include:

- The LLNA limit dose procedure (also referred to as the "reduced" LLNA[1])
- The ability of the traditional LLNA to test mixtures, metals, and aqueous solutions (i.e., a re-evaluation of the applicability domain for the traditional LLNA)
- Three modifications of the traditional LLNA not requiring the use of radioactive materials:
  - LLNA: DA (Local Lymph Node Assay: Daicel Adenosine Triphosphate)
  - LLNA: BrdU-FC (Local Lymph Node Assay: Bromodeoxyuridine detected by flow cytometry)
  - LLNA: BrdU-ELISA (Local Lymph Node Assay: Bromodeoxyuridine detected by ELISA)
- The traditional LLNA as a stand-alone assay for potency determinations (including severity) for the purpose of hazard classification

The Panel also evaluated the draft ICCVAM LLNA Performance Standards and considered whether they were adequate for assessing the accuracy and reliability of alternative test methods that are functionally and mechanistically similar to the traditional LLNA.

LLNA Limit Dose Procedure

The Panel agreed that the LLNA limit dose procedure, which normally allows for testing at one dose level, should be routinely recommended for hazard identification when used for testing purposes which do not require dose response information, because it would offer time, cost, throughput and logistical benefits as well as using fewer animals. In instances when a necessity to measure relative skin sensitization potency for the purpose of risk assessment was present, then the traditional LLNA should be used in order to generate dose response information. Still, the Panel recommended use of the LLNA limit dose procedure as

[1] As described in this report, the Panel agreed that consideration should be given to applying the same term to the LLNA limit dose procedure since in various places throughout the draft BRD it was referred to differently as either the "cut-down", the "limit dose", or the "reduced LLNA" (i.e., "rLLNA"). Since the European Centre for the Validation of Alternative Methods (ECVAM) has already established a naming convention of "rLLNA", the Panel recommended adopting the ECVAM terminology to harmonize the terminology used among the international validation agencies. However, because the ICCVAM documents that were reviewed use "LLNA limit dose procedure" that term is retained in this report.

the initial testing procedure to identify sensitizers and non-sensitizers before conducting the traditional LLNA even when dose response information *is* required since if the test substance were negative in the limit dose procedure, it would not be necessary to conduct a multiple-dose LLNA test.

The draft background review document (BRD) for the LLNA limit dose procedure provides a comprehensive review of available data and information for assessing the usefulness and limitations of this modified version of the LLNA for the purpose of skin sensitization hazard classification. The Panel evaluated the draft BRD for completeness, errors, and omissions and recommended that it be updated to reflect their suggestions/corrections relating to general, statistical, and specific editorial issues. In particular, the Panel noted that the differences in terminology used for this procedure caused confusion and recommended that an internationally harmonized term be adopted. They suggested referring to the procedure as the "reduced LLNA" (i.e. "rLLNA") since that is being used by the European Centre for the Validation of Alternative Methods (ECVAM).

The Panel concluded that the stimulation index (SI) based on the ratio of 3.0 as the cutoff value was indicative of a response that was sufficiently greater than the control and would be considered an immunologically relevant response, but recommended that statistical analyses be used to definitively establish that a response induced by a test substance is significantly different from the vehicle control. The Panel agreed that the LLNA protocol recommended by ICCVAM (ICCVAM 1999; Dean et al. 2001) should be the standard protocol for all future LLNA limit dose studies using the traditional LLNA protocol. Specifically, prospective LLNA limit dose procedure studies should require that lymph nodes be collected from individual animals instead of pooling them with other animals in a treatment group, which is also currently permitted by the Organisation for Economic Co-operation and Development (OECD) Test Guideline 429 (OECD 2002). Individual animal response data are necessary in order to statistically analyze for differences between treated and control data. In addition, having data from individual animals also allows for identification of technical problems and outlier animals within a dose group. Based on power calculations provided as supplemental information, the Panel agreed that five animals per dose group is an appropriate number to recommend for LLNA limit dose studies following the traditional LLNA protocol. It should be noted that the Panel's analysis of the LLNA limit dose dataset was not restricted to studies with confirmed individual animal data, and that the Panel considered data known to have been generated using pooled group data. The Panel stated that, internationally, both individual and pooled animal data have likely been used both for regulatory decisions and for in-house decisions relating to product development and risk management. In addition, the fact that the retrospective data analysis set out in the draft LLNA limit dose procedure BRD did not distinguish between individual or pooled animal data suggested that both met the quality standards for inclusion in the draft BRD.

Although they did not reach consensus, the Panel suggested that for laboratories in which the LLNA is "routinely" performed and have demonstrated the ability to consistently obtain positive results, hexyl cinnamic aldehyde (HCA) or another positive control (e.g., a substance that matches the chemical class of the test substances) could be run at intervals for quality control purposes rather than concurrent with each experiment. The Panel cited Kimber et al. (2006), which describes "routine" use of the "rLLNA" utilizing only a vehicle and a high-dose group, as a rationale for this suggestion. However, the Panel does not recommend

omitting the concurrent positive control in laboratories that perform the LLNA only "occasionally".

Based on the analyses presented in the draft BRD, the Panel considered the accuracy of the LLNA limit dose procedure to have been adequately evaluated and compared to the traditional LLNA, mindful of the limitations associated with a retrospective evaluation. For instance, it cannot be assumed that the compounds tested in the retrospective studies were always tested at the highest possible dose unless such information was explicitly indicated. In this regard, the Panel recommended that a more detailed description of what is considered "*avoidance of excessive irritation*" and "*evidence of systemic toxicity*" be included in any LLNA protocol in order to aid in choosing the most appropriate high (i.e., limit) dose, although specific indicators of "*systemic toxicity or excessive irritation*" were not formally discussed.

The Panel agreed that it was appropriate to assume that the intra- and inter-laboratory reproducibility of the LLNA limit dose procedure and the traditional LLNA would be similar, because reproducibility is more dependent on the method than on the number of dose groups. However, reducing the number of test substances dose groups from three to one might reduce the sensitivity of the assay. The traditional LLNA may have a greater chance of correctly identifying a sensitizer even in the presence of one or more technical errors since data from three dose groups are being considered and an SI $\geq$3.0 at any dose group would result in the substance being classified as a sensitizer. However, for the purpose of adopting an assay that uses fewer animals and provides increased throughput for testing purposes, these hypothetical considerations are not a sufficient reason to argue against use of the limit dose LLNA procedure.

LLNA for Testing Aqueous Solutions, Metals, and Mixtures

The draft ICCVAM recommendations state that, although more data are needed to assess the use of the LLNA for testing for mixtures and aqueous solutions before a recommendation can be made, the traditional LLNA appears to be useful for the testing of metal compounds, with the exception of nickel. The Panel agreed with these draft ICCVAM recommendations. Regarding the use of the LLNA for testing mixtures, the Panel acknowledged that the ability of ICCVAM to develop draft test method recommendations was limited not only by the amount of data available, but the relatively poor concordance of traditional LLNA outcomes in comparison to those obtained in guinea pig tests, and recommended that this be noted in the final ICCVAM recommendations. The term "mixtures" can represent an infinite number of materials and it would be more beneficial to specify types or formulations of mixtures that are being examined.

Regarding metals, the Panel concluded that the accuracy statistics for the traditional LLNA when compared to results obtained from evaluation in humans supported use of the traditional LLNA as a hazard identification tool for metals, with the exception of nickel, which produces variable responses. One minority opinion stated that the results for nickel compounds were not entirely questionable and that the traditional LLNA might also be suitable for testing nickel compounds. Thus, the Panel recommended further evaluation of the variable results obtained for nickel in the context of the available literature on allergic contact dermatitis to nickel in humans.

Regarding substances tested in aqueous solutions, the Panel suggested expanding the brief section of the draft test method recommendations discussing the test method protocol for the traditional LLNA to specifically point out how the conclusions of the applicability domain evaluation may affect the standard traditional LLNA protocol. For instance, it could be suggested that aqueous test solutions be avoided due to problems associated with skin application. It would be preferable for a hierarchy of organic solvents to be considered as dosing vehicles, with emphasis on using a vehicle to which humans may actually be exposed in circumstances linked to occupational sensitization.

The Panel agreed with the draft ICCVAM recommendation for continued accrual of information from traditional LLNA evaluations of mixtures, metals, and aqueous solutions with comparative data for guinea pig (i.e., guinea pig maximization test [GPMT] or Buehler test [BT]) and human (i.e., human maximization test [HMT] or human repeat insult patch test [HRIPT]) tests. However, the Panel suggested that, given resource limitations, it would be important to organize the recommendations based on relative priority.

The draft Addendum to the original validation report for the traditional LLNA (ICCVAM 1999) provided a comprehensive review of currently available data and information for evaluating the usefulness and limitations of the traditional LLNA for assessing the skin sensitization potential of mixtures, metal compounds, and substances tested in aqueous solutions. The Panel evaluated the draft Addendum for completeness, errors, and omissions and concluded that there were no apparent errors or omissions, although they did state that the term "mixtures" was used too broadly (i.e., can represent an infinite number of materials) and it would be more beneficial to specify types or formulations of mixtures that are being examined.

The Panel did not identify any classes of chemicals missing from the dataset used to review the utility of the traditional LLNA for testing aqueous solutions. However, while they did not propose an alternative, the Panel expressed concern over the most appropriate definition for an aqueous solution (defined in the draft Addendum as any solution containing ≥20% water). For the mixtures included in the analysis, the Panel noted that quantitative compositions had not been provided and therefore they could not comment on whether these mixtures were representative of the types of mixtures typically tested in the traditional LLNA. With respect to metals (none of which are mixtures), there was a paucity of important representatives of commercially useful metals such as platinum, palladium, iron, zinc, manganese and silver in the data set. The Panel suggested that to enlarge the group of metal non-sensitizers, substances used as cosmetic ingredients (e.g., titanium dioxide) and aluminum compounds currently used in antiperspirants might be considered.

The Panel agreed that, although it was important to identify data obtained according to GLP guidelines, data obtained from non-GLP studies should not be excluded automatically from this retrospective analysis. The Panel concluded that other factors could be used to identify high quality data. Examples include data published in peer-reviewed journals or obtained from a study conducted in a laboratory that has GLP capabilities.

The Panel concluded that, considering the limited comparative data that were available, particularly for mixtures and aqueous solutions, the accuracy assessment of the traditional LLNA for testing mixtures, metals, and aqueous solutions when compared to available human and/or guinea pig test results was as comprehensive as was possible. The limited

amount of comparative data made it unfeasible to draw definitive conclusions for mixtures and aqueous solutions from the available accuracy statistics.

Non-Radioactive LLNA Protocol - The LLNA: DA Test Method

The Panel concluded that the available data and test method performance support the ICCVAM draft recommendations for the LLNA: Daicel Adenosine Triphosphate test method (LLNA: DA), and that the test method may be useful for identifying substances as potential skin sensitizers and non-sensitizers, but that this recommendation is contingent upon receipt, review, and analyses of additional existing data and information from the test method developer. Therefore, this non-radioactive version of the traditional LLNA cannot currently be recommended for the hazard identification of skin sensitizing substances, regardless of whether or not there are restrictions on the use of radioactive materials, until such time as this existing data has been received and confirmed.

The draft LLNA: DA BRD was compiled to provide a comprehensive review of available data and information evaluating the usefulness and limitations of the LLNA: DA test method to assess the allergic contact dermatitis potential of chemicals and other substances. The Panel evaluated the draft BRD for completeness, errors, and omissions and recommended that their suggestions/corrections relating to general, statistical, and specific editorial issues be incorporated into future revisions.

The Panel agreed that five animals per dose group should be recommended for validation of modified LLNA test methods. The Panel, however, noted that supplemental power calculations for the LLNA: DA test method indicated that the power for detecting a three-fold increase in the treatment group was estimated to be 95% for a sample size of three mice per dose group. Thus, the Panel identified the use of three animals per dose group as a potential opportunity to reduce animal number when using modified assays in the future, assuming all essential validation requirements can be successfully met. A minority opinion expressed by five Panel members was that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from a least four animals per dose group could be considered.

Generally, the Panel viewed the difference in treatment schedule between the LLNA: DA and the traditional LLNA to be potentially significant if the LLNA: DA induced the elicitation phase of skin sensitization. The Panel was concerned that the 1% sodium lauryl sulfate (SLS) pretreatment step in the LLNA: DA might modify the inherent sensitivity of the LLNA. They recommended that the test method developer (Daicel Chemical Industries, Ltd.) justify the use of 1% SLS or consider an alternative decision criterion (i.e., an SI threshold other than 3.0) such that the 1% SLS pretreatment is no longer necessary.

The Panel considered the database of substances tested in the LLNA: DA to be representative of a sufficient range of chemicals expected to be tested for skin sensitization potential, and concluded that the accuracy analysis had made appropriate comparisons to the traditional LLNA, guinea pig tests, and human data/experience. The Panel could not identify specific characteristics associated with the one false negative (i.e., 2-mercaptobenzothiazole) or the one false positive (i.e., benzalkonium chloride), but reemphasized that the potential impact of pretreatment with 1% SLS in this context needed to be considered.

With regard to test method reliability, the Panel concluded the intralaboratory reproducibility of the LLNA: DA had not been adequately evaluated. They noted that the two sensitizers tested had similar chemical structures (i.e., eugenol and isoeugenol) and that it was unclear if the tests were truly independent. The Panel also noted that the interlaboratory reproducibility of the assay could not be adequately evaluated given the lack of original laboratory data and limitations in the study design. In particular, they cited the use of pooled lymph nodes from the mice in each treatment group and the testing of each substance at predetermined dose levels established by the lead laboratory as study design limitations. Still, a Panel minority considered pooled data acceptable and the setting of dose levels for all laboratories based on results from the lead laboratory to be reasonable.

The Panel also commented that ideally, test substances should be coded during the validation of a new assay, although they did not feel that a lack of coding constituted a reason for rejecting the current LLNA: DA dataset. The Panel also commented that although GLP compliance is highly recommended for validation studies, the current studies should not be rejected solely on the basis of a lack of GLP compliance. However, the Panel considered it important to obtain the original records for all validation studies (which have been requested by NICEATM) in order to confirm that the reported data were the same as the data recorded in the laboratory notebooks.

With regard to the 5% (1/19) false negative and 10% (1/10) false positive rates obtained with the LLNA: DA, the Panel commented that it was important to identify reasons why the substances gave "false" results, taking into consideration factors such as intended use of the substances and the target population. They agreed that it might be useful to follow a suspected inaccuracy with an investigation of the mechanistic basis for the discordance since it may help to establish a biologically-based rationale for the discordance.

The Panel noted that the available LLNA: DA data did not support all of the ICCVAM draft recommendations in the proposed test method standardized LLNA: DA protocol. First, although the Panel agreed with the ICCVAM protocol that recommends five animals per dose group, they noted that supplemental statistical information provided for the LLNA: DA test method implied that using less than five animals per dose group was acceptable (e.g., a 3.0-fold increase in the SI value would likely be detected with 99% confidence when using four animals per dose group). In addition, the Panel considered it important to adequately characterize the effect of the 1% SLS pretreatment step in the LLNA: DA, and it should be demonstrated that the day 8 applications do not induce a skin reaction that could be indicative of the onset of the elicitation phase of skin sensitization. Keeping these points in mind, the Panel agreed that if the limit dose procedure was applicable to the traditional LLNA, then it would also be applicable to the LLNA: DA in order to further reduce the number of animals used.

The Panel also stated that the available data supported the ICCVAM draft recommendations for the LLNA: DA in terms of future studies, which included performing a more comprehensive evaluation using more non-sensitizers within and across laboratories. A minority opinion stated by one Panel member was that although testing more sensitizers might be warranted for interlaboratory validation studies, a sufficient number of non-sensitizers (n=11) had already been tested within the same laboratory.

The Panel also commented that the protocol differences between the LLNA: DA and the traditional LLNA could not clearly be constituted as "major" or "minor" changes. However, they considered this issue largely irrelevant if a test method was able to correctly predict the dermal sensitization potential of a test substance. Consequently, the Panel concluded that the current draft ICCVAM Performance Standards could be applicable to the LLNA: DA as a mechanistically and functionally similar test method.

Non-Radioactive LLNA Protocol - The LLNA: BrdU-FC Test Method

Overall, the Panel concluded that the available data and test method performance of the LLNA with bromodeoxyuridine (BrdU) detected by flow cytometry (LLNA: BrdU-FC) supported the draft ICCVAM recommendations that it may be useful for identifying substances as potential skin sensitizers and non-sensitizers, but that more information and existing data must be made available before the LLNA: BrdU-FC can be recommended for routine use. The Panel concluded that the test method usefulness and limitations identified in the draft ICCVAM recommendations accurately summarized the limits of the information supplied and the additional information that would need to be generated or provided for further consideration of the test method. As a result, the Panel concluded that the LLNA: BrdU-FC could not currently be considered as a scientifically valid replacement alternative to the traditional LLNA. Still, the Panel suggested that the test method recommendation should clearly state that the test method was not "invalid", but simply that there was currently not sufficient evidence and information to state that it had been adequately validated.

The draft LLNA: BrdU-FC BRD was compiled to provide a comprehensive review of available data and information evaluating the usefulness and limitations of the LLNA: BrdU-FC test method to assess the ACD-inducing potential of substances. The Panel evaluated the draft BRD for completeness, errors, and omissions and recommended that their recommendations/corrections relating to general, statistical, and specific editorial issues be incorporated into future revisions.

The LLNA: BrdU-FC included routine measurements of ear swelling as an indicator of excessive skin irritation. The Panel viewed that this, or any other quantitative measurement of skin irritation, should be carefully considered for inclusion in all LLNA protocols. The Panel considered inclusion of optional quantification of immunophenotypic markers as an additional mechanism for distinguishing irritants from sensitizers to be useful, as it might reduce the frequency of false positives (i.e., substances which are actually skin irritants) and improve comparisons with human data. However, they considered application of immunological markers too detailed and costly for routine LLNA use (i.e., for hazard classification purposes) and more suited for research purposes.

The Panel noted that the substances tested in the LLNA: BrdU-FC seemed representative of a sufficient range of chemical classes and physical chemical properties, and thus that the test method appeared applicable to many of the types of chemicals and products that are typically tested for skin sensitization potential. However, the Panel considered the total database available for evaluation of the validation status of the LLNA: BrdU-FC to be relatively small compared to the large number of substances assessed in the traditional LLNA. Therefore, the Panel recommended caution when making conclusions related to its concordance with the

traditional LLNA. Still, the accuracy of the LLNA: BrdU-FC was considered adequately evaluated and comparable to the traditional LLNA.

The Panel concluded that intralaboratory reproducibility was not adequately assessed and it should be better evaluated in order to support the validation of this test method. The Panel suggested that although the studies evaluated in the draft BRD were not GLP-compliant, this should not affect acceptance of the data for an evaluation of the validation status of this test method. However, some sources of variability in the intralaboratory data, such as failure to appreciate differences in composition of dosing solutions between experiments caused by test article instability or other phenomena, might be obscured if not in complete compliance with GLP guidelines. Thus, the Panel suggested that any additional studies undertaken to validate the test method should ideally be GLP-compliant.

The Panel agreed that the available data supported the ICCVAM draft recommendations for the LLNA: BrdU-FC procedure in terms of the proposed test method standardized protocol. They suggested that the utility of ear swelling or other methods to detect inflammation appeared warranted in every variation of the LLNA (including the traditional LLNA), but should be further investigated before routine inclusion in the protocol is recommended. The Panel also concluded that the traditional LLNA limit dose procedure could be applied to the LLNA: BrdU-FC, keeping in mind the limitations associated with a "limit dose" procedure.

The Panel further agreed that the ICCVAM draft recommendations for future studies highlighted the unanswered questions raised by the available data set. Specifically, conducting interlaboratory studies as a part of the validation process is important. The Panel considered the immunological markers suggested for the LLNA: BrdU-FC to be acceptable, but that additional immunological markers for discrimination of irritant versus sensitization phenomena were also possible. In general, for any future work, efforts should be made to decrease the variability and to thereby increase the power of the test in order to ensure that more animals were not needed relative to the traditional LLNA or other alternative LLNA protocols.

The Panel considered the protocol differences between the LLNA: BrdU-FC and the traditional LLNA to be "minor" changes, and therefore concluded that assessment of the validity of this test method could be based on the draft ICCVAM LLNA Performance Standards. The Panel also cautioned, however, that a clear definition of what constituted a "major" versus a "minor" change, or a different protocol altogether could be better addressed once the recommendations for the current draft ICCVAM LLNA Performance Standards were finalized.

Non-Radioactive LLNA Protocol - The LLNA: BrdU-ELISA Test Method

The Panel concluded that the available data and test method performance for the LLNA with BrdU detected by enzyme-linked immunosorbent assay (LLNA: BrdU-ELISA) support the ICCVAM draft recommendations that it may be useful for identifying substances as potential skin sensitizers and non-sensitizers, but that more information and existing data must be made available before the LLNA: BrdU-ELISA can be recommended for use. The Panel also stated that a detailed protocol was needed, in addition to sufficient quantitative data for broader analysis on a larger set of balanced reference substances that take into account

physicochemical properties and sensitization potency, as well as an appropriate evaluation of interlaboratory reproducibility.

The draft LLNA: BrdU-ELISA BRD was compiled to provide a comprehensive review of available data and information evaluating the usefulness and limitations of the LLNA: BrdU-ELISA test method to assess the ACD-inducing potential of chemicals and other substances. The Panel evaluated the draft BRD for completeness, errors, and omissions and recommended that their suggestions/corrections relating to general, statistical and specific editorial issues be incorporated into the final document.

The Panel's main concern with the test method was that the accuracy of the LLNA: BrdU-ELISA at SI ≥3.0 was inadequate and not equivalent to the traditional LLNA. Furthermore, although using a decision criterion of SI ≥1.3 improved the test's performance in identifying sensitizers from non-sensitizers, it did not resolve concerns about the test method. Based on a power analysis for the LLNA: BrdU-ELISA, which was provided to the Panel as supplemental information, the Panel concluded that it was difficult to justify using a SI ≥1.3 as the cutoff value, given the much larger number of animals that would be required to detect a 1.3-fold increase above vehicle controls with similar power to the traditional LLNA when five animals per dose group are used. For a three-fold increase, the supplemental statistical analyses indicated that a sample size of four was sufficient. Still, the Panel agreed with the ICCVAM recommendation to use five animals per dose group and to collect individual animal data. They concluded that this would allow for more robust calculations in the event that an outlier prevented some of the data from being included in the analysis. A minority opinion by five Panel members was stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from a least four animals could be considered.

The Panel noted that in organizations where the use or disposal of radioactive materials was restricted, the potential to use the LLNA: BrdU-ELISA could reduce the number of animals needed per test compared to the traditional LLNA and would result in less pain and suffering compared to using traditional guinea pig test methods. However, if the SI ≥1.3 was chosen as the decision criterion because of its improved accuracy compared to SI ≥3.0, the Panel stated that the number of mice needed to perform the LLNA: BrdU-ELISA test should be compared to the number of guinea pigs that would be needed for skin sensitization tests in order to assess if the LLNA: BrdU-ELISA actually reduced overall animal use for skin sensitization testing.

In general, the Panel considered the number of substances tested in the LLNA: BrdU-ELISA too few, and that data from more substances tested using the traditional LLNA, guinea pig tests, and human tests should have been included. The Panel also did not consider the available data from the LLNA: BrdU-ELISA to be representative of a sufficient range of chemical classes and physical chemical properties. The limited dataset prevents an evaluation of whether the test method would be considered applicable to any of the types of chemicals and products typically tested for skin sensitization potential.

However, the Panel concluded that the appropriate comparisons between the traditional LLNA, guinea pig test and human data had been made. The Panel agreed that the false negative rate for hazard identification using the SI ≥3.0 in the LLNA: BrdU-ELISA was excessive (i.e., using this SI threshold value, the LLNA: BrdU-ELISA misclassified 29% and

39% of the substances classified as sensitizers in the traditional LLNA or in humans, respectively).

The Panel also considered that the intralaboratory reproducibility of the LLNA: BrdU-ELISA was not adequately evaluated and compared to the traditional LLNA. The Panel indicated that the number of substances was too few, and in some cases there was a wide variation in repeat tests of the same substance. The Panel recommended a more comprehensive evaluation of the intralaboratory reproducibility of the test method, using different SI values, and that the analysis of the variability of the estimated concentration needed to produce a positive SI value (ECt values) be conducted on a log scale.

The Panel also noted that interlaboratory reproducibility for the LLNA: BrdU-ELISA could not be evaluated because neither the design of the study sponsored by the Japanese Center for Validation of Alternative Methods nor any of the resulting data had been provided in advance of their evaluation. The Panel agreed that a multi-laboratory validation study using a balanced set of chemicals would adequately characterize the interlaboratory reproducibility of the LLNA: BrdU-ELISA.

In general, the Panel agreed that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA procedure in terms of the proposed test method standardized protocols. However, as noted above, a minority opinion by five Panel members was that there could be circumstances in which pooled data from at least four animals could also be acceptable. The Panel also stated that if the LLNA: BrdU-ELISA was found to be equivalent to the traditional LLNA in the future that it would be appropriate to apply the LLNA limit dose procedure to the test. The Panel also agreed with ICCVAM's test method recommendations for future studies and emphasized that more data were needed in order to determine the appropriate threshold value for the decision criterion. The Panel concluded that it might be more appropriate to use a statistically-based decision criterion rather than a stimulation index to classify substances as sensitizers, and that this should be further investigated.

The Panel agreed that the LLNA: BrdU-ELISA protocol differed from the traditional LLNA only in the method used to assess lymphocyte proliferation and as such concluded that this represented a "minor" change (as defined in the current draft ICCVAM LLNA Performance Standards) and separate performance standards for the LLNA: BrdU-ELISA were not needed.

Draft ICCVAM LLNA Performance Standards

The draft ICCVAM LLNA Performance Standards are intended to evaluate the acceptability of proposed test methods that are mechanistically and functionally similar to the traditional LLNA. ICCVAM proposed that the applicability of the draft ICCVAM LLNA Performance Standards be restricted to protocols that incorporate "minor" modifications to the traditional LLNA procedure, defined as changes only to the method for measuring lymphocyte proliferation. The Panel agreed that different methods of measuring lymphocyte proliferation represent "minor" modifications, but recommended that, instead of trying to define "minor" modifications, a better strategy might be to define criteria that would need to be satisfied in order to ensure that the alternative test method was mechanistically and functionally similar to the traditional LLNA (e.g., only measure cell proliferation associated with the induction

phase of a skin sensitization reaction). The Panel considered that the draft performance standards were also appropriate for evaluating other modifications. Examples of acceptable modifications included test animal sex, strain, the use of rats rather than mice, the number of animals per group, and timing of test article treatment. One minority opinion considered the potential impact of changes to protocol components other than the method of measuring lymphocyte proliferation to be significant and therefore would require more extensive validation, which was not defined.

The Panel indicated that alternative LLNA protocols that are undergoing validation should contain essential test method components that follow the ICCVAM-recommended protocol (ICCVAM 1999; Dean et al. 2001), unless adequate scientific rationale for deviating from this protocol was provided.

The Panel also identified aspects of the LLNA that should be required as part of the test method validation process: (1) application of the test substance to the skin with sampling of the lymph nodes draining that site, (2) measurement of cell proliferation in the draining lymph node, (3) absence of a skin reaction that could be indicative of the onset of the elicitation phase of skin sensitization, (4) data collected at the level of the individual animal to allow for an estimate of the variance within control and treatment groups (using this variance, a power analysis needs to be conducted to demonstrate that the modified method is utilizing a sufficient number of animals per treatment group to permit hazard identification with at least 95% power), and (5) if dose response information is needed, there are an adequate number of dose groups ($n \geq 3$) with which to accurately characterize the dose response for a given test substance.

The Panel noted that the list of substances included in the draft ICCVAM LLNA Performance Standards was sufficiently representative of the types of materials that are likely to be tested for skin sensitization. However, among the 13 sensitizers in the list of "required" substances, only five were considered to have robust data (i.e., traditional LLNA data based on at least three independent studies).

To evaluate performance for use in hazard identification, the Panel concluded that all 22 substances in the draft ICCVAM-recommended list should be tested and accuracy statistics calculated (Note: this list of substances includes "required" substances as well as "optional" false negative and false positive substances, of which only 8/22 have "robust" datasets [$n \geq 3$ as defined by the Panel]). To the extent possible, a rationale for any discordant results should be provided. However, the most potent sensitizers (e.g., dinitrochlorobenzene [DNCB]) should always be identifiable. Also, considerable weight should be given to the balance between animal welfare and human safety when considering the adequacy of test method accuracy. Based on the limited data available for the sensitizers on the list and the lack of standardization of test methods from which the results were obtained, the current database does not support inclusion of ECt values as a component of the accuracy evaluation.

The Panel agreed with the draft ICCVAM recommendations for evaluating test method reliability. These recommendations included obtaining ECt values that are generally within 0.5x to 2.0x of the mean historical EC3 (i.e., estimated concentrations needed to produce an SI of 3) values for hexyl cinnamic aldehyde (HCA) (intralaboratory, n=4 experiments in one laboratory), or HCA and DNCB (interlaboratory, n=1 experiment in three laboratories). However, the Panel recommended that the criteria for independent tests should be specified

(e.g., different animal shipment, different reagents, different operator). The Panel concluded that the proposed criteria for acceptability appeared to be appropriate in this case, because only one or two substances were being evaluated (i.e., a statistical multiple comparisons[2] problem does not exist). The Panel also suggested that historical control data using HCA and DNCB in the same vehicle could be used to demonstrate adequate intra- and/or inter-laboratory reproducibility.

The Panel also recommended that statistical tests to analyze the data might allow for a more accurate interpretation. They recommended that a suitable variance-stabilizing transformation (e.g., log transformation, square root transformation) be applied in all statistical analyses and in reporting summary standard deviations. The Panel also recommended that a more rigorous evaluation be conducted of what would be considered an appropriate range of ECt values to include as a requirement. This would be a statistical evaluation that considers the variability of ECt values generated among the sensitizers included on the performance standards reference substances list and the statistical multiple comparisons problem.

Use of the LLNA for Potency Determinations

The Panel agreed with the draft ICCVAM recommendation that the LLNA should not be used as a stand-alone assay for categorizing skin sensitizers as strong vs. weak, but that it could be used as part of a weight-of-evidence evaluation (e.g., along with quantitative structure-activity relationships, peptide reactivity, human evidence, historical data from other experimental animal studies) for this purpose. The Panel also agreed with the draft ICCVAM recommendation that any LLNA studies conducted for the purpose of evaluating skin sensitization potency should use the ICCVAM-recommended LLNA protocol. In addition, the Panel viewed that the relevant testing guidelines for the traditional LLNA should be revised to include the procedure for calculating an EC3 value.

A draft BRD was compiled by ICCVAM that provided a comprehensive review of available data and information and an evaluation of the usefulness and limitations of the traditional LLNA for the categorization of substances with regard to skin sensitization potency. The Panel evaluated the draft BRD for completeness, errors, and omissions and noted alternative analyses that would allow for a more complete evaluation of the use of the traditional LLNA for skin sensitization potency categorizations (see below).

The Panel agreed that the database of substances evaluated for potency determinations was sufficient and represented a range of chemical classes and physicochemical properties applicable to products typically tested for skin sensitization potential. The Panel also concluded that since the database was compiled from existing data, the lack of substance coding likely had no impact on the retrospective evaluation presented in the draft BRD. Still, the Panel recommended the coding of test substances in any future validation studies. The

[2] When multiple experiments are conducted and multiple observations, comparisons, or hypothesis tests are conducted, the chance of observing rare events increases. Suppose, for example, that an interval is established such that 5% of observations from a particular population of data are outside that interval. Then if *k* independent experiments generate data from this population (e.g., a standard normal distribution), the chances that all 20 results will lie inside the interval is $(1.0 - 0.05)^k$ (N. Flournoy, personal communication).

Panel generally agreed that potency determinations based on traditional LLNA results should ideally be limited to data from studies that evaluated lymph node proliferation in individual animals so that outliers and technical errors could be identified. However, they also agreed that pooled animal data should not be excluded automatically from a retrospective analysis.

The Panel indicated that the relevance of the LLNA for potency determinations had been adequately compared and evaluated to human (i.e., HMT or HRIPT) and guinea pig (i.e., GPMT or BT) data. A minority opinion stated by one Panel member was that the relevance of the traditional LLNA to human clinical observations had not been sufficiently determined.

In general, the Panel agreed that the proposed two-level categorization scheme (weak vs. strong sensitizers) for both human and guinea pig data was appropriate. However, a minority opinion stated by two Panel members was that a moderate category should be included since certain compounds might be on the border between weak and strong sensitizers. Thus, they suggested that the five-category scheme proposed by Kimber et al. (2003), which includes non-sensitizers, might be recommended.

The Panel concluded that the decision criteria providing the best overall performance was the use of <250 $\mu g/cm^2$ to distinguish between strong and weak sensitizers in humans and the use of an LLNA EC3 $\leq 9.4\%$ to distinguish between strong and weak sensitizers in the LLNA. The Panel stated that more data would be needed to determine if values different from these two would be more appropriate. The Panel also recommended that safety factors other than 10 for the lowest observed effect level (LOEL) be evaluated to determine if improved results could be obtained. The Panel also suggested an analysis that directly compares the LOEL values without using a safety factor (i.e., using LOEL data only) and an analysis that only uses no observed effect level data. The Panel further stated that traditional LLNA tests based on pooled or individual lymph nodes for a dose group should be evaluated independently to assess the impact of using pooled data on the accuracy analysis for skin sensitization potency. Finally, the Panel stated that the effect of different vehicles should be recognized as a limitation in the current data analysis and a likely contributor to the variability observed within and across laboratories.

The Panel stated that data from studies that could not be confirmed as being GLP-compliant, but that were from peer-reviewed literature or sources with high-quality laboratory management practices, were still appropriate to include in the accuracy analysis. However, the Panel stated that, ideally, GLP compliance should be the standard, as it is clearly the only objective way to judge the credibility of the data.

The Panel recommended that more data should be collected to determine the optimal threshold in humans for distinguishing between strong and weak sensitizers. In addition, the Panel discouraged conducting additional animal studies unless such studies would be expected to lead to an overall reduction in animal use. The Panel recommended that the LOELs from Akkan et al. (2003) be used instead of the DSA_{05} (i.e., the dose per skin area leading to a sensitization incidence of 5%) values from Schneider and Akkan (2004) in all of the potency analyses. A minority opinion by one Panel member stated that it was acceptable to use the DSA_{05} values from Akkan et al. (2003) as LOEL values in the evaluation. This panelist mentioned that the DSA_{05} value is a LOEL value adjusted to 5% incidence of induction in order to correct for human studies leading to different inductions. Furthermore, the panelist stated that because the DSA_{05} is corrected for an induction rate of 5%, it would

be better to compare with the traditional LLNA EC3 than to use the default uncorrected LOEL.

1.0 Murine Local Lymph Node Assay (LLNA) Limit Dose Procedure[1]

1.1 Comments on the Draft Background Review Document (BRD) for Completeness, Errors and Omissions

1.1.1 General Comments

The international independent scientific peer review panel (hereafter, Panel) was asked if there were errors in the draft LLNA limit dose procedure BRD that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. The Panel agreed that consideration should be given to applying the same term to the LLNA limit dose procedure since in various places throughout the draft BRD it was referred to differently as either the "cut-down", the "limit dose", or the "reduced LLNA" (rLLNA). Since the European Centre for the Validation of Alternative Methods (ECVAM) has already established a naming convention of "rLLNA", the Panel recommended adopting the ECVAM terminology to harmonize the terminology used among the international validation agencies.

The Panel recommended that since the validation of the LLNA limit dose procedure encompassed data that was analyzed retrospectively, a discussion of the limitations of a retrospective evaluation of previously published LLNA results should be included in the final BRD. In particular, the assumption that the highest dose in the retrospective dose-response study would be equivalent to the highest possible dose tested in the limit dose procedure should be addressed. Discussing such a limitation would be important since it bears directly on the validation of the limit dose procedure.

Further, since determination of the appropriate "limit dose" is critical to the LLNA limit dose procedure, the Panel suggested that a discussion of how to arrive at the maximal concentration for test substance dosing should be included in the final BRD. The final BRD should also specifically define what is meant by the terms "avoidance of excessive irritation" and "systemic toxicity" to aid in choosing the most appropriate maximum dose. In this regard, the Panel suggested that a systematic and quantitative measurement of ear thickness and systemic toxicity be considered or evaluated for routine inclusion in the LLNA protocol.

The Panel discussed modifying the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) LLNA protocol requirement for testing concurrent positive controls (ICCVAM 1999; Dean et al. 2001) as a means of further streamlining the LLNA limit dose procedure (i.e., reducing animal number, cost, etc.). Although the Panel did not reach consensus, a suggestion was made that for laboratories in which the LLNA is "routinely" performed and which had demonstrated the ability to consistently obtain positive results, hexyl cinnamic aldehyde (HCA) or another positive control (e.g., a substance that matches the chemical class of the test substances) could be run at intervals for quality control purposes rather than concurrent with each experiment. The Panel noted that Kimber et al. (2006) have described the "routine" use of the "rLLNA" utilizing only a vehicle and a high-dose group. The Panel also recommended that for laboratories that perform the LLNA only "occasionally", a concurrent positive control should be used. However, in their discussions,

[1] Also referred to as the "reduced" LLNA.

the Panel was not able to conclude what would constitute "routine" or "occasional" LLNA use or what would be an appropriate interval between positive control testing when a concurrent positive control is not used.

The Panel also noted that including the following additional information in the final BRD might prove informative if included:

- An indication of any procedural problems reported for the tests
- An indication of the range of historical values obtained with the negative and positive controls (the positive control historical range might give insight into the need for a concurrent positive control)
- Any discussion of global harmonization should expand on why the draft BRD did not place more reliance upon Organisation for Economic Co-operation and Development (OECD) Test Guideline (TG) 429 (OECD 2002) as a normative reference
- For two of the substances tested multiple times (HCA and linalool alcohol), different doses were used and discordant results were obtained. It should be noted for which (if either) of these tests, the highest dose tested was consistent with the dose selection principles set out in the draft BRD

1.1.2 General Statistical Comments

The Panel also had some statistical comments related to the LLNA limit dose procedure. First, the Panel viewed that a reference to the supplemental statistical information in which Dr. Joseph Haseman performed power calculations on the traditional LLNA would be useful in determining if the sample size used in the LLNA limit dose procedure was adequate for evaluating skin sensitization potential. Also, the Panel concluded that although a stimulation index (SI) based on the ratio of 3.0 as the decision criterion for a sensitizer is informative, statistical analysis determining if the test substance is significantly different from the control substance should be recommended.

1.1.3 Comments with Specific References to the Text

The Panel also identified the following minor formatting and grammatical errors, and information gaps in the draft BRD:

- The manner of notating numerical data in the draft BRD tables was not consistent (e.g., in some places the value "one" was shown as 1, elsewhere as 1.0; in a few places the SI was shown over the percent concentration used).
- Lines 291-294: The statement was made that "chemical class information is included to provide an indication of the variety of structural elements present in the substances that were evaluated in this analysis, but it is not intended to suggest an impact of structure on sensitization potential". The latter concept is not entirely correct; the portion of the sentence stating "but it is not intended to suggest an impact of structure on sensitization potential" should be omitted. Consideration should be given to using the large database of chemicals to selectively modify structure-activity relationship (SAR) software for improved predictivity. This could likely be accomplished by communication with

software developers to point out availability of the newly expanded ICCVAM dataset developed for evaluation of the LLNA limit dose procedure.

- Lines 299-300: The sentence is incomplete; “non-sensitizers” should be inserted at the end of the sentence.
- Line 358: The citation to Sailstad et al. (2001) was not listed in Section 12.0 (References) and should be included.
- Lines 365-384 (Section 1.1.2): Consideration should be given to expanding the background on the mechanism and natural history of allergic contact dermatitis (ACD). Some additional detail regarding the biochemistry and cell biology of immune induction and elicitation would be useful as an orientation to how the LLNA functions as an integrated method of detection for ACD.
- Lines 366–368: The introductory sentence on the prevalence of ACD as an occupational health issue would benefit from amplification to also indicate that ACD is of variable severity with some potentially severe ramifications, and that ACD is recurrent upon rechallenge possibly requiring workplace accommodation or change of employment.
- Lines 366-368: There is no reference provided for the statistic from the U.S. Department of Labor Bureau of Labor Statistics cited in Section 1.1.2 of the draft BRD.
- Line 433: The reason for further evaluation of negative results with concentrations less than 10% should be clarified (Kimber et al. 2006).
- Lines 435-436: This bullet point re-plays conclusions made in the summary of the “rLLNA” issued by the scientific advisory committee of the European Centre for the Validation of Alternative Methods (ECVAM). However, in the draft BRD the statement has been altered and should instead read "..., as appropriate, per OECD TG 429 (OECD 2002)." rather than citing ICCVAM (1999) and Dean et al. (2001).
- Lines 452-453: The intent of the sentence would be clarified by modifying to read "...to identify potential human skin sensitizers through quantification of lymphocyte proliferation in the test method."
- Lines 496-500: Reading of the Kimber et al. (2006) citation does not indicate a recommendation for a concurrent positive control group. Thus, the sentence in the draft BRD that reports use of vehicle and positive control groups in the limit dose procedure based on the Kimber et al. (2006) paper is incorrect.
- Lines 509-510: This sentence should also mention that the LLNA limit dose procedure, as published by Kimber et al. (2006), is rationalized not only as a means of bioresource economy but as a valid means of streamlining the LLNA for regulatory screening assessment purposes under regulation such as Registration, Evaluation and Authorisation of Chemicals (REACH).
- Lines 520-523: A footnote might be needed to explain why OECD TG 429 (OECD 2002) is not referenced here.

- Lines 629-636: Data donated by GlaxoSmithKline (GSK) to the National Toxicology Program Interagency Center for the Validation of Alternative Toxicological Methods (NICEATM)-ICCVAM LLNA review were generated under GLP conditions at a clinical research organization (CRO) repeatedly audited for GLP compliance by GSK. This information should be added to the text.
- Section 6.1: The presentation of data and associated discussion regarding limitations in accepting a 10% dose concentration cut-off should be repositioned for emphasis. This information is important in developing suggestions for a standard LLNA limit dose procedure. Data from Appendix D could be reduced to a small table or figure and be integrated into the body of the final BRD.
- Lines 722-723: This appears to be an incomplete sentence.
- Lines 815-822: Data donated by GSK to the NICEATM-ICCVAM LLNA review were generated under GLP conditions at a CRO repeatedly audited for GLP compliance by GSK. This information should be added to the text.
- Table following line 1126: There is a typographical error in the 2 x 2 table. It appears that the cell for Negative (New Test) x Total (Reference Test) should read "c + d" rather than "a + d".

1.2 Comments on the Validation Status of the LLNA Limit Dose Procedure

1.2.1 Test Method Protocol

For the proposed LLNA limit dose procedure, ICCVAM recommended that the number of animals used in each group should be the same as that recommended by ICCVAM for the traditional LLNA based on its 1998 evaluation (i.e., at least five animals per group), and that individual animal data should be collected and reported (ICCVAM 1999; Dean et al. 2001). The Panel was asked whether they agreed that these are appropriate protocol requirements for the limit dose procedure. The Panel agreed that, based on the supplemental power calculations for the traditional LLNA performed by Dr. Joseph Haseman (see **Table 1-1**), a minimum of five animals per treatment group should be recommended for all future studies employing the limit dose procedure. In addition, the collection of individual animal data, as recommended by ICCVAM for the traditional LLNA (ICCVAM 1999; Dean et al. 2001), should also apply to all future studies following the LLNA limit dose procedure. Similarly, application of the LLNA limit dose procedure to a modified LLNA protocol would require adherence to a validated protocol with the exception of omitting the middle and low dose groups. Respective power calculations would indicate if application of the LLNA limit dose procedure to a validated modified LLNA protocol would allow fewer animals per dose group.

Table 1-1 Power Calculations for the Traditional LLNA

	3.0-fold increase	2.5-fold increase	2.0-fold increase	1.5-fold increase	1.3-fold increase
Mean Rx response	1034.4	862.0	689.6	517.2	448.24
Log (Mean Rx response)	6.942	6.759	6.536	6.248	6.105
Difference (log scale)	1.099	0.916	0.693	0.405	0.262
Difference/SD	2.40	2.00	1.51	0.88	0.57
Power for N=5	95%	80-90%	50-80%	<50%	<50%
Power for N=4	90%	80%	50%	<50%	<50%
Power for N=3	50-80%	50-80%	<50%	<50%	<50%
Other Power	–	–	95% (N=11)	95% (N=29)	95% (N=68)
Other Power	–	–	90% (N=9)	90% (N=23)	90% (N=54)

The power calculations above are based on a one-sided $p<0.05$ Student's t test applied to log-transformed data from vehicle control LLNA tests.
Abbreviations: Fold-increase=Required increase above the vehicle control for a positive response (i.e., the stimulation index); N=number of animals; Rx=Treatment; SD=standard deviation.

The primary rationale for both provisions is to underpin robust statistical analysis of LLNA results. Furthermore, the use of individual animal data would allow for the evaluation of dosing errors or other anomalies that might be masked by the use of pooled animal data.

1.2.2 Substances Used for the Validation Studies

The Panel was asked whether they considered the traditional LLNA database representative of a sufficient range of chemical classes and physical chemical properties such that it would be applicable to any of the types of chemicals and products that are typically tested for skin sensitization potential. If not, the Panel was asked which relevant chemical classes/properties (other than those identified as limitations in the traditional LLNA) should be tested with caution, or not evaluated using the limit dose procedure, and which chemicals or products should be evaluated to fill this data gap. The Panel agreed that, in general, the traditional LLNA database included in the LLNA limit dose procedure evaluation was representative of a sufficient range of chemical classes and physical chemical properties and that it should be applicable to any of the types of chemicals and products that are typically tested in the traditional LLNA for skin sensitization potential. It was notable that the substances included in this evaluation provided a diverse chemical database. Since much is known about the mechanism of sensitization, the LLNA should theoretically identify any chemical that works by migration of haptens to the lymph node. However, the Panel noted that substance classes that are sometimes problematic in the LLNA (i.e., metals) would also likely be problematic in the LLNA limit dose procedure. There were also some substance classes that had limited or no representation in the draft BRD (i.e., mixtures/formulations, higher molecular weight biopharmaceuticals, and medical device materials). Thus, in general, the LLNA (and the LLNA limit dose procedure) is best used as part of a weight-of-evidence appraisal in which

attributes such as physical chemical parameters, SAR evaluation, and indications of other biological activity involving potential chemical-to-biological macromolecule interactions, are carefully considered along with LLNA results to evaluate dermal sensitizing potential.

Because the LLNA limit dose procedure was based on a retrospective evaluation of existing data, most of which was not generated using coded chemicals, the Panel was asked whether a lack of coding of test substances adversely impacted or biased the current evaluation. The Panel considered that although coding of chemicals should be recommended for prospective validation studies, this evaluation was likely not adversely impacted or biased because of a retrospective evaluation of existing data. This is supported by the fact that many of the chemical structures included in the analyses do not appear to contain known structural motifs associated with ACD or chemical hypersensitivity and therefore there was no a priori expectation that the chemical tested would be a sensitizer. The Panel viewed it important to consider the issue of coding or bias in prospective validation studies.

For some substances submitted using the traditional LLNA test method, it was not possible to confirm whether the data were generated using pooled animal data for each dose group (as allowed in OECD TG 429 [OECD 2002]). ICCVAM (1999), Dean et al. (2001), and EPA (2003) recommend the use of statistical analyses to help interpret LLNA study results, which necessitates data collected at the level of the individual animal. Furthermore, Cockshott et al. (2006) reported that using individual animal data allowed for technical problems or other outliers during an experiment to be identified. The Panel was asked what impact the inclusion of pooled animal data might have on the accuracy analysis of the LLNA limit dose procedure. The Panel concluded that, although it would be important to note whether individual or pooled animal data were reported, the retrospective analysis of the LLNA limit dose procedure versus the traditional LLNA should not be limited to studies with confirmed individual animal data. The Panel stated that internationally, both individual and pooled animal data have likely been used both for regulatory decisions and for in-house decisions relating to product development and risk management. Also, the fact that the retrospective data analysis presented in the draft LLNA limit dose procedure BRD did not distinguish between individual or pooled animal data suggested that both met the quality standards for inclusion in the draft BRD.

1.2.3 Test Method Accuracy

The Panel was asked whether the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA limit dose procedure had been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft ICCVAM BRD). The Panel concluded that the relevance of the LLNA limit dose procedure had been adequately evaluated and compared to the traditional LLNA. Comparisons resulting in an accuracy of 98.9% (461/466), a sensitivity of 98.4% (308/313) and a specificity of 100% (153/153) for the LLNA limit dose procedure when compared to the traditional LLNA were sufficient to consider it adequately validated for use in the evaluation of skin sensitization, mindful of its known limitations that are described elsewhere. Still, the Panel noted that it was important to keep in mind that a prospective analysis may not have the same accuracy as this retrospective analysis.

Furthermore, there were five substances for which the highest concentration tested produced an SI <3.0, while lower concentrations of these substances produced an SI >3.0 (see Table 6-

2 of the draft ICCVAM BRD). These substances were classified as false negatives compared to what was obtained in the traditional LLNA. The Panel was asked to identify any characteristics associated with these or other substances that might signal that this type of abnormal dose response might occur, and therefore, that using the LLNA limit dose procedure would not be appropriate. The Panel could not identify any common characteristics associated with the five false negative substances that would explain the non-linear dose response obtained. It was not known if any procedural problems were reported with these studies or what values were returned by the negative/positive control groups (in relation to other historical positive control values).

Thus, the Panel suggested that it might be worthwhile to examine whether LLNA results with these five false negative substances should be repeated. If the difference turned out to be repeatable, there could be hypothetical reasons to explain why the higher doses did not pass the SI threshold of 3.0. For example, under certain experimental conditions, the target lymphocytes may be selectively induced to a highly sensitive state by some chemicals at higher doses and may undergo either induction of apoptosis or inhibition of cell proliferation. Still, there was no evidence that these substances were immunomodulators that might have differentially stimulated or depressed the immune response depending on the dose and exposure. In any case, understanding false negatives is encouraged to ensure adequate protection of public health.

The Panel was asked whether the draft BRD adequately characterized the usefulness and limitations of the LLNA limit dose procedure based on the accuracy analyses. Overall, the Panel agreed that the draft BRD adequately characterized the usefulness and limitations of the LLNA limit dose procedure based on the accuracy analyses. Since the LLNA limit dose procedure and the traditional LLNA have close concordance, there was no need for detailed discussion in the draft BRD. However, it was not explicitly stated in the draft BRD that compared to a fully conducted traditional LLNA, a false positive result in the LLNA limit dose procedure is not possible (i.e., if the single dose used in the proposed limit dose procedure gives an SI $\geq$3.0, then so would the top dose in the traditional LLNA). Furthermore, prospective testing with the LLNA limit dose procedure to predict the sensitization potential of an unknown chemical was not discussed.

1.2.4 Test Method Reliability

The Panel was asked if it was appropriate to assume that the intra- and inter-laboratory reproducibility of the LLNA limit dose procedure and the traditional LLNA would be similar, based on the fact that they use identical protocols with the exception of the number of doses used (i.e., would reducing the number of test substance dose groups from three to one reduce the reliability of the assay?). The Panel agreed that it was appropriate to assume that the intra- and inter-laboratory reproducibility of the LLNA limit dose procedure and the traditional LLNA would be similar, because reproducibility is more dependent on the method than on the number of dose groups. However, reducing the number of test substances dose groups from three to one could reduce the sensitivity of the assay (i.e., the ability to correctly identify sensitizers). The traditional LLNA may have a greater chance of correctly identifying a sensitizer even in the presence of one or more technical errors since there are data from three dose groups for consideration and an SI $\geq$3.0 at any dose group would result in the substance being classified as a sensitizer. However, for the purpose of adopting an assay that uses fewer animals and provides increased throughput for screening purposes,

these hypothetical considerations are not a sufficient reason to argue against use of the limit dose LLNA procedure.

1.2.5 Data Quality

For some studies included in the draft BRD, it was not possible to determine whether or not they had been conducted in accordance with GLP guidelines. Furthermore, original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data was the same as the data recorded in laboratory notebooks. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. The Panel was asked whether the results of such studies (all of which are currently included) should be excluded from the performance analyses. The Panel concluded that it was important to note if the data were obtained from studies conducted according to GLP guidelines, as ideally this should be the case. However, the Panel concluded that the data resulting from the retrospective studies that could not be confirmed as GLP-compliant should not be excluded from the performance analysis. Since there was not an indication that the reliability of the data presented for consideration may have been compromised, omitting any data would likely lessen the impact of the analysis. Furthermore, data obtained from peer-reviewed literature or final reports were likely of sufficient quality.

1.2.6 Consideration of All Available Data and Relevant Information

The Panel was asked if all the relevant data identified in published or unpublished studies conducted using the traditional LLNA had been adequately considered in the draft BRD. If not, the Panel was asked what other traditional LLNA data needed to be considered and how such data could be obtained. The Panel considered that the draft BRD had taken into account a large majority of the relevant data identified in published and unpublished traditional LLNA studies. The data received as a result of the *Federal Register (FR)* notices and the key literature citations seemed to be inclusive of the relevant data for this analysis. Although additional data that could have been included might exist, it was deemed unlikely that the current outcome (which is based on 466 substances) would be altered given the very small change in accuracy statistics relative to Kimber et al. (2006), which was based on 211 substances.

1.3 Comments on the Draft ICCVAM Test Method Recommendations on the LLNA Limit Dose Procedure

1.3.1 Test Method Usefulness and Limitations

The Panel was asked to comment on whether the available data supported the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed test method usefulness and limitations (i.e., that the LLNA limit dose procedure should be routinely recommended for hazard identification when dose response information is not required). The Panel considered that, based on the available information, the draft recommendations appeared valid, but made the following suggestions:

- Further emphasis should be given to using the LLNA limit dose procedure as a part of a comprehensive weight-of-evidence evaluation of dermal sensitizing potential (e.g., including physical chemical evaluation, SAR information,

including likelihood of dermal penetration, ability of materials to adduct biomacromolecules).

- Such information in addition to LLNA results might also be useful in confirming or questioning LLNA outcomes terms of in human hazard identification, since it should be emphasized that a major application of the method is to prospectively detect harmful chemicals.
- Solubility or thermodynamic activity data, beyond visual assessment (e.g., use of chemically-specific methods to document solubility), should be used to confirm the appropriateness of the maximum dose tested.
- Vehicle selection for the LLNA can affect the results and may not allow accurate comparisons between chemicals applied in different vehicles. In choosing the best vehicle, consider measured solubility information for the potential vehicle. Then, it would be important to take into account how the vehicle affects the amount of the chemical that can be applied to the ear. More importantly, the impact that vehicle selection has on the amount of applied chemical that actually gets into the mouse to induce the sensitization response should be evaluated. Some of the recommended LLNA vehicles (e.g., 4:1 acetone:olive oil (AOO), dimethylsulfoxide (DMSO), methyl ethyl ketone) could be expected to disrupt the barrier properties of the skin. Additionally, although propylene glycol might allow an increased amount of chemical to be applied, it might also inhibit the penetration of a chemical by enhancing partitioning in the vehicle relative to the skin.

The Panel was asked whether the LLNA limit dose procedure should be routinely recommended for the hazard identification of skin sensitizing chemicals when dose response information *is not* required. With the points noted above in mind, the Panel agreed that it should be routinely recommended since the LLNA limit dose procedure offers time, cost, throughput, and logistical benefits as well as using fewer animals. Still, the investigator should keep in mind what is known of the chemical regarding general toxicity and note scenarios where abnormal dose-response relationships in the traditional LLNA might result in false negatives in the limit dose procedure (see Table 6-2 of the ICCVAM BRD).

The Panel was then asked whether the LLNA limit dose procedure should be routinely recommended as the initial test to identify sensitizers before conducting the traditional LLNA, as a way to further reduce animal use, even if dose response information *is* required, since negative results would not require further testing. The Panel agreed that use of the LLNA limit dose procedure, as the initial testing procedure to identify sensitizers and non-sensitizers before conducting the traditional LLNA, is justifiable even when dose response information is required. This is applicable in the occupational and public health setting where obtaining hazard information is of critical importance. There is a benefit since dose-response information generated in subsequent testing in the traditional LLNA for substances that were positive in the limit dose procedure then gives further assurance of detecting hazardous substances and allowing a potency estimate. The benefits of screening out the negatives (which do not require dose response information) is clear; however the animal welfare gains will depend on the proportion of test substances in any class that turn out to be non-

sensitizers and there might be possible consequences of the delays resulting from a further round of testing for those materials that are identified as sensitizers.

Based on the existing database, there is a false negative rate of 1.6% (5/313 positive compounds) for the LLNA limit dose approach compared to the results obtained in the traditional LLNA. The Panel was asked whether they considered that this is adequately addressed by the proposed cautionary language and weight-of-evidence consideration for negative substances. The Panel agreed that the small rate of false negatives was adequately addressed in the draft test method recommendations by giving cautionary and weight-of-evidence consideration to the negative substances (and any possible false positive results). Furthermore, given that the dose responses for these five materials were rather unusual, it was not known whether these studies were repeatable, whether any procedural problems were reported with these studies, or what values were returned by the negative/positive control groups (in relation to other historical positive control values). In general, the Panel viewed that the false negative rate of 1.6% would likely be unimportant when the larger differences between the animal model and humans are considered.

1.3.2 Test Method Protocol

The Panel was asked whether they agreed that the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed standardized test method protocol. The Panel agreed and recommended adherence to the ICCVAM (1999) LLNA protocol for future studies of the LLNA limit dose procedure with the exception of omitting the middle and low dose groups. Similarly, application of the LLNA limit dose procedure to a modified LLNA protocol would require adherence to the modified LLNA protocol with the exception of omitting the middle and low dose groups. Adhering to the use of individual animals for future studies was specifically stressed because it would allow for an estimate of inter-animal variability.

The recommended ICCVAM protocol (ICCVAM 1999; Dean et al. 2001; EPA 2003), as well as OECD TG 429 (OECD 2002), specifies that the highest dose tested should be the highest soluble concentration that does not induce systemic toxicity and/or excessive skin irritation. However, Kimber et al. (2006) concluded that negative results obtained from studies where the highest concentration tested was below 10% should be considered invalid, and adopted a 10% application concentration as a threshold of confidence for categorization of a chemical as being negative while noting that the figure should not be considered as inviolable. The Panel was asked whether the data presented in the draft BRD (i.e., 51/313 positive substances in the NICEATM database were negative at concentrations equal to or above 10%, but were positive at even higher concentrations) were adequate to conclude that this threshold concentration is not appropriate. The Panel viewed that this point should be clarified. ICCVAM recommended that no threshold should be used to determine the validity of conduct of the LLNA limit dose procedure. Instead, formal attempts to maximize dose delivery including documentation of solubility of the test substance in the vehicle used should be undertaken.

The Panel was asked whether additional testing should be required if a negative result was obtained for a test substance in a study where the highest concentration that could be tested (based on systemic toxicity or excessive local irritation, as described in ICCVAM [1999], Dean et al. [2001], and EPA [2003]) was <10%. The Panel considered that, if a negative

result was obtained for a test substance under these conditions, additional testing should not be required, because at that point it would likely be a toxic effect and not sensitization. In contrast, the imperative should be to minimize the number of false negatives. For this purpose, rigorous examination of maximum solubility or other parameters to ensure testing at maximum concentration should be employed. In addition, weight-of-evidence considerations such as SAR and physicochemical characteristics should be documented. More animal testing to verify negative results should only be undertaken if the weight-of-evidence suggests that it would be appropriate.

The Panel was asked if the current approach for selecting the "limit dose" was appropriate or whether there is a threshold concentration for the LLNA at which a negative result could always be considered as an acceptable result. The Panel agreed that the current recommendation to select a maximum applied dose in the LLNA limit dose procedure is appropriate. However, the data presented in the draft BRD implied that at present it is not possible to establish a uniform concentration threshold for the "limit dose". Thus, it seemed justifiable that preliminary experimentation (as would be typically performed during a dose range finding study) should be conducted on vehicle selection, test substance solubility, and stability in the vehicle.

1.3.3 Future Studies

The Panel was asked if they agreed that the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed future studies. Although limited in scope, the Panel considered that the available data supported the ICCVAM draft recommendations for additional studies. The Panel agreed that attempts be made to investigate if maximum solubility was achieved (e.g., use of chemically-specific methods to document solubility). For hazard assessment, it was considered troublesome that there were so many vehicle choices, because the vehicle could have a significant effect on whether (and how much) a test substance penetrated the skin barrier. Observed vehicle effects may relate to dermal penetration as well as to immunomodulation. The Panel considered it desirable to follow the hierarchy of vehicles recommended in the ICCVAM (ICCVAM 1999; Dean et al. 2001) protocol. In addition, dedicated attempts must be made to investigate solubility in AOO mixtures before using other vehicles. Regardless of the vehicle used, it is important to ensure that a vehicle does not promote lymph node cell proliferation. The Panel also suggested that it might be informative to test both known mild and severe sensitizers concurrently in all recommended vehicles to evaluate if a specific vehicle choice(s) might influence the results.

Although the false negative rate in the current analysis was small, a need exists to better understand factors that could lead to false negative results with future use of the LLNA limit dose procedure. Thus, consideration should also be given to formal statistical assessments to verify group size and use of individual animal data in routine performance of the LLNA limit dose procedure. Criteria should be established to verify proficiency with the LLNA limit dose procedure. Such criteria could be used to answer questions about the necessity to perform concurrent positive controls.

1.3.4 Comments with Specific References to the Text

The Panel also identified the following comments and/or corrections to the draft ICCVAM test method recommendations document on the LLNA limit dose procedure that should be considered by ICCVAM:

- Lines 26-28: Conclusions given here regarding the relative potency ratings of the five materials classified as false negative in the analysis in Section 6.2 of the draft BRD were newly introduced. This assessment should also have been considered for inclusion in Section 6.2 of the draft BRD.
- Line 28: The citation of Gerberick et al. (2004) was not accompanied by a reference.
- Lines 62-70: The listing of substances not amenable to test in the LLNA could have been expanded to also include agents with anticipated pharmacodynamic action as immune suppressants.
- Line 69: The citation of Gaspari et al. (2007) was not accompanied by a reference.
- Lines 75-79: Dependent upon other considerations, this portion of the text could have been modified to (1) clarify recommendations regarding routine use of concurrent positive control (i.e., possible exception for laboratories conducting a high volume of LLNA work in which periodic positive control for quality control purposes might suffice), and (2) expand on the details regarding indications of excessive irritation and/or systemic toxicity to aid in choice of maximal test dose.

2.0 LLNA for Testing Aqueous Solutions, Metals, and Mixtures

2.1 Comments on the Draft Addendum for Completeness, Errors, and Omissions

In regard to the draft Addendum to the traditional LLNA BRD, the Panel was asked to comment on any errors that should be corrected or omissions of relevant data/information that should have been included. The Panel concluded that there were no apparent errors or omissions to the draft Addendum.

2.2 Comments on the Validation Status of the Traditional LLNA for Testing Aqueous Solutions, Metals, and Mixtures

2.2.1 Substances Used for the Validation Studies

The Panel was asked whether the database of substances evaluated was representative of a sufficient range of mixtures, metal compounds, and substances in aqueous solutions that are typically tested for skin sensitization potential. While there were limited data available on the effects of mixtures, metals, and aqueous solutions on skin sensitization potential, the Panel considered the database to be generally representative. The Panel indicated that there did not seem to be obvious classes of chemicals missing from the data set used to evaluate the utility of the traditional LLNA for testing aqueous solutions. However, quantitative compositions for the mixtures included in the analysis had not been provided. Thus, it was difficult to determine if those mixtures were representative of the types of mixtures typically tested in the traditional LLNA. With respect to metals, there was a paucity of commercially useful metals such as platinum, palladium, iron, zinc, manganese, and silver compounds. To enlarge the group of metal non-sensitizers, substances used as cosmetic ingredients (e.g., titanium dioxide) and aluminum compounds currently used in antiperspirants might be considered. However, the Panel considered that the inclusion of an array of other metals and at least one zinc and manganese salt likely weighted the data set appropriately and it appeared sufficiently broad to support conclusions about the utility of the traditional LLNA for testing the skin sensitization potential of metals.

Substances or mixtures that were tested in an aqueous or an organic:aqueous vehicle were labeled as aqueous solutions. For the purpose of this evaluation, a substance or mixture containing at least 20% water was defined as an aqueous solution. The Panel was asked whether this criterion was appropriate for defining an aqueous solution. The Panel was uncertain about the appropriateness of this definition of an aqueous solution, but did not offer an alternative definition. However, the Panel indicated that an organic:aqueous solution that is not miscible would likely produce varying results because of partitioning of the chemical into either phase.

The Panel was asked whether the lack of coding of test substances might adversely impact or bias the current evaluation. While coding of chemicals is recommended for prospective validation studies, the retrospective evaluations in the draft Addendum were based on existing data, most of which were not generated using coded chemicals. However, the Panel agreed that the lack of chemical coding was not likely to bias the evaluation since this study was retrospective. This is supported by the fact that many of the chemical structures included in the analyses did not contain known structural motifs associated with allergic contact dermatitis/chemical hypersensitivity and therefore there was no a priori expectation that the

chemical tested would be a sensitizer. Furthermore, many of the substances tested were apparently evaluated for hazard assessment purposes rather than to test the predictive ability of the traditional LLNA. Thus, there does not appear to be any bias in chemical selection for the expanded dataset considered in the study of applicability domain for the traditional LLNA.

For some substances submitted using the traditional LLNA test method, it was not possible to confirm whether the data were generated based on pooled lymph nodes among animals within a dose group, as allowed in OECD TG 429 (OECD 2002), or individual animal responses, as recommended by ICCVAM (1999) and required by EPA (2003). ICCVAM (1999) and EPA (2003) both recommend the use of statistical analyses to aid in the interpretation of traditional LLNA study results; such analyses necessitate data collected from individual animals. Additionally, Cockshott et al. (2006) reported that using individual animal data allowed for outlier animal results within a dose group to be identified. The Panel was asked whether the analysis of the performance of the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions should be limited to data from studies that collected individual animal data, and then to comment on the potential impact on the accuracy analysis of including results from studies in which pooled animal data were collected. The Panel concluded that, although individual animal data were preferred, pooled animal data should not be excluded automatically from this retrospective analysis.

2.2.2 Test Method Accuracy

The Panel was asked whether the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions had been adequately evaluated and compared to the human and guinea pig test results. The Panel agreed that the comparative assessment of the relevance of the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions appeared to be as comprehensive as was feasible. However, because of the limited number of comparisons available, the accuracy statistics probably do not give a complete picture of the usefulness and limitations of the traditional LLNA for identifying skin sensitizers among these types of substances or when using an aqueous vehicle.

When multiple traditional LLNA studies were available for the same substance, the "majority call" (among studies using the same vehicle and generally tested over the same concentration range) was used by ICCVAM to assign an overall classification for the purposes of the accuracy analysis. For example, if chemical X was tested five times and was positive in three studies and negative in two, the overall classification was positive. The Panel was asked whether they agreed with this approach. They expressed their concern about the approach in the following way; if all nickel-containing compounds in the analysis were viewed as a group, there were four positive calls and four negative calls (see Appendix C2 of the draft Addendum). Using the "majority call" approach, the overall call would be determined by the next available study, which may not provide the correct call. More data would be needed to confirm whether the classification was appropriate. For this dataset, most of the "negative calls" had SI values that approached 3.0. Thus, a more suitable method might be to base the overall call on the SI data, while giving greater positive call consideration/weight to SI values just below 3.0. It may also be useful to perform a meta-analysis. It is important for the

Addendum to mention the potential impact of using the "majority call" decision, rather than relying on a weight-of-evidence approach, on the accuracy analyses.

2.2.3 Data Quality

For some studies included in the draft Addendum, it was not possible to determine whether or not they had been conducted in accordance with GLP guidelines. Original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data was the same as the data recorded in laboratory notebooks. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. The Panel was asked to discuss what impact this lack of information might have on the evaluation of the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions and whether such studies should be excluded from an analysis of test method accuracy. The Panel considered it important to note if the data were obtained from studies conducted according to international GLP guidelines, since ideally this should be the process followed. However, the Panel viewed that data from studies that could not be confirmed as being GLP-compliant were still appropriate to include in the accuracy analysis, provided that the data were from the peer-reviewed literature or from sources with high quality laboratory management practices. Much of the value for this draft Addendum was the potential to supplement the data available at the time of the ICCVAM (1999) analysis. Additional information on test substance identification would clearly be useful in the continued evaluation of the applicability domain of the traditional LLNA, but omitting data on mixtures, metals, or use of aqueous solutions based solely on the lack of GLP compliance would lessen the impact of the current retrospective analysis and did not seem warranted. However, if the original data were not available, it would be appropriate to note this in the final version of the Addendum.

2.2.4 Consideration of All Available Data and Relevant Information

The Panel was asked whether the draft Addendum included all of the relevant data for studies conducted using the traditional LLNA for testing mixtures, metal compounds, and substances in aqueous solutions. The Panel considered that, although it was possible that there might be a few studies in the literature to augment the analysis, it seemed that the relevant data had been identified and the response to the *FR* notice and the literature citations examined had included the most relevant studies.

2.3 Comments on the Draft ICCVAM Test Method Recommendations on the Traditional LLNA for Testing Aqueous Solutions, Metals, and Mixtures

2.3.1 Test Method Usefulness and Limitations

ICCVAM stated that more data would be needed before a recommendation on the usefulness and limitations of the traditional LLNA for testing mixtures could be made, due to the limitations associated with the available mixtures database (i.e., unknown formulae, lack of human data). The Panel was asked whether they agreed that the available data supported the ICCVAM draft recommendations for the traditional LLNA with regard to testing mixtures in terms of the proposed test method usefulness and limitations. The Panel agreed that ICCVAM's draft recommendation with respect to the traditional LLNA testing of mixtures appeared valid based on the limitations inherent in the available data set. Still, the Panel

urged that the ICCVAM recommendation indicate that the approach may be viable. The Panel further recommended that the test method recommendations summary should indicate that the limitations include relatively poor concordance of traditional LLNA outcomes for mixtures with to those obtained in guinea pig tests. Routine comparisons of accuracy according to classification criteria may not be sufficient to evaluate the concordance for mixtures, and furthermore, the guinea pig tests are not necessarily valid for mixtures. The Panel also indicated that the term "mixtures" was used too broadly (i.e., can represent an infinite number of materials) and it would be more beneficial to specify types or formulations of mixtures that are being examined.

ICCVAM recommended that, based on the available data for metals, the traditional LLNA was useful for the testing of metal compounds, with the exception of nickel. The Panel was asked whether they agreed that the available data supported the ICCVAM draft recommendations for the traditional LLNA with regard to testing metals in terms of the proposed test method usefulness and limitations. Based on the available information, the Panel agreed that the draft recommendations with regard to testing metals appeared to be valid. In particular, the evidence for most metals (e.g., accuracy of 86% (12/14), sensitivity of 100% (9/9), specificity of 60% (3/5), 0% (0/9) false negatives) when comparing traditional LLNA results to those obtained from evaluations in humans supported the use of the traditional LLNA as a hazard identification tool for metals, excluding nickel. However, the Panel recommended that it would be worthwhile to study further the variable results obtained for nickel since there is a wealth of literature on allergic contact dermatitis of nickel in humans.

In a minority opinion, Dr. Dagmar Jírová stated that it should not be concluded that the traditional LLNA was not suitable for testing nickel compounds, because the different vehicles used may have had a significant impact on the ability of nickel to penetrate the skin and be bioavailable. She noted that nickel chloride and nickel sulfate were both positive in aqueous solutions, and negative only when non-aqueous vehicles were used. In human exposures, nickel compounds were applied in aqueous solutions. Thus, this may serve as sufficient justification to use aqueous vehicles when nickel, and perhaps also other substances, are tested and evaluated in the traditional LLNA. When DMSO was used as the vehicle, the SI value increased with increasing nickel concentration. Unfortunately, no data were available for concentrations over 5% for either nickel compound in DMSO. Nickel chloride as 10% in aqueous solution reached an SI of 6.6. Inconsistent test results due to the vehicle have also occurred in other *in vitro* studies (e.g., phototoxicity). Thus, Dr. Jírová concluded that the traditional LLNA could be used even for testing nickel compounds when other vehicles (in particular aqueous) are used.

Due to the limited number of substances tested in aqueous solutions, ICCVAM recommended that more data would be needed before a recommendation on the usefulness and limitations of the traditional LLNA for testing substances in aqueous solutions could be made. The Panel was asked whether they agreed that the available data supported this ICCVAM draft recommendation for the traditional LLNA with regard to the testing of substances in aqueous solutions. The Panel agreed that the draft ICCVAM recommendation was appropriate and that more data were required before an adequate evaluation of the use of the traditional LLNA with aqueous solutions could be conducted.

2.3.2 Test Method Protocol

The Panel was asked whether they agreed that the available data supported the ICCVAM draft recommendations for the traditional LLNA in terms of the proposed test method standardized protocol. The Panel agreed that, in general, the results of the assessment in the draft Addendum supported the proposals for standardized conduct of the traditional LLNA. However, this conclusion depended on a side-by-side reading of the draft Addendum and the ICCVAM (1999) protocol. The Panel suggested expanding the brief section of the draft test method recommendations dealing with test method protocol for the traditional LLNA (Section 2.0) to specifically point out how the conclusions of the applicability domain evaluation may affect the standard traditional LLNA protocol. For example, the evaluation of aqueous solutions apparently resulted in the methodological recommendation that aqueous test solutions be avoided and the further recommendation of a hierarchy of organic solvents to be considered as dosing vehicles. The emphasis might be on using a vehicle to which humans may actually be exposed.

2.3.3 Future Studies

The Panel was asked whether they agreed that the available data support the ICCVAM draft recommendations for the traditional LLNA in terms of the proposed future studies. The Panel agreed that the ICCVAM recommendation for continued accrual of information from traditional LLNA evaluations of mixtures, metals, and aqueous solutions with comparative data for guinea pig and human tests was appropriate. The traditional LLNA accuracy for metals of 86% and sensitivity of 100% (0% false negative) was excellent; a specificity of 60% (40% false positive) was considered acceptable as over-classification maintains safe human use. The Panel encouraged the use of the traditional LLNA to acquire further information on mixtures, metals, and aqueous solutions. However, the Panel suggested that, given resource limitations, it would be important to prioritize the recommendations in order to focus on what is most important.

[This Page Intentionally Left Blank]

3.0 Non-Radioactive LLNA Protocol - The LLNA: Daicel Adenosine Triphosphate (LLNA: DA) Test Method

3.1 Comments on the Draft BRD for Completeness, Errors, and Omissions

3.1.1 General Comments

The Panel was asked if there were any errors in the draft LLNA: DA BRD that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. As a general comment, the Panel noted that the draft BRD clearly and succinctly provided an overview of the LLNA: DA test method and the relevant validation study data. The draft BRD indicated that the LLNA: DA differs from the traditional LLNA in the method of measuring proliferation (measures levels of adenosine triphosphate [ATP] instead of radioactivity), substance treatment (pretreating the test site with 1% sodium lauryl sulfate [SLS] prior to test substance application and an additional treatment on day 7), and sampling time (draining auricular lymph nodes are collected on day 8 rather than on day 6). Because the traditional LLNA evaluates the induction phase only, the relevance of results with the LLNA: DA (and any other LLNA protocol) should always be considered in the context of human experimental sensitization data, human epidemiologic data, and elicitation in the clinical setting.

3.1.2 Comments with Specific References to the Text

The Panel noted the following text that should be clarified or corrected in the final version of the LLNA: DA BRD:

- Line 428: The text should read "1% SLS", not "1% SDS"; the same terminology should be used throughout rather than going back and forth between SLS and sodium dodecyl sulfate (SDS).
- Line 449: The text and formula in lines 448-450 appear misplaced and instead seem to belong in Section 7.0. Additionally, X and Y should be defined and the "Var (ln SI)" formula should be clarified.
- Table 3-1: There were 33 substances in the table, yet the discussion of the table in the text referred to 31 substances. Although the reason for this apparent discrepancy becomes more evident later on, this should be discussed up front.
- Table 3-2 and 3-3: The interlaboratory distribution and testing of the sensitizers versus non-sensitizers should be indicated here.
- Table 6-1: Although the table clearly provided a comparison of the different methods, it would be useful if the footer for this table also indicated the basis for the differences in substances included in each analysis (i.e., n=25, 26, or 29) as stated in the text.
- Table 7-1 was only moderately helpful because the standard deviations (SD) were not calculated on a log scale. Given the skewness in the data, the ranges given were misleading indicators of increases due to the vehicles. Without the samples sizes, an analysis of variance was impossible to calculate, and that

would have been the appropriate measure of differences between the experiments. In order to make recommendations regarding needed reproducibility experiments, it would have been helpful to have a power analysis for this situation. That is, for each vehicle with each chemical, using two, three, four and five animals per dose group per experiment, how many experiments need to be run to detect significant differences between the experiments?

3.2 Comments on the Validation Status of the LLNA: DA

3.2.1 Test Method Protocol

Based on its 1998 evaluation of the traditional LLNA procedure, ICCVAM recommended that at least five animals be used in each test group (ICCVAM 1999). The LLNA: DA validation studies presented in the draft BRD were performed using four animals per dose group. Thus, the Panel was asked to comment on the potential impact of using fewer than five mice per dose group. The Panel noted that supplemental statistical information they were provided with indicated that the power for detecting a three-fold increase in the SI value in the treatment group for the LLNA: DA dataset evaluated in the draft BRD was estimated to be 95% for a sample size of three mice per dose group (see **Table 3-1**). Since an increase of false negatives may not be an issue, the potential opportunity exists for utilizing this smaller group size. The Panel cautioned, however, that using less than five animals per group might result in a less precise estimate of the mean response, which, in turn, will impact accuracy. Also, if technical errors further reduce the sample size, accuracy is further reduced. Thus, the Panel recommended that all initial validation studies adhere to the ICCVAM-recommended protocol (ICCVAM 1999; Dean et al. 2001) of five animals per dose group until sufficient information is generated to indicate that the use of fewer animals per dose group is statistically valid.

Table 3-1 Power Calculations for the LLNA: DA

Parameter	3.0-fold Increase	2.5-fold Increase	2.0-fold Increase	1.5-fold Increase	1.3-fold Increase
Mean Rx response	8835	7362.5	5890	4417.5	3828.5
Log (mean Rx response)	9.086	8.904	8.681	8.393	8.250
Difference from control (log scale)	1.098	0.916	0.693	0.405	0.262
Difference/SD	3.95	3.29	2.49	1.46	0.94
Power for N=5	99%	99%	95%	50-80%	<50%
Power for N=4	99%	95-99%	90%	50%	<50%
Power for N=3	95%	90-95%	80%	<50%	<50%
Other power	–	–	–	95% (N=11)	95% (N=25)
Other power	–	–	–	90% (N=9)	90% (N=20)

The power calculations above are based on a one-sided $p<0.05$ Student's t-test applied to log-transformed data from vehicle control LLNA: DA tests.
Abbreviations: Fold-increase=Required increase above the vehicle control for a positive response; Rx=Treatment; N=Number; SD=Standard deviation.

The data generated for the substances analyzed in the LLNA: DA interlaboratory validation studies came from auricular lymph nodes that were pooled across animals in each treatment group. The Panel was asked to comment on the potential impact of including pooled animal data on the accuracy analysis of the LLNA: DA. The Panel noted that a statistical analysis of differences between treatment-related and vehicle control ATP levels could not be determined without measures of variability. Individual animal data highlights technical issues and allows for consideration of dose-response information and statistical analyses.

The LLNA: DA differs from the traditional LLNA in the treatment schedule and by including a pretreatment step with 1% SLS just prior to application of the test substance. The Panel was asked to comment on the appropriateness of these protocol differences. The Panel did not consider these differences to be significant, as long as it could be demonstrated that the 1% SLS pretreatment step and the additional test substance treatment on day 7 did not induce a skin reaction indicative of the elicitation phase of skin sensitization. Although it was being used at a lower concentration than the estimated concentration needed to produce a stimulation index of 3 (EC3), the Panel expressed concern about pretreating the mouse ear with 1% SLS since SLS is an irritant and positive in the traditional LLNA. Consequently, the inherent sensitivity of the LLNA may be modified by the 1% SLS pretreatment step. To demonstrate that these concerns are not justified, the Panel concluded that weak irritants and weak sensitizers needed to be tested in the LLNA: DA assay with and without pretreatment with 1% SLS. The test method developer might also consider using decision criteria other than SI $\geq$3.0 such that 1% SLS pretreatment is no longer necessary.

3.2.2 Substances Used for the Validation Studies

The Panel was asked if they considered the substances tested in the LLNA: DA to be representative of a sufficient range of chemical classes and physicochemical properties that the test method would be applicable to any of the types of chemicals and products that are typically tested for skin sensitization potential. The Panel considered the database of test substances tested in the LLNA: DA representative of a sufficient range of chemicals. The selected substances included solids and liquids and a range of solvents/vehicles. The database also represented a range of sensitizing potency, a variety of different chemical classes and substances with differing requirements for metabolic activation. However, it might have been useful to have also included substances with clearly different protein reaction mechanisms (protein binding), as well as dyes, natural extracts, and mixtures.

3.2.3 Test Method Accuracy

The accuracy analysis in the draft LLNA: DA BRD was based on overall concordance with the traditional LLNA. Accuracy statistics compared to the guinea pig tests and human data/experience were also provided. The Panel, when asked if they considered these comparisons appropriate for assessing the accuracy of the LLNA: DA, agreed that the comparisons to the traditional LLNA performance and also to the guinea pig and human sensitization data were important. The Panel also stressed that, because the traditional LLNA only evaluates the induction phase, the relevance of the LLNA: DA results should always be considered in the context of human experimental sensitization data, human epidemiologic data, and elicitation in the clinical setting.

The Panel was asked if they considered the evaluation of the relevance of the LLNA: DA and the comparison to the traditional LLNA to be adequate. The Panel noted that Table 6-1 of the draft LLNA: DA BRD clearly provided a comparison of the different reference methods (i.e., traditional LLNA, human tests, and guinea pig tests). Thus, the Panel concluded that the relevance of the LLNA: DA had been adequately evaluated. However, including data on more substances is likely to further strengthen confidence in the concordance data.

One substance, 2-mercaptobenzothiazole, produced a false negative response compared to the traditional LLNA when tested using the LLNA: DA. The Panel was asked if they could identify any characteristics associated with this or similar substances, compared to the correctly identified sensitizers, that might signal that this type of discordant response would occur, and therefore using the LLNA: DA to test such substances would not be appropriate (or that negative results for substances with such properties may warrant additional testing). The Panel could not identify specific characteristics that might explain the false negative response using the LLNA: DA. Although understanding the solubility and stability of the test substance in different vehicles is important, the differences in response did not seem to be explained by the vehicle differences (AOO and dimethylformamide [DMF]) between the two tests. In addition, the impact of 1% SLS pretreatment on the negative response in the LLNA: DA is not known but should be considered.

One substance, benzalkonium chloride, produced a false positive response compared to the traditional LLNA and guinea pig test when tested using the LLNA: DA. The Panel was asked if they could identify any characteristics associated with this or similar substances, compared

to the correctly identified non-sensitizers, that might signal that this type of discordant response would occur, and therefore using the LLNA: DA to test such substances would not be appropriate (or that positive results for substances with such properties may warrant additional testing). The Panel could not identify specific characteristics that might explain the false positive response for this substance in the LLNA: DA. The Panel viewed that it was important to note, however, that this chemical is a well-known skin irritant, and on occasion it had also been considered a human sensitizer, typically on the basis of positive diagnostic patch test data. Thus, the Panel reiterated that the relevance of LLNA: DA results should always be considered in the context of human experimental sensitization data, human epidemiologic data, and elicitation in the clinical setting. The actual impact of the 1% SLS pretreatment step on the LLNA: DA has not been well established, although van Och et al. (2000) and De Jong et al. (2002) have reported that 1% SLS pretreatment enhances the response in the traditional LLNA.

3.2.4 Test Method Reliability

The Panel was asked if they considered the intralaboratory reproducibility of the LLNA: DA to have been adequately evaluated and compared to the traditional LLNA (refer to Table 7-1 of the draft LLNA: DA BRD). The Panel noted that only eugenol and isoeugenol, two sensitizers with similar chemical structures, were tested. The Panel recommended testing a positive control commonly used in the traditional LLNA (e.g., HCA) for a more complete evaluation of intralaboratory reproducibility. In addition, it was unclear if the tests were truly independent. Factors that might indicate independence should have been documented (e.g., time interval between experiments, different animal shipment, different reagents, different operator).

The Panel was also asked if they considered the interlaboratory reproducibility of the LLNA: DA to have been adequately evaluated and compared to the traditional LLNA. The Panel noted that the interlaboratory reproducibility of the assay could not be adequately evaluated given the lack of original study data and limitations in the study design. Study design limitations included:

- Pooled lymph nodes were used from mice within a dose group. This precluded an analysis of variation between laboratories.
- The lead laboratory established the dose levels to be tested by the other laboratories participating in the interlaboratory validation effort. In a minority opinion, Drs. Nathalie Alépée and Michael Woolhiser asserted that for an effective and efficient interlaboratory evaluation, it seemed reasonable to set dose levels for all laboratories based on results from the lead laboratory.

In addition, the Panel considered that the interlaboratory studies could benefit by performing more than one test on two commonly used positive controls (i.e., HCA and DNCB).

The draft LLNA: DA BRD contained an analysis of data from two interlaboratory reproducibility validation studies that used coded substances, as well as an intralaboratory accuracy validation study with 31 substances that were not coded. The Panel was asked if they considered the lack of coding of the test substances to have adversely impacted or biased the intralaboratory accuracy evaluation. The Panel commented that, in the validation of a new assay, it is better to avoid the potential for bias by testing coded substances.

However, the Panel concluded the data already generated for the LLNA: DA test method should be considered and not be rejected in the current validation evaluation.

The lead laboratory established the dose levels tested by the participating laboratories in the two interlaboratory validation studies. The Panel was asked if this adversely impacted or biased the evaluation. The Panel considered that the choice of the maximum test substance concentration is crucial for the proper performance of the traditional LLNA as well as any modified LLNA. Thus, predetermining the dose levels to be tested for each substance might have reduced variability between the two interlaboratory studies. In a minority opinion, Drs. Nathalie Alépée and Michael Woolhiser asserted that for an effective and efficient interlaboratory evaluation, it seemed reasonable to set dose levels for all laboratories based on results from the lead laboratory.

3.2.5 Data Quality

The studies evaluated in the draft BRD for the LLNA: DA were not conducted in accordance with GLP guidelines although they were reportedly done in laboratories that conduct GLP studies, and were conducted "in the spirit" of GLP (K. Idehara, personal communication). Furthermore, the original records for the interlaboratory studies were requested but have not yet been obtained. As a result, an independent audit could not be conducted to confirm that the reported data was the same as the data recorded in laboratory notebooks. The Panel was asked to comment on the potential impact this might have had on the evaluation of the LLNA: DA. The Panel commented that, ideally, GLP compliance is recommended for validation studies, but the current studies should not be rejected based on the lack of GLP compliance alone. However, all the raw data obtained through the validation process should be made available and audited for accuracy. The Panel further commented that since the original records for the interlaboratory studies have not yet been provided, recommendations from ICCVAM should be contingent upon receiving these data. Obtaining original laboratory records is a necessary step to confirm that all data generated during the validation studies have been provided, and that the reported data are the same as the data recorded in laboratory notebooks.

3.2.6 Consideration of All Available Data and Relevant Information

The Panel was asked to comment on whether all of the relevant data identified in published or unpublished studies that employed the LLNA: DA had been adequately compared. The Panel viewed that, generally, it seemed that all of the relevant results had been adequately identified and considered. However, as mentioned above, all of the original data supporting these results have not been provided. The Panel again expressed concern related to the effect of pretreating the mouse ear with 1% SLS and the Panel therefore recommended that the results from van Och et al. (2000) and De Jong et al. (2002) should be considered.

3.3 Comments on the Draft ICCVAM Test Method Recommendations on the LLNA: DA

3.3.1 Test Method Usefulness and Limitations

The Panel was asked to comment on whether the available data supported the ICCVAM draft recommendations for the LLNA: DA procedure in terms of the proposed test method usefulness and limitations. The Panel agreed with ICCVAM's recommendation, which stated that the LLNA: DA might be useful for identifying substances as potential skin sensitizers

and non-sensitizers, but this recommendation was contingent upon the receipt of additional data and information. The Panel further added that information on the possibility of skin reactions suggestive of the onset of the elicitation phase and the impact of the 1% SLS pretreatment step on the performance of the LLNA: DA should be evaluated. The Panel also considered that the ICCVAM proposed limitations needed to be more clearly defined, as it was not clear from the draft recommendations what points were considered as limitations. For instance, limitations that are known for the traditional LLNA would likely apply to this modified protocol as well and these should be noted.

The Panel was asked whether restrictions on using radioactive materials would warrant that the LLNA: DA be routinely recommended for hazard identification of skin sensitizing substances in lieu of having to possibly use guinea pig tests. The Panel noted that, based on gaps in the currently available dataset and information described in this report, the LLNA: DA could not yet be recommended for the routine use for hazard identification of skin sensitizing substances, regardless of whether restrictions on using radioactive materials were present or not. Generally, non-radioactive LLNA test methods are preferred in lieu of using guinea pig tests because fewer animals are used and animal pain and distress is reduced. However, policy issues regarding restrictions on radioactivity should have no impact on this science-based conclusion.

The Panel was asked if, from a public health perspective, the recommended guidance for evaluating negatives were sufficient to address concerns associated with the false negative rate of 5% (1/19 substances) calculated for the LLNA: DA. The Panel noted that this was not a scientific question, rather a risk characterization issue, and could not be answered without considering other factors such as intended use, target population, etc. The Panel was also asked if, from a testing strategy perspective, the ICCVAM guidance addressed concerns associated with the false positive rate of 10% (1/10 substances) calculated for the LLNA: DA and/or if they had other suggestions for additional guidance or limitations. The Panel again commented that this was not a scientific question but a risk characterization issue and could not be answered without considering other factors such as intended use, target population, etc. Furthermore, the Panel noted that it would be difficult to generalize the finding of one test substance being a "false" result. Instead, they considered it better to identify reasons why a substance was a "false" result. Certainly, if a "false" result is suspected, confirmatory testing with another mouse LLNA method was not recommended. It might be important to follow a suspected inaccuracy with an investigation of the mechanistic basis for the discordance.

3.3.2 Test Method Protocol

The Panel was asked if they agreed that the available data supported the ICCVAM draft recommendations for the LLNA: DA procedure in terms of the proposed test method standardized protocols or what recommendations they would make. The Panel noted that available data did not support all of the ICCVAM draft recommendations in the LLNA: DA standardized protocol. First, the ICCVAM protocol (ICCVAM 1999; Dean et al. 2001) for the traditional LLNA recommends using at least five animals per dose group. Although the Panel agreed that five animals per dose group should be recommended for validation studies, they suggested that power calculations would be useful in determining if subsequent use of the modified test method could use fewer animals per dose group. For the LLNA: DA test method, the Panel noted that based on statistical power calculations that were provided as

supplemental information, using four animals per group instead of five did not appear to be a limitation (i.e., detecting a 3.0-fold increase in the SI with four animals per group was estimated to have a 99% confidence level). In addition, the Panel generally agreed with the recommendation in the ICCVAM protocol (ICCVAM 1999; Dean et al. 2001) that individual animal data should be collected. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable.

Of greater importance, the Panel concluded that pretreatment with 1% SLS should not be accepted until its impact on the performance of the LLNA: DA has been adequately characterized. Although used at a concentration below its EC3, the Panel was concerned about pretreating the mouse ear with an irritant reported as positive in the traditional LLNA. To demonstrate that these concerns are not justified, the Panel recommended that substances that are weak irritants and weak sensitizers be tested in the LLNA: DA with and without pretreatment with 1% SLS. It also needed to be demonstrated that the 1% SLS pretreatments, as well as the additional test substance treatment on day 7, did not induce a skin reaction that could be indicative of the onset of the elicitation phase of skin sensitization.

The Panel was asked to comment on whether the limit dose procedure could be applied to the LLNA: DA. The Panel concluded that if the limit dose procedure is considered applicable to the traditional LLNA, then it should also be applicable to the LLNA: DA, in order to reduce the number of animals used. This would require adherence to the LLNA: DA test method protocol, with the exception that the middle and low dose groups would be omitted in the limit dose version.

3.3.3 Future Studies

The Panel was asked if they agreed that the available data support the ICCVAM draft recommendations for the LLNA: DA in terms of the proposed future studies or, if not, what recommendations they would make. The Panel stated that the available data supported the ICCVAM draft recommendations for the LLNA: DA in terms of a more comprehensive evaluation using more non-sensitizers within and across laboratories. In a minority opinion, Dr. Thomas Gebel stated that although testing of more non-sensitizers might be warranted for interlaboratory validation studies, a sufficient number of non-sensitizers had been tested within the same laboratory (Table 6.3 in the draft BRD).

However, the Panel viewed that there were additional studies that ICCVAM might consider. As previously mentioned, the Panel recommended that the 1% SLS pretreatment step should not be accepted until its impact on the performance of the LLNA: DA had been adequately characterized. Furthermore, it should be demonstrated that such pretreatments did not induce a skin reaction that could be indicative of the onset of the elicitation phase of skin sensitization. It might also be of interest to evaluate ATP as a marker of lymph node proliferation using the traditional LLNA dosing scheme and lymph node collection schedule. Lastly, the Panel considered that studies on the reliability of outlier analysis in small sample sizes and the effects of reduced sample size on the power of the LLNA: DA test method should be proposed.

3.3.4 Performance Standards

The draft LLNA: DA BRD indicated that the LLNA: DA protocol differed from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. In addition, there are differences between the two protocols that relate to how and when the test substance is applied and when the lymph nodes are collected (Table 2-1 and Appendix A in the draft LLNA: DA BRD). According to the proposed draft ICCVAM LLNA Performance Standards for the traditional LLNA, any change to the LLNA protocol other than the method used to assess lymphocyte proliferation would be considered a "major" change. The Panel was asked if they agreed that these should be considered "major" changes and therefore the usefulness and limitations of the LLNA: DA should not be assessed using the draft ICCVAM LLNA Performance Standards. The Panel commented that answering this question depended on having a clear definition of what constitutes a "major" versus a "minor" change, and what may constitute a different protocol altogether. Depending on the goal of the assay, whether a change is "major" versus "minor" may not be relevant. Ultimately, if a test method is able to make the correct prediction with regard to the dermal sensitization potential of a test substance, then the issue of "major" versus "minor" modifications might not apply. Considering the robust nature of the current draft ICCVAM LLNA Performance Standards, it is difficult to identify the need for additional requirements for methods like the LLNA: DA. Thus, the draft ICCVAM LLNA Performance Standards could be used to evaluate the LLNA: DA as a mechanistically and functionally similar test method.

The Panel was asked, even if the draft ICCVAM LLNA Performance Standards were not found applicable to the LLNA: DA, whether an analysis based on 13 of the 18 proposed required reference substances in the performance standards would impact the overall evaluation of the test method accuracy. The Panel commented that the accuracy analysis based on 13 of the 18 proposed required reference substances in the performance standards (with one false negative substance) should have no impact on the overall evaluation of test method accuracy as 31 substances have been tested. However, given the concern regarding pretreatment with 1% SLS, the Panel stated that testing of substances with and without 1% SLS was needed to characterize the effect of this pretreatment on the performance of the assay. The Panel concluded that as described above, the idea of "major" versus "minor" changes might be reconsidered, thus the current draft ICCVAM LLNA Performance Standards could be applicable to the LLNA: DA as a mechanistically and functionally similar test method.

[This Page Intentionally Left Blank]

4.0 Non-Radioactive LLNA Protocol – The LLNA: Bromodeoxyuridine Detected by Flow Cytometry (BrdU-FC) Test Method

4.1 Comments on the Draft BRD for Completeness, Errors, and Omissions

4.1.1 General Comments

The Panel was asked if there were any errors in the draft LLNA: BrdU-FC BRD that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. The Panel noted that overall, errors and omissions in the draft LLNA: BrdU-FC BRD were few. The majority of omissions relating to the data records were identified in the text, and all reasonable efforts to obtain additional information from MB Research Labs, the developer of the LLNA: BrdU-FC, appear to have been made.

The following describes the identified errors, omissions, and/or information gaps in the draft LLNA: BrdU-FC BRD that should be addressed:

- Data are available in the peer-reviewed literature on the application of BrdU in the LLNA with histochemical or enzyme-linked immunosorbent assay (ELISA) detection. This could be briefly mentioned in the final BRD for the LLNA: BrdU-FC method, simply as a means of indicating the utility of non-radiolabeled tracer methods in the LLNA.
- It should be noted that a potential reason why nickel chloride was negative in the LLNA: BrdU-FC may be due to oral tolerance in the mice that was induced by nickel-containing nipples of drinking bottles and nickel cages (Van Hoogstraten et al. 1993).
- The vehicle(s) used with the test substances should be stated.
- Information on experience of the inter-laboratory transferability of other technologies that depend upon flow cytometry technology as the key data read-out should be included.
- All raw data for the LLNA: BrdU-FC and the enhanced LLNA (eLLNA): BrdU-FC should be made available.

4.1.2 Comments with Specific References to the Text

The Panel stated that the following comments and/or suggested corrections relevant to specific parts of the draft LLNA: BrdU-FC BRD text should be addressed:

- Line 226: Citation was made to a reference dated 2001 by MB Research Labs which established their development of the LLNA: BrdU-FC; however, no reference was included in Section 12.0 (References) of the draft BRD.
- Lines 232-233: For a sensitizer, the SI should be greater than *or equal to* three.
- Line 246: “i.e., positive” should be explained.
- Line 254: 11% should be 17%.

- Line 263: For purposes of completeness, it may be worthwhile to add a brief description of the comparative accuracy of the available traditional LLNA, LLNA: BrdU-FC, and eLLNA: BrdU-FC results versus human maximization/patch test data. This information is of importance and displayed in Tables 6-1 and 6-2.
- Lines 286-288: The issue of the refinement/reduction in animal use that might follow the availability of a scientifically validated non-radioactive variant of the LLNA was mentioned in the draft BRD as a benefit but it was not quantified, and no authoritative reference was cited in support.
- Line 288: The final LLNA: BrdU-FC BRD should explain why the BrdU method would result in less pain and distress to the animals (i.e., does the route of injection of BrdU vs. ^{3}H-methyl thymidine produce less discomfort?).
- Line 335: Reference was made to a citation dated 2001 by MB Research Labs which established their development of the LLNA: BrdU-FC; however, no citation was included in Section 12.0 (References) of the draft BRD.
- Line 356: The sentence starting "To evaluate excessive skin…" implies that evaluation of excessive skin irritation by measuring ear thickness is recommended by the ICCVAM LLNA protocol although it is only recommended in the LLNA: BrdU-FC protocol.
- Lines 365-366: Consider supplementing the list of abbreviations for Figure 2-1 with B220+, B:T, CD69+, and IAk+. Also, the figure shows I-Ak+ while all other text uses IAk+. The MB Research Labs protocol shows I-Ak+.
- Figure 2-1: Should be redrawn to show the SI decision point lines coming off of the "Analyze Proliferating LNC (lymph node cells)" box rather than the "Inject BrdU and Excise…" box.
- Lines 500-503: Classification of "equivocal results" was unclear without data comparison (i.e., benzocaine produced divergent results in both tests). Were these results unlike what was expected from human data? What were the data for salicylic acid and mercaptobenzothiazole?
- Lines 552-558: It may be useful for comparative purposes to add summary accuracy data for the traditional LLNA versus human maximization/patch test data from the larger data set reported in the 1999 ICCVAM LLNA report to the section of the final BRD which discusses performance of the LLNA: BrdU-FC method.
- Table 6-3: Benzocaine was missing and it seems that salicylic acid was the same in both traditional LLNA and LLNA: BrdU-FC. Mercaptobenzothiazole was not reported for the LLNA: BrdU-FC. This needs to be corrected or explained.
- Table 6-5: The human outcome for benzalkonium chloride and ethylene glycol methacrylate should be negative. See also Table 6-6.

4.2 Comments on the Validation Status of the LLNA: BrdU-FC

4.2.1 Test Method Protocol

The LLNA: BrdU-FC protocol includes routine measurements of ear swelling as an indicator of excessive dermal irritation. The Panel was asked if they considered this procedure to be an appropriate approach and if this measurement should be recommended for routine inclusion into all LLNA protocols. The Panel stated that, as a quantitative parameter associated with inflammatory cell influx and fluid retention near the site of test substance application, ear swelling (or other quantitative measurements) should be carefully considered for inclusion into all LLNA protocols. This might assist in differentiating between sensitizers and irritants, assist in the interpretation of equivocal results, and possibly detect other procedure-related problems that might require further exploration/consideration.

The LLNA: BrdU-FC protocol also includes optional quantification of immunophenotypic markers as an additional mechanism for distinguishing irritants from sensitizers. The Panel was asked if they considered this to be an appropriate approach to reduce false positives, and if the correct markers were being considered. The Panel was also asked if these measurements should be recommended for routine inclusion in the LLNA: BrdU-FC. The Panel agreed that the use of immunological markers would be appropriate for detailed studies, as it might reduce the frequency of false positives (irritants) and improve comparisons with human data. However, since the primary use of the LLNA is for discrimination of human hazard from direct chemical contact, it could be argued that some false positives are acceptable (especially for methods which have relatively lower rates of false negatives). Given this dominant use, application of immunological markers would likely be too detailed and costly for routine LLNA use. Thus, the Panel suggested that results of ear swelling measurements be compared with the more technically complex flow cytometry markers to determine if similar results might be obtained. Furthermore, alternative immunological markers for discriminating between irritants and sensitizers may be available, although the draft LLNA: BrdU-FC BRD did not contain information allowing any informed decision on whether other markers might be more predictive. Thus, based on current knowledge, the current markers suggested in the draft LLNA: BrdU-FC BRD seemed acceptable to the Panel. Two other possibilities suggested were a surface marker relating to CD4 T-helper cells (Th) or Th1 cells (interferon-γ).

The Panel was also asked to comment on the appropriateness of the "sequential strategy" used in the eLLNA: BrdU-FC (see Figure 2-1 of the draft LLNA: BrdU-FC BRD). Generally, the Panel viewed that incorporation of immune parameters improved the value of a predictive assay and may also help explain mechanisms, which is important. Still, the "sequential strategy" used in the eLLNA: BrdU-FC for discriminating irritation from sensitization might be more sensible for research studies because of resource and cost considerations, and may not be appropriate for routine use of the LLNA in hazard identification. For human hazard detection, more simplified methods should be available for discrimination of irritants.

4.2.2 Substances Used for the Validation Studies

The Panel was asked to consider if the substances tested in the LLNA: BrdU-FC were representative of a sufficient range of chemical classes and physicochemical properties such

that the test method would be applicable to any of the types of chemicals and products that are typically tested for skin sensitization potential. The Panel agreed that if the proviso that the applicability domain limitations published for the traditional LLNA remained in force, the substances tested in the LLNA: BrdU-FC seemed representative of a sufficient range of chemical classes and physical chemical properties and it would likely be applicable to many of the types of chemicals and products that are typically tested for skin sensitization potential. However, the available LLNA: BrdU-FC database was relatively small compared to the large number of substances assessed in the traditional LLNA and this implied some caution in assuming that assay performance was concordant with the traditional LLNA.

4.2.3 Test Method Accuracy

The accuracy analysis presented in the draft LLNA: BrdU-FC BRD was based on overall concordance with the traditional LLNA. Accuracy statistics compared to the guinea pig tests and human data/experience were also provided. The Panel was asked if these comparisons were appropriate for assessing the accuracy of the LLNA: BrdU-FC. The Panel viewed that since the traditional LLNA is used to provide human hazard identification and information relevant to human health, the accuracy statistics compared to human data/experience were important. Since the LLNA: BrdU-FC is fairly similar to the traditional LLNA, guinea pig comparisons might not have been necessary. However, taken together, the availability of both human data/experience and guinea pig data allowed additional insights that might have expanded the applicability domain of the LLNA: BrdU-FC, or indicated improved performance with respect to LLNA false negatives and positives.

The Panel was then asked if the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: BrdU-FC had been adequately evaluated and compared to the traditional LLNA. The Panel agreed that the relevance of the LLNA: BrdU-FC was adequately evaluated and compared to the traditional LLNA, and supported the inclusion of accuracy analyses with and without equivocal materials.

Three substances (benzalkonium chloride, resorcinol, and Tween 80) produced a false positive response compared to the traditional LLNA and guinea pig test when tested using the LLNA: BrdU-FC (based on immunophenotyping, benzalkonium chloride was subsequently classified as an irritant rather than a sensitizer). The Panel was asked if they could identify any characteristics associated with these or similar substances that might suggest that using the LLNA: BrdU-FC to test such substances would not be appropriate or that positive results for substances with such properties may warrant additional testing. Overall, the Panel stated that there were not any patterns or unifying concepts that explained the three false positive results in the available data set. They noted that only a single laboratory is using the LLNA: BrdU-FC method and recommended that the raw data on which the reports were prepared be made available in order to allow further investigation. The Panel also suggested that additional studies be conducted to determine whether LLNA: BrdU-FC results with these three substances are repeatable.

Dr. Raymond Pieters stated that benzalkonium chloride and Tween 80 are considered aggressive irritants, but both published data (Manetz and Meade 1999; Varani et al. 2008) and unpublished data from his laboratory has shown that benzalkonium chloride (5%) is more potent than SLS in the stimulation of lymph node cell proliferation and may therefore may actually be considered a sensitizer. However, in the traditional LLNA these compounds

did not increase the SI above the threshold for a positive response (i.e., SI ≥3), so they were identified as non-sensitizers.

4.2.4 Test Method Reliability

The Panel was asked if the intralaboratory reproducibility of the LLNA: BrdU-FC had been adequately evaluated and compared to the traditional LLNA and if any limitations were apparent based on this assessment. The draft LLNA: BrdU-FC BRD analyzed data from repeat testing of HCA in six different vehicles and intralaboratory reproducibility was assessed by a coefficient of variation (CV) evaluation. The calculated CVs ranged from 30% to 53%. The Panel agreed that the relatively large SD and associated CV values raised questions about the extent of experiment-to-experiment variability. There was less concern about vehicle choice and effects on the range of group means than about the CVs greater than 50% for the group means of HCA tested in DMSO and AOO. The large number of repeated experiments for these tests would have been expected to dramatically reduce variability. The Panel concluded that the results suggested that key elements of assay standardization were not yet developed. Further evaluation using other positive control substances would have been valuable to more adequately characterize reproducibility.

4.2.5 Data Quality

The studies evaluated in the draft BRD for the LLNA: BrdU-FC were not all conducted in accordance with GLP guidelines although they were done in a laboratory that routinely conducts GLP studies (G. DeGeorge, personal communication). The Panel was asked to discuss what impact this might have on the evaluation of the LLNA: BrdU-FC. The Panel considered that, even without formal GLP compliance, the current LLNA: BrdU-FC results appeared to reflect a sincere attempt to perform work of high quality. The only area in which a lack of full GLP compliance may have been a source of assay variability was in the quantitative analysis of dosing solutions. For instance, failure to appreciate differences in composition of dosing solutions between experiments caused by test article instability or other phenomena may account for the relative large variability in intralaboratory data and possibly of some of the discordant results (i.e., false negatives and differences in LLNA: BrdU-FC results between repeat studies for the same substance). Thus, the Panel viewed that any additional studies undertaken to validate the test method should ideally be GLP-compliant.

Furthermore, the original records for these studies were requested but had not yet been obtained at the time of the Panel review. As a result, an independent audit could not be conducted to confirm that the reported data was the same as the data recorded in laboratory notebooks. The Panel was asked if they agreed that any recommendations from ICCVAM should be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription. The Panel agreed that, although a request for original data had been made, it was good practice to hold final recommendations until an independent audit could be performed. While it would be expected that no serious errors would be uncovered which would alter the current findings, an audit would confirm assay performance to date and position ICCVAM for further consideration of the LLNA: BrdU-FC.

4.2.6 Consideration of All Available Data and Relevant Information

The Panel was also asked if, based on the draft LLNA: BrdU-FC BRD, all the relevant data identified in published or unpublished studies that employ this test method had been adequately considered. Furthermore, they were asked that if there were other comparative test method data that were not considered in the draft BRD, how such data might be obtained. Overall, the Panel considered that all the relevant data identified in published or unpublished studies that employed this test method had been adequately considered in the draft LLNA: BrdU-FC BRD. However, some additional information was available in the peer-reviewed literature on application of BrdU in the LLNA with other methods of detection (e.g., histochemistry, ELISA). The Panel felt that these could have been briefly mentioned in the draft BRD for the LLNA: BrdU-FC method, simply as a means of indicating the utility of non-radiolabeled tracer methods in the LLNA. Furthermore, if an analysis of the CV for the traditional LLNA was undertaken, a more direct comparison with the LLNA: BrdU-FC could have been performed.

4.3 Comments on the Draft ICCVAM Test Method Recommendations on the LLNA: BrdU-FC

4.3.1 Test Method Usefulness and Limitations

The Panel was asked if they agreed that the available data and test method performance support the ICCVAM draft recommendations for the LLNA: BrdU-FC in terms of the proposed test method usefulness and limitations (i.e., that it may be useful for identifying substances as potential skin sensitizers and non-sensitizers, but that more information and data are needed before a recommended use of the LLNA: BrdU-FC can be made). The Panel agreed that the available data and test method performance of the LLNA: BrdU-FC support the draft ICCVAM recommendations. They considered the proposed test method usefulness and limitations to have well summarized the limits of the information supplied and the additional information that would need to be generated or provided for further consideration of this test method. As a result, the LLNA: BrdU-FC could not at this stage be considered scientifically validated as a replacement alternative to the traditional LLNA. Still, the test method recommendation should clearly state that the test method was not "invalid" but simply that there was currently not sufficient evidence and information to affirm that it had been adequately validated by ICCVAM. Instead, the Panel considered that the LLNA: BrdU-FC could be recommended in instances where mechanistic information about a sensitizer is required.

The Panel was asked if restrictions on using radioactive materials were or were not present, whether or not the LLNA: BrdU-FC should be routinely recommended for hazard identification of skin sensitizing substances in lieu of having to possibly use guinea pig tests. The Panel agreed that it is preferable to use alternative methods for the LLNA (i.e., ELISA detection of BrdU or histochemical detection of BrdU-labeled cells), as opposed to application of guinea pig test methods, if a limitation on radioisotope use exists (e.g., the lack of a radioactivity use license). This rationale is based on avoidance of the less quantitative guinea pig test methods, which may employ adjuvant treatment with associated animal stress and harm. Still, at this time, the Panel considered that data gaps in the LLNA: BrdU-FC method precluded recommending it for routine hazard identification of skin sensitizing

substances in lieu of the traditional LLNA, whether or not limitations on using radioactive materials exist. Policy issues regarding restrictions on radioactivity should have no impact on this science-based conclusion.

The Panel was asked if the ICCVAM recommendations adequately addressed concerns associated with the false positive rate of 17% (3/18 substances) calculated for the LLNA: BrdU-FC and if there were other suggestions for additional guidance or limitations that should be considered. The Panel agreed that the relatively high false positive rate was adequately identified and discussed, and that no mechanistic reason could be identified for these results based on available information. The Panel noted that it might be worthwhile to point out in the final BRD the impact on human health of false positive results versus false negative results in the context of hazard screening and identification. Consideration of factors such as intended use and target population of the false positive substances would further aid in characterizing human risk for these substances.

4.3.2 Test Method Protocol

The Panel was asked whether or not they agreed that the available data supported the ICCVAM draft recommendations for the LLNA: BrdU-FC procedure in terms of the proposed test method standardized protocol and if not, what recommendations would they make. The Panel noted that the draft ICCVAM recommendations for conduct of a standardized method for the LLNA: BrdU-FC variant were relatively brief and stated only that all applicable portions of the 1999 ICCVAM procedure be carefully followed. The Panel agreed that the available data supported the ICCVAM draft recommendations for the LLNA: BrdU-FC procedure in terms of adhering to the ICCVAM LLNA protocol (ICCVAM 1999; Dean et al. 2001). In particular, the Panel agreed that at least five animals per dose group should be used, particularly in light of Dr. Haseman's power analysis (see **Table 4-1**). The Panel did note however that power calculations could be undertaken to determine if fewer animals per dose group might be adequate for post-validation studies, though Dr. Haseman's power analyses suggest that this is unlikely. The majority of the Panel also agreed with the ICCVAM-recommended protocol to use individual animal data although a minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD TG 429 guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals per dose group could be considered acceptable.

Further, the Panel considered the methodological description of the LLNA: BrdU-FC procedure supplied by MB Research Labs (Appendix A to the draft BRD) to be comprehensive. The utility of ear swelling or other methods to detect inflammation/excessive local irritation appear to be warranted in every variation of the LLNA (including the traditional LLNA), but should be further investigated before routine inclusion in any protocol is recommended.

The Panel was asked whether the LLNA limit dose procedure could be applied to the LLNA: BrdU-FC. The Panel agreed that the LLNA limit dose procedure could be applied to the LLNA: BrdU-FC as long as the limitations associated with the limit dose procedure were appreciated. Furthermore, application of the limit dose procedure to the LLNA: BrdU-FC would require adherence to a validated LLNA: BrdU-FC test method protocol with the exception that the middle and low dose groups would be excluded. Furthermore, it would

need to be confirmed that the number/pattern of sensitizers that would have been identified/missed from the "high dose" group would mirror that of the traditional LLNA.

Table 4-1 Power Calculations for the LLNA: BrdU-FC

	3.0-fold Increase	2.5-fold Increase	2.0-fold Increase	1.5-fold Increase	1.3-fold Increase
Mean Rx response	30279	25232.5	20186	15139.5	13120.9
Log (Mean Rx response)	10.318	10.136	9.913	9.625	9.482
Difference from control (log scale)	1.098	0.916	0.693	0.405	0.262
Difference/SD	1.75	1.46	1.10	0.65	0.42
Power for N=5	80%	50-80%	<50%	<50%	<50%
Power for N=4	50-80%	50%	<50%	<50%	<50%
Power for N=3	50%	<50%	<50%	<50%	<50%
Other Power	95% (N=9)	95% (N=12)	95% (N=19)	95% (N=52)	95% (N>100)
Other Power	90% (N=7)	90% (N=10)	90% (N=15)	90% (N=42)	90% (N>100)

The power calculations above are based on a one-sided $p<0.05$ Student's t-test applied to log-transformed data from vehicle control LLNA: BrdU-FC tests.
Abbreviations: Fold-increase=Required increase above the vehicle control for a positive response; N=Number; Rx=Treatment; SD=Standard deviation.

4.3.3 Future Studies

Finally, the Panel was asked whether they agreed that the available data supported the ICCVAM draft recommendations for the LLNA: BrdU-FC in terms of the proposed future studies. The Panel agreed that the ICCVAM draft recommendations for future studies highlighted the unanswered questions raised by the available database. Specifically, conducting interlaboratory studies as a part of the validation process was considered important. As mentioned previously, the Panel viewed that the immunological markers suggested for the LLNA: BrdU-FC in the draft BRD were acceptable but that additional immunological markers for discrimination of irritant versus sensitization phenomena might also be identified. A suggestion for a future study was to use the surface marker relating to CD4 Th cells or internal marker relating to Th1 cells (interferon-γ).

In general, for any future work, the Panel considered that efforts should be made to decrease the variability and thereby increase the power of the test in order to ensure that more animals were not needed relative to the traditional LLNA or other alternative LLNA protocols. For instance, further optimization of the LLNA: BrdU-FC method should include kinetic studies to demonstrate that the optimal protocol was being used.

4.3.4 Performance Standards

The draft BRD indicated that the LLNA: BrdU-FC protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. According to the proposed draft ICCVAM LLNA Performance Standards for the traditional LLNA, any change to the LLNA protocol other than the method used to assess lymphocyte proliferation

was considered a "major" change. According to this criterion, the Panel considered the protocol differences between the LLNA: BrdU-FC and the traditional LLNA to be "minor" changes, and therefore considered that the validity of this test method could be based on the draft ICCVAM LLNA Performance Standards. However, the Panel also recognized that this depended on a clear definition of what constituted a "major" versus a "minor" change, or a different protocol altogether. Thus, further consideration of this topic could be addressed once the recommendations for the current draft ICCVAM LLNA Performance Standards were finalized. The Panel found it difficult to identify any additional requirements for methods like the LLNA: BrdU-FC.

Even if the draft ICCVAM LLNA Performance Standards were not found to apply to the LLNA: BrdU-FC, the Panel considered that the impact of the LLNA: BrdU-FC accuracy analysis based on 13 of the 18 proposed required reference substances in the draft ICCVAM LLNA Performance Standards should not have a major impact on the overall evaluation of test method accuracy, as 45 substances, representative of an appropriate range, were tested. However, based on consideration for development of LLNA performance standards, it would be desirable for validation purposes that the substances missing from the range of 18 standard materials be assessed in the LLNA: BrdU-FC protocol.

The draft LLNA: BrdU-FC BRD also indicated that three out of six sensitizers for which EC3 data were available had EC3 values that were outside of the proposed 0.5x to 2.0x EC3 acceptability range, which was developed based on the traditional LLNA. The Panel viewed that the primary concern seemed to have less to do with the variation in the response than with a concern that the range of response would skew the interpretation of any LLNA: BrdU-FC results used for sensitization potency estimates. Furthermore, it was not known if the same vehicle was used to derive both EC3 values/ranges. The proposed 0.5 x to 2.0 x range seemed to be based upon empirical/goodness of fit rather than any biological constant. The appropriateness of this range should be considered further when the finalized ICCVAM LLNA Performance Standards document is considered. In general, if the vehicles were different the question is irrelevant.

4.3.5 Comments with Specific References to the Text

The Panel made the following comment with specific reference to the text in the draft ICCVAM test method recommendations on the LLNA: BrdU-FC and suggested that it be addressed:

- Lines 39-41: For parallel construction of this sentence with the preceding sentence, suggest substituting the following "One of the other equivocal substances, salicylic acid, is one of the recommended reference standard materials used as a non-sensitizer in the draft ICCVAM LLNA Performance Standards…". This is based on the assumption that salicylic acid was the substance intended for discussion and that it was used in the draft ICCVAM LLNA Performance Standards as a model non-sensitizer.

[This Page Intentionally Left Blank]

5.0 Non-Radioactive LLNA Protocol - The LLNA: Bromodeoxyuridine Detected by ELISA (BrdU-ELISA) Test Method

5.1 Comments on the Draft BRD for Completeness, Errors, and Omissions

5.1.1 General Comments

The Panel was asked if there were any errors in the draft LLNA: BrdU-ELISA BRD that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. The Panel noted that, in general, all of the data included were relevant, and that it was apparent that considerable effort had been involved in carefully developing the comprehensive database. The Panel noted that they would have preferred to have the original papers by Dr. Takeyoshi included in the review materials, but they were easily retrieved from the journal websites. The Panel indicated that raw data (i.e., the actual optical density at 370 nm [OD_{370}] readings for the triplicates and the SD of the triplicates) are necessary for a thorough evaluation. Additionally, the Panel noted that only a relatively small number of substances had been tested in the LLNA: BrdU-ELISA.

When considering the animal welfare impact of implementing the LLNA: BrdU-ELISA, the Panel agreed that it would be less painful than guinea pig tests in those circumstances where the use of radioactive materials are restricted. Thus, the Panel agreed that the test represents a potential refinement. The Panel further stated that, if there is not an option to replace the guinea pig test with a non-animal test, decreasing the extent of pain and distress should be the first animal welfare priority. The Panel cautioned that at some point, however, the numbers of animals being utilized must be considered. An eventual recommendation that the LLNA: BrdU-ELISA be routinely used instead of guinea pig test methods where the use of radioactive substances are restricted would apparently require a significant increase in the number of mice killed per test (to increase the statistical power of the test method - see **Section 5.1.2** below) if an SI ≥1.3 is deemed the appropriate criterion to use for determining a positive response. The Panel stated that it would be helpful to know how many guinea pigs are currently being used nationally and internationally for skin sensitization tests, and how many mice would be used in the LLNA: BrdU-ELISA with the SI ≥1.3 criterion. Even an order of magnitude estimate would help the Panel judge whether the increase in numbers of mice needed is justified as the quest to relieve pain in guinea pigs is pursued.

5.1.2 General Statistical Comments

The Panel was concerned about using an SI of ≥1.3 to optimize the performance of the LLNA: BrdU-ELISA method. One Panel member's extensive experience with ELISA protocols was cited as evidence that the difference between the OD_{370} of the vehicle and the positive test at 1.3 would not likely be statistically significant. The Panel recommended that the raw data must be reviewed to evaluate this. In addition, based on Dr. Joseph Haseman's power analysis (see **Table 5-1**), the Panel stated that it was difficult to justify using a SI ≥1.3 as the decision criterion since it would result in a significant increase in the number of animals needed to obtain an acceptable confidence level. In this regard, the Panel

recommended that power calculations should be routinely recommended to ensure that the appropriate number of animals per dose group is being analyzed.

Table 5-1 Power Calculations for the LLNA: BrdU-ELISA

Parameter	3.0-fold Increase	2.0-fold Increase	1.3-fold Increase
Mean Rx response	0.399	0.266	0.173
Log (mean Rx response)	-0.92	-1.32	-1.75
Difference from control (log scale)	1.10	0.70	0.27
Difference/SD	3.64	2.32	0.89
Power for N=4	99%	80-90%	<50%
Other power	95% (N=3)	95% (N=5)	50% (N=8)
Other power	–	50-80% (N=3)	80% (N=16)
Other power	–	–	90% (N=22)

The power calculations above are based on a one-sided $p<0.05$ Student's t-test applied to log-transformed data from vehicle control LLNA: BrdU-ELISA tests.
Abbreviations: Fold-increase=Required increase above the vehicle control for a positive response; N=Number; Rx=Treatment; SD=Standard deviation.

5.1.3 Comments with Specific References to the Text

The Panel also identified the following minor formatting and grammatical errors, as well as information gaps, in the draft BRD:

- The Panel noted a discrepancy between the draft LLNA: BrdU-ELISA BRD and the draft ICCVAM LLNA Performance Standards in the vehicle used for testing 2-mercaptobenzothiazole. Table 6-2 of the draft LLNA: BrdU-ELISA BRD indicated that the vehicle was AOO but the revised draft ICCVAM LLNA Performance Standards indicated that the vehicle was DMF (see page C15, C22 of September 7, 2007, draft and page B-6 of January 7, 2008, revised draft). Additionally, Table 1 on page C-7 of the revised draft ICCVAM LLNA Performance Standards listed AOO as the vehicle for 2-mercaptobenzothiazole. For both vehicles, the revised draft ICCVAM LLNA Performance Standards indicated that the EC3 value is 2.5%, although the text on page 10 of the draft LLNA: BrdU-ELISA BRD stated "the NICEATM database of traditional LLNA studies indicates that 2-mercaptobenzothiazole has a higher EC3 value when tested in AOO (mean EC3=9.8%) compared with DMF (mean EC3=2.5%)..."
- The Panel noted that Table 6-1 of the draft LLNA: BrdU-ELISA BRD indicated that, when compared to the guinea pig and human test data, the sensitivity and specificity of the LLNA: BrdU-ELISA is lower than that of the traditional LLNA. In fact, depending on the SI threshold value used, the

sensitivity and specificity of the LLNA: BrdU-ELISA can be higher than that of the traditional LLNA.

The Panel recommended that the draft LLNA: BrdU-ELISA BRD be updated to rectify these errors and omissions.

5.2 Comments on the Validation Status of the LLNA: BrdU-ELISA

5.2.1 Test Method Protocol

The data generated for the substances analyzed in the LLNA: BrdU-ELISA test method came from auricular lymph nodes from four individual mice in each dose group. The ICCVAM-recommended LLNA protocol (ICCVAM 1999; Dean et al. 2001) and OECD TG 429 (OECD 2002) recommend a minimum of five animals per dose group when collecting individual animal data. The Panel was asked what impact might the use of four animals per dose group have on the accuracy and reliability of the LLNA: BrdU-ELISA, and if the Panel agreed with the ICCVAM recommendation that future use of this test method protocol should include five animals per dose group. The Panel majority agreed with the ICCVAM recommendation that future use of this test method should use five animals per dose group and collect individual animal data, as per the ICCVAM-recommended protocol. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable. Based on the supplemental data provided by Dr. Haseman, the power to detect a three-fold increase with a sample size of four was determined to be 99%. These calculations, however, assume that a sample size of four is always obtained. If a sample size of four was planned and fewer usable data values were obtained, then the experiment might be compromised. Furthermore, the Panel concluded that testing for and eliminating "outliers" from experiments with small sample sizes is questionable. A reduction in sample size from five to four was not recommended unless data was provided on the frequency with which "outliers" occurred and an analysis is performed that establishes that a reduction in the nominal sample size from five to four would not compromise the performance of the test method. The Panel stated that the handling of suspected "outliers" and the use of robust statistics are issues that need to be addressed in such an analysis. For example, robust procedures may compensate for apparent "outliers" and eliminate the impulse to discard data. An example is calculating the mean values used in the SI on a log scale and then exponentiating the result to construct the SI.

The Panel also indicated that it was important to routinely include a positive control group in test method validation experiments (e.g., HCA), which was likely not the case for most of the LLNA: BrdU-ELISA validation experiments. Although the Panel did not reach consensus, they did consider the suggestion that for laboratories in which the LLNA is "routinely" performed, positive controls (e.g., HCA or a substance that matches the chemical class of the test substances) could be run at intervals for quality control purposes rather than concurrent with each experiment in which substances are tested. The Panel also discussed that omitting the concurrent positive control should not be recommended for laboratories that perform the LLNA only "occasionally". In their discussions, the Panel was not able to conclude what should constitute "routine" or "occasional" LLNA use or what would be an appropriate interval between positive control testing when a concurrent positive control is not used.

5.2.2 Substances Used for the Validation Studies

The Panel was asked whether the LLNA: BrdU-ELISA database was representative of a sufficient range of chemical classes and physicochemical properties such that the test method would be applicable to any of the types of chemicals and products typically tested for skin sensitization potential. The Panel indicated that the ratio of solids to liquids was not comparable; more solids should be included. The Panel further indicated that more substances for which traditional LLNA data are available should be tested, and that compounds including metals (e.g., nickel, cobalt), mixtures, and substances in aqueous solutions should be included.

5.2.3 Test Method Accuracy

The current accuracy analysis using an SI ≥3.0 or SI ≥1.3 to identify sensitizers is based on overall concordance with the traditional LLNA. Accuracy statistics compared to the guinea pig tests and human data/experience were also provided to the Panel. The Panel was asked whether these comparisons were appropriate for assessing the accuracy of the LLNA: BrdU-ELISA. The Panel indicated that comparing the LLNA: BrdU-ELISA performance to the traditional LLNA and the guinea pig tests were appropriate. Comparisons between the LLNA: BrdU-ELISA and human data were considered particularly valuable because the traditional LLNA doesn't match human data with 100% accuracy. For this reason, the Panel considered comparing the performance of the LLNA: BrdU-ELISA with that of the traditional LLNA with respect to predicting the human outcomes to be the best method of comparing these two LLNA protocols. The Panel concluded that in moving forward with any test method recommendation, key importance should be placed on interpreting the test results and making them clinically applicable to humans.

Takeyoshi et al. (2007) performed an accuracy analysis using decision criteria other than an SI ≥3.0 to classify substances as sensitizers. Maximal accuracy for the LLNA: BrdU-ELISA occurred when an SI ≥1.3 was used to distinguish between sensitizers and non-sensitizers. Using this decision criterion, the LLNA: BrdU-ELISA achieved an accuracy of 91% (21/23), with a sensitivity of 100% (16/16) and a specificity of 71% (5/7) (i.e., there were no false negatives and two false positives). The Panel was asked whether this analysis supported a recommendation that the decision criteria be based on an SI ≥1.3, and if there were concerns with using such a small increase (i.e., 1.3-fold) above the vehicle control response as the basis for identifying a positive response. The Panel did not support using an SI ≥1.3 as the criterion for positive results. An SI=1.0 means there was no difference between the vehicle control and the test substance. An SI=1.3 represents a 30% increase from the vehicle control. The difference between the OD_{370} of the vehicle and the positive test at 1.3 may not be statistically significant. An SI=3.0, which represents a three-fold difference between the vehicle and a positive test, would be a more believable positive difference. If the positive test criteria must be reduced to 1.3, then the Panel questioned whether the protocol is useful in its current state. The supplemental information that provided power calculations indicated that it would not be realistic to expect to detect a 1.3 fold increase in the control response without a significant addition of animals. Although using SI ≥1.3 increases the accuracy of the test, it comes at an increased cost to animals, which merits consideration. Furthermore, the ICCVAM (1999) report stated that an irritating chemical might induce proliferation, but that the response seldom exceeds an SI ≥3.0 (page 6). The Panel concluded that this might

provide further justification against using a low SI (e.g., 1.3) as a threshold for a positive response.

The Panel was asked if the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: BrdU-ELISA, using the SI ≥3.0 criterion, had been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft LLNA: BrdU-ELISA BRD). If not, the Panel was asked what other analyses should be performed. The Panel agreed that the relevance of the LLNA: BrdU-ELISA, using the SI ≥3.0 criterion, had been adequately evaluated. The Panel further stated that a better evaluation could be performed, however, if the database for the LLNA: BrdU-ELISA included more substances with traditional LLNA, guinea pig, and human data. The Panel considered the false negative rate of the test method to be excessive when results are compared with that obtained in the traditional LLNA (29/33/27% for the various datasets) or with human data (39%) – the results should be at least comparable with the traditional LLNA.

Using the SI ≥3.0 criterion, there were four substances (aniline, 4-chloroaniline, 2-mercaptothiazole, and hydroxycitronellal) that produced false negative responses when tested using the LLNA: BrdU-ELISA compared to the traditional LLNA. The Panel was asked whether it could identify any characteristics associated with these or similar substances, compared to the correctly identified sensitizers, that might indicate that such substances should not be tested in the LLNA: BrdU-ELISA or that negative results for such substances should indicate a need for confirmatory testing. The Panel could not identify any characteristics associated with these substances that might allow the identification of these substances as false negatives prior to testing. The Panel stated that the LLNA: BrdU-ELISA test, using the standard SI ≥3.0 to indicate positive results, simply does not perform well for identifying sensitizers.

5.2.4 Test Method Reliability

The Panel was asked whether the intralaboratory reproducibility of the LLNA: BrdU-ELISA had been adequately evaluated and compared to the traditional LLNA, and whether any limitations were apparent based on this intralaboratory reproducibility assessment. The Panel indicated that the number of substances evaluated for intralaboratory reproducibility was too few and, in some cases, there was a wide variation in repeat test results for the same substance. Only six substances (five sensitizers and only one non-sensitizer) were tested multiple times. The non-sensitizer, propylene glycol, was tested only twice and opposite results were obtained. The Panel considered the results of an intralaboratory reproducibility evaluation that was based on two discordant results only to be unacceptable. The numbers calculated in Table 7-1 of the draft LLNA: BrdU-ELISA BRD are correct, but the Panel questioned the dependability of the data since only two to three values were available for calculating the mean and CV. The Panel considered the CV values (over 30%) high, compared to the traditional LLNA (draft LLNA: BrdU-ELISA BRD Tables 7-1 to 7-3). The Panel stated that at least four independent tests with three concentrations tested represent a solid basis for calculation. The Panel considered the number of tests for intralaboratory concordance analysis to be insufficient, and stated that more intralaboratory testing is needed. The Panel recommended an evaluation of the intralaboratory reproducibility of the EC ≥1.3 and that the analysis of the variability of the ECt be conducted on a log scale.

The substances evaluated for intralaboratory reproducibility of the LLNA: BrdU-ELISA study were not coded. The Panel was asked whether the lack of coding of test substances adversely impacts or biases the current evaluation. The Panel stated that, although coding of substances is preferred for independent testing and evaluation of test results, the current data should not be rejected from consideration because the substances tested were not coded.

The Japanese Center for Validation of Alternative Methods (JaCVAM) has implemented a multi-laboratory validation study of the LLNA: BrdU-ELISA. The Panel was asked whether the study design was appropriate to adequately determine the extent of interlaboratory reproducibility for the LLNA: BrdU-ELISA. If not, the Panel was asked what other studies should be performed. The Panel stated that they had insufficient time to evaluate the study design and that they could not evaluate interlaboratory reproducibility because the study data were not available at the time of their evaluation.

5.2.5 Data Quality

The studies evaluated in the draft BRD for the LLNA: BrdU-ELISA were not conducted in strict accordance with GLP guidelines, although there were reportedly performed in laboratories that conduct GLP studies (M. Takeyoshi, personal communication). In other words, an audit report was not available. Also, the raw data were unavailable for an independent audit. The Panel was asked to discuss what impact this might have on the evaluation of the LLNA: BrdU-ELISA. The Panel concluded that ideally, validation studies should be performed in accordance with GLP guidelines. Although the systems employed for tests (i.e., test facilities, staff, reagents, etc.) were identical to those for GLP-compliant studies, the data quality may be questioned and therefore should at least be available for a retrospective independent audit. However, in this case, the Panel concluded that the lack of GLP compliance was not likely the reason for the poor results obtained with the LLNA: BrdU-ELISA.

The original records for these studies were requested but had not been received by the time the Panel convened. As a result, an independent audit could not be conducted to confirm that the reported data in peer reviewed publications and a poster presentation is the same as the raw data. The Panel was asked whether any recommendations from ICCVAM should be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription. The Panel concluded that, to have confidence in data quality, ICCVAM recommendations should be contingent upon the completion of an independent audit. Moreover, if an SI $\geq$1.3 is used as the criterion for positive results, review of the raw data is necessary to confirm statistically significant differences. The Panel concluded that this test, as described, had poor accuracy, poor sensitivity, and poor specificity. The Panel stated that changing the SI decision criterion from 3.0 to improve test performance, especially to such a drastic change as SI $\geq$1.3, is a mistake and sets a dangerous precedent.

5.2.6 Consideration of All Available Data and Relevant Information

Based on the draft LLNA: BrdU-ELISA BRD, the Panel was asked whether all the relevant data identified in published or unpublished studies that employ this test method had been adequately considered, and if other comparative test method data that were not considered were available. If yes, the Panel was asked to suggest how to obtain such data. The Panel believed that all of the relevant data, with the exception of the interlaboratory reproducibility study, were presented and that they were not aware of any omissions.

5.3 Comments on the Draft ICCVAM Test Method Recommendations on the LLNA: BrdU-ELISA

5.3.1 Test Method Usefulness and Limitations

The Panel was asked whether the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA in terms of the proposed test method usefulness and limitations. The Panel agreed with the ICCVAM recommendation that the LLNA: BrdU-ELISA may be useful for identifying substances as potential skin sensitizers and non-sensitizers but that, at this time, more information and data are needed before a recommended use of the LLNA: BrdU-ELISA can be made. The Panel also stated that a detailed protocol is needed, in addition to sufficient quantitative data for a more comprehensive analysis based on a larger set of balanced reference substances with regard to physicochemical properties and sensitization potency, as well as an evaluation of interlaboratory reproducibility.

The Panel was asked whether the LLNA: BrdU-ELISA should be routinely recommended for hazard identification of skin-sensitizing substances in lieu of using guinea pig tests if restrictions on using radioactive materials *are* present, due to the fact that fewer animals might be used and because pain and distress would be avoided. The Panel stated that if the accuracy of the test method was at least similar to the traditional LLNA, the LLNA: BrdU-ELISA might be routinely recommended for hazard identification of skin-sensitizing substances in terms of reduction of animals and refinement of the pain and distress associated with guinea pig tests. Clearly, using the LLNA: BrdU-ELISA instead of the traditional LLNA or guinea pig test methods would also offer advantages for the environment due to the use of a non-radioactive probe chemical. However, the Panel stated that the accuracy of the current LLNA: BrdU-ELISA dataset at SI ≥3.0 was inadequate and not equivalent to the traditional LLNA. The Panel also noted that if an SI ≥1.3 was used because of its apparent increased accuracy, additional mice (over and above the number needed in the standard LLNA test) would apparently be needed (see **Table 5-1**). Thus, the Panel stated that reduction of animals would not be achieved. In this regard, the Panel noted that some quantification of the total animal use numbers would be useful as it is not clear whether the increased number of mice used would outweigh the avoidance of pain and distress in guinea pigs.

The Panel was asked whether the LLNA: BrdU-ELISA procedure or other valid and accepted non-radioactive method could be routinely recommended for hazard identification of skin sensitizing substances instead of the traditional LLNA if limitations in using radioactive materials *are not* present. The Panel stated that the LLNA: BrdU-ELISA procedure could not be routinely recommended for hazard identification of skin sensitizing substances instead of the traditional LLNA, because the accuracy of this test at SI ≥3.0 was inadequate. In other words, the current dataset available for the LLNA: BrdU-ELISA did not predict the guinea pig or human outcomes as accurately as the traditional LLNA. Thus, the Panel acknowledged that there is the possibility that additional data might impact on the accuracy statistics and eliminate this concern. The Panel stated that factors that weigh on a decision of replacement of the LLNA with a non-radioactive method would include:

- Are more animals needed?
- Is the replacement test safer and less complex?

- Is the replacement test more efficient?
- Is the replacement test less costly?

The Panel stated that additional factors to consider might exist, but overall recommended that whether or not restrictions on radioactivity exist, a test that causes the least pain and uses the fewest number animals should be preferred, as long as adequate test method performance is maintained. Clearly, policy issues regarding restrictions on radioactivity should have no impact on this science-based conclusion.

The Panel was asked whether using a decision criterion of SI ≥1.3 instead of SI ≥3.0 resolved any concerns with respect to potential false positives or false negatives that may occur in this test method. The Panel was also asked for other suggestions for additional guidance or limitations that should be considered. The Panel stated that using a decision criterion of SI ≥1.3 instead of SI ≥3.0 would not itself resolve any concerns; more raw data are needed for a broader set of reference positive and negative sensitizers, including metals, mixtures, and aqueous solutions. The Panel also stated that a detailed protocol is needed, as is an evaluation of interlaboratory reproducibility. The Panel considered the current database to be inadequate, but based on the limited database, concluded that it might be more appropriate to use a statistically based decision criteria than a stimulation index.

5.3.2 Test Method Protocol

The ICCVAM draft recommendations state that the LLNA: BrdU-ELISA protocol should adhere to the ICCVAM LLNA protocol (ICCVAM 1999; Dean et al. 2001), except for measurement of lymphocyte proliferation. The Panel was asked whether the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA procedure in terms of the proposed test method standardized protocols. In general, the Panel agreed that the available data support the ICCVAM draft recommendations for this test method in terms of the standardized protocol. As stated previously, the Panel majority agreed with the ICCVAM recommendation that future studies should use five animals per dose group and collect individual animal data, as per the ICCVAM-recommended protocol. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable. The Panel further noted that using an SI <3.0 would require more animals to achieve adequate statistical power (**Table 5-1**) and therefore any considerations of reducing the SI to improve test method accuracy should include this point.

The Panel was asked whether the traditional LLNA limit dose procedure could be applied to the LLNA: BrdU-ELISA. The Panel stated that, if the LLNA: BrdU-ELISA was considered equivalent to the traditional LLNA, then it would be appropriate to apply the LLNA limit dose procedure to this test method. The Panel explained that, as in the case of the traditional LLNA, the protocol would be the same except for testing the maximum dose only, so applying the limit dose procedure would appear to have the same opportunity to reduce the number of animals needed to perform the test. However, using an SI ≥3.0 would not be appropriate because of the associated low accuracy in identifying sensitizers.

5.3.3 Future Studies

The Panel was asked whether the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA in terms of the proposed future studies. The Panel stated that the proposed future studies were justified. The Panel concluded that it is important to consider non-radioactive methods because, in some laboratories, it is difficult or not permissible to use radioactivity. The Panel also stated that, if more data were available and there was less variability in this test method, it might warrant re-evaluation. The Panel concluded that more data are needed, especially for determination of the appropriate threshold value for the decision criterion, and that interlaboratory reproducibility should be also evaluated (which presumably will occur once the Japanese interlaboratory validation effort is complete).

5.3.4 Performance Standards

The LLNA: BrdU-ELISA protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. According to the proposed draft ICCVAM LLNA Performance Standards for the traditional LLNA, any change to the LLNA protocol other than the method used to assess lymphocyte proliferation is considered a "major" change. The Panel was asked whether protocol differences between the LLNA: BrdU-ELISA and the traditional LLNA should be considered only "minor" changes and therefore if the validity of this test method should be based on the draft ICCVAM LLNA Performance Standards. In general, the Panel agreed that the LLNA: BrdU-ELISA protocol differs only in the method used to assess lymphocyte proliferation. Thus, based on the current draft ICCVAM LLNA Performance Standards, it should be considered as having only "minor" changes and therefore the validity of this test method could be based only on the draft ICCVAM LLNA Performance Standards.

However, the Panel concluded also that the answer to this question might differ depending on what the draft ICCVAM LLNA Performance Standards ultimately constitutes as a "major" change, a "minor" change, or a different protocol altogether. The Panel further stated that, depending on the goal of the assay, these distinctions may not be relevant. Ultimately, if a test method is able to make the correct prediction with regard to the sensitization potential of a test substance, then the issue of "major" versus "minor" changes in the protocol should not apply.

The Panel was asked, even if the draft ICCVAM LLNA Performance Standards do not apply to the LLNA: BrdU-ELISA, what impact should the accuracy analysis based on eight of the 18 proposed required reference substances in the draft ICCVAM LLNA Performance Standards (only one false negative and no false positives) have on the overall evaluation of test method accuracy. The Panel concluded that the accuracy analysis based only on eight of the 18 proposed required reference substances had a significant impact on the evaluation of test method accuracy. The low number of experiments provided data that resulted in unacceptable test method performance.

The Panel was asked whether there were concerns that 4/4 sensitizers, for which EC3 data were available, had EC3 values that were outside of the proposed recommended 0.5x to 2.0x EC3 acceptability range developed based on the traditional LLNA. The Panel concluded that

the EC3 values outside the recommended 0.5x to 2.0x EC3 acceptability range raised concerns related to test reproducibility and reliability.[2]

The Panel was asked whether separate performance standards should be developed for the LLNA: BrdU-ELISA. The Panel concluded that separate performance standards for the LLNA: BrdU-ELISA were not needed because the test principles are identical to the traditional LLNA.

[2] During their public meeting on March 4-6, 2008, the Panel's discussion, conclusions and recommendations on the LLNA: BrdU-ELISA took place prior to the discussion, conclusions and recommendations on the draft ICCVAM LLNA performance standards. Following their discussion of the draft LLNA performance standards, the Panel concluded that that an evaluation of test method accuracy should be based on overall accuracy statistics when compared to the traditional LLNA, and not on a chemical-by-chemical match that is based on obtaining an EC3 value within a specified range of EC3 values.

6.0 Draft ICCVAM LLNA Performance Standards

6.1 Comments on the Proposed Purpose and Applicability

ICCVAM proposed that these performance standards should only be applicable to versions of the LLNA that incorporate "minor" modifications to the traditional LLNA. Currently, *this is limited to the use of non-radioactive reagents to measure lymphocyte proliferation*. It is considered essential that the modified LLNA should otherwise adhere to all other aspects of the traditional LLNA protocol, as defined by ICCVAM (1999) and Dean et al. (2001). This includes aspects such as: the sex and strain of mouse used, the number of mice per dose group, the timing and site of test article treatment, the duration between the last treatment and lymph node collection, the inclusion of concurrent negative and positive control groups, the measured endpoint (i.e., lymphocyte proliferation in the draining auricular lymph node), and the collection of data at the level of the individual mouse. The Panel was asked if they agreed that the use of non-radioactive reagents for measuring cell proliferation in the lymph nodes constitutes a "minor" modification to the traditional LLNA protocol.

The Panel noted that the draft ICCVAM LLNA Performance Standards are proposed for evaluating the acceptability of test methods that are functionally and mechanistically similar to the traditional LLNA (i.e., measuring the same biological effect), and understood that ICCVAM proposed that these performance standards should only be applicable to protocols that incorporate "minor" modifications to the traditional LLNA, as defined above. The Panel unanimously agreed that based on ICCVAM's definition, the use of non-radioactive reagents for measuring cell proliferation is a "minor" modification of the traditional LLNA protocol. However, the Panel also agreed that other modifications may be considered "minor" and that a better strategy for the performance standards might be to define criteria that need to be satisfied to insure that the method is mechanistically and functionally similar (see criteria listed under essential test method components). Examples of potentially acceptable modifications identified by the Panel include sex, strain, the use of rats rather than mice, number of animals per group, and timing of test article treatment.

Regardless of the modification, the Panel stated that the modified test method should be designed to measure only the induction phase of the immune response. This is crucial, since the traditional LLNA is intended for hazard identification with the underlying principle that stimulation of cell proliferation in the draining lymph node suggests that sensitization (i.e., induction) is occurring. Using only the induction phase as the method to identify hazardous substances involves a short time frame, and reduces pain and distress in treated animals (i.e., no dermatitis response). Furthermore, the Panel stated that the performance standards should not imply that the traditional LLNA, or any alternative LLNA protocol, is capable of specifically distinguishing a type IV hypersensitivity reaction (as might be inferred from the text beginning with line 342 of the draft ICCVAM LLNA Performance Standards document). Therefore, reference to type IV hypersensitivity reaction should be removed from the document.

The Panel was asked if they considered it necessary that a modified LLNA keep the same decision criteria for distinguishing between sensitizers and non-sensitizers as the traditional LLNA (i.e., an SI $\geq$3.0). The Panel considered it unnecessary for a modified LLNA to keep this same decision criteria as a different method for measuring cellular proliferation might

have better concordance with the human data at a SI different than 3.0. Thus, with any modified LLNA, the SI threshold defining a sensitizer would need to be established (i.e., it is important to consider if the results are biologically relevant to humans).

The Panel was asked if other procedural modifications could be identified as "minor", based on the description in the draft ICCVAM LLNA Performance Standards document, and therefore could be evaluated for equivalence to the traditional LLNA using the proposed performance standards. The Panel reiterated that sex, strain, the use of rats rather than mice, animals per group, and timing of test article treatment are also potentially "minor" modifications. Furthermore, the proposed performance standards appear robust; therefore, regardless of the modification (i.e., "major" or "minor"), there is the same expectation for test method performance. Dr. James McDougal offered a minority opinion to express his concern about the potential impact that allowing alternative LLNA protocols with modifications other than the method by which lymphocyte proliferation was measured would have.

The Panel was asked if they considered the draft ICCVAM LLNA Performance Standards applicable to the LLNA limit dose procedure. The Panel noted that the current draft ICCVAM LLNA Performance Standards could be applicable to the LLNA limit dose procedure as long as it is recognized that this procedure can only be used for a yes/no hazard classification (i.e., an ECt estimate is not feasible).

6.2 Comments on the Essential Test Method Components

The essential test method components are based on the ICCVAM-recommended protocol (ICCVAM 1999; Dean et al. 2001), which is the basis for the current EPA (2003) test guideline. There are some notable differences between these protocols and OECD TG 429 (OECD 2002) for the LLNA. The Panel was asked to comment on, when evaluations of non-radioactive versions of the traditional LLNA are conducted using these performance standards, whether it is necessary that the validation studies follow the ICCVAM-recommended protocol. The Panel indicated that ideally, there would be one globally recognized set of performance standards (ICCVAM, ECVAM, JaCVAM). However, when validating versions of the traditional LLNA where the only difference is in the use of a non-radioactive method to measure cell proliferation, the ICCVAM-recommended protocol should be used. If more extensive changes to the protocol are being considered, the following requirements should be considered during modifications of the LLNA:

- Application of the test substance should be to the skin, with sampling of the lymph nodes draining that site.
- Cell proliferation should be measured in the draining lymph node.
- No skin reaction should be present, since presence of a skin reaction might indicate the onset of the elicitation phase of skin sensitization.
- Data should be collected at the level of the individual animal to allow for an estimate of the variance within control and treatment groups. Using this variance, a power analysis needs to be conducted to demonstrate that the modified method is utilizing a sufficient number of animals per treatment group to permit hazard identification with at least 95% power.

- If dose response information is needed, there should be an adequate number of dose groups (n ≥3) with which to adequately characterize the dose response for a given test substance.

The Panel was asked to comment on whether validation studies should include a concurrent positive control with each test substance and if so, whether the concurrent testing of the positive control and test substance should be conducted in the same vehicle or if different vehicles were acceptable. The Panel noted that a concurrent positive control should be included in each validation study to ensure that the test system was operating as expected and technical errors were not occurring. A concurrent positive control would be especially useful when an unknown test material was being tested or when a laboratory was collecting a dataset to serve as historical control data. However, if a known sensitizer was being tested, a concurrent positive control might not be needed, thus reducing animal use. Finally, the Panel concluded that the positive control should be tested in the same vehicle as the test substance. Using a different vehicle for the positive control would require an additional set of vehicle control animals.

The Panel was also asked whether the validation studies should use a minimum of five animals per dose group and collect lymph node data from individual animals. The Panel commented that until sufficient data were collected to enable a reliable power calculation to be conducted to determine the optimal number of animals per dose group, at least five animals per dose group should be used. The Panel also agreed that when validating a modified LLNA protocol, lymph node proliferation should be evaluated at the level of the individual animal within each dose group. Variance is only measurable if lymph nodes from individual animals are assessed. If the variability within a dose group of a modified LLNA protocol was substantially less than the traditional LLNA, reducing the number of animals per dose group might yield similar results. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable.

6.3 Comments on the Proposed Reference Substances

The Panel was asked if they agreed with the selection and prioritization criteria used to select the performance standards reference substances. The Panel noted that the rationale for selection of the reference substances included in the draft ICCVAM LLNA Performance Standards was well documented (taking into account the physicochemical characteristics, the purity, the stability, the quality of the *in vivo* data, and the chemical classes covered). The substances also appeared to be distributed over a wide range of EC3 values. However, the available database for some of the substances was insufficient. Among the 13 sensitizers in the "required" list, only five appear to have a robust database (i.e., have been tested in at least three independent studies). Thus, consideration should be given to revising the list of substances and/or making the data for the substances on the current list more robust. Ideally, the reference list should be based only on substances with robust data for LLNA, human, and guinea pig tests.

The rationale for the number of substances included on the "required" list of substances (n=18) was provided in the draft ICCVAM LLNA Performance Standards. In addition, there

were four additional substances that were described as problematic in the traditional LLNA (i.e., false negatives and false positives). The Panel was asked if they considered 18 "required" substances to be an adequate number upon which to evaluate the performance of non-radioactive LLNA test methods, where the only protocol modification is the method for assessing cell proliferation in the auricular lymph nodes, and if not, how many reference chemicals should be tested. The Panel commented that ideally, one would like to be able to demonstrate that an assay is equivalent to the traditional LLNA. However, with the small number of reference substances available, establishing equivalence will be extremely difficult. Therefore, the Panel recommended that, for use in hazard identification, a modified method should be evaluated with all 22 substances (including false negatives and false positives) and accuracy statistics calculated. To the extent possible, rationale for any discordant results should be provided, but the most potent sensitizers (e.g., DNCB) should always be identifiable. There also should be considerable weight given to the balance between animal welfare and human safety when considering the adequacy of test method accuracy.

The panel considered it noteworthy that 19 of the 22 substances on the draft ICCVAM list are in common with the ECVAM performance standards list. The Panel also considered it important that substances be coded during validation studies.

It is also relevant to note that the Panel discussed the value of GLP procedures on several occasions during the meeting. In each instance, the Panel agreed that data collected under GLP conditions would be greatly preferred, particularly for reasons of data quality and the associated reliability of any interpretations. However, they noted that GLP compliance would not be considered a requirement that would automatically exclude data from consideration. The Panel concluded that other factors could be used to identify high quality data. Examples would include published in a peer-reviewed journal or obtained from a study conducted in a laboratory that routinely conducts GLP studies. Data generated under non-GLP conditions would be subject to a critical quality review, and as such the Panel considered it important to obtain the original records in order to confirm the reported data.

The Panel was asked if they considered the types of substances included in the reference substance list, with regard to relative sensitization potency, physicochemical characteristics, and vehicles, to be representative of the overall diversity of substances that are likely to be tested for skin sensitization. The Panel concluded that although the list should not be considered all-inclusive, it was sufficiently representative.

The Panel was also asked if there were other types of information relevant to skin sensitization that should be considered in order to demonstrate an adequately diverse reference list. The Panel commented that identifying concentrations of each of the substances that are known to cause excessive local irritation or overt systemic toxicity would be useful.

The Panel was asked if there were other substances that they considered to be more appropriate for assessing the sensitivity (i.e., ability of the test method to correctly identify sensitizing substances) and specificity (i.e., ability of the test method to correctly identify non-sensitizing substances) of non-radioactive LLNA test methods, and for which there is available LLNA, guinea pig, and human data. The Panel could not identify such substances given the time frame for consideration but reiterated that substances in the reference list should have robust data.

A subset of "discordant chemicals" (i.e., false negative or two false positive compared to guinea pig tests or human data) were included as "optional" substances that could be studied to evaluate if the proposed modifications might provide improved performance relative to the traditional LLNA. The Panel was asked to comment on the appropriateness of including these specific substances in the reference list, whether they should be required, whether different substances should be included, and if more false negative/positive substances should be tested. As mentioned previously, the Panel commented that it was appropriate to include such substances in the reference list and that they should be required and evaluated during the validation of alternative LLNA assays that are functionally and mechanistically similar to the traditional LLNA assay. The Panel noted that the substances that were considered discordant depended on the species to which comparisons are made (i.e., LLNA vs. guinea pig or LLNA vs. human). Still, since the "discordant compounds" were false negatives or positives in the traditional LLNA, they would provide an opportunity to determine if modifications to the traditional LLNA may even have increased accuracy.

Finally, the Panel was asked if "correct" results with these discordant chemicals would be sufficient to consider the alternative test method to be more predictive of skin sensitization than the traditional LLNA. The Panel concluded that correct results with the "discordant chemicals" would not be sufficient to consider the alternative test method to be more predictive of skin sensitization, but it could provide supporting evidence to indicate further testing with additional compounds would be of value.

6.4 Comments on the Test Method Accuracy Standards

The draft ICCVAM LLNA Performance Standards state that the non-radioactive proposed LLNA test method should exactly match the accuracy of the traditional LLNA when evaluated with the minimum set of 18 reference substances. The Panel was asked if they agreed that test method accuracy should be based on a chemical-by-chemical match with regard to identifying the chemicals as sensitizers or non-sensitizers. The Panel commented that although an assay that is able to predict the same hazard classification for the reference substances as the traditional LLNA is desired, with the small number of reference substances available (n=18), clearly establishing equivalence will be extremely difficult. Furthermore, even with this small number, there is a statistical multiple comparisons[3] problem because more than one chemical is being tested. The likelihood that a modified LLNA will fail to demonstrate equivalence to the traditional LLNA will increase with the number of chemicals that must be identified correctly. A statistical measure of concordance should be calculated so that accuracies can be compared between methods.

The Panel reiterated their recommendation that, for use in hazard identification, a modified test method should be evaluated with all 22 substances (including false negatives and false positives) and accuracy statistics calculated. A statistical measure of concordance should be calculated so that accuracies can be compared between methods. To the extent possible,

[3] When multiple experiments are conducted and multiple observations, comparisons or hypothesis tests are conducted, the chance of observing rare events increases. Suppose, for example, that an interval is established such that 5% of observations from a particular population of data are outside that interval. Then if k independent experiments generate data from this population (e.g., a standard normal distribution), the chances that all 20 results will lie inside the interval is $(1.0 - 0.05)^k$ (N. Flournoy, personal communication).

rationale for any discordant results should be provided. However, the most potent sensitizers (e.g., DNCB) should always be identifiable. Considerable weight should be given to the balance between animal welfare and human safety when considering the adequacy of test method accuracy.

The draft ICCVAM LLNA Performance Standards recommend that, for each sensitizer, the threshold concentration that induces a positive SI response should be within 0.5x to 2.0x of the concentration obtained for the EC3 in the traditional LLNA. As described in the draft ICCVAM LLNA Performance Standards, statistical approaches have been used in an attempt to identify an appropriate range, but these calculated ranges do not appear to be the most practical. The Panel was asked to comment on the appropriateness of using this criterion to judge the equivalency of a non-radioactive version of the traditional LLNA and, if this approach was not acceptable, to suggest an alternative along with the basis for this approach. The Panel commented that the usefulness and limitations of the traditional LLNA for determining relative potency have not been definitively established, and therefore equivalence should not be based strictly on potency. Furthermore, the current database does not support the inclusion of EC3 values as a component of the accuracy evaluation. The range of 0.5x to 2x EC3 value suggested in the draft ICCVAM LLNA Performance Standards document are based on the experience with a range of known skin sensitizers tested in the standard LLNA. However, based on the available data provided, the 0.5x to 2.0x EC3 range may be too restrictive *if a strict interpretation of equivalence is applied*. The chances of a failure to achieve an EC3 within this range would vary from chemical to chemical depending upon the inherent underlying variability and robustness in the estimation of the EC3. For those chemicals for which the EC3 can be accurately estimated, the failure rate may be close to zero. For other, more variable chemicals, perhaps with fewer data points, the failure rate for a single chemical will be much higher.

The Panel reiterated their concern with regard to EC3 values (i.e., the statistical multiple comparisons problem). The likelihood that a modified LLNA will fail to demonstrate equivalence to the traditional LLNA will increase with the number of chemicals tested, the extent to which the new test must obtain the same EC3 value, and how independent the results are for different chemicals in the same lab.

For five of the 13 sensitizers on the draft ICCVAM reference substances list, the reference EC3 value was based on a single LLNA study (Table C1 of the draft ICCVAM LLNA Performance Standards). The Panel was asked to comment on the appropriateness of including such chemicals in the list of recommended reference substances and whether or not the 0.5x to 2.0x criteria should be applied to such substances. The Panel concluded that the appropriateness of the 0.5x to 2.0x ECt range had not been adequately justified. It was inappropriate to include chemicals represented by only one LLNA study on a list of recommended reference substances, as there was insufficient data by which to calculate a robust mean ECt value. Thus, those compounds should either (1) be exchanged for compounds with sufficient EC3 data (i.e., have been tested in at least three independent studies using the same solvent), or (2) retained but not considered to be part of the ECt criterion until such data has been collected.

6.5 Comments on the Test Method Reliability Standards

The draft ICCVAM LLNA Performance Standards state that acceptable intralaboratory reproducibility will be indicated by a laboratory obtaining, in each of four independent experiments conducted with at least one week between each experiment, ECt values (the estimated concentration needed to produce an SI of a defined threshold [e.g., EC3]) for HCA that are generally within 0.5x to 2.0x (i.e., 5% to 20%) of the historical mean EC3 concentration (10%) for this substance, based on existing available traditional LLNA data. The Panel was asked if they considered four repeat experiments to be adequate. The Panel concluded that four experiments would be adequate, as requiring four independent experiments is similar to the original LLNA submission, as is a one-week interval between experiments. Therefore, these requirements were appropriate for a comparison of modified methods to the traditional LLNA. However, it would be useful to have this number evaluated statistically (see **Section 6.7**).

The Panel was asked if they considered testing HCA adequate for demonstrating intralaboratory reproducibility and if not, which substance(s) should be tested. The Panel concluded that HCA testing would be adequate for demonstrating intralaboratory reproducibility and would allow an effective comparison to the traditional LLNA.

The Panel was asked to comment on whether the required one-week interval between independent tests was adequate and/or appropriate. The Panel concluded that the minimum one-week interval seemed logical and that the more important clarifying information might be the elements that define independent tests (e.g., different animal shipment, different reagents, different operator, blind testing).

The Panel was asked to comment on the appropriateness of the criteria for acceptability (generally within 0.5x to 2.0x EC3 for HCA), or to describe another criteria and explain the basis for their recommendation. The Panel concluded that the criteria for acceptability appeared to be appropriate because the statistical multiple comparisons issue does not exist. However, given that there is so much data and experience with HCA and the fact that only one compound is being tested (not 18), it is reasonable to evaluate reproducibility using the mean ± 3 standard deviations rather than the 0.5x to 2.0x EC3 range to account for a single comparison (see **Section 6.7** regarding data transformation recommendations). The Panel noted that historical control data using HCA in the same vehicle could be used to demonstrate adequate intralaboratory reproducibility.

The draft ICCVAM LLNA Performance Standards state that acceptable interlaboratory reproducibility will be indicated by each of three laboratories obtaining ECt values for HCA and DNCB from a single experiment that are generally within 0.5x to 2.0x (5% to 20% and 0.025 to 0.1%, respectively) of the mean historical EC3 concentration (10% and 0.05%, respectively) obtained for these two substances in the traditional LLNA. The Panel was asked if they considered the single experiment per substance in each laboratory to be adequate. The Panel concluded that, considering the overall validation plan for a given laboratory, multiple experiments (n=3) within each laboratory should be conducted.

The Panel was asked if they considered testing HCA and DNCB to be adequate for demonstrating interlaboratory reproducibility and if not, which substance(s) should be tested. The Panel concluded that, since there is a great deal of data and experience with HCA and

DNCB, and many laboratories have successfully worked with them in the traditional LLNA, they should be considered adequate for this purpose.

The Panel was asked if they considered the criteria for acceptability to be appropriate. The Panel concluded that the criteria for acceptability (i.e., generally within 0.5x to 2.0x ECt for HCA and DNCB) appeared to be appropriate because the statistical multiple comparisons problem was relatively minor given that only two substances are being tested. However, given that there is so much data and experience with HCA and DNCB and the fact that two compounds are being tested (not 18), it is reasonable to evaluate reproducibility using the mean ± 4.5 standard deviations to account for statistical multiple comparisons (see **Section 6.7** regarding data transformation recommendations). The Panel also noted that historical control data using HCA and DNCB in the same respective vehicle could be used to demonstrate adequate interlaboratory reproducibility.

6.6 Summary

The Panel was asked what criteria should be used to evaluate the equivalence of a radioactive or non-radioactive LLNA method to the traditional LLNA, if one were proposed with a "major" change, as defined in the draft ICCVAM LLNA Performance Standards (e.g., different mouse strain or use of male mice, change in the schedule for test article administration, change in schedule for lymph node excision, etc.). The Panel commented that the idea of what is a "major" and a "minor" change should be re-considered (refer to Question 2 regarding essential test components). The final version of the performance standards should be adequate to evaluate any protocol modifications.

The Panel was asked if a new set of performance standards would be required for a modified version of the LLNA that incorporated one or more "major" protocol changes. Based on the above response, the Panel concluded that a new set would not be required.

The Panel was asked to comment on how many reference substances might be considered adequate for evaluating the validity of a modified version of the LLNA with a "major" protocol change; specifically, if the 18 minimum reference substances in the draft ICCVAM LLNA Performance Standards would be sufficient. The Panel concluded that additional substances should not be considered necessary. However, since eight of the proposed sensitizers had limited data (i.e., EC3 values based on ≤2 LLNA studies), other substances with more robust data should be considered as replacements. Furthermore, if the goal is to evaluate a specific applicability domain, additional test substances might be needed.

The Panel was asked to comment, regardless of the number of reference substances, whether the alternative LLNA with a "major" change should be required to obtain the same "call" (and potency for sensitizers) as the traditional LLNA for the 18 minimum reference substances in the draft ICCVAM LLNA Performance Standards. The Panel reiterated that an assay that is equivalent to the traditional LLNA is desired, but with the small number of reference substances available, clearly establishing equivalence will be extremely difficult. They also reiterated their concern regarding the statistical multiple comparisons problem.

For use in hazard identification, a proposed modified LLNA should be evaluated with all 22 substances (including false negatives and false positives) and accuracy statistics calculated so that accuracies can be compared between the modified test method and the traditional LLNA. To the extent possible, rationale for any discordant results should be provided. However, the

most potent sensitizers (e.g., DNCB) should always be identifiable. Considerable weight should be given to the balance between animal welfare and human safety when considering the adequacy of test method accuracy.

The Panel was asked to identify any additional specific substances that should be used. The Panel concluded that while additional substances should not be needed, it would be useful to identify replacements for the eight proposed sensitizers with limited test data. If the goal is to evaluate a specific applicability domain, additional test substances might be needed.

6.7 Additional Statistical Comments

During the evaluation of the draft ICCVAM LLNA Performance Standards, the Panel noted a number of statistical issues that should be addressed. They suggested that in order to achieve a normal distribution of the data and to reduce differences between groups, a suitable variance stabilizing transformation (e.g., log transformation, square root transformation) should be applied in all statistical analyses and in reporting summary standard deviations. The Panel also suggested that there should be a more rigorous evaluation of what would be considered an appropriate range of ECt values to include as a requirement. This would be a statistical evaluation that takes into consideration the variability of ECt values generated among the sensitizers included on the performance standards reference substances list and the statistical multiple comparisons problem and the fact that sample sizes that are less than 30 invalidate statistics based on the normal distribution (Young 2007).

Furthermore, bioequivalence models have been developed (Berger and Hsu 1996) and should be applied to the LLNA. Probability values can be used as descriptive statistics and as such provide a summary measure of weight-of-evidence that would be useful for comparison of performance standards across test methods. In this context, it would be informative to have statistical tests of data generated for these purposes. A test of concordance for measuring the accuracy of classification should be done.

Intralaboratory tests should include analysis of variance (ANOVA)-like tests with a test for no trend, with the null hypothesis being that there is a difference and the alternative being that the difference is bioequivalent. Interlaboratory tests should include ANOVA-like tests with the null hypothesis being that there is a difference and the alternative hypothesis being that the difference is bioequivalent. The reliability tests require “bioequivalence” to be defined (i.e., what is acceptable to be considered equivalent).

It is not known whether these specific statistical tests can be identified in the literature or if they need to be developed. If they do need to be developed, this should be given a priority. Prior to running reliability studies, these statistical methods should be used to determine the appropriate number of substances and the number of times each substance needs to be tested within and among laboratories in the study design (see also ISO 5725 [ISO 1994] and ASTM Standard E691 [ASTM 2005]). The power for the traditional LLNA should be established for comparison purposes.

[This Page Intentionally Left Blank]

7.0 Use of the LLNA for Potency Determinations

7.1 Comments on the Draft BRD for Completeness, Errors, and Omissions

The Panel was asked if there were any errors in the draft BRD on the use of the LLNA for potency determinations that should be corrected, if omissions of existing relevant data had been identified, or if there was additional information that should be included. The Panel noted alternative analyses that would better help evaluate the use of the traditional LLNA for skin sensitization potency (see the discussion of the categorization scheme in **Section 7.2** and the discussion of future studies in **Section 7.3**).

7.2 Comments on the Validation Status of the Traditional LLNA to Determine Skin Sensitization Potency

7.2.1 Substances Used for the Validation Studies

The Panel was asked to consider whether the validation status of the traditional LLNA for potency categorization (i.e., "strong" vs. "weak" sensitizers) has been adequately characterized, and if the traditional LLNA is sufficiently accurate and reliable to be used as a stand-alone assay for characterizing the potency of sensitizing substances, based on the comparison to human and guinea pig responses. The Panel agreed that the LLNA database of 170 substances with comparative guinea pig (i.e., Guinea Pig Maximization Test or Buehler Test) and/or human data (i.e., Human Maximization Test [HMT] and/or Human Repeat Insult Patch Test [HRIPT], but not human clinical observations) is sufficient in number and well balanced for this evaluation. The database included 112 substances (97 sensitizers, 15 non-sensitizers) with comparative human data and 105 substances (52 sensitizers, 53 non-sensitizers) with comparative guinea pig data. Known contact sensitizers of public health concern from various chemical groups are included. The Panel further agreed that these substances were representative of a sufficient range of chemical classes and physical chemical properties so that it would be applicable to the types of chemicals and products typically tested for skin sensitization potential.

While coding of chemicals to reduce bias is recommended for validation studies, this evaluation was based on a retrospective evaluation of existing data, most of which were generated using chemicals that were not coded. The Panel was asked whether the lack of coding of test substances adversely impacted or biased the current evaluation. Given the nature of the studies (i.e., the testing was not conducted to demonstrate the ability of the LLNA to be used for potency characterization), the Panel stated that the lack of coding likely had no impact on the current evaluation.

For some substances tested for sensitization using the traditional LLNA, it was not possible to determine whether the data were generated using pooled or individual animal lymph node samples within a dose group (the former allowed in OECD TG 429 [OECD 2002]; the latter as recommended in the ICCVAM 2001 protocol and required in the EPA 2003 skin sensitization test guideline). Cockshott et al. (2006) reported that using individual animal data allowed for technical problems during an experiment and outlier animals within a dose group to be identified. Considering this, the Panel was asked whether the analysis of the performance of the traditional LLNA for potency determinations should be limited to data from studies that can be confirmed as using individual animal data collection procedures.

A majority of the Panel agreed that, ideally, future traditional LLNA potency determinations should be based on data from studies that use individual data collection procedures, as this would allow for the identification of outliers that might skew the average group stimulation index. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable.

7.2.2 Test Method Accuracy

The Panel was further asked what impact the inclusion of pooled animal data might have on the accuracy analysis included in Section 6.0 of the draft ICCVAM LLNA potency BRD. With regard to this retrospective dataset, the Panel agreed that pooled data should not be excluded from the current analysis to assess potency determinations for the traditional LLNA. The Panel stated that it is impossible to assess the impact of using pooled data without a separate analysis of the ability of the traditional LLNA to be used for characterizing skin sensitization potency using pooled vs. individual data, which the Panel recommended be done (see the discussion of future studies in **Section 7.3**).

A minority opinion from Dr. Dagmar Jírová stated that, since OECD TG 429 (OECD 2002) allows the use of both pooled and individual animal data, the analysis that includes both types of data is appropriate. Even with the diversity of data sources (the vehicle is not known for 43% of substances tested in the traditional LLNA; human data were obtained by different, even undefined methods, etc.), the outcome of the evaluation was good, which documents the strength and robustness of the traditional LLNA.

The Panel was asked whether the correct classification, as well as the over- and under-classification, rates of the traditional LLNA for sensitization potency determinations had been adequately compared and appropriately evaluated based on the corresponding human and guinea pig data (refer also to Section 6.0 of the draft ICCVAM LLNA potency BRD). The Panel agreed that the two approaches used in the draft BRD for analyzing the ability of the traditional LLNA to discriminate between strong and weak skin sensitizers were appropriate and correct. In these two approaches, the traditional LLNA was evaluated, after identifying the optimal EC3 value, for its ability to correctly classify strong and weak sensitizers as defined by human or guinea pig threshold values based on: (1) sensitizers only, and (2) sensitizers combined with false positives, false negatives, and non-sensitizers.

A minority opinion from Dr. Howard Maibach stated that the relevance of the traditional LLNA to human clinical observations has not been sufficiently determined and should be.

The accuracy analysis (see Section 6.0 of draft ICCVAM BRD) focuses on a proposed two-level categorization scheme (weak sensitizers vs. strong sensitizers) for both human and guinea pig data. The Panel was asked whether this was an appropriate categorization scheme, or if other categorization schemes should be considered. The Panel agreed that the two-level categorization scheme was appropriate, especially considering the fact that, for human situations, risk assessment should be performed, and therefore more categories are not needed. Even a weak sensitizer under heavy exposure and individual circumstances may reach a comparable risk level as a strong sensitizer under conditions of low exposure.

A minority opinion from Drs. Raymond Pieters and Michael Woolhiser recommended the addition of at least a moderate category since certain compounds will always be on the border between weak and strong. Dr. Pieters specifically recommended the categorization scheme of Kimber et al. (2003), which is based on five categories if non-sensitizers are included.

Of the two human threshold concentrations that are proposed in this two-category categorization scheme (i.e., <250 μg/cm^2 or <500 μg/cm^2), the Panel was asked which threshold was the most appropriate for categorizing sensitizing substances as strong vs. weak for humans, or if another threshold was more appropriate for this purpose. The Panel noted that this validation was based on comparison to guinea pig and HMT/HRIPT information. These data relate only to induction and do not permit an assessment of risk in humans for elicitation.

For the data provided, the Panel concluded that the best results were obtained using the decision criterion of 250 μg/cm^2 and the corresponding optimal traditional LLNA EC3 value of 9.4%. Using this cut-off when traditional LLNA false negative and false positive substances are included in the analysis, in addition to sensitizers in both the traditional LLNA and in humans using the HRIPT and/or HMT, correct classification of strong sensitizers was 79% and underclassification was 21%. Underclassification of substances in this context means classification as weak instead of strong sensitizers (i.e., they are not missed as sensitizers regarding the labeling and safety of consumers). The Panel stated that more data are needed to determine if another threshold is more appropriate.

When the potency categorization analysis was based on sensitizers only, the guinea pig tests predicted weak sensitizers with higher accuracy than did the LLNA (89% vs. 75% for the 250 μg/cm^2 cutoff and 83% vs. 60% for the 500 μg/cm^2 cutoff), which is logical because the guinea pig test methodology involves all phases of the sensitization process and usually involves adjuvants. However, the guinea pig tests were less accurate for the prediction of strong sensitizers compared to LLNA (48% vs. 71% for the 250 μg/cm^2 cutoff and 42% vs. 63% for the 500 μg/cm^2 cutoff), which represents a higher risk for consumers. For the protection of public health, it is more important to correctly identify strong sensitizers than weak sensitizers.

The Panel was asked whether the draft BRD adequately characterized the usefulness and limitations of the LLNA for potency categorizations. If not, the Panel was asked what additions or changes should be made to the description of usefulness and limitations in the draft BRD. The Panel stated that additional evaluations should be conducted to determine the impact on potency categorization if the human threshold data are evaluated differently (e.g., alternative lowest observed effect level [LOEL] safety factors other than 10, using LOEL data only, using no observed effect level [NOEL] data only), and if this might improve the correlation between the LLNA and the human results. According to the Panel, the approach of directly comparing the LOEL values without using a safety factor compares values of similar significance in humans and in the LLNA. In other words, the LOEL in humans describes the threshold induction area dose in humans and the EC3 value in the traditional LLNA is the threshold induction area dose and thus could be the analogous value to the human LOEL. The Panel further stated that traditional LLNA tests based on pooled or individual lymph nodes for a dose group should be evaluated independently to assess the

impact of using pooled data on the accuracy analysis for skin sensitization potency. Finally, the Panel stated that the effect of different vehicles should be recognized as a limitation in the data analysis given the demonstrated variability of results.

7.2.3 Test Method Reliability

The Panel was asked whether the reliability (e.g., intralaboratory repeatability, intra- and inter-laboratory reproducibility) of the traditional LLNA for potency determinations had been adequately evaluated. If not, the Panel was asked what other analyses should be performed. Similar to their recommendations for test method accuracy, the Panel stated that additional evaluations of reliability should be conducted based on using different approaches for human threshold data (e.g., using alternative LOEL safety factors other than 10, using LOEL data only, using NOEL data only). The Panel further stated that the reliability of LLNA based on using pooled or individual animal data should be evaluated independently. Finally, the Panel stated that the effect of different vehicles should be recognized as a limitation in the data analysis, as a source of increased variability.

7.2.4 Data Quality

It was not possible to determine whether or not all studies included in the draft LLNA potency BRD had been conducted in accordance with GLP guidelines, nor was it possible to obtain the results of GLP audits for all studies determined to be GLP-compliant. The Panel was asked to discuss what impact this might have on the evaluation of the LLNA for potency determinations and whether any of the non-GLP studies should be excluded from the analyses. The Panel concluded that it was important to note if the data were obtained from studies conducted according to GLP guidelines, as ideally this should be the case. However, the Panel concluded that data from studies that could not be confirmed as being GLP-compliant, but that were from peer-reviewed literature or other sources with high-quality laboratory management practices were still appropriate to include in this retrospective analysis.

As described in the draft BRD, original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Considering this, the Panel was asked whether the results of these studies (all of which are currently included) be excluded from any of the performance analyses. The Panel considered the data to have been generated by repeatedly published and reliable laboratories and therefore did not question the adequacy/quality of the retrospective data analysis. Thus, although data should be checked when available, exclusion of data was not deemed necessary, in this case.

7.2.5 Consideration of All Available Data and Relevant Information

Based on the draft BRD, the Panel was asked whether all the relevant data identified in published or unpublished studies conducted using the traditional LLNA had been adequately considered. If not, the Panel was asked what other studies should be considered. The Panel recommended that the LOELs from Akkan et al. (2003) be used instead of the DSA_{05} values from Schneider and Akkan (2004) in all of the potency analyses. A minority opinion by Dr. Thomas Gebel stated that it was acceptable to use the DSA_{05} values from Akkan et al. (2003) as LOEL values in the evaluation. Dr. Gebel mentioned that the DSA_{05} value is a LOEL

value adjusted to 5% incidence of induction. Akkan et al. (2003) used the DSA_{05} value to correct for different human studies leading to different inductions. Dr. Gebel further stated that as the DSA_{05} is corrected for an induction rate of 5%, it would be better to compare with the traditional LLNA EC3 than to use the default uncorrected LOEL.

7.3 Comments on the Draft ICCVAM Test Method Recommendations for the Use of the LLNA for Potency Determination

7.3.1 Test Method Usefulness and Limitations

With regard to the use of the LLNA for potency categorization (i.e., strong vs. weak sensitizers), the ICCVAM draft recommendation is that the traditional LLNA should not be considered as a stand-alone test method for predicting sensitization potency, but must instead be used as part of a weight-of-evidence evaluation to discriminate between strong and weak sensitizers. This is based on the fact that, although there is a significant positive correlation between traditional LLNA EC3 values and human sensitization threshold doses, this correlation is not strong [see detailed discussion in the draft ICCVAM recommendations]. The Panel agreed that the traditional LLNA should not be considered a stand-alone assay for categorization of skin sensitization potency, but it could be used in a weight-of-evidence evaluation (e.g., along with quantitative structure-activity relationship [QSAR], peptide reactivity, human evidence) to discriminate between strong and weak sensitizers. The Panel further stated that there are additional studies proposed that may provide a better correlation and improve prediction of potency categorization (see the discussion of future studies below).

A minority opinion from Drs. Thomas Gebel and Dagmar Jírová stated that there is a significant positive correlation between EC3 values and human threshold values. It is likely that limitations in estimating human threshold values and the inclusion of human NOEL values in the current evaluation contributed negatively to the resulting R^2 value of 0.405 (when LLNA EC3 data vs. human threshold data were compared, see Table 6-2 of the draft ICCVAM BRD). Thus, the R^2 value may improve when the additional analyses that have been suggested by the Panel are conducted.

The Panel stated that the effect of different vehicles should be recognized as a limitation in the data analysis and a likely source of within and between laboratory variability.

7.3.2 Test Method Protocol

The Panel was asked whether the ICCVAM-recommended LLNA protocol (ICCVAM 1999; EPA 2003) should be used when generating data that will or might be considered for sensitization potency categorization decisions. The Panel agreed that this protocol should be used. A minority opinion by Drs. Nathalie Alépée, Thomas Gebel, Dagmar Jírová, Raymond Pieters, and Michael Woolhiser stated that if laboratories were operating under OECD guidance (OECD 2002) and a reliable validation dataset had been generated, then pooled data from at least four animals could be considered acceptable.

The Panel was asked whether the relevant testing guidelines for the traditional LLNA should be updated to include the calculation of an EC3 value. The Panel agreed with this recommendation. The calculation of an EC3 value is briefly described in the draft ICCVAM

LLNA Performance Standards for specific situations with references to Basketter et al. (2000) and Ryan et al. (2007).

7.3.3 Future Studies

The Panel was asked whether the available data support the ICCVAM draft recommendations for the traditional LLNA in terms of the proposed future studies. The Panel agreed and concluded that more data are needed to determine the optimal threshold in humans for distinguishing between strong and weak sensitizers. However, the Panel discouraged conducting new animal studies unless it was likely that results from such studies would lead to an overall reduction in animal use. The Panel stated further that the traditional LLNA appears to be a robust rodent assay for the quantification of the induction of cell-mediated immunity. Thus, use of the traditional LLNA for potency determination can be used in conjunction with QSAR information, guinea pig assays, HRIPT/HMT, and the quantitative data of elicitation and frequency of positive response in humans in a weight-of-evidence approach. The Panel further stated that additional evaluations should be conducted to determine the impact on potency categorization if the human threshold data are evaluated differently (e.g., alternative LOEL safety factors other than 10, using LOEL data only, using NOEL data only). This might improve the correlation between LLNA and human data. The Panel further stated that LLNA tests based on pooled or individual animal data should be evaluated independently to assess the impact of using pooled data on the accuracy for determining skin sensitization potency.

The Panel recommended a statistical analysis to determine where an appropriate cutoff value between weak or strong sensitizers might be best defined for traditional LLNA data. For example, receiver operating characteristic curves could be used to identify the optimum cut-off for determining the difference between weak and strong sensitizers.

Finally, the Panel stated that the effect of different vehicles should be recognized as a limitation in the current data analysis, that this was a source of variability within and between laboratories, and that its impact should be considered in future analyses.

8.0 References

Akkan Z, Kalberlah F, Oltmanns J, Schneider K. 2004. Beurteilung der Wirkst€arke hautsensibilisierender Chemikalien anhand des Local Lymph Node Assay (LLNA) [in German]. Bremerhaven, Germany: Federal Institute for Occupational Safety and Health.

ASTM. 2005. ASTM Standard E691-05 Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method. In: Annual Book of ASTM Standards. West Conshohocken: ASTM International. www.astm.org.

Basketter DA, Blaikie L, Dearman RJ, Kimber I, Ryan CA, Gerberick GF, et al. 2000. Use of the local lymph node assay for the estimation of relative contact allergenic potency. Contact Dermatitis 42:344-348.

Berger RL, Hsu JC. 1996. Bioequivalence trials, intersection-union tests and equivalence confidence sets. Stat Sci 11:283-319.

Cockshott A, Evans P, Ryan CA, Gerberick GF, Betts CJ, Dearman RJ, et al. 2006. The local lymph node assay in practice: a current regulatory perspective. Hum Exp Toxicol 25:387-394.

Dean JH, Twerdok LE, Tice RR, Sailstad DM, Hattan DG, Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay: II. Conclusions and recommendations of an independent scientific peer review panel. Regul Toxicol Pharmacol 34:258-273.

De Jong WH, van Och FM, Den Hartog Jager CF, Spiekstra SW, Slob W, Vandebriel RJ, et al. 2002. Ranking of allergenic potency of rubber chemicals in a modified local lymph node assay. Toxicol Sci 66:226–232.

EPA. 2003. Health Effects Test Guideline OPPTS 870.2600. Skin Sensitization. Research Triangle Park, NC: Environmental Protection Agency.

Gaspari AA. 2007. Mechanism of action and other potential roles of an immune response modifier. Cutis 79(4 SUPPL.):36-45.

Gerberick GF, Ryan CA, Kern PS, Dearman RJ, Kimber I, Patlewicz GY, et al. 2004. A chemical dataset for evaluation of alternative approaches to skin-sensitization testing. Contact Dermatitis 50:274-288.

Gerberick GF, Vassallo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin JP. 2007. Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. Toxicol Sci 97:417-27.

ICCVAM. 1999. The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals/Compounds. The Results of an Independent Peer Review Evaluation Coordinated by the Interagency Coordinating Committee on the Validation of Alternative Toxicological Methods (ICCVAM) and the National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication No. 99-4494. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No: 03-4508. Research Triangle Park, NC: National Toxicology Program.

ISO. 1994. ISO 5725. Accuracy (trueness and precision) of measurement methods and results. Geneva: International Organization for Standardization.

ISO. 2002. ISO 10993-10: Tests for Irritation and Sensitization. Geneva: International Organization for Standardization.

Jowsey IR, Clapp CJ, Safford R, Gibbons BT, Basketter D. 2008. The impact of vehicle on the relative potency of skin sensitising chemicals in the local lymph node assay. In Press.

Kimber I, Basketter DA, Butler M, Gamer A, Garrigue JL, Newsome C, et al. 2003. Classification of allergens according to potency: proposals. Food Chem Toxicol 41:1799-1809.

Kimber I, Dearman RJ, Betts CJ, Gerberick GF, Ryan CA, Kern PS, et al. 2006. The local lymph node assay and skin sensitization: a cut-down screen to reduce animal requirements? Contact Dermatitis 54:181-185.

Manetz TS, Meade BJ. 1999. Development of a flow cytometry assay for the identification and differentiation of chemicals with the potential to elicit irritation, IgE-mediated, or T cell-mediated hypersensitivity responses. Toxicol Sci. 48:206-17.

OECD. 2002. Guideline for the Testing of Chemicals 429 - Skin Sensitisation: Local Lymph Node Assay. Paris: OECD.

Ryan CA, Chaney JG, Gerberick GF, Kern PS, Dearman RJ, Kimber I, Basketter DA. 2007. Extrapolating local lymph node assay EC3 values to estimate relative sensitizing potency. Cutan Ocul Toxicol 26:135-145.

Schneider K, Akkan Z. 2004. Quantitative relationship between the local lymph node assay and human skin sensitization assays. Regul Toxicol Pharmacol 39:245-55.

Sailstad DM, Hattan D, Hill RN, Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay: I. The ICCVAM review process. Regul Toxicol Pharmacol 34:249-257.

Takeyoshi M. 2007. Promising endpoint and assay performance of non-radioisotopic local lymph node assay based on BrdU incorporation. Sixth World Congress on Alternatives and Animal Use in the Life Sciences. August 21-25, 2007. Tokyo, Japan.

Van Hoogstraten IM, Boos C, Boden D, Von Blomberg ME, Scheper RJ, Kraal G. 1993. Oral induction of tolerance to nickel sensitization in mice. J Invest Dermatol 101:26-31.

Van Och FMM, Slob W, De Jong WH, Vandebriel RJ, Van Loveren H. 2000. A quantitative method for assessing the sensitizing potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of the uncertainty margins. Toxicology 146:49-59.

Varani J, Perone P, Spahlinger DM, Singer LM, Diegel KL, Bobrowski WF, et al. 2008. Human skin in organ culture and human skin cells (keratinocytes and fibroblasts) in monolayer culture for assessment of chemically induced skin damage. Toxicol Pathol 35:693-701.

Young SS. 2007. Gaming the system: chaos from multiple testing. Institute of Mathematical Statistics Bulletin 36:13

[This Page Intentionally Left Blank]

.

Appendix A

Peer Review Panel Member Biosketches

[This Page Intentionally Left Blank]

Panel Member Biosketches

Nathalie Alépée, Ph.D.

Dr. Alépée performed research leading to a Ph.D. in Medical Virology and Microbiology at the Centre National de la Recherché Scientifique research institute, Gif sur Yvette, France. She is currently the Global Pfizer Leader for photosafety, including the global portfolio support and Associate Research Fellow in Investigative Toxicology, at Pfizer Global Research and Development, Amboise, France. As a laboratory manager in the Molecular and Cellular Toxicology Group with Pfizer, she implemented the Local Lymph Node Assay (LLNA) in the laboratory. She serves on the European Centre for the Validation of Alternative Methods (ECVAM) Scientific Advisory Committee (ESAC), representing the European Federation of Pharmaceutical Industries Associations (EFPIA). She is also the Pfizer representative to the European Partnership on Alternative to Animal Testing (EPAA), in two working groups; Identification of Opportunities, Including R&D (working group 2), and Validation and Acceptance (working group 5). She served as a peer reviewer of the reduced LLNA test protocol and prediction model for ESAC in 2007 and has been designated as an ESAC peer reviewer for ECVAM's performance standards for the standard LLNA.

Anne Marie Api, Ph.D.

Dr. Api received a Ph.D. from Aston University in Birmingham, England and is currently Vice President of Human Health Sciences at the Research Institute for Fragrance Materials (RIFM), as well as the Scientific Director. She is responsible for the human health scientific program, and the investigation and initiation of new research and testing projects for RIFM. She is also Adjunct Assistant Professor at the University of Medicine and Dentistry of New Jersey. She is a member of 10 professional organizations, including the American Contact Dermatitis Society, the European Society of Contact Dermatitis, and the Society of Investigative Dermatology. She participated in the World Health Organization (WHO) International Workshop in Skin Sensitization in Chemical Risk Assessment held in Berlin, Germany in 2006. She is author of over 100 publications and presentations relevant to dermatology and dermatotoxicology.

Nancy Flournoy, M.S., Ph.D.

Dr. Flournoy received a M.S. degree in Biostatistics from the University of California at Los Angeles, and a Ph.D. in Biomathematics from the University of Washington. She is Professor and Chair of the Department of Statistics at the University of Missouri-Columbia. Her research interests include adaptive designs, bioinformatics, chemometrics, clinical trials, and environmetrics. She has an extensive list of edited volumes and papers on statistical theory, statistical genetics and immunology, epidemiology in immune suppressed subjects, clinical trials for prevention and treatment of viral infection, transplantation biology and its effects on digestion, lungs, eyes, mouth, and central nervous system, optimization of statistical processing, and additional papers, interviews, and technical reports. She has editorial responsibilities for numerous statistical journals, serves on numerous advisory boards, and nominating committees. She is a member and past Chair of the Council of Sections of the American Statistical Association, and served in various other statistical, medical and toxicological societies or programs as Chair or as a member of the Board of Directors. She is a former member of the Scientific Advisory Committee on Alternative Toxicological

Methods. She also served on the Expert Panels for the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) that evaluated the Revised Up-and-Down Procedure; the Current Validation Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants; and Five *In Vitro* Pyrogen Test Methods.

Thomas Gebel, Ph.D.

Dr. Gebel received a Ph.D. in Toxicology from the University of Mainz and is certified as a toxicologist by the German Society of Toxicology. His scientific interests are in biomonitoring, genetic toxicology, environmental hygiene, and occupational toxicology. He has published over 40 papers in peer-reviewed scientific journals. He is employed by the German Federal Institute for Occupational Safety and Health, and is an Associate Professor at the University of Goettingen. Dr. Gebel is currently a member of the Organisation for Economic Co-operation and Development (OECD) Globally Harmonized System of Classification and Labeling of Chemicals (GHS) expert group on sensitization and head of the German advisory committee on classification and labeling of existing substances and biocides. Dr. Gebel also is head of the German Delegations to the United Nations Economic and Social Council Sub-Committee of Experts on the GHS, and to the OECD Task Force on Harmonisation of Classification and Labeling. He participated in the WHO International Workshop in Skin Sensitization in Chemical Risk Assessment held in Berlin, Germany in 2006.

Sidney Green Ph.D., F.A.T.S.

Dr. Green received a Ph.D. in Biochemical Pharmacology from Howard University. His research interests include toxicology, mutagenic assay systems, and alternatives to animals in toxicology. He is currently Graduate Professor of Pharmacology at Howard University and a faculty member at the Centers for Alternatives to Animal Testing at the Johns Hopkins University School of Public Health. Previously, he has been Director of the Department of Toxicology at Covance Laboratories Inc. and the Director of the Division of Toxicological Research at the U.S. Food and Drug Administration (FDA). Dr. Green is a Fellow of the Academy of Toxicological Sciences (F.A.T.S.). He has served on numerous expert panels and committees. He was a participant in an International Workshop organized by ICCVAM and NICEATM on *In Vitro* Methods for Assessing Acute Systemic Toxicity in 2000. He served on the ICCVAM/NICEATM Expert Panels that evaluated the Corrositex® Test Method for Assessing Dermal Corrosivity Potential of Chemicals, and *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants. He is a former member of the ICCVAM Advisory Committee on Alternative Toxicological Methods (ACATM) and of SACATM. He has authored over 60 publications for peer-reviewed journals.

Kim Headrick, B.Admin., B.Sc.

Kim Headrick received Bachelor of Administration and B.Sc. degrees from the University of Ottawa, Canada. She is currently International Harmonization and Senior Policy Advisor for Health Canada, and Chair of the UN Sub-Committee of Experts on GHS. She manages the overall strategy for the implementation of the GHS in Canada. She was awarded the Queen Elizabeth Commemorative Golden Jubilee Medal in 2002, which focuses on the achievements of people who, over the past 50 years, have created the Canada of today. She is

a member of the OECD Task Force on Harmonization of Classification and Labelling and the OECD Expert Group Meeting on Sensitization Hazards.

Dagmar Jírová, M.D., Ph.D.

Dr. Jírová received a Ph.D. from the Medical Faculty of Hygiene at Charles University in Prague. She is currently the Head of the Reference Center for Cosmetics, and Head of National Reference Laboratory for Experimental Immunotoxicology at the National Institute of Public Health in the Czech Republic. Her main responsibilities include safety assessment of consumer products, particularly cosmetics and their ingredients, performance of toxicological methods *in vivo* in animals, human patch testing for local toxicity assessment, and introduction of *in vitro* techniques for screening of toxicological endpoints using cell and tissue cultures. She represents the Czech Republic in the Standing Committee on Cosmetics of the European Commission. She is an ESAC-ECVAM member and was involved in Peer Review Panel for Skin Irritation Validation Study and LLNA test protocol and prediction model. She is author of more than 100 publications and presentations relevant to dermatotoxicology including a recent presentation at the 6th World Congress on Alternatives & Animal Use in the Life Sciences, held in Tokyo, 2007, titled "Comparison of Human Skin Irritation and Photoirritation Patch Test Data with Cellular *in vitro* Assays and Animal *in vivo* data".

David Lovell, Ph.D., B.Sc. (Hons), F.S.S., FIBiol, CStat, CBiol

Dr. Lovell received a Ph.D. from the Department of Human Genetics and Biometry, University College, London. He is currently Reader in Medical Statistics at the Postgraduate Medical School at the University of Surrey. Previously, he was Associate Director and Head of Biostatistics support to Clinical Pharmacogenomics at Pfizer Global Research and Development in Sandwich, Kent providing data management and statistical support to pharmacogenetics and genomics. He joined Pfizer in 1999 as the Biometrics Head of Clinical Pharmacogenetics. Before joining Pfizer, Dr. Lovell was the Head of the Science Division at BIBRA International, Carshalton, which included Molecular Biology, Genetic Toxicology, Biostatistics and Computer Services. At BIBRA, Dr. Lovell managed the statistical and computing group providing specialized statistical support to BIBRA's Clinical Unit and contract research work. He conducted and managed research programs on genetics, statistics and quantitative risk assessment for the European Union (EU) and U.K. Government Departments. His research interests at BIBRA were in the use of mathematical and statistical methods together with genetic models in the understanding of toxicological mechanisms and risk assessment problems. Dr. Lovell had previously been a Senior Research Officer with the U.K. Medical Research Council (MRC) Experimental Embryology and Teratology Unit, a visiting Postdoctoral Fellow at the National Institute of Environmental Health Sciences (NIEHS) in North Carolina, U.S., a Geneticist at the MRC Laboratories, Carshalton, and a Research Assistant in Cytogenetics at Birmingham University. He has acted as a consultant to a number of organizations, has considerable experience of working with Regulatory Authorities, has many publications related to his work and has wide experience of making presentations to a wide range of audiences. He is a member of the U.K. Government's advisory Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment (COM) and the Independent Scientific Advisory Committee for Medicines and Healthcare Products Regulatory Agency database research. He served on the NICEATM-ICCVAM Expert Panels

that evaluated the Frog Embryo Teratogenesis Assay - Xenopus, *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants, and Five *In Vitro* Pyrogen Test Methods.

Michael Luster, Ph.D.

Dr. Luster received a Ph.D. in Immunology from Loyola University of Chicago. He was formerly Chief, Toxicology and Molecular Biology Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health (NIOSH), and currently serves as a senior advisor to the Director of the Health Effects Laboratories and the staff of Toxicology and Molecular Biology Branch at NIOSH. Program areas include neuroscience, dermatology, molecular carcinogenesis, molecular epidemiology, molecular toxicology, molecular epidemiology, and inflammation/immunotoxicology. In addition, Dr. Luster conducts basic and applied research in immunotoxicology including its application in risk assessment. Current research activities include molecular epidemiology studies of genetic polymorphism involved in workplace-related diseases and experimental studies involving occupational allergic rhinitis. Dr. Luster is also working with various staff at the U.S. Environmental Protection Agency (EPA) through the Risk Assessment Forum to develop immunotoxicity testing guidelines. He also directed two studies for the NTP on the Toxicology and the Carcinogenesis of Promethazine and Ortho-phenylphenol, in 1990 and 1986, respectively. He is a co-author of over 300 publications in peer-reviewed journals.

Howard Maibach, M.D.

Dr. Maibach received an M.D. from Tulane University. He is currently a professor in the Department of Dermatology at the University of California, San Francisco (USCF), where he is also Chief of the Occupational Dermatology Clinic. In his 35 years at UCSF, Dr. Maibach has written and lectured extensively on dermatotoxicology and dermatopharmacology. His current research programs include defining the chemical-biologic faces of irritant dermatitis and the study of percutaneous penetration. Dr. Maibach served on the 1998 ICCVAM Peer Panel that evaluated the Murine LLNA. Dr. Maibach has been on the editorial boards of over 30 scientific journals and is a member of 19 professional societies including the American Academy of Dermatology, San Francisco Dermatological Society, and the International Commission on Occupational Health. He has co-authored over 1500 publications related to dermatology.

James McDougal, Ph.D., F.A.T.S.

Dr. McDougal earned a Ph.D. in Pharmacology/Toxicology at the University of Arizona. He is currently Professor and Director of Toxicology Research in the Department of Pharmacology and Toxicology at Wright State University's Boonshoft School of Medicine. Prior to his appointment at Wright State, he worked in the Air Force toxicology research organization for about 17 years. He has active skin research programs related to dermal pharmacokinetics, molecular biology of skin irritation, dermal risk assessment, and biologically-based mathematical modeling. He has served on many national committees, published more than 75 manuscripts, and consults for a wide variety of government and industry organizations. Dr. McDougal is a member of the National Academy of Sciences (National Research Council) Committee on Toxicology and the American Congress of Governmental Industrial Hygienists Threshold Limit Value Committee for Chemical substances. Dr. McDougal is also past president of the Dermal Toxicology Specialty Section of the Society of Toxicology.

Michael Olson, Ph.D., A.T.S.

Dr. Olson received a Ph.D. in Toxicology from the University of Arkansas for Medical Sciences, with dissertation research conducted at the FDA National Center for Toxicological Research. Following graduate training, he served as NIEHS National Research Service Award Post-doctoral Fellow in the Department of Pharmacology, School of Medicine - University of North Carolina. Currently he is Director, Occupational Toxicology, Corporate Environment Health and Safety for GlaxoSmithKline. Dr. Olson is a Fellow of the Academy of Toxicological Sciences (A.T.S.). His research interests include mechanisms of chemically-induced toxicity; genetic toxicity; xenobiotic metabolism; alternative methods in toxicology; hazard evaluation, risk assessment, and communication. Dr. Olson has authored a number of peer-reviewed manuscripts and book chapters in these areas as well as preparing many occupational health effects reviews for pharmaceutical active ingredients, isolated intermediates, and associated chemicals. He has served as an editorial board member and *ad hoc* referee for numerous toxicology and biosciences journals. In addition, he has worked as a Visiting Scientist, EPA, as well as advisor to EPA Risk Assessment Forum, U.S. National Institutes of Health (NIH) (Toxicology Study Section I), U.S. Air Force, Transportation Research Board, and the National Research Council - National Academy of Sciences. A member of several biomedical professional societies, Dr. Olson has served in elective and appointed positions in the Society of Toxicology, including Chairman of the Society of Toxicology (SOT) Occupational Health Specialty Section.

Raymond Pieters, Ph.D.

Dr. Pieters received a Ph.D. at Utrecht University and is currently an Associate Professor at the Institute for Risk Assessment Sciences, and Group Leader for Immunotoxicology at that institution. In 2007, he presented a paper on Development of Strategies to Assess Drug Hypersensitivity at the Congress of the European Societies of Toxicology. He was involved in the development of the Reporter Antigen Popliteal Lymph Node Assay, an assay to assess the immunomodulating potential of chemicals, which enables differentiation between immunosensitizing chemicals (sensitizers), immunostimulating chemicals (irritants), and chemicals that have no apparent immunological effects. He has published over 70 papers on sensitization and other subjects in immunotoxicology in peer-reviewed journals, including a review article, *Murine Models of Drug Hypersensitivity*, in 2005.

Jean Regal, Ph.D.

Dr. Regal received a Ph.D. in Pharmacology from the University of Minnesota. She is currently a Professor in the Department of Pharmacology, Department of Biochemistry & Molecular Biology and Associate Dean of Faculty Affairs, Medical School Duluth, University of Minnesota. Her current research is focused on respiratory allergy, especially asthma. She has served on multiple NIH review panels regarding asthma, as an immunotoxicologist in 2000 for an Institute of Medicine Committee on Health Effects Associated with Exposures Experienced during the Persian Gulf War, as well as on the 1998 ICCVAM Peer Panel that evaluated the Murine LLNA. In 2007 she served as an ad hoc reviewer for the NTP Board of Scientific Counselors for two nominations: Artificial Butter Flavoring Mixture & O-phthalaldehyde, at NIEHS. Also in 2007, she served on an NIEHS Center in Environmental Toxicology pilot project program for the University of Texas Medical Branch at Galveston. She is currently Vice-President-elect of the Immunotoxicology

Specialty Section of SOT and Associate Editor of the Journal of Immunotoxicology. Dr. Regal has authored over 50 research articles and reviews in peer-reviewed journals and holds two patents on pulmonary administration of sCR1 and other complement inhibitory proteins.

Jonathan Richmond, B.Sc. (Hons) Med.Sci., MB ChB, FRCSEd, FRMS

Dr. Richmond received a Bachelor of Science in Medical Science with Honors (B.Sc. [Hons] Med.Sci.) and Bachelor of Medicine and Bachelor of Surgery (MB ChB) degrees with Distinction in Medicine and Therapeutics from Edinburgh University. Presently, he is head of the Animals Scientific Procedures Division at the Home Office. He is a Fellow of the Royal College of Surgeons of Edinburgh (FRCSEd) and a Fellow of the Royal Society of Medicine (FRMS). Other appointments include convener of the U.K. interdepartmental group on the 3Rs, board member U.K. National Centre for the 3Rs, convener of the International Standards Organization Technical Corrigendum 194/Working Group 3 (*Biocompatibility of Medical Device Materials*), and member of related expert working groups. He is a former member of the EU Committee on Scientific and Technical Progress and past Chairman of the European Commission Technical Expert Working Group on ethical review. He served as chair of the peer review panel for the reduced LLNA test protocol and prediction model for ESAC in 2007 and has been designated as an ESAC peer reviewer for ECVAM's performance standards for the standard LLNA. He served on the ICCVAM/NICEATM Expert Panel that evaluated Five *In Vitro* Pyrogen Test Methods. He has a variety of publications in peer-reviewed journals and national and international meetings, on the principles and practice of surgery, regulation of biomedical research, principles of humane research, bioethics, and public policy.

Peter Theran, V.M.D.

Dr. Theran holds a Doctor of Veterinary Medicine degree from the University of Pennsylvania. He has had many years of experience both as a veterinary internal medicine specialist at the Massachusetts Society for Prevention of Cruelty to Animals' Angell Memorial Animal Hospital in Boston, and as the director of Boston University Medical Center's Laboratory Animal Science Center. He presently serves on a number of government committees as an animal welfare member, and is a member of the Board of Directors of the Institute for *In Vitro* Sciences in Gaithersburg, MD and Chimp Haven in Shreveport, Louisiana. He served on the NICEATM-ICCVAM Expert Panels that evaluated the *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants, and Five *In Vitro* Pyrogen Test Methods. He is a former member of ACATM and SACATM. He is presently working as a consultant.

Stephen Ullrich, Ph.D.

Dr. Ullrich received a Ph.D. in Microbiology from Georgetown University. He is currently the Dallas/Fort Worth Living Legends Professor, and Professor of Immunology at the University of Texas, M.D. Anderson Cancer Center, where he is also Associate Director, The Center for Cancer Immunology Research. He is also a member of the Animal Research Strategic Advisory Committee. He has served numerous national review committees and panels, including the 1998 ICCVAM Peer Panel that evaluated the Murine LLNA. Dr. Ullrich has authored over 75 peer-reviewed publications, over 30 invited articles, and he holds four patents in the U.S., E.U., and Australia for a UV-induced Immunosuppressive Substance. He is the past President of the American Society for Photobiology.

Michael Woolhiser, Ph.D.

Dr. Woolhiser received a Ph.D. in Pharmacology and Toxicology from the Medical College of Virginia at Virginia Commonwealth University. He is a specialist in immunotoxicology and is currently a toxicologist for the Dow Chemical Company where he serves as a Technical Leader for Immunotoxicology, and Polyurethane Business Toxicology Consultant. Dr. Woolhiser is also an Adjunct Professor at the Center for Integrative Toxicology, Michigan State University. He is a member of the Program Committee of the Society of Toxicology's Immunotoxicology Specialty Section. He has served on numerous working groups, including an LLNA Expert Working Group under the European Crop Protection Agency's Toxicology Expert Group, a European Centre for Ecotoxicology and Toxicology of Chemicals LLNA Task Force. He has authored 29 peer-reviewed publications.

Takahiko Yoshida, M.D., Ph.D.

Dr. Yoshida earned his M.D. and a Ph.D. in Medical Science from Tokai University. He is currently Professor in the Department of Health Science at Asahikawa Medical College. Prior to this appointment, he held the posts of Instructor, Assistant Professor and Associate Professor at the Tokai University School of Medicine. He has also been a Guest Researcher at NIEHS. He has also worked as an occupational physician for major Japanese corporations, including Toyota and Sony. Dr. Yoshida’s research interests include occupational health, public health, environmental health and preventative medicine. He is a member of the International Congress of Occupational Health, the Japanese Society of Hygiene, the Japanese Society of Immunotoxicology, the Japanese Society of Clinical Ecology, and the SOT.

[This Page Intentionally Left Blank]

Appendix B

Questions for the Peer Review Panel

[This Page Intentionally Left Blank]

Appendix B1

Questions for the Peer Review Panel: LLNA Limit Dose Procedure

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: LLNA Limit Dose Procedure

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the murine local lymph node assay (LLNA) limit dose procedure as a substitute for the traditional LLNA for assessing the allergic contact dermatitis potential of chemicals and other substances.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA limit dose procedure. You are first asked to review the information in the draft ICCVAM LLNA limit dose procedure Background Review Document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in this BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[1]) have been appropriately addressed for the proposed use of the LLNA limit dose procedure. Adequate validation[2] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. This validation process characterizes the usefulness and limitations of a test method for its intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA limit dose procedure (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft LLNA limit dose procedure BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA limit dose procedure.

These include questions prepared by the ICCVAM IWG to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method, and/or adequate guidance for organizations that may be involved in conducting or supporting further development, standardization, and/or validation.

The overall question to consider is whether the validation status of the LLNA limit dose procedure has been adequately characterized for its intended purpose, and is it sufficiently accurate and reliable to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure when there is not a need for dose response information, in order to reduce the number of animals required for such testing.

I. Questions to the Panel: Review for Errors and Omissions

---

[1] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[2] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

1. Are there any errors or omissions of existing relevant data or information that should be included in the draft BRD?

II. Questions to the Panel: LLNA Limit Dose Procedure Draft BRD

1. For the proposed LLNA limit dose procedure, ICCVAM recommends that the number of animals used in each group should be the same as that recommended by ICCVAM for the traditional LLNA based on its 1998 evaluation of the LLNA, and that individual animal data should be collected and reported (ICCVAM, 1999). Do you agree that these are appropriate provisions for the limit dose procedure? Please explain your answer.
2. Do you consider the traditional LLNA database representative of a sufficient range of chemical classes and physical chemical properties that it would be applicable to any of the types of chemicals and products that are typically tested for skin sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the limit dose procedure? What chemicals or products should be evaluated to fill this data gap? Please explain your answer.
3. While coding of chemicals is recommended for prospective studies, this evaluation is based on a retrospective evaluation of existing data, most of which was not generated using coded chemicals to reduce the potential for bias. Does the lack of coding of test substances adversely impact or bias the current evaluation? Please explain your answer.
4. For some substances submitted using the traditional LLNA test method, it was not possible to confirm whether the data were generated using pooled animal data for each dose group (as allowed in Test Guideline [TG] 429 of the Organisation for Economic Co-operation and Development [OECD]). ICCVAM (1999), Dean et al. (2001), and EPA (2003) recommend the use of statistical analyses to help interpret LLNA study results, which necessitates data collected at the level of the individual animal, while Cockshott et al. (2006) reported that using individual animal data allowed for technical problems during an experiment to be identified. Considering this, should the analysis of the performance of the LLNA limit dose procedure against the traditional LLNA be limited to data from studies that can be confirmed as using individual animal data collection procedures? What impact might the inclusion of pooled animal data have on the accuracy analysis of the LLNA limit procedure? Please explain your answer.
5. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA limit dose procedure been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft ICCVAM BRD)? If not, what other analyses should be performed?
6. There were five substances for which the highest concentration tested produced an SI of less than 3.0, while lower concentrations of these substances produced an SI of greater than 3.0 (see Table 6-2 of the draft ICCVAM BRD). These

substances are classified as "false negatives" compared to what was obtained in the traditional LLNA. Can you identify any characteristics associated with these or other substances that might signal that this type of abnormal dose response might occur, and therefore using the LLNA limit dose procedure would not be appropriate? Please explain your answer.

7. Does the BRD adequately characterize the usefulness and limitations of the LLNA limit dose procedure based on the accuracy analyses? If not, what additions or changes should be made to the current usefulness and limitations? Please explain your answer.
8. Is it appropriate to assume that the intra- and inter-laboratory reproducibility of the LLNA limit dose procedure and the traditional LLNA will be similar, based on the fact that they use identical protocols with the exception of the number of doses used? Do you agree? Does reducing the number of test substances dose groups from three to one reduce the reliability of the assay? Please explain your answer.
9. For some studies included in the draft BRD, it was not possible to determine whether or not they had been conducted in accordance with Good Laboratory Practice (GLP) guidelines. Original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. Considering this, should the results of these studies (all of which are currently included) be excluded from any of the performance analyses? Please explain your answer.
10. Based on the draft BRD, have all the relevant data identified in published or unpublished studies conducted using the traditional LLNA been adequately considered? If not, what other traditional LLNA data needs to be considered and how can it be obtained?

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA Limit Dose Procedure

1. Do the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed test method usefulness and limitations? If not, what recommendations would you make? Please explain your answer.

   - Should the LLNA limit dose procedure be routinely recommended for the hazard identification of skin sensitizing chemicals when potency information is not required? Please explain your answer.
   - If potency information is required, should the LLNA limit dose procedure be routinely recommended as the initial test to identify sensitizers before conducting the traditional LLNA as a way to further reduce animal use, since negative results would not require further testing? Please explain your answer.
   - Based on the existing database, there is a false negative rate of 1.6% (5/313 positive compounds) for the LLNA limit dose approach compared to the results obtained in the traditional LLNA. Do you consider that this is adequately addressed by the proposed cautionary language and weight of evidence consideration for negative substances? Please explain your answer.

2. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed test method standardized protocol? If not, then what recommendations would you make? Please explain your answer.

   - The recommended ICCVAM protocol (ICCVAM 1999; Dean et al. 2001; EPA 2003), as well as OECD TG 429, specifies that the highest dose tested should be the highest soluble concentration that does not induce systemic toxicity and/or excessive skin irritation. However, Kimber et al. (2006) concluded that negative results obtained from studies where the highest concentration tested was below 10% should be considered invalid, and adopted a 10% application concentration as a threshold of confidence for categorization of a chemical as being negative while noting that the figure should not be considered as inviolable. Are the data presented in the draft BRD (i.e., 5/313 positive substances in the NICEATM database were negative at concentrations ≤10%, but were positive at higher concentrations) adequate to conclude that this threshold concentration is not appropriate? If a negative result was obtained for a test substance in a study where the highest concentration that could be tested (based on systemic toxicity or excessive local irritation, as described in ICCVAM [1999], Dean et al. [2001], and EPA [2003]) was <10%, should additional testing be required? Do you agree that the current approach for selecting the "limit" dose is appropriate or do you conclude that there is a threshold concentration for the LLNA at which a negative result could always be considered as an acceptable result? If so, what is that concentration? Please explain your answer.

3. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA limit dose procedure in terms of the proposed future studies? If not, then what recommendations would you make? Please explain your answer.

[This Page Intentionally Left Blank]

Appendix B2

Questions for the Peer Review Panel: LLNA for Testing Aqueous Solutions, Metals, and Mixtures

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: LLNA for Testing Aqueous Solutions, Metals, and Mixtures

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is currently updating the original validation report of the LLNA (ICCVAM 1999) based on a comprehensive review of available data and information regarding the current validity of the LLNA for assessing the skin sensitizing potential of mixtures, metal compounds, and substances tested in aqueous solutions. The information is based on a retrospective review of LLNA data derived from a database of over 500 substances (including mixtures) tested in the LLNA and builds on the previous ICCVAM evaluation of the LLNA, which was based on 209 substances (ICCVAM 1999). In the original ICCVAM report, the performance of the LLNA was compared to 1) the results from guinea pig tests and 2) information about sensitizers in humans (e.g., human maximization test [HMT] results, substances used in human repeat insult patch test [HRIPT], clinical data), where available. This addendum updates the LLNA performance analyses for mixtures, metal compounds, and substances tested in aqueous solutions when compared to human and guinea pig results.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA. You are first asked to review the information in the draft Addendum to the ICCVAM (1999) report for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in this Addendum to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[3]) have been appropriately addressed for the proposed use of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions. Adequate validation[4] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. This validation process characterizes the usefulness and limitations of a test method for its intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft Addendum.

The questions relating to the draft Addendum that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions.

---

[3] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[4] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

These include questions prepared by the ICCVAM IWG to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method, and/or adequate guidance for organizations that may be involved in conducting or supporting further development, standardization, and/or validation.

The overall question to consider is whether the validation status of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions has been adequately characterized, and is it sufficiently accurate and reliable to be used for the identification of sensitizing substances based on a comparison to either human or guinea pig responses.

I. Questions to the Panel: Review for Errors and Omissions

1. In the draft Addendum, are there any errors that need to be corrected or omissions of existing relevant data or information that should be included?

II. Questions to the Panel: Updated LLNA Applicability Domain Addendum

1. Do you consider the database of substances evaluated representative of a sufficient range of mixtures, metal compounds, and substances in aqueous solutions that are typically tested for skin sensitization potential? Please explain your answer.
2. For the purpose of this evaluation, aqueous solutions were defined by the proportion of water (at least 20%) (i.e., substances or mixtures that were tested in an aqueous or an organic:aqueous vehicle were labeled as aqueous solutions). Do you consider this to be an appropriate criterion for defining aqueous solutions? If not, what would be more appropriate? Please explain your answer.
3. While coding of chemicals is recommended for prospective studies, this evaluation is based on a retrospective evaluation of existing data, most of which was not generated using coded chemicals to reduce bias. Does the lack of coding of test substances adversely impact or bias the current evaluation? Please provide a rationale for your answer.
4. For some substances submitted using the LLNA, it was not possible to confirm whether the data were generated using pooled animal data for each dose group (as allowed in Test Guideline [TG] 429 of the Organisation for Economic Co-operation and Development [OECD]) rather than individual animal data (as recommended in the ICCVAM 2001 protocol)? Cockshott et al. (2006) reported that using individual animal data allowed for technical problems during an experiment to be identified. Considering this, should the analysis of the performance of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions be limited to data from studies that can be confirmed as using individual animal data collection procedures? What impact might the inclusion of pooled animal data have on the accuracy analysis included in Section 5.0 of the draft Addendum? Please explain your answer.
5. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions been adequately evaluated and compared to the human and guinea pig (refer also to Section 5.0 of the draft

Addendum)? If not, what other analyses should be performed? Please explain your answer.

6. When multiple LLNA studies were available for the same substance, the majority call (where all studies used the same vehicle and the same concentration range) was used to assign an overall classification for the purposes of the accuracy analysis. For example, if chemical X was tested 5 times and was positive in 3 studies and negative in two, the overall classification was positive. Do you agree with the approach to assigning overall classifications? If not, how would you propose that this be accomplished? Please explain your answer.
7. Does the Addendum adequately characterize the usefulness and limitations of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions based on the accuracy analyses? If not, what additions or changes should be made to the current usefulness and limitations? Please explain your answer.
8. For some studies included in the draft Addendum, it was not possible to determine whether or not they had been conducted in accordance with Good Laboratory Practice (GLP) guidelines. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. Please discuss what impact this lack might have on the evaluation of the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions and whether such studies should be excluded from any analysis.
9. As described in the draft Addendum, original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Considering this, should the results of these studies (all of which are currently included) be excluded from any of the performance analyses? Please explain your answer.
10. Based on the draft Addendum, have all the relevant data identified in published or unpublished studies conducted using the LLNA for testing mixtures, metal compounds, and substances in aqueous solutions been adequately considered? If not, what other studies should to be considered?

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA for Testing Aqueous Solutions, Metals, and Mixtures

1. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA with regard to testing mixtures, metal compounds, and substances in aqueous solutions in terms of the proposed test method usefulness and limitations? Please explain your answer. If not, what recommendations would you make?
2. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA in terms of the proposed test method standardized protocol? Please explain your answer. If not, then what recommendations would the Panel make?

3. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA in terms of the proposed future studies? Please explain your answer. If not, then what recommendations would you make?

Appendix B3

Questions for the Peer Review Panel:
Non-Radioactive LLNA Protocol - LLNA: DA Test Method

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Non-Radioactive LLNA Protocol: LLNA: DA Test Method

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the LLNA: DA (Local Lymph Node Assay-Daicel adenosine triphosphate [ATP]) for assessing the allergic contact dermatitis (ACD) potential of chemicals and other substances. This test method, developed by Daicel Chemical Industries, Ltd. (Tokyo, Japan), is a non-radiolabeled version of the traditional LLNA, and is based on measuring levels of ATP in the auricular lymph nodes as an indicator of increased cell proliferation.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA: DA. You are first asked to review the information in the draft ICCVAM LLNA: DA Background Review Document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in the BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[5]) have been appropriately addressed for the proposed use of the LLNA: DA. Adequate validation[6] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. The validation process characterizes the usefulness and limitations of a test method for a specific intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA: DA (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft LLNA: DA BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA: DA.

These include questions prepared by the ICCVAM Immunotoxicity Working Group (IWG) to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method. The questions are also intended to obtain guidance that will be helpful to federal agencies and other organizations that may be involved in conducting or supporting further development, standardization, and/or validation studies.

The overall question to consider is whether the validation status of the LLNA: DA has been adequately characterized for its intended purpose, and is it sufficiently accurate and reliable

---

[5] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[6] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure.

I. Questions to the Panel: Review of the Draft LLNA: DA BRD for Errors and Omissions

1. In the draft LLNA: DA BRD, are there any errors in the BRD that should be corrected, or omissions of existing relevant data or information that should be included?

II. Questions to the Panel: Draft LLNA: DA BRD

1. Test Method Protocol
   i. The traditional LLNA protocol (ICCVAM 1999; Dean et al. 2001) recommends a minimum of five successfully treated animals per dose group. Current validation of the LLNA: DA was performed using four animals per dose group. What impact might using fewer mice have on the accuracy analysis of the LLNA: DA? Please explain your answer.
   ii. The data generated for the substances analyzed in the LLNA: DA test method come from auricular lymph nodes that were pooled across mice in each dose group rather than being analyzed on an individual animal data. What impact might the inclusion of pooled animal data have on the accuracy analysis of the LLNA: DA? Please explain your answer.
   iii. The LLNA: DA differs from the traditional LLNA in the treatment schedule and by including a pre-treatment with 1% SLS prior to application of the test substance. Do you consider these changes to be appropriate? Please explain your answer.
2. Substances Used for the Validation Studies
   i. Do you consider the LLNA: DA database representative of a sufficient range of chemical classes and physicochemical properties that it would be applicable to any of the types of chemicals and products that are typically tested for skin sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the LLNA: DA? What chemicals or products should be evaluated to fill this data gap? Please explain your answer.
3. Test Method Accuracy
   i. The current accuracy analysis is based on overall concordance with the traditional LLNA. Accuracy statistics compared to the guinea pig tests and human data/experience have also been provided. Are these comparisons appropriate for assessing the accuracy of the LLNA: DA? Please explain your answer.
   ii. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: DA been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft

ICCVAM BRD)? If not, what other analyses should be performed? Please explain your answer.

iii. There was one substance (2-mercaptobenzothiazole) that produced a "false negative" response compared to the traditional LLNA when tested using the LLNA: DA. The mean EC3 in the traditional LLNA for this substance is 2.5 (n=2), it is positive in both the guinea pig and human, and has "high" peptide reactivity as per Gerberick et al. (2007). Can you identify any characteristics associated with this or similar substances, compared to the correctly identified sensitizers, that might signal that this type of discordant response might occur, and therefore using the LLNA: DA to test such substances would not be appropriate or that negative results for such substances should indicate a need for confirmatory testing? Please explain your answer.

iv. There was one substance (benzalkonium chloride) that produced a "false positive" response compared to the traditional LLNA and guinea pig test when tested using the LLNA: DA. Can you identify any characteristics associated with this or similar substances, compared to the correctly identified non-sensitizers that might signal that this type of discordant response might occur, and therefore using the LLNA: DA to test such substances would not be appropriate, or that positive results for substances with such properties may warrant additional testing? Please explain your answer.

4. Test Method Reliability (Intra- and Inter-laboratory reproducibility)

   i. Has the intralaboratory reproducibility of the LLNA: DA been adequately evaluated and compared to the traditional LLNA (refer also to Table 7-1 of the draft LLNA: DA BRD)? If not, what other analyses should be performed? Are any limitations apparent based on this intralaboratory reproducibility assessment? Please explain your answer.

   ii. Has the interlaboratory reproducibility of the LLNA: DA been adequately evaluated and compared to the traditional LLNA (refer also to Tables 7-2 and 7-3 of the draft LLNA: DA BRD)? If not, what other analyses should be performed? Are any limitations apparent based on this interlaboratory reproducibility assessment? Please explain your answer.

   iii. The draft LLNA: DA BRD analyzes data from two interlaboratory validation studies that used coded substances, as well as an intralaboratory validation study with 31 substances that were not coded. Does the lack of coding of test substances adversely impact or bias the current evaluation? In addition, it appears that the lead laboratory established the dose levels tested in the two interlaboratory validation studies and the participating laboratories did not determine their own dose levels for testing. Does this adversely impact or bias the current evaluation? Please explain your answer.

5. Data Quality

   i. The studies evaluated in the draft BRD for the LLNA: DA were not conducted in accordance with Good Laboratory Practices (GLP) guidelines although there were reportedly done in laboratories that conduct GLP studies and were

conducted "in the spirit" of GLP (K. Idehara, personal communication). Please discuss what impact this might have on the evaluation of the LLNA: DA.

ii. The original records for these studies were requested but have not yet been obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Should any recommendations from ICCVAM be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription? Please explain your answer.

6. Consideration of all available data and relevant information

i. Based on the draft LLNA: DA BRD, have all the relevant data identified in published or unpublished studies that employ this test method been adequately considered? Are there other comparative test method data that were not considered in the draft BRD, but are available for consideration? If yes, please explain how to obtain such data.

IV. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA: DA

1. Test Method Usefulness and Limitations

i. Do you agree that the available data and test method performance (accuracy and reliability) support the ICCVAM draft recommendations for the LLNA: DA in terms of the proposed test method usefulness and limitations? Please explain your answer.

ii. If restrictions on using radioactive materials are present, should the LLNA: DA be routinely recommended for hazard identification of skin sensitizing substances in lieu of having to possibly use guinea pig tests? Please explain your answer.

iii. Even if limitations in using radioactive materials are not present, should the LLNA: DA procedure or other valid and accepted non-radioactive method be routinely recommended for hazard identification of skin sensitizing substances instead of the traditional LLNA? Please explain your answer.

iv. From a public health perspective, is the recommended guidance for evaluating negatives sufficient to address concerns associated with the false negative rate of 5% (1/19 substances) calculated for the LLNA: DA? Do you have suggestions for additional guidance or limitations? Please explain your answer.

v. From a testing strategy perspective, does the ICCVAM guidance address concerns associated with the false positive rate of 10% (1/10 substances) calculated for the LLNA: DA? Are there other suggestions for additional such guidance or limitations? Please explain your answer.

2. Test method protocol

i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: DA procedure in terms of the proposed test method standardized protocols? If not, what recommendations would you make? Please explain your answer.

ii. Can the traditional LLNA limit dose procedure be applied to the LLNA: DA? Please explain your answer.

3. Future Studies

i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: DA in terms of the proposed future studies? If not, then what recommendations would you make? Please explain your answer.

4. Performance Standards

i. The LLNA: DA protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001; EPA 2003) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. In addition, there are differences between the two protocols that relate to how and when the test substance is applied and when the lymph nodes are collected (Table 2-1 and Appendix A in the draft LLNA: DA BRD). According to the proposed draft ICCVAM Performance Standards for the traditional LLNA (http://iccvam.niehs.nih.gov/methods/immunotox/llna_PerfStds.htm), any change to the LLNA protocol other than the method used to assess lymphocyte proliferation is considered a major change. Do you agree that these should be considered major changes and therefore the usefulness and limitations of the LLNA: DA should not be assessed using the draft ICCVAM Performance Standards? Please explain your answer.

ii. Even if the draft ICCVAM Performance Standards do not apply to the LLNA: DA, what impact should the accuracy analysis based on 13 of the 18 required performance standards substances (only one false negative and no false positives) have on the overall evaluation of test method accuracy? Please explain your answer.

iii. Should separate performance standards be developed for the LLNA: DA? Please explain your answer.

[This Page Intentionally Left Blank]

Appendix B4

Questions for the Peer Review Panel: Non-Radioactive LLNA Protocol - LLNA: BrdU-FC Test Method

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Non-Radioactive LLNA Protocol: LLNA: BrdU-FC Test Method

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the LLNA: BrdU-FC (Local Lymph Node Assay: Bromodeoxyuridine detected by flow cytometry) for assessing the allergic contact dermatitis potential of chemicals and other substances. This test method, developed by MB Research Labs (Spinnerstown, PA), is a non-radiolabeled version of the traditional LLNA, and is based on measuring the incorporation of bromodeoxyuridine (BrdU) into the DNA of dividing lymphocytes using flow cytometry as in indicator of cell proliferation.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA: BrdU-FC. You are first asked to review the information in the draft ICCVAM LLNA: BrdU-FC Background Review Document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in the BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[7]) have been appropriately addressed for the proposed use of the LLNA: BrdU-FC. Adequate validation[8] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. The validation process characterizes the usefulness and limitations of a test method for a specific intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA: BrdU-FC (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft LLNA: BrdU-FC BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA: BrdU-FC.

These include questions prepared by the ICCVAM Immunotoxicity Working Group (IWG) to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method. The questions are also intended to obtain guidance that will be helpful to federal agencies and other organizations that may be involved in conducting or supporting further development, standardization, and/or validation studies.

The overall question to consider is whether the validation status of the LLNA: BrdU-FC has been adequately characterized for its intended purpose, and is it sufficiently accurate and

---

[7] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[8] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

reliable to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure.

I. Questions to the Panel: Review of the Draft LLNA: BrdU-FC BRD for Errors and Omissions

1. In the draft LLNA: BrdU-FC BRD, are there any errors in the BRD that should be corrected, or omissions of existing relevant data or information that should be included?

II. Questions to the Panel: Draft LLNA: BrdU-FC BRD

1. Test Method Protocol
   i. The LLNA: BrdU-FC includes routine measurements of ear swelling as an indicator of excessive dermal irritation. Do you consider this procedure to be an appropriate approach? Do you think that this measurement should be recommended for routine inclusion into all LLNA protocols? Please explain your answers.
   ii. The LLNA: BrdU-FC also includes optional quantification of immunophenotypic markers as an additional mechanism for distinguishing irritants from sensitizers. Do you consider this to be an appropriate approach to reduce false positives? Are the correct markers being considered or do you recommend other/additional markers? Should these measurements be recommended for routine inclusion in the LLNA: BrdU-FC? Please explain your answers.
   iii. Please comment on the appropriateness of the "sequential strategy" used in the eLLNA: BrdU-FC (see Figure 2-1 of the draft BRD).
2. Substances Used for the Validation Studies
   i. Do you consider the LLNA: BrdU-FC database representative of a sufficient range of chemical classes and physicochemical properties that it would be applicable to any of the types of chemicals and products that are typically tested for skin sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the LLNA: BrdU-FC? What chemicals or products should be evaluated to fill this data gap? Please explain your answers.
3. Test Method Accuracy
   i. The current accuracy analysis is based on overall concordance with the traditional LLNA. Accuracy statistics compared to the Guinea Pig tests and human data/experience have also been provided. Are these comparisons appropriate for assessing the accuracy of the LLNA: BrdU-FC? Please explain your answer.
   ii. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: BrdU-FC been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of

the draft ICCVAM BRD)? If not, what other analyses should be performed? Please explain your answer.

iii. Three substances (benzalkonium chloride, resorcinol, and Tween 80) produced a "false positive" response compared to the traditional LLNA and guinea pig test when tested using the LLNA: BRDU-FC (Based on immunophenotyping, benzalkonium chloride was subsequently classified as an irritant rather than a sensitizer). Can you identify any characteristics associated with these or similar substances, compared to the correctly identified non-sensitizers that might signal that this type of discordant response might occur, and therefore using the LLNA: BrdU-FC to test such substances would not be appropriate, or that positive results for substances with such properties may warrant additional testing? Please explain your answer.

4. Test Method Reliability (Intra- and Inter-laboratory reproducibility)

i. Has the intralaboratory reproducibility of the LLNA: BrdU-FC been adequately evaluated and compared to the traditional LLNA (refer also to Table 7-1 of the draft LLNA: BrdU-FC BRD)? If not, what other analyses should be performed? Are any limitations apparent based on this intra-laboratory reproducibility assessment? Please explain your answer.

ii. The draft LLNA: BrdU-FC BRD analyzes data from repeat testing of hexyl cinnamic aldehyde (HCA) in six different vehicles and intralaboratory reproducibility is assessed by coefficient of variation (CV). The calculated CVs ranged from 30% to 53%. Based on these data, are there concerns with the intralaboratory reproducibility of the LLNA: BrdU-FC? Please explain your answer.

5. Data Quality

i. The studies evaluated in the draft BRD for the LLNA: BrdU-FC were not all conducted in accordance with Good Laboratory Practices (GLP) guidelines although there were done in a laboratory that routinely conducts GLP studies (G. DeGeorge, personal communication). Please discuss what impact this might have on the evaluation of the LLNA: BrdU-FC.

ii. The original records for these studies were requested but have not yet been obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Do you agree that any recommendations from ICCVAM should be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription? Please explain your answer. Consideration of all available data and relevant information

iii. Based on the draft LLNA: BrdU-FC BRD, have all the relevant data identified in published or unpublished studies that employ this test method been adequately considered? Are there other comparative test method data that were not considered in the draft BRD, but are available for consideration? If yes, please explain how to obtain such data.

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA: BrdU-FC

1. Test Method Usefulness and Limitations

    i. Do you agree that the available data and test method performance (accuracy and reliability) support the ICCVAM draft recommendations for the LLNA: BrdU-FC in terms of the proposed test method usefulness and limitations? Please explain your answer.

    ii. If restrictions on using radioactive materials are present, should the LLNA: BrdU-FC be routinely recommended for hazard identification of skin sensitizing substances in lieu of having to possibly use guinea pig tests? Please explain your answer.

    iii. Even if limitations in using radioactive materials are not present, should the LLNA: BrdU-FC procedure be routinely recommended for hazard identification of skin sensitizing substances instead of the traditional LLNA? If not, then why? Please explain your answer.

    iv. Do the ICCVAM recommendations adequately address concerns associated with the false positive rate of 17% (3/18 substances) calculated for the LLNA: BrdU-FC? Are there other suggestions for additional such guidance or limitations that should be considered? Please explain your answer.

2. Test Method Protocol

    i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-FC procedure in terms of the proposed test method standardized protocol? If not, then what recommendations would you make? Please explain your answer.

    ii. Can the traditional LLNA limit dose procedure be applied to the LLNA: BrdU-FC? Please explain your answer.

3. Future Studies

    i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-FC in terms of the proposed future studies? What other recommendations would you make? Please explain your answer.

4. Performance Standards

    i. The LLNA: BrdU-FC protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. According to the proposed draft ICCVAM Performance Standards for the traditional LLNA (http://iccvam.niehs.nih.gov/methods/immunotox/llna_PerfStds.htm), any change to the LLNA protocol other than the method used to assess lymphocyte proliferation is considered a major change. Do you agree that protocol differences between the LLNA: BrdU-FC and the traditional LLNA

should be considered only minor changes and therefore the validity of this test method should be based only on the draft ICCVAM Performance Standards? Please explain your answer.

ii. Even if the draft ICCVAM Performance Standards do not apply to the LLNA: BrdU-FC, what impact should the accuracy analysis based on 13 of the 18 required performance standards substances have on the overall evaluation of test method accuracy? Please explain your answer.

iii. Are there concerns that 3/6 sensitizers, for which EC3 data were available, had EC3 values that were outside of the proposed 0.5x to 2.0x EC3 acceptability range developed based on the traditional LLNA? Please explain your answer.

[This Page Intentionally Left Blank]

Appendix B5

Questions for the Peer Review Panel:
Non-Radioactive LLNA Protocol - LLNA: BrdU-ELISA Test Method

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Non-Radioactive LLNA Protocol: LLNA: BrdU-ELISA Test Method

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the LLNA: BrdU-ELISA (Local Lymph Node Assay with bromodeoxyuridine [BrdU] detected by ELISA) for assessing the allergic contact dermatitis potential of chemicals and other substances. This test method, developed by Dr. Masahiro Takeyoshi (Tokyo, Japan), is a non-radiolabeled version of the traditional LLNA based on measuring levels of incorporated BrdU in the auricular lymph nodes as an indicator of increased cell proliferation.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA: BrdU-ELISA. You are first asked to review the information in the draft ICCVAM LLNA: BrdU-ELISA Background Review Document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in the BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[9]) have been appropriately addressed for the proposed use of the LLNA: BrdU-ELISA. Adequate validation[10] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. The validation process characterizes the usefulness and limitations of a test method for a specific intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA: BrdU-ELISA (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft LLNA: BrdU-ELISA BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA: BrdU-ELISA.

These include questions prepared by the ICCVAM Immunotoxicity Working Group (IWG) to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method. The questions are also intended to obtain guidance that will be helpful to federal agencies and other organizations that may be involved in conducting or supporting further development, standardization, and/or validation studies.

---

[9] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[10] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

The overall question to consider is whether the validation status of the LLNA: BrdU-ELISA has been adequately characterized for its intended purpose, and is it sufficiently accurate and reliable to be used for the identification of sensitizing substances and non-sensitizing substances in place of the traditional LLNA procedure.

I. Questions to the Panel: Comments on the Draft LLNA: BRDU-ELISA BRD for Errors and Omissions

1. In the draft LLNA: BrdU-ELISA BRD, are there any errors in the BRD that should be corrected, or omissions of existing relevant data or information that should be included?

II. Questions to the Panel: Draft LLNA: BrdU-ELISA BRD

1. Test Method Protocol

   i. The data generated for the substances analyzed in the LLNA: BrdU-ELISA test method come from auricular lymph nodes from four individual mice in each dose group. The recommended ICCVAM LLNA protocol and OECD Test Guideline 429 recommend a minimum of five animals per dose group for collecting individual animal data. What impact might the use of four animals per dose group have on the accuracy and reliability of the LLNA: BrdU-ELISA? Do you agree with the ICCVAM recommendation that future use of this test method protocol should include five animals per dose group? Please explain your answer.

2. Substances Used for the Validation Studies

   i. Do you consider the LLNA: BrdU-ELISA database representative of a sufficient range of chemical classes and physicochemical properties that the test method would be applicable to any of the types of chemicals and products that are typically tested for skin sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the LLNA: BrdU-ELISA? What chemicals or products should be evaluated to fill this data gap? Please explain your answers.

3. Test Method Accuracy

   i. The current accuracy analysis using a Stimulation Index (SI) ≥3 or SI ≥1.3 to identify sensitizers is based on overall concordance with the traditional LLNA. Accuracy statistics compared to the Guinea Pig tests and human data/experience have also been provided. Are these comparisons appropriate for assessing the accuracy of the LLNA: BrdU-ELISA? Please explain your answer.

   ii. Takeyoshi et al. (2007) performed an accuracy analysis using decision criteria other than SI ≥3 to classify substances as sensitizers. Maximal accuracy for the LLNA: BrdU-ELISA occurred when an SI ≥1.3 was used to distinguish between sensitizers and non-sensitizers. Using this decision criteria, they achieved an accuracy of 91% (21/23), with a sensitivity of 100% (16/16) and a specificity of 71% (5/7) (i.e., there were no false negatives and two false

positives). Does this analysis support a recommendation that the decision criteria be based on an SI ≥1.3? Are there concerns with using a small increase (i.e., 1.3-fold) above the vehicle control response as the basis for identifying a positive response? Please explain your answers.

iii. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA: BrdU-ELISA, using the SI ≥3 criterion, been adequately evaluated and compared to the traditional LLNA (refer also to Table 6-1 of the draft ICCVAM BRD)? If not, what other analyses should be performed? Please explain your answer.

iv. Using the SI ≥3 criterion, there were four substances (aniline, 4-chloroaniline, 2-mercaptothiazole, and hydroxycitronellal) when tested using the LLNA: BrdU-ELISA that produced "false negative" responses compared to the traditional LLNA. 4-Chloroanline and aniline are amines. 2-Mercaptobenzothiazole is a heterocylic compound and hydroxycitronellal is a hydrocarbon. 2-Mercaptobenzothiazole is a liquid, but the other three substances are solids. Can you identify any characteristics associated with these or similar substances, compared to the correctly identified sensitizers, that might signal that this type of discordant response might occur, and therefore using the LLNA: BrdU-ELISA to test such substances would not be appropriate or that negative results for such substances should indicate a need for confirmatory testing? Please explain your answer.

4. Test Method Reliability (Intra- and Inter-laboratory reproducibility)

   i. Has the intralaboratory reproducibility of the LLNA: BrdU-ELISA been adequately evaluated and compared to the traditional LLNA (refer also to Tables 7-1 through 7-3 of the draft LLNA: BrdU-ELISA BRD)? If not, what other analyses should be performed? Are any limitations apparent based on this intra-laboratory reproducibility assessment? Please explain your answers.

   ii. The substances evaluated for intralaboratory reproducibility of the LLNA: BrdU-ELISA study were not coded. Does the lack of coding of test substances adversely impact or bias the current evaluation? Please explain your answer.

   iii. The Japanese Center for the Validation of Alternative Methods (JaCVAM) has implemented a multi-laboratory validation study of the LLNA: BrdU-ELISA. Although the results from this study have yet to be reported, we are hoping to obtain information on the study design (i.e., with regard to number and types of chemicals tested and the number of laboratories involved). If we do, do you consider the design appropriate to adequately determine the extent of interlaboratory reproducibility for the LLNA: BrdU-ELISA? If not, what other analyses should be performed? Are any limitations apparent based on this study design? Please explain your answer.

5. Data Quality

   i. The studies evaluated in the draft BRD for the LLNA: BrdU-ELISA were not conducted in strict accordance with all provisions of the Good Laboratory Practice (GLP) guidelines, although there were reportedly performed in

laboratories that conduct GLP studies (M. Takeyoshi, personal communication). Please discuss what impact this might have on the evaluation of the LLNA: BrdU-ELISA.

ii. The original records for these studies were requested but were not available. As a result, an independent audit could not be conducted to confirm that the reported data in peer reviewed publications and a poster presentation is the same as the data recorded in laboratory notebooks. Should any recommendations from ICCVAM be contingent upon the completion of such an audit and findings that there were no significant errors in data transcription? Please explain your answer.

6. Consideration of all available data and relevant information

   i. Based on the draft LLNA: BrdU-ELISA BRD, have all the relevant data identified in published or unpublished studies that employ this test method been adequately considered? Are there other comparative test method data that were not considered in the draft BRD, but are available for consideration? If yes, please explain how to obtain such data.

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA: BrdU-ELISA

1. Test Method Usefulness and Limitations

   i. Do you agree that the available data and test method performance (accuracy and reliability) support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA in terms of the proposed test method usefulness and limitations? Please explain your answer.

   ii. If restrictions on using radioactive materials are present, should the LLNA: BrdU-ELISA be routinely recommended for hazard identification of skin sensitizing substances in lieu of using guinea pig tests due to the advantages of fewer animals and the avoidance of pain and distress? Please explain your answer.

   iii. Even if limitations in using radioactive materials are not present, should the LLNA: BrdU-ELISA procedure or other valid and accepted non-radioactive method be routinely recommended for hazard identification of skin sensitizing substances instead of the traditional LLNA? Please explain your answer.

   iv. Does using a decision criterion of SI $\geq$1.3 instead of SI $\geq$3.0 resolve any concerns with respect to potential false positives or false negatives that may occur in this test method? Are there other suggestions for additional such guidance or limitations that should be considered? Please explain your answer.

2. Test Method Protocol

   i. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA procedure in terms of the proposed test method standardized protocols? If not, then what recommendations would you make? Please explain your answer.

   ii. Can the traditional LLNA limit dose procedure be applied to the LLNA: BrdU-ELISA? Please explain your answer.

3. Future Studies

   a. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA: BrdU-ELISA in terms of the proposed future studies? If not, then what recommendations would you make? Please explain your answer.

i. Performance Standards

   i. The LLNA: BrdU-ELISA protocol differs from the ICCVAM-recommended protocol for the traditional LLNA (ICCVAM 1999; Dean et al. 2001) in the method used to assess lymphocyte proliferation in the auricular lymph nodes. According to the proposed draft ICCVAM Performance Standards for the traditional LLNA (http://iccvam.niehs.nih.gov/methods/immunotox/llna_PerfStds.htm), any change to the LLNA protocol other than the method used to assess

lymphocyte proliferation is considered a major change. Do you agree that protocol differences between the LLNA: BrdU-ELISA and the traditional LLNA should be considered only minor changes and therefore the validity of this test method should be based only on the draft ICCVAM Performance Standards? Please explain your answer.

ii. Even if the draft ICCVAM Performance Standards do not apply to the LLNA: BrdU-ELISA, what impact should the accuracy analysis based on eight of the 18 required performance standards substances (only one false negative and no false positives) have on the overall evaluation of test method accuracy? Please explain your answer.

iii. Are there concerns that 4/4 sensitizers, for which EC3 data were available, had EC3 values that were outside of the recommended 0.5x to 2.0x EC3 acceptability range developed based on the traditional LLNA? Please explain your answer.

iv. Should separate performance standards be developed for the LLNA: BrdU-ELISA? Please explain your answer.

Appendix B6

Questions for the Peer Review Panel: Draft ICCVAM Performance Standards for the LLNA

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Draft ICCVAM Performance Standards for the LLNA

ICCVAM has developed draft LLNA performance standards consisting of essential test method components, a minimum list of reference substances, and expected accuracy and reliability values. These are proposed for evaluating the acceptability of proposed test methods that are mechanistically and functionally similar to the traditional LLNA test method protocol previously recommended by ICCVAM.

The overall question for the Panel is whether these performance standards are considered adequate for assessing the accuracy and reliability of test method protocols that are based on similar scientific principles and that measure the same biological effect as the traditional LLNA.

1) Purpose and Applicability

a) ICCVAM proposes that these performance standards should only be applicable to versions of the LLNA that incorporate minor modifications to the traditional LLNA. Currently, this is limited to the use of non-radioactive reagents to measure lymphocyte proliferation. It is considered essential that the modified LLNA should otherwise adhere to all other aspects of the traditional LLNA protocol, as defined by ICCVAM (1999) and Dean et al. (2001). This includes aspects such as: the sex and strain of mouse used, the number of mice per dose group, the timing and site of test article treatment, the duration between the last treatment and lymph node collection, the inclusion of concurrent negative and positive control groups, the measured endpoint (i.e., lymphocyte proliferation in the draining auricular lymph node), and the collection of data at the level of the individual mouse. Do you agree that the use of non-radioactive reagents for measuring cell proliferation in the lymph nodes, if that is the only difference, constitutes a minor modification to the traditional LLNA protocol? Is it necessary to keep the same decision criteria for distinguishing between sensitizers and non-sensitizers (i.e., an SI of 3)? Please explain your answer.

b) Are there other procedural modifications that you consider minor and therefore could be evaluated for equivalence to the traditional LLNA using the proposed performance standards? If yes, please explain what they are and why.

c) Do you consider these performance standards to also be applicable to the LLNA limit dose procedure? Please explain your answer.

2) Essential Test Method Components

a) The essential test method components are based on the ICCVAM recommended protocol (ICCVAM 1999; Dean et al. 2001), which is the basis for the current U.S. Environmental Protection Agency (EPA 2003) test guideline (TG). There are some notable differences between these protocols and the Organisation for Economic Co-operation and Development TG 429 for the LLNA (OECD 2002). When evaluations of non-radioactive versions of the traditional LLNA are conducted using these performance standards, is it necessary that the validation studies follow the ICCVAM recommended protocol? Specifically, should the studies include: 1) a concurrent positive control with each test substance; 2) using a minimum of five animals per dose group; and 3) measuring proliferation in lymph nodes from individual animals

rather than pooling lymph nodes across all animals in a dose group? Please explain your answers.

b) Should the concurrent testing of the positive control and test substance be conducted in the same vehicle or can different vehicles be used? Please explain your answer.

3) Proposed Reference Substances

a) Do you agree with the selection and prioritization criteria used to select the performance standards reference substances? Please explain your answer.

b) The rationale for the number of substances included on the "required" list of substances (n=18) is provided in Appendix C of the draft ICCVAM Performance Standards. Do you consider this to be an adequate number upon which to evaluate the performance of non-radioactive LLNA test methods, where the only protocol modification is the method for assessing cell proliferation in the auricular lymph nodes? If not, how many reference chemicals should be tested? Please explain your answer.

c) Do you consider the types of substances included in the reference substance list, with regard to relative sensitization potency, physicochemical characteristics, and vehicles, to be representative of the overall diversity of substances that are likely to be tested for skin sensitization? Please explain your answer.

d) Are there other types of information relevant to skin sensitization that should be considered in order to demonstrate an adequately diverse reference list? If yes, please explain what additional information should be included.

e) Are there other substances that you consider to be more appropriate for assessing the sensitivity (ability of the test method to correctly identify sensitizing substances) and specificity (ability of the test method to correctly identify non-sensitizing substances) of non-radioactive LLNA test methods, and for which there is available LLNA, guinea pig, and human data? If yes, please name the substances and explain why.

f) Four "discordant chemicals" (i.e., two LLNA false negatives and two false positives compared to guinea pig tests or human data) are included as optional substances that could be studied to evaluate if the proposed modifications might provide improved performance relative to the traditional LLNA.

- Please comment on the appropriateness of including these specific substances in the reference list. Should different substances be included? Should more false negative/positive substances be tested? If so, what are they? Please explain your answers.
- Do you consider their "optional" status appropriate, or should testing these substances be required? Please explain your answer.
- Would "correct" results with these four discordant chemicals be sufficient to consider the alternative test method to be more predictive of skin sensitization than the traditional LLNA? Please explain your answer.

4) Test Method Accuracy Standards

a) The draft ICCVAM Performance Standards state that the non-radioactive proposed LLNA test method should exactly match the accuracy of the traditional LLNA when evaluated with the minimum set of 18 reference substances. Do you agree that test method accuracy should be based on a chemical-by-chemical match with regard to identifying the chemicals as sensitizers or non-sensitizers? Please explain your answer.

b) The draft ICCVAM LLNA Performance Standards recommend that, for each sensitizer, the threshold concentration that induces a positive SI response should be within 0.5x to 2.0x of the concentration obtained for the EC3 in the traditional LLNA. As described in Appendix D of the draft ICCVAM Performance Standards, statistical approaches have been used in an attempt to identify an appropriate range, but these calculated ranges do not appear to be the most practical. In contrast, the NICEATM LLNA database demonstrates that EC3 values from replicate tests for a sensitizing chemical when tested using the same solvent are rarely outside of this proposed 0.5x to 2.0x acceptability EC3 range. Please comment on the appropriateness of using this criterion to judge the equivalency of a non-radioactive version of the traditional LLNA. If this approach is not acceptable, please explain why, and present an alternative approach along with the basis for this approach.

c) For five of the 13 sensitizers on the draft ICCVAM reference substances list, the reference EC3 value is based on a single LLNA study (see Table C1 of the draft ICCVAM Performance Standards). Please comment on the appropriateness of including such chemicals in the list of recommended reference substances and whether or not the 0.5x to 2.0x criteria should be applied to such substances. Please explain your answer.

5) Test Method Reliability Standards

a) The draft ICCVAM Performance Standards state that acceptable intralaboratory reproducibility will be indicated by a laboratory obtaining, in each of four independent experiments conducted with at least one week between each experiment, ECt values (the estimated concentration needed to produce an SI of a defined threshold [e.g., EC3]) for hexyl cinnamic aldehyde (HCA) that are generally within 0.5x to 2.0x (i.e., 5% to 20%) of the historical mean EC3 concentration (10%) for this substance, based on existing available traditional LLNA data.

- Do you consider the number of repeat experiments (n=4) to be adequate? Please explain your answer.
- Do you consider testing HCA adequate for demonstrating intralaboratory reproducibility? If not, which substance(s) should be tested? Please explain your answer.
- Is the required one-week interval between independent tests adequate and/or appropriate? If not, please provide an alternative schedule and explain the basis for your recommendation.
- Do you consider the criteria for acceptability to be appropriate? If not, please describe another criteria and explain the basis for your recommendation.

b) The draft ICCVAM Performance Standards state that acceptable interlaboratory reproducibility will be indicated by each of three laboratories obtaining ECt values for HCA and 2,4-dinitrochlorobenzene (DNCB) from a single experiment that are generally within 0.5x to 2.0x (5% to 20% and 0.025 to 0.1%, respectively) of the mean historical EC3 concentration (10% and 0.05%, respectively) obtained for these two substances in the traditional LLNA.

- Do you consider the single experiment per substance in each laboratory to be adequate? If not, please provide an alternative approach and explain why.
- Do you consider testing HCA and DNCB to be adequate for demonstrating interlaboratory reproducibility? If not, which substance(s) should be tested? Please explain your answer.
- Do you consider the criteria for acceptability to be appropriate? If not, please describe another criteria and explain the basis for your recommendation.

6) Summary Question

a) If a radioactive or non-radioactive LLNA method were proposed with a "major change" (e.g., different mouse strain or use of male mice, change in the schedule for test article administration, change in schedule for lymph node excision, etc.), what criteria should be used to evaluate the equivalence of this method to the traditional LLNA?

- Would a new set of performance standards be required for this method? Please explain your answer.
- How many reference substances might be considered adequate? Would the 18 minimum reference substances in the draft ICCVAM LLNA Performance Standards be sufficient? If more substances are considered necessary, how many should there be tested and what should their characteristics be? Please explain your answer.
- Regardless of the number of reference substances, should the alternative LLNA be required to obtain the same call (and potency for sensitizers) as the traditional LLNA for the 18 minimum reference substances in the draft ICCVAM LLNA Performance Standards? Please explain your answer.
- Are there additional specific substances that should be used? If yes, what are they? Please explain your answer.
- What, if any, additional information would be considered necessary and why?

Appendix B7

Questions for the Peer Review Panel: Use of the LLNA for Potency Determinations

[This Page Intentionally Left Blank]

Instructions for the Peer Review Panel: Use of the LLNA for Potency Determinations

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in conjunction with National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is evaluating the validation status of the murine local lymph node assay (LLNA) for assessing the potential potency of substances to cause allergic contact dermatitis.

The following materials are intended to guide you in your review of the ICCVAM evaluation of the LLNA for potency determinations. You are first asked to review the information in the draft ICCVAM LLNA for potency determinations Background Review Document (BRD) for completeness, and to identify any errors or omissions of existing relevant data or information that should be included. You are then asked to evaluate the information in this BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[11]) have been appropriately addressed for the proposed use of the LLNA for potency determinations. Adequate validation[12] is a prerequisite for a test method to be considered for use in regulatory decision-making by U.S. Federal agencies. This validation process characterizes the usefulness and limitations of a test method for its intended use.

Lastly, you are asked to consider the ICCVAM draft test method recommendations for the LLNA for potency determinations (i.e., the proposed test method use, the proposed recommended standardized protocol, the proposed test method performance standards, and any proposed additional studies) and comment on whether the recommendations are supported by the information provided in the draft BRD.

The questions relating to the draft BRD that must be addressed are provided in **Sections I** and **II** of this guidance, while **Section III** contains questions relating to the draft ICCVAM test method recommendations on the LLNA for potency determinations.

These include questions prepared by the ICCVAM IWG to ensure that the assessment provides adequate information to facilitate U.S. Federal agency decisions on the regulatory acceptability of this test method, and/or adequate guidance for organizations that may be involved in conducting or supporting further development, standardization, and/or validation.

The overall question to consider is whether the validation status of the LLNA for potency determinations has been adequately characterized, and is it sufficiently accurate and reliable to be used for the identification of sensitizing substances according to their relative potency classification based on a comparison to either human or guinea pig responses.

I. Questions to the Panel: Review for Errors and Omissions

1. Are there any errors or omissions of existing relevant data or information that should be included in the draft BRD?

---

[11] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC. NIEHS (Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

[12] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

II. Questions to the Panel: LLNA for Potency Determinations Draft BRD

1. Do you consider the database of substances evaluated representative of a sufficient range of chemical classes and physical chemical properties that it would be applicable to any of the types of chemicals and products that are typically tested for skin sensitization potential? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the traditional LLNA) that should be tested with caution, or not evaluated using the for potency determinations? What chemicals or products should be evaluated to fill this data gap? Please explain your recommendation.
2. While coding of chemicals is recommended for prospective studies, this evaluation is based on a retrospective evaluation of existing data, most of which was not generated using coded chemicals to reduce bias. Does the lack of coding of test substances adversely impact or bias the current evaluation? Please provide a rationale for your answer.
3. For some substances submitted using the LLNA, it was not possible to confirm whether the data were generated using pooled animal data for each dose group (as allowed in Test Guideline [TG] 429 of the Organisation for Economic Co-operation and Development [OECD]) rather than individual animal data (as recommended in the ICCVAM 2001 protocol)? Cockshott et al. (2006) reported that using individual animal data allowed for technical problems during an experiment to be identified. Considering this, should the analysis of the performance of the LLNA for potency determinations be limited to data from studies that can be confirmed as using individual animal data collection procedures? What impact might the inclusion of pooled animal data have on the accuracy analysis included in Section 6.0 of the draft ICCVAM BRD? Please explain your answer.
4. Has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive and false negative rates) of the LLNA for potency determinations been adequately evaluated and compared to the human and guinea pig (refer also to Section 6.0 of the draft ICCVAM BRD)? If not, what other analyses should be performed?
5. The accuracy analysis (see Section 6.0 of the draft ICCVAM BRD) focuses on the two-level categorization scheme proposed by the United Nations Globally Harmonized System for Classification and Labelling for both human and guinea data. Should other categorization schemes be considered?
6. Does the BRD adequately characterize the usefulness and limitations of the LLNA for potency determinations based on the accuracy analyses? If not, what additions or changes should be made to the current usefulness and limitations?
7. Has the reliability (e.g., intralaboratory repeatability, intra- and inter-laboratory reproducibility) of the LLNA for potency determinations been adequately evaluated (refer also to Section 7.0 of the draft ICCVAM BRD)? If not, what other analyses should be performed?

8. For some studies included in the draft BRD, it was not possible to determine whether or not they had been conducted in accordance with Good Laboratory Practice (GLP) guidelines. Neither was it possible to obtain the results of GLP audits for all studies conducted in accordance with GLP guidelines. Please discuss what impact this lack might have on the evaluation of the LLNA for potency determinations and whether such studies should be excluded from any analysis.
9. As described in the draft BRD, original records for some of the non-GLP studies included in this evaluation could not be obtained. As a result, an independent audit could not be conducted to confirm that the reported data is the same as the data recorded in laboratory notebooks. Considering this, should the results of these studies (all of which are currently included) be excluded from any of the performance analyses? If yes, please explain.
10. Based on the draft BRD, have all the relevant data identified in published or unpublished studies conducted using the LLNA for potency determinations been adequately considered? If not, what other studies should to be considered?

III. Questions to the Panel: Draft ICCVAM Test Method Recommendations on the LLNA for Potency Determinations

1. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA for potency determinations in terms of the proposed test method usefulness and limitations? Please explain your answer. If not, why recommendations would you make?
    - Should the LLNA be routinely recommended for the hazard classification of the skin sensitization potency of chemicals?
2. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA in terms of the proposed test method standardized protocol? Please explain your answer. If not, then what recommendations would you make?
3. Should the relevant testing guidelines for the LLNA be updated to include the calculation of an EC3 value?
4. Do you agree that the available data support the ICCVAM draft recommendations for the LLNA in terms of the proposed future studies? Please explain your answer. If not, then what recommendations would you make?