(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Regulatory Toxicology and Pharmacology 64 (2012) 9–16

Contents lists available at SciVerse ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Workshop Report

Optimised testing strategies for skin sensitization – The LLNA and beyond

David Basketter [a,*], Jonathan Crozier [b], Bruno Hubesch [c], Irene Manou [b], Annette Mehling [d], Julia Scheel [e]

[a] DABMEB Consultancy Ltd., Sharnbrook, UK
[b] European Partnership for Alternative Approaches to Animal Testing, Brussels, Belgium
[c] CEFIC, Brussels, Belgium
[d] BASF Personal Care and Nutrition GmbH, Dusseldorf, Germany
[e] Henkel, Dusseldorf, Germany

ARTICLE INFO

Article history:
Received 18 April 2012
Available online 17 June 2012

Keywords:
Skin sensitization
Local lymph node assay
In vitro alternatives
Risk assessment

ABSTRACT

As toxicology in the 21st century progresses towards a future which aims at avoiding the use of *in vivo* testing, the endpoint of skin sensitisation can now be found in the front line. Accordingly, it was appropriate for several industry sectors to meet and review what has been learned from the currently most widely used *in vivo* method, the local lymph node assay (LLNA), and to consider the status of progress as we attempt to move beyond that test. No toxicology test is perfect, an experience brought into focus by issues of false positives and, to a lesser extent, false negatives in the LLNA. Use of weight of evidence arguments for classification and labelling, as well as for risk assessment was emphasised and it was also noted that a sufficient body of evidence now exists for conduct of methods other than the LLNA for carefully defined chemical classes. In terms of *in vitro* alternatives, progress towards methods which will deliver mainly hazard identification is being made, with some entering the final stages of validation, whereby (Q)SAR tools still need *improvement* to be used on a large scale in practise. As various other challenges also remain, e.g. testing lipophilic substances, as well as the development of non-animal methods which deliver reliable information on potency for risk assessment, these will remain a topic for continuing research and development.

1. Introduction

Skin sensitisation remains a key endpoint in the toxicological evaluation of chemicals and of the preparations into which they are placed. Accordingly, the test methods used to identify these effects in terms of hazard and risk remain an important focus for toxicologists. In the last decade, in the EU the previous dominance of guinea pig methods (reviewed in Buehler, 1965; Magnusson and Kligman, 1970; Andersen and Maibach, 1985) has been overtaken by a murine method, the local lymph node assay (LLNA) (Kimber and Basketter, 1992; Basketter and Chamberlain, 1995; Dean et al., 2001). This change has come about because the LLNA offers refinement/reduction of animal use as well as delivering an objective, quantitative endpoint (Balls and Hellsten, 2000) enabling a more precise estimate of potency. Consequently, it has become the preferred test in many regulatory areas, including REACH (Commission of the European Communities, 2006). However, with the drive towards non-animal alternatives and the identification of viable options for that purpose, there is also the real probability that the LLNA will in its turn also be replaced in the medium term (Adler et al., 2011; Kimber et al., 2011).

1.1. About LRI

Launched 13 years ago, the Long-Range Research Initiative (LRI) is one of the major voluntary initiatives of the European chemical industry to support its competitiveness and innovation potential. LRI aims to identify and fill gaps in our understanding of the hazards posed by chemicals and to improve the methods available for assessing the associated risks. LRI sponsors high quality research, published in peer reviewed journals, and seeks to provide sound scientific advice on which industry and regulatory bodies will draw to respond more quickly and accurately to the public's concerns. More details can be found on the organisation's website: www.cefic-lri.org.

1.2. About EPAA

Founded in 2005, the European Partnership for Alternative Approaches to Animal Testing (EPAA) is a unique platform gathering industry & European Commission across different sectors. The EPAA benefits from the strong political support of all commissioners involved in policy areas relevant to alternative approaches.

* Corresponding author. Address: DABMEB Consultancy Ltd., Sharnbrook MK44 1PR, UK. Fax: +44 123478944.
E-mail address: dabmebconsultancyltd@me.com (D. Basketter).

0273-2300/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.yrtph.2012.06.003

During its second mandate covering a 5 years period (2011–2015), the EPAA aims to further promote the application of Russell and Burch's 3Rs principles in the framework of regulatory compliance procedures. The EPAA contributes to the uptake of 3Rs oriented approaches by assessing their implementation and scope throughout different industry sectors, mainly at European level. The EPAA promotes a holistic approach, from research & development and the use of 3Rs approaches in regulation to communication and dissemination. The EPAA has an Action Plan, focusing on selected priority items in which the EPAA, as a public private partnership, can make a valuable contribution, likely to produce significant progress. More details can be found on the organisation's website: www.epaa.eu.com.

For the time being, in terms of regulatory toxicology within the EU, the LLNA remains the assay of first choice in most regulatory sectors, with typically a significant scientific case being required for the conduct of an *in vivo* alternative, such as one of the guinea pig methods (e.g. under REACH – Commission Regulation, 2011). Where this is not the case, e.g. the EU Biocides Directive which is over 10 years old and currently under revision, then there is more flexibility. However, the LLNA is also the most common assay selection when it is necessary to measure the relative potency of an identified skin sensitiser, e.g. in the cosmetic industry for safety assessment (Maxwell et al., 2011; Adler et al., 2011). Similar trends towards use of the LLNA for hazard characterisation have in fact occurred in a number of industry sectors, e.g. pesticides and agrochemicals (Vohr and Ahr, 2005).

In the first part of this report, the practical experiences associated with the adoption of the LLNA as the skin sensitisation test of first choice are reviewed and considered in terms of their didactic value as the progression to *in vitro* alternatives gathers pace and experiences with *in silico* tools are becoming available. The workshop explored how different industry sectors deal with some of the problems and how they have deployed variations and additions to the standard LLNA protocol. In the second part it will be elucidated how they are developing, assessing and implementing *in vitro* and *in silico* alternatives. Each of these topics is explored in some detail, presenting recommendations for further work or action at each stage, and then an effort is made to summarise the key points and main recommendations.

2. Experience of use of the LLNA in practise

A separation is often made between regulatory hazard identification on the one hand, and the process of risk assessment on the other. Risk assessment utilises information obtained from hazard assessments and combines it with an evaluation of exposure. The basis for hazard classification and labelling are generally very clearly regulated, e.g. via the globally harmonised system (GHS) or other national legislation. Depending on the area of use, e.g. as a cosmetic, pharmaceutical or agrochemical ingredient, substances and formulations may be subject to different assessments of risk depending on the regulatory setting and use. Consequently, in this section the use of the LLNA from hazard identification through to final risk assessment is considered, taking insights from different sectors, and with an eye firmly on what the future may hold.

The LLNA was the first alternative toxicology test to successfully pass formal independent validation by ECVAM and ICCVAM (Kimber and Basketter, 1992; Gerberick et al., 2000). This has led to the development of a range of practical experiences, from both industry and regulatory bodies. As the LLNA is now the prescribed method under REACH, for the first of these groups, a dominant issue has been the reliable identification of a sensitization hazard without the generation of an excessive number of false positive results (Basketter et al., 1998, 2009a; Vohr and Ahr, 2005; Kreiling et al., 2008; Garcia et al., 2010; Ball et al., 2011). For the latter, there has been a more general sharing of the experience with the variety of challenges presented by regulatory submissions (Cockshott et al., 2006; McGarry, 2007; Angers-Loustau et al., 2011). This body of work has also, on occasion, prompted response from the originators of the LLNA (Basketter et al., 1998, 2006, 2009b; Cockshott et al., 2006; Basketter and Kimber, 2007). Elements of these experiences were further discussed in this workshop. However, it is not the purpose of this section to review the published body of work, but rather to examine some further, new, material which addressed both the strengths and limitations of the LLNA.

To exemplify the newer data, a number of case studies were presented. The preliminary evaluation of polyfunctional silicones described in an earlier publication (Basketter et al., 2009a,b) has been extended to a series of five. In every case the GPMT was negative, but 4/5 LLNA studies were positive. In regulatory terms, both assays are equally acceptable, therefore the way to judge which outcome represents the truth has to be by application of a weight of evidence review (as clearly indicated in REACH regulations – Commission Regulation (EU) No 286/2011 of 10 March 2011). The arguments included:

- Low dermal penetration due to high molecular weight (>10,000 Daltons)
- Absence of any functional groups known to be associated with skin sensitisation
- Absence of occupational allergic contact dermatitis in workers with daily skin exposures for a period of more than 10 years
- A history of safe use by consumers, e.g. via use in shampoos at up to 2%
- The absence of sub-classification effects in the GPMT at high test concentrations
- Weak LLNA responses with no evidence of dose response

Weight of evidence arguments of this type are quite generally applicable. A presentation from the surfactant industry illustrated a similar theme. A total of seven substances were positive in the standard LLNA but negative in the GPMT, whereas only one was positive in both the GPMT and LLNA. The first material, sodium lauryl sulphate (SLS) is the classic historical false positive from the LLNA and is so again in the study presented. It is much more irritant than other materials tested (compare the challenge concentrations selected for the GPMT, which are the maximum non-irritant levels) and it duly causes a >20% increase in ear weight in the mouse, used as a marker of irritant effect (Kirk et al., 2007; Basketter et al., 2012). Data on other glycol ethers was presented in which the LLNA yielded positive results and the guinea pig tests negative results. Based on the additional data presented (e.g. *in vitro* tests) along with a history of safe use, we know that substances such as these glycol ethers are in reality devoid of sensitising effects and that these are false positives in the LLNA. These results and a possible role of irritation and IL-1 alpha are discussed in more detail in Ball et al., 2011. Further examples were presented for fatty acid-like substances, in which double-bonds may be implicated in the causative mechanism. The results of these studies are discussed in more detail elsewhere (Kreiling et al., 2009).

Hazard and risk assessments not only impact the consumer but also workers in the occupational setting. This is not only true for the cosmetic and household product sector but also for the pharmaceutical and agrochemical sector. Based on the hazard or risk, risk management measures need to be implemented which can entail wearing protective clothing or using dilutions of a material or even a prohibition of use. The LLNA is to date the "gold standard" for the assessment of the potency of a sensitizer which essential for risk assessments. Its applicability to assess the sensitization

potential of formulations is still being discussed. Examples of challenges generated by the transition from hazard to risk assessments by regulatory bodies were presented as were the implications for safe use of formulations. The development of suitable exposure assessments and the quantitative transfer of animal data to humans will need to be addressed in more detail in the future.

Ultimately, the most important conclusion from this element of the workshop was that where the limitations of the LLNA have become more evident, then the use of the guinea pig model for certain types of substance are justified. In addition, whichever test is applied, the overall evaluation of a substance must be made on the basis of the weight of all the evidence, not only on the outcome of a single test. Unfortunately, this may prove difficult when assessing the sensitization potential of a new class of substances which will be a further challenge in the future.

2.1. Conclusions

There is now an extensive body of practical experience with the standard LLNA which has served to clarify both its strengths and its weaknesses. Details of the primary conclusions are presented at the end of the following section on LLNAs with modified protocols.

2.2. Recommendations

These follow at the end of the subsequent section.

3. Experience with modifications to the standard LLNA

Although the standard LLNA using tritiated thymidine uptake as the endpoint was the assay adopted into an OECD Test Guideline approximately a decade ago (OECD, 2002), there has been a prolonged search for non-radioactive alternatives (reviewed in Gerberick et al., 2008). Recently, this has led to the adoption into OECD Test Guidelines of two non-radioactive LLNA alternatives that had the benefit of a sufficient body of supporting evidence concerning their sensitivity and specificity when judged against internationally agreed Performance Standards (Basketter et al., 2008a,b; OECD, 2010a,b). A further consideration was whether it was possible to address photo effects in a modified LLNA.

A presentation, under the title "Modification of the LLNA – a never ending story" provided insight into the history of the LLNA and its modifications. An overview was given on the studies published during the development of the radioactive LLNA, progress in the development nonradioactive alternative endpoints and how to address the topic of the specificity of the assay (Kimber et al., 1989; Vohr and Ahr, 2005; Ehling et al., 2005, Omori et al., 2008). An important point made was that lymph node cell counts offer a valuable alternative to the radioactive endpoint (e.g. Gamer et al., 2008). The potential utility of cell counting was reinforced in a recent publication (Basketter et al., 2012, in press) and substantially expanded in data presented at this workshop by other speakers. Importantly, measurement of the threshold concentration required to produce a positive result from the cell count method, the EC1.5, showed excellent correlation with the LLNA EC3 value.

One modification to the LLNA which cropped up in several presentations was the use of measures of irritancy and/or non-immune inflammatory responses as supplementary endpoints which might enhance the specificity of the LLNA (both radioactive and non-radioactive variants). Measures such as ear swelling (see above) and ear weight have been proposed (Ehling et al., 2005). Although sometimes indicative, an irrefutable correlation between the presence of ear swelling and false positive results was not found based on the data presented in this workshop. One potential factor complicating this may be the chronological differences in the time course of the irritant and sensitisation responses (Ehling et al., 2005; Ulrich and Vohr, in preparation). However, based on ear swelling, no irritation was found when testing fatty acid like substances. Potential alternate explanations included the role of unsaturation in fatty acid chains (see Kreiling et al., 2008). For other materials (e.g. polyfunctional silicones) there was also little evidence for skin irritancy as an underlying cause of false positive LLNA results. For certain substances, the presence of a degree of skin irritation was associated with a positive LLNA skin sensitisation result for which the weight of evidence pointed towards this being an incorrect result (e.g. certain surfactants). In the case of surfactants, there was a certain degree of correlation between ear swelling and lymph node proliferation yet there was a clear indication of potential skin irritation in at least 3 of 4 different tests conducted with these substances and not based on ear swelling (Ball et al., 2011). In particular, there was a high correlation between IL-1α, a key marker of irritation, and the false positive results found in the LLNA. An influence between ear swelling and lymph node hyperplasia was also found in other cases. Data presented also looked into possible effects via immunophenotyping of the draining lymph node cells to identify shifts in lymphocyte subsets that could be of help in discriminating sensitisers and irritants. No definitive correlations were found and this did not improve the predictive capacity of the LLNA (Ulrich et al., 2001a,b).

A further adaption of the LLNA, potentially used in combination with modifications discussed above, has been to explore the utility of the method for the predictive identification of photosensitising chemicals. Early efforts in this area were not successful, failing to differentiate photosensitisers from photoirritants (Scholes et al., 1992). Further work indicated the usefulness of the UV-LLNA in differentiating between photoirritants and photosensitizers (Vohr et al., 1994, 2000; Neumann et al., 2005). Continued efforts to identify additional means to enhance the specificity of the assay have been applied. This has particular resonance for the pharmaceutical industry where over 250 drug substances have been associated with confirmed photosensitisation of which 90 are of therapeutic relevance, a number which greatly exceeds the problem associated with other industry sectors, e.g. cosmetics, household products. An additional issue in the pharmaceutical field is the oral uptake instead of dermal uptake. Developments of a biphasic (challenge) LLNA protocol including an UV-LLNA protocol were presented. In addition to the need to ensure a suitable and properly calibrated solar simulator light source and avoiding the influence of increased temperature due to irradiation, several other key features were detailed, e.g. the impact of the photo exposure on baseline/control measurement; identifying suitable dose–response relationships for the induction phase; use of lymphocyte markers and cytokine profiles to support interpretation and the importance of taking the individual drug photo reactivity and pharmacokinetic profile into account when designing the biphasic UV-LLNA which means that there cannot be a standard protocol. The applicability of this approach outside the pharmaceutical domain has yet to be established, but in principle there is no reason why it should not work.

3.1. Conclusions

- there is now a sufficient body of evidence for specific substance groups to provide a sound scientific argument for conduct of an guinea pig *in vivo* test (OECD 406) as a substitute for the LLNA
- there is similarly a sufficient body of evidence (and a regulatory requirement) that weight of evidence considerations rather than data from individual methods should be the basis for assessment

- the potency assessment of identified skin sensitisers is a critical issue for complete replacement of the current *in vivo* tests
- there are now validated non-radioactive radioactive endpoint LLNA options, as well as further variants capable of consideration
- the specific causes of false positive LLNA results remains unclear; skin irritation or certain molecular structures, e.g. unsaturated bonds, may sometimes be involved
- variants of the LLNA which measure irritation via ear swelling, ear weight, etc. have been established, but not yet formally validated
- variants of the LLNA which have the potential to detect photosensitisers (allergens) have been developed and are in regular use

3.2. Recommendations

- consider a review publication focused on the basic principles and core requirements associated with the first two bullets
- consider an expert review to extend the human potency dataset and to gain its broad acceptance, including via peer reviewed publication
- develop a strategy for gaining acceptance of variants of the LLNA or other methods incorporating an assessment of markers which discriminate true from false positives
- continue to publish the work on false positive results in the LLNA; where irritation may be and where irritation may not be the root cause
- consider the value of a review paper to promote the strengths and limitations of photo-LLNA protocol(s), including provision of a detailed overview of their use and interpretation

4. Experience with the development and use of alternatives to the LLNA

Although the LLNA offers two valuable elements of the 3Rs, refinement and reduction, as a murine *in vivo* method it fails to address the complete replacement of animal usage, a desirable goal which has now become a requirement in EU cosmetic legislation (EU, 2003). There have been a number of recent reviews of the status of the development of non-animal methods (Aeby et al., 2010; Vandebriel and van Loveren, 2010; Kimber et al., 2011; Peiser et al., 2012). The second day of the workshop focused on examples of *in silico* and *in vitro* approaches to the evaluation of skin sensitisation, including their status in the validation process, the key points from which are summarised below.

Considering *in silico* approaches first, an important focus on the workshop was on the utility of the OECD toolbox (OECD, 2012). This freely available system encompasses a considerable number of individual tools and is updated and/or supplemented regularly. However, this does tend to result in different outcomes being obtained by different users or by the same user addressing the same question but on different dates. All of this implies potential for uncertainty in practical use, at least for skin sensitisation, coupled with a need for expert interpretation. In terms of the prediction of the sensitising activity of 249 pharmaceutical chemicals, the overall prediction accuracy when using OECD Toolbox Version 1.1 was

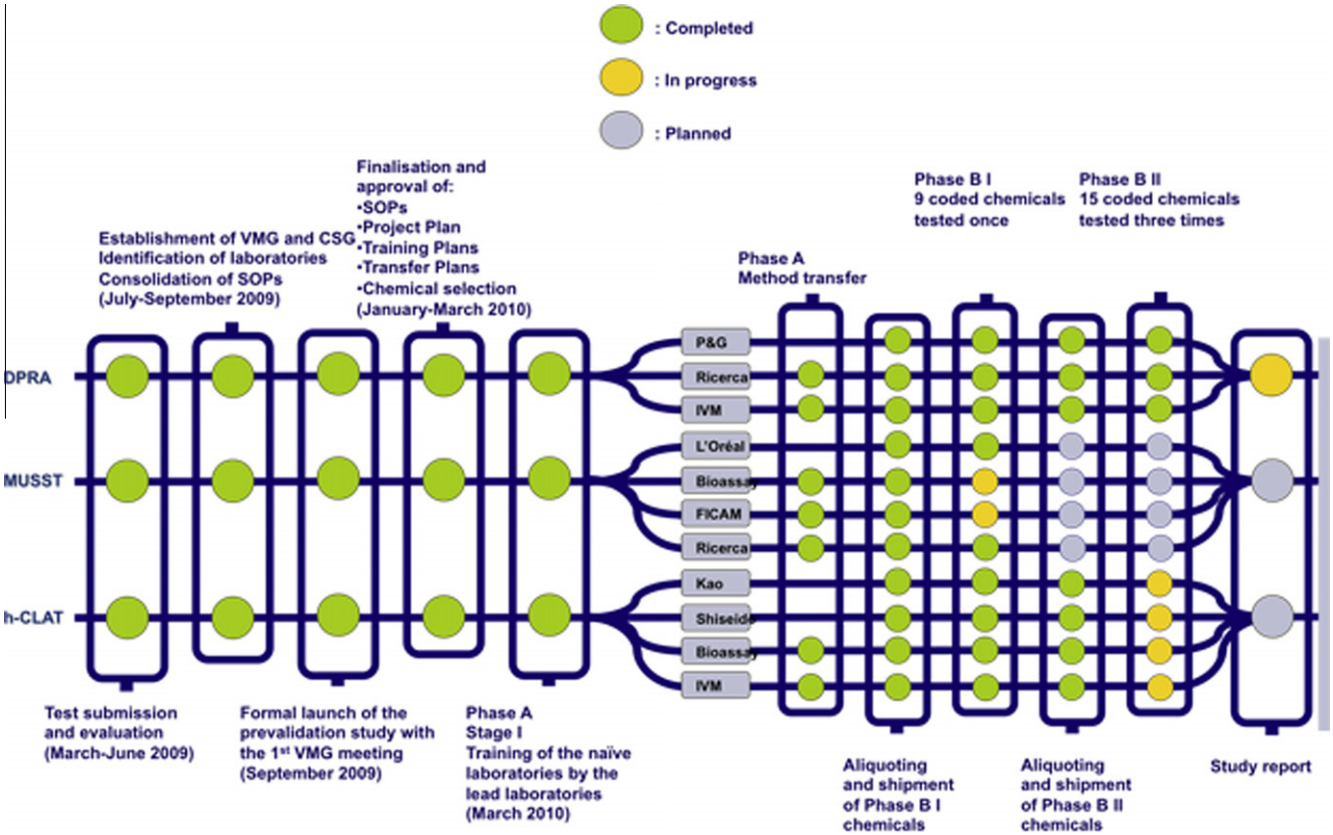

Fig. 1. Status of ECVAM skin sensitisation validation study at December 2011.

Table 1
Examples of experience with LLNA.

Chemical classes/	Sector	Issues	Reference
Surfactants	Home care, cosmetic ingredients, agroformulations	Large number of false positives; irritation as potential confounder	Ball et al., 2011
Fatty acid like substances	Cosmetic ingredients	Large number of false positives	Kreiling et al., 2008
Siloxanes	Cosmetic ingredients	Large number of false positives	Basketter et al., 2009a,b; Eigler et al., in preparation
Rosins + Resins	HARRPA	Clarification of true positives and read across	Botham et al., 2008; Illing et al., 2009
Photosensitizers	Pharma	Discriminating between photoirritants and photosensitizers	Ulrich et al., 1998; Neumann et al., 2005
Surfactants, performance standards	Home care, cosmetic ingredients, agroformulations	Immunophenotyping does not increase predictivity of LLNA	Ball et al., 2011; Kolle et al., in press

Table 2
Examples of experience with *in vitro/in silico* approaches.

System tested	Basis of the commentary	Focus of the work	Commentary[a]	Reference/ Presenter
OECD toolbox	Experiences made with (Q)SAR methods	Predictive capacity	Insufficient accuracy with unstructured use; offers valuable insights when used on well characterised chemical parameter space or for WoE evaluations	Seaman/ Keller/ Mehling
Keratinosens + LC-MS peptide reactivity	KeratinoSens: a cell based method for detection of reactive chemistry; LC-MS peptide reativity: optimised *in chemico* method, in which protein adducts are also identified	Sensitising impurities	KeratinoSens: promising method, which seems likely to be a key contributor to hazard identification, but misses specific lysine reactive chemistry. LC-MS peptide reactivity: in contrast to the DPRA, has the benefit of identifying adducts. Study here: novel use to identify sensitising impurities	Natsch
In vitro test battery	A parallel assessment of promising *in vitro* methods	Predictive capacity, alone or in combination	A combination of assays predict hazard, but not potency; neither the assays nor the prediction model are yet validated, but several are close	Ramirez Hernandez/ Kolle
Sens-it-iv	A range of newly developed cell based *in vitro* methods	Predictive capacity, alone or in combination	Although promising, assays are at an early stage in the validation process. One method distinguishes skin and respiratory sensitisers; another provides limited potency information	Roggen
COLIPA	Cosmetic industry view	Predictive capacity, alone or in combination	Current methods do not offer the prospect of prediction of sensitising potency, hence their ongoing substantial research programme	Taalman
CESIO	Surfactant industry view	Predictive capacity, alone or in combination	Currently, *in vitro* assays can contibute to a weight of evidence analysis and may help to identify false positives from *in vivo* assays	Ball

[a] This excludes issues generally common to all methods, including the difficulty of presenting very hydrophobic substances to aqueous systems, the absence of a "gold standard" database against which comparisons can reliably be made, but on a positive note there is the option to use multiple/repeated experiments in investigative mode.

only 53%, which was slightly worse than that of DEREK, v10, where the accuracy was 63%. By adjusting the rules used, selective evaluations based on specific structural alerts, adding data from additional tests and above all adding new good quality data to the data bases used, can improve predicitivities.

Whilst it is widely accepted that chemical structure activity relationships will continue to play an important role in reducing *in vivo* testing for skin sensitisation, it is also accepted that more biologically based test systems will also be necessary. Currently, ECVAM has underway the validation testing and review of four methods, Keratinosens (Natsch et al., 2009), the direct peptide reactivity assay (DPRA) (Gerberick et al., 2004, 2007) the human cell line activation test (h-CLAT) (Sakaguchi et al., 2009) and the myeloid U937 skin sensitisation test (MUSST) (Aeby et al., 2010; Maxwell et al., 2011). The first of these methods represents a full independent submission, whereas the remaining three are the subject of a blinded interlaboratory ring trial to demonstrate inter and intra laboratory reproducibility (Fig. 1). At the time of writing, the DPRA experimental work had been completed and the report was in preparation; h-CLAT and MUSST experimental work was expected to be finished within the following six months table 1 and 2.

In addition to the ECVAM work, a substantive body of work progressing towards *in vitro* assays for skin sensitisation has been completed within the Sens-it-iv 6th Framework programme (http://www.sens-it-iv.eu). Several potential methods have been submitted to ECVAM for consideration (Erwin Roggen, personal communication). Moreover, the combinations of existing test methods have been proposed to increase the predictivity of the *in vitro* systems (Bauch et al., 2011 and Bauch et al., 2012 under revision). Others are undergoing continuing development, with consideration being given not only to hazard identification, but also to potency assessment (e.g. McKim et al., 2010). It is to this last aspect that the final part of the workshop was addressed.

A great deal of attention in the development of *in vitro* alternatives is devoted to the third "R", the replacement of the *in vivo* methods that are represented in the framework of regulatory toxicology and specified in guidelines, notably those of the OECD. A number of industry sectors combined efforts to illustrate how even prior to validation, current *in vitro* assays were more likely to deliver correct classifications than the LLNA (Ball et al., 2011). Presentations were given on strategies to combine existing tests to improve predictivities, the chemical sector proposed a testing strategy using three different tests that even surpassed the

accuracy of the LLNA for the over 50 tested chemicals (Bauch et al., 2012, under revision). However, for many industry sectors, including those involved in the surfactant study, this by no means encapsulates the full importance of the assays detailed in guidelines for chemical testing. Assays are also used to understand the impact of formulation/product matrix on sensitisation. They are also used to develop an appreciation of the relative potency of an identified skin sensitiser, seen most often as the LLNA EC3 value mentioned earlier in this article. For the cosmetic industry, the research strategies developed? by COLIPA were presented (Aeby et al., 2010; Maxwell et al., 2011). These are directed towards potency characterisation without animal testing and address many of the knowledge gaps which currently limit our ability to address this question *in vitro*.

Thus, to a large extent the basic problem remains – how to use non-animal methods for the estimation of the potency of newly identified skin sensitising materials. Those at the workshop did not provide any instant solutions in this respect. Potential *in vitro* alternatives of various types were thoroughly discussed (see later section on this topic), but the primary point that was raised concerned the absence of a "gold standard" potency dataset. This is a particularly important point. Just as those involved in the validation of hazard identification alternatives have understood the importance of using all the available evidence for validation datasets (e.g. Casati et al., 2009), the same point is also true for the evaluation of methods to define sensitisation potency. However, the reality is that most of the groups working on this tend only to refer to the LLNA EC3 value, a valuable but clearly imperfect surrogate (Basketter et al., 2008a,b). What is needed is a substantial effort to collate and to extend initial efforts to combine all the experimental and clinical knowledge on the relative skin sensitising potency of a wide range of substances. This would provide an essential resource for further development of non-animal alternatives.

Whatever methods arise from the present and future research concerning the *in vitro* characterisation of sensitisation potency, it is certain that proving the utility of those methods (remember that potency does not require formal validation) will depend also on the bringing into existence a gold standard dataset of substances whose relative human potency is well characterised.

4.1. Conclusions

- SAR/QSAR systems remain in a developmental phase, with the consequence that their practical use remains in hands of specialists that recognise and understand the strengths and limitations
- *In vitro* alternatives for skin sensitisation hazard identification are already quite advanced, such that validated systems can be anticipated as replacements for the LLNA in the next few years
- Some of these non-animal tests are already finding use in screening assays, or for weight of evidence review
- Properly characterised *in vitro* alternatives which deliver information on the potency of a sensitiser are not yet well developed, although some initial approaches show promise.

4.2. Recommendations

- Continue to remind regulators that full replacement of the *in vivo* tests requires more than hazard identification
- Consider how to develop an inter-industry based strategy on how to use *in vitro* data to predict potency
- Investigate whether there is a need for research to identify what aspects of sensitisation chemistry contribute to potency
- consider an expert review to extend the potency dataset and to gain broad acceptance, including via peer reviewed publication

5. Risk assessment strategies/quantitative risk assessment (QRA)

Although the primary purpose of a skin sensitisation test, whether *in vivo, in vitro* or *in silico*, is to identify those chemicals which possess this intrinsic hazard, to protect human health it is necessary to undertake risk assessment and risk management. For these to be done to an adequate standard it is essential to develop an appreciation of the relative sensitising potency of a positive substance (Api et al., 2008; Basketter, 2008). Consequently, the workshop also considered concepts of risk assessment for skin sensitisation, bringing into the frame the key topic of how we currently identify relative skin sensitising potency and what strategies might be adopted when the *in vivo* methods, notably the LLNA, are no longer available. Importantly, this included the question of what represents the "gold standard" dataset for skin sensitising potency against which alternatives can be assessed. It was also noted that in contrast to hazard identification methods, risk assessment strategies do not require formal validation, although gaining widespread acceptance was judged to be appropriate.

Two primary themes were developed – how to identify a "safe" threshold for human exposure (which remains a generic question whether one uses *in vivo* or *in vitro* methods), and the vexed question of the quality of the currently available information on the relative skin sensitising potency of substances. Taking the first of these, the topic of the "threshold of sensitization concern (TSC) concept" was presented. This has formed the subject of a trio of recent publications which offer an increasingly refined view on defining an exposure level where there is a very high degree of confidence that no sensitisation risk will be present (Safford, 2008; Keller et al., 2009; Safford et al., 2011). However, there is no doubt that this approach is limited to situations where human exposure is actually low and other conservative assumptions can be applied. One of the assumptions relates to the chemistry of the substance, in particular either knowledge of its lack of reactivity with proteins, or, if reactive, its mechanism of action. This information potentially can be derived from peptide binding studies or by use of the OECD toolbox (OECD, 2011). The approach remains at an early stage of development and in need of more widespread testing and acceptance.

5.1. Conclusions

- Any approach to the definition of a safe exposure limit for skin sensitising chemicals benefits from wide ranging support. Thus repeated publication of examples and a transparent methodology are essential.

5.2. Recommendations

- Continue to publish worked examples of risk assessment, focusing on those where there is the support from clinical data showing the presence/absence of a human health problem.

6. Summary of main recommendations for potential follow-up

From the material presented and discussed at the Workshop, priority recommendations/actions have been identified:

1. The weight of evidence must always take precedence in classification and labelling conclusions. ACTION: define best practise examples, then arrange a training workshop with regulatory agencies.
2. Testing strategies should be optimised for balanced accuracy, avoiding a bias towards false negatives or false positives. ACTION: consider organising a workshop to explore integrated testing strategies for skin sensitisation hazard identification and characterization.
3. Validation of *in vitro* tests (and associated integrated testing strategies) requires a consistent gold standard database which ideally reflects humans, not rodents. This is of special importance for the prediction of relative sensitising potency. ACTION: collaborate with others on the agreement of a suitable dataset.
4. *In silico* approaches require a more extended period of research to enhance their more widespread utility, followed by a period of stable application. ACTION: PROPOSE increased funding, possibly for example via associations such as LRI and/or EPAA, in this area. Also, it is critical to add to datasets for e.g. OECD toolbox
5. The process of validation should be accelerated and integrated more fully with systems for regulatory acceptance. ACTION: propose improved schemes for pre-validation standards; involve regulators early on.
6. Experience with validated alternatives should be communicated to review their strengths and limitations. ACTION: organise opportunities to exchange data / information from various sectors on a regular basis.

Conflict of interest

DAB was paid by the EPAA for the preparation of this workshop report. The remaining authors are paid employees of their respective organisations.

Funding source

None other than that reflected in the conflict of interest statement.

Acknowledgments

The authors would like to thank the following individuals who were speakers at the workshop: Nicholas Ball, David Basketter, Juan-Carlos Carillo, Dorothea Eigler, Naveed Honarvar, Detlef Keller, Susanne Kolle, Reinhard Kreiling, Jon Lalko, Annette Mehling, Andreas Natsch, Albrecht Poth, Tzutzuy Ramirez, Erwin Roggen, Chris Seaman, Julia Scheel, Rob Taalman, Peter Ulrich, Hans-Werner Vohr, Adam Woolley.

References

Adler, S., Basketter, D.A., Creton, S., Pelkonen, O., van Benthem, J., Zuang, V., et al., 2011. Alternative (non-animal) methods for cosmetics testing: current status and future prospects—2010. Arch. Toxicol. 85, 367–485.

Aeby, P., Ashikaga, T., Bessou-Touya, S., Schepky, A., Gerberick, F., Kern, P., Marrec-Fairley, M., Ovigne, J.-M., Sakaguchi, H., Reisinger, K., Tailhardat, M., Martinozzi-Teissier, S., Winkler, P., 2010. Identifying and characterizing chemical skin sensitizers without animal testing: Colipa's research and method development program. Toxicol. In Vitro 24, 1465–1473.

Andersen, K.E., Maibach, H I., 1985. Guinea pig sensitization assays: An overview. In: Contact Allergy Predictive Tests in Guinea Pigs, Current Problems in Dermatology. Eds: K.E. Andersen and H.I. Maibach, Vol. 14, pp 59–106. New York: Karger.

Angers-Loustau, A., Tosti, L., Casati, S., 2011. The regulatory use of the local lymph node assay for the notification of new chemicals in Europe. Regul. Toxicol. Pharmacol. 60, 300–307.

Api, A.M., Basketter, D.A., Cadby, P.A., Cano, M.-F., Ellis, G., Gerberick, G.F., Griem, P., McNamee, P.M., Ryan, C.A., Safford, B., 2008. Dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. Regul. Toxicol. Pharmacol. 52, 3–23.

Ball, N., Cagen, S., Carrillo, J.C., Certa, H., Eigler, D., Emter, R., Faulhammer, F., Garcia, C., Graham, C., Haux, C., Kolle, S.N., Kreiling, R., Natsch, A., Mehling, A., 2011. Evaluating the sensitization potential of surfactants: integrating data from the local lymph node assay, guinea pig maximization test, and in vitro methods in a weight-of-evidence approach. Regul. Toxicol. Pharmacol. 60, 389–400.

Balls, M., Hellsten, E., 2000. Statement on the validity of the local lymph node assay for skin sensitization testing. ECVAM Joint Research Centre, European Commission. Ispra. Altern. Lab. Animals 28, 366–367.

Basketter, D.A., 2008. Skin sensitisation: strategies for risk assessment and risk management. Brit. J. Dermatol. 159, 267–273.

Basketter, D.A., Chamberlain, M., 1995. The validation of skin sensitisation assays. Food Chem. Toxicol. 33, 1057–1059.

Basketter, D.A., Kimber, I., 2007. Information derived from sensitisation test methods: test sensitivity, false positives and false negatives. Contact Dermatitis 56, 1–4.

Basketter, D.A., Gerberick, G.F., Kimber, I., 1998. Strategies for identifying false positive responses in predictive skin sensitization tests. Food Chem. Toxicol. 36, 327–333.

Basketter, D.A., McFadden, J., Evans, P., Andersen, K.E., Jowsey, I., 2006. Identification and classification of skin sensitisers: identifying false positives and false negatives. Contact Dermatitis 55, 268–273.

Basketter, D.A., Cockshott, A., Corsini, E., Gerberick, G.F., Idehara, K., Kimber, I., van Loveren, H., Takeyoshi, M., Matheson, J., Mehling, A., Omori, T., Rovida, C., Sozy, T., Stokes, W., Casati, S., 2008a. An evaluation of performance standards and non-radioactive endpoints for the local lymph node assay. ATLA 36, 243–257.

Basketter, D.A., Darlenski, R., Fluhr, J., 2008b. Skin irritation and sensitization: mechanisms and new approaches for risk assessment. Part II: skin sensitization. Skin Pharmacol. Physiol. 21, 191–202.

Basketter, D.A., Ball, N., Cagen, S., Carrillo, J.-C., Certa, H., Eigler, D., Esch, H., Graham, C., Haux, D., Kreiling, R., Mehling, A., 2009a. Application of a weight of evidence approach to analysing discordant sensitization datasets: implication for REACH. Regul. Toxicol. Pharmacol. 55, 90–96.

Basketter, D.A., McFadden, J.F., Gerberick, G.F., Cockshott, A., Kimber, I., 2009b. Nothing is perfect, not even the local lymph node assay. A commentary and the implications for REACH. Contact Dermatitis 60, 65–69.

Basketter, D.A., Kolle, S.N., Schrage, A., Honarvar, N., Gamer, A.O., van Ravenzwaay, B., Landseidel, R., in press. Experience with local lymph node assay performance standards using standard radioactivity and non-radioactive cell count measurements. J. Appl. Toxicol.

Basketter, D.A., Clewell, H., Kimber, I., Rossi, A., Blaauboer, B., Burrier, R., Daneshian, M., et al., 2012. A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing. ALTEX 29, 5–91.

Bauch, C., Kolle, S.N., Fabian, E., Pachel, C., Ramirez, T., Wiench, B., Wruck, C.J., van Ravenzwaay, B., Landsiedel, R., 2011. Intralaboratory validation of four in vitro assays for the prediction of the skin sensitizing potential of chemicals. Toxicol. In Vitro 25, 1162–1168.

Bauch, C., Kolle, S.N., Ramirez, T., Eltze, T., Fabian, E., Mehling, A., Teubner, W., van Ravenzwaay, B., Landsiedel, R., 2012. Putting the parts together: Combining in vitro methods to test for skin sensitizing potentials (under revision).

Botham, P.A., Lees, D., Illing, H.P., Malmfors, T., 2008. On the skin sensitisation potential of rosin and oxidised rosin. Regul. Toxicol. Pharmacol. 52, 257–263.

Buehler, E.V., 1965. Delayed contact hypersensitivity in the guinea pig. Arch. Dermatol. 91, 171–177.

Casati, S., Aeby, P., Kimber, I., Maxwell, G., Ovigne, J.M., Roggen, E., Rovida, C., Tosti, L., Basketter, D.A., 2009. Selection of chemicals for the development and evaluation of in vitro methods for skin sensitisation testing. Altern. Lab. Animals 37, 305–312.

Cockshott, A., Evans, P., Ryan, C.A., Gerberick, G.F., Betts, C.J., Dearman, R.J., Kimber, I., Basketter, D.A., 2006. The local lymph node assay in practice: a current regulatory perspective. Hum. Exp. Toxicol. 25, 387–394.

Commission of the European Communities. Regulation (EC)., 2006. No. 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC,93/105/EC and 2000/21/EC. Off J Eur Union, L 396/1 of 30.12.2006.

Commission Regulation (EU) No 286/2011 of 10 March 2011 amending, for the purposes of its adaptation to technical and scientific progress, Regulation (EC) No 1272/2008 of the European Parliament and of the Council on classification, labelling and packaging of substances and mixtures. Off J Eur Union, L 83, 54.

Dean, J.H., Twerdok, L.E., Tice, R.R., Sailstad, D.M., Hattan, D.G., Stokes, W.S., 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.

Ehling, G., Hecht, M., Heusener, A., Huesler, J., Gamer, A.O., van Loveren, H., Maurer, T., Riecke, K., Ullmann, L., Ulrich, P., Vandebriel, R., Vohr, H.-W., 2005. An European inter-laboratory validation of alternative endpoints of the murine local lymph node assay: first round. Toxicology 212, 60–68.

EU, 2003. Directive 2003/15/EC of the European Parliament and of the Council of 27 February 2003 amending Council Directive 76/768/EEC on the approximation of

the laws of the Member States relating to cosmetic products (“7th Amendment to the European Cosmetics Directive”). Official Journal of the European Union L 66, 26-35.

Gamer, A.O., Nies, E., Vohr, H.-W., 2008. Local lymph node assay (LLNA): comparison of different protocols by testing skin-sensitizing epoxy resin system components. Regul. Toxicol. Pharmacol. 52, 290–298.

Garcia, C., Ball, N., Cagen, S., Carrillo, J.C., Certa, H., Eigler, D., Esch, H., Graham, C., Haux, C., Kreiling, R., Mehling, A., 2010. Comparative testing for the identification of skin-sensitizing potentials of nonionic sugar lipid surfactants. Regul. Toxicol. Pharmacol. 58, 301–307.

Gerberick, G.F., Ryan, C.A., Kimber, I., Dearman, R.J., Lea, L.J., Basketter, D.A., 2000. Local lymph node assay: validation assessment for regulatory purposes. Am. J. Contact Derm. 11, 3–18.

Gerberick, G.F., Vassallo, J.D., Bailey, R.E., Chaney, J.G., Morrall, S.W., Lepoittevin, J.-P., 2004. Development of a peptide reactivity assay for screening contact allergens. Toxicol. Sci. 81, 332–343.

Gerberick, G.F., Vassallo, J.D., Foertsch, L.M., Price, B.B., Chaney, J.G., Lepoittevin, J.-P., 2007. Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. Toxicol. Sci. 97, 417–427.

Gerberick, G.F., Aleksic, M., Basketter, D.A., Casati, S., Karlberg, A.-T., Kern, P., Kimber, I., Lepoittevin, J.-P., Natsch, A., Ovigne, J.-M., Rovida, C., Sakaguchi, H., Schultz, T., 2008. Chemical reactivity measurement and the predictive identification of skin sensitisers. ATLA 36, 215–242.

Illing, H.P., Malmfors, T., Rodenburg, L., 2009. Skin sensitization and possible groupings for ‘read across’ for rosin based substances. Regul. Toxicol. Pharmacol. 54, 234–241.

Keller, D., Krauledat, M., Scheel, J., 2009. Feasibility study to support a threshold of sensitization concern concept in risk assessment based on human data. Arch. Toxicol. 83, 1049–1060.

Kimber, I., Basketter, D.A., 1992. The murine local lymph node assay: a commentary on collaborative trials and new directions. Food Chem. Toxicol. 30, 165–169.

Kimber, I., Basketter, D.A., Gerberick, G.F., Ryan, C.A., Dearman, R.J., 2011. Chemical allergy: translating biology into hazard characterisation. Toxicol. Sci. 120, 238–268.

Kirk, M.K., Broomhead Y, DiDonato L., DeGeorge, G.L., 2007. Use of an Enhanced Local Lymph Node Assay to Correctly Classify Irritants and False Positive Substances. <http://www.mbresearch.com/posters/2007/ELLNA_SOT_2007.pdf>.

Kolle, S., Basketter, D., Schrage, A., Gamer, A.O., van Ravenzwaay, B., Landseidel, R., in press. Further experience with the local lymph node assay using stand radioactive and non-radioactive cell count measurements. J. Appl. Toxicol.

Kreiling, R., Hollnagel, H.M., Hareng, L., Eigler, D., Lee, M.S., Griem, P., Dreesen, B., Kleber, M., Albrecht, A., Garcia, C., Wendel, A., 2008. Comparison of the skin sensitizing potential of unsaturated compounds and assessed by the murine local lymph node assay (LLNA) and the guinea pig maximization test (GPMT). Food Chem. Toxicol. 46, 1896–1904.

Magnusson, B., Kligman, A.M., 1970. Allergic Contact Dermatitis in the Guinea Pig: Identification of Contact Allergens, 141pp. Springfield, IL, USA: Charles C. Thomas.

Maxwell, G., Aeby, P., Ashikaga, T., Bessou-Touya, S., Diembeck, W., Gerberick, F., Kern, P., Marrec-Fairley, M., Ovigne, J.-M., Sakaguchi, H., Schroeder, K., Tailhardat, M., Teissier, S., Winkler, P., 2011. Skin sensitisation: the Colipa strategy for developing and evaluating non-animal test methods for risk assessment. ALTEX 28, 50–55.

McGarry, H.F., 2007. The murine local lymph node assay: Regulatory and potency consideration under REACH. Toxicology 238, 71–89.

McKim, J.M., Keller, D.J., Gorski, J.R., 2010. A new in vitro method for identifying chemical sensitizers combining peptide binding with ARE/EpRE-mediated gene expression in human skin cells. Cut. Ocul. Toxicol. 29, 171–192.

Neumann, N.J., Blotz, A., Wasinska-Kempka, G., Rosenbruch, M., Lehmann, P., Ahr, H.J., Vohr, H.W., 2005. Evaluation of phototoxic and photoallergic potentials of 13 compounds by different in vitro and in vivo methods. J. Photochem. Photobiol., B 79, 25–34.

OECD, 2002. Local Lymph Node Assay. Test Guideline No 429, Organisation for Economic Cooperation and Development, Paris.

OECD, 2010a. Organisation for Economic Cooperation and Development. Test Guideline 429: The Local Lymph Node Assay. Paris, France.

OECD, 2010b. Organisation for Economic Cooperation and Development. Guidelines for Test of Chemicals No 442a and 442b. Paris, France.

OECD, 2012. The OECD QSAR Toolbox, Version 2.2 July 18, 2011. Paris. <http://www.oecd.org/document/54/0,3746,en_2649_34379_42923638_1_1_1_1,00.html>. Last accessed, 27/01/2012.

Omori, T., Idehara, K., Kojima, H., Sozu, T., Arima, K., Goto, H., Hanada, T., Ikarashi, Y., Inoda, T., Kanazawa, Y., Kosaka, T., Maki, E., Morimoto, T., Shinoda, S., Shinoda, N., Takeyoshi, M., Tanaka, M., Uratani, M., Usami, M., Yamanaka, A., Yoneda. T., Yoshimura, I., Yuasa, A., 2008. Interlaboratory validation of the modified murine local lymph node assay based on adenosine triphosphate measurement. J. Pharmacol. Toxicol. Methods 58, 11–26. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18593646>.

Peiser, M., Tralau, T., Heidler, J., Api, A.M., Arts, J.H.E., Basketter, D.A., English, J., Diepgen, T.L., Fuhlbrigge, R.C., Gaspari, A.A., Johansen, J.D., Karlberg, A.T., Kimber, I., Lepoittevin, J.-P., Liebsch, M., Maibach, H.I., Martin, S.F., Merk, H.F., Platzek, T., Rustemeyer, T., Schnuch, A., Vandebriel, R.J., White, I.R., Luch, A., 2012. Allergic contact dermatitis: epidemiology, molecular mechanisms, in vitro methods & regulatory aspects. Cell. Mol. Life Sci. 69, 763–781.

Safford, R.J., 2008. The dermal sensitisation threshold - a TTC approach for allergic contact dermatitis. Regul. Toxicol. Pharmacol. 51, 195–200.

Safford, R.J., Aptula, A.O., Gilmour, N., 2011. Refinement of the dermal sensitisation threshold (DST) approach using a larger dataset and incorporating mechanistic chemistry domains. Regul. Toxicol. Pharmacol. 60, 218–224.

Sakaguchi, H., Ashikaga, T., Miyazawa, M., Kosaka, N., Ito, Y., Yoneyama, K., Sono, S., Itagaki, H., Toyoda, H., Suzuki, H., 2009. The relationship between CD86/CD54 expression and THP-1 cell viability in an in vitro skin sensitization test- human cell line activation test (h-CLAT). Cell Biol. Toxicol. 25, 109–126.

Scholes, E.W., Basketter, D.A., Lovell, W.W., Sarll, A.E., Pendlington, R.U., 1992. The identification of photoallergic potential in the local lymph node assay. Photodermatol. Photoimmunol. Photomed. 8, 249–254.

Ulrich, P., Homey, B., Vohr, H.W., 1998. A modified murine local lymph node assay for the differentiation of contact photoallergy from phototoxicity by analysis of cytokine expression in skin-draining lymph node cells. Toxicology 125, 149–168.

Ulrich, P., Grenet, O., Bluemel, J., Vohr, H.W., Wiemann, C., Grundler, O., Suter, W., 2001a. Cytokine expression profiles during murine contact allergy: T helper 2 cytokines are expressed irrespective of the type of contact allergen. Arch Toxicol. 75: 470–479. Erratum in. Arch. Toxicol. 2002 (76), 62.

Ulrich, P., Streich, J., Suter, W., 2001b. Intralaboratory validation of alternative endpoints in the murine local lymph node assay for the identification of contact allergic potential: primary ear skin irritation and ear-draining lymph node hyperplasia induced by topical chemicals. Arch. Toxicol. 74, 733–744.

Vandebriel, R.J., van Loveren, H., 2010. Non-animal sensitization testing: state of the art. Crit. Rev. Toxicol. 40, 389–404.

Vohr, H.W., Blümel, J., Blotz, A., Homey, B., Ahr, H.J., 2000. An intra-laboratory validation of the Integrated Model for the Differentiation of Skin Reactions (IMDS): discrimination between (photo)allergic and (photo)irritant skin reactions in mice. Arch. Toxicol. 73 (10–11), 501–509.

Vohr, H.-V., Ahr, A.H.J., 2005. The local lymph node assay being too sensitive? Arch. Toxicol. 79, 721–728.

Vohr, H.-W., Homey, B., Schuppe, H., Kind, P., 1994. Photoreactions detected in a modified local lymph node assay in the mouse. Photoderm. Photoimmunol. Photomed. 10, 57–64.