



Application of a weight of evidence approach to assessing discordant sensitisation datasets: Implications for REACH

David Basketter^{a,*}, Nicholas Ball^b, Stuart Cagen^c, Juan-Carlos Carrillo^d, Hans Certa^e, Dorothea Eigler^f, Christine Garcia^g, Harald Esch^h, Cynthia Grahamⁱ, Carl Haux^j, Reinhard Kreiling^k, Annette Mehling^l

^a DABMEB Consultancy Ltd, Sharnbrook, Bedford MK44 1PR, UK

^b Dow Europe GmbH, Horgen, Switzerland

^c Shell Health, Houston, TX, USA

^d Shell International BV, The Hague, The Netherlands

^e SASOL Germany GmbH, Marl, Germany

^f Evonik Goldschmidt GmbH, Essen, Germany

^g SEPPIC, Paris, France

^h BASF SE, Ludwigshafen, Germany

ⁱ Huntsman Corporation, The Woodlands, TX, USA

^j Akzo Nobel Surface Chemistry AB, Stenungsund, Sweden

^k Clariant Produkte (Deutschland) GmbH, Sulzbach, Germany

^l Cognis GmbH, Düsseldorf, Germany

ARTICLE INFO

Article history:

Received 24 April 2009

Available online 10 June 2009

Keywords:

Skin sensitisation

LLNA

False positive

False negative

REACH

ABSTRACT

The local lymph node assay (LLNA) is the assay of choice in European regulatory toxicology. As with other toxicology/sensitisation assays, it has a potential for false results, the anionic surfactant sodium lauryl sulphate (SLS) representing a classic example. In the work reported here, examples of false positives in the LLNA are compared to published “benchmarks” such as SLS. Clear false positives (e.g. oleic acid) are also contrasted with examples where data interpretation is more challenging. As the LLNA will be applicable to >30,000 chemicals under REACH, and in the light of animal welfare considerations to do no more than the absolute minimum of animal testing, results from a single LLNA often represent the only available data on sensitisation. This reinforces the need to ensure data from this assay are interpreted intelligently, using scientific analysis of results and considering the weight of evidence, before decisions are made on which substances should be classified as representing a skin sensitisation hazard. In chemical classes where the LLNA has been shown to be an inappropriate assay other standardised methods (e.g. the Buehler or Magnusson and Kligman guinea pig tests [OECD 406]) should be employed as the first choice assays.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

The local lymph node assay (LLNA) was developed as an alternative to the long standing guinea pig assays (OECD 406; Magnusson and Kligman or Buehler assays) used for the detection of potential skin sensitising chemical hazards (Kimber and Basketter, 1992; Kimber et al., 1994, 2002; Basketter et al., 2007a,b,c). Following its formal validation in the EU (Balls and Hellsten, 2000) and the USA (Dean et al., 2001), it was adopted as a complete alternative test method within the European Union (EU, 2004) and via its acceptance as Organisation for Economic Cooperation and Development (OECD) Test Guideline 429 (OECD, 2002). Furthermore, due to its technical and animal welfare advantages over the guinea

pig tests, the LLNA has been promoted as the method of first choice, initially by the United Kingdom (UK HSE, 2002) and subsequently within new legislation, Registration, Evaluation, Authorisation and restriction of Chemicals (REACH), in the EU (Commission of the European Communities, 2006). This relatively high profile of the LLNA has been supported also by the reporting of good experience with the assay by regulatory groups (Cockshott et al., 2006; McGarry, 2007; Oberemm et al., 2008).

As might be expected when a new toxicology assay finds increasing use post-validation and acceptance, the experience gained with the assay expands exponentially. This experience, largely deriving from testing a much wider range of industrial chemical classes, has probed areas of applicability not fully evaluated during the formal validation of the LLNA, and serves to highlight both strengths and potential areas of weakness. For example, the pesticide industries are required to assess aqueous formulations,

* Corresponding author. Fax: +44 01234782944.

E-mail address: david.basketter@ukonline.co.uk (D. Basketter).

an area which received little examination during validation. Their data suggests that, as indicated in OECD 429, wholly aqueous systems offer technical challenges (e.g. poor retention on skin) to the LLNA which demand some effort to overcome (Ryan et al., 2002; Woolhiser et al., 2007). In addition, as expressed in OECD 429 the LLNA may not be suitable to test irritants. Furthermore, although approximately 200 chemicals were tested during the original validation, some categories of substances were not well represented, e.g. surfactants, and others were not represented at all, e.g. fatty acids. A number of these issues have been the subject of a recent review by the Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM) in the USA (Interagency Coordinating Committee for the Validation of Alternative Methods, 2008).

However, a major issue of debate for the LLNA (indeed for most toxicology assays) is the degree to which the assay is associated with false positive and false negative outcomes. More than 10 years ago, a description of the topic, with examples of both the responses to a range of irritants in the LLNA and exposition of a strategy for the identification of potential false positives was published (Basketter et al., 1998). This has been followed more recently by several publications dealing with the topic from a largely theoretical viewpoint (Basketter et al., 2006; Basketter and Kimber, 2007) and/or proposing the use of benchmark data as a guide to correct interpretation of problematic data (Basketter, 2008). Others have been able to produce limited datasets to exemplify the issue, giving examples of chemicals which might be regarded as false positive in the LLNA (Vohr and Ahr, 2005). Very recently, a more collated set of data discussing this point with a focus on fatty acids as well as surfactants and related chemicals have been reported (Kreiling et al., 2008; Mehling et al., 2008).

The underlying problem is that all of the material which is related to the question of false positives and false negatives requires that the whole of the scientific evidence is carefully evaluated and a weight of evidence decision is made on the presence or absence of skin sensitisation hazard, which often will require an element of expert judgement (Basketter, 2008). To date, this has not been well accepted by regulatory agencies, but within the REACH legislation as well as in the Global Harmonisation Scheme (GHS), there is a requirement that classification decisions are taken on a weight of evidence (Globally Harmonised System of Classification and Labelling of Chemicals, 2007). Accordingly, in this present work, datasets on several chemicals are reviewed in the context of regulatory hazard classification to exemplify how such decisions might be made, but with emphasis on the identification of false positives. Critical in this analysis is that such classification decisions aim to distinguish significantly sensitising chemicals from those which are either non-sensitising or which lack significant sensitising potential (Commission of the European Communities, 2006; Basketter, 2008; Globally Harmonised System of Classification and Labelling of Chemicals, 2007). Whether a substance is to be classified as a sensitizer is largely determined by a defined minimum number of animals showing positive responses in comparison to vehicle-treated control animals in methods accepted by regulators (e.g. OECD 406; OECD 429). The outcome of the use of weight of evidence shows that, as with all toxicology assays, false positive and false negative results can and do occur with the LLNA (Basketter et al., 2009).

2. Materials and methods

2.1. Guinea pig tests

These procedures were carried out according to OECD Test Guideline 406 (Organisation for Economic Cooperation and Development, 1992).

Individual test details are contained in the appropriate table for each chemical and include commentary on any deviation from the standard protocol. Historical details can be accessed for the Buehler test (Buehler, 1965) and for the guinea pig maximisation test (GPMT) (Magnusson and Kligman, 1970). Normally, for a substance to be judged positive in the Buehler test at least 15% of tested guinea pigs must have positive skin reactions after challenge with the substance. For the GPMT, at least 30% of the guinea pigs are required to have positive skin reactions after challenge before a substance is considered to be positive.

2.2. Local lymph node assay

This method was carried out according to OECD Test Guideline 429 (Organisation for Economic Cooperation and Development, 2002). Individual test details are contained in the appropriate table for each chemical and include commentary on any deviation from the standard protocol. Normally a substance is considered positive if there is a 3-fold or greater increase in tritiated thymidine incorporation into the auricular lymph nodes of test animals compared to those of the concurrent vehicle-treated control group in the induction phase. Where data is sufficient, the dose response has been used to determine the threshold positive concentration, the EC3 value, using the approach published (Basketter et al., 1999).

2.3. Substances tested

These were all commercial grade raw materials (i.e. the grade of chemical upon which testing is required to be carried out to meet EU regulations). For commercial reasons, the exact chemical structures are not disclosed, but they comprise glyceryl thioglycolate, oleic acid and an oleic acid ester, two fatty alcohols and a polyaminofunctional siloxane. These materials were supplied by the authors' companies.

2.4. Data analysis

This was carried out using the weight of evidence approaches described in various regulations (Commission of the European Communities, 2006; Globally Harmonised System of Classification and Labelling of Chemicals, 2007), which suggest that data from all available sources should be integrated (ECHA, 2008, pp. 256–288) as well as in the light of the advice published by expert groups (Cockshott et al., 2006; McGarry, 2007; Basketter et al., 1998, 2006; Basketter and Kimber 2007; Basketter, 2008; Kligman and Basketter, 1995; Frankild et al., 1996; Schleder and Eppler, 1995). The final conclusion based on this type of analysis lead to a decision on whether the substance should be classified as a skin sensitizer and would thus merit the risk phrase R43: May cause sensitisation by skin contact (EC, 2001). It should be noted however that the intent of this analysis is not to promote specific classification decisions on these substances, but rather to exemplify the critical analysis which is generally necessary to evaluate skin sensitisation potential.

3. Results and discussion

Each of the substances has been considered in turn and where possible, evidence of the quality of the test procedure noted, paying particular attention to any data from positive controls. For example, the sensitivity of the LLNA can be demonstrated using a positive control substance, e.g. hexyl cinnamic aldehyde, and deriving an EC3 value to show the assay has been conducted with an appropriate degree of sensitivity (Basketter et al., 2007a,b,c). The LLNA assesses the induction/sensitisation phase of allergic

responses only via incorporation of radiolabelled thymidine into the regional lymph node. The guinea pig tests assess both the induction/sensitisation phase and the elicitation phase of allergic responses via visual assessment of skin reactions following challenge with the test substance. The protocol according to Magnusson Kligman also incorporates an intradermal induction step in order to overcome the skin barrier. This step also entails use of a powerful adjuvant to create conditions conducive to sensitisation. It should be noted that newer test methods designed to differentiate between irritant and sensitization responses, e.g. B220 ratios (Betts et al., 2007), were not applied to these examples as the tests were performed prior to the development of these tests. Furthermore, they are still in the development phase and have other constraints, e.g. evaluation of the B220 ratio can usually only be performed if non-radioactive methods are used in the LLNA, relevant threshold scores are still being evaluated and a sufficient number of chemicals has still not been tested. In the following examples, discordant results in the two tests are discussed. In making a judgment on the weight of evidence, while it is accepted that the biological data should take precedence, emphasis is placed also on the chemistry.

3.1. Sodium lauryl sulphate

Before exploring the specific results in detail, it is worth remembering the response in the LLNA obtained with the anionic surfactant sodium lauryl sulphate (SLS) (Loveless et al., 1996). SLS is widely used in cosmetics and household products, with the highest concentration of 50% used in non-colouring shampoos (CIR, 2005). Due to its moderate/strong irritant properties, it is routinely used by dermatologists and others as a positive control in human patch testing for irritants. SLS is not a human skin sensitizer. The basis for this conclusion came from experience from its widespread use in consumer products and negative guinea pig and human data (Magnusson and Kligman, 1970). Loveless et al (1996) tested SLS the LLNA at concentrations up to 20% in and showed a positive dose response which exceeded the 3-fold criterion for identification of a skin sensitizer (at 20%, the stimulation index was in the region of 4–8). Were this information submitted today for review by a regulatory authority it is certainly possible that SLS could be classified as a skin sensitizer solely on the basis of the positive LLNA results as this is the preferred method. However, SLS is widely regarded as a true false positive in the LLNA since it is well recognised as a non-sensitising, but is known as a classic and relatively strong skin irritant (reviewed in Basketter, 2008; CIR, 1983). A number of reports have provided evidence that irritancy augments the induction response and/or can sometimes lead to non-specific proliferation responses in the LLNA (Montelius et al., 1998; Woolhiser et al., 1998; Basketter et al., 2007a,b,c). Thus, SLS provides an excellent benchmark against which to judge other chemicals which are significantly irritating and give relatively low stimulation indices in the LLNA.

Table 1
Sensitisation test data for fatty acid glutamate.

Buehler test ^a concentrations and results		LLNA test ^b concentration and results	
Induction patch	50% in water	% in water	Stimulation index
Challenge patch	50% in water		
Challenge 1 result	13/20 (65%) positive	25	1.8
		50	1.2
		100	4.8

^a Mercaptobenzothiazole positive control gave a 100% positive response at 0.03%.

^b Hexyl cinnamic aldehyde positive control gave an EC3 value of approximately 8%.

3.2. Fatty acid glutamate

The detailed information on this substance is contained in Table 1. This substance was assessed in a Buehler test, using very high induction and challenge concentrations. At challenge, 65% of the guinea pigs were judged positive. Rechallenge results confirmed this outcome (data not shown). In the LLNA, the same substance gave a very weakly positive result only when tested undiluted and with no evidence of a dose response. The sensitivity of the LLNA was demonstrated by concurrent conduct of a positive control substance, hexyl cinnamic aldehyde which produced an EC3 value of about 8%, entirely consistent with the values indicated by other laboratories (Dearman et al., 2001). Although some evidence of skin irritation was noted in the top dose group of the LLNA, the generally low irritant potential of this type of substance renders it difficult in the face of the guinea pig evidence to regard the LLNA result as a false positive irritant. The LLNA results could be interpreted as being negative due to lack of a response other than at the highest dose at which irritation was also observed. The fact that the responses were seen at very high concentration and the absence of structural alerts (i.e. chemical substructural elements known to be associated with skin sensitisation—Langton et al., 2006) with this general type of chemical suggests that the positive responses are most likely due to a low level sensitising impurity. Analytical work and challenge with purified fatty acid glutamate in the guinea pig maximisation test revealed this to be the case. Due to animal welfare laws now in place, these experiments would be difficult to repeat.

3.3. Glyceryl thioglycolate

The detailed information on this substance is contained in Table 2. This substance was assessed in a GPMT at very high concentrations, but with no evidence of reactions in any of the 20 test animals (nor in the 10 controls). In contrast, in the LLNA glyceryl thioglycolate produced a very clear positive dose response, with even the lowest dose tested (10%) clearly exceeding the 3-fold criterion for positivity. In this case therefore, the results from the tests appear to be in conflict. A first question must be whether either of the tests might be regarded as invalid or at least inadequate for any reason. The GPMT does not have the benefit of any concurrent positive control data to assess its quality, neither does the LLNA, although careful examination of the test reports does not reveal any evidence of a failure to follow the proper protocol (data not shown). However, two other considerations can be applied. Firstly, glyceryl thioglycolate is a reactive chemical, chosen specifically for its ability to act as a reducing agent, particularly for application in hair styling and/or hair removal. Reactive chemistry in a substance of this molecular weight would be consistent with a positive skin sensitisation result. This is then supported by the second consideration, clinical evidence, where reports of both occupational and consumer sensitisation to glyceryl thioglycolate exist (Guerra et al., 1992; Storrs, 1984; Conde-Salazar et al., 1995). Whilst the clinical evidence might be insufficient by itself to trigger

Table 2
Sensitisation test data for glyceryl thioglycolate.

GPMT concentrations and results		LLNA test concentration and results	
Induction injection	2.5%	% in acetone/olive oil	Stimulation index
Induction patch	50%	(4:1, v/v)	
Challenge patch	100%		
Challenge 1 result	0/20 (0%)	10	8.0
		25	14.0
		50	31.1

classification of this substance as a skin sensitiser, it adds to the argument on reactive chemistry and fits well with the positive LLNA data. Thus, the weight of evidence considered here for glyceryl thioglycolate is that it would be classifiable as a skin sensitiser, meaning that the GPMT result is actually a false negative.

3.4. Oleic acid

The detailed information on this substance is contained in Table 3. Oleic acid was tested in the GPMT at very high concentration, with challenge at 1% demonstrating absolutely no evidence of skin sensitisation. A positive control GPMT using hexyl cinnamic aldehyde demonstrated that the procedure was of an acceptable degree of sensitivity (Table 2). In contrast, the LLNA with oleic acid gave a modest positive dose response at all three test concentrations. A concurrent positive control study with hexyl cinnamic aldehyde confirmed the validity of LLNA test conduct, giving a response within the normal range. So, as with glyceryl thioglycolate, the results from the two predictive tests are in conflict. The GPMT was carried out using an upper limit concentration of 5% for induction injection which can be argued to conflict with OECD Guideline requirements (OECD, 1992) but is nevertheless in line with common regulatory test practice (Schlede and Eppler, 1995). Also, it is a little surprising that whilst 100% oleic acid could be used for the induction patch, only 1% was used for the challenge patch. Nevertheless, although these represent a potential criticism of the assay, the concentrations were fully justified in the report, positive control data were acceptable and the result was unequivocally negative, i.e. 0%, not just some degree below the 30% classification limit. Interestingly, use of higher challenge concentrations in a separate GPMT reported on a different sample of oleic acid had led to confounded data (Kreiling et al., 2008). What of other evidence? Unlike glyceryl thioglycolate, oleic acid is not a protein reactive chemical (although still of a molecular weight which would allow skin penetration), which is fully consistent with the outcome of the GPMT. Oleic acid is an endogenous fatty acid present in many animal and vegetable products, for example olive oil. It has been used in cosmetics for decades with concentrations used ranging up to 20% (CIR 2006). Oleic acid was evaluated by the US agency Cosmetic Ingredient Review (CIR) and was not considered to be a sensitiser (CIR, 2006). Clinical data on irritant/allergic contact dermatitis from oleic acid is largely lacking, but it is of interest that oleic acid has been recommended as part of skin lipid mixture for clinical treatment of this type of skin condition, as well as for treatment of chronic atopic dermatitis (Berardesca et al., 2001). This argues particularly strongly for an absence of sensitising effects. As already indicated, oleic acid is not irritating to any great degree, so there is not really an option here to explain the result as an irritant false positive, as would be done for the more strongly

irritating anionic surfactant sodium lauryl sulphate, even though the level of proliferative response is somewhat similar (Basketter, 2008). Nevertheless, the weight of evidence here must be that the LLNA has delivered a false positive and that oleic acid should not be regarded as a skin sensitiser.

3.5. Oleic acid ester

The detailed information on this substance is contained in Table 4. The GPMT gave a 0% positive result following induction and challenge at very high concentrations. However, the interpretation was made complex by grade 1 reactions occurring in every control as well as every test animal at challenge 1 (data not shown). Rechallenge at 15% still produced scattered grade 1 reactions in test and control, but at 10% controls were clear as were the test group, hence the straightforward conclusion that no animals were in fact sensitised. Given the relatively low irritancy of this chemical, the LLNA was conducted at concentrations from 25% to 100%, the same as used for oleic acid itself, and indeed the result was also similar, a partial dose response with the two highest concentration exceeding the 3-fold threshold, meaning the LLNA was positive. The concurrent positive control, hexyl cinnamic aldehyde, was the same as that for the oleic acid LLNA. Again, the two test results are in conflict, but weighing additional considerations leads readily to the conclusion that the LLNA must represent a false positive. An ester of a fatty acid is no more a reactive organic chemical than the fatty acid from which it was derived and this is borne out by a clear negative result from a GPMT conducted at high concentrations and which should not be overridden by a slight LLNA positive result. Clinical evidence of dermatitis from oleic acid esters is lacking, but given their use in cosmetic products, this does represent consistent evidence that oleic acid esters are not significant skin sensitisers. Thus the weight of evidence is that oleic acid ester is not a classifiable skin sensitiser and the result of the LLNA has to be regarded as a false positive.

3.6. Fatty alcohol 1

The detailed information on this substance is contained in Table 5. The GPMT gave a positive response rate of 20% to challenge at 100%, a result which was confirmed by rechallenge at 60% in one of the two animals positive at first challenge. Recent positive control data obtained with mercaptobenzothiazole confirmed the quality of test conduct. The LLNA, also carried out at high concentrations, showed weak positive responses at all concentrations, but with no more than partial evidence of a dose response. The positive control data for the LLNA used *p*-phenylenediamine at 1%, which gave a stimulation index of 4.0. Given that experience shows PPD to be a potent allergen, with an EC3 value of around 0.1%, this

Table 3
Sensitisation test data for oleic acid.

GPMT ^a concentrations and results		LLNA ^b test concentration and results	
Induction injection	100%	% in acetone/olive oil (4:1, v/v)	Stimulation index
Induction patch	100%		
Challenge patch	1% in PEG300 ^c		
Challenge 1 result	0/10 (0%)	25	3.4
Challenge 2 result	0/10 (0%)	50	5.7
		100	6.5

^a Hexyl cinnamic aldehyde positive control gave a 40% positive response to challenge at 10%.

^b Hexyl cinnamic aldehyde positive control gave an EC3 value of approximately 8%.

^c PEG = polyethylene glycol.

Table 4
Sensitisation test data for oleic acid ester.

GPMT ^a concentrations and results		LLNA ^b test concentration and results	
Induction injection	100%	% in acetone/olive oil (4:1, v/v)	Stimulation index
Induction patch	100%		
Challenge patch	25% in PEG300 ^c		
Challenge 1 result	0/10 (0%)	25	2.8
		50	5.2
		100	4.7

^a Hexyl cinnamic aldehyde positive control gave a 40% positive response to challenge at 10%.

^b Hexyl cinnamic aldehyde positive control gave an EC3 value of approximately 8%.

^c PEG = polyethylene glycol.

Table 5
Sensitisation test data for fatty alcohol 1.

GPMT ^a concentrations and results		LLNA ^b test concentration and results	
Induction injection	5% in saline ^c	% in acetone/olive oil (4:1, v/v)	Stimulation index
Induction patch	100%		
Challenge patch	100%		
Challenge 1 result	2/10 (20%)	25	3.1
		50	4.8
		100	4.4

^a Mercaptobenzothiazole positive control gave a 80% positive response to challenge at 15%.

^b *p*-Phenylenediamine positive control at 1% gave a stimulation index of 4.0.

^c Isotonic saline.

degree of proliferation seems rather lower than might normally be expected (Basketter et al., 2007a,b,c). In this case, the results are consistent in the sense that both indicate a limited degree of skin sensitising activity, but they differ in that the GPMT outcome is below the classification threshold, i.e. is not sensitising, whereas the LLNA result is above it. Fatty alcohols are not chemicals associated with significant degree of allergic contact dermatitis, although it is not unknown. Given their use in medicaments, as well as cosmetic and household products, this is consistent with them possessing either no sensitising capacity, or possessing it only to a very limited degree. Where fatty alcohols are reported to be cause of allergic contact dermatitis, classically it is represented by cetostearyl alcohol allergy in association with stasis ulcers, but is also occasionally reported with other fatty chain lengths (Rademaker et al., 1997; Andersen and Broesby-Olsen, 2006). This provides reassurance, that even uncommon allergens are frequently discovered by clinical investigators, strengthening the observation of the absence of evidence for fatty alcohols being sensitizers. Consequently, in the absence of any further data, then the balance of regulatory judgement might well be that this material should be classified as a skin sensitizer, the positive LLNA (even in the absence of a dose response, but noting the weak positive control data) adding to the below classification threshold GPMT result. However, it is also important to consider the information below on a second example of this class.

3.7. Fatty alcohol 2

The detailed information on this second fatty alcohol (and which is of closely similar chemical structure to the first) is contained in Table 6. In this case, the GPMT was carried out with a higher induction concentration than the previous example and the very limited irritancy seen also meant that 10% SLS pretreatment had to be used prior to the induction patch. Nevertheless,

Table 6
Sensitisation test data for fatty alcohol 2.

GPMT ^a concentrations and results		LLNA ^b test concentration and results	
Induction injection	30% in saline ^c	% in acetone/olive oil (4:1, v/v)	Stimulation index
Induction patch	100%		
Challenge patch	100%		
Challenge 1 result	0/10 (0%)	25	2.2
		50	2.7
		100	4.1

^a Formaldehyde positive control gave a positive response to challenge at 10%.

^b Hexyl cinnamic aldehyde positive control gave an EC3 value of approximately 15%.

^c Isotonic saline.

the result was entirely negative. In contrast, the LLNA was positive only at the highest (100%) concentration tested and with a modest dose response. The same positive controls applied to these tests as to the previous fatty alcohol discussed above. When considered in isolation, the data for this second fatty alcohol are partly in conflict, but the high exposure levels in the GPMT, the limited dose response in the LLNA which was only positive at 100% and the lack of association of this chemistry with reactivity or with a common clinical history would balance towards non-classification as a skin sensitizer. Taken together with the information on the fatty alcohol 1, it is possible to speculate that this might be an indication of a potential sensitising impurity, but it is perhaps a better indication that this class of material may be a true false positive in this LLNA. Although the first fatty alcohol did give a low level of response in the GPMT, it must be remembered that the LLNA threshold for positivity was set up to be equivalent to classification thresholds in the GPMT and Buehler tests. Therefore a substance which is not classified as a sensitizer in a well conducted GPMT ought generally to be negative in the LLNA. Fatty alcohols have a long history of safe use in cosmetics along with the extensive OECD SIDS evaluation of long chain fatty alcohols in which the panel stated that these alcohols are not sensitizers (OECD, 2006). Comparative patch testing with an exemplary fatty alcohol, myristyl alcohol, in two centres revealed that a considerable number of doubtful reactions due to irritation and not sensitisation develop at concentrations of 10% (Geier et al., 2006). This group also stated that the complete lack of clinical relevance also demonstrates that the reactions observed are of irritant nature and not weakly allergic reactions. Combined with other inputs from chemistry, e.g. lack of structural alerts and low reactivity and other clinical experience, the discordances within and between these two fatty alcohol datasets, generally should balance away from regulatory classification.

3.8. Polyaminofunctional siloxane

The detailed information on the last substance to be considered is contained in Table 7. The GPMT was conducted following standard guidelines and practices that included pre-test evaluation to determine appropriate dose concentrations. The challenge dose level was set at 15% test article in olive oil due to the irritancy observed at higher concentrations. The irritancy observed in this GPMT study is consistent with the irritancy potential observed in previously conducted eye and skin irritancy studies with this substance. Sensitization was not observed in this GPMT employing mild-moderate irritating and non-irritating dose levels at induction and challenge, respectively. In sharp contrast, the LLNA gave a clear positive dose response, with the 2 higher concentrations producing substantial increases in the stimulation index compared to concurrent vehicle-treated controls. In this case, the 1% PPD positive control used for the LLNA gave a stimulation index of 12.1, in

Table 7
Sensitisation test data for siloxane.

GPMT ^a concentrations and results		LLNA ^b test concentration and results	
Induction injection	5% in olive oil	% in acetone/olive oil (4:1, v/v)	Stimulation index
Induction patch	75% in olive oil		
Challenge patch	15% in olive oil		
Challenge 1 result	0/10 (0%)	5	1.6
		10	3.5
		25	18.9
		50	22.8

^a 2-Mercaptobenzothiazole positive control gave a 90% (24 h reading) and 100% (48 h reading) positive response to challenge at 10% in mineral oil.

^b *p*-Phenylenediamine positive control at 1% gave a stimulation index of 12.1.

line with published data (Warbrick et al., 1999). Consequently, reconciling these contradictory datasets is more challenging. Evaluation of this irritant within the LLNA demonstrated marked increases in ear thickness and hair loss at the application site at the two highest dose levels (25% and 50%) and a marginal increase in ear thickness at 10%. In concert, the stimulation index increased in a dose response manner with values of 1.6, 3.5, 18.9 and 22.8 for the dose levels of 5%, 10%, 25% and 50%, respectively. Irritancy remains as a confounding characteristic regardless of the assay methodology for skin sensitization and in this situation one cannot be certain what role it plays in this particular assay. The relatively high concentrations used in the GPMT and the increases in ear thickness in the LLNA at high concentration are consistent with no more than a moderate degree of irritation potential, so the positive LLNA outcome, at least above the 10% concentration, cannot be explained in the same way as SLS, that is, a false positive due to irritancy. Taking a weight of evidence approach one may consider the lack of a sensitization response in the GPMT at the 15% concentration, the irritancy potential, as well as the lack of structural alerts for skin sensitization, as sufficient evidence to consider the LLNA assay result a false positive, or at least equivocal with regard to assessing the sensitization potential of this polyaminofunctional siloxane. Furthermore, based on its molecular structure, the substance is regarded as being unreactive and of a molecular weight that is not likely to readily penetrate skin. The current findings highlight the importance for defining appropriate test method strategies as well as applying a weight of evidence approach for assessing a material's potential skin sensitization hazard. There is no evidence that this type of siloxane has been associated with any degree of occupational or consumer allergic contact dermatitis, so ultimately, this particular siloxane probably does not possess significant skin sensitisation hazard potential, but such a conclusion for an entirely unknown chemical would be difficult.

4. Summary

Correct conclusions on the regulatory hazard classification of skin sensitising chemicals can only be made by a careful and scientific consideration of the weight of all the evidence. The material presented here shows that it is unwise to make a definitive assessment from any single piece of information as no individual assay or other assessment, e.g. based on structural chemistry, is 100% accurate on every occasion (Table 8). This means that from time to time, one piece of conflicting data has to be set aside; this may be a LLNA or GPMT result, it may be the absence of a structural alert or something else. Furthermore, it may be necessary to do this even though it is not possible to arrive at a convincing explanation of why the conflicting data may be a false positive or a false negative (but a potential impetus for investigative work).

Table 8

Summary of weight of evidence judgment on sensitising potential compared to results of individual tests.

Substance	Sensitising results ^a		Weight of evidence judgment for classification as skin sensitising
	GPMT	LLNA	
Fatty acid glutamate with impurity	Positive	Negative	Should be classified
Glycerol thioglycolate	Negative	Positive	Should be classified
Oleic acid	Negative	Positive	Not classified
Oleic acid ester	Negative	Positive	Not classified
Fatty alcohol 1	Negative	Positive	Not classified
Fatty alcohol 2	Negative	Positive	Not classified
Polyaminofunctional siloxane	Negative	Positive	Not classified

^a Results of tests given in more detail in Tables 2–7.

It has to be recognised that, at least in the EU, new chemicals legislation (REACH) in combination with animal welfare requirements, will normally mean that only a single skin sensitisation test is completed, and that, unless there is an adequate scientific justification, this test must be the LLNA. This raises two questions: what represents a scientific justification for using a test method other than the LLNA?; what other information contributes to the weight of evidence? To address the first of these, then clearly an important component of the scientific rationale for not conducting a LLNA is likely to include evidence, such as that in this paper and in other publications (Vohr and Ahr, 2005; Kreiling et al., 2008), that the type of chemistry being evaluated is known not to be correctly assessed in this assay. This may also be the case when testing certain subsets of chemicals that are irritants (OECD 429; Montelius et al., 1998; Woolhiser et al., 1998). It may well also be the case that this type of chemistry was not used in the original validation dataset (Dean et al., 2001) and/or that a suitable vehicle cannot be found (Dean et al., 2001; OECD, 2002).

Which brings us the general question of weight of evidence, as for many individual chemicals there may only be data derived from a single LLNA. To the LLNA output should be added insights from our knowledge of structure–activity relationships and information from the general class of chemistry. Both of those will exist for many of the thousands of chemicals to be assessed within REACH. Only for a limited number will there be significant evidence from human exposure, but where this is available and of good quality, it also should contribute substantially to the determination of the presence (or absence, if exposure has been substantial) of skin sensitisation hazard. Furthermore, where possible, the information should be supplemented by data on impurities or monomeric residues, since the latter may sometimes provide the explanation for weak sensitisation responses, as in the present case with fatty acid glutamate. Of course, it must be the case that where a weight of evidence is used, then the information and rationale involved must be justified and properly documented and the material in this paper illustrates the approach to this.

Ultimately, skin sensitisation tests, as all other toxicology assays are imperfect reflections of the reality of human toxicology. Validation indicated that the LLNA, when correlated to the guinea pig sensitisation tests, approached not more than a 90% level of accuracy and tests in either species performed less well than this in terms of prediction for humans (Dean et al., 2001). This means that they will be incorrect on a proportion of occasions and more so with individual classes of chemistry for which they are, for whatever reason, demonstrated to be poor predictors. Given the prominence of the LLNA in REACH, this does represent a situation in which those responsible for making both the type of regulatory hazard assessments discussed here, as well as assessments for occupational and/or consumer safety, should ensure that they are fully cognizant of all factors and use the weight of evidence approach when limited information is available (Basketter et al., 2007a,b,c; Gerberick et al., 2005; Basketter and Kimber, 2006; Api et al., 2008).

Acknowledgments

The authors would like to thank the European Committee of Organic Surfactants and their Intermediates (CESIO) and Centre Européen des Silicoes (CES) for their support.

References

- Andersen, K.E., Broesby-Olsen, S., 2006. Allergic contact dermatitis from oleyl alcohol in Elidel cream. *Contact Dermatitis* 55, 354–356.
- 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. *Reg. Toxicol. Pharmacol.* 52, 3–23.
- Balls, M., Hellsten, E., 2000. Statement on the validity of the local lymph node assay for skin sensitisation testing. ECVAM Joint Research Centre, European Commission, Ispra. *Alternatives to Laboratory 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., Kimber, I., 2006. Predictive test for irritants and allergens and their use in quantitative risk assessment. In: Frosch, P.J., Menné, T., Lepoittevin, J.-P. (Eds.), *Contact Dermatitis*, fourth ed. Springer Verlag, Heidelberg, pp. 179–188.
- 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 sensitisation tests. *Food Chem. Toxicol.* 36, 327–333.
- Basketter, D.A., Lea, L., Cooper, K., Dickens, A., Briggs, D., Pate, I., Dearman, R.J., Kimber, I., 1999. A comparison of statistical approaches to derivation of EC3 values from local lymph node assay dose responses. *J. Appl. Toxicol.* 19, 261–266.
- Basketter, D.A., McFadden, J., Evans, P., Andersen, K.E., Jowsey, I., 2006. Identification and classification of skin sensitizers: identifying false positives and false negatives. *Contact Dermatitis* 55, 268–273.
- Basketter, D.A., Gerberick, G.F., Kimber, I., 2007a. The local lymph node assay: current position in regulatory classification of skin sensitizing chemicals. *Cut. Ocul. Toxicol.* 26, 293–301.
- Basketter, D.A., Kan-King-Yu, D., Dierkes, P., Jowsey, I.R., 2007b. Does irritation potency contribute to the skin sensitization potency of contact allergens. *Cutan. Ocul. Toxicol.* 26, 279–286.
- Basketter, D.A., Gerberick, G.F., Kimber, I., 2007c. The local lymph node assay EC3 value: status of validation. *Contact Dermatitis* 57, 70–75.
- Basketter, D.A., McFadden, J.F., Gerberick, G.F., Cockshott, A., Kimber, I., 2009. Nothing is perfect, not even the local lymph node assay. A commentary and the implications for REACH. *Contact Dermatitis* 60, 65–69.
- Berardesca, E., Barbareschi, M., Veraldi, S., Pimpinelli, N., 2001. Evaluation of efficacy of a skin lipid mixture in patients with irritant contact dermatitis, allergic contact dermatitis or atopic dermatitis: a multicenter study. *Contact Dermatitis* 45, 280–285.
- Betts, C.J., Dearman, R.J., Kimber, I., Ryan, C.A., Gerberick, G.F., Lalko, J., Api, A.M., 2007. B220 analysis with the local lymph node assay: proposal for a more flexible prediction model. *J. Appl. Toxicol.* 27, 506–510.
- Buehler, E.V., 1965. Delayed contact hypersensitivity in the guinea pig. *Arch. Dermatol.* 91, 171–177.
- CIR Sodium Lauryl Sulfate, 1983. Final report on the safety assessment of sodium lauryl sulfate and ammonium lauryl sulfate. *J. Am. Coll. Toxicol.* 2, 127–181.
- CIR Sodium Lauryl Sulfate, 2005. Annual Assessment of Cosmetic Ingredient Safety Assessments—2002/2003. *Int. J. Toxicol.* 24 (Suppl. 1), 1–102.
- CIR Oleic acid, 2006. Annual Review of Cosmetic Ingredient Safety Assessments – 2004/2005. *Int. J. Toxicol.* 25 (Suppl. 2), 1–89.
- Cockshott, A., Evans, P., Gerberick, G.F., Betts, C.J., Dearman, R.J., Kimber, I., Basketter, D.A., 2006. Use and abuse of the local lymph node assay: a regulatory perspective. *Human Exp. Toxicol.* 25, 387–394.
- Commission of the European Communities, 2006. Regulation (EC) 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/EEC and 2000/21/EC. *Off. J. Eur. Union*, L 396/1 of 30.12.2006. Available from: <http://europa.eu.int/comm/enterprise/reach/index_en.htm>.
- Conde-Salazar, L., Baz, M., Guimaraens, D., Cannavo, A., 1995. Contact dermatitis in hairdressers: patch test results in 379 hairdressers. *Am. J. Contact Dermatitis* 6, 19–23.
- 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. *Reg. Toxicol. Pharmacol.* 34, 258–273.
- Dearman, R.J., Wright, Z.M., Basketter, D.A., Ryan, C.A., Gerberick, G.F., Kimber, I., 2001. The suitability of hexyl cinnamic aldehyde as a calibrant for the using local lymph node assay. *Contact Dermatitis* 44, 357–361.
- EC, 2001. Commission Directive 2001/59/EC of 6 August 2001 adapting to technical progress for the 28th time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. *Official Journal of the European Union* L225, 1–333.
- ECHA, 2008. Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.7a: Endpoint Specific Guidance. European Chemicals Agency, Helsinki, pp. 256–288.
- EU, 2004. EU Test Method B.42: Skin Sensitisation: Local Lymph Node Assay. Directive 2004/73/EC. *Official J. European Communities* L152.
- Frankild, S., Basketter, D.A., Andersen, K.E., 1996. The value and limitations of rechallenge in the guinea pig maximisation test. *Contact Dermatitis* 35, 135–140.
- Geier, J., Lessmann, H., Fuchs, T., Andersen, K.E., 2006. Patch testing with myristyl alcohol. *Contact Dermatitis* 55, 357–366.
- Gerberick, G.F., Ryan, C.A., Kern, P.S., Schlatter, H., Dearman, R.J., Kimber, I., Patlewicz, G., Basketter, D.A., 2005. Compilation of historical local lymph node assay data for the evaluation of skin sensitization alternatives. *Dermatitis* 16, 157–202.
- Globally Harmonised System of Classification and Labelling of Chemicals, 2007. Globally Harmonised System of Classification and Labelling of Chemicals, second revised edition. United Nations, New York and Geneva. Available from: <http://www.unece.org/trans/danger/publi/ghs/ghs_rev02/English/03e_part3.pdf>.
- Guerra, L., Bardazzi, F., Tosti, A., 1992. Contact dermatitis in hairdressers' clients. *Contact Dermatitis* 26, 108–111.
- Interagency Coordinating Committee on the Validation of Alternative Methods, 2008. Validation status of the murine local lymph node assay (LLNA) for the assessment of the contact dermatitis potential of chemicals and products, March 4–6, 2008. Available from: <iccvam.niehs.nih.gov/meetings/LLNAPanelMtg08/LLNAPanelMtg08.htm>.
- Kimber, I., Basketter, D.A., 1992. The murine local lymph node assay: collaborative studies and new directions: a commentary. *Food Chem. Toxicol.* 30, 165–169.
- Kimber, I., Dearman, R.J., Scholes, E.W., Basketter, D.A., 1994. The local lymph node assay: developments and applications. *Toxicology* 93, 13–31.
- Kimber, I., Dearman, R.J., Basketter, D.A., Ryan, C.A., Gerberick, G.F., 2002. The local lymph node assay: past, present and future. *Contact Dermatitis* 47, 315–328.
- Kligman, A.M., Basketter, D.A., 1995. A critical commentary and updating of the guinea pig maximisation test. *Contact Dermatitis* 32, 129–134.
- 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.
- Langton, K., Patlewicz, G.Y., Long, A., Marchant, C.A., Basketter, D.A., 2006. Structure activity-relationships for skin sensitisation: recent improvements to DEREK for windows. *Contact Dermatitis* 55, 342–347.
- Loveless, S.E., Ladics, G.S., Gerberick, G.F., Ryan, C.A., Basketter, D.A., Scholes, E.W., House, R.V., Hilton, J., Dearman, R.J., Kimber, I., 1996. Further evaluation of the local lymph node assay in the final phase of an international collaborative trial. *Toxicol.* 108, 141–152.
- Magnusson, B., Kligman, A.M., 1970. Allergic Contact Dermatitis in the Guinea Pig: Identification of Contact Allergens. Charles C. Thomas, Springfield, IL, USA. 141 pp.
- McGarry, H.F., 2007. The murine local lymph node assay: regulatory and potency consideration under REACH. *Toxicol.* 238, 71–89.
- Mehling, A., Eigler, D., Ball, N., Certa, H., Skirda, K., Esch, H., Haux, C., Kreiling, R., Veenstra, G., 2008. Discordant results between the LLNA and Guinea Pig Tests: specificity vs. selectivity. *The Toxicologist (Supplement to Toxicological Sciences)* 102, 8.
- Montelius, J., Wahlkvist, H., Boman, A., Wahlberg, J.E., 1998. Murine local lymph node assay for predictive testing of allergenicity: two irritants caused significant proliferation. *Acta Derm Venereol.* 78, 433–437.
- Oberemm, A., Richter-Reichhelm, H.B., Gundert-Remy, U., 2008. Use of the local lymph node assay for sensitization hazard identification. A five year experience from a regulator's point of view. *The Toxicologist* 102, 7–8.
- Organisation for Economic Cooperation and Development, 1992. Guidelines for Testing of Chemicals No. 406. Skin sensitization. Organisation for Economic Cooperation and Development, Paris, France.
- Organisation for Economic Cooperation and Development, 2002. OECD Guideline for the Testing of Chemicals No. 429: Skin Sensitisation: Local Lymph Node Assay. Organisation for Economic Cooperation and Development, Paris, France, 7 pp.
- OECD SID Initial Assessment Report: Long chain alcohols, 2006. Available from: <http://www.oecd.org/document/63/03343en_2649_34379_1897983_1_1_1_100.html>.
- Rademaker, M., Wood, B., Greig, D., 1997. Contact dermatitis from cetostearyl alcohol. *Australasian J. Dermatol.* 38, 220–221.
- Ryan, C.A., Cruse, L.W., Skinner, R.A., Dearman, R.J., Kimber, I., Gerberick, G.F., 2002. Examination of a vehicle for use with water soluble materials in the murine local lymph node assay. *Food Chem. Toxicol.* 40, 1719–1725.
- Schlede, E., Eppler, R., 1995. Testing for skin sensitization according to the notification procedure for new chemicals. The Magnusson and Kligman test. *Contact Dermatitis* 32, 1–4.
- Storrs, F., 1984. Permanent wave contact dermatitis: contact allergy to glyceryl monothioglycolate. *J. Am. Acad. Dermatol.* 11, 74–85.
- Vohr, H.-V., Ahr, J.H., 2005. The local lymph node assay being too sensitive. *Arch. Toxicol.* 79, 721–728.
- Warbrick, E.V., Dearman, R.J., Lea, L.J., Basketter, D.A., Kimber, I., 1999. Local lymph node assay responses to paraphenylenediamine: intra- and inter-laboratory evaluations. *J. Appl. Toxicol.* 19, 255–260.
- Woolhiser, M.R., Hayes, B.R., Meade, B.J., 1998. A combined murine local lymph node and irritancy assay to predict sensitization and irritancy potential of chemicals. *Toxicol. Meth.* 8, 245–256.
- Woolhiser, M., Wiescinski, C., Botham, P., Lees, D., Debruyne, E., Repetto-Larsay, M., Ladics, G., Hoban, D., Gamer, A., Remmele, M., Wang-Fan, W., Ullmann, L., Billington, R., Mehta, J., 2007. ECPA interlaboratory study investigating the suitability of an aqueous vehicle in the mouse local lymph node assay. *The Toxicologist* 96, 237.