KeratinoSens test submission - Attachment 12d_PC-surfactants.

Copy of manuscript submitted for publication – please treat confidential

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

Nicholas Ball[1], *Stuart Cagen*[2], *Juan-Carlos Carrillo*[3], *Hans Certa*[4], *Dorothea Eigler*[5], *Roger Emter*[6], *Frank Faulhammer*[7], *Christine Garcia*[8],*Cynthia Graham*[9], *Carl Haux*[10], *Susanne N. Kolle*[7], *Reinhard Kreiling*[11], *Andreas Natsch*[6] *and Annette Mehling (chair)*[12*]

[1]Dow Europe GmbH, Horgen, Switzerland; [2]Shell Health, Houston, Texas, USA; [3]Shell International BV, The Hague, Netherlands; [4]SASOL Germany GmbH, Marl, Germany; [5]Evonik Goldschmidt GmbH, Essen, Germany; [6]Givaudan Schweiz AG, Dübendorf, Switzerland; [7]BASF SE, Ludwigshafen, Germany; [8]SEPPIC, Paris, France; [9]Huntsman Corporation, The Woodlands, Texas, USA; [10]Akzo Nobel Surface Chemistry AB, Stenungsund, Sweden; [11]Clariant Produkte (Deutschland) GmbH, Sulzbach, Germany; [12]Cognis GmbH (now part of BASF), Düsseldorf, Germany

Corresponding author: Annette Mehling
Email: Annette.Mehling@Cognis.com

Running title: Comparative sensitization testing of surfactants

Key words: Alternative methods, skin sensitization, guinea pig maximization test (GPMT), hCLAT, irritation, KeratinoSens assay, local lymph node assay (LLNA), peptide reactivity assay, regulatory tests, weight of evidence

Abstract

An integral part of hazard and safety assessments is the estimation of a chemical's potential to cause skin sensitization. Currently, only animal tests (OECD 406 and 429) are accepted in a regulatory context. Nonanimal test methods are being developed and formally validated. In order to gain more insight into the responses induced by eight exemplary surfactants, a battery of *in vivo* and *in vitro* tests were conducted using the same batch of chemicals. In general, the surfactants were negative in the GPMT, KeratinoSens and hCLAT assays and none formed covalent adducts with test peptides. In contrast, all but one was positive in the LLNA. Most were rated as being irritants by the EpiSkin assay with the additional endpoint, IL1-α. The weight of evidence based on this comprehensive testing indicates that, with one exception, they are non-sensitizing skin irritants, confirming that the LLNA tends to overestimate the sensitization potential of surfactants. As results obtained from LLNAs are considered as the gold standard for the development of new nonanimal alternative test methods, results such as these highlight the necessity to carefully evaluate the applicability domains of test methods in order to develop reliable nonanimal alternative testing strategies for sensitization testing.

Introduction

An integral part of hazard and safety assessments for consumer and occupational health is the estimation of a chemical's potential to cause allergic contact dermatitis. Currently animal tests are typically used in a regulatory context to assess a chemical's potential to induce skin sensitization. Both the murine Local Lymph Node Assay (LLNA; OECD 429) and guinea pig based tests (GPTs, OECD 406; Guinea pig maximization test (GPMT) or Buehler tests) are test methods accepted by the regulatory bodies to assess this endpoint. As a 3R method, the LLNA has become the preferred method for sensitization testing in the European Union (EU) and increasingly in other countries. Within the EU, the new chemicals legislation on the registration, evaluation, authorization and restriction of chemicals (REACH) requires the submission of information on human health effects of chemicals. With few exceptions, all substances registered in accordance with REACH will require skin sensitization data. Within the framework of REACH, the local lymph node assay (OECD, 2010) is the preferred method for generating data on skin sensitizing potential. Use of other methods, including the traditionally used guinea pig tests (OECD, 1992)may only be performed under exceptional circumstances when sufficient scientific justification warrants their use.

Following the validation of the LLNA, the observation was made that the LLNA often overestimates the sensitization potential for some substances, e.g. surfactants, fatty acids, fatty alcohols and siloxanes (Basketter et al., 2009a; Garcia et al., 2010; Kreiling et al., 2008; Penninks, 2006). Indeed, the classic example of a substance eliciting false positive responses in the LLNA is the surfactant sodium lauryl sulphate (SLS). SLS was one of the substances included in the set of chemicals used in the validation of the LLNA (Dean et al., 2001). The existing data in humans and guinea pigs indicate that this surfactant is irritating but not a sensitizer in these two species. However the LLNA identified it as a sensitizer, subsequently leading to the understanding that this was a true false positive in the assay (Basketter et al., 2009b; Dean et al., 2001). Clearly no biological assay is perfect, and in the majority of cases the LLNA appears to be accurately

predictive of whether a chemical can trigger an induction of the immune system indicating a potential for being a sensitizer (Dean et al., 2001). However, if this assay is to become the only permitted assay for the future assessment of sensitizing potentials, as is by in large stipulated by REACH, it is important to establish if there are chemical classes that are incompatible with the assay as, in general, only one animal test may be conducted per endpoint due to animal welfare considerations. Based on the increased awareness within the scientific community, the recently revised version of the OECD No. 429 guideline (adopted on July 22, 2010) has taken certain aspects of applicability into account and now reads "…Despite the advantages of the LLNA over TG 406, it should be recognised that there are certain limitations that may necessitate the use of TG 406 (e.g. false negative findings in the LLNA with certain metals, false positive findings with certain skin irritants [such as some surfactant type chemicals], or solubility of the test substance). In addition, test substance classes or substances containing functional groups shown to act as potential confounders may necessitate the use of guinea pig tests (i.e. TG 406)…. " (OECD, 2010).

The increasing deliberation on the ethics of animal testing has manifested itself in a regulatory context in REACH but even more so in the European Cosmetics Directive. REACH calls for alternative test methods to be used wherever possible. The Cosmetics Directive foresees a progressive phasing out of animal tests for the purpose of safety assessments of cosmetics. Marketing and testing bans apply as alternative methods are validated and adopted through EU legislation with the goal of phasing out animal tests for cosmetics by 2013. This has motivated the development of a number of alternative non-animal test methods. However, few have gone beyond intralaboratory validation, even less have been formally validated according to the ECVAM validation procedure and only a small number have achieved regulatory acceptance at this time (http://ecvam.jrc.ec.europa.eu/).

With the current innovation in the field of non-animal test methods, a number of *in vitro* assays have been developed to specifically assess skin sensitizing potentials, the accuracy of which are

usually assessed using the LLNA as the gold standard. This is due to the ability of the assay to yield objective measurements via scintillation counting and to give information on dose responses with which an evaluation of potency is possible. Currently, the direct peptide reactivity assay (DPRA), the human Cell Line Activation Test (hCLAT) and the Myeloid U937 Skin Sensitization Test (MUSST) methods are in the prevalidation phase at ECVAM and the KeratinoSens assay will be submitted in the near future. The advantage of many of these *in vitro* assays is that the endpoints measured are linked to key stages in the mechanism leading to a skin sensitizing response. Following skin penetration, protein reactivity and the triggering of specific signalling pathways, e.g. via interleukins (Wang et al., 1999) are involved in the activation of the Langerhans cells (LC) of the skin which in turn are essential for triggering the proliferation of antigen specific T-cells. Whereas the animal tests include all these steps, the *in vitro* tests can only assess specific stages in the sensitization process. Peptide reactivity assays (Gerberick et al., 2008; Natsch and Gfeller, 2008) have been developed to assess whether the chemical can interact with synthetic peptides to mimic the formation of hapten/skin protein complexes necessary for T-cell recognition of the allergen. The keratinocytes of the epidermis are essential for generation of “danger signals”, such as interleukins IL-18, IL-1β and IL-1α, in response to irritants and/or sensitizers which are required for the activation of antigen presenting cells such as the LC of the skin. One assay measuring keratinocyte activation is the KeratinoSens assay (Emter et al.2010) which was developed following the observation that sensitizers appear to trigger the Keap1-Nrf2-ARE regulatory pathway (Natsch et al., 2010) leading to induction of genes under the control of the Antioxidant Response Element (ARE). The activation of the antigen presenting cells themselves can be assessed by using the hCLAT or MUSST assays, both of which assess the expression of specific cell surface markers of antigen presenting cells as a measure of cell activation. A certain level of skin irritation appears to facilitate skin sensitization reactions, but on the other hand skin irritation may also be a confounding factor in animal tests and in human patch tests. The skin irritation potential can now be assessed *in vitro* with the validated Episkin™ irritation assay with reduction of cell viability in a 3-D

skin model as read-out. As an option, the secretion of the danger signal interleukin-1α (IL-1α) as a measure of irritation can be used as an additional endpoint in this test. The use of an array of *in vitro* assays allows a more 'mechanistic' analysis of the various stages in the sensitizing response and generates a comprehensive dataset which can feed into a 'weight of evidence' approach. The determination of whether or not a substance is a potential human sensitizer can therefore be made more objectively.

In the current study, a comparative testing program was conducted. As the sensitization potentials of surfactants are often overestimated in the LLNA (Garcia et al., 2010; Mehling A. et al., 2008); the first part of the testing program was necessary to identify if the two animal assays gave the same predictions. The standard GPMT and LLNA were used to assess the sensitization potential of eight exemplary surfactants including five commercially available surfactants of high purity. In addition to the standard LLNA endpoints, additional parameters, including ear thickness and flow cytometry to measure the number of lymph node cells carrying the B220 marker, were included (Gerberick et al., 2002). In the second part of the testing program, the surfactants were tested in *in vitro* sensitization assays, namely the peptide reactivity assay, KeratinoSens assay and the hCLAT assay. To address irritancy, the often discussed confounding factor leading to overestimations of sensitization potentials in the LLNA, the Episkin™ assay with concurrent IL-1α quantitation was also included into the test program.

Materials & Methods

Test materials

The surfactants were selected to reflect nonionic and anionic surfactant types. Selection was also based e.g. on chain length (C12 – C16) and degree of ethoxylation (EO2 to EO6). Five of the

chemicals were obtained from Sigma Aldrich and were of analytical grade purity (>97%): sodium lauryl sulfate (SLS; Nr. 71729); tetraethylene glycol monotetradecyl ether (C14EO4; Nr. 86697), hexaethylene glycol monododecyl ether (C12EO6; Nr. 52044); n-heptyl β-d-thioglucopyranoside (Nr: H3264; thioglucopyranoside); and 1-nonane sulfonic acid sodium salt (nonane sulfonate; Nr. 74318). Due to the limited quantities commercially available in analytical grade quality, the following three surfactants were synthesized on a laboratory scale and are of the following purities with respect to the surfactant content: n-decylphenol-polyethylene glycol ether (decylphenol ethoxylate; 97% purity), hexadecan-1-ol ethoxylated-2 EO (C16EO2; 79% purity; impurity: appr. 20% free alcohol)) and iso-nonyl- β -d glucopyranoside (isononyl glucoside; 47% purity; various impurities not all identified). All tests were performed with the same batch of chemicals.

Animal tests were conducted according to Good Laboratory practice (GLP) in AALAC-approved laboratories thereby taking animal welfare laws into account. The tests were performed prior to the implementation of REACH and prior to the testing ban imposed by the European Cosmetics Directive (7^{th} amendment). Due to the numbers of experiments routinely conducted by the test institutes, historical controls were used to confirm the validity of the testing procedure and to minimize animal usage where possible.

Guinea Pig Maximization Test (GPMT)

The guinea pig maximization tests (Magnusson and Kligman, 1969) were conducted in accordance with OECD test guideline No. 406 (OECD, 1992) (OECD, 1992), EC guideline B.6 (EC, 1996). The vehicle was selected based on preliminary solubility tests performed before the study initiation date. The dissolved test substance was the highest technically achievable concentration, from which the intradermal, epidermal and challenge dilutions were prepared. Vehicles used were corn oil (Carl Roth GmbH & Co., Germany) deionized water, and polyethylene glycol 300 (PEG 300- FLUKA Chemie; CH).

In brief, 4 to 6 week old albino guinea pigs were used to assess the sensitization potential. In the pre-tests (using 2-3 animals) the concentration producing mild irritation and the maximum non-irritant concentration (MNIC) were determined. The concentrations causing mild irritation were used for intradermal and dermal induction and the MNIC was used for elicitation/challenge. Vehicle control groups were treated with the vehicle during the induction phase and the test substance during the elicitation/challenge phase. Positive controls were treated with 25% α-hexylcinnamaldehyde in dimethylformamide (historical data, conducted annually). Studies were performed on 10 controls and 20 treated animals. Dose preparations of the chemicals were made immediately prior to dosing. For induction, 0.1 mL of the test sample was applied with Freund's complete adjuvant (50%/50%) via intradermal injection into the shaved scapular region. Approximately 1 week later, 0.5 mL of the test substance was applied onto the injection sites under an occluded patch (area: 8 cm^2). After approximately 17 days, the animals were challenged with 0.2 mL (area: 9 cm^2) of the test sample applied under an occluded patch to the shaved skin of the flank of the animal. Skin reactions were evaluated 24 and 48 hours after patch removal according to a grading scale ranging from 0 (no visible damage) to 3 (intense erythema and swelling). In cases where the results at the first challenge were ambiguous, a second challenge was performed at a lower dose concentration approximately 1 week later and reassessed at 24 and 48 hours after the patch removal.

According to the EC criteria for classification and labelling requirements for dangerous substances, a product is considered to be a skin sensitizer if the tested concentration results in a positive skin reaction in at least 30% of the animals 24 - 48 h after challenge patch removal when compared to the negative controls.

Local Lymph Node Assay (LLNA)

The LLNAs were carried out in accordance with OECD guideline No. 429 (version 2002), the regulation Council Directive 86/609/EEC and OECD Environmental Health and Safety publications No. 19. In addition to thymidine measurements, flow cytometric analyses using aliquots of the lymph node cell suspensions were performed for determination of lymphocyte subpopulations. As a measure of irritation, ear thickness was measured on days 0, 2 and 5 with a gauge device. The test substance was considered to be slightly irritating if the ear thickness was increased by 10 to 25% and as irritating if ear thickness was increased by 25% or more (Kirk et al., 2007).

For the main study, 5 animals per dose group of CBA/J mouse (8 to 12 weeks old) were treated with the test substance. SLS was dissolved in DMSO; methyl ethyl ketone was used as the vehicle for all other surfactants. In these solvents and based on visual assessments, emulsions or homogeneous suspensions were obtained. Negative control animals were treated identically with the vehicle only. Positive controls were treated with α-hexylcinnamaldehyde in dimethylformamide (historical controls). The test substance was applied (25 µL per ear) to the dorsum of both ears on three consecutive days (study days 0, 1 and 2). Following a two day rest period, 250 µL [20 mCi] ^{3}H- thymidine was injected into the tail vein of the mice (day 5). The auricular drain lymph nodes were excized five hours later. Cell suspensions of the pooled lymph nodes of each animal were prepared. The lymph node response was evaluated by measuring the incorporation of ^{3}H- thymidine in the cells by β-scintillation counting and cell count. In addition, flow cytometric analyses of B220 (CD45RA) were performed using fluorophore labelled antibodies obtained from Becton Dickinson, (Heidelberg) and a FACs calibur flow cytometer. The Stimulation Index (SI) was calculated according to the following formula: SI = dpm of treated group/dpm of vehicle control group. According to OECD 429 guideline and the EC criteria for classification and labelling requirements for dangerous substances, a substance was classified and labelled as a skin sensitizer if the SI for any of the dose groups was > 3. An increase in B220 positive lymphocytes of over 1.25 fold was considered to indicate a skin sensitizing potential (Gerberick et al, 2002). In cases where a SI of >3

was observed, the EC3 (the concentration that produced the SI of 3) was determined using linear extrapolation between doses.

Peptide reactivity assay

The synthetic peptide Cor1-C420 (Ac-NKKCDLF) was obtained from Genscript Inc. (purity of 95.7%; Piscataway, NJ, USA). The reaction conditions used are described in (Natsch and Gfeller, 2008b). The assay was performed in the presence and absence of metabolic activation. To determine peptide depletion and peptide-adduct formation in the absence of metabolic activation, reactions between 1 mM test compound and 0.1 mM test peptide were carried out in a final volume of 1 mL in HPLC vials for 24 h. LC–MS analysis was performed as described previously (Natsch and Gfeller, 2008a; Natsch and Gfeller, 2008b) on a Finnigan LCQ classic Mass spectrometer (Thermo Finnigan, San Jose, CA, USA) operated in the ESI(+) mode. Mass spectra were recorded from 200 to 2000 amu. The peptide reactivity assays conducted in the presence of metabolic activation (human liver microsomes; HLM) were run under the following conditions: 0.5 mM of test peptide, 2.5 mM NADPH and 0.2 mM of test compound were dissolved in a total volume of 200 µl of Na-Phosphate buffer (20 mM, pH 7.4) and amended with 10 µl of HLM (Gentest, Woburn, MA, USA). Samples were mixed by inversion, then incubated at 37°C without shaking for 2 h. The reaction was stopped by adding 60 µl of acetonitrile. Samples were centrifuged, filtered and analyzed by LC-MS for new adduct peaks. As controls, parallel samples were prepared without the addition of HLM. Cinnamic aldehyde (a direct-acting hapten) was used as positive control in the direct reactivity assay, and the pro-hapten eugenol was the positive control for the assay with metabolic activation. All assays were run in triplicates. A substance is considered to be protein reactive if a covalent adduct is formed with an apparent molecular mass higher than the test peptide. Depletion of peptide through the formation of protein dimers is not considered to be sufficient evidence of direct protein reactivity.

ARE-Assay (KeratinoSens)

The KeratinoSens cell line is derived from the human keratinocyte culture HaCaT (Boukamp et al., 1988). It contains a stable insertion of a Luciferase gene under the control of the ARE-element of the gene AKR1C2. The optimization of this cell line has been described in detail (Emter et al., 2010). All tests were run according to the method published previously. Briefly, cells were grown for 24 h in 96-well plates. The medium was then replaced with medium containing the test chemical and the solvent DMSO at a final level of 1% of. Each compound was tested at 12 binary dilutions in the range from 0.98 to 2000 µM. Each test plate contained 7 test chemicals, 6 wells with the solvent control, 1 well with no cells for background value and 5 wells with the positive control, cinnamic aldehyde, in five different concentrations. In each repetition, three parallel replicate plates were run with this same set-up and a fourth parallel plate was prepared for cytotoxicity determination using an MTT assay. Cells were incubated for 48 h with the test agents, and then luciferase activity and cytotoxicity were determined. This full procedure was repeated at least three times for each chemical. For each chemical in each repetition and at each concentration, the gene induction compared to DMSO controls and the wells with statistically significant induction over the threshold of 1.5 (i.e. 50% enhanced gene activity) were determined. Furthermore the maximal fold-induction (I_{max}) and the EC1.5 value (concentration in µM for induction above the threshold, based on linear extrapolation as done in the LLNA) were calculated. Chemicals are rated as positive in the assay if the following three criteria are fulfilled (i) The EC1.5 value is below 1000 µM in all three repetitions or in at least 2 repetitions, (ii) at the lowest concentration with a gene induction above 1.5 fold (i.e. at the EC 1.5 determining value), the cellular viability is above 70% and (iii) there is an apparent overall dose-response for luciferase induction, which is similar between the repetitions

EpiSkin™ assay

In order to compare the skin irritation potential, the test items were applied to EpiSkin™ tissues and the MTT-reducing capacity and the Interleukin-1α release by the tissue-samples were measured

after 48 h as indicated by the manufacturer (SkinEthics, Nice, France). In the validated EpiSkin™ assay only neat chemicals are used as the prime goal is classification and labelling. However, in order to gain more quantitative information on the irritation potential, dilutions of the test materials were used to generate a dose–response curve. The test items were diluted in dipropylene glycol, which is non-irritating to the tissue-samples, or in PBS and these dilutions were tested according to the SOP. DPG rather than the rapidly evaporating MEK used in the LLNA was chosen as vehicle, since it is non-volatile and thus actual concentration dose-response curves can be measured. IL-1α was determined using an ELISA kit obtained from Diaclone/Biotest (Rupperswil, Switzerland) and the IL-1α reference standard obtained from R&D systems (Minneapolis, MN, USA). According to the SOP for the assay, reduction of viability > 50% and/or release of > 80 pg/mL IL-1α would suffice to rate a chemical as a potential irritant, although only the viability endpoint was included in the formally validated method.

Human Cell Line Activation Test (hCLAT)

The dendritic cell activation assay using THP-1 cells was performed as described by Ashikaga and co-workers (Ashikaga et al., 2006) and Sakaguchi and co-workers (Sakaguchi et al., 2006) with some modifications. Briefly, THP-1 cells (DSMZ, Braunschweig, Germany, ACC 16) were cultured in RPMI1640 1x GlutaMAX™ medium (Gibco, Darmstadt, Germany) supplemented with 10% FBS (Biochrom, Berlin, Germany) 0.05 mM 2-mercaptoethanol (Invitrogen, Darmstadt, Germany) and 1% penicillin/streptomycin (Biochrom, Berlin, Germany). A range finding test was conducted prior to the main assay to determine the concentration associated with 75% THP-1 cell viability (CV75). For the main test, THP-1 cells were incubated in 24well plates with test chemicals for 24 h using eight concentrations selected from preliminary range-finding studies. After 24 h test substance incubation, cells were stained with anti-human CD86 (BD Pharmingen, Heidelberg, Germany) or CD54 (DAKO, Eching, Germany) antibodies coupled with the fluorescent dye FITC. Concurrently, cells were stained with propidium iodide for evaluating cell viability at each

concentration. The mean fluorescence intensity (MFI) was measured by flow cytometry using a FC 500 MPL (Beckman Coulter, Krefeld, Germany) and the MXP Software (Beckman Coulter, Krefeld, Germany). Analysis was performed using the CXP Software (Beckman Coulter, Krefeld, Germany) and results were expressed as fold induction of CD86 or CD54 expression compared to the respective vehicle control.

A chemical was predicted to be a potential sensitizer if it stimulated CD86 or CD54 expression above the set threshold of 1.5 at sufficiently non toxic concentrations (cell viability >70%) in at least two experiments. The strong sensitizer DNCB (3.0 µg/mL, Sigma Aldrich, Germany) was used as positive and lactic acid (LA, 500 µg/mL, Sigma Aldrich, Germany) as nonsensitizing negative control. The final concentration of DMSO, when used as solvent for the test substances was 0.2%.

Identification of structural alerts

The OECD QSAR Toolbox (Version 2.0) was used to identify whether the chemical structures contained any DNA or protein binding structures.

Results

In the first part of the testing program the intention was to clarify if the two *in vivo* assays gave the same predictions. The standard GPMT was used, however in addition to the standard LLNA endpoints, additional endpoints such as ear thickness and flow cytometry to measure the number of lymph node cells carrying the cell surface marker B220 (CD45RA) were included (Table 1). In the second part of the testing program, *in vitro* tests were conducted to assess how the surfactants performed in these tests and to verify the sensitization potentials based on the *in vivo* tests used namely the peptide reactivity assay, the KeratinoSens assay, the hCLAT assay and the Episkin ™

assay. The results of the KeratinoSens and Episkin ™ assays are depicted in Tables 2 and 3. The results from the hCLAT test were all negative at or above 70% viability. All tests were carried out with the same batches of chemicals to avoid the presence of possible impurities or other sample variations which could possibly lead to confounding results. Results from *in vivo* and *in vitro* studies were used to classify the substances for sensitizing potential using a weight of evidence approach, which is presented in Table 4.

***In vivo* tests**

Seven of eight surfactants were assessed as being non-sensitizers in the GPMT. Of these, six chemicals elicited no responses in the GPMT (Table 1). C16EO2 elicited responses in 75% of the treated animals (15 of 20 animals) at a concentration of 10% for the challenge dose. As 70% (7 of 10 animals) of the animals in the control group also developed reactions after having been challenged with this concentration, these reactions were assessed as being caused by irritation and not sensitization. A rechallenge at a lower, non-irritating dose of 1% elicited no reactions in either group and therefore C16EO2 was assessed as being a non-sensitizer. Isononyl glucoside (laboratory scale; purity 47%) elicited reactions in 85% (17 of 20 animals) of the animals and 20% of the controls when the challenge was conducted at a concentration of 100%. Following rechallenge with a concentration of 50%, half the test animals still developed reactions whereas none of the controls did. Isononyl glucoside was synthesized on a laboratory scale and only had a purity of 47% with respect to surfactant content. Although this specific preparation containing isononyl glucoside was assessed as being a sensitizer in guinea pigs, the surfactant itself need not be the sensitizer but sensitization may be caused by an impurity (see below).

The SIs for seven of eight chemicals tested in the LLNA exceeded the threshold of three for at least one concentration and would therefore be classified as being sensitizers (Table 1). The SIs of SLS and isononyl glucoside only barely exceeded three (3.09 and 3.42, respectively) and were considered to be positive/equivocal. Only the nonane sulfonate was negative. When using ear

swelling as a measure of irritation, six of eight would be considered irritants as ear swelling increased by over 25% for at least one concentration. Ear swelling was not observed when testing nonane sulfonate and decylphenol ethoxylate.

The B-cell cell surface marker B220 (CD45RA) has been discussed as a marker to differentiate between irritants and sensitizers in the LLNA (Gerberick et al., 2002). The cells of the lymph nodes were therefore subjected to flow cytometric analyses. The model used in the initial study defined that a B220 test:vehicle ratio cut off of 1.25 for discriminating between allergens ($\geq$1.25) and irritants (<1.25). Based on this cut off, six of eight surfactants induced a B-cell proliferation that, in conjunction with an SI of $\geq$3, would rate them as being sensitizers. Only the nonane sulfonate and the isononyl glucoside would not be considered to be sensitizers whereas SLS and thioglucopyranoside would be assessed as being borderline sensitizers. The correlation index between the measured stimulation indices (DPM) and B220 scores was not conclusive suggesting that B220 would not be a suitable additional marker to differentiate between irritation and sensitization. Furthermore, SLS, a well known irritant and not a sensitizer, would be considered a borderline sensitizer based on the B220 results.

When assessing the sensitizing potentials of substances, the dose per area plays a substantial role (Kimber et al., 2008). Although the GPMT and LLNA assess different endpoints and stages in the sensitization process, have different application protocols and use different species, a tentative estimate of the dose/cm^2 can be made. As reported by Garcia et al, 2010, an estimate of the amount applied in a single application during epidermal induction can be made based on the following calculation: GPMT (8 cm^2 application area; 0.50 mL = approx. 0.50 g applied (not taking density into account) which results in an application volume of 0.0625 g/cm^2; LLNA (ear assumed to be 1 cm^2 area; 25 µL = approx. 0.025 g applied (not taking density into account) which resulted in an application volume of 0.025 g/cm^2). Although the amount applied per square area is somewhat

higher in the GPMT, the dose metrics are of a similar order of magnitude and rough comparisons of the dose metrics via the concentration expressed as the percentage of the substance tested can be made. With the exception of SLS and nonane sulfonate, the concentrations used in the induction phase and in most cases the challenge phase of the GPMT (Table 1) were higher than the concentrations applied in the LLNAs, suggesting that the dose metrics are not the cause for the conflicting results in the LLNA and GPMTs.

Peptide reactivity assay

An LC-MS based peptide reactivity assay was used to simultaneously determine peptide depletion, adduct formation and peptide oxidation. Since the covalent reaction of sensitizers with proteins is considered a hallmark in the induction of skin sensitization, the formation of covalent adducts in this assay is considered as evidence to rate a chemical as a skin sensitizer (Natsch and Gfeller, 2008b). None of the eight test chemicals formed a covalent adduct with the test peptide (Table 2). Significant depletion of the test peptide was only recorded for C12EO6, yet this was attributed to peptide dimerisation, as the dimer peak increased accordingly. In contrast, the positive control, cinnamic aldehyde, depleted the peptide by 48% with the simultaneous formation of a direct adduct peak with the mass of 1040.3 and without significant peptide dimerisation being observed.

It is often argued that chemicals, which are not directly reactive with peptides or proteins, may become reactive by metabolic activation. These are termed prohaptens. The LC-MS assays were thus additionally conducted under modified conditions in which human liver microsomes (HLM) were added. In the absence of HLM, no adduct formation for the prohapten eugenol, which served as positive control, was observed (data not shown). When eugenol was added to the test peptide in the presence of HLM a single new chromatographic peak with a base ion *m/z* 1071.4 in the ESI-spectrum was observed (data not shown). This result can be explained by enzymatic oxidation of eugenol to a quinone methide and subsequent adduct formation. Nevertheless, for all the 8 tested

surfactants no new adduct peaks were detected following incubation in the presence of HLM. Based on the results obtained in these experiments, there is no indication that the surfactants are prohaptens.

KeratinoSens assay

Table 2 gives the results as maximal gene induction (I_{max}), concentration for significant gene induction (EC 1.5) and IC50 values for cytotoxicity. Figure 1 depicts examples of dose-response curves. Only chemicals that induce a dose-dependent luciferase induction at non-cytotoxic concentrations are considered to be sensitizers. C12EO6 and SLS induced gene activity in several replicates but only at partially cytotoxic concentrations (see Fig. 1 for the example of SLS). A borderline induction in 6 of 12 repetitions was recorded for isononyl glucoside. The remaining test items did not significantly induce the luciferase gene, however the nonionic surfactants in particular exhibited a pronounced cytotoxicity. In summary, isononyl glucoside would be considered a questionable/borderline sensitizer and the remaining test items would be rated as non-sensitizers, as they do not induce the luciferase gene at non-cytotoxic concentrations.

EpiSkin™ Assay

To characterize the skin irritation potential of the test chemicals *in vitro,* they were tested in a dose-response analysis in the EpiSkin™ assay. The results are summarized in Table 3 and Figure 2 gives examples of dose-response curves obtained. The irritation potential of SLS is reflected by a strong cytotoxicity down to concentrations of 3.12% and high induction of IL-1α even at the 1.56% dose. Compared to SLS, nonane sulfonate has a clearly lower cytotoxicity and induces a much lower release of IL-1α but it would still be rated as an irritant. Isononyl glucoside is only irritating at 100%, with weak IL-1α release at lower doses. Compared to isononyl glucoside, thioglucopyranoside is rated as a much stronger irritant. At first sight, the results of the four non-ionic ethoxylated surfactants are somewhat surprising. Given the fact that these chemicals exhibit a

very pronounced cytotoxicity in the monolayer cultures in the KeratinoSens assay, they have a relatively low cytotoxicity when tested on the intact epidermis which concurs with available human data from other chemicals (Jirova et al., 2010). On the other hand, their capacity to trigger very pronounced release of IL-1α clearly underlines their irritation potential: Thus decylphenol ethoxylate triggered the formation of up to 800 pg/mL of IL-1α, whereby 80 pg/mL IL-1α would suffice to rate a chemical as a potential irritant according to the original SOP of the EpiSkin™ assay. High levels of IL-1α were also observed for C14EO4 and C12EO6. A low cytotoxicity and lower IL-1α release were observed for C16EO2; which is unexpected given its high cytotoxicity in monolayer cultures. The irritancy as assessed by the MTT part of the assay is in agreement with the IL-1α for five out of the seven surfactants. For the three where there was a difference in prediction, for two test items the MTT indicated non-irritant and for one chemical the IL-1α predicted non-irritant.

Human Cell Line Activation Test (hCLAT)

Lactic acid (negative control) and DNCB (positive control) were not cytotoxic to the THP-1 cells. CD54 and CD86 expression was not induced after 24 hour treatment with lactic acid and but was induced after 24 hours exposure to DNCB thus confirming the validity of the assay. The rate of expression was within the range of the historical negative and positive control data (data not shown). None of the test substances induced the expression of the CD54 or CD86 markers within the range of a viability of 70% and above. There was no evidence of dendritic cell activation and therefore the test materials were assessed as being unlikely to cause skin sensitization.Although no changes in the classification would occur, it should be noted that a new prediction model has been proposed (Sakaguchi et al., 2010). The prediction model indicates the threshold for CD54 should be increased to 2.0 (CD 54 RFI > 200).

(Quantitative) structure activity relationships

(Q)SARs are computer-based models which are designed to predict the physico-chemical properties, potential human health and environmental effects of a substance from knowledge of its chemical structure. The surfactant structures were investigated for structural alerts using the OECD toolbox (Version 2.0). None of the surfactants displayed any DNA or protein binding potentials. Structural alerts were found for the impurity (5-((nonyloxy)methyl)furan-2-carbaldehyde) found in the isononyl glucoside and which was identified using the KeratinoSens assay (Fig. 3, see below). This substance was identified as having structural alerts associated with Schiff-base formation with aldehydes or as being a direct acting Schiff-base former itself.

Identification of potentially sensitizing impurity in isononyl glucoside

A borderline induction in 6 of 12 repetitions was noted for isononyl glucoside in the KeratinoSens assay. Since this preparation has a low purity and isononyl glucoside has no structural alert for skin sensitization, the sensitization in the GPMT might be due to a potentially sensitizing impurity. The isononyl glucoside preparation was fractionated by hexane extraction. It contained 0.75% of hexane extractable materials. This fraction as well as the aqueous fraction and the original preparation itself were compared in the KeratinoSens assay (Figure 4). A shallow dose response was observed for the original preparation, but it was (on the average of all 12 repetitions) not above the threshold of 1.5. The aqueous phase remaining after hexane extraction gave no dose response and there was no induction observed even at the highest test dose. On the other hand, a clear gene-induction above the threshold of 1.5 was noted in 6 out of 7 repetitions for the residue from the hexane phase (Figure 4). This was indicative for this fraction containing the sensitizing substance found in the isononyl glucoside preparation. This fraction was therefore further analyzed by GC/MS. It contained low levels of the isononanol isomers used for synthesis. It did not contain the isononyl glucoside itself. The main constituents were a number of isomers, all with apparent molecular ions of *m/z* of 252. Based on the observed fractionation pattern, high resolution GC/MS analysis and database comparisons they were proposed to be 5-((nonyloxy)methyl)furan-2-carbaldehyde. A reference of

this compound (i.e. the n-nonyl isomer) was thus synthesized and compared to the unknown peaks in the hexane fraction. It showed identical retention times and an identical fragmentation pattern in the MS-spectrum with one of the isomer peaks and GC/MS analysis confirmed that the other peaks with a molecular weight of 252 are all isomers of this molecule with the typical methyl branching in the alkyl chain of the isononanol used in the synthesis (data not shown). The synthetic compound was then tested in the KeratinoSens assay and found to be positive in 3 of 6 repetitions with an average I_{max}. of 1.56. Interestingly, a synthetic n-heptyl analogue gave a more clear-cut result with an I_{max}. of 2.7-fold and all repetitions positive, confirming that these 5-((alkyloxy)methyl)furan-2-carbaldehydes can induce the gene activity in the KeratinoSens assay. The synthetic references were further analyzed in the LC-MS peptide reactivity assay. The n-nonyl derivative formed a main adduct with the test peptide with an apparent mass of 1164.5 (data not shown). When the heptyl-homologue was tested, it formed an adduct with an apparent mass of 1136.5 (i.e 28 daltons less), thus indicating that the observed peaks clearly are adducts of the 5-((alkyloxy)methyl)furan-2-carbaldehyde. This molecular mass, however, does not conform to either Schiff-base formation nor Michael-addition, and these 5-((alkyloxy)methyl)furan-2-carbaldehydes appear to react with the test peptide by an unknown mechanism. Nevertheless, they are clearly peptide-reactive and thus potential skin sensitizers based on adduct formation. Taken together, these results indicate that it is likely that they form the sensitizing impurity responsible for the clearly positive result in the guinea pig assay and if not due to irritation, possibly also for the slightly positive result in the LLNA.

Discussion

The potential of a chemical to induce allergic contact hypersensitivity is an integral part of hazard assessments in a regulatory context and/or in risk assessments, therefore reliable methods are needed to assess a chemical's intrinsic potential to trigger an immunological response in terms of skin sensitization. Skin sensitization is a complex biological phenomenon entailing numerous

sequential steps including skin penetration, induction of danger signals, peptide binding, possible metabolic steps and activation of antigen presenting cells. Much of the sensitization data available to date has been generated using animal models, i.e. the guinea pig (OECD guideline 406) or mouse models (OECD guideline 429). The LLNA was developed to assess skin sensitization by assessing lymph node cell proliferation in mice during the first phase of the allergic response, induction, while avoiding the second phase of the allergy development, elicitation, thereby reducing animal stress. It implements certain aspects of the 3R approach (reduce, refine, replace) by reducing animal stress and, depending on the study design, reducing animal numbers. During its formal validation at ICCVAM, the accuracy of the results was compared to the guinea pig tests used previously and was found to have more than 80% accuracy (Dean et al., 2001). The LLNA also allows an assessment of the potency of a sensitizer (determination of an EC3 value; (Kimber et al., 2001) which can optimize risk assessments. It has therefore become the method of choice under REACH and is used as the gold standard for validation of new non-animal alternative test methods. The development of nonanimal alternative methods to replace animal testing is the final goal and its importance is reflected in both the European Cosmetics Directive in which a phasing out of animal tests is required and in REACH.

In the first phase of this testing program the GPMT was compared to the LLNA for the set of eight exemplary surfactants. Overall there was very poor concordance between the results of these two assays for the set of surfactants. The two assays agreed for only one out of the eight surfactants. For a second surfactant the LLNA result was in partial agreement with the GPMT (isononyl glucoside) giving an equivocal/positive response compared to a clear positive in the GMPT. There are a number of possible reasons that these two assays could give different results, including species differences, test material differences between batches, vehicle choice (e.g. Wright et al., 2001), and test concentrations used. Considering species differences, in the absence of any additional information it would be difficult to determine which species is more predictive for humans. Of the group of surfactants, the only one where human data exists is SLS and this supports the negative

results from the GPMT (Basketter et al., 2009a). The same batches of test materials were used in both assays so it is unlikely that variations in the batches were the cause of any discordance. The vehicle choice for the two assays was different, however the vehicles chosen were recommended by the test guidelines. It is unlikely that the test concentrations used in the two assays were responsible for any difference in the results since the percentage concentration at the EC3 in the LLNA was in most cases lower than or equal to those used for the induction and challenge in the GPMT (Table 1). It has been reported previously (Gerberick et al. 2002) that using flow cytometry to assess the number of lymphocytes carrying the B220 marker can assist in determining when a substance is a sensitizer or an irritating, non-sensitizer. In this study however, the inclusion of flow cytometry analysis in the LLNA protocol did not aid in interpretation of the discordance between these two assays since the B220 marker was consistent with the SI for incorporation of ^{3}H-thymidine for seven out of the eight surfactants. Therefore, if considering only the data from the two animal assays, it is not possible to draw a firm conclusion on which is most predictive of the sensitizing potential in humans.

Although surfactants themselves may not be sensitizers, animal studies conducted by Karlberg et al. (2003) revealed a possible allergenic potential of oxidation products formed by long term exposure of an ethoxylated surfactant to air (10 months under continual stirring). It is unlikely that oxidation products of the ethoxylated surfactants were the cause of the discordant results reported in this paper since the test materials were not continually exposed to air for a significant period during the testing program and the LLNA and GPMT assays were conducted at the same time. Additionally, it seems plausible that any oxidation products that could have been formed would have influenced the results of both *in vivo* and *in vitro* assays rather than just the LLNA.

Taking into consideration information from humans; although consumers are often exposed to surfactants in cosmetic formulations, e.g. as shampoos or body washes, as well as in household cleaners, reports of sensitization are rare. Due to their irritancy, surfactants are difficult to assess

for sensitizing potential in diagnostic patch testing and false positive assessments are not unusual . An example of this is the surfactant cocoamidopropyl betaine (CAPB), which is often included in the patch test series for hairdressers. However, an extensive evaluation of studies by the IVDK (Information Network of Departments of Derrmatology, Germany) has shown that CAPB is not considered to be a relevant sensitizer and that sensitization may often have been caused by impurities (Schnuch et al., 2011). .In a recent study published by Corazza et al. (2011) eight different types of surfactants (anionic, non-ionic, amphoteric and cationic) were tested for their sensitization potentials via human repeated insult patch testing (hRIPT) and no sensitization was observed. This work corroborates the observation that human evidence of sensitization following surfactant exposure is rare.

In order to build the weight of evidence case, the second phase of the testing program used a selection of the available *in vitro* assays currently or in the near future undergoing prevalidation at ECVAM. At the time the program was designed, no suggestion had been made on how to include *in vitro* data in such an assessment therefore this program used a selection of four assays that represented different stages in the mechanism of the sensitising response. The four assays chosen were the peptide reactivity assay, the KeratinoSens assay, the EpiSkin™ assay, and the hCLAT assay. The KeratinoSens assay gives information on the responses induced by a sensitizer in the keratinocytes. The peptide reactivity assay provides information on the potential of a hapten to bind to proteins and in turn become an allergen. The hCLAT assay measures the responses elicited in the antigen presenting cells which are necessary for activating T-cells. The EpiSkin™ assay assesses the potential irritancy of substance and has not been studied in great detail with respect to correlations with sensitizing potentials as such. However one school of thought regarding false positives in the LLNA is that some irritants appear to confound the assay. This was one explanation for the positive response of SLS in the LLNA (Cumberbatch et al., 2002; Jacobs et al., 2006; Woolhiser et al., 1998). The EpiSkin™ assay was therefore included to determine whether any of the eight surfactants could be considered as irritants thus giving some insight into whether this is a

potential confounder of the LLNA. With the exception of the peptide reactivity assay all the assays use human tissues, reducing some of the uncertainty in extrapolating from the assay predictions to the situation in man. One important consideration however is that these assays are all undergoing (or about to) pre-validation at ECVAM. This entails that they have already undergone extensive method validation including intra- and inter-laboratory assessments. An eventual replacement of the *in vivo* assays with a single *in vitro* assay may be difficult and it is more likely that a battery of assays will be required and a weight of evidence assessment then made (Jowsey et al., 2006).

For seven of the eight surfactants the results of the peptide reactivity assay, KeratinoSens assay and hCLAT assay were consistent with the GPMT results, i.e. they all supported a prediction of no sensitizing potential. The exception was isononyl glucoside. In the peptide reactivity assay and the hCLAT assay this was not considered to be a potential sensitizer. However in the KeratinoSens assay there was some indication for sensitizing potential in the absence of cytotoxicity. This test material had the lowest purity (47%) and subsequent work identified the presence of an impurity that appears to be a potential sensitizer. Use of *in vitro* tests is a novel approach to characterize sensitizing preparations as they may be used to identify impurities causing sensitization (Natsch et al., 2010). In this study, a fraction which did not contain the isononyl glucoside itself was identified as containing a potential sensitizer by the KeratinoSens assay. In light of this data, it is plausible that the equivocal/positive response seen in the LLNA was due to the test concentration, as in this case the GPMT was performed at a higher concentration. The concentration of the impurity (0.75%) may have only been sufficient at the 100% application in the GPMT to cause the positive response.

It is evident from Table 3 that the irritation potential of these surfactants is not consistent across the different assays. In the LLNA six out of eight of the surfactants produced an increase in ear thickness. In the KeratinoSens assay only five were strongly cytotoxic which may indicate irritation. In the EpiSkin^{TM} assay the two different endpoints (MTT and IL-1α) disagreed for three

of the surfactants. All surfactants elicited signs of irritation in at least one of the tests. However, especially based on the release of IL-1α, all but nonane sulfonate would be considered to be clear irritants which would be in good accord with the positive results obtained in the LLNA. This would suggest that irritation, and in particular the release of IL-1α, may be the reason for, or a contributing factor in the false positive results. Proliferation of the lymph node cells is assumed to be a direct correlation to the sensitization process, yet irritation has also been reported to induce Langerhans cells (LC) to migrate to the regional lymph nodes which may induce nonspecific lymph node proliferation. Cumberbatch et al. (2002) reported that the irritant SLS was able to induce the migration of LC to the draining lymph nodes and that this process was IL-1α dependent. This would agree well with the results obtained in this study, as high levels of secreted IL-1α concentrations appeared to coincide with high stimulation indices (SI) in the LLNA. Indeed, a tentative correlation was also found when the Episkin™ assay was being optimized (Coquette et al., 1999) where the sensitizer DNCB did not induce increased levels of IL-1α but SLS did. Although the data set for the surfactants reported here is limited, and much work would need to be done in the future to verify this correlation, IL-1α may well be a marker which would assist in the discrimination of irritating non-sensitizers and sensitizers.

No experimental assay is perfect, however since the aim of performing these assays is to predict the potential hazard to man, it is important to understand if there are groups of substances where the available assays may over or underestimate the potential hazard. Since its introduction, there have been publications reporting apparent false positive results from the LLNA for certain groups of chemicals, including surfactants (Garcia et al., 2010; Kreiling et al., 2008; Woolhiser M.R. et al., 1998), the conclusion being that the LLNA may in some cases be misrepresenting the sensitizing potential of the tested substances. The basis for claims that the assay gives ‘false positives’ comes from contradictory data from experimental results obtained from other animal models, e.g. the GPMT, human data (hRIPT, worker/consumer experience), and the lack of structural alerts within

the substance that are associated with sensitizing potential. The classic example of a chemical which produces false positive results in the LLNA is the surfactant SLS. In this study, eight surfactants (including five commercially available surfactants) were assessed using both *in vivo* and *in vitro* studies to build a weight of evidence case to ascertain their sensitizing potential. Although the LLNA and the GPMT results are contradictory for six of the surfactants, the weight-of-evidence supports the conclusion that the positive results in the LLNA are likely to be 'false-positives' or in other words, not predictive of the skin sensitizing potential (Table 4). This conclusion is based on following evidence

(i) the molecules identified as positives in the LLNA are irritating as observed at higher doses in the LLNA (in most cases) and based on the EpiSkin™ *in vitro* irritation result (in particular IL-1α release)

(ii) they have no structural alerts for sensitizing potential

(iii) they are not peptide-reactive and there is no evidence for them acting as prohaptens

(iv) they are negative in the KeratinoSens assay

(v) they are negative in the hCLAT assay

(vi) surfactants are rarely clinically relevant allergens

Conclusion

Incorporating the available *in vitro* assays into an overall weight-of-evidence assessment of the human skin sensitizing potential of eight exemplary surfactants has made it possible to conclude that they would not be considered as potential human skin sensitizers. This conclusion is driven primarily by the concordance between the *in vitro* data, the lack of structural alerts and the results of the GPMT. In addition, an *in vitro* assay was successfully used to identify a potentially sensitizing impurity - a new approach which is not possible using animal studies due to animal welfare aspects. Based on the results presented in this study, future *in vivo* testing of surfactants for sensitizing potentials should be done using one of the available guinea pig test guidelines to ensure the highest

relevance of the results to humans. In addition, although more work needs to be done, the use of the EpiSkin™ IL-1α measure (possibly in conjunction with an additional *in vitro* test, e.g. with a hCLAT assay) may help in defining which *in vivo* test is more suitable when testing an unknown or an uncharacterized substance or to interpret data from substances for which conflicting data is available from a LLNA or GPT.

In the future, the use of animals to assess the potential for skin sensitization in humans is likely to be replaced by one or more of the *in vitro* assays currently being developed. The assays used in this work are all promising and going through prevalidation with ECVAM. Their use in a tiered or combined screening set or as a stand alone method will need to be defined in the future. This testing program has demonstrated how a selection of assays can be used to characterize the sensitizing potential of a substance. However, it is not the aim of this work to dictate the choice of assay(s) in the future. Additionally, the results of this work indicate the importance of not validating a new assay against a single existing assay when assessing its predictive power. Rather, the new assays should be assessed against all available data for a substance in a weight of evidence approach to ensure that errors from one assay are not carried through into the next generation of assays.

Conflict of interest:

Other than employment, the authors are not aware of any conflicts of interest.

Acknowledgements

The authors would like to thank the European Committee of Organic Surfactants and their Intermediates (CESIO) and Givaudan Schweiz AG for the support given, Adama Traore for his help with the structural alerts, Nicole Frijus-Plessen for helpful comments, Hans Gfeller and Joachim Schmid (Givaudan analytical department) for LC/MS and GC/MS analysis and structure elucidation and Thierry Granier (Givaudan Bioscience) for synthesizing 5-((nonyloxy)methyl)furan-2-carbaldehyde.

References

Ashikaga, T., et al., 2006. Development of an in vitro skin sensitization test using human cell lines: the human Cell Line Activation Test (h-CLAT). I. Optimization of the h-CLAT protocol. Toxicol In Vitro. 20**,** 767-73.

Basketter, D., et al., 2009a. Application of a weight of evidence approach to assessing discordant sensitisation datasets: implications for REACH. Regul Toxicol Pharmacol. 55**,** 90-6.

Basketter, D. A., et al., 2009b. Nothing is perfect, not even the local lymph node assay: a commentary and the implications for REACH. Contact Dermatitis. 60**,** 65-9.

Boukamp, P., et al., 1988. Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. J Cell Biol. 106**,** 761-71.

Coquette, A., et al., 1999. Differential expression and release of cytokines by an in vitro reconstructed human epidermis following exposure to skin irritant and sensitizing chemicals. Toxicol In Vitro. 13**,** 867-77.

Corazza M, Lauriola MM, Bianchi A, Zappaterra M, Virgili A. 2010. Irritant and sensitizing potential of eight surfactants commonly used in skin cleansers: an evaluation of 105 patients. Dermatitis. 21):262-8.

Cumberbatch, M., et al., 2002. Differential regulation of epidermal langerhans cell migration by interleukins (IL)-1alpha and IL-1beta during irritant- and allergen-induced cutaneous immune responses. Toxicol Appl Pharmacol. 182**,** 126-35.

Dean, J. H., et al., 2001. ICCVAM evaluation of the murine local lymph node assay. Conclusions and recommendations of an independent scientific peer review panel. Regul Toxicol Pharmacol. 34**,** 258-73.

Emter, R., et al., 2010. Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers in vitro. Toxicol Appl Pharmacol. 245**,** 281-290.

Garcia, C., et al., 2010. Comparative testing for the identification of skin-sensitizing potentials of nonionic sugar lipid surfactants. Regul Toxicol Pharmacol.

Gerberick, F., et al., 2008. Chemical reactivity measurement and the predicitve identification of skin sensitisers. The report and recommendations of ECVAM Workshop 64. Altern Lab Anim. 36**,** 215-42.

Gerberick, G. F., et al., 2002. Use of a B cell marker (B220) to discriminate between allergens and irritants in the local lymph node assay. Toxicol Sci. 68**,** 420-8.

Jírová D, Basketter D, Liebsch M, Bendová H, Kejlová K, Marriott M, Kandárová H. 2010. Comparison of human skin irritation patch test data with in vitro skin irritation assays and animal data. Contact Dermatitis. 62:109-16

Jacobs, J. J., et al., 2006. Skin irritants and contact sensitizers induce Langerhans cell migration and maturation at irritant concentration. Exp Dermatol. 15**,** 432-40.

Jowsey, I., Basketter, D.A., Westmoreland, C. & Kimber, I. (2006). A future approach to measuring relative skin sensitising potency. Journal of Applied Toxicology 26, 341 - 350.

Kimber, I., et al., 2001. Skin sensitization testing in potency and risk assessment. Toxicol Sci. 59**,** 198-208.

Kimber, I., et al., 2008. Dose metrics in the acquisition of skin sensitization: thresholds and importance of dose per unit area. Regul Toxicol Pharmacol. 52**,** 39-45.

Kirk, M.K., Broomhead Y, DiDonato L., DeGeorge, G.L. 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

Kreiling, R., et al., 2008. Comparison of the skin sensitizing potential of unsaturated compounds as 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., 1969. The identification of contact allergens by animal assay. The guinea pig maximization test. J Invest Dermatol. 52**,** 268-76.

Mehling A., et al., 2008. Discordant Results Between the LLNA and Guinea Pig Tests: Specificity vs. Selectivity. The Toxicologist (Supplement to Toxicological Sciences. 102**,** 8.

Natsch, A., Gfeller, H., 2008. LC-MS-based characterization of the peptide reactivity of chemicals to improve the in vitro prediction of the skin sensitization potential. Toxicol Sci. 106**,** 464-78.

Natsch, A., 2010. The Nrf2-Keap1-ARE toxicity pathway as a cellular sensor for skin sensitizers--functional relevance and a hypothesis on innate reactions to skin sensitizers. Toxicol Sci 113, 284-92.

Natsch, A., Gfeller, H., Emter, R.Ellis, G., 2010. Use of in vitro testing to identify an unexpected skin sensitizing impurity in a commercial product: a case study. Toxicol In Vitro 24, 411-6.

OECD, OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Skin Sensitization:. In: OECD/OCDE, (Ed.), 406, Paris, 1992.

OECD, OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Skin Sensitization: Local Lymph Node Assay. In: OECD/OCDE, (Ed.), 429, Paris, 2010.

Penninks, A. H., Limitations of the LLNA as preferred test for skin sensitization: concerns about false positive and false negative test results. . Vol. V 7217. TNO, Zeist, 2006.

Sakaguchi, H., et al., 2006. Development of an in vitro skin sensitization test using human cell lines; human Cell Line Activation Test (h-CLAT). II. An inter-laboratory study of the h-CLAT. Toxicol In Vitro. 20**,** 774-84.

Sakaguchi H, Ryan C, Ovigne JM, Schroeder KR, Ashikaga T. 2010 Predicting skin sensitization potential and inter-laboratory reproducibility of a human Cell Line Activation Test (h-CLAT) in the European Cosmetics Association (COLIPA) ring trials. Toxicol In Vitro. 24:1810-20..

Schnuch A, Lessmann H, Geier J, Uter W. 2011. Is cocamidopropyl betaine a contact allergen? Analysis of network data and short review of the literature. Contact Dermatitis. 64:203-11.

Wang, B., et al., 1999. Role of cytokines in epidermal Langerhans cell migration. J Leukoc Biol. 66**,** 33-9.

Woolhiser M.R., et al., 1998. A combined murine local lymph node and irritancy assay to predict sensitization and irritancy potential of chemicals. Toxicol. Meth. . 8**,** 245-256.

Wright ZM, Basketter PA, Blaikie L, Cooper KJ, Warbrick EV, Dearman RJ, Kimber I. 2001. Vehicle effects on skin sensitizing potency of four chemicals: assessment using the local lymph node assay. Int J Cosmet Sci. 23:75-83.

Table 1: Results of the *in vivo* assays: LLNA (OECD 429) and GPMT (OECD 406)

Substance (purity)	**LLNA**				**GPMT**		**LLNA vs. GPMT**
	Test Conc. [% in vehicle]	**Result DPM [SI] [EC3]**	**B220 [SI]**	**Ear thickness Day 5 [SI]**	**Intradermal Epidermal Challenge (rechallenge) [% in vehicle]**	**Number of animals with positive reactions**	**EC3 vs. epidermal induction [or challenge] in % concentration applied***
SLS (≥98%)	3 10 30 DMSO	1.36 3.09 2.73 EC3: 9.6	0.88 1.54 1.26	1.01 1.01 1.48	15 0.1 0.05 Corn oil	0/20	Appr. 0.1 vs. 15 [0.05]
C16EO2 (79%)	2.5 5 10 MEK	5.3 7.36 16.89 EC3: 2.2	1.66 2.03 2.80	1.00 1.45 1.49	10 50 10 (1% r.c.) water	Challenge 10% 15/20 controls 7/10 rechallenge 1% 0/20	2.2 vs. 50 [10 and 1]
C14EO4 (>99%)	2.5 5 10 MEK	2.09 9.45 8.54 EC3: 2.8	0.87 2.28 1.57	1 1.03 1.67	5 25 10 water	0/20	2.8 vs. 25 [10]
C12EO6 (≥98%)	5 10 25 MEK	2.19 3.85 10.28 EC3: 7.4	1.09 2.03 4.14	1.0 1.45 1.50	0.1 25 15 water	0/20	7.4 vs. 25 [15]
Isononyl glucoside (47%)	3 10 30 MEK	1.23 1.44 3.42 EC3: 25.8	0.62 0.61 0.63	1.0 1.0 1.47	15 100 100 water	Challenge 100% 17/20 controls 2/10 rechallenge 100% 18/20 50% 10/20	100 vs. 100 [100]
Thioglucopyranoside (>99%)	5 10 25 MEK	1.46 3.79 4.81 EC3: 8.3	0.53 1.02 1.25	1 1.18 1.46	10 50 10 PEG	0/20	8.3 vs. 50 [10]
Nonane sulfonate (≥97%)	3 10 40 MEK	1.06 1.51 1.32 EC3: n.a.	0.86 0.79 0.55	1 1 1.01	0.1 10 1 water	0/20	40 vs. 10 [1]
Decylphenol ethoxylate (>97%)	5 10 50 MEK	4.40 5.45 25.37 EC3: 2.9	2.18 1.68 2.88	1.01 1.01 1.01	1 100 100 PEG	0/20	2.9 vs. 100 [100]

* This column has been added to allow a quick comparison of the concentration calculated to elicit positive responses in the LLNA (EC3) and the concentration used for induction/challenge in the GPMT

Table 2: Results of the peptide reactivity and KeratinoSens assays

Substance (purity)	**Peptide reactivity assay [% Peptide depletion based on LC-MS]***	**KeratinoSens: Cytotoxicity [IC50]**	**KeratinoSens: ARE induction**	**Conclusions**
SLS (≥98%)	No peptide depletion	50.1 µM	Imax = 4.0 EC 1.5 = 35.4	ARE-dependent luciferase-induction in 3 of 5 experiments, but **only at cytotoxic concentrations**. High cytotoxicity indicates irritation potential
C16EO2 (79%)	2.6% peptide depletion, no adduct formation	20.5 µM	Imax = 1.3 EC 1.5 = no induction	No ARE-dependent luciferase-induction. High cytotoxicity indicates irritation potential
C14EO4 (>99%)	No peptide depletion	12.2 µM	Imax = 1.59 EC 1.5 = no induction	ARE-dependent luciferase-induction in 1 of 5 experiments, and **only at cytotoxic concentrations**, very high cytotoxicity indicates irritation potential
C12EO6 (≥98%)	32% peptide depletion due to peptide dimerization only, no adduct formation	22.3 µM	Imax = 1.67 EC 1.5 = 11.8 µM	ARE-dependent luciferase-induction in 3 of 5 experiments**, but only at cytotoxic concentrations**. High cytotoxicity indicates irritation potential
Isononyl glucoside (47%)	7.5 % peptide depletion, no adduct formation	1279 µM	Imax = 1.49 EC 1.5 = 785.3	ARE-dependent luciferase-induction in 6 of 12 experiments at non-cytotoxic concentration, borderline result. Low cytotoxicity.
Thioglucopyranoside (>99%)	No peptide depletion	> 2000 µM	Imax = 1.26, EC 1.5 = no induction	ARE-dependent luciferase-induction in 1 of 5 experiments, **but only at cytotoxic concentrations** and > 1000µM, Low cytotoxicity. No indication for sensitization or irritation
Nonane sulfonate (≥97%)	No peptide depletion	> 2000 µM	Imax = 0.94 EC 1.5 = no induction	No induction of ARE-dependent luciferase activity. Low cytotoxicity. No indication for sensitization or irritation
Decylphenol ethoxylate (>97%)	3% peptide depletion, no adduct formation	11.4 µM	Imax = 1.42, EC 1.5 = no induction	No ARE-dependent luciferase-induction, very high cytotoxicity indicates irritation potential

* Peptide depletion in this assay is not indicative of sensitization if due to dimerization or if no adducts are formed.

Table 3: Irritation potentials found in the various test methods (bold type indicates an irritatnt)

Substance (purity)	Concen-tration tested	Episkin: MTT Viability[(1)]	Classification based on MTT	Episkin: IL1-α secretion	Classification based on IL-1 α	Cytotoxicity KeratinoSens	Irritation seen in LLNA (concentration)*
SLS (≥98%)	1.56 3.125 6.25	60.1 ± 11.1 26.9 ± 2.4 4.0 ± 0.7	**Irritant**	332.6 ± 54.4 569.1 ± 47.1 371.8 ± 31.8	**Irritant**	High: **possible irritant**	**30%**
C16EO2 (79%)	12.5 25 50 100	91.0 ± 0.3 94.5 ± 3.2 95.4 ± 2.8 86.1 ± 3.9	Non irritant	121.0 ± 7.0 222.1 ± 23.1 175.5 ± 37.1 215.7 ± 96.3	**Irritant**	High: **possible irritant**	**5%**
C14EO4 (>99%)	12.5 25 50 100	89.9 ± 2.4 87.3 ± 4.0 77.6 ± 7.4 65.6 ± 6.8	Non irritant	291.3 ± 13.0 292.3 ± 63.3 363.5 ± 10.6 309.7 ± 7.2	**Irritant**	Very high: **possible irritant**	**10%**
C12EO6 (≥98%)	12.5 25 50 100	65.4 ± 5.0 59.6 ± 0.4 31.50 ± 9.4 9.2 ± 1.6	**Irritant**	305.8 ± 28.3 310.5 ± 104.8 480.8 ± 39.7 685.7 ± 41.8	**Irritant**	High: **possible irritant**	**10%**
Isononyl glucoside (47%)	25 50 100	102.5 ± 1.5 97.1 ± 0.2 47.9 ± 8.2	**Irritant**	80.6 ± 25.1 99.5 ± 37.5 358.5 ± 75.0	**Irritant**	Low	**30%**
Thioglucopyranoside (>99%)	6.25 12.5 25	88.8 ± 3.0 43.4 ± 22.4 17.2 ± 3.1	**Irritant**	157.6 ± 77.7 307.4 ± 39.0 285.9 ± 9.1	**Irritant**	Low	**25%**
Nonane sulfonate (≥97%)	6.25 12.5 25	60.6 ± 14.6 29.0 ± 10.3 6.1 ± 2.3	**Irritant**	100.1 ± 12.2 63.4 ± 22.8 53.8 ± 19.3	Nonirritant	Low	Nonirritant
Decylphenol ethoxylate (>97%)	12.5 25 50 100	62.6 ± 5.8 52.1 ± 4.3 9.1 ± 0.6 8.7 ± 0.5	**Irritant**	394.1 ± 36.4 515.3 ± 22.5 804.4 ± 84.5 767.9 ± 155.8	**Irritant**	Very high: **possible irritant**	Nonirritant
Dipropyleneglycol (DPG, vehicle)	100	100 ± 2.8	Non irritant	37.3 ± 13.4	Non irritant		
PBS [(1)]	100	111.9 ± 8.8	Non irritant	12.2 ± 9.3	Nonirritant		

* Irritation *in vivo*: LLNA: ear thickness >1.25

[(1)] The values for DPG were used as the 100% control, as this vehicle was used in most of the test, hence PBS is not 100% but relative to DPG

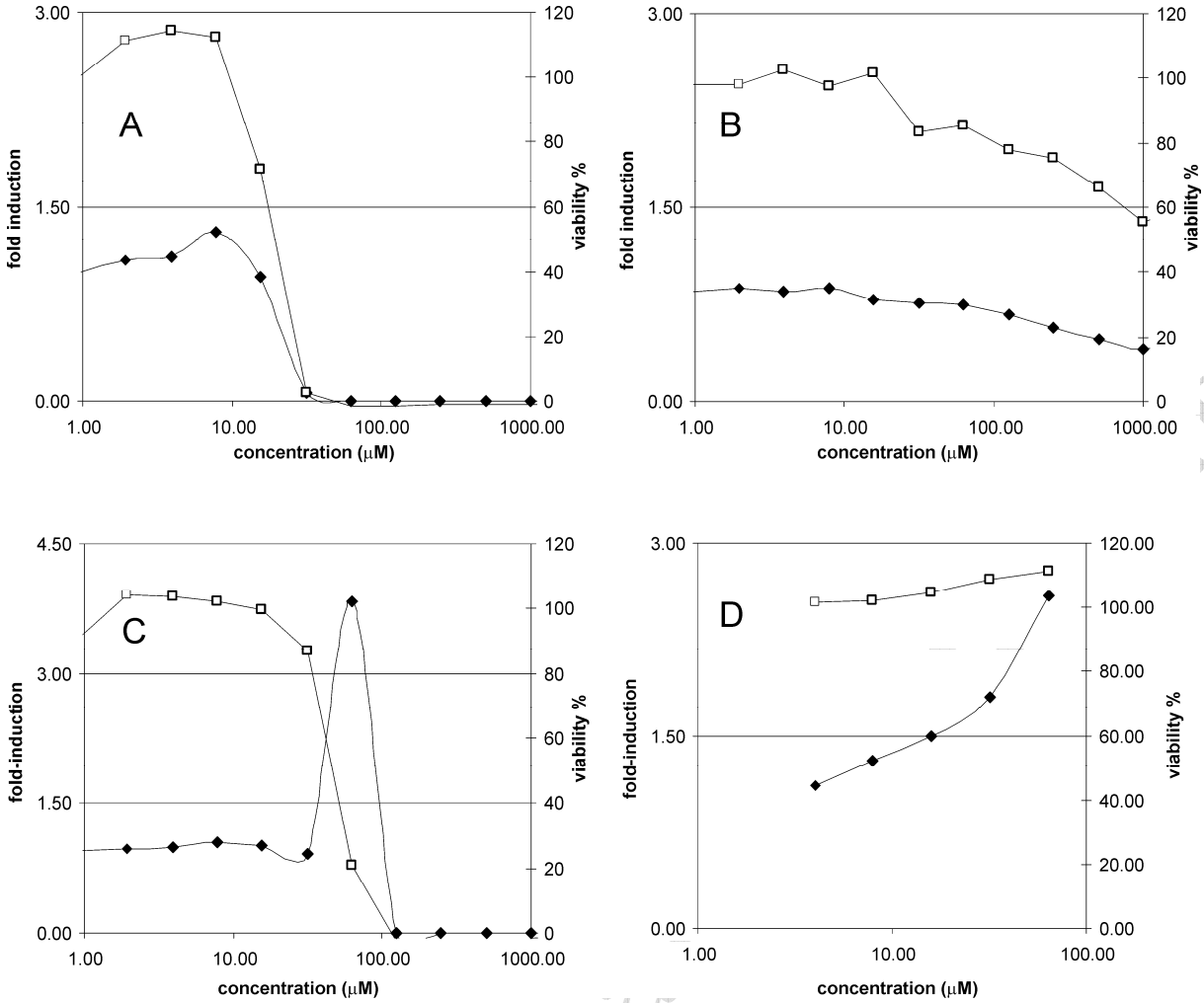

Figure 1. KeratinoSens assay: Induction of luciferase activity (closed diamonds) and cellular viability (open squares) for (A) C162EO (B) nonane sulfonate (C) SLS and (D) the positive control cinnamic aldehyde included in all assay plates. Note the gene induction at non-cytotoxic concentrations for cinnamic aldehyde and the induction at a cytotoxic concentration only for SLS.

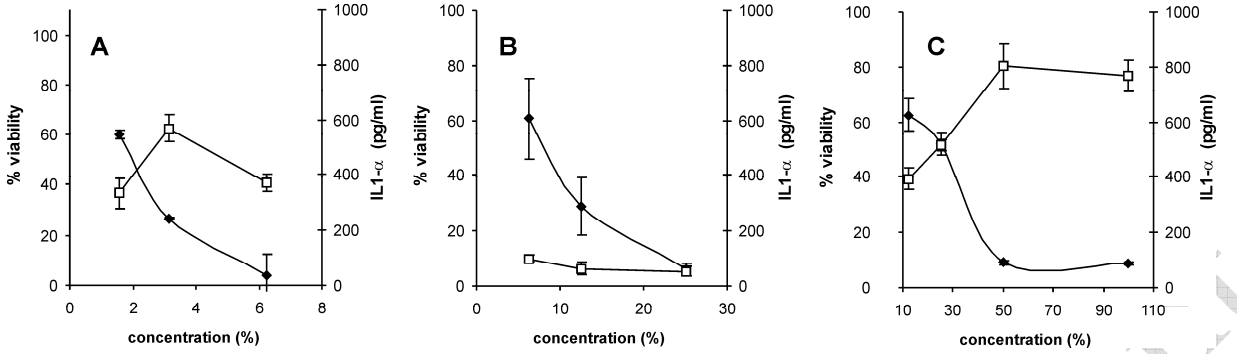

Figure 2. EpiSkin Assay: Cellular viability (closed diamonds) and IL-1α release (open squares) for (A) SLS (B) nonane sulfonate and (C) decylphenol ethoxylate. The y axis on the left indicates % viability, the axis on the right IL-1α release. Note the very pronounced IL-1α release for decylphenol ethoxylated, which had the highest SI in the LLNA and the low IL-1α release for nonane sulfonate which was negative in the LLNA.

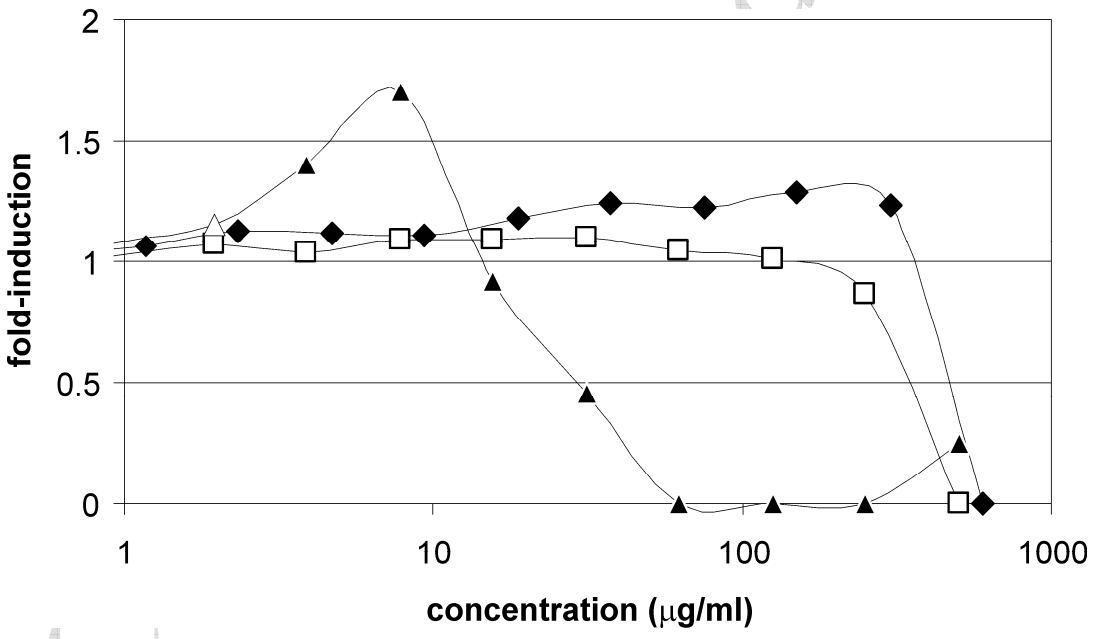

Figure 3. The KeratinoSens result for isononyl glucoside. 400 mg of the preparation were extracted with hexane. The water phase after hexane extraction (open squares), the residue (3 mg) from the hexane phase (closed triangles) and the original isononyl glucoside preparation (filled diamonds) were compared in parallel in the KeratinoSens assay in 4 repetitions (each with three replicates). The fold increase in the induction of the Luciferase gene is depicted.

Supplementary figure:

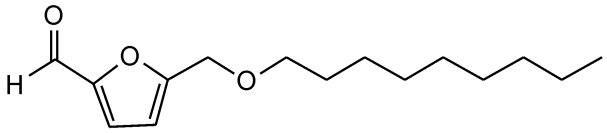

Figure S1: Identified impurity. This substance was identified as the key component of the hexane phase. Based on the results of the KeratinoSens assays (figure 4) it is likely to be the sensitizing impurity. This substance was identified as having structural alerts associated with Schiff base formation with aldehydes or as being a direct acting Schiff base former, although LC-MS analysis indicates peptide-adduct formation by an unknown mechanism.