In-vitro Skin sensitisation testing – What are the benefits/limitations

Presentation to NVT Meeting on10th October 2017

David Hart, Senior Toxicologist, Akzo Nobel Specialty Chemicals

AkzoNobel

Background – Test methods

The pressure caused by the animal testing ban for cosmetic ingredients and the need to reduce animal usage for REACH has resulted in the development and approval of currently 3 approved (OECD guideline) in-chemico/in-vitro tests for skin sensitisation.

These test have been developed against the background of the development of an AOP for skin sensitisation.

Our experience so far has been with the three methods with OECD Guidelines.

These test method represent the three Key Events (KE) from the AOP after absorption.

KE1 (MIE Molecular Initiating Event) is the binding of the chemical to proteins in the skin, such that it will be identified as foreign by the immune system. The DPRA (Direct Peptide Reactivity Assay) OECD TG 442C. There are two predilection models which can be used depending on the peptide depletion seen in the cysteine and lysine peptides.

KE2 Activation of inflammatory cytokines in the Keratinocytes. This is identified using the KeratinoSens® OECD TG 442D or LuSens (under validation assessment) assays. These monitor the activation of the Keap-Nrf2-ARE pathway, linked to a gene for luciferase so that the degree of activation can be measured by the increase in fluorescence.

Background – Test methods

KE3 The activation of the dendritic cells, which would then migrate to the draining lymph node to activate the T cells. This is identified using the h-CLAT human Cell Line Activation Test OECD TG 442E. Also in the process of gaining regulatory acceptance is the U-SENS assay developed from the MUSST assay, for which there is draft OECD TG in-vitro Skin Sensitisation: U937 Cell Line Activation test (U-SENS™)

In the h-CLAT the activity is monitored by measuring the increase in the CD54 and CD86 cell surface markers expression while the U-SENS only monitors expression of the CD86 cell surface marker. In both case the expression is by flow cytometry following cell staining with labelled antibodies.

All these methods have some limitation as there is no metabolic activity in the DPRA and little or no metabolic activity in the cell based KeratinoSens® and h-CLAT assays.

ECHA has produced updated guidelines in July 2017 (REACH Guidance chapter 7a). This was designed to allow the use of these alternative methods to avoid where possible the use of animals in the LLNA (Local Lymph Node Assay).

ECHA Guidance - Limitations

In the ECHA webinar earlier this year, it was clear that to get a clear negative result you needed three negative tests and QSAR indicating a lack of any potential for skin sensitisation.

They leave the decision on classification based on in-vitro data to the registrant, they do indicate that the DPRA could possibly alone be sufficient for a substance to be considered a skin sensitiser, as protein binding in the MIE.

They suggest a weight of evidence approach, but it may be difficult to not classify if more than one test is positive and difficult to consider a result as negative if all tests are not negative.

The ECHA guidelines make it clear that the only way you can conclude that a substance is category 1 is based on existing animal test data from before 10 May 2017 which does not allow the use of the CLP (GHS) criteria to differentiate if it is 1A a potent skin sensitiser or 1B a less potent skin sensitiser

ECHA Guidance - Limitations

It is now a specific requirement for classifications based on new data to specify if it is category 1A or 1B.

The ECHA guidelines further indicate that should you conclude based on in-vitro and QSAR data via a weight of evidence that a substance is a skin sensitiser then the default classification would be category 1A. It is possible to try to justify a 1B based on such data as the reactivity in the DPRA but this may not be accepted.

While the guidelines state that testing in animals must only be last resort, where the results based on in-vitro testing are equivocal an LLNA seems to be the only way to get a definitive results.

If you conclude that an LLNA is required, I recently noticed a limitation in the guidelines concerning the protocol to be used

ECHA Guidance - Limitations

I received a quotation for an LLNA from a US lab and found it was for OECD442B BrdU-ELISA none radioactive method. Initially I assumed this would be OK until I saw the following the ECHA Guidance.

f) Note: for the LLNA variants, i.e. EU B.50/OECD TG 442A and B.51/442B, there are currently no CLP criteria available for predicting skin sensitisation potency. However, dose-response relationship information may provide some information on skin sensitisation potency that can be used within a Weight-of-Evidence approach. **It is recommended that, when new in vivo data are generated, the "standard" LLNA according to EU B.42 / OECD TG 429 be used**, if possible.

This indicates that for REACH the OECD TG 429 original test is advisable.

Akzo Nobel's Experience so far with these test methods

AkzoNobel

Now to discuss our experience with these tests so far.

We have had some problems with the DPRA for several surfactants of two different chemical classes, for these we saw some shifting of the HPLC peaks for the lysine peptide. This was not interpreted as depletion and could have been an effect on polarity or on the column. Due to this the Cysteine prediction model was used and all were negative.

For another substance we had an apparent depletion of 84.9% for the lysine peptide. A second peak was seen eluting at 13 minutes which was identical to the lysine peak area of the control.

Further investigation of a retained sample confirmed it was a false positive result due to the surfactant affecting the HPLC column causing ion-pairing effects between the lysine peptide and the HPLC column. Therefore the lysine result was discarded and the cysteine result compared to the cysteine only model and a negative result was obtained.

The use of HPLC to monitor for peptide depletion is a limitation, as depletion can be caused by dimerisation of the peptide. Some authors use GC/MS methods for screening purposes as this allows you to confirm if there are adducts formed between the peptide and the test substance, avoiding false positives.

Akzo Nobel's Experience so far with these test methods

AkzoNobel

The same surfactant where we had the high false depletion result for the lysine peptide, was negative in the KeratinoSens assay but positive in the h-CLAT.

As the surfactant was neutralised with an amine which is classified as a skin sensitiser, this result was considered to require a confirmatory animal test.

Due to experience with false positive results from the LLNA for such irritant surfactants, which is recognised in the LLNA OECD guideline 429, a Guinea Pig Maximisation test (GPMT) was done. This test showed no positive reactions in any of the treated guinea pigs confirming that it was not a skin sensitiser.

The DPRA can be influenced by the solvent, I have an example which was weakly positive in the range that the OECD TG states needs to be repeated. This was clearly negative when tested in a much more experience laboratory. Investigation found the first used acetonitrile as the solvent while the second used water as the substance was highly water soluble.

I have another example of a cosmetic ingredient being registered for REACH. For this again the DPRA and KeratinoSens® assay were negative.

Further observations from our experience

The h-CLAT however was positive, but showed very variable results between the three runs in which the second showed no increase for either CD54 or CD86. The first and third showed a small increase in CD54 at the top concentration, but four concentrations were over the 150% threshold for CD86 in the third run but well under 150% in the other two runs. The two runs with increased CD54 makes a positive result.

I will show you the data from the three runs to illustrate the issue.

The data from this study clearly shows that even in a lab with a lot of experience of the test, how variable the h-CLAT results can be.

Further observations from our experience – h-CLAT Experiment 1

AkzoNobel

Sample	Conc. [µg/mL]	Cell Viability [%]			Mean Fluorescence Intensity			corrected Mean Fluorescence Intensity		Relative Flourescence Intensity (RFI)		Ratio Isotype IgG1 to [%]	
		CD86	CD54	Isotype IgG1	CD86	CD54	Isotype IgG1	CD86	CD54	CD86	CD54	CD86	CD54
Medium Control	-	96.7	97.1	97.4	1452	965	824	628	141	**100**	**100**	176	117
DMSO Control	0.20%	96.3	96.4	95.7	1659	1032	809	850	223	**135**	**158**	205	128
DNCB	4.00	86.0	86.7	86.4	3400	1857	831	2569	1026	**302**	**460**	409	223
N-[2-(2- ethyl)acetamide	5000.00	96.1	96.3	96.5	1344	1012	656	688	356	**110**	**252**	205	154
	4166.67	96.1	96.3	96.3	1615	1174	769	846	405	**135**	**287**	210	153
	3472.22	95.6	95.0	95.9	1560	964	690	870	274	**139**	**194**	226	140
	2893.52	95.7	95.5	95.8	1464	916	647	817	269	**130**	**191**	226	142

Further observations from our experience – h-CLAT Experiment 2

AkzoNobel

Sample	Conc. [µg/mL]	Cell Viability [%]			Mean Fluorescence Intensity			corrected Mean Fluorescence Intensity		Relative Flourescence Intensity (RFI)		Ratio Isotype IgG1 to [%]	
		CD86	CD54	IgG Isotype	CD86	CD54	Isotype IgG1	CD86	CD54	CD86	CD54	C86	CD54
Medium Control	-	96.2	96.1	96.3	2007	1584	972	1035	612	**100**	**100**	206	163
DMSO Control	0.20%	95.9	96.3	95.7	2323	1606	954	1369	652	**132**	**107**	244	168
DNCB	4.0	88.3	88.7	88.7	6095	3811	955	5140	2856	**375**	**438**	638	399
N-[2-(2-Hy etl	5000.00	97.0	96.5	96.5	2025	1551	791	1234	760	**119**	**124**	256	196
	4166.67	96.1	95.7	95.7	1998	1534	829	1169	705	**113**	**115**	241	185
	3472.22	96.5	95.7	96.1	1972	1518	822	1150	696	**111**	**114**	240	185
	2893.52	96.1	96.5	96.3	2059	1427	852	1207	575	**117**	**94**	242	167
	2411.27	96.3	96.1	96.1	2062	1535	900	1162	635	**112**	**104**	229	171

Further observations from our experience – h-CLAT Experiment 3

AkzoNobel

Sample	Conc. [µg/mL]	Cell Viability [%]			Mean Fluorescence Intensity			corrected Mean Fluorescence Intensity		Relative Flourescence Intensity (RFI)		Ratio Isotype IgG1 to [%]	
		CD86	CD54	IgG Isotype	CD86	CD54	Isotype IgG1	CD86	CD54	CD86	CD54	C86	CD54
Medium Control	-	96.9	97.0	97.0	1314	1093	896	418	197	100	100	147	122
DMSO Control	0.20%	96.7	96.7	96.4	1351	1124	826	525	298	126	151	164	136
DNCB	4.0	86.1	87.4	86.3	2870	2341	789	2081	1552	**396**	**521**	364	297
N-[2-(2-	5000.0	96.3	96.4	96.4	1423	1187	726	697	461	**167**	**234**	196	163
	4166.67	96.4	96.4	96.7	1442	1126	756	686	370	**164**	188	191	149
	3472.22	96.1	96.3	96.4	1416	1117	761	655	356	**157**	181	186	147
	2893.52	96.3	95.8	96.2	1499	1104	795	704	309	**168**	157	189	139
	2411.27	96.8	96.8	96.7	1413	1100	812	601	288	144	146	174	135

Further observations from our experience

I find it hard to have much faith in an assay which gives such variability between runs.

So how to interpret these results, the h-CLAT is positive despite the variability making me question it.

How to interpret the overall weight of evidence?

QSAR does not indicate a skin sensitiser but does indicate some potential to metabolise to have some protein reactivity

One option is to use the approach developed by BASF as published by Bausch et.al. (2012) , with a prediction model where you need at least 2 out of three of the studies positive to consider the test substance to be a skin sensitiser.

Using the 2 out of 3 PM I can conclude that it is not a skin sensitiser, but would ECHA agree. A more scientific weight of evidence argument may be more successful?

Akzo Nobel's Experience in an EFfCI research project

AkzoNobel

I have been part of a working group in EFfCI (The European Cosmetic Ingredient Manufactures Trade Association) where we tested 8 unsaturated and 1 saturated lipophilic compounds which had previously been used to compare the LLNA with the GPMT.

The reason for this testing was a concern that the new in-vitro test have been validated mainly against the LLNA, and these 9 compounds had previous shown 6 false positive results compared to the GPMT. The aim of this project was to see how the results of the three in-chemico/in-vitro tests would compare with the results from the LLNA and GPMT and how best to integrate the results from the new test methods.

As none had absolutely negative results in all three tests, with some being negative but inconclusive, the proposed prediction model (PM) of Bausch et.al. (2012) of requiring two or more out of three tests positive to give an overall skin sensitiser result was applied to try to minimise false positive responses.

The lack of any clearly recognised prediction model from ECHA guidelines limits the use of these tests.

Akzo Nobel's Experience so far with these test methods –EFfCI Project

AkzoNobel

	Oleic acid	Linoleic acid	Linolenic acid	Undecylenic acid	Fumaric acid	Maleic acid	Succinic acid	Squalene	1-Octyn-3-ol
Overall test result in the LLNA [b]	Positive	Positive	Positive	Positive	Negative	Positive	Negative	Positive	Positive
Overall test result in the GPMT [b]	Negative	Negative	Negative	Positive	Negative	Negative	Negative	Negative	Negative
DPRA PM 2 (cys only)	[Negative (minimal)] [c]	Positive (low)	Positive (moderate)	[Negative (minimal)] [c]	Negative (minimal)	Positive (moderate)	Negative (minimal)	[Negative (minimal)] [c]	Positive (moderate)
KeratinoSens™ assay	Negative [d]	Positive	Positive	Positive	Negative	Negative	Negative	[Negative] [e]	Positive
h-CLAT	[Negative] [f]	Positive	Positive	Positive	Positive	Negative	Negative	[Negative] [f]	Positive
2-out-of-3 PM [a]	Non-sensitiser	Sensitiser	Sensitiser	Sensitiser	Non-sensitiser	Non-sensitiser	Non-sensitiser	[Non-sensitiser]	Sensitiser
Evaluation as compared to GPMT	True negative	False positive	False positive	True positive	True negative	True negative	True negative	True negative	False positive

Grey shade: Positive test result or positive outcome of 2-out-of-3 PM.

a: The 2-out-of-3 prediction model implies classification as skin sensitiser if a test substances yields positive results in 2 of the 3 assays (Bauch *et al.*, 2012).

b: As published by Kreiling *et al.* (2008)

c: Result inconclusive since precipitation or phase separation was observed prior to the HPLC analysis, so that reactivity might have been underestimated.

d: Even though oleic acid showed impaired solubility in cell culture medium and has a log K_{ow} >7, the negative test result is assessed as valid since oleic acid elicited cytotoxicity in the concordant MTT assay, thereby providing evidence that the test substance reached the cells within the incubation period.

e: Squalene showed impaired solubility in cell culture medium and has a log K_{ow} >7. Further, it did not elicit any effects in the MTT assay. Therefore, it cannot be ensured that squalene reached the cells within the incubation period, and the test result is assessed as inconclusive.

f: Log K_{ow} >3.5, i.e. in accordance with OECD TG 442E, this negative test result should not be considered to predict (non) skin sensitisation. Further, due to impaired solubility in the cell culture medium, it cannot be ensured that squalene reached the cells

Discussion of EFfCI project

The oleic acid, linoleic acid and squalene all have octanol water partition coefficients of just over 7 (over 14 for squalene) which is outside the applicability domain of the KeratinoSens® range of ≤7 and for the h-CLAT (LogKow >3.5 tend to false negative). Squalene was difficult in most systems due to it poor solubility in the available solvents. Despite the high Log Kow for the others, positive results were considered as acceptable as cytotoxicity was seen which confirm that the cells were exposed to the test substance.

As was shown in the table the LLNA gave false positive responses with six of the 9 substances, which using the 2 out of 3 PM only resulted in 3 out of 9 false positives with undecylenic acid as a true positive and the others as true negatives against the GPMT data.

Based on this experience we made the following observation of the three methods

The DPRA has simple methodology, has a large variety of solvent that can be used and give quick results.

However difficulty in getting solution of lipophilic substances, the lysine peptide was not stable in isopropanol one of the allowed solvents. The peptides vary in their stability and the acetonitrile has to be checked with each new batch, also the cysteine peptide was also not very stable

Discussion of EFfCI project

The KeratinoSens assay was simple and had some flexibility of solvents. However there is no dose ranging before the main study which can result in excessive cytotoxicity, solubility of very lipophilic substance can be problematic.

The h-CLAT does have dose ranging so excessive cytotoxicity can be avoided and there is some flexibility of solvents. However it was found to be the most complex and problematic assay. It is very time consuming as there has to be an activation test of every batch of cells including dose ranging and the main study.

Not possible to do multiple samples at the same time.

Test substance interference with Fluorescence-activated cell sorting measurement can occur

Solubility of very lipophilic substances problematic

Differences in target cell line between different suppliers observed.

Further observations from our experience

The paper describing this work has Dr Reinhard Kreiling as the main Author and has been published last week in Toxicology and Pharmacology the Title: In chemico, in vitro and in vivo comparison of the skin sensitizing potential of eight unsaturated and one saturated lipophilic compounds. Volume 90 (2017) Pages 262-276

https://www.ncbi.nlm.nih.gov/labs/articles/28958912/

Other Alternatives;

U-SENS is an alternative to the h-CLAT I have no experience of this but if more reliable than the h-CLAT it would be of interest.

For cosmetic ingredients to build a weight of evidence without using animal testing the SENS-IS test is looking promising (using Epiderm tissues). As the reconstructed skin has a strartum corneum, then for substances like our cosmetic ingredient with very low <1% dermal penetration in human skin in vitro, a more realistic exposure and assessment of the hazard could be obtained. So far we have no experience with this. The question is would a negative in this assay be sufficient to conclude an overall negative that would satisfy ECHA.

Conclusions

It is important that we get more experience with these new test methods in a wider variety of chemistries, so we can establish the applicability domain and identify where specific test methods are not suitable.

For the 2018 registrations there will be many smaller companies registering who do not have internal expertise, so doing weight of evidence scientific justification for classification decisions based on none animal data may be beyond them.

We urgently need some guidance from ECHA on how to integrate the results from these three test and conclude if a test substance should be considered a sensitiser and when it can be considered not to be a sensitiser, i.e. not requiring classification.

Sharing as much information as is possible on experience with these test will help all concerned.

I can take questions now.