

## **OECD TEST GUIDELINE FOR THE TESTING OF CHEMICALS BASED ON KEY EVENTS**

### **In Vitro Skin Sensitisation assays addressing the Key Event on activation of dendritic cells on the Adverse Outcome pathway for Skin Sensitisation**

#### GENERAL INTRODUCTION

##### **Activation of dendritic cells Key Event based Test Guideline**

1. A skin sensitiser refers to a substance that will lead to an allergic response following skin contact as defined by the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS) (1). There is general agreement on the key biological events underlying skin sensitisation. The current knowledge of the chemical and biological mechanisms associated with skin sensitisation has been summarised as an Adverse Outcome Pathway (AOP) (2), starting with the molecular initiating event through intermediate events to the adverse effect, namely allergic contact dermatitis. In this instance, the molecular initiating event (i.e. the first key event) is the covalent binding of electrophilic substances to nucleophilic centres in skin proteins. The second key event in this AOP takes place in the keratinocytes and includes inflammatory responses as well as changes in gene expression associated with specific cell signalling pathways such as the antioxidant/electrophile response element (ARE)-dependent pathways. The third key event is the activation of dendritic cells (DC), typically assessed by expression of specific cell surface markers, chemokines and cytokines. The fourth key event is T-cell activation and proliferation, which is indirectly assessed in the murine Local Lymph Node Assay (LLNA) (3).

2. This Test Guideline (TG) describes *in vitro* assays that address mechanisms described under the Key Event on activation of dendritic cells of the AOP for skin sensitisation (2). The TG comprises test methods to be used for supporting the discrimination between skin sensitisers and non-sensitisers in accordance with the UN GHS (1).

The test methods described in this TG are:

- Human Cell Line Activation test (h-CLAT)
- U937 cell line activation Test (U-SENS<sup>TM</sup>)

- Interleukin-8 Reporter Gene Assay (IL-8 Luc assay)

3. The test methods included in this Test Guideline may differ in relation to the procedure used to generate the data and the readouts measured but can be used indiscriminately to address countries' requirements for test results on the Key Event on activation of dendritic cells of the AOP for skin sensitisation while benefiting from the Mutual Acceptance of Data.

### **Background and principles of the test methods included in the Key Event based Test Guideline**

4. The assessment of skin sensitisation has typically involved the use of laboratory animals. The classical methods that use guinea-pigs, the Guinea Pig Maximisation Test (GPMT) of Magnusson and Kligman, and the Buehler Test (TG 406) (4), assess both the induction and elicitation phases of skin sensitisation. The murine tests, the LLNA (TG 429) (3) and its two non-radioactive modifications, LLNA: DA (TG 442 A) (5) and LLNA: BrdU-ELISA (TG 442 B) (6), all assess the induction response exclusively, and have also gained acceptance, since they provide an advantage over the guinea pig tests in terms of animal welfare together with an objective measurement of the induction phase of skin sensitisation.

5. Recently mechanistically-based *in chemico* and *in vitro* test methods addressing the first key event (OECD TG 442C; Direct Peptide Reactivity Assay (7)), and second key event (OECD TG 442D; ARE-Nrf2 Luciferase Test Method (8)) of the skin sensitisation AOP have been adopted for contributing to the evaluation of the skin sensitisation hazard potential of chemicals.

6. Test methods described in this TG either quantify the change in the expression of cell surface marker(s) associated with the process of activation of monocytes and DC following exposure to sensitisers (e.g. CD54, CD86) or the changes in IL-8 expression, a cytokine associated with the activation of DC. Skin sensitisers have been reported to induce the expression of cell membrane markers such as CD40, CD54, CD80, CD83, and CD86 in addition to induction of proinflammatory cytokines, such as IL-1 $\beta$  and TNF- $\alpha$ , and several chemokines including IL-8 (CXCL8) and CCL3 (9) (10) (11) (12), associated with DC activation (2).

7. However, as DC activation represents only one key event of the skin sensitisation AOP (2) (13), information generated with test methods measuring markers of DC activation alone may not be sufficient to conclude on the presence or absence of skin sensitisation potential of chemicals. Therefore data generated with the test methods described in this Test Guideline are proposed to support the discrimination between skin sensitisers (i.e. UN GHS Category 1) and non-sensitisers when used within Integrated Approaches to Testing and Assessment (IATA), together with other relevant complementary information, e.g. derived from *in vitro* assays addressing other key events of the skin sensitisation AOP as well as non-testing methods, including read-across from chemical analogues (13). Examples of the use of data generated with these methods within Defined Approaches, i.e. approaches standardised both in relation to the set of information sources used and in the procedure applied to the data to derive predictions, have been published (13) and can be employed as useful elements within IATA.

8. The test methods described in this Test Guideline cannot be used on their own, neither to sub-categorise skin sensitisers into subcategories 1A and 1B as defined by UN GHS (1), for authorities implementing these two optional subcategories, nor to predict potency for safety assessment decisions. However, depending on the regulatory framework, positive results generated with these methods may be used on

their own to classify a chemical into UN GHS category 1.

9. The term "test chemical" is used in this Test Guideline to refer to what is being tested<sup>1</sup> and is not related to the applicability of the test methods to the testing of mono-constituent substances, multi-constituent substances and/or mixtures. Limited information is currently available on the applicability of the test methods to multi-constituent substances/mixtures (14) (15). The test methods are nevertheless technically applicable to the testing of multi-constituent substances and mixtures. However, before use of this Test Guideline on a mixture for generating data for an intended regulatory purpose, it should be considered whether, and if so why, it may provide adequate results for that purpose<sup>2</sup>. Such considerations are not needed when there is a regulatory requirement for the testing of the mixture. Moreover, when testing multi-constituent substances or mixtures, consideration should be given to possible interference of cytotoxic constituents with the observed responses.

## LITERATURE

1. United Nations UN (2015), Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Sixth revised edition. New York & Geneva: United Nations Publications. ISBN: 978-92-1-117006-1. Available at: [https://www.unece.org/trans/danger/publi/ghs/ghs\\_rev06/06files\\_e.html](https://www.unece.org/trans/danger/publi/ghs/ghs_rev06/06files_e.html)].
2. OECD (2012), The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Series on Testing and Assessment No. 168. Available at: [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO\(2012\)10/PART1&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)10/PART1&docLanguage=En)
3. OECD (2010), The Local Lymph Node Assay. OECD Guideline for the Testing of Chemicals No. 429, Organisation for Economic Cooperation and Development, Paris. Available at: <http://www.oecd.org/env/testguidelines>
4. OECD (1992), Skin Sensitisation. OECD Guideline for the Testing of Chemicals No. 406, Organisation for Economic Cooperation and Development, Paris. Available at: <http://www.oecd.org/env/testguidelines>
5. OECD (2010), Skin Sensitisation: Local Lymph Node Assay: DA. OECD Guideline for the Testing of Chemicals No. 442A, Organisation for Economic Cooperation and Development, Paris. Available at: <http://www.oecd.org/env/testguidelines>
6. OECD (2010), Skin sensitisation: Local Lymph Node Assay: BrdU-ELISA. OECD Guideline for the Testing of Chemicals No. 442B, Organisation for Economic Cooperation and Development, Paris. Available at: <http://www.oecd.org/env/testguidelines>
7. OECD (2015), OECD Guideline for the Testing of Chemicals No. 442C: In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA). Paris, France: Organisation for Economic

---

<sup>1</sup> In June 2013, the Joint Meeting agreed that where possible, a more consistent use of the term "test chemical" describing what is being tested should be applied in new and updated Test Guidelines.

<sup>2</sup> This sentence was proposed and agreed at the April 2014 WNT meeting

Cooperation and Development. Available at: <http://www.oecd.org/env/testguidelines>

8. OECD (2015), OECD Guideline for the Testing of Chemicals No. 442D: In Vitro Skin Sensitisation: ARE-Nrf2 Luciferase Test Method. Paris, France: Organisation for Economic Cooperation and Development. Available at: <http://www.oecd.org/env/testguidelines>
9. Steinman RM. 1991. The dendritic cell system and its role in immunogenicity. *Annu Rev Immunol* 9:271-96.
10. Caux C, Vanbervliet B, Massacrier C, Azuma M, Okumura K, Lanier LL, and Banchereau J. 1994. B70/B7-2 is identical to CD86 and is the major functional ligand for CD28 expressed on human dendritic cells. *J Exp Med* 180:1841-7.
11. Aiba S, Terunuma A, Manome H, and Tagami H. 1997. Dendritic cells differently respond to haptens and irritants by their production of cytokines and expression of co-stimulatory molecules. *Eur J Immunol* 27:3031-8.
12. Aiba S, Manome H, Nakagawa S, Mollah ZU, Mizuashi M, Ohtani T, Yoshino Y, and Tagami. H. 2003. p38 mitogen-activated protein kinase and extracellular signal-regulated kinases play distinct roles in the activation of dendritic cells by two representative haptens, NiCl<sub>2</sub> and DNCB. *J Invest Dermatol* 120:390-8.
13. OECD (2016). Series on Testing & Assessment No. 256: Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated Approaches To Testing And Assessment (IATA) For Skin Sensitisation, Annex 1 and Annex 2. [ENV/JM/HA\(2016\)29](#). Organisation for Economic Cooperation and Development, Paris. Available at: [<https://community.oecd.org/community/iatass>].
14. Ashikaga T, Sakaguchi H, Sono S, Kosaka N, Ishikawa M, Nukada Y, Miyazawa M, Ito Y, Nishiyama N, Itagaki H. (2010), A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). *Altern. Lab. Anim.* 38, 275-284.
15. Piroird, C., Ovigne, J.M., Rousset, F., Martinozzi-Teissier, S., Gomes, C., Cotovio, J., Alépée, N. (2015). The Myeloid U937 Skin Sensitization Test (U-SENS) addresses the activation of dendritic cell event in the adverse outcome pathway for skin sensitization. *Toxicol. In Vitro* 29, 901-916.

## **Annex I: *In Vitro* Skin Sensitisation: human Cell Line Activation** **Test (h-CLAT)**

### **INITIAL CONSIDERATIONS AND LIMITATIONS**

1. The h-CLAT method quantifies changes in the expression of cell surface markers associated with the process of activation of monocytes and dendritic cells (DC) (i.e. CD86 and CD54), in the human monocytic leukaemia cell line THP-1, following exposure to sensitisers (1) (2). The measured expression levels of CD86 and CD54 cell surface markers are then used for supporting the discrimination between skin sensitisers and non-sensitisers.

2. The h-CLAT method has been evaluated in a European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM)-coordinated validation study and subsequent independent peer review by the EURL ECVAM Scientific Advisory Committee (ESAC). Considering all available evidence and input from regulators and stakeholders, the h-CLAT was recommended by EURL ECVAM (3) to be used as part of an IATA to support the discrimination between sensitisers and non-sensitisers for the purpose of hazard classification and labelling. Examples of the use of h-CLAT data in combination with other information are reported in the literature (4) (5) (6) (7) (8) (9) (10) (11).

3. The h-CLAT method proved to be transferable to laboratories experienced in cell culture techniques and flow cytometry analysis. The level of reproducibility in predictions that can be expected from the test method is in the order of 80% within and between laboratories (3) (12). Results generated in the validation study (13) and other published studies (14) overall indicate that, compared with LLNA results, the accuracy in distinguishing skin sensitisers (i.e. UN GHS Cat.1) from non-sensitisers is 85% (N=142) with a sensitivity of 93% (94/101) and a specificity of 66% (27/41) (based on a re-analysis by EURL ECVAM (12) considering all existing data and not considering negative results for chemicals with a Log Kow greater than 3.5 as described in paragraph 4). False negative predictions with the h-CLAT are more likely to concern chemicals showing a low to moderate skin sensitisation potency (i.e. UN GHS subcategory 1B) than chemicals showing a high skin sensitisation potency (i.e. UN GHS subcategory 1A) (4) (13) (15). Taken together, this information indicates the usefulness of the h-CLAT method to contribute to the identification of skin sensitisation hazards. However, the accuracy values given here for the h-CLAT as a stand-alone test method are only indicative, since the test method should be considered in combination with other sources of information in the context of an IATA and in accordance with the provisions of paragraphs 7 and 8 in the General introduction. Furthermore, when evaluating non-animal methods for skin sensitisation, it should be kept in mind that the LLNA test as well as other animal tests may not fully reflect the situation in humans.

4. On the basis of the data currently available, the h-CLAT method was shown to be applicable to test chemicals covering a variety of organic functional groups, reaction mechanisms, skin sensitisation potency (as determined in *in vivo* studies) and physicochemical properties (3) (14) (15). The h-CLAT method is applicable to test chemicals soluble or that form a stable dispersion (i.e. a colloid or suspension in which the test chemical does not settle or separate from the solvent/vehicle into different phases) in an appropriate solvent/vehicle (see paragraph 14). Test chemicals with a Log Kow greater than 3.5 tend to produce false negative results (14). Therefore negative results with test chemicals with a Log Kow greater than 3.5 should not be considered. However, positive results obtained with test chemicals with a Log Kow greater than 3.5 could still be used to support the identification of the test chemical as a skin sensitiser.

Furthermore, because of the limited metabolic capability of the cell line used (16) and because of the experimental conditions, pro-haptens (i.e. substances requiring enzymatic activation for example via P450 enzymes) and pre-haptens (i.e. substances activated by oxidation) in particular with a slow oxidation rate may also provide negative results in the h-CLAT (15). Fluorescent test chemicals can be assessed with the h-CLAT (17), nevertheless, strong fluorescent test chemicals emitting at the same wavelength as fluorescein isothiocyanate (FITC) or as propidium iodide (PI), will interfere with the flow cytometric detection and thus cannot be correctly evaluated using FITC-conjugated antibodies or PI. In such a case, other fluorochrome-tagged antibodies or other cytotoxicity markers, respectively, can be used as long as it can be shown they provide similar results as the FITC-tagged antibodies (see paragraph 24) or PI (see paragraph 18) e.g. by testing the proficiency substances in Appendix II. In the light of the above, negative results should be interpreted in the context of the stated limitations and together with other information sources within the framework of IATA. In cases where there is evidence demonstrating the non-applicability of the h-CLAT method to other specific categories of test chemicals, it should not be used for those specific categories.

5. As described above, the h-CLAT method supports the discrimination between skin sensitisers from non-sensitisers. However, it may also potentially contribute to the assessment of sensitising potency (4) (5) (9) when used in integrated approaches such as IATA. Nevertheless, further work, preferably based on human data, is required to determine how h-CLAT results may possibly inform potency assessment.

6. Definitions are provided in Appendix I.

#### PRINCIPLE OF THE TEST

7. The h-CLAT method is an *in vitro* assay that quantifies changes of cell surface marker expression (i.e. CD86 and CD54) on a human monocytic leukemia cell line, THP-1 cells, following 24 hours exposure to the test chemical. These surface molecules are typical markers of monocytic THP-1 activation and may mimic DC activation, which plays a critical role in T-cell priming. The changes of surface marker expression are measured by flow cytometry following cell staining with fluorochrome-tagged antibodies. Cytotoxicity measurement is also conducted concurrently to assess whether upregulation of surface marker expression occurs at sub-cytotoxic concentrations. The relative fluorescence intensity of surface markers compared to solvent/vehicle control are calculated and used in the prediction model (see paragraph 26), to support the discrimination between sensitisers and non-sensitisers

#### DEMONSTRATION OF PROFICIENCY

8. Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency, using the 10 Proficiency Substances listed in Appendix II. Moreover, test method users should maintain an historical database of data generated with the reactivity checks (see paragraph 11) and with the positive and solvent/vehicle controls (see paragraphs 20-22), and use these data to confirm the reproducibility of the test method in their laboratory is maintained over time.

## PROCEDURE

9. This test method is based on the h-CLAT DataBase service on ALternative Methods to animal experimentation (DB-ALM) protocol no. 158 (18) which represents the protocol used for the EURL ECVAM-coordinated validation study. It is recommended that this protocol is used when implementing and using the h-CLAT method in the laboratory. The following is a description of the main components and procedures for the h-CLAT method, which comprises two steps: *dose finding assay* and *CD86/CD54 expression measurement*.

### *Preparation of cells*

10. The human monocytic leukaemia cell line, THP-1, should be used for performing the h-CLAT method. It is recommended that cells (TIB-202™) are obtained from a well-qualified cell bank, such as the American Type Culture Collection.

11. THP-1 cells are cultured, at 37°C under 5% CO<sub>2</sub> and humidified atmosphere, in RPMI-1640 medium supplemented with 10% foetal bovine serum (FBS), 0.05 mM 2-mercaptoethanol, 100 units/mL penicillin and 100 µg/mL streptomycin. The use of penicillin and streptomycin in the culture medium can be avoided. However, in such a case users should verify that the absence of antibiotics in the culture medium has no impact on the results, for example by testing the proficiency substances listed in Appendix II. In any case, in order to minimise the risk of contamination, good cell culture practices should be followed independently of the presence or not of antibiotics in the cell culture medium. THP-1 cells are routinely seeded every 2-3 days at the density of 0.1 to 0.2 × 10<sup>6</sup> cells/mL. They should be maintained at densities from 0.1 to 1.0 × 10<sup>6</sup> cells/mL. Prior to using them for testing, the cells should be qualified by conducting a reactivity check. The reactivity check of the cells should be performed using the positive controls, 2,4-dinitrochlorobenzene (DNCB) (CAS n. 97-00-7, ≥ 99% purity) and nickel sulfate (NiSO<sub>4</sub>) (CAS n. 10101-97-0, ≥ 99% purity) and the negative control, lactic acid (LA) (CAS n. 50-21-5, ≥ 85% purity), two weeks after thawing. Both DNCB and NiSO<sub>4</sub> should produce a positive response of both CD86 and CD54 cell surface markers, and LA should produce a negative response of both CD86 and CD54 cell surface markers. Only the cells which passed the reactivity check are to be used for the assay. Cells can be propagated up to two months after thawing. Passage number should not exceed 30. The reactivity check should be performed according to the procedures described in paragraphs 20-24.

12. For testing, THP-1 cells are seeded at a density of either 0.1 × 10<sup>6</sup> cells/mL or 0.2 × 10<sup>6</sup> cells/mL, and pre-cultured in culture flasks for 72 hours or for 48 hours, respectively. It is important that the cell density in the culture flask just after the pre-culture period be as consistent as possible in each experiment (by using one of the two pre-culture conditions described above), because the cell density in the culture flask just after pre-culture could affect the CD86/CD54 expression induced by allergens (19). On the day of testing, cells harvested from culture flask are resuspended with fresh culture medium at 2 × 10<sup>6</sup> cells/mL. Then, cells are distributed into a 24 well flat-bottom plate with 500 µL (1 × 10<sup>6</sup> cells/well) or a 96-well flat-bottom plate with 80 µL (1.6 × 10<sup>5</sup> cells/well).

### *Dose finding assay*

13. A *dose finding assay* is performed to determine the CV75, being the test chemical concentration that results in 75% cell viability (CV) compared to the solvent/vehicle control. The CV75 value is used to determine the concentration of test chemicals for the *CD86/CD54 expression measurement* (see paragraphs 20-24).

*Preparation of test chemicals and control substances*

14. The test chemicals and control substances are prepared on the day of testing. For the h-CLAT method, test chemicals are dissolved or stably dispersed (see also paragraph 4) in saline or medium as first solvent/vehicle options or dimethyl sulfoxide (DMSO,  $\geq 99\%$  purity) as a second solvent/vehicle option if the test chemical is not soluble or does not form a stable dispersion in the previous two solvents/vehicles, to final concentrations of 100 mg/mL (in saline or medium) or 500 mg/mL (in DMSO). Other solvents/vehicles than those described above may be used if sufficient scientific rationale is provided. Stability of the test chemical in the final solvent/vehicle should be taken into account.

15. Starting from the 100 mg/mL (in saline or medium) or 500 mg/mL (in DMSO) stock solutions of the test chemicals, the following dilution steps should be taken:

- For saline or medium as solvent/vehicle: Eight stock solutions (eight concentrations) are prepared, by two-fold serial dilutions using the corresponding solvent/vehicle. These stock solutions are then further diluted 50-fold into culture medium (working solutions). If the top final concentration in the plate of 1000  $\mu\text{g/mL}$  is non-toxic, the maximum concentration should be re-determined by performing a new cytotoxicity test. The final concentration in the plate should not exceed 5000  $\mu\text{g/mL}$  for test chemicals dissolved or stably dispersed in saline or medium.
- For DMSO as solvent/vehicle: Eight stock solutions (eight concentrations) are prepared, by two-fold serial dilutions using the corresponding solvent/vehicle. These stock solutions are then further diluted 250-fold into culture medium (working solutions). The final concentration in plate should not exceed 1000  $\mu\text{g/mL}$  even if this concentration is non-toxic.

The working solutions are finally used for exposure by adding an equal volume of working solution to the volume of THP-1 cell suspension in the plate (see also paragraph 17) to achieve a further two-fold dilution (usually, the final range of concentrations in the plate is 7.81–1000  $\mu\text{g/mL}$ ).

16. The solvent/vehicle control used in the h-CLAT method is culture medium (for test chemicals solubilised or stably dispersed (see paragraph 4) either with medium or saline) or DMSO (for test chemicals solubilised or stably dispersed in DMSO) tested at a single final concentration in the plate of 0.2%. It undergoes the same dilution as described for the working solutions in paragraph 15.

*Application of test chemicals and control substances*

17. The culture medium or working solutions described in paragraphs 15 and 16 are mixed 1:1 (v/v) with the cell suspensions prepared in the 24-well or 96-well flat-bottom plate (see paragraph 12). The treated plates are then incubated for  $24 \pm 0.5$  hours at  $37^\circ\text{C}$  under 5%  $\text{CO}_2$ . Care should be taken to avoid evaporation of volatile test chemicals and cross-contamination between wells by test chemicals, e.g. by sealing the plate prior to the incubation with the test chemicals (20).

*Propidium iodide (PI) staining*

18. After  $24 \pm 0.5$  hours of exposure, cells are transferred into sample tubes and collected by centrifugation. The supernatants are discarded and the remaining cells are resuspended with 200  $\mu\text{L}$  (in case of 96-well) or 600  $\mu\text{L}$  (in case of 24-well) of a phosphate buffered saline containing 0.1% bovine serum albumin (staining buffer). 200  $\mu\text{L}$  of cell suspension is transferred into 96-well round-bottom plate (in case of 96-well) or micro tube (in case of 24-well) and washed twice with 200  $\mu\text{L}$  (in case of 96-well) or 600  $\mu\text{L}$  (in case of 24-well) of staining buffer. Finally, cells are resuspended in staining buffer (e.g. 400  $\mu\text{L}$ ) and PI solution (e.g. 20  $\mu\text{L}$ ) is added (for example, final concentration of PI is 0.625  $\mu\text{g/mL}$ ). Other

cytotoxicity markers, such as 7-Aminoactinomycin D (7-AAD), Trypan blue or others may be used if the alternative stains can be shown to provide similar results as PI, for example by testing the proficiency substances in Appendix II.

*Cytotoxicity measurement by flow cytometry and estimation of CV75 value*

19. The PI uptake is analysed using flow cytometry with the acquisition channel FL-3. A total of 10,000 living cells (PI negative) are acquired. The cell viability can be calculated using the following equation by the cytometer analysis program. When the cell viability is low, up to 30,000 cells including dead cells should be acquired. Alternatively, data can be acquired for one minute after the initiation of the analysis.

$$\text{Cell Viability} = \frac{\text{Number of living cells}}{\text{Total Number of acquired cells}} \times 100$$

The CV75 value (see paragraph 13), i.e. a concentration showing 75% of THP-1 cell survival (25% cytotoxicity), is calculated by log-linear interpolation using the following equation:

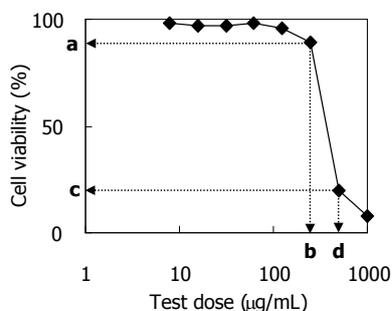
$$\text{Log CV75} = \frac{(75 - c) \times \text{Log (b)} - (75 - a) \times \text{Log (d)}}{a - c}$$

Where:

a is the minimum value of cell viability over 75%

c is the maximum value of cell viability below 75%

b and d are the concentrations showing the value of cell viability a and c respectively



Other approaches to derive the CV75 can be used as long as it is demonstrated that this has no impact on the results (e.g. by testing the proficiency substances).

*CD86/CD54 expression measurement**Preparation of the test chemicals and control substances*

20. The appropriate solvent/vehicle (saline, medium or DMSO; see paragraph 14) is used to dissolve or stably disperse the test chemicals. The test chemicals are first diluted to the concentration corresponding to 100-fold (for saline or medium) or 500-fold (for DMSO) of the  $1.2 \times CV75$  determined in the *dose finding assay* (see paragraph 19). If the CV75 cannot be determined (i.e. if sufficient cytotoxicity is not observed in the *dose finding assay*), the highest soluble or stably dispersed concentration of test chemical prepared with each solvent/vehicle should be used as starting concentration. Please note that the final concentration in the plate should not exceed 5000 µg/mL (in case of saline or medium) or 1000 µg/mL (in case of DMSO). Then, 1.2-fold serial dilutions are made using the corresponding solvent/vehicle to obtain the stock solutions (eight concentrations ranging from  $100 \times 1.2 \times CV75$  to  $100 \times 0.335 \times CV75$  (for saline or medium) or from  $500 \times 1.2 \times CV75$  to  $500 \times 0.335 \times CV75$  (for DMSO)) to be tested in the h-CLAT method (see DB-ALM protocol No. 158 for an example of dosing scheme). The stock solutions are then further diluted 50-fold (for saline or medium) or 250-fold (for DMSO) into the culture medium (working solutions). These working solutions are finally used for exposure with a further final two-fold dilution factor in the plate. If the results do not meet the acceptance criteria described in the paragraphs 29 and 30 regarding cell viability, the *dose finding assay* may be repeated to determine a more precise CV75. Please note that only 24-well plates can be used for CD86/CD54 expression measurement.

21. The solvent/vehicle control is prepared as described in paragraph 16. The positive control used in the h-CLAT method is DNCB (see paragraph 11), for which stock solutions are prepared in DMSO and diluted as described for the stock solutions in paragraph 20. DNCB should be used as the positive control for *CD86/CD54 expression measurement* at a final single concentration in the plate (typically 4.0 µg/mL). To obtain a 4.0 µg/mL concentration of DNCB in the plate, a 2 mg/mL stock solution of DNCB in DMSO is prepared and further diluted 250-fold with culture medium to a 8 µg/mL working solution. Alternatively, the CV75 of DNCB, which is determined in each test facility, could be also used as the positive control concentration. Other suitable positive controls may be used if historical data are available to derive comparable run acceptance criteria. For positive controls, the final single concentration in the plate should not exceed 5000 µg/mL (in case of saline or medium) or 1000 µg/mL (in case of DMSO). The run acceptance criteria are the same as those described for the test chemical (see paragraph 29), except for the last acceptance criterion since the positive control is tested at a single concentration.

*Application of test chemicals and control substances*

22. For each test chemical and control substance, one experiment is needed to obtain a prediction. Each experiment consists of at least two independent runs for *CD86/CD54 expression measurement* (see paragraphs 26-28). Each independent run is performed on a different day or on the same day provided that for each run: a) independent fresh stock solutions and working solutions of the test chemical and antibody solutions are prepared and b) independently harvested cells are used (i.e. cells are collected from different culture flasks); however, cells may come from the same passage. Test chemicals and control substances prepared as working solutions (500 µL) are mixed with 500 µL of suspended cells ( $1 \times 10^6$  cells) at 1:1 ratio, and cells are incubated for  $24 \pm 0.5$  hours as described in paragraphs 20 and 21. In each run, a single replicate for each concentration of the test chemical and control substance is sufficient because a prediction is obtained from at least two independent runs.

*Cell staining and analysis*

23. After 24±0.5 hours of exposure, cells are transferred from 24 well plate into sample tubes, collected by centrifugation and then washed twice with 1mL of staining buffer (if necessary, additional washing steps may be done). After washing, cells are blocked with 600 µL of blocking solution (staining buffer containing 0.01% (w/v) globulin (Cohn fraction II, III, Human: SIGMA, #G2388-10G)) and incubated at 4°C for 15 min. After blocking, cells are split in three aliquots of 180 µL into a 96-well round-bottom plate or micro tube.

24. After centrifugation, cells are stained with 50 µL of FITC-labelled anti-CD86, anti-CD54 or mouse IgG1 (isotype) antibodies at 4°C for 30 min. The antibodies described in the h-CLAT DB-ALM protocol no. 158 (18) should be used by diluting 3:25 (v/v, for CD86 (BD-PharMingen, #555657; Clone: Fun-1)) or 3:50 (v/v, for CD54 (DAKO, #F7143; Clone: 6.5B5) and IgG1 (DAKO, #X0927)) with staining buffer. These antibody dilution factors were defined by the test method developers as those providing the best signal-to-noise ratio. Based on the experience of the test method developers, the fluorescence intensity of the antibodies is usually consistent between different lots. However, users may consider titrating the antibodies in their own laboratory's conditions to define the best concentrations for use. Other fluorochrome-tagged anti-CD86 and/or anti-CD54 antibodies may be used if they can be shown to provide similar results as FITC-conjugated antibodies, for example by testing the proficiency substances in Appendix II. It should be noted that changing the clone or supplier of the antibodies as described in the h-CLAT DB-ALM protocol no. 158 (18) may affect the results. After washing twice or more with 150 µL of staining buffer, cells are resuspended in staining buffer (e.g. 400 µL), and the PI solution (e.g. 20 µL to obtain a final concentration of 0.625 µg/mL) or another cytotoxicity marker's solution (see paragraph 18) is added. The expression levels of CD86 and CD54, and cell viability are analysed using flow cytometry.

**DATA AND REPORTING***Data evaluation*

25. The expression of CD86 and CD54 is analysed with flow cytometry with the acquisition channel FL-1. Based on the geometric mean fluorescence intensity (MFI), the relative fluorescence intensity (RFI) of CD86 and CD54 for positive control (ctrl) cells and chemical-treated cells are calculated according to the following equation:

$$\text{RFI} = \frac{\text{MFI of chemical-treated cells} - \text{MFI of chemical-treated isotype control cells}}{\text{MFI of solvent/vehicle-treated ctrl cells} - \text{MFI of solvent/vehicle-treated isotype ctrl cells}} \times 100$$

The cell viability from the isotype control (ctrl) cells (which are stained with mouse IgG1 (isotype) antibodies) is also calculated according to the equation described in paragraph 19.

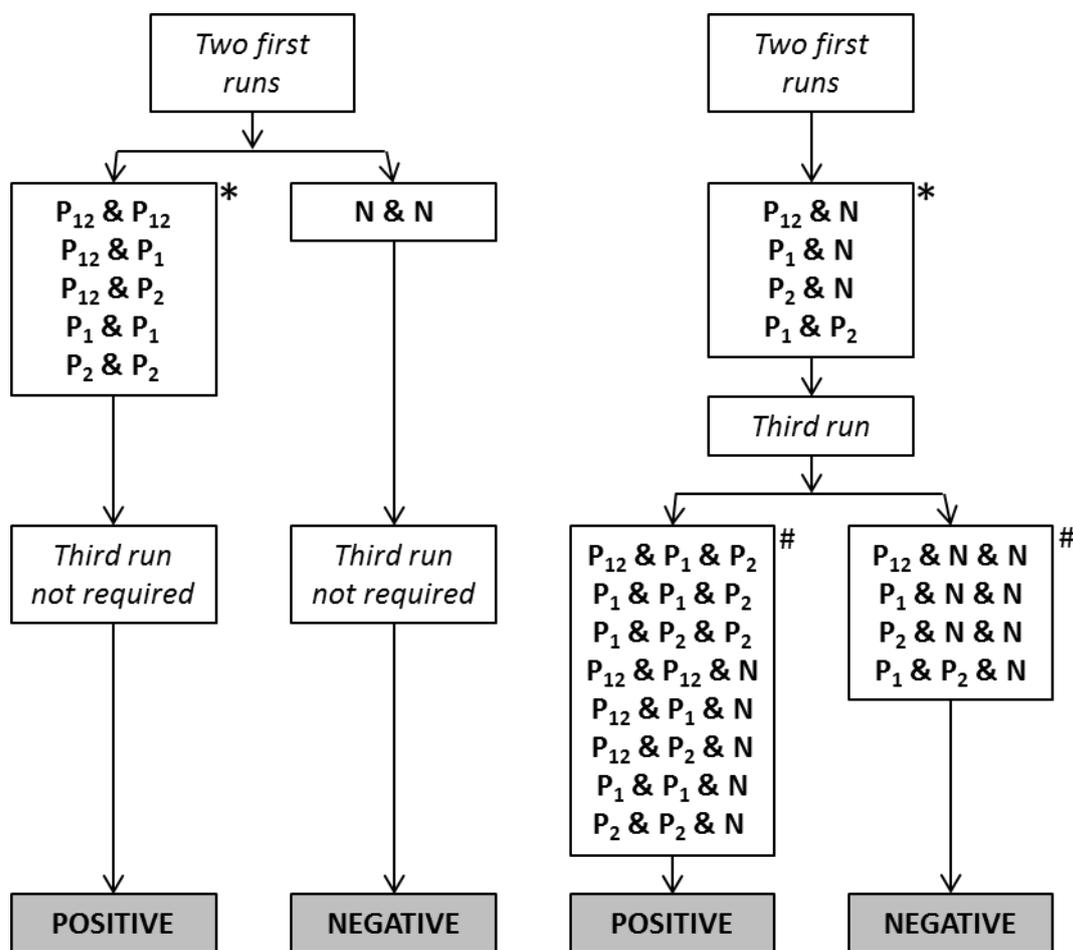
*Prediction model*

26. For *CD86/CD54 expression measurement*, each test chemical is tested in at least two independent runs to derive a single prediction (POSITIVE or NEGATIVE). An h-CLAT prediction is considered POSITIVE

if at least one of the following conditions is met in 2 of 2 or in at least 2 of 3 independent runs, otherwise the h-CLAT prediction is considered NEGATIVE (Figure 1):

- The RFI of CD86 is equal to or greater than 150% at any tested concentration (with cell viability  $\geq$  50%);
- The RFI of CD54 is equal to or greater than 200% at any tested concentration (with cell viability  $\geq$  50%).

27. Based on the above, if the first two runs are both positive for CD86 and/or are both positive for CD54, the h-CLAT prediction is considered POSITIVE and a third run does not need to be conducted. Similarly, if the first two runs are negative for both markers, the h-CLAT prediction is considered NEGATIVE (with due consideration of the provisions of paragraph 30) without the need for a third run. If however, the first two runs are not concordant for at least one of the markers (CD54 or CD86), a third run is needed and the final prediction will be based on the majority result of the three individual runs (i.e. 2 out of 3). In this respect, it should be noted that if two independent runs are conducted and one is only positive for CD86 (hereinafter referred to as  $P_1$ ) and the other is only positive for CD54 (hereinafter referred to as  $P_2$ ), a third run is required. If this third run is negative for both markers (hereinafter referred to as N), the h-CLAT prediction is considered NEGATIVE. On the other hand, if the third run is positive for either marker ( $P_1$  or  $P_2$ ) or for both markers (hereinafter referred to as  $P_{12}$ ), the h-CLAT prediction is considered POSITIVE.



**Figure 1: Prediction model used in the h-CLAT test method.** An h-CLAT prediction should be considered in the framework of an IATA and in accordance with the provision of paragraphs 7 and 8 in the General introduction. P<sub>1</sub>: run with only CD86 positive; P<sub>2</sub>: run with only CD54 positive; P<sub>12</sub>: run with both CD86 and CD54 positive; N: run with neither CD86 nor CD54 positive. \*The boxes show the relevant combinations of results from the first two runs, independently of the order in which they may be obtained. #The boxes show the relevant combinations of results from the three runs on the basis of the results obtained in the first two runs shown in the box above, but do not reflect the order in which they may be obtained.

28. For the test chemicals predicted as POSITIVE with the h-CLAT, optionally, two Effective Concentrations (EC) values, the EC<sub>150</sub> for CD86 and EC<sub>200</sub> for CD54, i.e. the concentration at which the test chemicals induced a RFI of 150 or 200, may be determined. These EC values potentially could contribute to the assessment of sensitising potency (9) when used in integrated approaches such as IATA (4) (5) (6) (7) (8). They can be calculated by the following equations:

$$EC_{150} \text{ (for CD86)} = B_{\text{concentration}} + [(150 - B_{RFI}) / (A_{RFI} - B_{RFI}) \times (A_{\text{concentration}} - B_{\text{concentration}})]$$

$$EC_{200} \text{ (for CD54)} = B_{\text{concentration}} + [(200 - B_{RFI}) / (A_{RFI} - B_{RFI}) \times (A_{\text{concentration}} - B_{\text{concentration}})]$$

where

$A_{\text{concentration}}$  is the lowest concentration in  $\mu\text{g/mL}$  with  $\text{RFI} > 150$  (CD86) or 200 (CD54)

$B_{\text{concentration}}$  is the highest concentration in  $\mu\text{g/mL}$  with  $\text{RFI} < 150$  (CD86) or 200 (CD54)

$A_{\text{RFI}}$  is the RFI at the lowest concentration with  $\text{RFI} > 150$  (CD86) or 200 (CD54)

$B_{\text{RFI}}$  is the RFI at the highest concentration with  $\text{RFI} < 150$  (CD86) or 200 (CD54)

For the purpose of more precisely deriving the EC150 and EC200 values, three independent runs for *CD86/CD54 expression measurement* may be required. The final EC150 and EC200 values are then determined as the median value of the ECs calculated from the three independent runs. When only two of three independent runs meet the criteria for positivity (see paragraphs 26-27), the higher EC150 or EC200 of the two calculated values is adopted.

### ***Acceptance criteria***

29. The following acceptance criteria should be met when using the h-CLAT method (22) (27).

- The cell viabilities of medium and solvent/vehicle controls should be higher than 90%.
- In the solvent/vehicle control, RFI values of both CD86 and CD54 should not exceed the positive criteria (CD86  $\text{RFI} \geq 150\%$  and CD54  $\text{RFI} \geq 200\%$ ). RFI values of the solvent/vehicle control are calculated by using the formula described in paragraph 25 ("MFI of chemical" should be replaced with "MFI of solvent/vehicle", and "MFI of solvent/vehicle" should be replaced with "MFI of (medium) control").
- For both medium and solvent/vehicle controls, the MFI ratio of both CD86 and CD54 to isotype control should be  $> 105\%$ .
- In the positive control (DNCB), RFI values of both CD86 and CD54 should meet the positive criteria (CD86  $\text{RFI} \geq 150$  and CD54  $\text{RFI} \geq 200$ ) and cell viability should be more than 50%.
- For the test chemical, the cell viability should be more than 50% in at least four tested concentrations in each run.

30. Negative results are acceptable only for test chemicals exhibiting a cell viability of less than 90% at the highest concentration tested (i.e.  $1.2 \times \text{CV75}$  according to the serial dilution scheme described in paragraph 20). If the cell viability at  $1.2 \times \text{CV75}$  is equal or above 90% the negative result should be discarded. In such a case it is recommended to try to refine the dose selection by repeating the CV75 determination. It should be noted that when 5000  $\mu\text{g/mL}$  in saline (or medium or other solvents/vehicles), 1000  $\mu\text{g/mL}$  in DMSO or the highest soluble concentration is used as the maximal test concentration of a test chemical, a negative result is acceptable even if the cell viability is above 90%.

### ***Test report***

31. The test report should include the following information.

#### ***Test chemical***

- Mono-constituent substance

- Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
  - Physical appearance, Log Kow, water solubility, DMSO solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available;
  - Justification for choice of solvent/vehicle for each test chemical.
- Multi-constituent substance, UVCB and mixture
- Characterisation as far as possible by e.g. chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available;
  - Physical appearance, water solubility, DMSO solubility and additional relevant physicochemical properties, to the extent available;
  - Molecular weight or apparent molecular weight in case of mixtures/polymers of known compositions or other information relevant for the conduct of the study;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available;
  - Justification for choice of solvent/vehicle for each test chemical.

#### *Controls*

- Positive control
- Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
  - Physical appearance, Log Kow, water solubility, DMSO solubility, molecular weight, and additional relevant physicochemical properties, to the extent available and where applicable;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available;
  - Reference to historical positive control results demonstrating suitable run acceptance criteria, if applicable.

- Negative and solvent/vehicle control
  - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
  - Physical appearance, molecular weight, and additional relevant physicochemical properties in the case other control solvent/vehicle than those mentioned in the Test Guideline are used and to the extent available;
  - Storage conditions and stability to the extent available;
  - Justification for choice of solvent/vehicle for each test chemical.

#### *Test method conditions*

- Name and address of the sponsor, test facility and study director;
- Description of test method used;
- Cell line used, its storage conditions and source (e.g. the facility from which they were obtained);
- Flow cytometry used (e.g. model), including instrument settings, globulin, antibodies and cytotoxicity marker used;
- The procedure used to demonstrate proficiency of the laboratory in performing the test method by testing of proficiency substances, and the procedure used to demonstrate reproducible performance of the test method over time, e.g. historical control data and/or historical reactivity checks' data.

#### *Test acceptance criteria*

- Cell viability, MFI and RFI values obtained with the solvent/vehicle control in comparison to the acceptance ranges;
- Cell viability and RFI values obtained with the positive control in comparison to the acceptance ranges;
- Cell viability of all tested concentrations of the tested chemical.

#### *Test procedure*

- Number of runs used;
- Test chemical concentrations, application and exposure time used (if different than the one recommended)
- Duration of exposure (if different than the one recommended);
- Description of evaluation and decision criteria used;
- Description of any modifications of the test procedure.

## *Results*

- Tabulation of the data, including CV75 (if applicable), individual geometric MFI, RFI, cell viability values, EC150/EC200 values (if applicable) obtained for the test chemical and for the positive control in each run, and an indication of the rating of the test chemical according to the prediction model;
- Description of any other relevant observations, if applicable.

## *Discussion of the results*

- Discussion of the results obtained with the h-CLAT method;
- Consideration of the test method results within the context of an IATA, if other relevant information is available.

## *Conclusions*

## LITERATURE

1. Ashikaga T, Yoshida Y, Hirota M, Yoneyama K, Itagaki H, Sakaguchi H, Miyazawa M, Ito Y, Suzuki H, Toyoda H. (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, 767-773.
2. Miyazawa M, Ito Y, Yoshida Y, Sakaguchi H, Suzuki H. (2007), Phenotypic alterations and cytokine production in THP-1 cells in response to allergens. *Toxicol. In Vitro* 21, 428-437.
3. EC EURL-ECVAM (2013), Recommendation on the human Cell Line Activation Test (h-CLAT) for skin sensitisation testing. Accessible at: <https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations>
4. Takenouchi O, Fukui S, Okamoto K, Kurotani S, Imai N, Fujishiro M, Kyotani D, Kato Y, Kasahara T, Fujita M, Toyoda A, Sekiya D, Watanabe S, Seto H, Hirota M, Ashikaga T, Miyazawa M. (2015), Test battery with the human cell line activation test, direct peptide reactivity assay and DEREK based on a 139 chemical data set for predicting skin sensitizing potential and potency of chemicals. *J Appl Toxicol.* 35, 1318-1332.
5. Hirota M, Fukui S, Okamoto K, Kurotani S, Imai N, Fujishiro M, Kyotani D, Kato Y, Kasahara T, Fujita M, Toyoda A, Sekiya D, Watanabe S, Seto H, Takenouchi O, Ashikaga T, Miyazawa M. (2015), Evaluation of combinations of in vitro sensitization test descriptors for the artificial neural network-based risk assessment model of skin sensitization. *J Appl Toxicol.* 35, 1333-1347.
6. Bauch C, Kolle SN, Ramirez T, Fabian E, Mehling A, Teubner W, van Ravenzwaay B, Landsiedel R. (2012), Putting the parts together: combining in vitro methods to test for skin sensitizing potentials. *Regul Toxicol Pharmacol.* 63, 489-504.
7. Van der Veen JW, Rorije E, Emter R, Natch A, van Loveren H, Ezendam J. (2014), Evaluating the performance of integrated approaches for hazard identification of skin sensitizing chemicals. *Regul Toxicol Pharmacol.* 69, 371-379.
8. Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle S, Landsiedel R, Jaworska J, Kern PS, Gerberick F, Natsch A, Emter R, Ashikaga T, Miyazawa M, Sakaguchi H. (2015), Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul Toxicol Pharmacol.* 71, 337-351.
9. Jaworska JS, Natsch A, Ryan C, Strickland J, Ashikaga T, Miyazawa M. (2015), Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative weight of evidence and adaptive testing strategy. *Arch Toxicol.* 89, 2355-2383.
10. Strickland J, Zang Q, Kleinstreuer N, Paris M, Lehmann DM, Choksi N, Matheson J, Jacobs A, Lowit A, Allen D, Casey W. (2016), Integrated decision strategies for skin sensitization hazard. *J Appl Toxicol.* DOI 10.1002/jat.3281.
11. Nukada Y, Ashikaga T, Miyazawa M, Hirota M, Sakaguchi H, Sasa H, Nishiyama N. (2012), Prediction of skin sensitization potency of chemicals by human Cell Line Activation Test (h-CLAT)

and an attempt at classifying skin sensitization potency. *Toxicol. In Vitro* 26, 1150-60.

12. EC EURL ECVAM (2015), Re-analysis of the within and between laboratory reproducibility of the human Cell Line Activation Test (h-CLAT). Accessible at: <https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations/eurl-ecvam-recommendation-on-the-human-cell-line-activation-test-h-clat-for-skin-sensitisation-testing>
13. EC EURL ECVAM (2012), human Cell Line Activation Test (h-CLAT) Validation Study Report Accessible at: <https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations>
14. Takenouchi O, Miyazawa M, Saito K, Ashikaga T, Sakaguchi H. (2013), Predictive performance of the human Cell Line Activation Test (h-CLAT) for lipophilic with high octanol-water partition coefficients. *J. Toxicol. Sci.* 38, 599-609.
15. Ashikaga T, Sakaguchi H, Sono S, Kosaka N, Ishikawa M, Nukada Y, Miyazawa M, Ito Y, Nishiyama N, Itagaki H. (2010), A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). *Altern. Lab. Anim.* 38, 275-284.
16. Fabian E., Vogel D., Blatz V., Ramirez T., Kolle S., Eltze T., van Ravenzwaay B., Oesch F., Landsiedel R. (2013), Xenobiotic metabolizing enzyme activities in cells used for testing skin sensitization in vitro. *Arch Toxicol* 87, 1683-1969.
17. Okamoto K, Kato Y, Kosaka N, Mizuno M, Inaba H, Sono S, Ashikaga T, Nakamura T, Okamoto Y, Sakaguchi H, Kishi M, Kuwahara H, Ohno Y. (2010), The Japanese ring study of a human Cell Line Activation Test (h-CLAT) for predicting skin sensitization potential (6<sup>th</sup> report): A study for evaluating oxidative hair dye sensitization potential using h-CLAT. *AATEX* 15, 81-88.
18. DB-ALM (INVITTOX) (2014), Protocol 158: human Cell Line Activation Test (h-CLAT), 23pp. Accessible at: <http://ecvam-dbalm.jrc.ec.europa.eu/>
19. Mizuno M, Yoshida M, Kodama T, Kosaka N, Okamoto K, Sono S, Yamada T, Hasegawa S, Ashikaga T, Kuwahara H, Sakaguchi H, Sato J, Ota N, Okamoto Y, Ohno Y. (2008), Effects of pre-culture conditions on the human Cell Line Activation Test (h-CLAT) results; Results of the 4th Japanese inter-laboratory study. *AATEX* 13, 70-82.
20. Sono S, Mizuno M, Kosaka N, Okamoto K, Kato Y, Inaba H, , Nakamura T, Kishi M, Kuwahara H, Sakaguchi H, Okamoto Y, Ashikaga T, Ohno Y. (2010), The Japanese ring study of a human Cell Line Activation Test (h-CLAT) for predicting skin sensitization potential (7<sup>th</sup> report): Evaluation of volatile, poorly soluble fragrance materials. *AATEX* 15, 89-96.
21. OECD (2005), Guidance Document No.34 on “the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment”. OECD Series on Testing and Assessment. Organization for Economic Cooperation and Development, Paris, France, 2005, 96 pp.
22. OECD (2012), The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Series on Testing and Assessment No. 168. Available at: [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO\(2012\)10/PART1&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)10/PART1&docLanguage=En)

23. United Nations UN (2013), Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Fifth revised edition. New York & Geneva: United Nations Publications. ISBN: 978-92-1-117006-1. Available at: [[http://www.unece.org/trans/danger/publi/ghs/ghs\\_rev05/05files\\_e.html](http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html)].
24. ECETOC (2003). Contact sensitization: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals (Technical Report No. 87).
25. Ashikaga T, Sakaguchi H, Okamoto K, Mizuno M, Sato J, Yamada T, Yoshida M, Ota N, Hasegawa S, Kodama T, Okamoto Y, Kuwahara H, Kosaka N, Sono S, Ohno Y. (2008), Assessment of the human Cell Line Activation Test (h-CLAT) for Skin Sensitization; Results of the First Japanese Inter-laboratory Study. AATEX 13, 27-35.

APPENDIX IDEFINITIONS

**Accuracy:** The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with concordance to mean the proportion of correct outcomes of a test method (21).

**AOP (Adverse Outcome Pathway):** sequence of events from the chemical structure of a target chemical or group of similar chemicals through the molecular initiating event to an *in vivo* outcome of interest (22).

**CV75:** The estimated concentration showing 75% cell viability.

**EC150:** the concentrations showing the RFI values of 150 in CD86 expression

**EC200:** the concentrations showing the RFI values of 200 in CD54 expression

**Flow cytometry:** a cytometric technique in which cells suspended in a fluid flow one at a time through a focus of exciting light, which is scattered in patterns characteristic to the cells and their components; cells are frequently labeled with fluorescent markers so that light is first absorbed and then emitted at altered frequencies.

**Hazard:** Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub) population is exposed to that agent.

**IATA (Integrated Approach to Testing and Assessment):** A structured approach used for hazard identification (potential), hazard characterisation (potency) and/or safety assessment (potential/potency and exposure) of a chemical or group of chemicals, which strategically integrates and weights all relevant data to inform regulatory decision regarding potential hazard and/or risk and/or the need for further targeted and therefore minimal testing.

**Medium control:** An untreated replicate containing all components of a test system. This sample is processed with test chemical-treated samples and other control samples to determine whether the solvent/vehicle interacts with the test system.

**Mixture:** A mixture or a solution composed of two or more substances in which they do not react.

**Mono-constituent substance:** A substance, defined by its quantitative composition, in which one main constituent is present to at least 80% (w/w).

**Multi-constituent substance:** A substance, defined by its quantitative composition, in which more than one main constituent is present in a concentration  $\geq 10\%$  (w/w) and  $< 80\%$  (w/w). A multi-constituent substance is the result of a manufacturing process. The difference between mixture and multi-constituent

substance is that a mixture is obtained by blending of two or more substances without chemical reaction. A multi-constituent substance is the result of a chemical reaction.

**Positive control:** A replicate containing all components of a test system and treated with a substance known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the positive response should not be excessive.

**Pre-haptens:** chemicals which become sensitisers through abiotic transformation

**Pro-haptens:** chemicals requiring enzymatic activation to exert skin sensitisation potential

**Relative fluorescence intensity (RFI):** Relative values of geometric mean fluorescence intensity (MFI) in chemical-treated cells compared to MFI in solvent/vehicle-treated cells.

**Relevance:** Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (21).

**Reliability:** Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability (21).

**Run:** A run consists of one or more test chemicals tested concurrently with a solvent/vehicle control and with a positive control.

**Sensitivity:** The proportion of all positive/active chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (21).

**Staining buffer:** A phosphate buffered saline containing 0.1% bovine serum albumin.

**Solvent/vehicle control:** An untreated sample containing all components of a test system except of the test chemical, but including the solvent/vehicle that is used. It is used to establish the baseline response for the samples treated with the test chemical dissolved or stably dispersed in the same solvent/vehicle. When tested with a concurrent medium control, this sample also demonstrates whether the solvent/vehicle interacts with the test system.

**Specificity:** The proportion of all negative/inactive chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method (21).

**Substance:** Chemical elements and their compounds in the natural state or obtained by any production process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition.

**Test chemical:** The term "test chemical" is used to refer to what is being tested.

**United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS):** A system proposing the classification of chemicals (substances and mixtures) according to standardised types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (23).

**UVCB:** substances of unknown or variable composition, complex reaction products or biological materials.

**Valid test method:** A test method considered to have sufficient relevance and reliability for a specific purpose and which is based on scientifically sound principles. A test method is never valid in an absolute sense, but only in relation to a defined purpose (21).

APPENDIX IIPROFICIENCY SUBSTANCES

Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency by correctly obtaining the expected h-CLAT prediction for the 10 substances recommended in Table 1 and by obtaining CV75, EC150 and EC200 values that fall within the respective reference range for at least 8 out of the 10 proficiency substances. Proficiency substances were selected to represent the range of responses for skin sensitisation hazards. Other selection criteria were that the substances are commercially available, and that high-quality *in vivo* reference data as well as high quality *in vitro* data generated with the h-CLAT method are available. Also, published reference data are available for the h-CLAT method (3) (14).

Table 1: Recommended substances for demonstrating technical proficiency with the h-CLAT method

Proficiency substances	CASRN	Physical state	<i>In vivo</i> prediction <sup>1</sup>	CV75 Reference Range in µg/mL <sup>2</sup>	h-CLAT results for CD86 (EC150 Reference Range in µg/mL) <sup>2</sup>	h-CLAT results for CD54 (EC200 Reference Range in µg/mL) <sup>2</sup>
2,4-Dinitrochlorobenzene	97-00-7	Solid	Sensitiser (extreme)	2-12	Positive (0.5-10)	Positive (0.5-15)
4-Phenylenediamine	106-50-3	Solid	Sensitiser (strong)	5-95	Positive (<40)	Negative (>1.5) <sup>3</sup>
Nickel sulfate	10101-97-0	Solid	Sensitiser (moderate)	30-500	Positive (<100)	Positive (10-100)
2-Mercaptbenzothiazole	149-30-4	Solid	Sensitiser (moderate)	30-400	Negative (>10) <sup>3</sup>	Positive (10-140)
R(+)-Limonene	5989-27-5	Liquid	Sensitiser (weak)	>20	Negative (>5) <sup>3</sup>	Positive (<250)
Imidazolidinyl urea	39236-46-9	Solid	Sensitiser (weak)	25-100	Positive (20-90)	Positive (20-75)
Isopropanol	67-63-0	Liquid	Non-sensitiser	>5000	Negative (>5000)	Negative (>5000)
Glycerol	56-81-5	Liquid	Non-sensitiser	>5000	Negative (>5000)	Negative (>5000)
Lactic acid	50-21-5	Liquid	Non-sensitiser	1500-5000	Negative (>5000)	Negative (>5000)
4-Aminobenzoic acid	150-13-0	Solid	Non-sensitiser	>1000	Negative (>1000)	Negative (>1000)

Abbreviations: CAS RN = Chemical Abstracts Service Registry Number

<sup>1</sup> The *in vivo* hazard and (potency) prediction is based on LLNA data (3) (14). The *in vivo* potency is derived using the criteria proposed by ECETOC (24).

<sup>2</sup> Based on historical observed values (13) (25).

<sup>3</sup> Historically, a majority of negative results have been obtained for this marker and therefore a negative result is mostly expected. The range provided was defined on the basis of the few historical positive

results observed. In case a positive result is obtained, the EC value should be within the reported reference range.

## **Annex II: In Vitro Skin Sensitisation: U937 Cell Line Activation Test** **(U-SENS™)**

### **INITIAL CONSIDERATIONS AND LIMITATIONS**

1. The U-SENS™ method quantifies the change in the expression of a cell surface marker associated with the process of activation of monocytes and dendritic cells (DC) (i.e. CD86), in the human histiocytic lymphoma cell line U937, following exposure to sensitisers (1). The measured expression levels of CD86 cell surface marker in the cell line U937 is then used for supporting the discrimination between skin sensitisers and non-sensitisers.

2. The U-SENS™ method has been evaluated in a validation study (2) coordinated by L'Oreal and subsequently independent peer reviewed by the European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) Scientific Advisory Committee (ESAC) (3). Considering all available evidence and input from regulators and stakeholders, the U-SENS™ was recommended by EURL ECVAM (4) to be used as part of an IATA to support the discrimination between sensitisers and non-sensitisers for the purpose of hazard classification and labelling. In its guidance document on the reporting of structured approaches to data integration and individual information sources used within IATA for skin sensitisation, the OECD currently discusses a number of case studies describing different testing strategies and prediction models. One of the different defined approaches is based on the U-SENS assay (5). Examples of the use of U-SENS™ data in combination with other information, including historical data and existing valid human data (6), are also reported elsewhere in the literature (4) (5) (7).

3. The U-SENS™ method proved to be transferable to laboratories experienced in cell culture techniques and flow cytometry analysis. The level of reproducibility in predictions that can be expected from the test method is in the order of 90% and 84% within and between laboratories, respectively (8). Results generated in the validation study (8) and other published studies (1) overall indicate that, compared with LLNA results, the accuracy in distinguishing skin sensitisers (i.e. UN GHS Cat.1) from non-sensitisers is 86% (N=166) with a sensitivity of 91% (118/129) and a specificity of 65% (24/37). Compared with human results, the accuracy in distinguishing skin sensitisers (i.e. UN GHS Cat.1) from non-sensitisers is 77% (N=101) with a sensitivity of 100% (58/58) and a specificity of 47% (20/43). False negative predictions compared to LLNA with the U-SENS™ are more likely to concern chemicals showing a low to moderate skin sensitisation potency (i.e. UN GHS subcategory 1B) than chemicals showing a high skin sensitisation potency (i.e. UN GHS subcategory 1A) (1) (8) (9). Taken together, this information indicates the usefulness of the U-SENS™ method to contribute to the identification of skin sensitisation hazards. However, the accuracy values given here for the U-SENS™ as a stand-alone test method are only indicative, since the test method should be considered in combination with other sources of information in the context of an IATA and in accordance with the provisions of paragraphs 7 and 8 in the General introduction. Furthermore, when evaluating non-animal methods for skin sensitisation, it should be kept in mind that the LLNA test as well as other animal tests may not fully reflect the situation in humans.

4. On the basis of the data currently available, the U-SENS™ method was shown to be applicable to test chemicals (including cosmetics ingredients e.g. preservatives, surfactants, actives, dyes) covering a variety of organic functional groups, of physicochemical properties, skin sensitisation potency (as determined in in vivo studies) and the spectrum of reaction mechanisms known to be associated with skin sensitisation (i.e. Michael acceptor, Schiff base formation, acyl transfer agent, substitution nucleophilic bimolecular [SN2], or nucleophilic aromatic substitution [SNAr]) (1) (8) (9) (10). The U-SENS™ method is

applicable to test chemicals that are soluble or that form a stable dispersion (i.e. a colloid or suspension in which the test chemical does not settle or separate from the solvent/vehicle into different phases) in an appropriate solvent/vehicle (see paragraph 13). Chemicals in the dataset reported to be pre-haptens (i.e. substances activated by oxidation) or pro-haptens (i.e. substances requiring enzymatic activation for example via P450 enzymes) were correctly predicted by the U-SENS™ (1) (10). Membrane disrupting substances can lead to false positive results due to a non-specific increase of CD86 expression, as 3 out of 7 false positives relative to the *in vivo* reference classification were surfactants (1). As such positive results with surfactants should be considered with caution whereas negative results with surfactants could still be used to support the identification of the test chemical as a non-sensitiser. Fluorescent test chemicals can be assessed with the U-SENS™ (1), nevertheless, strong fluorescent test chemicals emitting at the same wavelength as fluorescein isothiocyanate (FITC) or as propidium iodide (PI), will interfere with the flow cytometric detection and thus cannot be correctly evaluated using FITC-conjugated antibodies (potential false negative) or PI (viability not measurable). In such a case, other fluorochrome-tagged antibodies or other cytotoxicity markers, respectively, can be used as long as it can be shown they provide similar results as the FITC-tagged antibodies or PI (see paragraph 18) e.g. by testing the proficiency substances in Appendix II. In the light of the above, positive results with surfactants and negative results with strong fluorescent test chemicals should be interpreted in the context of the stated limitations and together with other information sources within the framework of IATA. In cases where there is evidence demonstrating the non-applicability of the U-SENS™ method to other specific categories of test chemicals, it should not be used for those specific categories.

5. As described above, the U-SENS™ method supports the discrimination between skin sensitisers from non-sensitisers. However, it may also potentially contribute to the assessment of sensitising potency when used in integrated approaches such as IATA. Nevertheless, further work, preferably based on human data, is required to determine how U-SENS™ results may possibly inform potency assessment.

6. Definitions are provided in Appendix I.

## PRINCIPLE OF THE TEST

7. The U-SENS™ method is an *in vitro* assay that quantifies changes of CD86 cell surface marker expression on a human histiocytic lymphoma cell line, U937 cells, following 45±3 hours exposure to the test chemical. The CD86 surface marker is one typical marker of U937 activation. CD86 is known to be a co-stimulatory molecule that may mimic monocytic activation, which plays a critical role in T-cell priming. The changes of CD86 cell surface marker expression are measured by flow cytometry following cell staining typically with fluorescein isothiocyanate (FITC)-labelled antibodies. Cytotoxicity measurement is also conducted (e.g. by using PI) concurrently to assess whether upregulation of CD86 cell surface marker expression occurs at sub-cytotoxic concentrations. The stimulation index (S.I.) of CD86 cell surface marker compared to solvent/vehicle control is calculated and used in the prediction model (see paragraph 19), to support the discrimination between sensitisers and non-sensitisers.

## DEMONSTRATION OF PROFICIENCY

8. Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency, using the 10 Proficiency Substances listed in

Appendix II in compliance with the Good *in vitro* Method Practices (11). Moreover, test method users should maintain a historical database of data generated with the reactivity checks (see paragraph 11) and with the positive and solvent/vehicle controls (see paragraphs 15-16), and use these data to confirm the reproducibility of the test method in their laboratory is maintained over time.

## PROCEDURE

9. This test method is based on the U-SENS<sup>TM</sup> DataBase service on ALternative Methods to animal experimentation (DB-ALM) protocol no. 183 (12). The Standard Operating Procedures (SOP) should be employed when implementing and using the U-SENS<sup>TM</sup> method in the laboratory. An automated system to run the U-SENS<sup>TM</sup> can be used if it can be shown to provide similar results, for example by testing the proficiency substances in Appendix II. The following is a description of the main components and procedures for the U-SENS<sup>TM</sup> method.

### *Preparation of cells*

10. The human histiocytic lymphoma cell line, U937 (13) should be used for performing the U-SENS<sup>TM</sup> method. Cells (clone CRL1593.2) should be obtained from a well-qualified cell bank such as the American Type Culture Collection.

11. U937 cells are cultured, at 37°C under 5% CO<sub>2</sub> and humidified atmosphere, in RPMI-1640 medium supplemented with 10% foetal calf serum (FCS), 2 mM L-glutamine, 100 units/mL penicillin and 100 µg/mL streptomycin (complete medium). U937 cells are routinely passaged every 2-3 days at the density of 1.5 or 3 × 10<sup>5</sup> cells/mL, respectively. The cell density should not exceed 2 × 10<sup>6</sup> cells/mL and the cell viability measured by trypan blue exclusion should be ≥ 90% (not to be applied at the first passage after thawing). Prior to using them for testing, every batch of cells, FCS or antibodies should be qualified by conducting a reactivity check. The reactivity check of the cells should be performed using the positive control, picrylsulfonic acid (2,4,6-Trinitro-benzene-sulfonic acid: TNBS) (CASRN 2508-19-2, ≥ 99% purity) and the negative control lactic acid (LA) (CASRN 50-21-5, ≥ 85% purity), at least one week after thawing. For the reactivity check, six final concentrations should be tested for each of the 2 controls (TNBS: 1, 12.5, 25, 50, 75, 100µg/mL and LA: 1, 10, 20, 50, 100, 200µg/mL). TNBS solubilised in complete medium should produce a positive and concentration-related response of CD86 (e.g. when a positive concentration, CD86 S.I. ≥ 150, is followed by a concentration with an increasing CD86 S.I.), and LA solubilised in complete medium should produce negative response of CD86 (see paragraph 21). Only the batch of cells which passed the reactivity check 2 times should be used for the assay. Cells can be propagated up to seven weeks after thawing. Passage number should not exceed 21. The reactivity check should be performed according to the procedures described in paragraphs 18-22.

12. For testing, U937 cells are seeded at a density of either 3 × 10<sup>5</sup> cells/mL or 6 × 10<sup>5</sup> cells/mL, and pre-cultured in culture flasks for 2 days or 1 day, respectively. Other pre-cultured conditions than those described above may be used if sufficient scientific rationale is provided and if it can be shown to provide similar results, for example by testing the proficiency substances in Appendix II. In the day of testing, cells harvested from culture flask are resuspended with fresh culture medium at 5 × 10<sup>5</sup> cells/mL. Then, cells are distributed into a 96-well flat-bottom plate with 100 µL (final cell density of 0.5 × 10<sup>5</sup> cells/well).

*Preparation of test chemicals and control substances*

13. Assessment of solubility is conducted prior to testing. For this purpose, test chemicals are dissolved or stably dispersed at a concentration of 50 mg/mL in complete medium as first solvent option or dimethyl sulfoxide (DMSO,  $\geq 99\%$  purity) as a second solvent/vehicle option if the test chemical is not soluble in the complete medium solvent/vehicle. For the testing, the test chemical is dissolved to a final concentration of 0.4 mg/mL in complete medium if the chemical is soluble in this solvent/vehicle. If the chemical is soluble only in DMSO, the chemical is dissolved at a concentration of 50 mg/mL. Other solvents/vehicles than those described above may be used if sufficient scientific rationale is provided. Stability of the test chemical in the final solvent/vehicle should be taken into account.

14. The test chemicals and control substances are prepared on the day of testing. Because a dose finding assay is not conducted, for the first run, 6 final concentrations should be tested (1, 10, 20, 50, 100 and 200  $\mu\text{g/mL}$ ) into the corresponding solvent/vehicle either in complete medium or in 0.4% DMSO in medium. For the subsequent runs, starting from the 0.4 mg/mL in complete medium or 50 mg/mL in DMSO, solutions of the test chemicals, at least 4 working solutions (i.e. at least 4 concentrations), are prepared using the corresponding solvent/vehicle. The working solutions are finally used for treatment by adding an equal volume of U937 cell suspension (see paragraph 11 above) to the volume of working solution in the plate to achieve a further 2-fold dilution (12). The concentrations (at least 4 concentrations) for any further run are chosen based on the individual results of all previous runs (8). The usable final concentrations are 1, 2, 3, 4, 5, 7.5, 10, 12.5, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180 and 200  $\mu\text{g/mL}$ . The maximum final concentration is 200  $\mu\text{g/mL}$ . In the case of a CD86 positive value at 1  $\mu\text{g/mL}$  is observed, then 0.1  $\mu\text{g/mL}$  is evaluated in order to find the concentration of the test chemical that does not induce CD86 above the positive threshold. For each run, the EC150 (concentration at which a chemical reaches the CD86 positive threshold of 150%, see paragraph 19) is calculated if a CD86 positive concentration-response is observed. Where the test chemical induces a positive CD86 response not concentration related, the calculation of the EC150 might not be relevant as described in the U-SENS<sup>TM</sup> DB-ALM protocol no. 183 (12). For each run, CV70 (concentration at which a chemical reaches the cytotoxicity threshold of 70%, see paragraph 19) is calculated whenever possible (12). To investigate the concentration response effect of CD86 increase, any concentrations from the usable concentrations should be chosen evenly spread between the EC150 (or the highest CD86 negative non cytotoxic concentration) and the CV70 (or the highest concentration allowed i.e. 200  $\mu\text{g/mL}$ ). A minimum of 4 concentrations should be tested per run with at least 2 concentrations being common with the previous run(s), for comparison purposes.

15. The solvent/vehicle control used in the U-SENS<sup>TM</sup> method is complete medium (for test chemicals solubilised or stably dispersed) (see paragraph 4) or 0.4% DMSO in complete medium (for test chemicals solubilised or stably dispersed in DMSO).

16. The positive control used in the U-SENS<sup>TM</sup> method is TNBS (see paragraph 11), prepared in complete medium. TNBS should be used as the positive control for CD86 expression measurement at a final single concentration in plate (50  $\mu\text{g/mL}$ ) yielding  $> 70\%$  of cell viability. To obtain a 50  $\mu\text{g/mL}$  concentration of TNBS in plate, a 1 M (i.e. 293 mg/mL) stock solution of TNBS in complete medium is prepared and further diluted 2930-fold with complete medium to a 100  $\mu\text{g/mL}$  working solution. Lactic acid (LA, CAS 50-21-5) should be used as the negative control at 200  $\mu\text{g/mL}$  solubilised in complete medium (from a 0.4 mg/mL stock solution). In each plate of each run, three replicates of complete medium untreated control, solvent/vehicle control, negative and positive controls are prepared (12). Other suitable

positive controls may be used if historical data are available to derive comparable run acceptance criteria. The run acceptance criteria are the same as described for the test chemical (see paragraph 12).

#### *Application of test chemicals and control substances*

17. The solvent/vehicle control or working solutions described in paragraphs 14-16 are mixed 1:1 (v/v) with the cell suspensions prepared in the 96-well flat-bottom plate (see paragraph 12). The treated plates are then incubated for  $45 \pm 3$  hours at  $37^\circ\text{C}$  under 5%  $\text{CO}_2$ . Prior to incubation, plates are sealed with semi permeable membrane, to avoid evaporation of volatile test chemicals and cross-contamination between cells treated with test chemicals (12).

#### *Cell staining*

18. After  $45 \pm 3$  hours of exposure, cells are transferred into V-shaped microtiter plate and collected by centrifugation. Solubility interference is defined as crystals or drops observed under the microscope at  $45 \pm 3$  hours post treatment (before the cell staining). The supernatants are discarded and the remaining cells are washed once with 100  $\mu\text{L}$  of an ice-cold phosphate buffered saline (PBS) containing 5 % foetal calf serum (staining buffer). After centrifugation, cells are re-suspended with 100  $\mu\text{L}$  of staining buffer and stained with 5  $\mu\text{L}$  (e.g. 0.25  $\mu\text{g}$ ) of FITC-labelled anti-CD86 or mouse IgG1 (isotype) antibodies at  $4^\circ\text{C}$  for 30 min protected from light. The antibodies described in the U-SENS<sup>TM</sup> DB-ALM protocol no. 183 (12) should be used (for CD86: BD-PharMingen #555657 Clone: Fun-1, or Caltag/Invitrogen # MHCD8601 Clone: BU63; and for IgG1: BD-PharMingen #555748, or Caltag/Invitrogen # GM4992). Based on the experience of the test method developers, the fluorescence intensity of the antibodies is usually consistent between different lots. Other clones or supplier of the antibodies which passed the reactivity check may be used for the assay (see paragraph 11). However, users may consider titrating the antibodies in their own laboratory's conditions to define the best concentration for use. Other detection system e.g. fluorochrome-tagged anti-CD86 antibodies may be used if they can be shown to provide similar results as FITC-conjugated antibodies, for example by testing the proficiency substances in Appendix II. After washing with 100  $\mu\text{L}$  of staining buffer two times and once with 100  $\mu\text{L}$  of an ice-cold PBS, cells are resuspended in ice-cold PBS (e.g. 125  $\mu\text{L}$  for samples being analysed manually tube by tube, or 50  $\mu\text{L}$  using an auto-sampler plate) and PI solution is added (final concentration of 3  $\mu\text{g}/\text{mL}$ ). Other cytotoxicity markers, such as 7-Aminoactinomycin D (7-AAD) or Trypan blue may be used if the alternative stains can be shown to provide similar results as PI, for example by testing the proficiency substances in Appendix II.

#### *Flow cytometry analysis*

19. Expression level of CD86 and cell viability are analysed using flow cytometry. Cells are displayed within a size (FSC) and granularity (SSC) dot plot set to log scale in order to clearly identify the population in a first gate R1 and eliminate the debris. A targeting total of 10,000 cells in gate R1 are acquired for each well. Cells from the same R1 gate are displayed within a FL3 or FL4 / SSC dot plot. Viable cells are delineated by placing a second gate R2 selecting the population of propidium iodide-negative cells (FL3 or FL4 channel). The cell viability can be calculated using the following equation by the cytometer analysis program. When the cell viability is low, up to 20,000 cells including dead cells could be acquired. Alternatively, data can be acquired for one minute after the initiation of the analysis.

$$\text{Cell Viability} = \frac{\text{Number of living cells}}{\text{Total number of acquired cells}} \times 100$$

Percentage of FL1-positive cells is then measured among these viable cells gated on R2 (within R1). Cell surface expression of CD86 is analysed in a FL1 / SSC dot plot gated on viable cells (R2).

For the complete medium / IgG1 wells, the analysis marker is set close to the main population so that the complete medium controls have IgG1 within the target zone of 0.6 to 0.9%.

Colour interference is defined as a shift of the FITC-labelled IgG1 dot-plot (IgG1 FL1 Geo Mean S.I.  $\geq$  150%).

The stimulation index (S.I.) of CD86 for controls cells (untreated or in 0.4% DMSO) and chemical-treated cells are calculated according to the following equation:

$$\text{S.I.} = \frac{\% \text{ of CD86}^+ \text{ treated cells} - \% \text{ of IgG1}^+ \text{ treated cells}}{\% \text{ of CD86}^+ \text{ control cells} - \% \text{ of IgG1}^+ \text{ control cells}} \times 100$$

% of IgG1<sup>+</sup> untreated control cells: referred to as percentage of FL1-positive IgG1 cells defined with the analysis marker (accepted range of  $\geq$  0.6% and  $<$  1.5%, see paragraph 22) among the viable untreated cells.

% of IgG1<sup>+</sup>/CD86<sup>+</sup> control/treated cells: referred to as percentage of FL1-positive IgG1/CD86 cells measured without moving the analysis marker among the viable control/treated cells.

## DATA AND REPORTING

### *Data evaluation*

20. The following parameters are calculated in the U-SENS<sup>TM</sup> test method: CV70 value, i.e. a concentration showing 70% of U937 cell survival (30% cytotoxicity) and the EC150 value, i.e. the concentration at which the test chemicals induced a CD86 stimulation index (S.I.) of 150%.

CV70 is calculated by log-linear interpolation using the following equation:

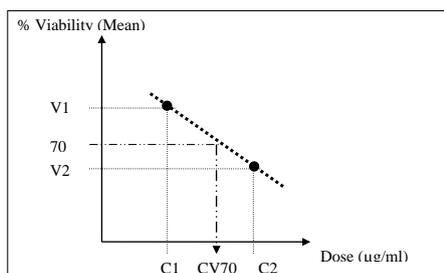
$$\text{CV70} = C1 + [(V1 - 70) / (V1 - V2) * (C2 - C1)]$$

Where:

V1 is the minimum value of cell viability over 70%

V2 is the maximum value of cell viability below 70%

C1 and C2 are the concentrations showing the value of cell viability V1 and V2 respectively.



Other approaches to derive the CV70 can be used as long as it is demonstrated that this has no impact on the results (e.g. by testing the proficiency substances).

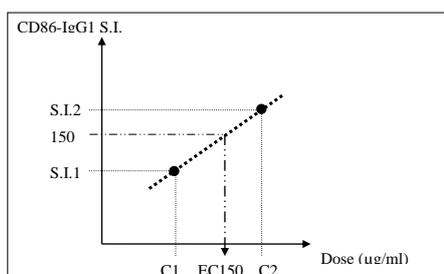
EC150 is calculated by log-linear interpolation using the following equation:

$$EC150 = C1 + [(150 - S.I.1) / (S.I.2 - S.I.1) * (C2 - C1)]$$

Where:

C1 is the highest concentration in µg/mL with a CD86 S.I. < 150% (S.I. 1)

C2 is the lowest concentration in µg/mL with a CD86 S.I. ≥ 150% (S.I. 2).



The EC150 and CV70 values are calculated

- for each run : the individual EC150 and CV70 values are used as tools to investigate the concentration response effect of CD86 increase (see paragraph 14),
- based on the average viabilities, the overall CV70 is determined (12) ,
- based on the average S.I. of CD86 values, the overall EC150 is determined for the test chemical predicted as POSITIVE with the U-SENS<sup>TM</sup> (see paragraph 21) (12).

### **Prediction model**

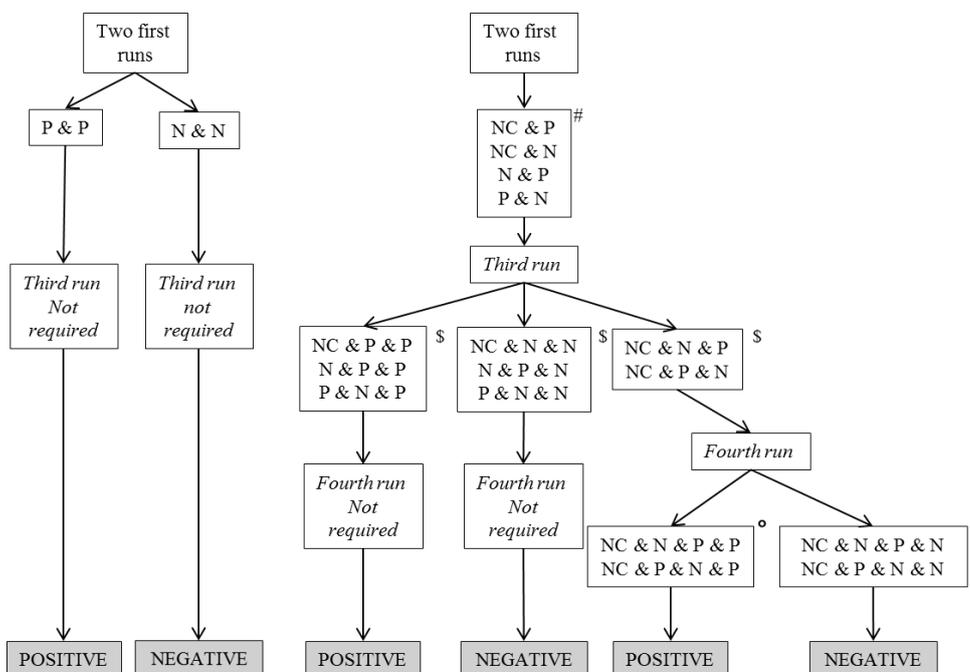
21. For CD86 expression measurement, each test chemical is tested in at least four concentrations and in at least two independent runs (performed on a different day) to derive a single prediction (NEGATIVE or POSITIVE).

- The individual conclusion of an U-SENS<sup>TM</sup> run is considered Negative (hereinafter referred to as N) if the S.I. of CD86 is less than 150% at all non-cytotoxic concentrations (cell viability ≥ 70%) and if no interference is observed (cytotoxicity, solubility: see paragraph 18 or colour: see paragraph 19 regardless of the non-cytotoxic concentrations at which the interference is detected). In all other cases: S.I. of CD86 higher or equal to 150% and/or interferences observed, the individual conclusion of an U-SENS<sup>TM</sup> run is considered Positive (hereinafter referred to as P).

- An U-SENS™ prediction is considered NEGATIVE if at least two independent runs are negative (N) (Figure 1). If the first two runs are both negative (N), the U-SENS™ prediction is considered NEGATIVE and a third run does not need to be conducted.

- An U-SENS™ prediction is considered POSITIVE if at least two independent runs are positive (P) (Figure 1). If the first two runs are both positive (P), the U-SENS™ prediction is considered POSITIVE and a third run does not need to be conducted.

- Because a dose finding assay is not conducted, there is an exception if, in the first run, the S.I. of CD86 is higher or equal to 150% at the highest non-cytotoxic concentration only. The run is then considered to be NOT CONCLUSIVE (NC), and additional concentrations (between the highest non cytotoxicity concentration and the lowest cytotoxicity concentration - see paragraph 20) should be tested in additional runs. In case a run is identified as NC, at least 2 additional runs should be conducted, and a fourth run in case runs 2 and 3 are not concordant (N and/or P independently) (Figure 1). Follow up runs will be considered positive even if only one non cytotoxic concentration gives a CD86 equal or above 150%, since the concentration setting has been adjusted for the specific test chemical. The final prediction will be based on the majority result of the three or four individual runs (i.e. 2 out of 3 or 2 out of 4) (Figure 1).



**Figure 1: Prediction model used in the U-SENS™ test method.** An U-SENS™ prediction should be considered in the framework of an IATA and in accordance with the provision of paragraph 4 and of the General introduction paragraphs 7, 8 and 9.

N: Run with no CD86 positive or interference observed;  
 P: Run with CD86 positive and/or interference(s) observed;  
 NC: Not Conclusive. First run with No Conclusion when CD86 is positive at the highest non-cytotoxic concentration only;

#: A Not Conclusive (NC) individual conclusion attributed only to the first run conducts automatically to the need of a third run to reach a majority of Positive (P) or Negative (N) conclusions in at least 2 of 3 independent runs.

§: The boxes show the relevant combinations of results from the three runs on the basis of the results obtained in the first two runs shown in the box above.

°: The boxes show the relevant combinations of results from the four runs on the basis of the results obtained in the first three runs shown in the box above.

### *Acceptance criteria*

22. The following acceptance criteria should be met when using the U-SENS™ method (12).

- At the end of the  $45 \pm 3$  hours exposure period, the mean viability of the triplicate untreated U937 cells had to be  $> 90\%$  and no drift in CD86 expression is observed. The CD86 basal expression of untreated U937 cells had to be comprised within the range of  $\geq 2\%$  and  $\leq 25\%$ .
- When DMSO is used as a solvent, the validity of the DMSO vehicle control is assessed by calculating a DMSO S.I. compared to untreated cells, and the mean viability of the triplicate cells had to be  $> 90\%$ . The DMSO vehicle control is valid if the mean value of its triplicate CD86 S.I. was smaller than 250% of the mean of the triplicate CD86 S.I. of untreated U937 cells.
- The runs are considered valid if at least two out of three IgG1 values of untreated U937 cells fell within the range of  $\geq 0.6\%$  and  $< 1.5\%$ .
- The concurrent tested negative control (lactic acid) is considered valid if at least two out of the three replicates were negative (CD86 S.I.  $< 150\%$ ) and non-cytotoxic (cell viability  $\geq 70\%$ ).
- The positive control (TNBS) was considered as valid if at least two out of the three replicates were positive (CD86 S.I.  $\geq 150\%$ ) and non-cytotoxic (cell viability  $\geq 70\%$ ).

### *Test report*

23. The test report should include the following information.

#### *Test Chemical*

- Mono-constituent substance
  - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
  - Physical appearance, complete medium solubility, DMSO solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;

- Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available;
  - Justification for choice of solvent/vehicle for each test chemical.
- Multi-constituent substance, UVCB and mixture:
- Characterisation as far as possible by e.g. chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available;
  - Physical appearance, complete medium solubility, DMSO solubility and additional relevant physicochemical properties, to the extent available;
  - Molecular weight or apparent molecular weight in case of mixtures/polymers of known compositions or other information relevant for the conduct of the study;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available;
  - Justification for choice of solvent/vehicle for each test chemical.

### *Controls*

- Positive control
- Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
  - Physical appearance, DMSO solubility, molecular weight, and additional relevant physicochemical properties, to the extent available and where applicable;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available;
  - Reference to historical positive control results demonstrating suitable run acceptance criteria, if applicable.
- Negative and solvent/vehicle control
- Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;

- Physical appearance, molecular weight, and additional relevant physicochemical properties in the case other control solvent/vehicle than those mentioned in the Test Guideline are used and to the extent available;
- Storage conditions and stability to the extent available;
- Justification for choice of solvent/vehicle for each test chemical.

#### *Test method Conditions*

- Name and address of the sponsor, test facility and study director;
- Description of test method used;
- Cell line used, its storage conditions and source (e.g. the facility from which they were obtained);
- Flow cytometry used (e.g. model), including instrument settings, antibodies and cytotoxicity marker used;
- The procedure used to demonstrate proficiency of the laboratory in performing the test method by testing of proficiency substances, and the procedure used to demonstrate reproducible performance of the test method over time, e.g. historical control data and/or historical reactivity checks' data.

#### *Test Acceptance Criteria*

- Cell viability and CD86 S.I values obtained with the solvent/vehicle control in comparison to the acceptance ranges;
- Cell viability and S.I. values obtained with the positive control in comparison to the acceptance ranges;
- Cell viability of all tested concentrations of the tested chemical.

#### *Test procedure*

- Number of runs used;
- Test chemical concentrations, application and exposure time used (if different than the one recommended)
- Duration of exposure;
- Description of evaluation and decision criteria used;
- Description of any modifications of the test procedure.

#### *Results*

- Tabulation of the data, including CV70 (if applicable), S.I., cell viability values, EC150 values (if applicable) obtained for the test chemical and for the positive control in each run, and an indication of the rating of the test chemical according to the prediction model;
- Description of any other relevant observations, if applicable.

## *Discussion of the Results*

- Discussion of the results obtained with the U-SENS™ method;
- Consideration of the test method results within the context of an IATA, if other relevant information is available.

## *Conclusions*

## LITERATURE

1. Piroird, C., Ovigne, J.M., Rousset, F., Martinozzi-Teissier, S., Gomes, C., Cotovio, J., Alépée, N. (2015). The Myeloid U937 Skin Sensitization Test (U-SENS) addresses the activation of dendritic cell event in the adverse outcome pathway for skin sensitization. *Toxicol. In Vitro* 29, 901-916.
2. EURL ECVAM (2017). The U-SENS™ test method Validation Study Report. Accessible at: [http://ihcp.jrc.ec.europa.eu/our\\_labs/eurl-ecvam/eurl-ecvam-recommendations](http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations)
3. EC EURL ECVAM (2016). ESAC Opinion No. 2016-03 on the L'Oréal-coordinated study on the transferability and reliability of the U-SENS™ test method for skin sensitisation testing. EUR 28178 EN; doi 10.2787/815737. Available at: [http://publications.jrc.ec.europa.eu/repository/handle/JRC103705].
4. EC EURL ECVAM (2017). EURL ECVAM Recommendation on the use of non-animal approaches for skin sensitisation testing. EUR 28553 EN; doi 10.2760/588955. Available at: [https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/eurl-ecvam-recommendation-use-non-animal-approaches-skin-sensitisation-testing].
5. Steiling, W. (2016). Safety Evaluation of Cosmetic Ingredients Regarding their Skin Sensitization Potential. doi:10.3390/cosmetics3020014. *Cosmetics* 3, 14.
6. OECD (2016). Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated Approaches To Testing And Assessment (IATA) For Skin Sensitisation, Series on Testing & Assessment No. 256, ENV/JM/MONO(2016)29. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm].
7. Urbisch, D., Mehling, A., Guth, K., Ramirez, T., Honarvar, N., Kolle, S., Landsiedel, R., Jaworska, J., Kern, P.S., Gerberick, F., Natsch, A., Emter, R., Ashikaga, T., Miyazawa, M., Sakaguchi, H. (2015). Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul. Toxicol. Pharmacol.* 71, 337-351.
8. Alépée, N., Piroird, C., Aujoulat, M., Dreyfuss, S., Hoffmann, S., Hohenstein, A., Meloni, M., Nardelli, L., Gerbeix, C., Cotovio, J. (2015). Prospective multicentre study of the U-SENS test method for skin sensitization testing. *Toxicol In Vitro* 30, 373-382.
9. Reisinger, K., Hoffmann, S., Alépée, N., Ashikaga, T., Barroso, J., Elcombe, C., Gellatly, N., Galbiati, V., Gibbs, S., Groux, H., Hibatallah, J., Keller, D., Kern, P., Klaric, M., Kolle, S., Kuehnl, J., Lambrechts, N., Lindstedt, M., Millet, M., Martinozzi-Teissier, S., Natsch, A., Petersohn, D., Pike, I., Sakaguchi, H., Schepky, A., Tailhardat, M., Templier, M., van Vliet, E., Maxwell, G. (2014). Systematic evaluation of non-animal test methods for skin sensitisation safety assessment. *Toxicol. In Vitro* 29, 259-270.
10. Fabian, E., Vogel, D., Blatz, V., Ramirez, T., Kolle, S., Eltze, T., van Ravenzwaay, B., Oesch, F., Landsiedel, R. (2013). Xenobiotic metabolizing enzyme activities in cells used for testing skin sensitization *in vitro*. *Arch. Toxicol.* 87, 1683-1696.
11. OECD. (2017). Draft Guidance document: Good *In Vitro* Method Practices (GIVIMP) for the Development and Implementation of *In Vitro* Methods for Regulatory Use in Human Safety Assessment. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/ehs/testing/OECD%20Draft%20GIVIMP\_v05%20-%20clean.pdf].
12. DB-ALM (2016). Protocol no. 183: Myeloid U937 Skin Sensitization Test (U-SENS™), 33pp.

Accessible at: [<http://ecvam-dbalm.jrc.ec.europa.eu/>].

13. Sundström, C., Nilsson, K. (1976). Establishment and characterization of a human histiocytic lymphoma cell line (U-937). *Int. J. Cancer* 17, 565-577.
14. OECD (2005). Series on Testing and Assessment No. 34: Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Organisation for Economic Cooperation and Development, Paris. Available at: [<http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm>].
15. United Nations UN (2015). Globally Harmonized System of Classification and Labelling of Chemicals (GHS). ST/SG/AC.10/30/Rev.6, Sixth Revised Edition, New York & Geneva: United Nations Publications. Available at: [[http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs\\_rev06/English/ST-SG-AC10-30-Rev6e.pdf](http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev06/English/ST-SG-AC10-30-Rev6e.pdf)].
16. OECD (2012). Series on Testing and Assessment No. 168: The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Organisation for Economic Cooperation and Development, Paris. Available at: [<http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm>].
17. ECETOC (2003). Technical Report No. 87: Contact sensitization: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals, Brussels. Available at: [[https://ftp.cdc.gov/pub/Documents/OEL/06.%20Dotson/References/ECETOC\\_2003-TR87.pdf](https://ftp.cdc.gov/pub/Documents/OEL/06.%20Dotson/References/ECETOC_2003-TR87.pdf)].

APPENDIX IDEFINITIONS

**Accuracy:** The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with concordance to mean the proportion of correct outcomes of a test method (14).

**AOP (Adverse Outcome Pathway):** sequence of events from the chemical structure of a target chemical or group of similar chemicals through the molecular initiating event to an *in vivo* outcome of interest (15).

**CD86 Concentration response:** There is concentration-dependency (or concentration response) when a positive concentration (CD86 S.I.  $\geq 150$ ) is followed by a concentration with an increasing CD86 S.I.

**CV70:** The estimated concentration showing 70% cell viability.

**Drift:** A drift is defined by i) the corrected %CD86<sup>+</sup> value of the untreated control replicate 3 is less than 50% of the mean of the corrected %CD86<sup>+</sup> value of untreated control replicates 1 and 2; and ii) the corrected %CD86<sup>+</sup> value of the negative control replicate 3 is less than 50% of mean of the corrected %CD86<sup>+</sup> value of negative control replicates 1 and 2.

**EC150:** the estimated concentrations showing the 150% S.I. of CD86 expression.

**Flow cytometry:** a cytometric technique in which cells suspended in a fluid flow one at a time through a focus of exciting light, which is scattered in patterns characteristic to the cells and their components; cells are frequently labeled with fluorescent markers so that light is first absorbed and then emitted at altered frequencies.

**Hazard:** Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub) population is exposed to that agent.

**IATA (Integrated Approach to Testing and Assessment):** A structured approach used for hazard identification (potential), hazard characterisation (potency) and/or safety assessment (potential/potency and exposure) of a chemical or group of chemicals, which strategically integrates and weights all relevant data to inform regulatory decision regarding potential hazard and/or risk and/or the need for further targeted and therefore minimal testing.

**Mixture:** A mixture or a solution composed of two or more substances in which they do not react.

**Mono-constituent substance:** A substance, defined by its quantitative composition, in which one main constituent is present to at least 80% (w/w).

**Multi-constituent substance:** A substance, defined by its quantitative composition, in which more than

one main constituent is present in a concentration  $\geq 10\%$  (w/w) and  $< 80\%$  (w/w). A multi-constituent substance is the result of a manufacturing process. The difference between mixture and multi-constituent substance is that a mixture is obtained by blending of two or more substances without chemical reaction. A multi-constituent substance is the result of a chemical reaction.

**Positive control:** A replicate containing all components of a test system and treated with a substance known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the positive response should not be excessive.

**Pre-haptens:** chemicals which become sensitisers through abiotic transformation, e.g. through oxidation.

**Pro-haptens:** chemicals requiring enzymatic activation to exert skin sensitisation potential.

**Relevance:** Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (14).

**Reliability:** Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability (14).

**Run:** A run consists of one or more test chemicals tested concurrently with a solvent/vehicle control and with a positive control.

**Sensitivity:** The proportion of all positive/active chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (14).

**S.I.:** Stimulation Index. Relative values of geometric mean fluorescence intensity in chemical-treated cells compared to solvent-treated cells.

**Solvent/vehicle control:** An untreated sample containing all components of a test system except of the test chemical, but including the solvent/vehicle that is used. It is used to establish the baseline response for the samples treated with the test chemical dissolved or stably dispersed in the same solvent/vehicle. When tested with a concurrent medium control, this sample also demonstrates whether the solvent/vehicle interacts with the test system.

**Specificity:** The proportion of all negative/inactive chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method (14).

**Staining buffer:** A phosphate buffered saline containing 5% foetal calf serum.

**Substance:** Chemical elements and their compounds in the natural state or obtained by any production process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition.

**Test chemical:** The term "test chemical" is used to refer to what is being tested.

**United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS):** A system proposing the classification of chemicals (substances and mixtures) according to standardized types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (16).

**UVCB:** substances of unknown or variable composition, complex reaction products or biological materials.

**Valid test method:** A test method considered to have sufficient relevance and reliability for a specific purpose and which is based on scientifically sound principles. A test method is never valid in an absolute sense, but only in relation to a defined purpose (14).

APPENDIX IIPROFICIENCY SUBSTANCES

Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency by correctly obtaining the expected U-SENS™ prediction for the 10 substances recommended in Table 1 and by obtaining CV70 and EC150 values that fall within the respective reference range for at least 8 out of the 10 proficiency substances. Proficiency substances were selected to represent the range of responses for skin sensitisation hazards. Other selection criteria were that the substances are commercially available, and that high-quality *in vivo* reference data as well as high quality *in vitro* data generated with the U-SENS™ method are available. Also, published reference data are available for the U-SENS™ method (1) (8).

Table 1: Recommended substances for demonstrating technical proficiency with the U-SENS™ method

Proficiency substances	CASRN	Physical state	<i>In vivo</i> prediction <sup>1</sup>	U-SENS™ Solvent/ Vehicle	U-SENS™ CV70 Reference Range in µg/mL <sup>2</sup>	U-SENS™ EC150 Reference Range in µg/mL <sup>2</sup>
4-Phenylenediamine	106-50-3	Solid	Sensitiser (strong)	Complete medium <sup>3</sup>	<30	Positive (≤10)
Picryl sulfonic acid	2508-19-2	Liquid	Sensitizer (strong)	Complete medium	>50	Positive (≤50)
Diethyl maleate	141-05-9	Liquid	Sensitiser (moderate)	DMSO	10-100	Positive (≤20)
Resorcinol	108-46-3	Solid	Sensitiser (moderate)	Complete medium	>100	Positive (≤50)
Cinnamic alcohol	104-54-1	Solid	Sensitiser (weak)	DMSO	>100	Positive (10-100)
4-Allylanisole	140-67-0	Liquid	Sensitiser (weak)	DMSO	>100	Positive (<200)
Saccharin	81-07-2	Solid	Non-sensitiser	DMSO	>200	Negative (>200)
Glycerol	56-81-5	Liquid	Non-sensitiser	Complete medium	>200	Negative (>200)
Lactic acid	50-21-5	Liquid	Non-sensitiser	Complete medium	>200	Negative (>200)
Salicylic acid	69-72-7	Solid	Non-sensitiser	DMSO	>200	Negative (>200)

Abbreviations: CAS RN = Chemical Abstracts Service Registry Number

<sup>1</sup> The *in vivo* hazard and (potency) prediction is based on LLNA data (1) (8). The *in vivo* potency is derived using the criteria proposed by ECETOC (17).

<sup>2</sup> Based on historical observed values (1) (8).

<sup>3</sup> Complete medium: RPMI-1640 medium supplemented with 10% foetal calf serum, 2 mM L-glutamine, 100 units/mL penicillin and 100 µg/mL streptomycin (8).

### **Annex III: *In Vitro* Skin Sensitisation: IL-8 Luc assay**

#### **INITIAL CONSIDERATIONS AND LIMITATIONS**

1. In contrast to assays analysing the expression of cell surface markers, the IL-8-Luc assay quantifies changes in IL-8 expression, a cytokine associated with the activation of dendritic cells (DC). In the THP-1-derived IL-8 reporter cell line (THP-G8, established from the human acute monocytic leukemia cell line THP-1), IL-8 expression is measured following exposure to sensitisers (1). The expression of luciferase is then used to aid discrimination between skin sensitisers and non-sensitisers.

2. The IL-8 Luc method has been evaluated in a validation study (2) conducted by the Japanese Centre for the Validation of Alternatives Methods (JaCVAM), the Ministry of Economy, Trade and Industry (METI), and the Japanese Society for Alternatives to Animal Experiments (JSAAE) and subsequently subjected to independent peer review (3) under the auspices of JaCVAM and the Ministry of Health, Labour and Welfare (MHLW) with the support of the International Cooperation on Alternative Test Methods (ICATM). Considering all available evidence and input from regulators and stakeholders, the IL-8 Luc assay is considered useful as part of IATA to discriminate sensitisers from non-sensitisers for the purpose of hazard classification and labelling. Examples of the use of IL-8 Luc assay data in combination with other information are reported in the literature (4) (5) (6).

3. The IL-8 Luc assay proved to be transferable to laboratories experienced in cell culture and luciferase measurement. Within and between laboratory reproducibilities were 87.7% and 87.5%, respectively (2). Data generated in the validation study (2) and other published work (1) (6) show that versus the LLNA, the IL-8 Luc assay judged 118 out of 143 chemicals as positive or negative and judged 25 chemicals as inconclusive and the accuracy of the IL-8 Luc assay in distinguishing skin sensitisers (UN GHS Cat. 1) from non-sensitisers (UN GHS No Cat.) is 86% (101/118) with a sensitivity of 96% (92/96) and specificity of 41% (9/22). Excluding substances outside the applicability domain described below (paragraph 5), the IL-8 Luc assay judged 113 out of 136 chemicals as positive or negative and judged 23 chemicals as inconclusive and the accuracy of the IL-8 Luc assay is 89% (101/113) with sensitivity of 96% (92/96) and specificity of 53% (9/17). Using human data cited in Urbisch et al. (7), the IL-8 Luc assay judged 76 out of 90 chemicals as positive or negative and judged 14 chemicals as inconclusive and the accuracy is 80% (61/76), sensitivity is 93% (54/58) and specificity is 39% (7/18). Excluding substances outside the applicability domain, the IL-8 Luc assay judged 71 out of 84 chemicals as positive or negative and judged 13 chemicals as inconclusive and the accuracy is 86% (61/71) with sensitivity of 93% (54/58) and specificity of 54% (7/13). False negative predictions with the IL-8 Luc assay are more likely to occur with chemicals showing low/moderate skin sensitisation potency (UN GHS subcategory 1B) than those with high potency (UN GHS subcategory 1A) (6). Together, the information supports a role for the IL-8 Luc assay in the identification of skin sensitisation hazards. The accuracy given for the IL-8 Luc assay as a standalone test method is only for guidance, as the method should be considered in combination with other sources of information in the context of an IATA and in accordance with the provisions of paragraphs 7 and 8 in the General introduction. Furthermore, when evaluating non-animal methods for skin sensitisation, it should be remembered that the LLNA and other animal tests may not fully reflect the situation in humans.

4. On the basis of the data currently available, the IL-8 Luc assay was shown to be applicable to test chemicals covering a variety of organic functional groups, reaction mechanisms, skin sensitisation potency (as determined in *in vivo* studies) and physicochemical properties (2) (6).

5. Although the IL-8 Luc assay uses X-VIVO™ 15 as a solvent, it correctly evaluated chemicals with a Log  $K_{ow}$  >3.5 and those with a water solubility of around 100 µg/ mL as calculated by EPI Suite™ and its performance to detect sensitisers with poor water solubility is better than that of the IL-8 Luc assay using dimethyl sulfoxide (DMSO) as a solvent (2). However, negative results for test chemicals that are not dissolved at 20 mg/ml may produce false negative results due to their inability to dissolve in X-VIVO™ 15. Therefore, negative results for these chemicals should not be considered. A high false negative rate for anhydrides was seen in the validation study. Furthermore, because of the limited metabolic capability of the cell line (8) and the experimental conditions, pro-haptens (substances requiring metabolic activation) and pre-haptens (substances activated by air oxidation) might give negative results in the assay. However, although negative results for suspected pre/prohaptens should be interpreted with caution, the IL-8 Luc assay correctly judged 11 out of 11 pre-haptens, 6/6 pro-haptens, and 6/8 pre/pro-haptens in the IL-8 Luc assay data set (2). Based on the recent comprehensive review on three non-animal methods (the DPRA, the KeratinoSens™ and the h-CLAT) to detect pre and prohaptens (9), and based on the fact that THP-G8 cells used in the IL-8 Luc assay is a cell line derived from THP-1 that is used in the h-CLAT, the IL-8 Luc assay may also contribute to increase the sensitivity of non-animal methods to detect pre and pro-haptens in the combination of other methods. Surfactants tested so far gave (false) positive results irrespective of their type (e.g. cationic, anionic or on-ionic). Finally, chemicals that interfere with luciferase can confound its activity/measurement, causing apparent inhibition or increased luminescence (10). For example, phytoestrogen concentrations higher than 1µM were reported to interfere with luminescence signals in other luciferase-based reporter gene assays due to over-activation of the luciferase reporter gene. Consequently, luciferase expression obtained at high concentrations of phytoestrogens or compounds suspected of producing phytoestrogen-like activation of the luciferase reporter gene needs to be examined carefully (11). Based on the above, surfactants, anhydrides and chemicals interfering with luciferase are outside the applicability domain of this assay. In cases where there is evidence demonstrating the non-applicability of the IL-8 Luc assay to other specific categories of test chemicals, the method should not be used for those specific categories.

6. As described above, the IL-8 Luc assay supports discrimination of skin sensitisers from non-sensitisers. Further work, preferably based on human data, is required to determine whether IL-8 Luc results can contribute to potency assessment when considered in combination with other information sources.

7. Definitions are provided in Appendix I.

## PRINCIPLE OF THE TEST

8. The IL-8 Luc assay makes use of a human monocytic leukemia cell line THP-1 that was obtained from the American Type Culture Collection (Manassas, VA, USA). Using this cell line, the Dept. of Dermatology, Tohoku University School of Medicine, established a THP-1-derived IL-8 reporter cell line, THP-G8, that harbours the Stable Luciferase Orange (SLO) and Stable Luciferase Red (SLR) luciferase genes under the control of the IL-8 and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) promoters, respectively (1). This allows quantitative measurement of luciferase gene induction by detecting

luminescence from well-established light producing luciferase substrates as an indicator of the activity of the IL-8 and GAPDH in cells following exposure to sensitising chemicals.

9. The dual-colour assay system comprises an orange-emitting luciferase (SLO;  $\lambda_{\text{max}} = 580 \text{ nm}$ ) (12) for the gene expression of the IL-8 promoter as well as a red-emitting luciferase (SLR;  $\lambda_{\text{max}} = 630 \text{ nm}$ ) (13) for the gene expression of the internal control promoter, GAPDH. The two luciferases emit different colours upon reacting with firefly D-luciferin and their luminescence is measured simultaneously in a one-step reaction by dividing the emission from the assay mixture using an optical filter (14) (Appendix II).

10. THP-G8 cells are treated for 16 hours with the test chemical, after which SLO luciferase activity (SLO-LA) reflecting IL-8 promoter activity and SLR luciferase activity (SLR-LA) reflecting GAPDH promoter activity are measured. To make the abbreviations easy to understand, SLO-LA and SLR-LA are designated as IL8LA and GAPLA, respectively. Table 1 gives a description of the terms associated with luciferase activity in the IL-8 Luc assay. The measured values are used to calculate the normalised IL8LA (nIL8LA), which is the ratio of IL8LA to GAPLA; the induction of nIL8LA (Ind-IL8LA), which is the ratio of the arithmetic means of quadruple-measured values of the nIL8LA of THP-G8 cells treated with a test chemical and the values of the nIL8LA of untreated THP-G8 cells; and the inhibition of GAPLA (Inh-GAPLA), which is the ratio of the arithmetic means of quadruple-measured values of the GAPLA of THP-G8 cells treated with a test chemical and the values of the GAPLA of untreated THP-G8 cells, and used as an indicator for cytotoxicity.

Table 1. Description of terms associated with the luciferase activity in the IL-8 Luc assay

Abbreviations	Definition
GAPLA	SLR luciferase activity reflecting GAPDH promoter activity
IL8LA	SLO luciferase activity reflecting IL-8 promoter activity
nIL8LA	IL8LA / GAPLA
Ind-IL8LA	nIL8LA of THP-G8 cells treated with chemicals / nIL8LA of untreated cells
Inh-GAPLA	GAPLA of THP-G8 treated with chemicals / GAPLA of untreated cells
CV05	The lowest concentration of the chemical at which Inh-GAPLA becomes < 0.05.

11. Performance standards (PS) (15) are available to facilitate the validation of modified *in vitro* IL-8 luciferase test methods similar to the IL-8 Luc assay and allow for timely amendment of this Test Guideline for their inclusion. Mutual Acceptance of Data (MAD) will only be guaranteed for test methods validated according to the PS, if these test methods have been reviewed and included in this Test Guideline by the OECD (16).

#### DEMONSTRATION OF PROFICIENCY

12. Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency, using the 10 Proficiency Substances listed in Appendix III in compliance with the Good *in vitro* Method Practices (17). Moreover, test method users should maintain a historical database of data generated with the reactivity checks (see paragraph 15) and

with the positive and solvent/vehicle controls (see paragraphs 21-24), and use these data to confirm the reproducibility of the test method in their laboratory is maintained over time.

## PROCEDURE

13. The Standard Operating Procedure (SOP) for the IL-8 Luc assay is available and should be employed when performing the test (18). Laboratories willing to perform the test can obtain the recombinant THP-G8 cell line from GPC Lab. Co. Ltd., Tottori, Japan, upon signing a Material Transfer Agreement (MTA) in line with the conditions of the OECD template. The following paragraphs provide a description of the main components and procedures of the assay.

### Preparation of cells

14. The THP-G8 cell line from GPC Lab. Co. Ltd., Tottori, Japan, should be used for performing the IL-8 Luc assay (see paragraphs 8 and 13). On receipt, cells are propagated (2-4 passages) and stored frozen as a homogeneous stock. Cells from this stock can be propagated up to a maximum of 12 passages or a maximum of 6 weeks. The medium used for propagation is the RPMI-1640 culture medium containing 10% foetal bovine serum (FBS), antibiotic/antimycotic solution (100U/mL of penicillin G, 100µg/mL of streptomycin and 0.25µg/mL of amphotericin B in 0.85% saline) (e.g. GIBCO Cat#15240-062), 0.15µg/mL Puromycin (e.g. CAS:58-58-2) and 300µg/mL G418 (e.g. CAS:108321-42-2).

15. Prior to use for testing, the cells should be qualified by conducting a reactivity check. This check should be performed 1-2 weeks or 2-4 passages after thawing, using the positive control, 4-nitrobenzyl bromide (4-NBB) (CAS:100-11-8, ≥ 99% purity) and the negative control, lactic acid (LA) (CAS:50-21-5, ≥85% purity). 4-NBB should produce a positive response to Ind-IL8LA (≥1.4), while LA should produce a negative response to Ind-IL8LA (<1.4). Only cells that pass the reactivity check are used for the assay. The check should be performed according to the procedures described in paragraphs 22-24.

16. For testing, THP-G8 cells are seeded at a density of 2 to 5 × 10<sup>5</sup> cells/mL, and pre-cultured in culture flasks for 48 to 96 hours. On the day of the test, cells harvested from the culture flask are washed with RPMI-1640 containing 10% FBS without any antibiotics, and then, resuspended with RPMI-1640 containing 10% FBS without any antibiotics at 1 × 10<sup>6</sup> cells/mL. Then, cells are distributed into a 96-well flat-bottom black plate (e.g. Costar Cat#3603) with 50µL (5 × 10<sup>4</sup> cells/well).

### Preparation of the test chemical and control substances

17. The test chemical and control substances are prepared on the day of testing. For the IL-8 Luc assay, test chemicals are dissolved in X-VIVO™ 15, a commercially available serum-free medium (Lonza, 04-418Q), to the final concentration of 20 mg/mL. X-VIVO™ 15 is added to 20 mg of test chemical (regardless of the chemical's solubility) in a microcentrifuge tube and brought to a volume of 1mL and then vortexed vigorously and shaken on a rotor at a maximum speed of 8 rpm for 30 min at an ambient temperature of about 20°C. Furthermore, if solid chemicals are still insoluble, the tube is sonicated until the chemical is dissolved completely or stably dispersed. For test chemicals soluble in X-VIVO™ 15, the solution is diluted by a factor of 5 with X-VIVO™ 15 and used as an X-VIVO™ 15 stock solution of the test chemical (4 mg/mL). For test chemicals not soluble in X-VIVO™ 15, the mixture is rotated again for at least 30 min, then centrifuged at 15,000 rpm (≈20,000g) for 5 min; the resulting supernatant is used as an X-VIVO™ 15 stock solution of the test chemical. A scientific rationale should be provided for the use of other solvents, such as DMSO, water, or the culture medium. The detailed procedure for

dissolving chemicals is shown in Appendix V. The X-VIVO™ 15 solutions described in paragraphs 18-23 are mixed 1:1 (v/v) with the cell suspensions prepared in a 96-well flat-bottom black plate (see paragraph 16).

18. The first test run is aimed to determine the cytotoxic concentration and to examine the skin sensitising potential of chemicals. Using X-VIVO™ 15, serial dilutions of the X-VIVO™ 15 stock solutions of the test chemicals are made at a dilution factor of two (see Appendix V) using a 96-well assay block (e.g. Costar Cat#EW-01729-03). Next, 50 µl/well of diluted solution is added to 50 µl of the cell suspension in a 96-well flat-bottom black plate. Thus for test chemicals that are soluble in X-VIVO™ 15, the final concentrations of the test chemicals range from 0.002 to 2 mg/mL (Appendix V). For test chemicals that are not soluble in X-VIVO™ 15 at 20 mg/mL, only dilution factors that range from 2 to 2<sup>10</sup>, are determined, although the actual final concentrations of the test chemicals remain uncertain and are dependent on the saturated concentration of the test chemicals in the X-VIVO™ 15 stock solution.

19. In subsequent test runs (i.e. the second, third, and fourth replicates), the X-VIVO™ 15 stock solution is made at the concentration 4 times higher than the concentration of cell viability 05 (CV05; the lowest concentration at which the Inh-GAPLA becomes <0.05) in the first experiment. If Inh-GAPLA does not decrease below 0.05 at the highest concentration in the first run, the X-VIVO™ 15 stock solution is made at the first run highest concentration. The concentration of CV05 is calculated by dividing the concentration of the stock solution in the first run by dilution factor for CV05 (X) (dilution factor CV05 (X)); the dilution factor required to dilute stock solution to CV05 (see Appendix V). For test substances not soluble in X-VIVO at 20 mg/ml, CV05 is determined by the concentration of the stock solution x 1/X. For run 2 to 4, a second stock solution is prepared as 4 x CV50 (Appendix V).

20. Serial dilutions of the X-VIVO™ 15 second stock solutions are made at a dilution factor of 1.5 using a 96-well assay block. Next, 50 µl/well of diluted solution is added to 50 µl of the cell suspension in the wells of a 96-well flat-bottom black plate. Each concentration of each test chemical should be tested in 4 wells. The samples are then mixed on a plate shaker and incubated for 16 hours at 37°C and 5% CO<sub>2</sub>, after which the luciferase activity is measured as described below.

21. The solvent control is the mixture of 50 µL/well of X-VIVO™ 15 and 50 µL/well of cell suspension in RPMI-1640 containing 10% FBS.

22. The recommended positive control is 4-NBB. 20 mg of 4-NBB is prepared in a 1.5-mL microfuge tube, to which X-VIVO™ 15 is added up to 1 mL. The tube is vortexed vigorously and shaken on a rotor at a maximum speed of 8 rpm for at least 30 min. After centrifugation at 20,000g for 5 min, the supernatant is diluted by a factor of 4 with X-VIVO™ 15, and 500 µl of the diluted supernatant is transferred to a well in a 96-well assay block. The diluted supernatant is further diluted with X-VIVO™ 15 at factors of 2 and 4, and 50 µl of the solution is added to 50 µl of THP-G8 cell suspension in the wells of a 96-well flat-bottom black plate (Appendix VI). Each concentration of the positive control should be tested in 4 wells. The plate is agitated on a plateshaker, and incubated in a CO<sub>2</sub> incubator for 16 hours (37°C, 5% CO<sub>2</sub>), after which the luciferase activity is measured as described in paragraph 29.

23. The recommended negative control is LA. 20 mg of LA prepared in a 1.5-mL microfuge tube, to which X-VIVO™ 15 is added up to 1 mL (20 mg/ mL). Twenty mg/mL of LA solution is diluted by a factor of 5 with X-VIVO™ 15 (4 mg/mL); 500 µl of this 4 mg/mL LA solution is transferred to a well of a 96-well assay block. This solution is diluted by a factor of 2 with X-VIVO™ 15 and then diluted again by a factor of 2 to produce 2 mg/mL and 1 mg/mL solutions. 50 µl of these 3 solutions and vehicle control (X-VIVO™ 15) are added to 50 µl of THP-G8 cell suspension in the wells of a 96-well flat-bottom black

plate. Each concentration of the negative control is tested in 4 wells. The plate is agitated on a plateshaker and incubated in a CO<sub>2</sub> incubator for 16 hours (37°C, 5% CO<sub>2</sub>), after which the luciferase activity is measured as described in paragraph 29.

24. Other suitable positive or negative controls may be used if historical data are available to derive comparable run acceptance criteria.

25. Care should be taken to avoid evaporation of volatile test chemicals and cross-contamination between wells by test chemicals, e.g. by sealing the plate prior to the incubation with the test chemicals.

26. The test chemicals and solvent control require 2 to 4 runs to derive a positive or negative prediction (see Table 2). Each run is performed on a different day with fresh X-VIVO™ 15 stock solution of test chemicals and independently harvested cells. Cells may come from the same passage.

### Luciferase activity measurements

27. Luminescence is measured using a 96-well microplate luminometer equipped with optical filters, e.g. Phelios (ATTO, Tokyo, Japan), Tristan 941 (Berthold, Bad Wildbad, Germany) and the ARVO series (PerkinElmer, Waltham, MA, USA). The luminometer must be calibrated for each test to ensure reproducibility (19). Recombinant orange and red emitting luciferases are available for this calibration.

28. 100µL of pre-warmed Tripluc® Luciferase assay reagent (Tripluc) is transferred to each well of the plate containing the cell suspension treated with or without chemical. The plate is shaken for 10 min at an ambient temperature of about 20°C. The plate is placed in the luminometer to measure the luciferase activity. Bioluminescence is measured for 3 sec each in the absence (F0) and presence (F1) of the optical filter. Justification should be provided for the use of alternative settings, e.g. depending on the model of luminometer used.

29. Parameters for each concentration are calculated from the measured values, e.g. IL8LA, GAPLA, nIL8LA, Ind-IL8LA, Inh-GAPLA, the mean ±SD of IL8LA, the mean ±SD of GAPLA, the mean ±SD of nIL8LA, the mean ±SD of Ind-IL8LA, the mean ±SD of Inh-GAPLA, and the 95% confidence interval of Ind-IL8LA. Definitions of the parameters used in this paragraph are provided in Appendices I and IV, respectively.

30. Prior to measurement, colour discrimination in multi-colour reporter assays is generally achieved using detectors (luminometer and plate reader) equipped with optical filters, such as sharp-cut (long-pass or short-pass) filters or band-pass filters. The transmission coefficients of the filters for each bioluminescence signal colour should be calibrated prior to testing, per Appendix II.

## DATA AND REPORTING

## Data evaluation

31. Criteria for a positive/negative decision require that in each run:

- an IL-8 Luc assay prediction is judged positive if a test chemical has a Ind-IL8LA  $\geq 1.4$  and the lower limit of the 95% confidence interval of Ind-IL8LA  $\geq 1.0$
- an IL-8 Luc assay prediction is judged negative if a test chemical has a Ind-IL8LA  $< 1.4$  and/or the lower limit of the 95% confidence interval of Ind-IL8LA  $< 1.0$

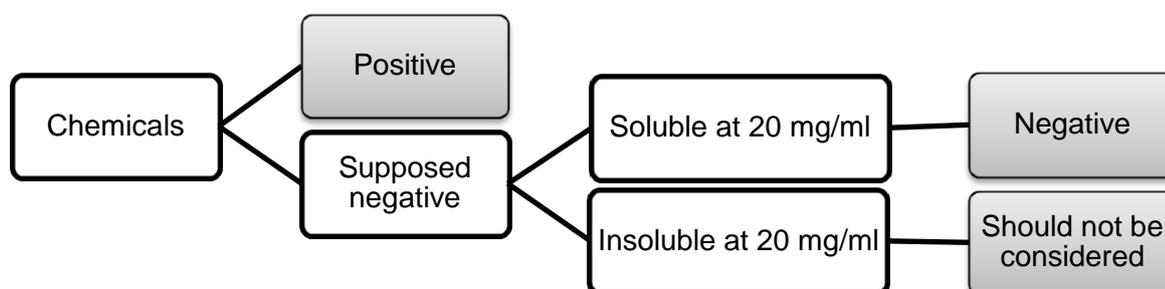
*Prediction model*

32. Test chemicals that provide two positive results from among the 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> or 4<sup>th</sup> runs are identified as positives whereas those that give three negative results from among the 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> or 4<sup>th</sup> runs are identified as supposed negative (Table 2). Among supposed negative chemicals, chemicals that are dissolved at 20 mg/ml of X-VOVO™ 15 are judged as negative, while chemicals that are not dissolved at 20 mg/ml of X-VOVO™ 15 should not be considered (Figure 1).

Table 2. Criteria for identifying positive and supposed negative

1st run	2nd run	3rd run	4th run	Final prediction
Positive	Positive	-	-	Positive
	Negative	Positive	-	Positive
		Negative	Positive	Positive
			Negative	Supposed negative
Negative	Positive	Positive	-	Positive
		Negative	Positive	Positive
			Negative	Supposed negative
	Negative	Positive	Positive	Positive
			Negative	Supposed negative
		Negative	-	Supposed negative

Figure 1. Prediction model for final judgment



### Acceptance criteria

33. The following acceptance criteria should be met when using the IL-8 Luc assay:

- Ind-IL8LA should be more than 5.0 at least in one concentration of the positive control, 4-NBB, in each run.
- Ind-IL8LA should be less than 1.4 at any concentration of the negative control, lactic acid, in each run.
- Data from plates for which the GAPLA of control wells with cells and Tripluc but without chemicals is less than 5 times of that of well containing test medium only (50  $\mu$ L/well of RPMI-1640 containing 10% FBS and 50  $\mu$ L/well of X-VIVO™ 15) should be rejected.
- Data from plates for which the Inh-GAPLA of all concentrations of the test or control chemicals is less than 0.05 should be rejected. In this case, the first test should be repeated so the highest final concentration of the repeated test is the lowest final concentration of the previous test.

### Test report

34. The test report should include the following information:

#### *Test chemicals*

- Mono-constituent substance:
  - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
  - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc;

- Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Solubility in X-VIVO™ 15. For chemicals that are insoluble in X-VIVO™ 15, whether precipitation or flotation are observed after centrifugation;
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available;
  - Justification for choice of solvent/vehicle for each test chemical if X-VIVO™ 15 has not been used.
- Multi-constituent substance, UVCB and mixture:
- Characterisation as far as possible by e.g. chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available;
  - Physical appearance, water solubility, and additional relevant physicochemical properties, to the extent available;
  - Molecular weight or apparent molecular weight in case of mixtures/polymers of known compositions or other information relevant for the conduct of the study;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Solubility in X-VIVO™ 15. For chemicals that are insoluble in X-VIVO™ 15, whether precipitation or flotation are observed after centrifugation;
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available.
  - Justification for choice of solvent/vehicle for each test chemical, if X-VIVO™ 15 has not been used.

### *Controls*

- Positive control:
- Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
  - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available and where applicable;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available;
  - Reference to historical positive control results demonstrating suitable acceptance criteria, if

applicable.

- Negative control:
  - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), and/or other identifiers;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc;
  - Physical appearance, molecular weight, and additional relevant physicochemical properties in the case other negative controls than those mentioned in the Test Guideline are used and to the extent available;
  - Storage conditions and stability to the extent available;
  - Justification for choice of solvent for each test chemical.

#### *Test method conditions*

- Name and address of the sponsor, test facility and study director;
- Description of test method used;
- Cell line used, its storage conditions, and source (e.g. the facility from which it was obtained);
- Lot number and origin of FBC, supplier name, lot number of 96-well flat-bottom black plate, and lot number of Tripluc reagent;
- Passage number and cell density used for testing;
- Cell counting method used for seeding prior to testing and measures taken to ensure homogeneous cell number distribution;
- Luminometer used (e.g. model), including instrument settings, luciferase substrate used, and demonstration of appropriate luminescence measurements based on the control test described in Appendix II;
- The procedure used to demonstrate proficiency of the laboratory in performing the test method (e.g. by testing of proficiency substances) or to demonstrate reproducible performance of the test method over time.

#### *Test procedure*

- Number of replicates and runs performed;
- Test chemical concentrations, application procedure and exposure time (if different from those recommended);
- Description of evaluation and decision criteria used;
- Description of study acceptance criteria used;
- Description of any modifications of the test procedure.

*Results*

- Measurements of IL8LA and GAPLA;
- Calculations for nIL8LA, Ind-IL8LA, and Inh-GAPLA;
- The 95% confidence interval of Ind-IL8LA;
- A graph depicting dose-response curves for induction of luciferase activity and viability;
- Description of any other relevant observations, if applicable.

*Discussion of the results*

- Discussion of the results obtained with the IL-8 Luc assay;
- Consideration of the assay results in the context of an IATA, if other relevant information is available.

*Conclusion*

## LITERATURE

1. Takahashi T, Kimura Y, Saito R, Nakajima Y, Ohmiya Y, Yamasaki K, and Aiba S. 2011. An in vitro test to screen skin sensitizers using a stable THP-1-derived IL-8 reporter cell line, THP-G8. *Toxicol Sci* 124:359-69.
2. OECD (2017), Validation report for the international validation study on the IL-8 Luc assay as a test evaluating the skin sensitizing potential of chemicals conducted by the IL-8 Luc Assay. Series on Testing and Assessment No. 267, ENV/JM/MONO(2017)19. Organisation for Economic Cooperation and Development, Paris. Available at: [<http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm>].
3. OECD (2017), Report of the Peer Review Panel for the IL-8 Luciferase (IL-8 Luc) Assay for in vitro skin sensitisation. Series on Testing and Assessment No. 258, ENV/JM/MONO(2017)20. Organisation for Economic Cooperation and Development, Paris. Available at: [<http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm>].
2. OECD (2016), Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated Approaches To Testing And Assessment (IATA) For Skin Sensitisation, Series on Testing & Assessment No. 256, ENV/JM/MONO(2016)29. Organisation for Economic Cooperation and Development, Paris. Available at: [<http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm>].
3. van der Veen JW, Rorije E, Emter R, Natsch A, van Loveren H, and Ezendam J. 2014. Evaluating the performance of integrated approaches for hazard identification of skin sensitizing chemicals. *Regul Toxicol Pharmacol* 69:371-9.
4. Kimura Y, Fujimura C, Ito Y, Takahashi T, Nakajima Y, Ohmiya Y, and Aiba S. 2015. Optimization of the IL-8 Luc assay as an in vitro test for skin sensitization. *Toxicol In Vitro* 29:1816-30.
5. Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle S, Landsiedel R, Jaworska J, Kern PS, Gerberick F, et al. 2015. Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul Toxicol Pharmacol* 71:337-51.
6. Ashikaga T, Sakaguchi H, Sono S, Kosaka N, Ishikawa M, Nukada Y, Miyazawa M, Ito Y, Nishiyama N, and Itagaki H. 2010. A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). *Alternatives to laboratory animals : ATLA* 38:275-84.
7. Patlewicz G, Casati S, Basketter DA, Asturiol D, Roberts DW, Lepoittevin J-P, Worth A and Aschberger K (2016) Can currently available non-animal methods detect pre and pro haptens relevant for skin sensitisation? *Regul Toxicol Pharmacol*, online.
8. Thorne N, Inglese J, and Auld DS. 2010. Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical biology. *Chem Biol* 17:646-57.

9. OECD (2016), Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264265295-en>.
10. Viviani V, Uchida A, Suenaga N, Ryufuku M, and Ohmiya Y. 2001. Thr226 is a key residue for bioluminescence spectra determination in beetle luciferases. *Biochem Biophys Res Commun* 280:1286-91.
11. Viviani VR, Bechara EJ, and Ohmiya Y. 1999. Cloning, sequence analysis, and expression of active *Phrixothrix* railroad-worms luciferases: relationship between bioluminescence spectra and primary structures. *Biochemistry* 38:8271-9.
12. Nakajima Y, Kimura T, Sugata K, Enomoto T, Asakawa A, Kubota H, Ikeda M, and Ohmiya Y. 2005. Multicolor luciferase assay system: one-step monitoring of multiple gene expressions with a single substrate. *Biotechniques* 38:891-4.
13. OECD (2017), To be published - Performance Standards for the assessment of proposed similar or modified *in vitro* skin sensitisation IL-8 luc test methods. OECD Environment, Health and Safety Publications, Series on Testing and Assessment. OECD, Paris, France.
14. OECD (2005), Guidance Document the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. OECD Environment, Health and Safety publications, OECD Series on Testing and Assessment No.34. OECD, Paris, France.
15. OECD (2017), Draft Guidance document: Good *In Vitro* Method Practices (GIVIMP) for the Development and Implementation of *In Vitro* Methods for Regulatory Use in Human Safety Assessment. Organisation for Economic Cooperation and Development, Paris. Available at: [[http://www.oecd.org/env/ehs/testing/OECD%20Draft%20GIVIMP\\_v05%20-%20clean.pdf](http://www.oecd.org/env/ehs/testing/OECD%20Draft%20GIVIMP_v05%20-%20clean.pdf)].
16. JaCVAM (2016), IL-8 Luc assay protocol, Available at. [http://www.jacvam.jp/en\\_effort/effort02.html](http://www.jacvam.jp/en_effort/effort02.html).
17. Niwa K, Ichino Y, Kumata S, Nakajima Y, Hiraishi Y, Kato D, Viviani VR, and Ohmiya Y. 2010. Quantum yields and kinetics of the firefly bioluminescence reaction of beetle luciferases. *Photochem Photobiol* 86:1046-9.
18. OECD (2012), The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins, Part 1: Scientific Evidence. OECD Environment, Health and Safety Publications, Series on Testing and Assessment No. 168. OECD, Paris, France.
19. United Nations (2015), Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Sixth revised edition. New York & Geneva: United Nations Publications. ISBN: 978-92-1-117006-1. Available at: [http://www.unece.org/trans/danger/publi/ghs/ghs\\_rev05/05files\\_e.html](http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html).

**APPENDIX I:  
DEFINITIONS**

**Accuracy:** The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with concordance to mean the proportion of correct outcomes of a test method (16).

**AOP (Adverse Outcome Pathway):** Sequence of events from the chemical structure of a target chemical or group of similar chemicals through the molecular initiating event to an *in vivo* outcome of interest (20).

**CV05:** Cell viability 05. Minimum concentration at which chemicals show less than 0.05 of Inh-GAPLA.

**FInSLO-LA:** Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to Ind-IL8LA. See Ind-IL8LA for definition.

**GAPLA:** Luciferase Activity of Stable Luciferase Red (SLR) ( $\lambda_{\max} = 630$  nm), regulated by GAPDH promoter and demonstrates cell viability and viable cell number.

**Hazard:** Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub) population is exposed to that agent.

**IATA (Integrated Approach to Testing and Assessment):** A structured approach used for hazard identification (potential), hazard characterisation (potency) and/or safety assessment (potential/potency and exposure) of a chemical or group of chemicals, which strategically integrates and weights all relevant data to inform regulatory decision regarding potential hazard and/or risk and/or the need for further targeted and therefore minimal testing.

**II-SLR-LA:** Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to Inh-GAPLA. See Inh-GAPLA for definition

**IL-8 (Interleukin-8):** A cytokine derived from endothelial cells, fibroblasts, keratinocytes, macrophages, and monocytes that causes chemotaxis of neutrophils and T-cell lymphocytes.

**IL8LA:** Luciferase Activity of Stable Luciferase Orange (SLO) ( $\lambda_{\max} = 580$  nm), regulated by IL-8 promoter.

**Ind-IL8LA:** Fold induction of IL8LA. It is obtained by dividing the nIL8LA of THP-G8 cells treated with chemicals by that of non-stimulated THP-G8 cells and represents the induction of IL-8 promoter activity by chemicals.

**Inh-GAPLA:** Inhibition of GAPLA. It is obtained by dividing GAPLA of THP-G8 treated with chemicals with GAPLA of non-treated THP-G8 and represents cytotoxicity of chemicals.

**Minimum induction threshold (MIT):** the lowest concentration at which a chemical satisfies the positive criteria

**Mixture:** A mixture or a solution composed of two or more substances in which they do not react.

**Mono-constituent substance:** A substance, defined by its quantitative composition, in which one main constituent is present to at least 80% (w/w).

**Multi-constituent substance:** A substance, defined by its quantitative composition, in which more than one of the main constituents is present in a concentration  $\geq 10\%$  (w/w) and  $< 80\%$  (w/w). A multi-constituent substance is the result of a manufacturing process. The difference between mixture and multi-constituent substance is that a mixture is obtained by blending of two or more substances without chemical reaction. A multi-constituent substance is the result of a chemical reaction.

**nIL8LA:** The SLO luciferase activity reflecting IL-8 promoter activity (IL8LA) normalised by the SLR luciferase activity reflecting GAPDH promoter activity (GALPA). It represents IL-8 promoter activity after considering cell viability or cell number.

**nSLO-LA:** Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to nIL8LA. See nIL8LA for definition

**Positive control:** A replicate containing all components of a test system and treated with a substance known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the positive response should not be excessive.

**Pre-haptens:** Chemicals which become sensitisers through abiotic transformation.

**Pro-haptens:** Chemicals requiring enzymatic activation to exert skin sensitisation potential.

**Relevance:** Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (16).

**Reliability:** Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability (16).

**Run:** A run consists of one or more test chemicals tested concurrently with a solvent/vehicle control and with a positive control.

**Sensitivity:** The proportion of all positive/active chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (16).

**SLO-LA:** Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to IL8LA. See IL8LA for definition.

**SLR-LA:** Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to GAPLA. See GAPLA for definition.

**Solvent/vehicle control:** An untreated sample containing all components of a test system except of the test chemical, but including the solvent/vehicle that is used. It is used to establish the baseline response for the samples treated with the test chemical dissolved or stably dispersed in the same solvent/vehicle. When tested with a concurrent medium control, this sample also demonstrates whether the solvent/vehicle interacts with the test system.

**Specificity:** The proportion of all negative/inactive chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method (16).

**Substance:** Chemical elements and their compounds in the natural state or obtained by any production process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition.

**Surfactant:** Also called surface-active agent, this is a substance, such as a detergent, that can reduce the surface tension of a liquid and thus allow it to foam or penetrate solids; it is also known as a wetting agent. (TG437)

**Test chemical:** The term "test chemical" is used to refer to what is being tested.

**THP-G8:** An IL-8 reporter cell line used in IL-8 Luc assay. The human macrophage-like cell line THP-1 was transfected the SLO and SLR luciferase genes under the control of the IL-8 and GAPDH promoters, respectively.

**United Nations Globally Harmonized System of Classification and Labeling of Chemicals (UN GHS):** A system proposing the classification of chemicals (substances and mixtures) according to standardised types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (21).

**UVCB:** substances of unknown or variable composition, complex reaction products or biological materials.

**Valid test method:** A test method considered to have sufficient relevance and reliability for a specific purpose and which is based on scientifically sound principles. A test method is never valid in an absolute sense, but only in relation to a defined purpose.

## APPENDIX II

**Principle of measurement of luciferase activity and determination of the transmission coefficients of optical filter for SLO and SLR**

MultiReporter Assay System -Tripluc- can be used with a microplate-type luminometer with a multi-colour detection system, which can equip an optical filter (e.g. Phelios AB-2350 (ATTO), ARVO (PerkinElmer), Tristar LB941 (Berthold)). The optical filter used in measurement is 600–620 nm long or short pass filter, or 600–700 nm band pass filter.

(1) Measurement of two-colour luciferases with an optical filter.

This is an example using Phelios AB-2350 (ATTO). This luminometer is equipped with a 600 nm long pass filter (R60 HOYA Co.), 600 nm LP, Filter 1) for splitting SLO ( $\lambda_{\max} = 580$  nm) and SLR ( $\lambda_{\max} = 630$  nm) luminescence.

To determine transmission coefficients of the 600 nm LP, first, using purified SLO and SLR luciferase enzymes, measure i) the intensity of SLO and SLR bioluminescence intensity without filter (F0), ii) the SLO and SLR bioluminescence intensity that passed through 600 nm LP (Filter 1), and iii) calculate the transmission coefficients of 600 nm LP for SLO and SLR listed below.

Transmission coefficients		Abbreviation	Definition
SLO	Filter 1 Transmission coefficients	$\kappa O_{R60}$	The filter's transmission coefficient for the SLO
SLR	Filter 1 Transmission coefficients	$\kappa R_{R60}$	The filter's transmission coefficient for the SLR

When the intensity of SLO and SLR in test sample are defined as O and R, respectively, i) the intensity of light without filter (all optical) F0 and ii) the intensity of light that transmits through 600 nm LP (Filter 1) F1 are described as below.

$$F0=O+R$$

$$F1=\kappa O_{R60} \times O + \kappa R_{R60} \times R$$

These formulas can be rephrased as follows:

$$\begin{pmatrix} F0 \\ F1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \kappa O_{R60} & \kappa R_{R60} \end{pmatrix} \begin{pmatrix} O \\ R \end{pmatrix}$$

Then using calculated transmittance factors ( $\kappa O_{R60}$  and  $\kappa R_{R60}$ ) and measured F0 and F1, you can calculate O and R-value as follows:

$$\begin{pmatrix} O \\ R \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \kappa O_{R60} & \kappa R_{R60} \end{pmatrix}^{-1} \begin{pmatrix} F0 \\ F1 \end{pmatrix}$$

**Materials and methods for determining transmittance factor**

(1) Reagents

· Single purified luciferase enzymes:

Lyophilised purified SLO enzyme

Lyophilised purified SLR enzyme

(which for the validation work were obtained from GPC Lab. Co. Ltd., Tottori, Japan with THP-G8 cell line)

· Assay reagent:

Tripluc<sup>®</sup> Luciferase assay reagent ( for example from TOYOBO Cat#MRA-301)

· Medium: for luciferase assay (30 ml, stored at 2 – 8°C)

Reagent	Conc.	Final conc. in medium	Required amount
RPMI-1640	-	-	27 ml
FBS	-	10 %	3 ml

(2) Preparation of enzyme solution

Dissolve lyophilised purified luciferase enzyme in tube by adding 200 µl of 10 ~ 100 mM Tris/HCl or Hepes/HCl (pH 7.5 ~ 8.0) supplemented with 10% (w/v) glycerol, divide the enzyme solution into 10 µl aliquots in 1.5 ml disposable tubes and store them in a freezer at -80°C. The frozen enzyme solution can be used for up to 6 months. When used, add 1 ml of medium for luciferase assay (RPMI-1640 with 10% FBS) to each tube containing the enzyme solutions (diluted enzyme solution) and keep them on ice to prevent deactivation.

(3) Bioluminescence measurement

Thaw Tripluc<sup>®</sup> Luciferase assay reagent (Tripluc) and keep it at room temperature either in a water bath or at ambient air temperature. Power on the luminometer 30 min before starting the measurement to allow the photomultiplier to stabilise. Transfer 100 µl of the diluted enzyme solution to a black 96 well plate (flat bottom) (the SLO reference sample to #B1, #B2, #B3, the SLR reference sample to #D1, #D2, #D3). Then, transfer 100 µl of pre-warmed Tripluc to each well of the plate containing the diluted enzyme solution using a pipetman. Shake the plate for 10 min at room temperature (about 25°C) using a plate shaker. Remove bubbles from the solutions in wells if they appear. Place the plate in the luminometer to measure the luciferase activity. Bioluminescence is measured for 3 sec each in the absence (F0) and presence (F1) of the optical filter.

Transmission coefficient of the optical filter was calculated as follows:

Transmission coefficient (SLO ( $\kappa O_{R60}$ ))= (#B1 of F1+ #B2 of F1+ #B3 of F1) / (#B1 of F0+ #B2 of F0+ #B3 of F0)

Transmission coefficient (SLR ( $\kappa R_{R60}$ ))= (#D1 of F1+ #D2 of F1+ #D3 of F1) / (#D1 of F0+ #D2 of F0+ #D3 of F0)

Calculated transmittance factors are used for all the measurements executed using the same luminometer.

**Quality control of equipment**

The procedures described in the IL-8 Luc protocol should be used (18).

.

## APPENDIX III

## PROFICIENCY SUBSTANCES

Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency by obtaining the expected IL-8 Luc assay prediction for the 10 substances recommended in Table 1 and by obtaining values that fall within the respective reference range for at least 8 out of the 10 proficiency substances (selected to represent the range of responses for skin sensitisation hazards). Other selection criteria were that the substances are commercially available, and that high-quality *in vivo* reference data as well as high quality *in vitro* data generated with the IL-8 Luc assay are available. Also, published reference data are available for the IL-8 Luc assay (6) (1).

Table 1: Recommended substances for demonstrating technical proficiency with the IL-8 Luc assay

Proficiency substances	CAS no.	State	Solubility in X- VIVO15 at 20 mg/mL	<i>In vivo</i> prediction <sup>1</sup>	IL-8 Luc prediction <sup>2</sup>	Reference range (µg/mL) <sup>3</sup>	
						CV05 <sup>4</sup>	IL-8 Luc MIT <sup>5</sup>
2,4-Dinitrochlorobenzene	97-00-7	Solid	Insoluble	Sensitiser (Extreme)	Positive	2.3-3.9	0.5-2.3
Formaldehyde	50-00-0	Liquid	Soluble	Sensitiser (Strong)	Positive	9-30	4-9
2-Mercaptobenzothiazole	149-30-4	Solid	Insoluble	Sensitiser (Moderate)	Positive	250-290	60-250
Ethylenediamine	107-15-3	Liquid	Soluble	Sensitiser (Moderate)	Positive	500-700	0.1-0.4
Ethyleneglycol dimethacrylate	97-90-5	Liquid	Insoluble	Sensitiser (Weak)	Positive	>2000	0.04-0.1
4-Allylanisole (Estragol)	140-67-0	Liquid	Insoluble	Sensitiser (Weak)	Positive	>2000	0.01-0.07
Streptomycin sulphate	3810-74-0	Solid	Soluble	Non- sensitiser	Negative	>2000	>2000
Glycerol	56-81-5	Liquid	Soluble	Non- sensitiser	Negative	>2000	>2000
Isopropanol	67-63-0	Liquid	Soluble	Non- sensitiser	Negative	>2000	>2000

Abbreviations: CAS no. = Chemical Abstracts Service Registry Number

<sup>1</sup> The *in vivo* potency is derived using the criteria proposed by ECETOC (19).

<sup>2</sup> Based on historical observed values (1) (6).

<sup>3</sup> CV05 and IL-8 Luc MIT were calculated using water solubility given by EPI Suite™.

<sup>4</sup> CV05: the minimum concentration at which chemicals show less than 0.05 of Inh-GAPLA.

<sup>5</sup> MIT: the lowest concentrations at which a chemical satisfies the positive criteria.

## APPENDIX IV

## Indexes and judgment criteria

**nIL8LA (nSLO-LA)**

The j-th repetition (j = 1-4) of the i-th concentration (i = 0-11) is measured for IL8LA (SLO-LA) and GAPLA (SLR-LA) respectively. The normalised IL8LA, referred to as nIL8LA (nSLO-LA), and is defined as:

$$nIL8LA_{ij} = nIL8LA_{ij} / GAPLA_{ij}$$

This is the basic unit of measurement in this assay.

**Ind-IL8LA (FInSLO-LA)**

The fold increase of the averaged nIL8LA (nSLO-LA) for the repetition on the i-th concentration compared with it at the 0 concentration, Ind-IL8LA, is the primary measure of this assay. This ratio is written by the following formula:

$$Ind-IL8LA_i = \left\{ (1/4) \times \sum_j nIL8LA_{ij} \right\} / \left\{ (1/4) \times \sum_j nIL8LA_{0j} \right\}$$

The lead laboratory has proposed that a value of 1.4 corresponds to a positive result for the tested chemical. This value is based on the investigation of the historical data of the lead laboratory. Data management team then used this value through all the phases of validation study. The primary outcome, Ind-IL8LA, is the ratio of 2 arithmetic means as shown in equation.

**95% confidence interval (95% CI)**

The 95% confidence interval (95% CI) based on the ratio can be estimated to show the precision of this primary outcome measure. The lower limit of the 95% CI  $\geq 1$  indicates that the nIL8LA with the i-th concentration is significantly greater than that with solvent control. There are several ways to construct the 95% CI. We used the method known as Fieller's theorem in this study. This 95% confidence interval theorem is obtained from the following formula:

$$\left[ \frac{-B - \sqrt{B^2 - 4AC}}{2A}, \frac{-B + \sqrt{B^2 - 4AC}}{2A} \right],$$

where  $A = \bar{x}_0^2 - t_{0.975(v)}^2 \times \frac{sd_0^2}{n_0}$ ,  $B = -2 \times \bar{x} \times \bar{y}$ ,  $C = \bar{y}_i^2 - t_{0.975(v)}^2 \times \frac{sd_{y_i}^2}{n_{y_i}}$ , and  $n_0 = 4$ ,

$$\bar{x}_0 = (1/n_0) \times \sum_j nIL8LA_{0j}, \quad sd_0^2 = \{1/(n_0 - 1)\} \times \sum_j (nIL8LA_{0j} - \bar{x}_0)^2,$$

$$n_{y_i}=4, \bar{y}_i = (1/n_{y_i}) \times \sum_j (nIL8LA_{ij}), \text{sd}_{y_i}^2 = \left\{ 1/(n_{y_i} - 1) \right\} \times \sum_j (nIL8LA_{ij} - \bar{y}_i)^2 .$$

$t_{0.975(v)}$  is 97.5 percentile of the central t distribution with the  $v$  of the degree of freedom, where

$$v = \left( \frac{\text{sd}_0^2}{n_0} + \frac{\text{sd}_{y_i}^2}{n_{y_i}} \right)^2 / \left\{ \left( \frac{\text{sd}_0^2}{n_0} \right)^2 / (n_0 - 1) + \left( \frac{\text{sd}_{y_i}^2}{n_{y_i}} \right)^2 / (n_{y_i} - 1) \right\} .$$

**Inh-GAPLA (II-SLR-LA)**

The Inh-GAPLA is a ratio of the averaged GAPLA (SLR-LA) for the repetition of the  $i$ -th concentration compared with that with solvent control, and this is written by

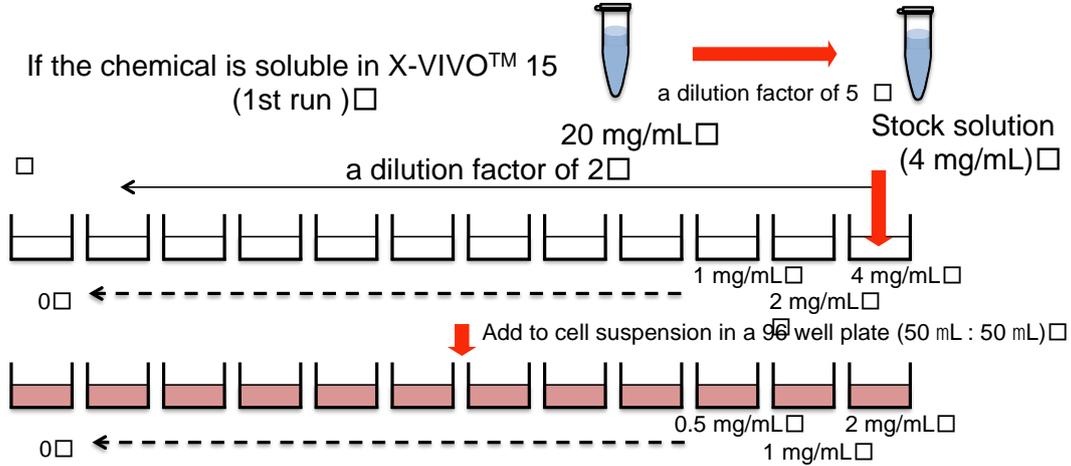
$$\text{Inh-GAPLA}_i = \left\{ (1/4) \times \sum_j \text{GAPLA}_{ij} \right\} / \left\{ (1/4) \times \sum_j \text{GAPLA}_{0j} \right\} .$$

Since the GAPLA is the denominator of the nIL8LA, an extremely small value causes large variation in the nIL8LA. Therefore, Ind-IL8LA values with an extremely small value of Inh-GAPLA (less than 0.05) might be considered poor precision.

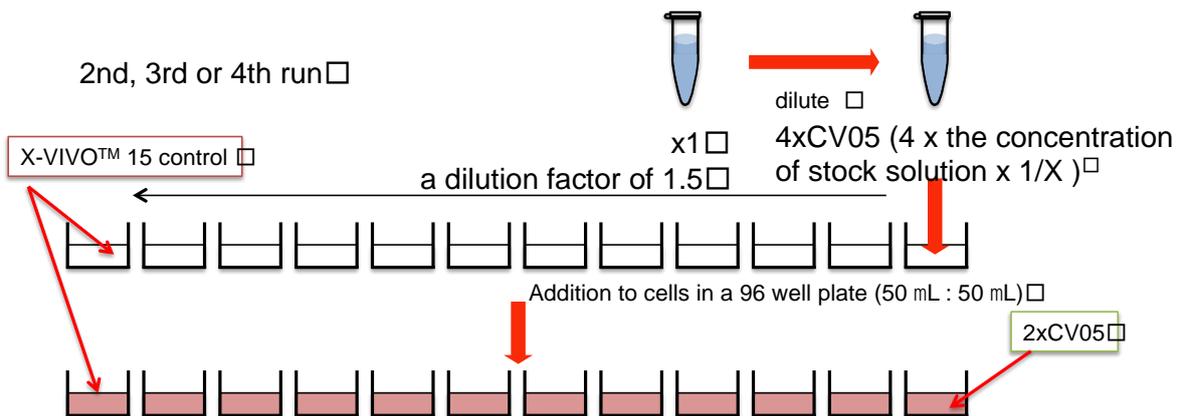
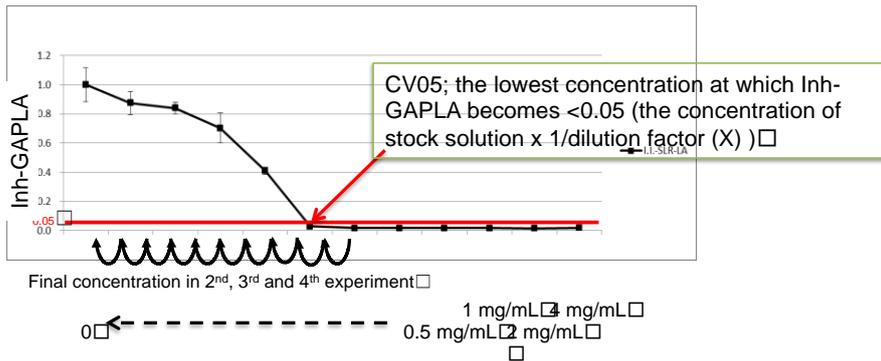
APPENDIX V

The scheme of the methods to dissolve chemicals for the IL-8 Luc assay.

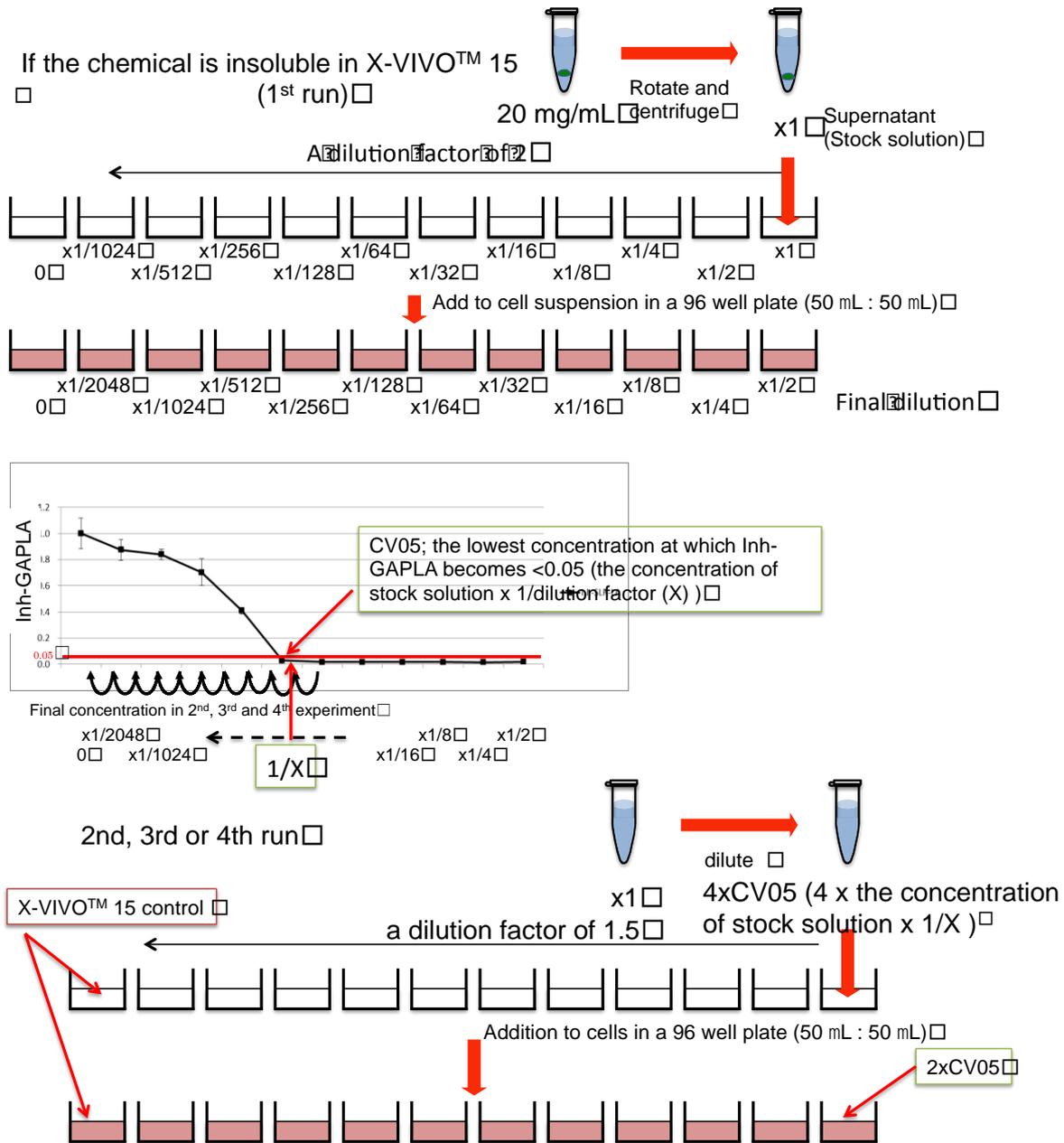
(a) For chemicals dissolved in X-VIVO™ 15 at 20 mg/mL



Determine the highest concentration of the following experiments



(b) For chemicals insoluble in X-VIVO™ 15 at 20 mg/mL



The scheme of the method to dissolve 4-NBB for the positive control of the IL-8 Luc assay.

The positive control : 4-NBB (insoluble in X-VIVO™ 15) □

