OECD GUIDELINE FOR THE TESTING OF CHEMICALS

DRAFT PROPOSAL FOR A NEW TEST GUIDELINE

In Vitro Skin Sensitisation: IL-8 Luc assay

INTRODUCTION

1. The term 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) (United Nations UN, 2013). This Test Guideline provides an *in vitro* procedure called the Interleukin-8 Reporter Gene Assay (IL-8 Luc assay), which is useful for distinguishing skin sensitisers from non-sensitisers in accordance with the UN GHS.

2. In 2012, the Organisation for Economic Co-operation and Development (OECD) published the adverse outcome pathway (AOP) for skin sensitisation that defines the key events in the sensitisation process (OECD, 2012a). In brief, the chemical sensitiser penetrates the stratum corneum, the uppermost layer of the skin, and subsequently binds covalently to proteins (key event 1) to form hapten-protein conjugates, which can be immunogenic. In parallel, keratinocytes are stimulated to release danger signals e.g. pro-inflammatory cytokines or ATP and, in addition, to activate anti-oxidative response genes (key event 2). Next, dendritic cells (DC) acquire mature phenotypes, such as the induction of several co-stimulatory molecules and production of proinflammatory cytokines and chemokines through the concerted recognition of hapten-protein conjugates by MHC (major histocompatibility complex) molecules (key event 3). The activated DCs mobilise and migrate from the skin to the draining lymph node to present the allergen to T cells. After stimulation by DCs, hapten-peptide-specific T cells expand (key event 4) to elicit the eventual adverse outcome after a second exposure to the chemical sensitiser. The establishment of this skin sensitisation AOP has promoted the development of a variety of non-animal test methods to distinguish sensitisers from non-sensitisers or to generate potency information (Adler et al., 2011, Reisinger et al., 2015).

3. Different laboratory animals have been used for the assessment of skin sensitisation. The classical methods using guinea-pigs, i.e. the Guinea Pig Maximisation Test (GPMT) of Magnusson and Kligman, and the Buehler Test - TG 406 (OECD, 1992) assess both the induction and elicitation phases of skin sensitisation. On the other hand, the murine tests, e.g. the local lymph node assay (LLNA) - TG 429 (OECD, 2010a) and its two non-radioactive modifications, LLNA: DA -TG 442A (OECD, 2010b) and LLNA: BrdU-ELISA - TG 442B

(OECD, 2010c) assess the induction response exclusively. The latter provide an advantage over the guinea pig tests in terms of animal welfare as well as with an objective measurement of the induction phase of skin sensitisation.

4. The OECD recently released new Test Guidelines for skin sensitisation testing using the Direct Peptide Reactivity Assay (DPRA) (TG 442C) (OECD, 2015a), the ARE-Nrf2 Luciferase (TG 442D) (OECD, 2015b) and the human Cell Line Activation Test (h-CLAT) (TG 442E) (OECD, 2016) test methods. DPRA examines protein-binding/haptenisation of chemicals, the ARE-Nrf2 Luciferase test method examines the induction of the Kelch-like ECH-associated protein 1 (Keap-1)/nuclear factor (erythroid-derived 2)-like factor 2 (Nrf2) pathways in keratinocytes and h-CLAT examines the induction of CD56 and CD84 in THP-1 cells. The DPRA, ARE-Nrf2 Luciferase, and h-CLAT test methods were designed to target three different key events in the AOP, i.e. key events 1, 2, and 3, respectively and to contribute to the evaluation of the skin sensitisation hazard potential of chemicals. However, it is unlikely that a single assay will be sufficient to adequately assess sensitisation potential because of the complexity of the sensitisation process (Bauch et al., 2012; Reisinger et al., 2015; Urbisch et al., 2015). Therefore, a combination of non-animal methods (*in silico, in chemico, in vitro*) within Integrated Approaches to Testing and Assessment (IATA) are needed to substitute adequately for the animal tests currently in use, given the restricted AOP mechanistic coverage of each of the currently available non-animal test methods (Jaworska et al., 2013; Nukada et al., 2013).

5. The IL-8 Luc assay is proposed to address the third key event (dendritic cell activation) of the skin sensitisation AOP by quantifying changes in the expression of cytokine associated with the process of activation of DC (i.e. IL-8), in the human monocytic leukemia cell line THP-1-derived IL-8 reporter cell line, THP-G8, following exposure to sensitisers (Takahashi et al., 2011). The measured expression levels of the luciferase activity are then used for supporting the discrimination between skin sensitisers and non-sensitisers.

6. The IL-8 Luc assay was developed in the New Energy and Industrial Technology Development Organisation (NEDO) project, and has been validated in a study conducted by the Japanese Center 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) as well as a subsequent independent peer review conducted by JaCVAM and the Ministry of Health, Labour and Welfare (MHLW) with the support of the International Cooperation on Alternative Test Methods (ICATM). Based on the available evidence as well as input from regulators and stakeholders, the IL-8 Luc assay is considered useful as part of IATA to distinguish sensitisers from non-sensitisers in accordance with the UN GHS (van der Veen et al., 2014). The possible use of IL-8 Luc assay data in combination with other information is reported in the

literature (Kimura et al., 2015).

7. Definitions are provided in Annex I.

INITIAL CONSIDERATIONS AND LIMITATIONS

8. It is well known that DCs such as epidermal Langerhans cells (LCs) maintain immature phenotype in the absence of hapten stimulation. Once stimulated by haptens either *in vivo* or *in vitro*, they drastically change their phenotype from immature to mature phenotype (Aiba and Katz, 1990; Aiba et al., 1997). Mature DCs increase expression of CD40, CD54, CD80, CD83, and CD83 in addition to induction of proinflammatory cytokines, such as IL-1β and TNF-α, and several chemokines including IL-8 (CXCL8) and CCL3 (Steinman, 1991; Caux et al., 1994; Aiba and Tagami, 1999; Aiba, 2007). Several studies have suggested that the basic molecular mechanism of DC maturation is mainly dependent on the activation of p38 mitogen activated protein kinase (p38 MAPK) (Sasaki et al., 2007; Mitjans et al., 2010). Therefore, biomarkers of DC activation should not be restricted to the increased expression of CD54 and CD86. The expression of other surface molecules such as CD40 and CD80 and the production of proinflammatory cytokines IL-1β and TNF-α and chemokines IL-8 and CCL3, which are also regulated by the activation of p38 MAPK, can also be good biomarkers for DC activation. In addition, besides augmenting their expression of CD54 and CD86, DCs need to produce proinflammatory cytokines to stimulate T cells and chemokines like IL-8 to recruit other inflammatory cells. IL-8 is well established as a potent chemotactic peptide for neutrophils, T lymphocytes, basophils (Leonard et al., 1990), and NK cells (Sebok et al., 1993). Recently, Weber et al (2015) clearly demonstrated that neutrophils are critically involved in both the sensitisation and elicitation phase of contact hypersensitivity.

9. Test methods such as the IL-8 Luc assay, which are based on DC-like cell lines and measure IL-8 as a marker of DC activation, are therefore considered relevant for the assessment of the skin sensitisation potential of chemicals. However, since DC activation represents only one key event of the skin sensitisation AOP, the information generated with test methods that measure markers of DC activation is not always sufficient for assessing the skin sensitisation potential of test chemicals. Therefore, data generated with the IL-8 Luc assay should be considered in the context of integrated approaches, such as IATA, and combined with other complementary information—e.g. data 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.

10. The test method described in this Test Guideline can be used to support the discrimination between skin sensitisers (i.e. UN GHS Category 1) and non-sensitisers in the context of IATA. This Test Guideline

cannot be used on its own to subcategorise skin sensitisers into UN GHS Category 1A and 1B (United Nations UN, 2013) by authorities implementing these two optional subcategories. Nor can it be used on its own to predict sensitisation potential for safety assessment. Depending on the regulatory framework, however, a positive result in the IL-8 Luc assay could be sufficient on its own to classify a chemical as UN GHS Category 1.

11. The IL-8 Luc assay proved to be transferable to laboratories experienced in cell culture techniques and luciferase measurement. Within- and between-laboratory reproducibility is roughly 80% (JaCVAM, 2016a). The data generated in the validation study (JaCVAM, 2016a) and other published studies (Kimura et al., 2015; Takahashi et al., 2011) shows that, when using LLNA results as a base, accuracy in distinguishing skin sensitisers (UN GHS Cat. 1) from non-sensitisers (UN GHS No Category) is 80% (N = 143) with a sensitivity of 86% (92/107) and a specificity of 64% (23/36). False negative predictions with the IL-8 Luc assay are more of a concern for chemicals showing 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) (Kimura et al., 2015). Taken together, this information indicates the usefulness of the IL-8 Luc assay to contribute to the identification of skin sensitisation hazards. However, the accuracy values given here for the IL-8 Luc assay 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 paragraph 9 above. 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 the species of interest, i.e. humans.

12. The term "test chemical" is used in this Test Guideline to refer to what is being tested[1] and is not related to the applicability of the IL-8 Luc assay to the testing of mono-constituent substances, multi-constituent substances and/or mixtures. On the basis of the current data 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 (JaCVAM, 2016a; Kimura et al., 2015). Limited information is currently available on the applicability of the IL-8 Luc assay to multi-constituent substances/mixtures. The test method is 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. 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, detection of their potential sensitisation properties may be affected by the cytotoxicity of the constituents.

[1] 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.

13. Negative results should be interpreted with caution for respiratory sensitisers such as anhydrides, including phthalic anhydride, trimellitic anhydride, maleic anhydride, hexahydrophthalic anhydride and methyltetrahydrophthalic anhydride. In general, although the IL-8 Luc assay can correctly judge pre/prohaptens, negative results for pre/prohaptens should be interpreted with caution. The IL-8 Luc assay correctly identified chemicals with a Log $K_{o/w}$ of greater than 8 and those with a water solubility of around 0.1 µg/mL. However, chemicals with a water solubility of less than 100 µg/mL and a minimum inhibition index for SLR-LA (I.I.-SLR-LA) of more than 0.8 might not dissolve in the solvent and their data should not be used in the assessment. A major limitation of the IL-8 Luc assay is that it identifies detergents as sensitisers irrespective of their classification. Therefore, detergents should be excluded from the applicability domain. Finally, test chemicals that interfere with the luciferase enzyme can confound the activity of luciferase in cell-based assays and cause either apparent inhibition or increased luminescence (Thorne et al., 2010). For example, phytoestrogen concentrations higher than 1 µM were reported to interfere with the luminescence signals in other luciferase-based reporter gene assays due to over-activation of the luciferase reporter gene. As a consequence, luciferase expression obtained at high concentrations of phytoestrogens or similar compounds suspected of producing phytoestrogen-like over-activation of the luciferase reporter gene need to be examined carefully (OECD, 2012b). In cases where evidence can be demonstrated on the non-applicability of the Test Guideline to other specific categories of test chemicals, the test method should not be used for those specific categories.

14. If the substances used in the validation study which are outside the applicability domain are excluded i.e. 1) detergents, such as hexadecyltrimethylammonium bromide, Tween-80, SLS, octanoic acid and benzalkonium chloride as well as anhydrides, 2) pre/prohaptens that were judged negative, and 3) chemicals with a water solubility of less than 100 µg/mL and a minimum I.I.-SLR-LA of more than 0.8, then the IL-8 Luc demonstrates an accuracy of 89%, a sensitivity of 94%, and a specificity of 74% for 129 chemicals. Furthermore, when human sensitisers are specifically considered, it yields an accuracy of 89%, a sensitivity of 91%, and a specificity of 80% for 129 chemicals.

PRINCIPLE OF THE TEST

16. The IL-8 Luc assay makes use of a human macrophage-like 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 harbors 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

(Takahashi et al., 2011). 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.

17. The dual-color assay system comprises an orange-emitting luciferase (SLO; λmax = 580 nm) (Viviani et al., 2001) for the gene expression of the IL-8 promoter as well as a red-emitting luciferase (SLR; λmax = 630 nm) (Viviani et al., 1999) for the gene expression of the internal control promoter, GAPDH. The two luciferases emit different colors 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 (Nakajima et al., 2005) (Annex II).

18. THP-G8 cells are treated for 16 hours with a test chemical, after which SLO luciferase activity (SLO-LA) and SLR luciferase activity (SLR-LA) are measured. Three parameters are derived from the measures: normalised SLO luciferase activity (nSLO-LA), which is the ratio of SLO-LA to SLR-LA; the fold induction of nSLO-LA (FInSLO-LA), which is the ratio of the nSLO-LA of chemically treated cells to that of non-treated cells, and Inhibition Index for SLR-LA (I.I.-SLR-LA), which is the ratio of the SLR-LA of chemically treated cells to that of non-treated cells.

19. Performance standards (PS) (in preparation) 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 (OECD, 2005).

DEMONSTRATION OF PROFICIENCY

20. Prior to routine use of the test method described in this Test Guideline, laboratories should demonstrate technical proficiency, using the 10 Proficiency Substances listed in Annex III. Moreover, test method users should maintain a historical database of data generated with the reactivity checks (see paragraph 23) and with the positive and solvent/vehicle controls (see paragraphs 27-30), and use these data to confirm the reproducibility of the test method in their laboratory is maintained over time.

PROCEDURE

21. Currently, the only test method covered by this Test Guideline is the scientifically valid IL-8 Luc assay

(Kimura et al., 2015; Takahashi et al., 2011). The Standard Operating Procedures (SOP) for the IL-8 Luc assay is available and should be employed when implementing and using the test method in the laboratory (JaCVAM, 2016b). Laboratories willing to implement the test method can obtain the recombinant cell line used in the IL-8 Luc assay by establishing a license agreement with the assay developer. The following paragraphs provide a description of the main components and procedures of the IL-8 luciferase assay.

Preparation of THP-G8

22. A transgenic cell line having a stable insertion of the luciferase reporter gene under the control of the SLO and SLR should be used (e.g. the THP-G8 cell line). Upon receipt, cells are propagated (e.g. 2 to 4 passages) and stored frozen as a homogeneous stock. Cells from this original stock can be propagated up to a maximum of 12 passages or for a maximum of 6 weeks and are employed for routine testing using a culture medium. The culture medium is RPMI-1640 contaiing 10% fetal bovine serum (FBS), antibiotic and antimycotic solution (100 U/mL of penicillin G, 100 µg/mL of streptomycin and 0.25 µg/mL of amphotericin B in 0.85% saline) (e.g. GIBCO Cat#15240-062), 0.15 µg/mL Puromycin (e.g. CAS:58-58-2) and 300 µg/mL G418 (e.g. CAS:108321-42-2).

23. Prior to testing, the cells should be qualified by conducting a reactivity check. The reactivity check should be performed one or two weeks 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 FIn-SLO-LA (≥1.4), while LA should produce a negative response to FIn-SLO-LA (<1.4). Only the cells that pass the reactivity check are used for the assay. The reactivity check is performed according to the procedures described in paragraphs 28-30.

24. For testing, THP-G8 cells are seeded at a density of 2-5 × 10^5 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 resuspended with RPMI-1640 containing 10% FBS, at 1 × 10^6 cells/mL. Then, cells are distributed into a 96 well flat-bottom black plate with 50 µL (5 × 10^4 cells/well) (e.g. a 96-well µclear-black plate (Greiner Bio-one Cat#655090 Costar Cat#3603 or Nunc Cat#165305)).

Preparation of the test chemical and control substances

25. Test chemicals and control substances are prepared on the day of the test. Each test chemical is dissolved in X-VIVO™ 15 that contains gentamycin, L-glutamine and phenol red (Lonza, 04-418Q) to the final desired concentration of 20 mg/mL. The incubation time is at most 16 hours and no sterile filtration is

needed. Test chemicals that are soluble in X-VIVO™ 15 are diluted 5 times in X-VIVO™ 15 and these solutions are used as the X-VIVO™ 15 stock solution of the test chemicals. Test chemicals that are not soluble in X-VIVO™ 15 are rotated for more than 30 minutes and centrifuged at 15,000 rpm (≈20,000 x g) for 5 minutes, after which these supernatants are used as the X-VIVO™ 15 stock solution of the test chemicals. A sufficient scientific rationale should be provided for the use of solvents other than X-VIVO™ 15, such as dimethyl sulfoxide (DMSO), water, or the culture medium.

26. In the first test, 11 serial dilutions of the X-VIVO™ 15 stock solutions of the test chemicals are made at a common ratio of 2 using X-VIVO™ 15. After this, 50 µL/well of the dilution series are added to the cells in a 96 well flat-bottom black plate. For test chemicals that are soluble in X-VIVO™ 15, the final concentrations of the test chemicals are from 0.002 to 2 mg/mL in the X-VIVO™ 15 solutions: culture medium=1:1. For test chemicals that are not soluble in X-VIVO™ 15, the final concentrations of the test chemicals are dependent on the saturated concentration of the test chemicals in the X-VIVO™ 15 stock solution. In the subsequent tests (e.g. the 2nd, 3rd, and 4th tests), the X-VIVO™ 15 stock solution is diluted so that the highest final concentration is double the concentration at which I.I.-SLR-LA falls below 0.05 in the 1st test. After this, the diluted stock solutions of the test chemicals are made at a common ratio of 1.5 using X-VIVO™ 15. Then, 50 µL/well of the dilution series are added to cells in a 96 well flat-bottom black plate. Each concentration of each chemical should be tested using four wells.

27. The solvent control is X-VIVO™ 15 tested at a single final concentration of 50% in the plate. It undergoes the same dilution as described in paragraph 26.

28. The 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 microfuge tube is vortexed vigorously and shaken on a rotor at a maximum speed of 8 rpm for more than 30 minutes just prior to centrifugation. After centrifugation at 15,000 rpm (≈20,000 x g) for 5 min, a 4-fold dilution of the supernatant in X-VIVO™ 15 is made, and 500 µL of the diluted supernatant is transferred to the well of a 96 well Assay Block. Two serial dilutions with X-VIVO™ 15 at a common ratio of 2 (1/8 and 1/16) and 50 µL of the solution are added to THP-G8 in a 96 well plate. The plate is shaken on a plate-shaker and incubated in a CO_2 incubator for 16 hours (37°C, 5% CO_2), after which the luciferase activity is measured as described in paragraph 34.

29. The negative control is LA. 20 mg of LA is prepared in a 1.5-mL microfuge tube, to which X-VIVO™ 15 is added up to 1 mL. A five-fold dilution of the supernatant in X-VIVO™ 15 is made, and 500 µL of the diluted supernatant is transferred to the well of a 96 well Assay Block. Two serial dilutions with X-VIVO™ 15 at a common ratio of 2 (1/10 and 1/20) and 50 µL of the solution are added to THP-G8 in a 96 well plate. The plate is shaken on a plate-shaker and incubated in a CO_2 incubator for 16 hours (37°C, 5% CO_2), after

which the luciferase activity is measured as described in paragraph 34.

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

31. The X-VIVO^{TM} 15 solutions described in paragraphs 26-29 are mixed 1:1 (v/v) with the cell suspensions prepared in a 96-well flat-bottomed plate (see paragraph 24). The treated plates are incubated for 16±0.5 hours at 37±1^{o}C in the presence of 5% CO_2. To avoid evaporation of volatile test chemicals contaminating other test chemicals, volatile chemicals should be tested alone on a single plate.

32. Each test chemical, and solvent control, requires between two and four runs to derive a prediction of sensitiser or non-sensitiser based on each of four runs. Each run is performed on a different day with the fresh X-VIVO^{TM} 15 stock solution of test chemicals and independently harvested cells. Cells may come from the same passage.

Luciferase activity measurements

33. Luminescence activity is measured using a 96-well microplate luminometer equipped with optical filters (for example, Phelios (ATTO, Tokyo, Japan), Tristan 941 (Berthold, Bad Wildbad, Germany), and the ARVO series (PerkinElmer, Waltham, MA). It is necessary to calibrate the luminometer for each test to ensure reproducibility (Niwa et al., 2010). Recombinant orange- and red-emitting luciferases are available for this calibration.

34. One hundred µL 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 25°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.

35. Parameters for each concentration are calculated from the measured values, e.g. SLO-LA, SLR-LA, nSLO-LA, FInSLO-LA, I.I.-SLR-LA, the mean ±SD of SLO-LA, the mean ±SD of SLR-LA, the mean ±SD of nSLO-LA, the mean ±SD of FInSLO-LA, the mean ±SD of I.I.-SLR-LA, and the 95% confidence interval of FInSLO-LA. Definitions of the parameters used in paragraph 35 are provided in Annex II.

36. Prior to measurement, color discrimination in multi-color 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 color should be calibrated prior to testing, per Annex II.

DATA AND REPORTING

Data evaluation

37. Criteria for identifying a test chemical as a sensitiser

In each run:

Criterion: Test chemicals showing a FInSLO-LA ≥ 1.4 as well as satisfying a statistical condition in which the lower limit of the 95% confidence interval of FInSLO-LA ≥ 1.0 are judged positive. On the other hand, test chemicals with a FInSLO-LA < 1.4 and/or the lower limit of the 95% confidence interval of FInSLO-LA < 1.0 are judged negative.(Annex IV).

Prediction model

38. Test chemicals that provide two positive results from among the 1st, 2nd, 3rd or 4th runs are identified as sensitisers and test chemicals that provide three negative results from among the 1st, 2nd, 3rd or 4th runs are identified as non-sensitisers. Once identification is made, no further run is necessary, as shown in the table below.

Table 1. Criteria for identifying sensitisers and non-sensitisers

1st run	2nd run	3rd run	4th run	Judgment
Positive	Positive	-	-	Sensitiser
	Negative	Positive	-	Sensitiser
		Negative	Positive	Sensitiser
			Negative	Non-sensitiser
Negative	Positive	Positive	-	Sensitiser
		Negative	Positive	Sensitiser
			Negative	Non-sensitiser
	Negative	Positive	Positive	Sensitiser
			Negative	Non-sensitiser
		Negative	-	Non-sensitiser

Acceptance criteria

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

- FInSLO-LA should be more than 5.0 at some concentration of the positive control, 4-NBB, in each run.

- FInSLO-LA should be less than 1.4 at some concentration of the negative control, lactic acid, in each run.
- Data from plates for which the SLR-LA of control wells without chemicals is less than 1,000 should be rejected.
- Data from plates for which the I.I.-SLR-LA 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 that the highest final concentration of the repeated test is the lowest final concentration of the previous test.

Test report

40. 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, in case X-VIVO™ 15 has not been used.

-Multi-constituent substance, UVCB and mixture:

- Characterisation by 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 to testing;
- 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, in case 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.

Assay 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 multi-well 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 Annex 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 runs performed;
- Application procedure and exposure time (if different from the 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 SLO-LA and SLR-LA;
- Calculations for nSLO-LA, FInSLO-LA, and I.I.-SLR-LA;
- The 95% confidence interval of FInSLO-LA;
- 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 within the context of an IATA, if other relevant information is available.

Conclusion

LITERATURE

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].

Adler, S., Basketter, D., Creton, S., Pelkonen, O., van Benthem, J., Zuang, V., Andersen, K.E., Angers-Loustau, A., Aptula, A., Bal-Price, A., Benfenati, E., Bernauer, U., Bessems, J., Bois, F.Y., Boobis, A., Brandon, E., Bremer, S., Broschard, T., Casati, S., Coecke, S., Corvi, R., Cronin, M., Daston, G., Dekant, W., Felter, S., Grignard, E., Gundert-Remy, U., Heinonen, T., Kimber, I., Kleinjans, J., Komulainen, H., Kreiling, R., Kreysa, J., Leite, S.B., Loizou, G., Maxwell, G., Mazzatorta, P., Munn, S., Pfuhler, S., Phrakonkham, P., Piersma, A., Poth, A., Prieto, P., Repetto, G., Rogiers, V., Schoeters, G., Schwarz, M., Serafimova, R., Tahti, H., Testai, E., van Delft, J., van Loveren, H., Vinken, M., Worth, A., Zaldivar, J.M., 2011. Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. Arch Toxicol 85, 367-485.

Aiba, S., 2007. Dendritic cells: importance in allergy. Allergol Int 56, 201-208.

Aiba, S., Katz, S.I., 1990. Phenotypic and functional characteristics of in vivo-activated Langerhans cells. J Immunol 145, 2791-2796.

Aiba, S., Tagami, H., 1999. Dendritic cell activation induced by various stimuli, e.g., exposure to microorganisms, their products, cytokines and simple chemicals as well as adhesion to extracellular matrix. J Dermatol Sci 20, 1-13.

Aiba, S., Terunuma, A., Manome, H., 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, 3031-3038.

Bauch, C., Kolle, S.N., Ramirez, T., Eltze, T., Fabian, E., Mehling, A., Teubner, W., van Ravenzwaay, B., Landsiedel, R., 2012. Putting the parts together: combining in vitro methods to test for skin sensitizing potentials. Regul Toxicol Pharmacol 63, 489-504.

Caux, C., Vanbervliet, B., Massacrier, C., Azuma, M., Okumura, K., Lanier, L.L., 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, 1841-1847.

ECETOC, 2003. Contact sensitisation: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals (Technical Report No. 87).

JaCVAM, 2016a. IL-8 Luc assay Validation Study Report, Accessible at. http://www.jacvam.jp/en/index.html.

JaCVAM, 2016b. IL-8 Luc assay protocol, Accessible at. http://www.jacvam.jp/en/index.html.

Jaworska, J., Dancik, Y., Kern, P., Gerberick, F., Natsch, A., 2013. Bayesian integrated testing strategy to assess skin sensitization potency: from theory to practice. J Appl Toxicol 33, 1353-1364.

Kimura, Y., Fujimura, C., Ito, Y., Takahashi, T., Nakajima, S., Ohmiya, Y., Aiba, S., 2015. Optimization of the IL-8 Luc assay as an in vitro test for skin sensitization. Toxicol In Vitro 29, 1816-1830.

Leonard, E.J., Skeel, A., Yoshimura, T., Noer, K., Kutvirt, S., Van Epps, D., 1990. Leukocyte specificity and binding of human neutrophil attractant/activation protein-1. J Immunol 144, 1323-1330.

Mitjans, M., Galbiati, V., Lucchi, L., Viviani, B., Marinovich, M., Galli, C.L., Corsini, E., 2010. Use of IL-8 release and p38 MAPK activation in THP-1 cells to identify allergens and to assess their potency in vitro. Toxicol In Vitro 24, 1803-1809.

Nakajima, Y., Kimura, T., Sugata, K., Enomoto, T., Asakawa, A., Kubota, H., Ikeda, M., Ohmiya, Y., 2005. Multicolor luciferase assay system: one-step monitoring of multiple gene expressions with a single substrate. Biotechniques 38, 891-894.

Niwa, K., Ichino, Y., Kumata, S., Nakajima, Y., Hiraishi, Y., Kato, D., Viviani, V.R., Ohmiya, Y., 2010. Quantum yields and kinetics of the firefly bioluminescence reaction of beetle luciferases. Photochem Photobiol 86, 1046-1049.

Nukada, Y., Miyazawa, M., Kazutoshi, S., Sakaguchi, H., Nishiyama, N., 2013. Data integration of non-animal tests for the development of a test battery to predict the skin sensitizing potential and potency of chemicals. Toxicol In Vitro 27, 609-618.

OECD, 1992, Skin Sensitisation, OECD Guidelines for Chemical Testing No. 457. OECD, Paris. http://www.oecd.org/env/testguidelines.

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.

OECD, 2010a. 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.

OECD, 2010b. 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.

OECD, 2010c. Skin Sensitisation: BrdU-ELISA: DA. OECD Guideline for the Testing of Chemicals No. 442B, Organisation for Economic Cooperation and Development, Paris. Available at: http://www.oecd.org/env/testguidelines.

OECD, 2012a. 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, 2012b. BG1Luc Estrogen Receptor Transactivation Test Method for Identifying Estrogen Receptor Agonists and Antagonists. OECD Guidelines for Chemical Testing No. 457. OECD, Paris. Available at: [http://www.oecd.org/env/testguidelines%5D.

OECD, 2015a. OECD Guideline for the Testing of Chemicals No. 442C: In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA). Paris, France: Organisation for Economic Cooperation and Development. Available at: http://www.oecd.org/env/testguidelines.

OECD, 2015b. 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.

OECD, 2016. OECD draft Guideline for the Testing of Chemicals: In Vitro Skin Sensitisation: h-CLAT Test Method. Paris, France: Organisation for Economic Cooperation and Development.

Reisinger, K., Hoffmann, S., Alepee, 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., 2015. Systematic evaluation of non-animal test methods for skin sensitisation safety assessment. Toxicol In Vitro 29, 259-270.

Sasaki, Y., Aiba, S., Aiba, S., 2007. Dendritic cells and contact dermatitis. Dendritic cells: importance in allergy. Clin Rev Allergy Immunol 33, 27-34.

Sebok, K., Woodside, D., al-Aoukaty, A., Ho, A.D., Gluck, S., Maghazachi, A.A., 1993. IL-8 induces the locomotion of human IL-2-activated natural killer cells. Involvement of a guanine nucleotide binding (Go) protein. J Immunol 150, 1524-1534.

Steinman, R.M., 1991. The dendritic cell system and its role in immunogenicity. Annu Rev Immunol 9, 271-296.

Takahashi, T., Kimura, Y., Saito, R., Nakajima, Y., Ohmiya, Y., Yamasaki, K., 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, 359-369.

Thorne, N., Inglese, J., Auld, D.S., 2010. Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical biology. Chem Biol 17, 646-657.

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.

van der Veen, J.W., Rorije, E., Emter, R., Natsch, 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.

Viviani, V., Uchida, A., Suenaga, N., Ryufuku, M., Ohmiya, Y., 2001. Thr226 is a key residue for bioluminescence spectra determination in beetle luciferases. Biochem Biophys Res Commun 280, 1286-1291.

Viviani, V.R., Bechara, E.J., Ohmiya, Y., 1999. Cloning, sequence analysis, and expression of active Phrixothrix railroad-worms luciferases: relationship between bioluminescence spectra and primary structures. Biochemistry 38, 8271-8279.

Weber, F.C., Nemeth, T., Csepregi, J.Z., Dudeck, A., Roers, A., Ozsvari, B., Oswald, E., Puskas, L.G., Jakob, T., Mocsai, A., Martin, S.F., 2015. Neutrophils are required for both the sensitization and elicitation phase of contact hypersensitivity. J Exp Med 212, 15-22.

ANNEX 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 (OECD, 2005).

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 (OECD, 2012a).

FInSLO-LA: The fold induction of normalised SLO-LA. It is obtained by dividing the nSLO-LA 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.

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.

I.I.-SLR-LA: The Inhibition Index of SLR-LA. It is obtained by dividing SLR-LA of THP-G8 treated with chemicals with SLR-LA of non-treated THP-G8 and represents cytotoxicity of chemicals.

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

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 are present in a concentration ≥ 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.

nSLO-LA: The SLO luciferase activity (SLO-LA) normalised by SLR luciferase activity (SLR-LA). It represents IL-8 promoter activity after considering cell viability or cell number.

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 (OECD, 2005).

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 (OECD, 2005).

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 (OECD, 2005).

SLO-LA: Luciferase Activity of Stable Luciferase Orange (SLO) (λmax = 580 nm), regulated by IL-8 promoter.

SLR-LA: Luciferase Activity of Stable Luciferase Red (SLR) (λmax = 630 nm), regulated by GAPDH promoter and demonstrates cell viability and viable cell number.

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 (OECD, 2005).

Substance: Chemical elements and their compounds in the natural state or obtained by any production process, inducing 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 it composition.

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 (UN, 2013).

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.

ANNEX II

Determination of the transmission coefficients of optical filter for SLO and SLR

Principal of measurement of luciferase activity

MultiReporter Assay System -Tripluc- can be used with a microplate-type luminometer with a multi-color 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-color 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 (λmax = 580 nm) and SLR (λmax = 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	kO_{R60}	The filter's transmission coefficient for the SLO
SLR	Filter 1 Transmission coefficients	kR_{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 transmit 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 (κO_{R60} and κR_{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

· Assay reagent:

Tripluc® Luciferase assay reagent (TOYOBO Cat#MRA-301)

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

Reagent	Company	Conc.	Final conc. in medium	Required amount
RPMI-1640	GIBCO #11875-093	-	-	27 mL
FBS	Biological Industries Cat#04-001-1E Lot:715004	-	10 %	3 mL

2 Preparation of luminescence reaction solution

Thaw Tripluc® 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.

Add 200 μL of 10–100 mM Tris/HCl or Hepes/HCl (pH 7.5–8.0) supplemented with 10% (w/v) glycerol to each tube of lyophilised purified luciferase sample to dissolve the enzymes, divide into 10 μL aliquots in 1.5 mL disposable tubes and store in a freezer at -80°C. The stored frozen solution of the reference samples can be used for up to 6 months.

Add 1 mL of B medium to each tube of frozen reference sample (10 μL sample per tube). Keep the reference samples on ice to prevent deactivation.

3 Bioluminescence measurement

Transfer 100 μL of the diluted reference samples to a black 96 well plate (flat bottom) (the SLO reference sample to #B1, #B2, #B3, the SLR reference sample to #D1, #D2, #D3).

Transfer 100 μL of pre-warmed Tripluc to each well of the plate containing the reference sample using a pipetman. Shake the plate for 10 minutes at an ambient temperature of about 25°C using a plate-shaker. Remove bubbles in 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 (κO_{R60}))= (#B1 of F1+ #B2 of F1+ #B3 of F1) / (#B1 of F0+ #B2 of F0+ #B3 of F0)

Transmission coefficient (SLR (κR_{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

In order to confirm the detector stability as the quality control, the reference luciferase sample, optical property, the protocol described hereafter should be performed at the beginning of the experiments every day using reference LED light source plates equipped with stabilised LEDs. LED plate data typically fluctuates up to 1.5% (σ). Disagreement to the old data should be less than 3×σ (= 4.5%). Here describes an example using a commercially available reference light source, TRIANT® (wSL-0001) by ATTO (Tokyo, Japan).

Start LED plate and select “PMT” mode.

Select three-color (BRG) mode and adjust light intensity to 1/10 (10E-1).

Place the LED plate into the luminometer. Light intensity is measured for 3 sec each in the absence (F0) and presence (F2) of the optical filter.

Read and record the data at the position of #F6, #E6, and #D6 that are LEDs of blue, green, and red LEDs, respectively.

LED plate data typically fluctuates up to 1.5% (σ). Disagreement to the old data should be less than 3×σ (= 4.5%).

ANNEX III

PROFICIENCY SUBSTANCES

Prior to routine use of a test method that adheres to this Test Guideline, laboratories should demonstrate technical proficiency by correctly 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. 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 IL-8 Luc assay are available. Also, published reference data are available for the IL-8 Luc assay (Kimura et al., 2015; Takahashi et al., 2011).

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

Proficiency Substances	**CASRN**	**Physical State**	***GHS* Prediction (1)**	**IL-8 Luc Prediction (2)**
2-Mercaptobenzothiazole	149-30-4	Solid	1A	Positive
2,4-Dinitrochlorobenzene	97-00-7	Solid	1A	Positive
4-Nitrobenzylbromide	100-11-8	Solid	1A	Positive
Ethyleneglycol dimethacrylate	97-90-5	Liquid	1B	Positive
4-Allylanisole (Estragol)	140-67-0	Liquid	1B	Positive
Ethylenediamine	107-15-3	Liquid	1B	Positive
Salicylic acid	69-72-7	Solid	NO	Negative
4-Aminobenzoic acid	150-13-0	Solid	NO	Negative
Glycerol	56-81-5	Liquid	NO	Negative
Isopropanol	67-63-0	Liquid	NO	Negative

Abbreviations: CAS RN = Chemical Abstracts Service Registry Number

[1] The *in vivo* potency is derived using the criteria proposed by ECETOC (ECETOC, 2003).

[2] Based on historical observed values (Kimura et al., 2015; Takahashi et al., 2011).

ANNEX IV

Indexes and judgment criteria

nSLO-LA

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

$nSLO\text{-}LA_{ij} = SLO\text{-}LA_{ij} / SLR\text{-}LA_{ij.}$

This is the basic unit of measurement in this assay.

FInSLO-LA

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

$$FInSLO\text{-}LA_i = \left\{(1/4)\times\sum_j nSLO\text{-}LA_{ij}\right\} \ / \ \left\{(1/4)\times\sum_j nSLO\text{-}LA_{0j}\right\}.$$

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 followed to use the value through all the phase of present validation study.

The primary outcome measure, FInSLO-LA, 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 ≥ 1 indicates that the nSLO-LA 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 kwon 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 $A=\bar{x}_0^2-t^2_{0.975(\nu)}\times\frac{sd_0^2}{n_0}$, $B=-2\times\bar{x}\times\bar{y}$, $C=\bar{y}_i^2-t^2_{0.975(\nu)}\times\frac{sd_{y_i}^2}{n_{y_i}}$, and

$\bar{x}_0^2=\left\{(1/4)\times\sum_j nSLO\text{-}LA_{0j}\right\}$, $sd_0^2=(1/3)\times\sum_j\left(nSLO-LA_{0j}-\bar{x}_0\right)^2$, n_0 =4,

$\bar{y}_0^2=\left\{(1/4)\times\sum_j nSLO\text{-}LA_{ij}\right\}^2$, $sd_{y_j}^2=(1/3)\times\sum_j\left(nSLO-LA_{ij}-\bar{y}_i\right)^2$, n_{y_i} =4,

$t_{0.975(\nu)}$ is 97.5 percentile of the central t distribution with the ν of the degree of freedom.

I.I.-SLR-LA

The I.I.-SLR-LA is a ratio of the averaged SLR-LA for the repetition of the i-th concentration compared with that with solvent control, and this is written by

$$\text{I.I.-SLR-LA}_i = \left\{(1/4)\times\sum_j \text{SLL-LA}_{ij}\right\} / \left\{(1/4)\times\sum_j \text{SLL-LA}_{0j}\right\}.$$

Since the SLR-LA is the denominator of the nSLO-LA, an extremely small value causes large variation of the nSLO-LA. Therefore, the i-th FInSLO-LA value with an extremely small value of I.I.- SLR-LA might be considered poor precision.