



Organisation for Economic Co-operation and Development

**ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND THE WORKING PARTY ON
CHEMICALS, PESTICIDES AND BIOTECHNOLOGY**

**DRAFT UPDATED Performance Standards for the assessment of proposed similar or
modified in vitro skin sensitisation DPRA and ADRA test methods**

FOREWORD

This document describes the Performance Standards (PS) for the assessment of proposed similar or modified methods to the Direct Peptide Reactivity Assay (DPRA) and the Amino acid Derivative Reactivity Assay (ADRA), both included in TG 442C for *in chemico* skin sensitisation assays addressing the Adverse Outcome Pathway Key Event on covalent binding to proteins. They are intended for the developers of new or modified similar test methods.

TG 442C was updated in 2019 with the inclusion of a new test method, the ADRA. This project was led by Japan, who also developed the present Performance Standards, with the collaboration of the expert group on skin sensitisation. The PS were circulated to the Working Group of the National Coordinators of the Test Guidelines Programme (WNT), together with the draft updated Test Guideline 442C, for review and comments in September and December 2018; they were revised accordingly.

The WNT approved the Performance Standards at its 31st meeting in April 2019. The Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology agreed to the declassification of the Performance Standards on 20 June 2019. This document is published under the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology.

APPENDIX II (ADRA) of TG 442C was revised in 2022 with the inclusion of a change in the concentration of the test chemical solution and the addition of a measurement method by fluorescence detection. The PS was also updated in line with this revision.

INTRODUCTION

1. Performance standards (PS) have been developed to facilitate the validation of proposed similar or modified test methods based on the Direct Peptide Reactivity Assay (DPRA) and the Amino acid Derivative Assay (ADRA) and to allow for their timely inclusion in the Test Guidelines (1) (2). Proposed similar or modified test methods based on in chemico covalent binding to proteins will only be added to the Test Guideline, however, after a review process to confirm that all criteria stipulated in the PS for similarity to the validated reference methods (VRM)—namely, DPRA and ADRA—have been met, that the proposed similar or modified test method includes all essential test method components, and that test performance achieves the target values for reproducibility and predictive capacity of the proposed reference chemicals. 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.
2. The purpose of these Performance Standards (PS) is to provide a basis by which proposed similar or modified test methods, both proprietary (i.e., copyrighted, trademarked, registered) and non-proprietary, can demonstrate sufficient reliability and relevance for testing purposes. The PS, based on a scientifically valid and accepted test method, can be used to evaluate the reliability and relevance of other analogous test methods (colloquially referred to as “me-too” test methods) that are based on similar scientific principles and measure or predict the same biological or toxic effect. (3). In addition, modified test methods which propose potential improvements to an approved test method should be evaluated to determine the effect of the proposed modifications on the test method’s performance and the extent to which such modifications affect the information available for the other components of the validated reference methods. Depending on the number and nature of the proposed modifications as well as the data and documentation available to support the modifications, proposed similar or modified test methods should either be subjected to the same validation process as any new test method or, where appropriate, to a limited assessment of reliability and relevance using established PS (3).
3. Similar (me-too) or modified test methods proposed for use under TG 442C for a test method based on in chemico covalent binding to proteins (1) (2) should be evaluated to determine their reliability and relevance using a set of reference chemicals (Table 1) that represent the full range of in vivo skin sensitisation effects. The proposed similar or modified test methods should demonstrate reliability, accuracy, sensitivity, and specificity values that are at least as good as those derived from the VRM—DPRA and ADRA—and as described below in paragraphs 8 to 12. The reliability of the proposed similar or modified test method as well as its ability to correctly predict the skin sensitisation potential of test chemicals should be validated prior to its use in testing chemicals.

These PS comprise the following three elements:

- I) Essential test method components
- II) Minimum list of reference chemicals
- III) Defined reliability and accuracy values

ESSENTIAL TEST METHOD COMPONENTS

4. The Essential Test Method Components comprise the essential structural, functional, and procedural elements of a VRM that should be included in the protocol of a proposed, mechanistically and functionally similar or modified test method. These components include unique characteristics of the test method, critical procedural details, and quality control measures. Adherence to essential test method components will help to assure that a similar or modified proposed test method is based on the same concepts as the corresponding VRM (3). The essential test method components are described in detail in the following paragraphs.
 - Proteins, peptides, amino acids, and their derivatives that are relevant to covalent binding to proteins in the skin sensitisation process should be used as nucleophilic reagents in the assay based on covalent binding to proteins.
 - This test method is based on the principle that, since skin sensitisers undergo *in vivo* covalent binding to proteins, skin sensitisation potential can be predicted by assessing whether or not a test chemical undergoes *in chemico* covalent binding with a nucleophilic reagent containing thiol groups like cysteine or amino groups like lysine.
 - Nucleophilic reagents containing thiol groups are susceptible to the formation of oxidative dimers, which can significantly compromise the quality of test results.

MINIMUM LIST OF REFERENCE CHEMICALS

5. Reference chemicals are used to determine if the reproducibility and predictive capacity of a proposed similar or modified test method that has been shown to be sufficiently similar, both structurally and functionally, to the VRM or represents only a minor modification of the VRM and are at least as good as that of the VRM (4) (5) (6). The recommended reference chemicals listed in Table 1 represent the full range of *in vivo* skin sensitisation effects that act via a variety of mechanisms and are representative of different chemical categories based on their functional groups. The chemicals included in this list include skin sensitisers of various potencies based on LLNA EC3 values—e.g., weak, moderate, strong, and extreme—as well as non-sensitizers. These chemicals were selected from those used in the validation studies of the VRM and evaluated during the independent peer reviews conducted by EURL ECVAM and JaCVAM (4) (5) (6).
6. The 20 reference chemicals listed in Table 1 represent the minimum number of chemicals that should be used to evaluate the reproducibility and predictive capacity of a proposed similar or modified test method to distinguish skin sensitizers from non-sensitizers. These 20 reference chemicals were selected from the 40 test chemicals used in the ADRA validation study with 1 mM test chemical solution and the 12 test chemicals used in the ADRA ring-study with 4 mM test chemical solution. 10 of 20 chemicals were also used in the DPRA validation study. The figures for reproducibility and predictive capacity given in paragraphs 10, 11, and 12, however, are based only on ADRA results. All 20 reference chemicals listed in Table 1 should be used to assess the predictive capacity and between-laboratory reproducibility (BLR) of the proposed similar or modified test method to distinguish skin sensitizers from non-sensitizers, including 13 sensitizers of various potencies and 7 non-sensitizers. In contrast to this, the within-laboratory reproducibility (WLR) should be assessed on the basis of a subset of 12 of the 20 reference chemicals, which are listed in Table 1 and include 8 sensitizers of various potencies and 4 non-sensitizers. The use of these reference chemicals for the development and optimisation of proposed similar test methods should be avoided. In situations where a listed

chemical is unavailable, it should be substituted with another chemical for which adequate *in vivo* reference data is available, preferably from the chemicals used in the validation of the VRM. To further evaluate the accuracy of the proposed test method, additional chemicals representing other chemical classes and for which adequate *in vivo* reference data are available may be added to the list of reference chemicals. Although benzyl salicylate (No. 6) and *m*-aminophenol (No.14) are known to be a moderate sensitisers, these two chemicals were both predicted to be non-sensitisers in some results for both DPRA and ADRA. Also, although benzyl cinnamate (No. 17) is known to be a weak sensitiser, this chemical was predicted to be non-sensitiser in both DPRA and ADRA. Yet their chemical structures are such that it is hard to conceive of either reacting strongly with thiol or amino groups. One possible explanation of their sensitization potential for benzyl salicylate and benzyl cinnamate is that, since both these chemicals have an ester structure in common, *in vivo* hydrolysis of these esters gives chemicals that become sensitisers after undergoing oxidative metabolism. *m*-Aminophenol is known as pre-/pro-hapten, and it considered to become a sensitiser by *in vivo* oxidative metabolism. Thus, although correctly predicted to be sensitisers in LLNA testing, both these chemicals gave false negative results when tested using DPRA and ADRA. Table 2 summarises the ranges of each depletion obtained from DPRA and ADRA performed and published in the past and the number of tests used for 20 test chemicals shown in Table 1. These ranges should be used as a reference when performing the DPRA and ADRA.

Table 1: List of reference chemicals for determination of reproducibility (12 chemicals for WLR, 20 chemicals for BLR) and predictive capacity (20 chemicals) in a proposed similar or modified protein reactivity assay

No.	Test chemicals	CAS No.	Physical state	Molecular weight	Mechanism	LLNA EC3 (%)	<i>in vivo</i> prediction ¹	DPRA prediction	ADRA prediction
12 Test chemicals for Within-Laboratory Reproducibility and Between-Laboratory Reproducibility									
1	Diphenylcyclopropenone	886-38-4	Solid	206.24	Acylation	0.0003	Sensitizer (extreme)	Pos ²	Pos ^{2,3}
2	Lauryl gallate	1166-52-5	Solid	338.44	pre-hapten, Michael acceptor	0.3	Sensitizer (strong)	Pos ²	Pos ²
3	2-Methyl-2H-isothiazol-3-one	2682-20-4	Solid	115.15	Michael acceptor, S _N 2	0.4	Sensitizer (strong)	Pos ²	Pos ^{2,3}
4	4-(Methylamino) phenol hemisulfate salt	55-55-0	Solid	221.23	pre-hapten, Michael acceptor	0.8	Sensitizer (strong)	Pos ²	Pos ²
5	2-Mercaptobenzothiazole	149-30-4	Solid	167.25	S _N 2, acylation	1.7	Sensitizer (moderate)	Pos ^{2,4}	Pos ²
6	Benzyl salicylate	118-58-1	Liquid	228.25	S _N 2, acylation	2.9	Sensitizer (moderate)	Pos/Neg ^{2,4}	Neg ²
7	Imidazolidinyl urea	39236-46-9	Solid	388.29	Acylation	24	Sensitizer (weak)	Pos ^{2,4}	Pos ^{2,3}
8	Ethyl acrylate	140-88-5	Liquid	100.12	Michael acceptor	28	Sensitizer (weak)	Pos ²	Pos ²
9	Salicylic acid	69-72-7	Solid	138.12	Non-reactive	-	Non-sensitizer	Pos/Neg ²	Neg ²
10	Propyl paraben	94-13-3	Solid	180.2	Non-reactive	-	Non-sensitizer	Neg ²	Neg ^{2,3}
11	Glycerol	56-81-5	Liquid	92.09	Non-reactive	-	Non-sensitizer	Neg ^{2,4}	Neg ²
12	Isopropanol	67-63-0	Liquid	60.1	Non-reactive	-	Non-sensitizer	Neg ^{2,4}	Neg ^{2,3}
8 Test chemicals for Between-Laboratory Reproducibility									
13	<i>p</i> -Benzoquinone	106-51-4	Solid	108.09	Michael acceptor	0.0099	Sensitizer (extreme)	Pos ^{2,4}	Pos ²
14	<i>m</i> -Aminophenol	591-27-5	Solid	109.13	pro-hapten, Michael acceptor	3.2	Sensitizer (moderate)	Neg ²	Pos/Neg ^{2,3}
15	Palmitoyl Chloride	112-67-4	Liquid	274.87	Acylation	8.8	Sensitizer (moderate)	Pos ²	Pos ^{2,3}
16	Famesal	19317-11-4	Liquid	220.35	Schiff base	12	Sensitizer (weak)	Pos ²	Pos ^{2,3}
17	Benzyl cinnamate	103-41-3	Solid	238.29	Michael acceptor, S _N 2	18	Sensitizer (weak)	Neg ^{2,4}	Neg ²
18	Dimethyl isophthalate	1459-93-4	Solid	194.19	Non-reactive	-	Non-sensitizer	Neg ^{2,4}	Neg ³
19	Methyl salicylate	119-36-8	Liquid	152.15	Non-reactive	-	Non-sensitizer	Pos/Neg ^{2,4}	Neg ²
20	4-Aminobenzoic acid	150-13-0	Solid	137.14	Non-reactive	-	Non-sensitizer	Neg ^{2,4}	Neg ²

Chemicals highlighted in pink were predicted to be sensitisers, those highlighted in blue were predicted to be non-sensitisers, and those highlighted in yellow had non-concordant results.

Molecular weight expressed in g · mol⁻¹.

¹Predictions of *in vivo* hazard (potency) are based on LLNA data (4) (7) (8) (9). *In vivo* potency is derived using criteria proposed by ECTOC. (10).

²Predictions based on published data (7) (8) (9) (11) (12).

³Result of ADRA ring-study (6).

⁴Result of DPRA validation study (4).

Chemicals were selected from the test chemicals used in validation of ADRA with 1 mM test chemical solution and ring-study of ADRA with 4 mM test chemical solution (5) (6). They were first sorted into non-sensitisers and skin sensitisers, then ranked on the basis of their testing purpose and skin sensitisation potency. The selection includes chemicals that

(i) are representative of the range of skin sensitisation potency tested with the VRM (e.g., weak, moderate, strong, and extreme sensitisers as well as non-sensitisers),

(ii) reflect the performance characteristics of the VRM for BLR and predictive capacity,

(iii) have chemical structures that are well-defined,

(iv) include a variety of mechanisms of action, (13) (14) (15),

(v) include a variety of chemical categories based on their organic functional groups,

(vi) induce to the extent possible definitive results in the *in vivo* reference test method,

(vii) are commercially available, and

(viii) are not prohibitively expensive to dispose of.

The *in vivo* categories are based on EC3 values from the LLNA test methods (weak: EC3 > 10%, moderate: EC3 ≥ 1%, strong: EC3 ≥ 0.1%, and extreme: EC3 < 0.1%).

Table 2: Reference range of depletion from DPRA and ADRA.

No.	Test chemicals	DPRA prediction ^{1,2}						ADRA prediction ^{1,3}					
		Cys-peptide		Lys-peptide		Mean depletion (%)	N	NAC		NAL		Mean depletion (%)	N
		Depletion (%)	N	Depletion (%)	N			Depletion (%)	N	Depletion (%)	N		
12 Test chemicals for Within-Laboratory Reproducibility and Between-Laboratory Reproducibility													
1	Diphenylcyclopropanone	99	1	0	1	50	1	64-76	16	2-8	16	33-41	16
2	Lauryl gallate	91	1	9	1	50	1	100	1	87	1	93	1
3	2-Methyl-2H-isothiazol-3-one	98	1	0	1	49	1	78-100	16	< 5	16	39-53	16
4	4-(Methylamino) phenol hemisulfate salt	100	1	45	1	72	1	100	1	19	1	60	1
5	2-Mercaptobenzothiazole	97-100	10	< 9	10	48-55	10	100	1	0	1	50	1
6	Benzyl salicylate	< 15	10	< 13	10	< 9	10	0	1	0	1	0	1
7	Imidazolidinyl urea	3-59	6	< 26	6	2-43	6	58-66	16	< 12	16	30-39	16
8	Ethyl acrylate	96-100	2	90-94	2	95	2	100	1	12	1	56	1
9	Salicylic acid	3-9	3	1-22	2	4-13	2	1	1	0	1	0	1
10	Propyl paraben	< 9	2	0	2	< 5	2	< 3	16	< 4	16	< 4	16
11	Glycerol	< 3	5	< 3	5	< 2	5	0	1	0	1	0	1
12	Isopropanol	< 11	5	< 3	5	< 6	5	< 3	16	< 2	16	< 2	16
8 Test chemicals for Between-Laboratory Reproducibility													
13	<i>p</i> -Benzoquinone	94-100	6	85-100	5	92-100	5	99	1	91	1	95	1
14	<i>m</i> -Aminophenol	7	1	1	1	4	1	2-34	16	< 3	16	2-17	16
15	Palmitoyl Chloride	26	1	27	1	26	1	2-37	16	82-100	16	45-69	16
16	Farnesal	16	1	9	1	12	1	77-100	10	10-31	16	43-66	16
17	Benzyl cinnamate	< 14	10	< 6	10	< 10	10	0	1	0	1	0	1
18	Dimethyl isophthalate	< 7	9	1-5	3	< 3	3	< 3	15	< 2	15	< 2	15
19	Methyl salicylate	< 12	11	< 25	11	< 13	11	1	1	1	1	1	1
20	4-Aminobenzoic acid	< 11	10	< 1	10	< 6	10	1	1	0	1	0	1

¹Predictions based on published data (7) (8) (9) (11) (12).

²Result of DPRA validation study (4).

³Result of ADRA ring-study (6).

N: the number of sources for reference ranges

DEFINED RELIABILITY AND ACCURACY VALUES

7. In order to assess the reliability and relevance of proposed similar or modified test methods based on *in chemico* covalent binding to proteins (1) (2), all reference chemicals listed in Table 1 should be tested. Validation studies based on performance standards should be assessed independently by internationally recognised validation bodies in agreement with international guidelines (3). The 20 reference chemicals should each be tested by at least three laboratories. Within-laboratory reproducibility should be evaluated using the subset of 12 reference chemicals listed in Table 1 to conduct three qualified tests resulting in three predictions at each laboratory. The remaining 8 reference chemicals should be used to conduct a single qualified test resulting in one prediction at each laboratory. Finally, results from all 20 reference chemicals should be used to assess predictive capacity. When the result for 8 reference chemicals used to conduct a single qualified test is closed to the threshold, each qualified test must comprise at least two qualified independent repetitions. If the first two repetitions are concordant, a third repetition is unnecessary. If the first two repetitions are non-concordant, a third repetition is needed to determine the outcome. Each repetition comprises three replicates of the test chemical solution, tested concurrently with three replicates of the negative and positive control reagents.
8. The calculation of values for within-laboratory reproducibility, between-laboratory reproducibility, accuracy, sensitivity, and specificity should be done according to the rules described below to ensure the use of a predefined and consistent approach.
 - a WLR should be calculated based on concordance of predictions made using only qualified test results obtained from the subset of 12 reference chemicals listed in Table 1 for which at least three qualified tests are available.
 - b BLR should be calculated based on concordance of predictions made using only qualified test results obtained from the 20 reference chemicals listed in Table 1 for which at least one qualified test per laboratory is available. For the subset of 12 chemicals that were tested three times each for assessing WLR, a single prediction should be derived based on the results of the three predictions and used to assess BLR.
 - c Values for accuracy should be calculated using all qualified test results obtained from the 20 reference chemicals at each laboratory. The calculations should be based on the individual predictions made for each qualified test result of each reference chemical in each laboratory. Accuracy is given as a percentage, calculated by dividing the sum of all sensitisers that were correctly predicted to be sensitisers and all non-sensitisers that were correctly predicted to be non-sensitisers by the total number (20) of chemicals tested.
9. The calculations should take into account the fact that the 12 chemicals used to assess both BLR and WLR were each tested nine times, whereas the 8 chemicals used to assess only BLR were tested three times each.
10. Test results are considered to be qualified test results if they satisfy the acceptance criteria. Acceptance criteria should be determined based on those from the test methods listed in the Test Guidelines (1) (2).

Within-laboratory reproducibility

11. Assessments of the WLR of the proposed similar or modified test method should demonstrate that at least 80.0% of the predictions obtained from three independent qualified test results for each chemical in the recommended subset of 12 reference chemicals listed in Table 1 are concordant (87% for DPRA per the validation dataset and 98.3% for ADRA per the ring-study dataset (8 proficiency substances and 4 substances with different results at 1 mM and 4 mM test chemical solution)) (4) (6).

Between-laboratory reproducibility

12. Assessments of the BLR of the proposed similar or modified test method should demonstrate that at least 80.0% of the predictions obtained for the 20 reference chemicals shown in Table 1 at a minimum of three laboratories are concordant (85.7% for DPRA per the validation dataset excluding for three substances out of the applicability domain and 100% for ADRA per the ring-study dataset (8 proficiency substances and 4 substances with different results at 1 mM and 4 mM test chemical solution)) (4) (6).

Predictive capacity

13. Assessments of the predictive capacity of the proposed similar or modified test method should be comparable to that of the VRM, and calculations should demonstrate an accuracy, sensitivity and specificity of at least 80.0% for the 20 reference chemicals listed in Table 1. (Accuracy of 84.1%, sensitivity of 79.5%, and specificity of 91.7% for DPRA per the validation dataset excluding for three substances out of the applicability domain and accuracy of 100%, sensitivity of 100%, and specificity of 100% for ADRA per the ring-study dataset (8 proficiency substances and 4 substances with different results at 1 mM and 4 mM test chemical solution)). Predictive capacities for both DPRA and ADRA were calculated on the basis of the full validation and ring-study dataset and are reported in the DPRA validation study and ADRA ring-study reports (4) (6). Also, a clear rationale should be given for any under-predictions (false negatives) of strong or extreme sensitizers.

LITERATURE

1. OECD (2015), Test Guideline on an In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA), OECD Guidelines for the Testing of Chemicals No 442C, OECD, Paris.
2. OECD (2022), Test Guideline on an In Chemico Skin Sensitisation: Amino acid Derivative Reactivity Assay (ADRA), OECD Guidelines for the Testing of Chemicals No 442C, OECD, Paris.
3. 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. Organisation for Economic Cooperation and Development, Paris, France, 2005, 96 pp.
4. EURL ECVAM (2012), Direct Peptide Reactivity Assay (DPRA) ECVAM Validation Study Report. Available at: <https://tsar.jrc.ec.europa.eu/system/files/Published-DPRA%20Validation%20Study%20Report.pdf>
5. OECD (2019), Validation report: Amino acid Derivative Reactivity Assay (ADRA) – JaCVAM Validation Study Report. Series on testing and Assessment n° 304. Organisation for Economic Cooperation and Development, Paris.
6. Fujita, et al. (2021), The within- and between-laboratories reproducibility and predictive capacity of Amino acid Derivative Reactivity Assay using 4 mM test chemical solution: Results of ring study implemented at five participating laboratories. *Journal of Applied Toxicology*, Dec 1 doi: 10.1002/jat.4268. (Epub ahead of print)
7. Natsch et al. (2013), A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation. *Journal of Applied Toxicology*, 33(11):1337-52.
8. G. F. Gerberick et al. (2007), Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. *Toxicological Sciences*, 97(2):417-427.
9. Urbisch D et al. (2015), Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regulatory Toxicology and Pharmacology*, 71(2):337-51
10. ECETOC (2003), Contact sensitization: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals (Technical Report No. 87).
11. Bauch et al. (2012), Putting the parts together: combining in vitro methods to test for skin sensitizing potentials. *Regulatory Toxicology and Pharmacology*, 63(3):489-504.
12. Imamura M et al. (2021), Improving predictive capacity of the Amino acid Derivative Reactivity Assay test method for skin sensitisation potential with an optimal molar concentration of test chemical solution. *Journal of Applied Toxicology*, 41, 303-329.
13. Aptula A.O., Patlewicz G., Roberts D.W., Schultz T.W. (2006), Non-enzymatic glutathione reactivity and in vitro toxicity: a non-animal approach to skin sensitization. *Toxicology In Vitro* 20, 239-247.
14. Casati S., Aeby P., Kimber I., Maxwell G., Ovigne J.M., Roggen E., Rovida C., Tosti L., Basketter D. (2009), Selection of chemicals for the development and evaluation of in vitro methods for skin sensitisation testing. *Altern Lab Anim.* 37, 305-312.
15. Rovida C., Martin S.F., Vivier M., Weltzien H.U., Roggen E. (2013), Advanced tests for skin and respiratory sensitization assessment. *ALTEX* 30, 231-352.