EUROPEAN COMMISSION
JOINT RESEARCH CENTRE

Institute for Health and Consumer Protection
European Centre for the Validation of Alternative Methods (ECVAM)

REPORTING TEMPLATE FOR

DRAFT

ESAC OPINION

on
the ECVAM-led study on the Direct Peptide Reactivity Assay (DPRA)

TABLE OF CONTENTS

Ispra, 22.10.2012

Summary of the ESAC Opinion

The ESAC WG considered the scientific work presented of good quality, despite (1) the fact that the WG had some concerns about the statistical calculations underlying the determination of an adequate sample size to analyse reproducibility as a primary study goal; (2) the WG's concern that possible limitations of the assay were not described in sufficient detail in the report.

Overall, the conclusions made by the WG correspond well with the conclusions formulated in the report by the VMG.

- The WLR of the test method with respect to concordance of classification (S/NS) met the target of 85% (as defined by the VMG).
- The data were considered strong enough to support transferability of the test to properly equipped, trained and staffed laboratories with the appropriate analytical capabilities.
- In spite of a BLR (75%) below the target of 80% (as defined by the VMG), the BLR of the test method with respect to concordance of classification was considered sufficient. On the other hand, the BLR assessment alone argued against the possibility to use the DPRA for assigning chemicals to one of the 4 reactivity classes (62.5%).

- As indicated by the VMG, the number of chemicals (N=24) was considered insufficient to draw firm conclusions on the predicative capacity of the test method. The preliminary data were, however, considered promising.

The predictive capacity of the test is not defined yet, but the preliminary data profile the test as a useful tool for early decision making during product development (screening) and a component in a weight-of-evidence approach or integrated testing strategy for safety/hazard assessment.

The WG recommends that the limitations of the DPRA that may lead to uncertainties concerning negative results be further investigated, either by additional prospective testing or through analysis of existing information.

---

1. Mandate of the ESAC

The opinion of ESAC should support ECVAM with respect to the development of recommendations regarding the reliability (transferability, within and between laboratory reproducibility) of the DPRA and the potential regulatory use of the test method.

1. Study design – transferability, reliability and relevance

- The ESAC was requested to review whether the prevalidation study was conducted appropriately in view of the objective of the study:
  - Reproducibility of the DPRA method within one laboratory (WLR);
  - Transferability to other laboratories;
  - Reproducibility in other laboratories (BLR);
  - Predictive capacity of the test method.

- With respect to the design and conduct of the study, the following issues were to be addressed:
  - Clarity of the test definition (module 1)
  - Clarity of the definition of the study objective
  - Appropriateness of the study design in view of study objective
  - Appropriateness of the study execution:
  - Appropriateness of the statistical analysis used for analysing WLR, transferability, BLR and (preliminary) predictive capacity.

2. Conclusions of the study

The ESAC was requested to assess the justification and plausibility of

- Reproducibility (WLR and BLR) and transferability;
- Preliminary predictive capacity;
- Possible gaps between study design and study conclusions which remain to be addressed in view of the suggested conclusions/use;
- Applicability and possible limitations of the test method, in particular in view of its potential use within an ITS for sensitisation.

3. The ESAC is requested (a) to evaluate, on the basis of the data submitted in the validation study, the possible use of the test method (also within a strategy) to identify skin sensitisers, (b) to make additional recommendations (as required) on the proper scientific use of the test method within such a strategy taking specific aspects of this method into account (e.g. applicability, limitations etc.) and (c) to identify possible further information required (i.e. are there gaps) to be able to conclude on the plausibility of the suggested use (including within an ITS).

2. Detailed opinion of the ESAC

1. Study design – transferability, reliability and relevance.

- The WLR was assessed at the level of concordance in prediction (S/NS). An average reproducibility of 87% met the 85% reproducibility target set by the VMG. The definition of the reproducibility target (85%) was based upon i) the background and specific objectives of the validation study; ii) the standards of performance that can realistically be expected from an *in vitro* test and standards of performance which have been considered acceptable in previous validation studies; iii) the proposed used of the *in vitro* tests (i.e. as a partial replacement method to become part of a toolbox of tests to be used in combination); and iv) the power of the design of the validation study.
- Transferability activities were divided into Training, Transferability and Quality Control. The WLR was formulated for each partner to include 1) concordance in prediction, 2) depletion values for cysteine and lysine, as well as 3) control values. The BLR was assessed in terms of 1) concordance in prediction and 2) depletion values for cysteine and lysine. IVMU did not meet the first preset criterion stipulating that all the runs had to meet the acceptance criteria defined in the SOP for assessment of success primarily because of the Reference Control C being outside the acceptance criteria. The causes of the difficulties to meet the criteria are still not understood. The VMG recommended in the VSR that these acceptance criteria should be relaxed in the future.
- Eighteen of the 24 chemicals were consistently classified (S/NS) by the three laboratories resulting in a BLR reproducibility of 75%, which is below the target (80%). The reproducibility assessment included 3 chemicals (beryllium sulphate, nickel chloride and dihydroeugenol) that were considered by the VMG as outside the applicability domain of the test. For 15 out of the 24 chemicals the laboratories assigned the same reactivity class resulting in a BLR of 62.5%. Data variability was observed to result for chemicals with low or no reactivity.
- The secondary goals included a <u>preliminary</u> evaluation of the ability of the test to discriminate skin sensitizers from non-sensitizers, and a preliminary consideration of the ability to contribute to sub-categorization of skin sensitising chemicals (GSH sub-category 1A and 1B). The VSR report did not present a summary of the predictive capacity based on all 24 chemicals tested, since the VMG judged three of them (beryllium sulphate, nickel chloride, dihydroeugenol) to fall outside the applicability domain. The two pre-haptens (4-phenylendeiamine and R(+)-Limonene) were included in the analysis as the VMG felt that there was insufficient evidence to exclude them from the evaluation of the predictive performance. The predictive capacity for all 24 substances was 77.8% (sensitive: 70.8%; specificity: 91,7%) while 82.4% (sensitive: 73.5%; specificity: 91,7%) for the 19 substances.
- The project was described and designed in clearly recognizable and well described phases including Test Definition (Module 1), Transferability (Module 3), Within

Laboratory Reproducibility (WLR) (Module 2), Between Laboratory Reproducibility (BLR) (Module 4) and Predictive Capacity (Module 5).

- Overall, the chosen statistical approach was considered appropriate, but a number of areas were unclear making it difficult to fully assess its validity.
  Within laboratory reproducibility: Similar calculations were performed for the WLR. However, in absence of experimental historical data on WLR, the VMG assumed, based on previous experiences in validation studies that the WLR concordance would be higher than that for BLR and was set at 95%. The statistical power chosen was 80%.
  Between laboratory reproducibility: The 'expected proportion' of concordant classifications (between laboratories) was calculated to be 90% on the basis of available data on BLR as submitted to ECVAM. However, it was not clear for the WG why a power of 75% rather than the more conventional 80% or 90% power had been applied. This power allows for detecting 25% changes in each direction and, as a consequence, leads to a broader confidence interval and thus more uncertainty about the 'real' value.

2. Conclusions of the study

- Overall, the test design and the quality of the selected chemicals (N=24) were considered appropriate for the purpose of addressing the first objective of the study: Assessing the WLR and BLR of the DPRA.
- Overall, the conclusions made by the WG correspond well with the conclusions drawn by the VMG as described in the VSR, indicating that these conclusions are supported by the results shown in the report.
  - The WLR of the test method with respect to concordance of classification (S/NS) met the target of 85% and was considered sufficient for the purpose of this study.
  - The data were considered strong enough to support transferability of the test to properly equipped, trained and staffed laboratories with the appropriate analytical capabilities.
  - In spite of a BLR (75%) below the target of 80%, the BLR of the test method with respect to concordance of classification was considered sufficient. On the other hand, the BLR assessment argued against the possibility to use the DPRA for potency classification (62.5% concordance). It was appreciated that reactivity classes do not necessarily correspond to potency classes, but the general feeling of the WG was that, as described in the section regarding the WLR, also the BLR results argue against the possibility to use the DPRA for potency classification as a stand-alone method. The number of chemicals (N=24) did not provide support for a firm conclusion about the predicitive capacity of the test method. The preliminary data were, however, considered promising.
  - The number of chemicals did not allow drawing a conclusion about the applicability domain of the test. Empirically the applicability domain seems to exclude pre-/pro-haptens and metal salts.

- Chemicals that preferably react with amino acids other then cysteine and lysine may fall outside the applicability domain. In addition, some pre-/pro-haptens were reported as correctly identified. Finally, the data seem to indicate that the test method has problems identifying weak sensitizers. The uncertainty about the applicability domain may result in an unacceptable level of false negative results.

3. Possible use of the test method (also within a strategy) to identify skin sensitisers, and additional recommendations (as required) on the proper scientific use of the test method within such a strategy.

As outlined in the VSR and the ECVAM request for ESAC advice, the DPRA cannot be used as a standalone in a regulatory context but should be considered for use in an Integrated Testing Strategy (ITS). On the basis of the present report, especially negative outcomes have to be considered with care.

- As pre-haptens are not consistently correctly predicted by the DPRA, there remains uncertainty about whether to consider pre-haptens as part of the applicability domain of the method or not.
- Unless there are sufficiently accurate assays available identifying chemicals as pre-/pro-haptens in view of excluding them from routine testing using the DPRA, such compounds will be tested in the DPRA and may cause false negative results.
- The selection of cysteine and lysine-containing peptides selects for the majority, but not all, reactive chemicals.

Regarding reactivity class, the data obtained did not support the possibility to use DPRA as a standalone assay for potency classification. This is in agreement with the statement of the VMG that the assay should be further evaluated for its capacity to "contribute" to a potency classification (VSR page 8).

Information generated by the DPRA can be used to support regulatory decision making when used in the context of a weight-of-evidence approach or Integrated Testing Strategy (ITS). It is important to use the test in a context that allows confident conclusions about the protein-reactivity of the chemical, especially when the chemical in question is negative in the DPRA. As such the method may be helpful to address testing requirements of the REACH legislation and the 7th Amendment of the Cosmetic Directive.

Its inclusion into future integrated testing strategies can be considered for the purpose of an eventual full replacement of current *in vivo* hazard identification assays.

Recommendations:

The DPRA addresses a key mechanism (haptenation) in the development of skin sensitization/allergic contact dermatitis. Overall the provided data support transferability and reproducibility of the test to qualified laboratories. The predictive capacity of the test is not defined yet, but the preliminary data profile the test as a useful tool for early decision making during product development (screening) and a component in a weight-of-evidence approach or integrated testing strategy for safety/hazard assessment.

The WG recommends that the limitations of the DPRA that may lead to uncertainties concerning negative results be further investigated, either by additional prospective testing or through analysis of existing information.

3. Informative background to the Mandate and Opinion

Skin sensitisation is the toxicological endpoint associated with substances that have the intrinsic ability to cause Allergic Contact Dermatitis, ACD in humans. ACD represents the most common manifestation of immunotoxicity in humans, i.e. adverse effects of xenobiotics involving the immune system. The identification of the ***skin sensitization potential*** represents an important component of

the safety assessment of any new substance and especially for those intended for topical application (e.g. cosmetics). Current regulatory predictive tests for skin sensitization rely on the use of animals, these include:

a) the traditional guinea pig tests: *Buehler Test* and *Guinea-pig Maximisation Test* (OECD TG 406, Ref.1),

b) the *Local Lymph Node Assay* (LLNA, OECD TG 429, Ref.2) and its recently OECD adopted non-radioactive variants (OECD TG 422A, Ref.3 and OECD TG 422B, Ref.4).

Despite the progress that has been made in the development of alternative methods for skin sensitisation hazard identification, there are currently no validated methods available. In addition none of the tests currently under development/evaluation is able to fully characterise the relative potency of sensitising substances and therefore, none of these assays is considered a stand-alone method, capable of fully replacing current animal procedures, in particular as regards to cosmetics.

The current view therefore is to combine different test methods in order to address different key mechanisms of skin sensitisation: skin bioavailability, haptenation (the protein binding of chemicals which triggers immunological responses), epidermal inflammation, dendritic cell activation and migration, T cell proliferation. Test methods are currently under development which have been specifically designed to address these key mechanistic steps involved in skin sensitisation. Before these test methods can be routinely used, e.g. in integrated testing strategies, their capacity to produce reproducible results needs to be demonstrated as a first step.

The *Direct Peptide Reactivity Assay* DPRA is addressing one of the key upstream events in the cascade of mechanisms leading to the induction of skin sensitisation. It measures the ability of chemicals to react with proteins (haptenation). There is good evidence that haptenation is a determinant step in the induction of skin sensitisation. Chemical allergens are usually low molecular weight chemicals which are not immunogenic per se. However, chemical allergens (or their metabolites, oxidation products) have electrophilic properties that allow them to bind covalently with the nucleophilic side chains of amino acids of skin proteins to form an immunogenic conjugate. Already in 1936 this correlation between the reactivity of chemicals with proteins and their skin sensitisation potential was described (Landsteiner and Jacobs, Ref.5) and has in the meantime been extensively described in the literature. This knowledge is being exploited for the development of several *in chemico* reactivity assays with relevance for the testing of sensitisation potential, amongst these the DPRA assay.

The DPRA is designed to screen the ***sensitisation potential*** of chemicals by measuring peptide depletion with UV-HPLC, following incubation of the test chemicals with synthetic heptapeptides containing either cysteine (peptide/chemical ratio in the reaction mixture 1:10) or lysine residues (peptide/chemical ratio in the reaction mixture 1:50) (Gerberick 2004, Ref.6). The average of peptide depletion values for cysteine and lysine are used to classify chemicals into four reactivity categories: minimal, low, moderate and high reactivity (Gerberick 2007, Ref.7). Based on the known correlation between haptenation/chemical reactivity and sensitisation potential, it is assumed that these reactivity classes as predicted by the DPRA may contribute to the characterisation of sensitiser potency.

The possible predictive capacity of the DPRA is supported by the data of the original DPRA submission. On the basis of 133 chemicals, the DPRA classified chemicals as sensitisers or non-sensitisers (in relation to LLNA data) with an accuracy of 86% (87% sensitivity, 83% specificity).

4. References

1. OECD, Organisation for Economic Cooperation and Development (1992) Skin Sensitisation Guidelines for Testing of Chemicals No. 406, Paris
2. OECD, Organisation for Economic Cooperation and Development (2002) The Local Lymph Node Assay. Guidelines for Testing of Chemicals No. 429, Paris
3. OECD, Organisation for Economic Cooperation and Development (2010a) Skin Sensitization: Local Lymph Node Assay: DA, Guidelines for Testing of Chemicals No. 442A, Paris
4. OECD, Organisation for Economic Cooperation and Development (2010b) Skin Sensitization: Local Lymph Node Assay: BrdU-ELISA, Guidelines for Testing of Chemicals No. 442B, Paris
5. Landsteiner K & Jacobs J (1936) Studies on the sensitisation of animals with simple chemical compounds. Journal of Experimental Medicine 64, 625-639
6. Gerberick GF, Vassallo JD, Bailey RE, Chaney JG, Morrall SW, Lepoittevin JP, (2004) Development of a peptide reactivity assay for screening contact allergens. Toxicol Sci. 81; 332-43.
7. Gerberick GF, Vassallo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin JP, (2007). Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. Toxicol Sci. 97, 417-27.