SCIENTIFIC OPINION

EFSA Journal

ADOPTED:
doi:10.2903/j.efsa.20YY.NNNN

DRAFT for public consultation

Reflection on interpretation of some aspects related to genotoxicity assessment

EFSA Scientific Committee:

Tony Hardy, Diane Benford, Thorhallur Halldorsson, Michael Jeger, Helle Katrine Knutsen, Simon More, Hanspeter Naegeli, Hubert Noteborn, Colin Ockleford, Antonia Ricci, Guido Rychen, Vittorio Silano, Roland Solecki, Dominique Turck, Maged Younes, Gabriele Aquilina, Riccardo Crebelli, Rainer Guertler, Karen Ildico Hirsch-Ernst, Elsa Nielsen, Maria Carfì, Nikolaos Georgiadis, Daniela Maurici, Juan Parra Morte, Josef Schlatter.

Abstract

The European Commission requested the European Food Safety Authority to provide advice on the following: (1) the suitability of the Unscheduled DNA synthesis in vivo assay to follow-up positive results in in vitro gene mutation tests; (2) the adequacy to demonstrate target tissue exposure in in vivo studies, particularly in the Mammalian Erythrocyte Micronucleus test; (3) the use of data in a weight of evidence approach to conclude on the genotoxic potential of substances and the consequent setting of health-based guidance values. The Scientific Committee (SC) concluded that the first question should be addressed in both a retrospective and a prospective way: for future assessments, it is recommended no longer performing the UDS test. For re-assessments, in cases with negative results, other more reliable tests such as Transgenic Rodent assay or in vivo comet assay would be needed to complete the assessment. Regarding the second question, the SC concluded that it should be addressed in both direct and indirect lines of evidence of bone marrow exposure: toxicity to the bone marrow in itself provides sufficient evidence to allow concluding on the validity of a negative outcome of a study. All other direct or indirect lines of evidence of target tissue exposure should be assessed within a weight-of-evidence approach. Regarding the third question, the Scientific Committee concluded that any available data that may assist in reducing the uncertainty in the assessment of the genotoxic potential of a substance should be taken into consideration. If the overall evaluation leaves no concerns for genotoxicity, health-based guidance values may be established. However, if concerns for genotoxicity remain, establishing health-based guidance values is not considered appropriate and additional information would be needed to complete the assessment.

Keywords: Genotoxicity, UDS assay, Micronucleus test, Weight of evidence, Bone marrow exposure

Requestor: European Commission

Question number: EFSA-Q-2017-00074

Correspondence: sc.secretariat@efsa.europa.eu

Acknowledgements: The Scientific Committee wishes to thank the following for the support provided to this scientific output: Aljcia Mortensen (ANS panel), Maria Vittoria Vettori (EFSA), Sandro Grilli (PPR panel) and Christiane Vleminckx (CONTAM panel). The Scientific Committee also wishes to acknowledge the following Hearing Experts for their input: David Kirkland and Jan van Benthem.

Suggested citation: EFSA Scientific Committee (Hardy T**,** Benford D, Halldorsson T, Jeger M, Knutsen H, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, , Silano V, Solecki R, Turck D, Younes M), Aquilina G, Crebelli R, Guertler R, Hirsch-Ernst KI, Nielsen E, Carfì M, Georgiadis N, Maurici D, Parra Morte J, Schlatter J, 2017. "Reflection on interpretation of some aspects related to genotoxicity assessment", Scientific opinion. EFSA Journal 2017;volume(issue):NNNN, 27 pp. doi:10.2903/j.efsa.20YY.NNNN

ISSN: 1831–4732

The EFSA Journal is a publication of the European Food Safety Authority, an agency of the European Union.

63 **Summary**

64 To be drafted at a later stage

Table of contents

1. Introduction

Information on genotoxicity is a key component in risk assessment of chemicals in general, including those used in food and feed, consumer products, human and veterinary medicines, and industry. Genotoxicity testing of substances used or proposed for use in food and feed has been routine for many years. Genotoxicity information is also essential for risk assessment of natural and environmental contaminants in food and feed. Many regulatory agencies and advisory bodies have made recommendations on strategies for genotoxicity testing (EFSA, 2011). While the strategies for different chemical sectors may differ in points of detail, to evaluate genotoxic potential, the majority recommend use of a basic test battery comprising two or more *in vitro* tests, or *in vitro* tests plus an *in vivo* test. This is followed up when necessary, in cases where the results of basic testing indicate that a substance is genotoxic *in vitro*, by further studies to assess whether the genotoxic potential is expressed *in vivo*. Follow-up usually comprises one or more *in vivo* tests.

Optimisation of testing batteries to minimise false positives may reduce the likelihood of detecting inherent genotoxic activity. Thus, in recommending strategies for genotoxicity testing for risk assessment purposes, a balance needs to be struck that ensures with reasonable certainty that genotoxic substances likely to be active *in vivo* are detected.

1.1. Background and Terms of Reference as provided by the European Commission

Commission Regulation (EU) No 283/2013[1] sets out the data requirements for active substances for which dossiers were submitted on or after 1 January 2014 (i.e. applicable to all substances submitted under Regulation (EU) No 844/2012[2]) in accordance with Regulation (EC) No 1107/2009,[3] and details the requirements in relation to genotoxicity testing in Section 5.4.

The requirements stipulate that the aim of genotoxicity testing is to

- predict genotoxic potential;
- identify genotoxic carcinogens at an early stage; and
- elucidate the mechanism of action of some carcinogens.

The requirements for *in vitro* studies are well defined, and certain studies must always be provided. Additionally, one *in vivo* study must always be provided even if all of the *in vitro* studies are negative, and furthermore, based on the results of the *in vitro* studies, further *in vivo* studies may also be required.

The data requirements stipulate that in cases where an equivocal or positive test result is obtained in any *in vitro* test, the nature of additional testing needed must be considered on a case-by-case basis, taking into account all relevant information using the same endpoint as considered in the *in vitro* test.

The choice of *in vivo* studies is not prescriptively defined in the data requirements for all situations; rather, examples of possible suitable studies are provided. This is particularly the case for follow-up of positive *in vitro* gene mutation tests whereby the Transgenic Rodent Somatic and Germ Cell Gene Mutation assays (TGR) are listed as a possible follow-up.

[1] Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, p. 1.
[2] Commission Implementing Regulation (EU) No 844/2012 of 18 September 2012 setting out the provisions necessary for the implementation of the renewal procedure for active substances, as provided for in Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 252, 19.9.2012, p. 26
[3] Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC**.** OJ L 309, 24.11.2009, p. 1.

Furthermore, the data requirements state that 'a tiered approach shall be adopted, with selection of higher tier tests being dependent upon interpretation of results at each stage.' Accepted tests and methods are listed in the Commission Communication,[4] which has been agreed by EFSA.

This therefore allows for some flexibility in terms of studies provided, depending on the characteristics of the substance. It should be further recalled that one of the general aims of EU law, also stated in Regulation 1107/2009, is to minimise the use of animal testing as far as possible. Testing on vertebrate animals should be replaced, restricted or refined, and only undertaken as a last resort.

The Scientific Opinion on genotoxicity testing strategies applicable to food and feed safety assessment[5] of the EFSA Scientific Committee (SC) provides recommendations for testing strategies for substances used in food and feed. The SC recommends a documented weight-of-evidence approach to the evaluation and interpretation of genotoxicity data, taking into account not only the quality and reliability of the data on genotoxicity itself, but also other relevant data that may be available. The Opinion also states that when considering pre-existing or non-standard data using a weight-of-evidence approach, a case-by-case approach using expert judgement is required.

During the evaluation of a number of active substances being considered under Regulation (EU) No 844/2012 (substances in the 'AIR3' renewal programme), a conclusion on the genotoxic potential of parent substances and/or metabolites (occurring in plants, livestock or groundwater) was not considered possible by EFSA based on the information submitted by the applicants in the dossiers considered in the peer review process. However, during the peer review process and thereafter, there have been strong divergences in opinion between some Member States, EFSA and applicants on this issue.

Taking into account the background above, the Commission requests consideration of the following three aspects to provide clarity and predictability for applicants and risk assessors carrying out assessments of genotoxicity and also for risk managers:

1. The adequacy of the Unscheduled DNA Synthesis (UDS) Assay to follow-up positive results in the *in vitro* gene mutation tests;

In light of this, the following questions are raised:

   a) Whilst it is clear that the Opinion of the SC does not recommend the UDS Assay as an appropriate follow-up study, in what situations and for what types of substances is the UDS Assay (carried out prior to the Opinion of the SC and submitted in a dossier for approval/renewal of approval) suitable to follow-up a positive *in vitro* gene mutation result?

   b) Does the current Opinion of the SC on testing strategies[6] satisfactorily highlight the limitations of using the UDS Assay in the assessment of genotoxicity? Does it fully consider how the UDS Assay may be used (when submitted as part of a data package that was developed prior to the Opinion)?

   c) Can any further indicators or advice be provided to better enable risk assessors to consider if a submitted UDS Assay is adequate to be used in concluding on the genotoxic potential of a substance?

2. The adequacy to demonstrate target tissue exposure in *in vivo* studies, particularly in the micronucleus (MN) test

   a) What type of measurements or information from other studies can be used to provide reassurance of bone marrow exposure in MN or other studies and enable the use of such studies to conclude on the genotoxic potential of the substance e.g. clinical signs of toxicity (particularly those considered systemic in nature), presence in the blood/plasma, ADME data?

[4] Commission communication in the framework of the implementation of Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ C 95, 3.4.2013, p. 1.

[5] EFSA Journal 2011;9(9):2379 [69 pp.].

[6] EFSA Journal 2011;9(9):2379 [69 pp.].

b) In the absence of information to definitively confirm bone marrow exposure, what additional data or information can be used to allow (if possible) for the use of such studies in the assessment of genotoxicity and provide reassurance that a negative result is not a false negative?

3. The use of data in a weight-of-evidence approach to conclude on the genotoxic potential of substances and the consequent setting of health-based reference values for use in human health risk assessment.

The EFSA Scientific Committee recommends in its Opinion:

'a documented weight-of-evidence approach to the evaluation and interpretation of genotoxicity data, in particular when considering heterogeneous and non-standard data sets' [particularly relevant for considering contaminants]. 'Such an approach should not only consider the quality and reliability of the data on genotoxicity itself, but also take into account other relevant data that may be available, such as physico-chemical characteristics, structure-activity relationships (including structural alerts for genotoxicity and read-across from structurally related substances), ADME, and the outcomes of any repeat-dose toxicity and carcinogenicity studies.'

a) It should be further considered and clarified as to what this means in terms of performing robust assessments of genotoxicity for pesticide active substances and metabolites in the absence of the standard or preferred battery of tests (and/or when the recommendations in the Opinion on genotoxicity testing are not followed), or where there may be positive results *in vitro* but where negative results in *in vivo* studies may have some limitations. In particular, this should be specifically considered for cases where there is a complete and robust data set available, including carcinogenicity and reproductive toxicity studies. This point should be considered in light of the need to minimise animal testing.

Taking this into account, the following point should be considered:

b) What considerations should be taken into account as part of a weight-of-evidence approach when determining if toxicological reference values for use in health-based risk assessments can be set when there are uncertainties or inadequacies with the assessment of genotoxicity (non-standard/non-preferred data set or non-guideline studies) and in which circumstances, if any, is it considered not possible to set health-based reference values, either fixed or provisional?

The Commission requests consideration of these three questions to provide clarity and predictability for applicants and risk assessors carrying out assessments of genotoxicity and also for risk managers when making decisions on approval of substances. The genotoxicity data requirements in the different EFSA frameworks, the Opinion of the SC on genotoxicity testing strategies and the wider scientific and technical knowledge on the assessment of genotoxicity shall be also taken into account.

1.2. Interpretation of the Terms of Reference

When addressing question 2 of the mandate, the SC assumes that the *in vivo* Micronucleus test (MN) has been selected as the appropriate follow-up of a positive *in vitro* outcome and will not address follow-up testing strategies of a positive *in vitro* test outcome. Hence, the interpretation of the question of the mandate is how to verify the exposure of the bone marrow when performing an *in vivo* MN test, either measuring micronuclei in bone marrow cells or in peripheral blood cells, and which elements (lines of evidence) should be taken into consideration to provide reassurance that a negative result in the *in vivo* MN test is not a false negative. The SC decided to categorise the elements in direct and indirect lines of evidence of exposure of the bone marrow.

When addressing question 3 of the mandate, *"setting of health-based reference values for use in human health risk assessment"* is understood as establishing Health Based Guidance Values (HBGV).

The SC notes that the quantitative assessment of dose-response relationships in genotoxicity studies is beyond the scope of the present Opinion.

1.3. Additional information

According to the genotoxicity testing strategies proposed by EFSA (EFSA, 2011), when all *in vitro* endpoints are clearly negative, it can be concluded that the substance is not genotoxic. If *in vivo* testing is considered necessary to follow-up a positive *in vitro* outcome, the choice of *in vivo* tests should relate to the genotoxicity endpoint(s) identified by the aforementioned positive *in vitro* tests as well as to the appropriate target tissues.

The SC notes that for different areas of risk assessments conducted by EFSA, the basis for evaluation of genotoxicity may differ with respect to the extent of the available dataset and its robustness. In the context of this Opinion, three different situations are considered:

I. Regulated "new" substances that are being subject to risk assessment in the context of a (first) authorisation procedure and for which a standard data package is required per regulation, e.g. new pesticide active substances and new food additives. For such substances, an extensive dataset, including studies that have been conducted according to the most recent standard of testing guidelines, can be expected.

II. Regulated "old" substances that have undergone assessment and approval in the past and are currently under re-evaluation with respect to extension of authorisation or renewal of approval, e.g. pesticide active substances under the renewal procedure and renewal of food/feed additives. Although a defined extensive data package is also required for such substances, the submitted dossiers may contain studies that were conducted under older protocols differing from updated protocols, or under provisions not reflecting the current state of scientific development.

III. Substances for which the available dataset is not based on data requirements laid down in specific legislation and thus may often contain information gaps and/or information from non-standard (non-guideline) studies (e.g. in the case of contaminants or metabolites in plants).

Taking into account the emphasis within the scope of the Terms of Reference, the current Opinion will focus mainly on situations II and III.

2. Data and Methodologies

2.1. Data

The evidence used for this mandate stems primarily from expert knowledge gathered by a working group of the EFSA SC dedicated to the work of this Opinion. Recent published literature and ongoing activities at different national and international levels were considered and discussed. The analyses presented in the appendix of this Opinion rely on the data from the EURL ECVAM database (https://eurl-ecvam.jrc.ec.europa.eu/databases/genotoxicity-carcinogenicity-db) as well as data retrieved from the literature. Established genotoxicity testing strategies are described in existing Scientific Opinion on genotoxicity testing strategies applicable to food and feed safety assessment (EFSA 2011). Following the EFSA policy on transparency and engagement in risk assessment (TERA), EFSA decided to publish the current draft Opinion for public consultation to get additional data and views from interested stakeholders.

2.2. Methodologies

The methodology used for this Opinion was to aggregate the information from the diverse EFSA areas where the assessment of genotoxicity is performed, considering different legislation applicable in the food and feed area, relevant EFSA scientific opinions and guidance documents and information from

the open literature. Ongoing activities in other EU or international organisations were considered. The working group of the SC was composed by members of the panels dealing with genotoxicity assessment. Ad hoc hearing experts were invited to provide additional data specifically in relation to questions 1 and 2 of the mandate. Consultation with the European Chemical Agency (ECHA) took place prior to the endorsement by the SC of the draft Opinion for public consultation. The public consultation was launched on the EFSA website on 24 July 2017 and ended on 6 September 2017. The report of the public consultation summarising the comments received and how they have been addressed will be published together with the finalised Opinion by the end of 2017.

3. Assessment

3.1. The adequacy of the Unscheduled DNA Synthesis (UDS) Assay to follow-up positive results in the *in vitro* gene mutation tests

The rat liver Unscheduled DNA Synthesis assay, referred hereafter as UDS, detects the induction of DNA repair synthesis in the liver of treated adult rats. To this aim, the incorporation *ex vivo* of radioactive labeled nucleotide, e.g. tritium-labeled thymidine (^{3}H-TdR), is measured in isolated hepatocytes by autoradiography. Because the rate of liver cell proliferation is very low in adult animals and the few replicating cells can be easily identified, the incorporation of radioactivity in nuclei of non-replicating cells indicates unscheduled DNA synthesis related to the repair of DNA by excision, removal and replacement of a damaged stretch of a DNA strand. The UDS is thus an indicator test detecting DNA damage, not a mutagenicity assay measuring stable genetic alterations. The test is designed to respond only to substances that induce a type of DNA damage that is repaired by excision repair; DNA damage processed by other mechanisms, as well as unrepaired genetic damage are not detected with this assay.

In principle, a UDS could be applied to any tissue not undergoing extensive cell proliferation, but in practice most experience is related to its application in the rat liver. A protocol for the detection of unscheduled DNA synthesis in rat liver cells was first proposed by Mirsalis et al. in the 1980s (Mirsalis et al., 1980; 1982). Based on the original study protocol (Butterworth et al., 1987), an OECD Test Guideline (TG 486: Unscheduled DNA synthesis (UDS) test with mammalian liver cells *in vivo*) was developed and first adopted in 1997. Since then, TG 486 has not been revised and is retained unchanged over the most recent round of revision of OECD Test Guidelines for genetic toxicology (2014–2015) (OECD, 2016).

The UDS was widely applied in studies performed for regulatory purpose until the adoption of the OECD test guidelines (TG) for the transgenic rodent somatic and germ cell gene mutation assays (OECD TG 488, 2011) and the *in vivo* mammalian alkaline comet assay (OECD TG 489, 2014). Both TGs have been updated in 2013 and 2016, respectively. As a matter of fact, until then, the UDS was the only test method applicable to somatic tissues other than the erythropoietic system, addressed by cytogenetic tests (OECD TGs 474 and 475). Thus, even though in principle the UDS should be considered appropriate only for substances targeting the liver and/or requiring liver metabolic activation, and inducing adducts removed by excision repair (as for many mutagens), it was formerly frequently recommended as the follow-up test for substances positive in genotoxicity tests *in vitro*. According to previous operational criteria underlying genotoxicity testing strategies, negative results in at least two *in vivo* assays in different target tissues were considered necessary to evaluate the biological relevance of *in vitro* positives (Carere et al., 1995, online) (see also Appendix A). Thus, to fulfill such a requirement, the UDS was often performed as default complement of cytogenetic assays in erythropoietic cells in the follow-up of *in vitro* positive results.

At present the UDS is still mentioned in several guidance documents or guidelines among *in vivo* tests applicable in the follow-up of *in vitro* positive results in spite of the limitations of the test (e.g. EFSA, 2011; ICH, 2011; EC, 2013; ECHA, 2016).

According to the EFSA Opinion on genotoxicity testing strategies (EFSA, 2011) '*the in vivo UDS test allows the investigation of genotoxic effects of substances in the liver ... However, UDS has a limited use for cells other than liver, and its sensitivity has been questioned (Kirkland and Speit, 2008).*' Thus, with reference to the question posed in the Terms of Reference, it can be said that the EFSA SC

Opinion on testing strategies (2011) adequately highlights the limitations of using the UDS in the assessment of genotoxicity.

According to the OECD overview of the set of OECD genetic toxicology test guidelines performed in 2014–2015 (OECD 2016), '*the UDS in vivo is a primary DNA damage test that identifies chemicals inducing DNA damage and subsequent repair (measured as unscheduled DNA synthesis versus normal S-phase scheduled synthesis) in liver cells … However, this test does not detect mutagenic consequences of the unrepaired genetic damage. Accordingly, the UDS test may be an appropriate test to detect DNA damage after exposure to chemicals that specifically target the liver and that were positive in the Ames test. The test responds positively only to chemicals that induce the type of DNA damage that is repaired by nucleotide excision (mainly bulky adducts) … The UDS test should not be considered as a surrogate for a gene mutation test and it may be less reliable than other primary DNA damage tests (Kirkland and Speit, 2008).*'

Finally, according to ECHA Guidance on information requirements and chemical safety assessment (ECHA, 2016), the UDS should be used '*only when liver is the target organ … A positive result in the UDS assay can indicate exposure of the liver DNA and induction of DNA damage by the substance under investigation, but it is not sufficient information to conclude on the induction of gene mutations by the substance. A negative result in a UDS alone is not a proof that a substance does not induce gene mutations.'*

In view of the inherent limitations of the UDS, it is widely accepted that negative results in this test may have limited relevance in the follow-up to *in vitro* positive results. Beyond theoretical considerations that restrict the field of applicability of the assay to certain classes of genotoxicants (i.e. those eliciting DNA excision repair), the comparison with the performance of other *in vivo* assays also highlights the overall low sensitivity of the assay, which casts doubts on its suitability for regulatory use.

The role of the UDS in regulatory testing strategies and its usefulness in the follow-up to positive *in vitro* results was questioned because of the low sensitivity it displayed to a set of rodent carcinogens negative or equivocal in the bone marrow MN test (Kirkland and Speit, 2008). In this analysis, less than 20 % (7 out of 41) of carcinogens tested positive in the UDS. Notably, those compounds testing negative in the UDS included 19 liver carcinogens, 6 of which were positive in the *in vivo* comet assay. Even narrowing the analysis to carcinogens testing *in vitro* as inducers of gene mutations – to which the UDS is expected to be more responsive – a high incidence of negative responses was observed with this assay (Kirkland and Speit, 2008).

A further analysis was performed to support this Opinion using data from the EURL ECVAM database (EURL_ECVAM, 2017). This recently constructed database currently compiles available genotoxicity and carcinogenicity data for 726 Ames-positive chemicals covering different sectors (industrial chemicals, cosmetics ingredients, plant protection and pharmaceutical products) and originating from different sources (regulatory agencies, industry and publicly available literature).[7] A rigorous methodology and defined criteria were applied for the selection and analysis of the data (Kirkland et al., 2014). Only chemicals with a known chemical identity and validated *in vitro* and *in vivo* results for the genotoxicity endpoints and/or carcinogenicity were selected. Data for the following tests were collected: *in vitro* tests [Ames, mouse lymphoma tk+/- (MLA) or gene mutation at the *hprt* locus, micronucleus (MN), chromosome aberration (CA)]; *in vivo* tests [MN, CA, UDS, TGR, DNA breakage (comet and alkaline elution assay)]; rodent carcinogenicity. 'Overall Calls' (Appendix A) were defined for each genotoxicity assay *in vitro* and *in vivo,* and carcinogenicity by following defined criteria for the reliability of each study and quality of data for those chemicals appearing in more than one source with different calls. Four categories were considered: positive, negative, equivocal and inconclusive. Where information was missing, even for those chemicals with one single data entry, scientific literature was consulted. The data were analysed by calculating the sensitivity and specificity of the respective tests.

The database comprises 88 Ames-positive chemicals that were tested both in the UDS and for carcinogenicity. Ten of these chemicals were not carcinogenic. Of the 78 Ames-positive that are

[7] It needs to be noted, however, that the database is biased towards positive carcinogenic and genotoxic chemicals.

carcinogens and tested for UDS, 31 (39.7%) were negative in UDS. Thus the sensitivity of the UDS to predict (genotoxic) carcinogenicity was 60.3% (Appendix A, Table 1). By contrast, when performing a similar analysis, the sensitivity of TGR and the comet assay to predict carcinogenicity was 88.2% (67 positives out of 76 carcinogens) (Appendix A, Table 2) and 91.2% (52 positives out of 57 carcinogens), respectively (Appendix A, Table 3).

Among the 31 substances negative in the UDS, 8 were tested in the TGR and 10 in the *in vivo* comet assay. Four out of 8 (50%)[8] were positive in the TGR, and 9/10 (90%) in the comet assay.

A further analysis was performed to compare the outcomes of the UDS with those of other *in vivo* genotoxicity assays. The database comprises 114 Ames-positive substances tested both in the UDS and in at least one other *in vivo* genotoxicity test (MN, CA, TGR, comet). 66 substances that were Ames-positive and *in vivo* positive (in either MN, CA, TGR, or comet assay) were also tested in the UDS: out of these, 39.4% (26/66) were negative in the UDS; conversely, among 48 Ames-positive substances that were negative *in vivo* (in either MN, CA, TGR, or comet assay), only 4 (8.3%) were positive in the UDS. Overall, this analysis shows an average concordance of results between UDS and other *in vivo* assays of 73.7% (84/114), with a remarkable fraction of *in vivo* positives (39.4%) not detected by the UDS (Appendix A, Table 4).

The overall low sensitivity of the UDS in detecting rodent carcinogens and/or *in vivo* genotoxicants highlighted by the analysis of the EURL ECVAM database confirms previous conclusions on the lower predictive value of the UDS compared with TGR and *in vivo* comet assays (Kirkland and Speit, 2008), supporting a more prominent role for the latter assays in regulatory testing strategies.

With reference to the specific question on the suitability of the UDS addressed in the Terms of Reference, the SC concluded that, taking into consideration the current EU legislation and guidance documents, this question should be addressed in both a retrospective and a prospective way.

- For future assessments (new submissions), the EFSA SC is not aware of situations or substances that can be identified where the UDS could be considered preferable to the TGR or comet assay (refer to point (a) of the Terms of Reference). Therefore the use of the UDS is no longer recommended
- For re-assessment, in cases where UDS data as a follow-up to a positive *in vitro* mutation test already exists, there might be positive or negative results: test results may be considered as adequate to assess genotoxic potential only in cases of positive results. If the outcome of the UDS is negative, however, other more reliable tests such as TGR or *in vivo* comet assay would be needed to complete the assessment (see also Section 3).

3.2. The adequacy to demonstrate target tissue exposure in *in vivo* studies, particularly in the Mammalian Erythrocyte Micronucleus test

As mentioned in Section 1.2, the SC assumes that the *in vivo* Mammalian Erythrocyte Micronucleus (MN) test (OECD TG 474) has been selected as the appropriate test to follow-up a positive *in vitro* outcome.

Regarding bone marrow (BM) exposure, the OECD Test Guideline (TG) 474 (version adopted in July 1997) states

> '*If there is evidence that the test substance, or a reactive metabolite, will not reach the target tissue, it is not appropriate to use this test.' And 'The likelihood that the test substance or its metabolites reach the general circulation or specifically the target tissue (e.g. systemic toxicity) should be discussed.'*

This TG was revised and the new version was adopted in July 2016. In relation to target tissue exposure, it states: 'A blood sample should be taken at appropriate time(s) in to permit investigation

[8] Dosing periods were shorter than recommended in some negative TGR assays, and this may have biased the results.

of the plasma level of the test substances for the purposes of demonstrating that exposure of the bone marrow occurred, where warranted and where other exposure data do not exist.'

Then, in relation to the evaluation and interpretation of the results:

> '*Evidence of exposure of the bone marrow to a test substance may include a depression of the immature erythrocyte ratio or measurement of the plasma or blood level of the substance. In case of intravenous administration, evidence of exposure is not needed. Alternatively, ADME data, obtained in an independent study using the same route and same species, can be used to demonstrate bone marrow exposure. Negative results indicate that, under the test conditions, the test chemical does not produce micronuclei in the immature erythrocytes of the test species.'*

If a positive result is observed in the *in vivo* Mammalian Erythrocyte MN test, demonstration of target tissue exposure is not needed. However, evidence of bone marrow exposure is needed to conclude that a substance is not genotoxic based on a negative Mammalian Erythrocyte Micronucleus MN test outcome.

Evidence of bone marrow exposure can be direct as well as indirect.

Lines of direct evidence of bone marrow exposure are:

- information from the Mammalian Erythrocyte MN test that the test substance can induce toxic effects in the bone marrow;
- information from toxicity studies that the test substance can induce toxic effects in the bone marrow; and
- information from an ADME study that the test substance (and/or its metabolites) can reach the bone marrow.

Lines of indirect evidence of bone marrow exposure are:

- information from the Mammalian Erythrocyte MN test that the test substance can induce systemic toxicity;
- information from a toxicity study that the test substance can induce systemic toxicity;
- information from ADME studies that the test substance (and/or its metabolites) can reach the systemic circulation; and
- information from a specific plasma analysis that the test substance can reach the systemic circulation.

These different lines of evidence of bone marrow exposure are described shortly below.

3.2.1. Direct evidence of bone marrow exposure

3.2.1.1. Toxicity to the bone marrow observed in the Mammalian Erythrocyte Micronucleus test

In the *in vivo* Mammalian Erythrocyte MN test, toxicity to the bone marrow is indicated if there is a reduction in the proportion of immature erythrocytes among total erythrocytes, i.e. a decrease in the PCE/(NCE+PCE) ratio (PCE: Polychromatic erythrocytes, i.e. immature erythrocytes, also named reticulocytes; NCE: Normochromatic erythrocytes, i.e. mature erythrocytes). A decrease in the ratio in such a study, if evaluated as being test-substance related, is direct evidence of bone marrow exposure**.**

3.2.1.2. Toxicity to the bone marrow observed in toxicity studies

In repeated-dose toxicity studies performed according with an internationally accepted test guideline such as the oral 28-day or 90-day studies (e.g. OECD TG 407, TG 408, TG 409) and the chronic toxicity studies (e.g. OECD TG 451, TG 452, TG 453), the bone marrow is one of the tissues that should be examined histopathologically. A histopathological change in the bone marrow observed in

such a study using the same route and preferably the same species as in the Mammalian Erythrocyte MN test, if evaluated as being test-substance related, is direct evidence of bone marrow exposure.

3.2.1.3. Test substance detected in the bone marrow in an ADME study

Toxicokinetic (ADME) studies (e.g., OECD TG 417) are conducted to obtain information on absorption, distribution, biotransformation (i.e. metabolism) and excretion of a test substance. Generally, the test substance is administered as the ^{14}C-radiolabelled compound (with the radiolabel located in a core portion of the molecule that is metabolically stable) and the radioactivity measured in the blood/plasma, organs and tissues, and excreta. Detection of radioactivity in the bone marrow in such a study, using the same route and preferably the same species, provides direct evidence of bone marrow exposure either to the test substance itself and/or its metabolites.

3.2.2. Indirect evidence of bone marrow exposure

The bone marrow is a well-perfused tissue and, therefore, systemic bioavailability of a test substance can be considered as indirect evidence of bone marrow exposure. Evidence on systemic bioavailability can be obtained directly from an ADME study or a specific plasma analysis that is recommended in the revised Mammalian Erythrocyte MN test (OECD TG 474, July 2016). Evidence on systemic bioavailability can also be obtained indirectly from toxicity studies when test-substance-related systemic toxicity is observed. These different lines of indirect evidence may be used in combination in a weight-of-evidence assessment to justify the assumption that the test substance is systemically available, as detailed in section 3.2.3.

3.2.2.1. Systemic toxicity observed in the bone marrow Micronucleus test

In the bone marrow MN test (OECD TG 474), general clinical observations of the test animals (cage-side) should be made and clinical signs recorded at least once a day. Certain clinical signs, which will usually only occur if the test substance is systemically bioavailable (such as certain signs related to the central nervous system (CNS), e.g. reduction of spontaneous activity, ataxia, sedation, staggering, salivation, abdominal position), if evaluated as being test-substance related, are direct evidence of systemic bioavailability, and consequently, indirect evidence of bone marrow exposure.

3.2.2.2. Systemic toxicity observed in a toxicity study

In all toxicity studies performed according to internationally accepted test guidelines, general clinical observations of the test animals (cage-side) should be made and clinical signs recorded at least once a day. In repeated-dose toxicity studies, such as the 28-day, 90-day and chronic toxicity studies, performed according to an internationally accepted test guideline, detailed clinical observations (outside the cage) should also be made, as well as haematological examinations, clinical biochemistry determinations and histopathological examinations of organs and tissues; urinalysis is a requirement or optional. These parameters could also be part of other kinds of repeated-dose toxicity studies (e.g., reproductive toxicity studies, neurotoxicity studies) as well as of repeated-dose toxicity studies not performed according to a test guideline. Several toxic effects such as certain clinical signs of toxicity, changes in haematological, clinical biochemistry and urinalysis parameters, and histopathological changes in organs and tissues, will usually only occur if the test substance is systemically bioavailable. Such systemic effects observed in a repeated-dose toxicity study, using the same route and preferably the same species as in the Mammalian Erythrocyte MN test, if evaluated as being test-substance related, are direct evidence of systemic bioavailability, and consequently, indirect evidence of bone marrow exposure. Some examples of endpoints which might contribute to the weight-of-evidence assessment are listed in Appendix B.

3.2.2.3. Test substance detected systemically in an ADME study

By analogy to what is described in section 3.2.1.3, detection of radioactivity in the blood/plasma, organs and tissues, or excreta (i.e. urine) in an ADME study using the same route and preferably the same species, provides direct evidence of systemic bioavailability of the test substance itself and/or its metabolites, and consequently, indirect evidence of bone marrow exposure either to the test substance itself and/or to its metabolites.

3.2.2.4. Test substance detected systemically in a specific plasma analysis

According to the revised version of the Mammalian Erythrocyte MN TG (OECD TG 474) adopted in July 2016, 'A blood sample should be taken at appropriate time(s) in to permit investigation of the plasma level of the test substances for the purposes of demonstrating that exposure of the bone marrow occurred, where warranted and where other exposure data do not exist.' Detection of the test substance in the plasma in a specific plasma analysis is direct evidence of systemic bioavailability of the test substance and, therefore, could be considered as indirect evidence of bone marrow exposure.

To follow OECD TG 474, the SC recommends consideration of the following aspects to evaluate whether the specific plasma analysis is suitable to indicate systemic exposure:

- Are the blood samples taken at the appropriate times, i.e. is the documentation for blood sampling times justified in terms of the toxicokinetics of the test substance – e.g. rapid elimination of the test substance and/or its metabolites from plasma and distribution to organs and tissues, rapid metabolism of the test substance – if only the parent compound is measured?
- Are the test substance itself and/or its metabolites measured?
- Can the analytical method only detect the free test substance or can it account for both the free and the bound (protein bound) test substance if a test substance is significantly bound to e.g. plasma proteins? In cases using radiolabelled substances in satellite group treated animals, information on plasma protein binding is key.
- Is the analytical method appropriate, i.e. can it detect and quantify the actual substances to be measured for the particular evaluation (are the limit of detection, the limit of quantification and the calibrated range reported for the analytical method used and are these parameters considered adequate)?
- Is the specific plasma analysis part of the actual Mammalian Erythrocyte MN test or is it performed as an additional analysis – maybe in relation to other studies on the test substance under evaluation?
- Is extrapolation between species necessary, e.g. is the specific plasma analysis performed in a rat study whereas the actual Mammalian Erythrocyte MN test is performed in mice?
- Is extrapolation between routes necessary, e.g. is the specific plasma analysis performed in a study via inhalation route whereas the actual Mammalian Erythrocyte MN test is performed via oral route?
- Are the plasma levels measured within or below the calibrated range?
- Are the plasma levels plausible considering, e.g. consistency among animals of the same group (i.e. how large is the variation) and among different sampling times?

3.2.3. Assessment of the different lines of evidence regarding bone marrow exposure

The SC notes that toxicity to the bone marrow (as detailed under 3.2.1.1 or 3.2.1.2) in itself may provide sufficient evidence of bone marrow exposure to conclude on the validity of a negative outcome of a study based on the *in vivo* Mammalian Erythrocyte MN assay.

All other direct or indirect lines of evidence of bone marrow exposure should be assessed within a weight-of-evidence approach (EFSA, 2017a, to be published in August 2017) to decide whether they might provide sufficient reassurance of a valid negative test result. In this context, it should be assessed on a case-by-case basis whether the detection of any (radioactive) test substance in bone marrow or the detected levels in blood/plasma might be biologically relevant to the target tissue (EFSA, 2017b, to be published in August 2017).

3.2.4. Additional information for the assessment of genotoxicity following a negative result in the Mammalian Erythrocyte Micronucleus test, in the absence of definitive confirmation of bone marrow exposure

Depending on the pattern of the results of the *in vitro* tests for which an *in* vivo Mammalian Erythrocyte MN test has been used as a follow-up, different scenarios might be distinguished, e.g.:

a) There is evidence that the substance is mainly aneugenic. In this scenario, the *in vivo* mammalian erythrocyte MN test is the only validated method to investigate aneugenicity.

b) The *in vitro* battery of tests has given negative results in a test for gene mutation (e.g. bacterial reverse mutation test) and a positive result in a mammalian cell MN test in the absence of rat liver S9 mix with evidence that the result is due to clastogenicity rather than aneugenicity. In this scenario, the *in vivo* MN test should be considered in combination with a comet assay in the first site of tissue contact (e.g. duodenum).

c) The *in vitro* battery of tests has given negative results in a test for gene mutation (e.g. bacterial reverse mutation test) and a positive clastogenic result in a mammalian cell MN test exclusively (or predominantly) in the presence of rat liver S9 mix. In this scenario, the MN test should be considered in combination with a comet assay in the liver (EFSA, 2011).

In addition, the SC is aware of international discussions about non-validated *in vivo* MN tests in tissues other than bone marrow and peripheral blood, such as in the liver and in tissues of the gastro-intestinal tract (e.g. stomach or colon). Although they could provide useful scientific information for substances that are positive in *in vitro* MN tests, further actions for development are needed to fully establish the methods and identify their sensitivity and specificity (Report of the 6th International Workshop on Genotoxicity Test Procedures; Martus *et al.,* 2015; Uno *et al.,* 2015a,b).

3.3. The use of data in a weight-of-evidence approach to conclude on the genotoxic potential of substances and the consequent setting of health-based reference values for use in human health risk assessment.

The weight-of-evidence approach is applied at different levels, e.g. in concluding on the validity or informative value of individual studies or in assessing specific endpoints such as genotoxicity, but also at a higher level, i.e. integrating all endpoints within the overall hazard assessment.

In line with the Scientific Opinion on genotoxicity testing strategies and Guidance on Weight of Evidence (EFSA 2011 and EFSA 2017a, publication foreseen in August), the SC recommends 'a clearly documented weight-of-evidence approach for the evaluation and interpretation of genotoxicity data', taking into account not only the quality and availability of the data on genotoxicity itself, but also all other relevant data that may be available. The SC notes that for different areas of regulatory risk assessment, the basis for evaluation of genotoxicity may differ with respect to the extent of the available dataset and its robustness. When considering pre-existing or non-standard data, a case-by-case approach using expert judgement is required. Any uncertainties (e.g. relating to reliability and relevance of studies, information gaps or conflicting data) should be clearly described.

The SC notes that lacking or inadequate genotoxicity data cannot generally be substituted with data from studies involving other endpoints including carcinogenicity and reproductive toxicity, as genotoxicity is an important endpoint *per se*. Besides providing information about potential irreversible DNA damage that may result in heritable or degenerative diseases, genotoxicity data are also used to predict potential carcinogenicity (EFSA, 2011).

3.3.1 Conclusion on the genotoxic potential of substances

Re-evaluation of all available *in vitro* and *in vivo* results of the standard test battery according to the published EFSA SC Opinion on genotoxicity testing strategies (EFSA, 2011) according to the EFSA SC Guidance on the use of the weight-of-evidence approach in scientific assessments (EFSA, 2017a, publication foreseen in August) is the first recommended step to conclude on the genotoxicity of substances. Particular studies belonging to the standard or preferred battery of genotoxicity tests (EFSA, 2011) may not have been conducted according to the revised OECD guidelines from 2016 or

may not be available (e.g. TGR and comet assay). If the dataset includes any of the two assays specifically addressed in questions 1 and 2 of the Terms of Reference (liver UDS assay *in vivo*, Mammalian Erythrocyte MN test *in vivo*), the advice provided in answering those questions should be taken into consideration for the assessment of the informative value of the resulting test data and of the appropriateness of these assays.

In cases where, based on the available genotoxicity studies, it is not possible to conclude on genotoxicity with confidence, i.e. there is high uncertainty (e.g. owing to the absence of the standard or preferred battery of tests, or where there may be positive results *in vitro* but where negative results in *in vivo* studies may have limitations), the assessor may take into consideration in a second step any additional available data that may assist in reducing the uncertainty. Data that might be considered include: studies to elucidate the mode of action of the substance under consideration, information on carcinogenicity testing, reproductive toxicity testing, structural alerts or read-across data from structurally related substances. If after assessing all the available data and considering animal welfare issues, it is still not possible to conclude on the genotoxicity, additional information would be needed to reduce the uncertainty.

3.3.2 Establishing health-based guidance values

Taking all available evidence into account, if the overall evaluation leaves no concern for genotoxicity *in vivo*, health-based guidance values (HBGV) may be established.

If, based on the overall assessment, concern for genotoxicity remains, derivation of a health-based guidance value is not considered appropriate. However, Chapter 8.1 of the SC Opinion on genotoxicity testing strategies (2011) describes some circumstances under which genotoxicity might occur only at doses resulting in saturation of detoxification pathways or in cases of substances that interact with molecular targets other than DNA (e.g. DNA polymerases, topoisomerases, and spindle proteins). In such cases, provided robust data on the underlying mode of action are available and taking into account all other relevant information, establishing a health-based guidance value might be possible.

In cases of re-evaluations (e.g. pesticidal active substances), if the uncertainties in the assessment of genotoxicity cannot be resolved, establishing an HBGV is not considered appropriate.

4. Conclusions and recommendations

Question 1:

Regarding the suitability of the UDS to follow-up positive results in *in vitro* gene mutation tests, the SC concluded that, taking into consideration the current EU legislation and guidance documents, this question should be addressed in both a retrospective and a prospective way.

- For future assessments (new submissions), the EFSA SC is not aware of situations or chemical classes that can be identified in which the UDS could be considered preferable to the TGR or comet assay. Therefore, it is recommended no longer performing the UDS test.
- For re-assessment, in cases of already existing UDS data as a follow-up of a positive *in vitro* mutation test, there might be positive or negative results:
    - Test results may be considered as adequate to assess genotoxic potential only in cases with positive results.
    - In cases with negative results, other more reliable tests such as TGR or *in vivo* comet assay would be needed to complete the assessment.

Question 2:

Regarding the adequacy to demonstrate target tissue exposure in *in vivo* studies, particularly in the Mammalian Erythrocyte MN test, the SC concluded that, taking into consideration the current EU legislation and guidance documents, this question should be addressed in both direct and indirect lines of evidence of bone marrow exposure.

a) Toxicity to the bone marrow in itself may provide sufficient evidence of bone marrow exposure to allow concluding on the validity of a negative outcome of a study based on the *in vivo* Mammalian Erythrocyte MN assay. All other direct or indirect lines of evidence of bone marrow exposure should be assessed within a weight-of-evidence approach to decide whether they might provide sufficient reassurance of a valid negative test result.

b) In the absence of information to definitively confirm bone marrow exposure, further testing would be required to conclude that the substance is not genotoxic (see section 3.2.4.).

Question 3:

Regarding the assessment of the genotoxic potential of substances in a weight-of-evidence approach and the consequent setting of health-based guidance values for use in human health risk assessment, the SC concluded that:

a) In case it is not possible to conclude on genotoxicity with confidence, the assessor may, in a second step, take into consideration any additional available data that may assist in reducing the uncertainty, including studies on mode of action, information on structural alerts or read-across data from structurally related substances. If it is still not possible to conclude on genotoxicity, additional information would be needed to reduce the uncertainty.

b) If the overall evaluation, taking all of the available evidence into account, leaves no concerns for genotoxicity, health-based guidance values may be established. However, if based on the overall assessment, concerns for genotoxicity remain, derivation of health-based guidance values is considered inappropriate.

References

Butterworth BE, Ashby J, Bermudez E, Casciano D, Mirsalis J, Probst G and Williams G, 1987. A protocol and guide for the in vivo rat hepatocyte DNA-repair assay. Mutation Research, 1987, 189(2):123–133.

Carere A, Mohn GR, Parry JM, Sors AI and Nolan CV, 1995. Methods and testing strategies for evaluating the genotoxic properties of chemicals. European Commission Report EUR 15945 EN, CEC, Luxembourg. Available online: https://publications.europa.eu/en/publication-detail/-/publication/832c2e1c-3b1e-4414-8789-b576418f2c89

EC (European Commission), 2013: Commission communication in the framework of the implementation of Commission Regulation (EU) 283/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. Official Journal of the European Union, C 95/7, 16pp. Available online: http://ctgb.nl/docs/default-source/gewas.-toetsingskader/evaluation-manual-ppp-1107-2.0/commission_communication_2013_c95_02---data-requirements-for-plant-protection-products.pdf?sfvrsn=2

ECHA (European Chemical Agency), 2016. Guidance on information requirements and chemical safety assessment, chapter R.7a: endpoint specific guidance. Version 5.0. Available online: https://chesar.echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf

EFSA Scientific Committee, 2011. Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA Journal, 9(9):2379. doi:10.2903/j.efsa.2011.2379

EFSA Scientific Committee, 2017a. Guidance on the use of the weight of evidence approach in scientific assessments (adopted by the Scientific Committee in its plenary meeting on 12 July 2017 and to be published in August 2017 (version published for public consultation: Available online: http://www.efsa.europa.eu/de/consultations/call/170306-0)

EFSA Scientific Committee, 2017b. Guidance on biological relevance (adopted by the Scientific Committee on 12 July 2017 and to be published in August 2017).

European Union Reference Laboratory for Alternatives to Animal Testing – EURL ECVAM, 2017. EURL EVCAM genotoxicity and carcinogenicity consolidated database of Ames positive chemicals. Available at: https://eurl-ecvam.jrc.ec.europa.eu/databases/genotoxicity-carcinogenicity-db

ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use), 2011. Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use: S2(R1). Available at: http://www.ich.org/products/guidelines/safety/article/safety-guidelines.html

Kirkland D and Speit G, 2008. Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens III. Appropriate follow-up testing in vivo. Mutation Research, 654(2):114–132. doi:10.1016/j.mrgentox.2008.05.002

Kirkland D, Zeiger E, Madia F and Corvi R, 2014. Can in vitro mammalian cell genotoxicity test results be used to complement positive results in the Ames test and help predict carcinogenic or in vivo genotoxic activity? II. Construction and analysis of a consolidated database. Mutation Research, 775–776:69–80. doi:10.1016/j.mrgentox.2014.10.006

Martus HJ, Hayashi M, Honma M, Kasper P, Gollapudi B, Mueller L, Schoeny R, Uno Y and Kirkland DJ, 2015. Summary of major conclusions from the 6th International Workshop on Genotoxicity Testing (IWGT), Foz do Iguaçu, Brazil. Mutation Research. Genetic Toxicology and Environmental Mutagenesis, 783:1–5. doi:10.1016/j.mrgentox.2015.02.004

Mirsalis JC and Butterworth BE, 1980. Detection of unscheduled DNA synthesis in hepatocytes isolated from rats treated with genotoxic agents: an in vivo-in vitro assay for potential carcinogens and mutagens. Carcinogenesis, 1(7):621–625. doi:10.1093/carcin/1.7.621

Mirsalis JC, Tyson CK and Butterworth BE, 1982. Detection of genotoxic carcinogens in the in vivo-in vitro hepatocyte DNA repair assay. Environmental and Molecular Mutagenesis, 4(5):553–562. doi:10.1002/em.2860040506

OECD (Organisation for Economic Co-operation and Development), 1997. Test No. 486: Unscheduled DNA Synthesis (UDS) Test with Mammalian Liver Cells in vivo, available http://www.oecd-ilibrary.org/environment/test-no-486-unscheduled-dna-synthesis-uds-test-with-mammalian-liver-cells-in-vivo_9789264071520-en;jsessionid=drtvsk7mctik.x-oecd-live-03

OECD (Organisation for Economic Co-operation and Development), 2008. Test No. 407: Repeated Dose 28-day Oral Toxicity Study in Rodents, available at http://www.oecd-ilibrary.org/environment/test-no-407-repeated-dose-28-day-oral-toxicity-study-in-rodents_9789264070684-en;jsessionid=drtvsk7mctik.x-oecd-live-03

OECD (Organisation for Economic Co-operation and Development), 1998. Test No. 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents, available at http://www.oecd-ilibrary.org/environment/test-no-408-repeated-dose-90-day-oral-toxicity-study-in-rodents_9789264070707-en;jsessionid=drtvsk7mctik.x-oecd-live-03

OECD (Organisation for Economic Co-operation and Development), 1998. Test No. 409: Repeated Dose 90-Day Oral Toxicity Study in Non-Rodents, available at http://www.oecd-ilibrary.org/environment/test-no-409-repeated-dose-90-day-oral-toxicity-study-in-non-rodents_9789264070721-en;jsessionid=drtvsk7mctik.x-oecd-live-03

OECD (Organisation for Economic Co-operation and Development), 2009. Test No. 451: Carcinogenicity Studies, available at http://www.oecd-ilibrary.org/environment/test-no-451-carcinogenicity-studies_9789264071186-en;jsessionid=drtvsk7mctik.x-oecd-live-03

OECD (Organisation for Economic Co-operation and Development 2009) Test No. 452: Chronic Toxicity Studies, available at http://www.oecd-ilibrary.org/environment/test-no-452-chronic-toxicity-studies_9789264071209-en;jsessionid=drtvsk7mctik.x-oecd-live-03

OECD (Organisation for Economic Co-operation and Development), 2009. Test No. 453: Combined Chronic Toxicity/Carcinogenicity Studies, available at http://www.oecd-ilibrary.org/environment/test-no-453-combined-chronic-toxicity-carcinogenicity-studies_9789264071223-en;jsessionid=drtvsk7mctik.x-oecd-live-03

OECD (Organisation for Economic Co-operation and Development), 2013. Test No. 488: Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays, available at http://www.oecd-ilibrary.org/environment/test-no-488-transgenic-rodent-somatic-and-germ-cell-gene-mutation-assays_9789264203907-en and replacing version published in 2011 (http://www.oecd-ilibrary.org/environment/test-no-488-transgenic-rodent-somatic-and-germ-cell-gene-mutation-assays_9789264122819-en)

OECD (Organisation for Economic Co-operation and Development), 2014. Test No. 475: Mammalian Bone Marrow Chromosome Aberration Test, available http://www.oecd-ilibrary.org/environment/test-no-475-mammalian-bone-marrow-chromosomal-aberration-test_9789264224407-en;jsessionid=drtvsk7mctik.x-oecd-live-03 and replacing the version published in 1997 (available http://www.oecd-ilibrary.org/environment/test-no-475-mammalian-bone-marrow-chromosome-aberration-test_9789264071308-en;jsessionid=drtvsk7mctik.x-oecd-live-03)

OECD (Organisation for Economic Co-operation and Development), 2016. Test No. 489: In Vivo Mammalian Alkaline Comet Assay, available at http://www.oecd-ilibrary.org/environment/test-no-489-in-vivo-mammalian-alkaline-comet-assay_9789264264885-en and replacing version published in 2014 (http://www.oecd-ilibrary.org/environment/test-no-489-in-vivo-mammalian-alkaline-comet-assay_9789264224179-en)

OECD (Organisation for Economic Co-operation and Development), 2016. Test No. 487: In Vitro Mammalian Cell Micronucleus Test, available at http://www.oecd-ilibrary.org/environment/test-no-487-in-vitro-mammalian-cell-micronucleus-test_9789264264861-en and replacing version published in 2014 (http://www.oecd-ilibrary.org/environment/test-no-487-in-vitro-mammalian-cell-micronucleus-test_9789264224438-en that was replacing version published in 2010

(http://www.oecd-ilibrary.org/environment/test-no-487-in-vitro-mammalian-cell-micronucleus-test_9789264091016-en)

OECD (Organisation for Economic Co-operation and Development), 2016. Test No. 474: Mammalian Erythrocyte Micronucleus Test, available http://www.oecd-ilibrary.org/environment/test-no-474-mammalian-erythrocyte-micronucleus-test_9789264264762-en;jsessionid=drtvsk7mctik.x-oecd-live-03 that replaces the version published in 2014 (available http://www.oecd-ilibrary.org/environment/test-no-474-mammalian-erythrocyte-micronucleus-test_9789264224292-en;jsessionid=drtvsk7mctik.x-oecd-live-03 that was replacing the version published in 1997 (available http://www.oecd-ilibrary.org/environment/test-no-474-mammalian-erythrocyte-micronucleus- test_9789264071285-en;jsessionid=drtvsk7mctik.x-oecd-live-03).

OECD (Organisation for Economic Co-operation and Development), 2016. Overview of the set of OECD Genetic Toxicology Test Guidelines and updates performed in 2014–2015. Series on Testing & Assessment No. 238, 58 pp. Available online: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2016)33&doclanguage=en

Uno Y, Morita T, Luijten M, Beevers C, Hamada S, Itoh S, Ohyama W and Takasawa H, 2015a. Recommended protocols for the liver micronucleus test: report of the IWGT working group. Mutation Research. Genetic Toxicology and Environmental Mutagenesis, 783:13–18. doi:10.1016/j.mrgentox.2014.10.010.

Uno Y, Morita T, Luijten M, Beevers C, Hamada S, Itoh S, Ohyama W and Takasawa H, 2015b. Micronucleus test in rodent tissues other than liver or erythrocytes: report of the IWGT working group. Mutation Research. Genetic Toxicology and Environmental Mutagenesis. 783:19–22. doi:10.1016/j.mrgentox.2015.03.001.

Abbreviations

ADME study: Absorption, Distribution, Metabolism, and Excretion study

Ames test: Bacterial Reverse Mutation Test (OECD TG 471)

CA: Chromosomal aberration

ECVAM: European Centre for the Validation of Alternative Testing

HBGV: Health-Based Guidance Values

MLA: Mouse Lymphoma Assay (OECD TG 476)

MN: Micronucleus test

NCE: Normochromatic erythrocytes, i.e. mature erythrocytes

OECD: Organisation for the Economic Cooperation and Development

PCE: Polychromatic erythrocytes, i.e. immature erythrocytes

TG: Test Guideline

TGR: Transgenic Rodent Assay

UDS: Unscheduled DNA Synthesis

Appendix A – Analysis of the sensitivity of UDS, TGR and comet assays to detect carcinogens

Analysis of EU reference laboratory (EU RL) ECVAM database

This analysis was performed using data from the EURL ECVAM database (https://eurl-ecvam.jrc.ec.europa.eu/databases/genotoxicitycarcinogenicity-db). A rigorous methodology and defined criteria were applied for the selection and analysis of the data (Kirkland et al., 2014).

The database includes:

- Data for 726 Ames-positive chemicals from different sources: regulatory agencies, industry and literature databases covering different sectors (industrial chemicals, cosmetics ingredients, plant protection and pharmaceutical products).
- Only chemicals with valid *in vitro* and *in vivo* results for the genotoxicity endpoints and/or for carcinogenicity.
- Only chemicals with a known chemical identity (structure, purity, molecular weight, CAS number).
- Combinations into single entries of free bases and respective simple acid salts or R- and S-isomers for those chemicals where a similar behavior was expected and/or proven.
- Data for the following tests: *in vitro* tests [Ames, mouse lymphoma tk+/- (MLA) or gene *Hprt* locus, micronucleus (MN), chromosome aberration (CA)]; *in vivo* tests [MN, CA, UDS, transgenic models, DNA breakage (comet and alkaline elution assay)]; rodent carcinogenicity.

Criteria for "overall calls" within the database

- "Overall Calls" were defined for each genotoxicity assay *in vitro* and *in vivo* and carcinogenicity by following defined criteria for the reliability of each study and quality of data for those chemicals appearing in more than one source with different calls. Four categories were considered: (+), (-), (E) and (I). Where information was missing, even for those chemicals with one single data entry, scientific literature was consulted.
- Overall calls were made for carcinogenicity as follows:
    - A positive (+) call was assigned if a positive response was reported in at least one sex of either rats or mice.
    - An equivocal (E) call was assigned if the substance was tested in both sexes of rats and mice and at least one equivocal call was noted, and the other calls were negative. A response is considered as equivocal when it is weak or not reproduced between experiments or between laboratories.
    - A negative (−) call was assigned if the substance was tested in both sexes of rats and mice and all four groups yielded negative responses.
    - An inadequate, and therefore inconclusive (I), call was assigned if the substance was tested in both sexes of rats and mice and the results in at least one of the groups was considered to be compromised because of inadequate dosing (too low), excess mortality, or a concurrent infection, and the other groups produced either negative or equivocal responses. If a substance was tested only in rats or mice and was not carcinogenic, that result was also considered inconclusive. Inconclusive (I) calls were considered "no valid data" and not included in the data analyses (see below).
- In arriving at overall calls for the various genotoxicity endpoints *in vitro* and *in vivo*, the quality of the study, robustness of the protocol, and quality of the data, where available, were

taken into account. For example, negative results from a recent Good Laboratory Practices (GLP) study conducted to current guidelines was considered more meaningful than a negative result from an old study that did not comply with current guidelines. Where conflicting results were reported in the different databases, the numbers of + and – calls were not considered as important as the quality and robustness of the tests, and whether the results had been obtained in different studies or from different publications, i.e. independent confirmations were more important than the same study result [e.g. from National Toxicology Program (NTP)] reported in different databases.

- For overall calls for each chemical and each genotoxicity endpoint *in vitro* or *in vivo*, the following criteria were adopted:

    - An overall positive (+) call for the *in vitro* studies was given regardless of whether the positive finding was noted only in the absence or only in the presence of S9, and regardless of the cell type. A positive call was made whether rat, hamster or mouse S9 was used.

    - An overall positive (+) call for the *in vivo* genotoxicity studies was given if there was clear evidence of a positive response from a single study (rats or mice, males or females). If there was clear weight of evidence from more than one study, or if a substance was positive in one species or sex and negative in the other, it was assigned a positive call. In the case of the latter, if two studies gave different results, but it was clear that systemic exposures were greater in the positive than in the negative study, an overall call of + was given.

    - An overall negative (–) call for both *in vitro* and *in vivo* studies was given when all the requirements of the current OECD guidelines or recommended best practices were fulfilled and there was no evidence of a positive or equivocal response. A negative call was made for *in vivo* genotoxicity studies only if there was evidence that the test substance reached the target tissue, otherwise it was considered inconclusive (see below).

    - An overall equivocal (E) call was given if results were ambiguous, doubtful, questionable, or inconsistent (e.g. a positive and a negative test) within a study, or if a dose-related increase in effects was noted close to the borderline of biological significance, but they were not biologically and/or statistically significant and no independent repeat experiment was done to check the response and produce a clear conclusion. An "E" call was also used where there were both positive and negative findings across different studies of apparent equal validity, and where the weight of evidence did not allow a clear positive or negative overall outcome to be concluded.

    - An overall inconclusive (I) call was given in cases of negative or unclear results, where no firm conclusion could be made in terms of meeting the requirements of the current OECD guidelines or recommended best practices.

Analysis of sensitivity of UDS, transgenics and comet assays to detect carcinogens - results

Table 1. Ames-positive chemicals which were tested both in the UDS and for carcinogenicity

	CARC		
UDS	**Postive**	**Negative**	**tot**
Postive/Eq	47	2	49
Negative	31	8	39
tot	78	10	
	60.26%	80.00%	
	sensitivity	**specificity**	

Table 2. Ames-positive chemicals which were tested both in the TGR and for carcinogenicity

	CARC		
Transgenic	**Postive**	**Negative**	**tot**
Postive/Eq	67	0	67
Negative	9	1	10
tot	76	1	
	88.16%	100.00%	
	sensitivity	**specificity**	

Table 3. Ames-positive chemicals which were tested both in the comet and for carcinogenicity

	CARC		
Comet	**Postive**	**Negative**	**tot**
Postive/Eq	52	4	56
Negative	5	6	11
tot	57	10	
	91.23%	60.00%	
	sensitivity	**specificity**	

Table 4. Prediction of carcinogenicity by UDS, TGR and comet assay for a common set of carcinogens

	Positive/Tested (sensitivity)
UDS	13/19 (68%)
TGR	14/19 (74%)
Comet	18/19 (95%)

930 **Table 5. Outcome of testing of Ames positive chemicals in UDS and other *in vivo* genotoxicity assays.**

	In vivo genotoxicity*		
UDS	Positive/Eq	Negative	
Positive/Eq	40	4	
Negative	26	44	
Total	66	48	
	60.6 %	91.6 %	73.7%
	Concordant results	Concordant results	Overall concordance

931 * In either Mammalian Erythrocyte MN, CA, TGR or comet assays

932 **Terms used to describe the performance of genotoxicity tests to predict rodent**
933 **carcinogenicity (from EFSA Opinion on Genotoxicity testing strategies, EFSA J. 2011)**

934

	Carcinogens	**Non-carcinogens**
Genotoxicity positive	A	B
Genotoxicity negative	C	D

935

Sensitivity	% correctly identified carcinogens	A/(A+C) * 100
Specificity	% correctly identified non-carcinogens	D/(B+D) * 100
Concordance	% correctly identified carcinogens and non-carcinogens	(A + D)/(A+B+C+D) * 100

936

937

938

939

940

941

942

943

944

945

946

947

948

949

Appendix B – Examples of endpoints which might contribute to the weight-of-evidence assessment

950
951 ### B.1.1. Clinical signs of toxicity

952 General clinical observations (cage-side) include certain CNS-related signs (e.g. reduction of
953 spontaneous activity, ataxia, sedation, staggering, salivation, abdominal position). Detailed clinical
954 observations (outside the cage), which should be made in OECD test guideline studies such as the 28-
955 day and 90-day studies (e.g. OECD TG 407, TG 408, TG 409) and the chronic toxicity studies (e.g.
956 OECD TG 452, TG 453), should include e.g. changes in autonomic activity (lacrimation, piloerection,
957 pupil size and unusual respiratory pattern), in gait, posture and response to handling as well as the
958 presence of clonic or tonic movements, stereotypies or bizarre behaviour. Such clinical signs would
959 usually occur only if the test substance is systemically bioavailable and are thus direct evidence of
960 systemic bioavailability, and consequently, indirect evidence of bone marrow exposure.
961
962 ### B.1.2. Haematology

963 The following haematological examinations should be made in OECD test guideline studies such as the
964 28-day and 90-day studies (e.g. OECD TG 407, TG 408, TG 409) and the chronic toxicity studies (e.g.
965 OECD TG 452, TG 453): haematocrit, haemoglobin concentration, erythrocyte count, total and
966 differential leukocyte count, platelet count and a measure of blood clotting time/potential. Other
967 haematology parameters such as Heinz bodies or other atypical erythrocyte morphology or
968 methaemoglobin may be measured as appropriate depending on the toxicity of the substance.
969 Haematological changes observed in a repeated-dose toxicity study such as e.g. decreased
970 erythrocyte count, haemoglobin concentration, haematocrit, leukocyte count, if evaluated as being
971 test-substance related and toxicologically significant, are direct evidence of systemic bioavailability,
972 and consequently, indirect evidence of bone marrow exposure.
973

974 ### B.1.3. Clinical biochemistry

975 Clinical biochemistry determinations to investigate major toxic effects in tissues and, specifically,
976 effects on kidney and liver should be made in OECD test guideline studies such as the 28-day and 90-
977 day studies (e.g. OECD TG 407, TG 408, TG 409) and the chronic toxicity studies (e.g., OECD TG 452,
978 TG 453) differ slightly between the different test guidelines. In general, the following parameters
979 should be determined: sodium, potassium, glucose, total cholesterol, urea, creatinine, total protein
980 and albumin, at least two enzymes indicative of hepatocellular effects (such as alanine
981 aminotransferase, aspartate aminotransferase, alkaline phosphatase, γ-glutamyl transpeptidase and
982 glutamate dehydrogenase), bile acids, bilirubin. Other clinical chemistry parameters such as fasting
983 triglycerides, specific hormones and cholinesterase may be measured as appropriate, depending on
984 the toxicity of the substance (chemicals in certain classes or on a case-by-case basis). Changes in
985 clinical biochemistry parameters observed in a repeated-dose toxicity study, if evaluated as being test-
986 substance related and toxicologically significant, are direct evidence of systemic bioavailability, and
987 consequently, indirect evidence of bone marrow exposure.
988

989 ### B.1.4. Urinalysis

990 The following urinalysis determinations should be made in OECD test guideline studies such as the 90-
991 day non-rodent study (OECD TG 409) and the chronic toxicity studies (e.g. OECD TG 452, TG 453):
992 appearance, volume, osmolality or specific gravity, pH, protein, and glucose. These determinations are
993 optional in the 28-day and 90-day rodent studies. Changes in urinalysis parameters observed in a
994 repeated-dose toxicity study, if evaluated as being test-substance related and toxicologically
995 significant, are direct evidence of systemic bioavailability, and consequently, indirect evidence of bone
996 marrow exposure.
997
998

B.1.5. Histopathology of organs and tissues

Several organs and tissues should be examined histopathologically in OECD test guideline studies such as the 28-day and 90-day studies (e.g. OECD TG 407, TG 408, TG 409) and the chronic toxicity studies (e.g. OECD TG 452, TG 453). In general, the following systemically exposed organs and tissues should be examined: brain, spinal cord, liver, kidneys, adrenals, spleen, heart, thymus, thyroid, gonads (testis and ovaries), accessory sex organs (uterus and cervix, epididymides, prostate, seminal vesicles), urinary bladder, lymph nodes, peripheral nerve, skeletal muscle, bone marrow. In the longer duration studies (90-day and chronic) the following additional systemically exposed organs and tissues should be examined: pancreas, mammary gland, pituitary, parathyroid, aorta, gall bladder (mouse). A histopathological change in any of these organs or tissues observed in a repeated-dose toxicity study, if evaluated as being test-substance related, is direct evidence of systemic bioavailability, and consequently, indirect evidence of bone marrow exposure.