OECD/OCDE 489
Adopted:
29 July 2016

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

In Vivo Mammalian Alkaline Comet Assay

INTRODUCTION

1. The *in vivo* alkaline comet (single cell gel electrophoresis) assay (hereafter called simply the comet assay) is used for the detection of DNA strand breaks in cells or nuclei isolated from multiple tissues of animals, usually rodents, that have been exposed to potentially genotoxic material(s). The comet assay has been reviewed and recommendations have been published by various expert groups (1) (2) (3) (4) (5) (6) (7) (8) (9) (10). This Test Guideline is part of a series of Test Guidelines on genetic toxicology. A document that provides succinct information on genetic toxicology testing and an overview of the recent changes that were made to these Test Guidelines has been developed (11).

2. The purpose of the comet assay is to identify substances that cause DNA damage. Under alkaline conditions (>pH 13), the comet assay can detect single and double stranded breaks, resulting, for example, from direct interactions with DNA, alkali labile sites or as a consequence of transient DNA strand breaks resulting from DNA excision repair. These strand breaks may be repaired, resulting in no persistent effect, may be lethal to the cell, or may be fixed into a mutation resulting in a permanent viable change. They may also lead to chromosomal damage which is also associated with many human diseases including cancer.

3. A formal validation trial of the *in vivo* rodent comet assay was performed in 2006-2012, coordinated by the Japanese Center for the Validation of Alternative Methods (JaCVAM), in conjunction with the European Centre for the Validation of Alternative Methods (ECVAM), the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) (12). This Test Guideline includes the recommended use and limitations of the comet assay, and is based on the final protocol (12) used in the validation trial, and on additional relevant published and unpublished (laboratories proprietary) data.

4. Definitions of key terms are set out in Annex 1. It is noted that many different platforms can be used for this assay (microscope slides, gel spots, 96-well plates etc.). For convenience the term “slide” is used throughout the remainder of this document but encompasses all of the other platforms.

© OECD, (2016)

You are free to use this material subject to the terms and conditions available at http://www.oecd.org/termsandconditions/.

This Guideline was adopted by the OECD Council by written procedure on 29 July 2016 [C(2016)103].

INITIAL CONSIDERATIONS AND LIMITATIONS

5. The comet assay is a method for measuring DNA strand breaks in eukaryotic cells. Single cells/nuclei embedded in agarose on a slide are lysed with detergent and high salt concentration. This lysis step digests the cellular and nuclear membranes and allows the release of coiled DNA loops generally called nucleoids and DNA fragments. Electrophoresis at high pH results in structures resembling comets, which, by using appropriate fluorescent stains, can be observed by fluorescence microscopy; DNA fragments migrate away from the "head" into the "tail" based on their size, and the intensity of the comet tail relative to the total intensity (head plus tail) reflects the amount of DNA breakage (13) (14) (15).

6. The *in vivo* alkaline comet assay is especially relevant to assess genotoxic hazard in that the assay's responses are dependent upon *in vivo* ADME (absorption, distribution, metabolism and excretion), and also on DNA repair processes. These may vary among species, among tissues and among the types of DNA damage.

7. To fulfil animal welfare requirements, in particular the reduction in animal usage (3Rs - Reduction, Refinement, Replacement - principles), this assay can also be integrated with other toxicological studies, e.g., repeated dose toxicity studies (10) (16) (17), or the endpoint can be combined with other genotoxicity endpoints such as the *in vivo* mammalian erythrocyte micronucleus assay (18) (19) (20). The comet assay is most often performed in rodents, although it has been applied to other mammalian and non-mammalian species. The use of non-rodent species should be scientifically and ethically justified on a case-by-case basis and it is strongly recommended that the comet assay only be performed on species other than rodents as part of another toxicity study and not as a standalone test.

8. The selection of route of exposure and tissue(s) to be studied should be determined based on all available/existing knowledge of the test chemicals e.g. intended/expected route of human exposure, metabolism and distribution, potential for site-of-contact effects, structural alerts, other genotoxicity or toxicity data, and the purpose of the study. Thus, where appropriate, the genotoxic potential of the test chemicals can be assayed in the target tissue(s) of carcinogenic and/or other toxic effects. The assay is also considered useful for further investigation of genotoxicity detected by an *in vitro* system. It is appropriate to perform an *in vivo* comet assay in a tissue of interest when it can be reasonably expected that the tissue of interest will be adequately exposed.

9. The assay has been most extensively validated in somatic tissues of male rats in collaborative studies such as the JaCVAM trial (12) and in Rothfuss et al. 2010 (10). The liver and stomach were used in the JaCVAM international validation trial. The liver, because it is the most active organ in metabolism of substances and also frequently a target organ for carcinogenicity. The stomach, because it is usually first site of contact for substances after oral exposure, although other areas of the gastro-intestinal tract such as the duodenum and jejunum should also be considered as site-of-contact tissues and may be considered more relevant for humans than the rodent glandular stomach. Care should be taken to ensure that such tissues are not exposed to excessively high test substance concentrations (21). The technique is in principle applicable to any tissue from which analysable single cell/nuclei suspensions can be derived. Proprietary data from several laboratories demonstrate its successful application to many different tissues, and there are many publications showing the applicability of the technique to organs or tissues other than liver and stomach, e.g., jejunum (22), kidney (23) (24), skin (25) (26)**,** or urinary bladder (27) (28), lungs and bronchoalveolar lavage cells (relevant for studies of inhaled substances) (29) (30), and tests have also been performed in multiple organs (31) (32).

10. Whilst there may be an interest in genotoxic effects in germ cells, it should be noted that the

© OECD, (2016)

standard alkaline comet assay as described in this guideline is not considered appropriate to measure DNA strand breaks in mature germ cells. Since high and variable background levels in DNA damage were reported in a literature review on the use of the comet assay for germ cell genotoxicity (33), protocol modifications together with improved standardization and validation studies are deemed necessary before the comet assay on mature germ cells (e.g. sperm) can be included in the test guideline. In addition, the recommended exposure regimen described in this guideline is not optimal and longer exposures or sampling times would be necessary for a meaningful analysis of DNA strand breaks in mature sperm. Genotoxic effects as measured by the comet assay in testicular cells at different stages of differentiation have been described in the literature (34) (35). However, it should be noted that gonads contain a mixture of somatic and germ cells. For this reason, positive results in whole gonad (testis) are not necessarily reflective of germ cell damage; nevertheless, they indicate that tested chemical(s) and/or its metabolites have reached the gonad.

11. Cross-links cannot be reliably detected with the standard experimental conditions of the comet assay. Under certain modified experimental conditions, DNA-DNA and DNA-protein crosslinks, and other base modifications such as oxidized bases might be detected (23) (36) (37) (38) (39). But further work would be needed to adequately characterize the necessary protocol modifications. Thus detection of cross linking agents is not the primary purpose of the assay as described here. The assay is not appropriate, even with modifications, for detecting aneugens.

12. Due to the current status of knowledge, several additional limitations (see Annex 3) are associated with the *in vivo* comet assay. It is expected that the Test Guideline will be reviewed in the future and if necessary revised in light of experience gained.

13. Before use of the Test Guideline on a mixture for generating data for an intended regulatory purpose, it should be considered whether, and if so why, it may provide adequate results for that purpose. Such considerations are not needed, when there is a regulatory requirement for testing of the mixture.

PRINCIPLE OF THE METHOD

14. Animals are exposed to the test chemical by an appropriate route. A detailed description of dosing and sampling is given in paragraphs 36-40. At the selected sampling time(s), the tissues of interest are dissected and single cells/nuclei suspensions are prepared (*in situ* perfusion may be performed where considered useful e.g. liver) and embedded in soft agar so as to immobilize them on slides. Cells/nuclei are treated with lysis buffer to remove cellular and/or nuclear membrane, and exposed to strong alkali e.g., pH≥13 to allow DNA unwinding and release of relaxed DNA loops and fragments. The nuclear DNA in the agar is then subjected to electrophoresis. Normal non-fragmented DNA molecules remain in the position where the nuclear DNA had been in the agar, while any fragmented DNA and relaxed DNA loops would migrate towards the anode. After electrophoresis, the DNA is visualized using an appropriate fluorescent stain. Preparations should be analysed using a microscope and full or semi-automated image analysis systems. The extent of DNA that has migrated during electrophoresis and the migration distance reflects the amount and size of DNA fragments. There are several endpoints for the comet assay. The DNA content in the tail (% tail DNA or % tail intensity) has been recommended to assess DNA damage (12) (40) (41) (42). After analysis of a sufficient number of nuclei, the data are analysed with appropriate methods to judge the assay results.

15. It should be noted that altering various aspects of the methodology, including sample preparation, electrophoresis conditions, visual analysis parameters (e.g. stain intensity, microscope bulb light intensity,

© OECD, (2016)

and use of microscope filters and camera dynamics) and ambient conditions (e.g. background lighting), have been investigated and may affect DNA migration (43) (44) (45) (46).

VERIFICATION OF LABORATORY PROFICIENCY

16. Each laboratory should establish experimental competency in the comet assay by demonstrating the ability to obtain single cell or nuclei suspensions of sufficient quality for each target tissue(s) for each species used. The quality of the preparations will be evaluated firstly by the % tail DNA for vehicle treated animals falling within a reproducible low range. Current data suggest that the group mean % tail DNA (based on mean of medians - see paragraph 57 for details of these terms) in the rat liver should be preferably not exceed 6%, which would be consistent with the values in the JaCVAM validation trial (12) and from other published and proprietary data. There are not enough data at this time to make recommendations about optimum or acceptable ranges for other tissues. This doesn't preclude the use of other tissues if justified. The test report should provide appropriate review of the performance of the comet assay in these tissues in relation to the published literature or from proprietary data. Firstly, a low range of %tail DNA in controls is desirable to provide sufficient dynamic range to detect a positive effect. Secondly, each laboratory should be able to reproduce expected responses for direct mutagens and pro-mutagens, with different modes of action as suggested in Table 1 (paragraph 29).

17. Positive substances may be selected, for example from the JaCVAM validation trial (12) or from other published data (see paragraph 9), if appropriate, with justification, and demonstrating clear positive responses in the tissues of interest. The ability to detect weak effects of known mutagens e.g. EMS at low doses, should also be demonstrated, for example by establishing dose-response relationships with appropriate numbers and spacing of doses. Initial efforts should focus on establishing proficiency with the most commonly used tissues e.g. the rodent liver, where comparison with existing data and expected results may be made (12). Data from other tissues e.g. stomach/duodenum/jejunum, blood etc. could be collected at the same time. The laboratory needs to demonstrate proficiency with each individual tissue in each species they are planning to study, and will need to demonstrate that an acceptable positive response with a known mutagen (e.g. EMS) can be obtained in that tissue.

18. Vehicle/negative control data should be collected so as to demonstrate reproducibility of negative data responses, and to ensure that the technical aspects of the assay were properly controlled or to suggest the need to re-establish historical control ranges (see paragraph 22).

19. It should be noted, that whilst multiple tissues can be collected at necropsy and processed for comet analysis, the laboratory needs to be proficient in harvesting multiple tissues from a single animal, thereby ensuring that any potential DNA lesion is not lost and comet analysis is not compromised. The length of time from euthanasia to removal of tissues for processing may be critical (see paragraph 44).

20. Animal welfare must be considered whilst developing proficiency in this test and therefore tissues from animals used in other tests can be used when developing competence in the various aspects of the test. Furthermore, it may not be necessary to conduct a full study during the stages of establishing a new test guideline method in a laboratory and fewer animals or test concentrations can be used when developing the necessary skills.

Historical control data

© OECD, (2016)

21. During the course of the proficiency investigations, the laboratory should build a historical database to establish positive and negative control ranges and distributions for relevant tissues and species. Recommendations on how to build and use the historical data (i.e. criteria for inclusion and exclusion of data in historical data and the acceptability criteria for a given experiment) can be found in the literature (47). Different tissues and different species, as well as different vehicles and routes of administrations, may give different negative control % tail DNA values. It is therefore important to establish negative control ranges for each tissue and species. Laboratories should use quality control methods, such as control charts (e.g. C-charts or X-bar charts (48)), to identify how variable their data are, and to show that the methodology is 'under control' in their laboratory. Selection of appropriate positive control substances, dose ranges and experimental conditions (e.g. electrophoresis conditions) may need also to be optimised for the detection of weak effects (see paragraph 17).

22. Any changes to the experimental protocol should be considered in terms of their consistency with the laboratory's existing historical control databases. Any major inconsistencies should result in the establishment of a new historical control database.

DESCRIPTION OF THE METHOD

Preparations

Selection of animal species

23. Common laboratory strains of healthy young adult rodents (6-10 weeks old at start of treatment though slightly older animals are also acceptable) are normally used. The choice of rodent species should be based on (i) species used in other toxicity studies (to be able to correlate data and to allow integrated studies), (ii) species that developed tumours in a carcinogenicity study (when investigating the mechanism of carcinogenesis), or (iii) species with the most relevant metabolism for humans, if known. Rats are routinely used in this test. However, other species can be used if ethically and scientifically justified.

Animal housing and feeding conditions

24. For rodents, the temperature in the experimental animal room ideally should be 22°C (±3°C). The relative humidity ideally should be 50-60%, being at least 30% and preferably not exceeding 70% other than during room cleaning. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The choice of diet may be influenced by the need to ensure a suitable admixture of a test chemical when administered by this route. Rodents should be housed in small groups (usually no more than five) of the same sex if no aggressive behaviour is expected. Animals may be housed individually only if scientifically justified. Solid floors should be used wherever possible as mesh floors can cause serious injury (49). Appropriate environmental enrichment must be provided.

Preparation of the animals

25. Animals are randomly assigned to the control and treatment groups. The animals are identified uniquely and acclimated to the laboratory conditions for at least five days before the start of treatment. The least invasive method of uniquely identifying animals must be used. Appropriate methods include ringing, tagging, micro-chipping and biometric identification. Toe and ear clipping are not scientifically justified in these tests. Cages should be arranged in such a way that possible effects due to cage placement are

© OECD, (2016)

minimized. At the commencement of the study, the weight variation of animals should be minimal and not exceed ± 20%.

© OECD, (2016)

Preparation of doses

26. Solid test chemicals should be dissolved or suspended in appropriate vehicles or admixed in diet or drinking water prior to dosing of the animals. Liquid test chemicals may be dosed directly or diluted prior to dosing. For inhalation exposures, test chemicals can be administered as gas, vapour, or a solid/liquid aerosol, depending on their physicochemical properties (50) (51).

27. Fresh preparations of the test chemical should be employed unless stability data demonstrate the acceptability of storage and define the appropriate storage conditions.

Test Conditions

Vehicle

28. The vehicle should not produce toxic effects at the dose volumes used, and should not be suspected of chemical reaction with the test substances. If other than well-known vehicles are used, their inclusion should be supported with reference data indicating their compatibility in terms of test animals, route of administration and endpoint. It is recommended that wherever possible, the use of an aqueous solvent/vehicle should be considered first. It should be noted that some vehicles (particularly viscous vehicles) can induce inflammation and increase background levels of DNA strand breaks at the site of contact, particularly with multiple administrations.

Controls

Positive controls

29. At this time, a group of a minimum of 3 analysable animals of one sex, or of each sex if both are used (see paragraph 32), treated with a positive control substance should normally be included with each test. In future, it may be possible to demonstrate adequate proficiency to reduce the need for positive controls. If multiple sampling times are used (e.g. with a single administration protocol) it is only necessary to include positive controls at one of the sampling times, but a balanced design should be ensured (see paragraph 48). It is not necessary to administer concurrent positive control substances by the same route as the test chemical, although it is important that the same route should be used when measuring site-of-contact effects. The positive control substances should be shown to induce DNA strand breaks in all of the tissues of interest for the test chemical, and EMS is likely to be the positive control of choice since it has produced DNA strand breaks in all tissues that have been studied. The doses of the positive control substances should be selected so as to produce moderate effects that critically assess the performance and sensitivity of the assay and could be based on dose-response curves established by the laboratory during the demonstration of proficiency. The % tail DNA in concurrent positive control animals should be consistent with the pre-established laboratory range for each individual tissue and sampling time for that species (see paragraph 16). Examples of positive control substances and some of their target tissues (in rodents) are included in Table 1. Substances other than those given in Table 1 can be selected if scientifically justified.

© OECD, (2016)

Table 1: Examples of positive control substances and some of their target tissues

Substances and CAS RN No.
Ethyl methanesulfonate (CAS RN 62-50-0) for any tissue
Ethyl nitrosourea (CAS RN 759-73-9) for liver and stomach, duodenum or jejunum
Methyl methanesulfonate (CAS RN 66-27-3) for liver, stomach, duodenum or jejunum, lung and bronchoalveolar lavage (BAL) cells, kidney, bladder, lung, testis and bone marrow/blood
N-Methyl-*N*′-nitro-*N*-nitrosoguanidine (CAS RN: 70-25-7) for stomach, duodenum or jejunum
1,2-Dimethylhydrazine 2HCl (CAS RN 306-37-6) for liver and intestine
N-methyl-*N*-nitrosourea (CAS RN 684-93-5) for liver, bone marrow, blood, kidney, stomach, jejunum, and brain.

Negative controls

30. A group of negative control animals, treated with vehicle alone, and otherwise treated in the same way as the treatment groups, should be included with each test for every sampling time and tissue. The % tail DNA in negative control animals should be within the pre-established laboratory background range for each individual tissue and sampling time for that species (see paragraph 16). In the absence of historical or published control data showing that no deleterious or genotoxic effects are induced by the chosen vehicle, by the number of administrations or by the route of administration, initial studies should be performed prior to conducting the full study, in order to establish acceptability of the vehicle control.

PROCEDURE

Number and Sex of Animals

31. Although there is little data on female animals from which to make comparison between sexes in relation to the comet assay, in general, other *in vivo* genotoxicity responses are similar between male and female animals and therefore most studies could be performed in either sex. Data demonstrating relevant differences between males and females (e.g. differences in systemic toxicity, metabolism, bioavailability, etc. including e.g. in a range-finding study) encourage the use of both sexes. In this case, it may be appropriate to perform a study in both sexes e.g. as part of a repeated dose toxicity study. It might be appropriate to use the factorial design in case both sexes are used. Details on how to analyse the data using this design are given in Annex 2.

32. Group sizes at study initiation (and during establishment of proficiency) should be established with the aim of providing a minimum of 5 analysable animals of one sex, or of each sex if both are used, per group (less in the concurrent positive control group – see paragraph 29). Where human exposure to chemicals may be sex-specific, as for example with some pharmaceuticals, the test should be performed with the appropriate sex. As a guide to maximum typical animal requirements, a study conducted according the parameters established in paragraph 33 with three dose groups and concurrent negative and positive controls (each group composed of five animals of a single sex) would require between 25 and 35 animals.

© OECD, (2016)

Treatment schedule

33. Animals should be given daily treatments over a duration of 2 or more days (i.e. two or more treatments at approximately 24 hour intervals), and samples should be collected once at 2-6 h (or at the Tmax) after the last treatment (12). Samples from extended dose regimens (e.g., 28-day daily dosing) are acceptable. Successful combination of the comet and the erythrocyte micronucleus test has been demonstrated (10) (19). However careful consideration should be given to the logistics involved in tissue sampling for comet analysis alongside the requirements of tissue sampling for other types of toxicological assessments. Harvest 24 hours after the last dose, which is typical of a general toxicity study, is not appropriate in most cases (see paragraph 40 on sampling time). The use of other treatment and sampling schedules should be justified (see annex 3). For example single treatment with multiple sampling could be used however, it should be noted that more animals will be required for a study with a single administration study because of the need for multiple sampling times, but on occasions this may be preferable, e.g. when the test chemical induces excessive toxicity following repeated administrations.

34. Whatever way the test is performed, it is acceptable as long as the test chemical gives a positive response or, for a negative study, as long as direct or indirect evidence supportive of exposure of, or toxicity to, the target tissue(s) has been demonstrated or if the limit dose is achieved (see paragraph 37):

35. Test chemicals also may be administered as a split dose, i.e., two treatments on the same day separated by no more than 2-3 hours, to facilitate administering a large volume. Under these circumstances, the sampling time should be scheduled based on the time of the last dosing (see paragraph 40).

Dose Levels

36. If a preliminary range-finding study is performed because there are no suitable data available from other relevant studies to aid in dose selection, it should be performed in the same laboratory, using the same species, strain, sex, and treatment regimen to be used in the main study according to current approaches for conducting dose range-finding studies. The study should aim to identify the maximum tolerated dose (MTD), defined as the dose inducing slight toxic effects relative to the duration of the study period (for example, clear clinical signs such as abnormal behaviour or reactions, minor body weight depression or target tissue cytotoxicity), but not death or evidence of pain, suffering or distress necessitating euthanasia. For a non-toxic test chemical, with an administration period of 14 days or more, the maximum (limit) dose is 1000 mg/kg bodyweight/day. For administration periods of less than 14 days the maximum (limit) dose is 2000 mg/kg bodyweight/day. For certain types of test chemicals (e.g. human pharmaceuticals) covered by specific regulations these limits may vary.

37. Substances that exhibit saturation of toxicokinetic properties, or induce detoxification processes that may lead to a decrease in exposure after long-term administration, may be exceptions to the dose-setting criteria and should be evaluated on a case-by-case basis.

38. For both acute and sub-acute versions of the comet assay, in addition to the maximum dose (MTD, maximum feasible dose, maximum exposure or limit dose) a descending sequence of at least two additional appropriately spaced dose levels (preferably separated by less than $\sqrt{10}$) should be selected for each sampling time to demonstrate dose-related responses. However, the dose levels used should also preferably cover a range from the maximum to one producing little or no toxicity. When target tissue toxicity is observed at all dose levels tested, further study at non-toxic doses is advisable (see paragraphs 54-55). Studies intending to more fully investigate the shape of the dose-response curve may require

© OECD, (2016)

additional dose group(s).

Administration of Doses

39. The anticipated route of human exposure should be considered when designing an assay. Therefore, routes of exposure such as dietary, drinking water, topical, subcutaneous, intravenous, oral (by gavage), inhalation, intratracheal, or implantation may be chosen as justified. In any case the route should be chosen to ensure adequate exposure of the target tissue(s). Intraperitoneal injection is generally not recommended since it is not a typical relevant route of human exposure, and should only be used with specific justification (e.g. some positive control substances, for investigative purposes, or for some drugs that are administered by the intraperitoneal route). The maximum volume of liquid that can be administered by gavage or injection at one time depends on the size of the test animal. The volume should not exceed 1 mL/100 g body weight, except in the case of aqueous solutions where 2 mL/100g body weight may be used. The use of volumes greater than this (if permitted by animal welfare legislation) should be justified. Wherever possible different dose levels should be achieved by adjusting the concentration of the dosing formulation to ensure a constant volume in relation to body weight at all dose levels.

Sampling Time

40. The sampling time is a critical variable because it is determined by the period needed for the test chemicals to reach maximum concentration in the target tissue and for DNA strand breaks to be induced but before those breaks are removed, repaired or lead to cell death. The persistence of some of the lesions that lead to the DNA strand breaks detected by the comet assay may be very short, at least for some substances tested *in vitro* (52) (53). Accordingly, if such transient DNA lesions are suspected, measures should be taken to mitigate their loss by ensuring that tissues are sampled sufficiently early, possibly earlier than the default times given below. The optimum sampling time(s) may be substance- or route-specific resulting in, for example, rapid tissue exposure with intravenous administration or inhalation exposure. Accordingly, where available, sampling times should be determined from kinetic data (e.g., the time (Tmax) at which the peak plasma or tissue concentration (Cmax) is achieved, or at the steady state for multiple administrations). In the absence of kinetic data a suitable compromise for the measurement of genotoxicity is to sample at 2-6 h after the last treatment for two or more treatments, or at both 2-6 and 16-26 h after a single administration, although care should be taken to necropsy all animals at the same time after the last (or only) dose. Information on the appearance of toxic effects in target organs (if available) may also be used to select appropriate sampling times.

Observations

41. General clinical observations related to the health of the animals should be made and recorded at least once a day preferably at the same time(s) each day and considering the peak period of anticipated effects after dosing (54). At least twice daily, all animals should be observed for morbidity and mortality. For longer duration studies, all animals should be weighed at least once a week, and at completion of the test period. Food consumption should be measured at each change of food and at least weekly. If the test chemical is administered via the drinking water, water consumption should be measured at each change of water and at least weekly. Animals exhibiting non-lethal indicators of excessive toxicity should be euthanized prior to completion of the test period, and are generally not used for comet analysis.

Tissue Collection

© OECD, (2016)

42. Since it is possible to study induction of DNA strand breaks (comets) in virtually any tissue, the rationale for selection of tissue(s) to be collected should be clearly defined and based upon the reason for conducting the study together with any existing ADME, genotoxicity, carcinogenicity or other toxicity data for the test substances under investigation. Important factors for consideration should include the route of administration (based on likely human exposure route(s)), the predicted tissue distribution and absorption, the role of metabolism and the possible mechanism of action of the test substances. The liver has been the tissue most frequently studied and for which there are the most data. Therefore, in the absence of any background information, and if no specific tissues of interest are identified, sampling the liver would be justified as this is a primary site of xenobiotic metabolism and is often highly exposed to both parent substance(s) and metabolite(s). In some cases examination of a site of direct contact (for example, for orally-administered substances the glandular stomach or duodenum/jejunum, or for inhaled substances the lungs) may be most relevant. Additional or alternative tissues should be selected based on the specific reasons for the test is being conducted but it may be useful to examine multiple tissues in the same animals providing the laboratory has demonstrated proficiency with those tissues and competency in handling multiple tissues at the same time.

Preparation of specimens

43. For the processes described in the following paragraphs (44-49) it is important that all solutions or stable suspensions should be used within their expiration date, or should be freshly prepared if needed. Also in the following paragraphs, the times taken to (i) remove each tissue after necropsy, (ii) process each tissue into cell/nuclei suspensions, and (iii) process the suspension and prepare the slides are all considered critical variables (see Definitions, Annex 1), and acceptable lengths of time for each of these steps should have been determined during establishment of the method and demonstration of proficiency.

44. Animals will be euthanised, consistent with effective animal welfare legislation and 3Rs principles, at the appropriate time(s) after the last treatment with a test chemical. Selected tissue(s) is removed, dissected, and a portion is collected for the comet assay, whilst at the same time a section from the same part of the tissue should be cut and placed in formaldehyde solution or appropriate fixative for possible histopathology analysis (see paragraph 55) according to standard methods (12). The tissue for the comet assay is placed into mincing buffer, rinsed sufficiently with cold mincing buffer to remove residual blood, and stored in ice-cold mincing buffer until processed. In situ perfusion may also be performed, e.g. for liver, kidney.

45. Many published methods exist for cell/nuclei isolation. These include mincing of tissues such as liver and kidney, scraping mucosal surfaces in the case of the gastro-intestinal tract, homogenization and enzymic digestion. The JaCVAM validation trial only studied isolated cells, and therefore in terms of establishing the method and being able to refer to the JaCVAM trial data for demonstration of proficiency, isolated cells are preferred. However, it has been shown that there was no essential difference in the assay result whether isolated cells or nuclei were used (8). Also different methods to isolate cells/nuclei (e.g., homogenizing, mincing, enzymic digestion and mesh filtration) gave comparable results (55). Consequently either isolated cells or isolated nuclei can be used. A laboratory should thoroughly evaluate and validate tissue-specific methods of single cell/nuclei isolation. As discussed in paragraph 40, the persistence of some of the lesions that lead to the DNA strand breaks detected by the comet assay may be very short (52) (53). Therefore, whatever method is used to prepare the single cell/nuclei suspensions, it is important that tissues are processed as soon as possible after the animals have been euthanised and placed in conditions that reduce the removal of lesions (e.g. by maintaining the tissue at low temperature). The cell suspensions should be kept ice-cold until ready for use, so that minimal inter-sample variation and appropriate positive and negative control responses can be demonstrated.

© OECD, (2016)

© OECD, (2016)

Preparation of slides

46. Slide preparation should be done as soon as possible (ideally within one hour) after single cell/nuclei preparation, but the temperature and time between animal death and slide preparation should be tightly controlled and validated under the laboratory's conditions. The volume of the cell suspension added to low melting point agarose (usually 0.5-1.0%) to make the slides should not reduce the percentage of low melting point agarose to less than 0.45%. The optimum cell density will be determined by the image analysis system used for scoring comets.

Lysis

47. Lysis conditions are also a critical variable and may interfere with the strand breaks resulting from specific types of DNA modifications (certain DNA alkylations and base adducts). It is therefore recommended that the lysis conditions be kept as constant as possible for all slides within an experiment. Once prepared, the slides should be immersed in chilled lysing solution for at least one hour (or overnight) at around 2-8^{o}C under subdued lighting conditions e.g. yellow light (or light proof) that avoid exposure to white light that may contain UV components. After this incubation period, the slides should be rinsed to remove residual detergent and salts prior to the alkali unwinding step. This can be done using purified water, neutralization buffer or phosphate buffer. Electrophoresis buffer can also be used. This would maintain the alkaline conditions in the electrophoresis chamber.

Unwinding and electrophoresis

48. Slides should be randomly placed onto the platform of a submarine-type electrophoresis unit containing sufficient electrophoresis solution such that the surfaces of the slides are completely covered (the depth of covering should also be consistent from run to run). In other type of comet assay electrophoresis units i.e. with active cooling, circulation and high capacity power supply a higher solution covering will result in higher electric current while the voltage is kept constant. A balanced design should be used to place slides in the electrophoresis tank to mitigate the effects of any trends or edge effect within the tank and to minimize batch-to-batch variability, i.e., in each electrophoresis run, there should be the same number of slides from each animal in the study and samples from the different dosage groups, negative and positive controls, should be included. The slides should be left for at least 20 minutes for the DNA to unwind, and then subjected to electrophoresis under controlled conditions that will maximize the sensitivity and dynamic range of the assay (i.e. lead to acceptable levels of % tail DNA for negative and positive controls that maximize sensitivity). The level of DNA migration is linearly associated with the duration of electrophoresis, and also with the potential (V/cm). Based on the JaCVAM trial this could be 0.7 V/cm for at least 20 minutes. The duration of electrophoresis is considered a critical variable and the electrophoresis time should be set to optimize the dynamic range. Longer electrophoresis times (e.g. 30 or 40 minutes to maximize sensitivity) usually lead to stronger positive responses with known mutagens. However longer electrophoresis times may also lead to excessive migration in control samples. In each experiment the voltage should be kept constant, and the variability in the other parameters should be within a narrow and specified range, for example in the JaCVAM trial 0.7 V/cm delivered a starting current of 300 mA. The depth of buffer should be adjusted to achieve the required conditions and maintained throughout the experiment. The current at the start and end of the electrophoresis period should be recorded. The optimum conditions should therefore be determined during the initial demonstration of proficiency in the laboratory concerned with each tissue studied. The temperature of the electrophoresis solution through unwinding and electrophoresis should be maintained at a low temperature, usually 2-10^{o}C (10). The temperature of the electrophoresis solution at the start of unwinding, the start of electrophoresis, and the end of electrophoresis should be recorded.

© OECD, (2016)

49. After completion of electrophoresis, the slides should be immersed/rinsed in the neutralization buffer for at least 5 minutes. Gels can be stained and scored "fresh" (e.g. within 1-2 days) or can be dehydrated for later scoring (e.g. within 1-2 weeks after staining) (56). However, the conditions should be validated during the demonstration of proficiency and historical data should be obtained and retained separately for each of these conditions. In case of the latter, slides should be dehydrated by immersion into absolute ethanol for at least 5 minutes, allowed to air dry, and then stored, either at room temperature or in a container in a refrigerator until scored.

Methods of Measurement

50. Comets should be scored quantitatively using an automated or semi-automated image-analysis system. The slides will be stained with an appropriate fluorescent stain e.g., SYBR Gold, Green I, propidium iodide or ethidium bromide and measured at a suitable magnification (e.g., 200x) on a microscope equipped with epi-fluorescence and appropriate detectors or a digital (e.g. CCD) camera.

51. Cells may be classified into three categories as described in the atlas of comet images (57), namely scorable, non-scorable and "hedgehog" (see paragraph 56 for further discussion). Only scorable cells (clearly defined head and tail with no interference with neighbouring cells) should be scored for % tail DNA to avoid artefacts. There is no need to report the frequency of non-scorable cells. The frequency of hedgehogs should be determined based on the visual scoring (since the absence of a clearly-defined head will mean they are not readily detected by image analysis) of at least 150 cells per sample (see paragraph 56 for further discussion) and separately documented.

52. All slides for analysis, including those of positive and negative controls, should be independently coded and scored "blinded" so the scorer is unaware of the treatment condition. For each sample (per tissue per animal), at least 150 cells (excluding hedgehogs – see paragraph 56) should be analysed. Scoring 150 cells per animal in at least 5 animals per dose (less in the concurrent positive control – see paragraph 29) provides adequate statistical power according to the analysis of Smith et al, 2008 (5). If slides are used, this could be from 2 or 3 slides scored per sample when five animals per group are used. Several areas of the slide should be observed at a density that ensures there is no overlapping of tails. Scoring at the edge of slides should be avoided.

53. DNA strand breaks in the comet assay can be measured by independent endpoints such as % tail DNA, tail length and tail moment. All three measurements can be made if the appropriate image software analyser system is used. However, the % tail DNA (also known as % tail intensity) is recommended for the evaluation and interpretation of results (12) (40) (41) (42), and is determined by the DNA fragment intensity in the tail expressed as a percentage of the cell's total intensity (13).

Tissue damage and cytotoxicity

54. Positive findings in the comet assay may not be solely due to genotoxicity, target tissue toxicity may also result in increases in DNA migration (12) (41). Conversely, low or moderate cytotoxicity is often seen with known genotoxins (12), showing that it is not possible to distinguish DNA migration induced by genotoxicity versus that induced by cytotoxicity in the comet assay alone. However, where increases in DNA migration are observed, it is recommended that an examination of one or more indicators of cytotoxicity is performed as this can aid in interpretation of the findings. Increases in DNA migration in the presence of clear evidence of cytotoxicity should be interpreted with caution.

© OECD, (2016)

55. Many measures of cytotoxicty have been proposed and of these histopathological changes are considered a relevant measure of tissue toxicity. Observations such as inflammation, cell infiltration, apoptotic or necrotic changes have been associated with increases in DNA migration, however, as demonstrated by the JaCVAM validation trial (12) no definitive list of histopathological changes that are always associated with increased DNA migration is available. Changes in clinical chemistry measures (e.g. AST, ALT), can also provide useful information on tissue damage and additional indicators such as caspase activation, TUNEL stain, Annexin V stain, etc. may also be considered. However, there are limited published data where the latter have been used for *in vivo* studies and some may be less reliable than others.

56. Hedgehogs (or clouds, ghost cells) are cells that exhibit a microscopic image consisting of a small or non-existent head, and large diffuse tails and are considered to be heavily damaged cells, although the etiology of the hedgehogs is uncertain (see Annex 3). Due to their appearance, % tail DNA measurements by image analysis are unreliable and therefore hedgehogs should be evaluated separately. The occurrence of hedgehogs should be noted and reported and any relevant increase thought to be due to the test chemical should be investigated and interpreted with care. Knowledge of the potential mode of action of the test substances may help with such considerations.

DATA AND REPORTING

Treatment of Results

57. The animal is the experimental unit and therefore both individual animal data and summarized results should be presented in tabular form. Due to the hierarchical nature of the data it is recommended that the median %tail DNA for each slide is determined and the mean of the median values is calculated for each animal (12). The mean of the individual animal means is then determined to give a group mean. All of these values should be included in the report. Alternative approaches (see paragraph 53) may be used if scientifically and statistically justified. Statistical analysis can be done using a variety of approaches (58) (59) (60) (61). When selecting the statistical methods to be used, the need for transformation (e.g. log or square root) of the data and/or addition of a small number (e.g. 0.001) to all (even non-zero) values to mitigate the effects of zero cell values, should be considered as discussed in the above references. Details of analysis of treatment/sex interactions when both sexes are used, and subsequent analysis of data where either differences or no differences are found is given in Annex 2. Data on toxicity and clinical signs should also be reported.

Acceptability Criteria

58. Acceptance of a test is based on the following criteria:

a. The concurrent negative control is considered acceptable for addition to the laboratory historical negative control database as described in paragraph 16

b. Concurrent positive controls (see paragraph 29) should induce responses that are compatible with those generated in the historical positive control database and produce a statistically significant increase compared with the concurrent negative control.

c. Adequate numbers of cells and doses have been analysed (paragraphs 50 and 36-38).

d. The criteria for the selection of highest dose are consistent with those described in paragraph 36.

© OECD, (2016)

Evaluation and Interpretation of Results

59. Providing that all acceptability criteria are fulfilled, a test chemical is considered to be clearly positive if:

a) at least one of the test doses exhibits a statistically significant increase compared with the concurrent negative control,

b) the increase is dose-related when evaluated with an appropriate trend test,

c) any of the results are outside the distribution of the historical negative control data for a given species, vehicle, route, tissue, and number of administrations.

When all of these criteria are met, the test chemical is then considered able to induce DNA strand breakage in the tissues studied in this test system. If only one or two of these criteria are satisfied, see paragraph 62.

60. Providing that all acceptability criteria are fulfilled, a test chemical is considered clearly negative if:

a) none of the test concentrations exhibits a statistically significant increase compared with the concurrent negative control,

b) there is no concentration-related increase when evaluated with an appropriate trend test

c) all results are inside the distribution of the historical negative control data for a given species, vehicle, route, tissue, and number of administrations

d) direct or indirect evidence supportive of exposure of, or toxicity to, the target tissue(s) has been demonstrated.

The test chemical is then considered unable to induce DNA strand breakage in the tissues studied in this test system.

61. There is no requirement for verification of a clearly positive or negative response.

62. In case the response is neither clearly negative nor clearly positive (i.e. not all the criteria listed in paragraphs 59 or 60 are met) and in order to assist in establishing the biological relevance of a result, the data should be evaluated by expert judgement and/or further investigations conducted, if scientifically justified. Scoring additional cells (where appropriate) or performing a repeat experiment possibly using optimised experimental conditions (e.g. dose spacing, other routes of administration, other sampling times or other tissues) could be useful.

63. In rare cases, even after further investigations, the data set will preclude making a conclusion of positive or negative results, and will therefore be concluded as equivocal.

64. To assess the biological relevance of a positive or equivocal result, information on cytotoxicity at the target tissue is required (see paragraphs 54-55). Where positive or equivocal findings are observed

© OECD, (2016)

solely in the presence of clear evidence of cytotoxicity, the study would be concluded as equivocal for genotoxicity unless there is enough information that is supportive of a definitive conclusion. In cases of a negative study outcome where there are signs of toxicity at all doses tested, further study at non-toxic doses may be advisable.

Test Report

65. The test report should include the following information:

Test chemical:

- source, lot number if available;
- stability of the test chemical, limit date for use, or date for re-analysis if known;

Mono-constituent substance:

- physical appearance, water solubility, and additional relevant physicochemical properties;
- chemical identification, such as IUPAC or CAS name, CAS number, SMILES or InChI code, structural formula, purity, chemical identity of impurities as appropriate and practically feasible, etc.

Multi-constituent substance, UVBCs and mixtures:

- characterised as far as possible by chemical identity (see above), quantitative occurrence and relevant physicochemical properties of the constituents.

Solvent/vehicle:

- justification for choice of solvent/vehicle;
- solubility and stability of the test chemical in the solvent/vehicle, if known;
- preparation of dose formulations;
- analytical determinations on formulations (e.g., stability, homogeneity, nominal concentrations);

Test animals:

- species/strain used and scientific and ethical justifications for the choice;
- number, age and sex of animals;
- source, housing conditions, diet, enrichment, etc.;
- individual weight of the animals at the start and at the end of the test, including body weight range, mean and standard deviation for each group;

Test conditions:

- positive and negative (vehicle/solvent) control data;
- results from the range-finding study (if conducted);
- rationale for dose level selection;
- details of test chemical preparation;
- details of the administration of the test chemical;
- rationale for route of administration;
- site of injection (for subcutaneous or intravenous studies);
- methods for sample preparation, where available, histopathological analyses, especially for a substance giving a positive comet response;

© OECD, (2016)

- rationale for tissue selection;
- methods for verifying that the test chemical reached the target tissue, or general circulation, if negative results are obtained;
- actual dose (mg/kg body weight/day) calculated from diet/drinking water test chemical concentration (ppm) and consumption, if applicable;
- details of diet and water quality;
- detailed description of treatment and sampling schedules and justifications for the choices (e.g. toxicokinetic data, where available);
- method of pain relief, analgesia;
- method of euthanasia;
- procedures for isolating and preserving tissues;
- methods for preparing single cell/nucleus suspension;
- source and lot numbers of all reagents (where possible);
- methods for evaluating cytotoxicity;
- electrophoresis conditions;
- staining techniques used; and
- methods for scoring and measuring comets;

Results:

- General clinical observations, if any, prior to and throughout the test period for each animal;
- evidence of cytotoxicity if performed;
- for studies longer than one week: Individual body weights during the study, including body weight range, mean and standard deviation for each group; food consumption;
- dose-response relationship, where evident;
- for each tissue/animal, the % tail DNA (or other measures, if chosen) and median values per slide, mean values per animal and mean values per group;
- concurrent and historical negative control data with ranges, means/medians and standard deviations for each tissue evaluated;
- concurrent and historical positive control data;
- for tissues other than liver, a dose-response curve using the positive control. This can be from data collected during the demonstration of proficiency (see paragraphs 16-17) and should be accompanied by a justification, with citations to current literature, for the appropriateness of the magnitude and scatter of the responses to the controls in that tissue;
- statistical analyses and methods applied; and
- criteria for considering a response as positive, negative or equivocal;
- frequency of hedgehogs in each group and per animal;

Discussion of the results

Conclusion

References

© OECD, (2016)

LITERATURE

(1) Kirkland, D., G. Speit (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*, Vol. 654/2, pp. 114-32.

(2) Brendler-Schwaab, S. et al. (2005), The *in vivo* Comet assay: use and status in genotoxicity testing, *Mutagenesis*, Vol. 20/4, pp. 245-54.

(3) Burlinson, B. et al. (2007), Fourth International Workgroup on Genotoxicity Testing: result of the *in vivo* Comet assay workgroup, *Mutation Research*, Vol. 627/1, pp. 31-5.

(4) Burlinson, B. (2012), The *in vitro* and *in vivo* Comet assays, *Methods in Molecular Biology*, Vol. 817, pp. 143-63.

(5) Smith, C.C, et al. (2008), Recommendations for design of the rat Comet assay, *Mutagenesis*, Vol. 23/3, pp. 233-40.

(6) Hartmann, A. et al. (2003), Recommendations for conducting the *in vivo* alkaline Comet assay, *Mutagenesis*, Vol. 18/1, pp. 45-51.

(7) McKelvey-Martin, V.J. et al (1993), The single cell gel electrophoresis assay (Comet assay): a European review, *Mutation Research*, Vol. 288/1, pp. 47-63.

(8) Tice, R.R. et al. (2000), Single cell gel/Comet assay: guidelines for *in vitro* and *in vivo* genetic toxicology testing, *Environmental and Molecular Mutagenesis*, Vol. 35/3, pp. 206-21.

(9) Singh, N.P. et al. (1988), A simple technique for quantitation of low levels of DNA damage in individual cells, *Experimental Cell Research*, Vol. 175/1, pp. 184-91.

(10) Rothfuss, A. et al. (2010), Collaborative study on fifteen compounds in the rat-liver Comet assay integrated into 2- and 4-week repeat-dose studies, *Mutation Research*, Vol., 702/1, pp. 40-69.

(11) OECD (2016). Overview of the set of OECD Genetic Toxicology Test Guidelines and updates performed in 2014-2015. ENV Publications. Series on Testing and Assessment, No. 234, OECD, Paris.

(12) OECD (2014), *Reports of the JaCVAM initiative international pre-validation and validation studies of the in vivo rodent alkaline comet assay for the detection of genotoxic carcinogens*, Series on Testing and Assessment, Nos. 195 and 196, OECD Publishing, Paris.

(13) Olive, P.L., J.P. Banath, R.E. Durand (1990), Heterogeneity in radiation-induced DNA damage and repair in tumor and normal cells using the "Comet" assay, *Radiation Research*, Vol. 122/1, pp. 86-94.

(14) Tice, R.R., G.H. Strauss (1995), The single cell gel electrophoresis/Comet assay: a potential tool for detecting radiation-induced DNA damage in humans, *Stem Cells*, Vol. 13/1, pp. 207-14.

(15) Collins, A.R (2004), The Comet assay for DNA damage and repair: principles, applications, and limitations, *Molecular Biotechnology*, Vol. 26/3, pp. 249-61.

(16) Rothfuss, A. et al. (2011), Improvement of *in vivo* genotoxicity assessment: combination of acute tests and integration into standard toxicity testing, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 723/2, pp. 108-20.

(17) Kushwaha, S. et al. (2010), Evaluation of multi-organ DNA damage by Comet assay from 28 days

© OECD, (2016)

repeated dose oral toxicity test in mice: A practical approach for test integration in regulatory toxicity testing, *Regulatory Toxicology and Pharmacology*, Vol. 58/1, pp. 145–54.

(18) Vasquez, M.Z. (2010), Combining the *in vivo* Comet and micronucleus assays: a practical approach to genotoxicity testing and data interpretation, *Mutagenesis*, Vol. 25/2, pp. 187-99.

(19) Bowen, D.E. (2011), Evaluation of a multi-endpoint assay in rats, combining the bone-marrow micronucleus test, the Comet assay and the flow-cytometric peripheral blood micronucleus test, *Mutation Research*, Vol. 722/1, pp. 7-19.

(20) Recio, L. et al. (2010), Dose-response assessment of four genotoxic chemicals in a combined mouse and rat micronucleus (MN) and Comet assay protocol, *The Journal of Toxicological Science*, Vol. 35/2, pp. 149-62.

(21) O'Donovan, M., B. Burlinson (2013), Maximum dose levels for the rodent comet assay to examine damage at the site of contact or to the gastrointestinal tract, *Mutagenesis*, Vol. 28/6, pp. 621-3.

(22) Hartmann, A. (2004), Use of the alkaline *in vivo* Comet assay for mechanistic genotoxicity investigations, *Mutagenesis*, Vol. 19/1, pp. 51-9.

(23) Nesslany, F. (2007), *In vivo* Comet assay on isolated kidney cells to distinguish genotoxic carcinogens from epigenetic carcinogens or cytotoxic compounds, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 630/1, pp. 28-41.

(24) Brendler-Schwaab, S.Y., B.A. Herbold (1997), A new method for the enrichment of single renal proximal tubular cells and their first use in the Comet assay, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*,Vol. 393/1-2, pp. 175-8.

(25) Toyoizumi, T. et al. (2011), Use of the *in vivo* skin Comet assay to evaluate the DNA-damaging potential of chemicals applied to the skin, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 726/2, pp. 175-80.

(26) Struwe, M. et al. (2008), Detection of photogenotoxicity in skin and eye in rat with the photo Comet assay, *Photochemical and Photobiological Sciences*, Vol. 7/2, pp. 240-9.

(27) Wada, K. et al. (2012), A comparison of cell-collecting methods for the Comet assay in urinary bladders of rats, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 742/1-2, pp. 26-30.

(28) Wang, A. et al. (2007), Measurement of DNA damage in rat urinary bladder transitional cells: improved selective harvest of transitional cells and detailed Comet assay protocols, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 634/ 1-2, pp. 51-9.

(29) Burlinson, B. et al. (2007), *In Vivo* Comet Assay Workgroup, part of the Fourth International Workgroup on Genotoxicity Testing. Fourth International Workgroup on Genotoxicity testing: results of the in vivo Comet assay workgroup, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 627/1, pp. 31-5.

(30) Jackson, P. et al. (2012), Pulmonary exposure to carbon black by inhalation or instillation in pregnant mice: effects on liver DNA strand breaks in dams and offspring, *Nanotoxicology*, Vol. 6/5, pp. 486-500.

(31) Sasaki, Y.F. et al. (2000), The comet assay with multiple mouse organs: comparison of Comet assay results and carcinogenicity with 208 chemicals selected from the IARC monographs and U.S. NTP Carcinogenicity Database, *Critical Reviews in Toxicology*, Vol. 30/6, pp. 629-799.

(32) Sekihashi, K. et al. (2002), Comparative investigations of multiple organs of mice and rats in the

© OECD, (2016)

Comet assay, *Mutation Research*, Vol. 517/1-2, pp. 53-74.

(33) Speit, G, M. Vasquez, A. Hartmann (2009), The comet assay as an indicator test for germ cell genotoxicity, *Mutation Research*, Vol. 681/1, pp. 3-12.

(34) Zheng, H., P.L. Olive (1997), Influence of oxygen on radiation-induced DNA damage in testicular cells of C3H mice, *International Journal of Radiation Biology*, Vol. 71/3, pp. 275-282.

(35) Cordelli, E. et al. (2003), Evaluation of DNA damage in different stages of mouse spermatogenesis after testicular X irradiation, *Journal of Radiation Research*, Vol. 160/4, pp. 443-451.

(36) Merk, O., G. Speit (1999), Detection of crosslinks with the Comet assay in relationship to genotoxicity and cytotoxicity, *Environmental and Molecular Mutagenesis*, Vol. 33/2, pp. 167-72.

(37) Pfuhler, S., H.U. Wolf (1996), Detection of DNA-crosslinking agents with the alkaline Comet assay, *Environmental and Molecular Mutagenesis*, Vol. 27/3, pp. 196-201.

(38) Wu, J.H., N.J. Jones (2012), Assessment of DNA interstrand crosslinks using the modified alkaline Comet assay, *Methods in Molecular Biology*, Vol. 817, pp. 165-81.

(39) Spanswick, V.J., J.M. Hartley, J.A. Hartley (2010), Measurement of DNA interstrand crosslinking in individual cells using the Single Cell Gel Electrophoresis (Comet) assay, *Methods in Molecular Biology*, Vol. 613, pp. 267-282.

(40) Kumaravel, T.S., A.N. Jha (2006), Reliable Comet assay measurements for detecting DNA damage induced by ionizing radiation and chemicals, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 605(1-2), pp. 7-16.

(41) Burlinson, B.et al. (2007), Fourth International Workgroup on Genotoxicity Testing: result of the *in vivo* Comet assay workgroup, *Mutation Research*, Vol.627/1, pp. 31-5.

(42) Kumaravel, T.S. et al. (2009), Comet Assay measurements: a perspective, *Cell Biology and Toxicology*, Vol. 25/1, pp. 53-64.

(43) Ersson, C., L. Möller (2011), The effects on DNA migration of altering parameters in the Comet assay protocol such as agarose density, electrophoresis conditions and durations of the enzyme or the alkaline treatments, *Mutagenesis*, Vol. 26/6, pp. 689-95.

(44) Møller, P. et al. (2010), Assessment and reduction of Comet assay variation in relation to DNA damage: studies from the European Comet Assay Validation Group, *Mutagenesis*, Vol. 25/2, pp. 109-11.

(45) Forchhammer, L. et al. (2010), Variation in the measurement of DNA damage by Comet assay measured by the ECVAG inter-laboratory validation trial, *Mutagenesis*, Vol. 25/2, pp. 113-23.

(46) Azqueta, A. et al. (2011), Towards a more reliable comet assay: Optimising agarose concentration, unwinding time and electrophoresis conditions, *Mutation Research*, Vol. 724/1-2, pp. 41-45.

(47) Hayashi, M. et al. (2011), Compilation and use of genetic toxicity historical control data, *Mutation Research*, Vol. 723/2, pp. 87-90.

(48) Ryan, T. P. (2000), *Statistical Methods for Quality Improvement,* John Wiley and Sons, New York 2nd ed.

(49) Appendix A of the European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (ETS No. 123)

(50) OECD (2009), *Test No. 412: Subacute Inhalation Toxicity: 28-Day Study*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris.

© OECD, (2016)

(51) OECD (2009), *Test No. 413: Subchronic Inhalation Toxicity: 90-day Study*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris.

(52) Blakey, D.H., G.R. Douglas (1984), Transient DNA lesions induced by benzo[a]pyrene in Chinese hamster ovary cells, *Mutation Research*, Vol. 140/2-3, pp. 141-45.

(53) Blakey, D.H., G.R. Douglas (1990), The role of excision repair in the removal of transient benzo[a]pyrene-induced DNA lesions in Chinese hamster ovary cells, *Mutation Research*, Vol. 236/1, pp. 35-41.

(54) OECD (2002), "Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation", OECD Environment, Health and Safety Publications (EHS), Series on Testing and Assessment, No. 19, OECD Publishing, Paris.

(55) Nakajima, M. (2012), Tissue sample preparation for *in vivo* rodent alkaline Comet assay, *Genes and Environment*, Vol. 34/1, pp. 50-4.

(56) Hartmann, A. et al. (2003), Recommendations for conducting the in vivo alkaline Comet assay, *Mutagenesis*, Vol.18/1, pp.45–51.

(57) Atlas of Comet Assay Images, Scientist Press Co., Ltd., Tokyo, Japan.

(58) Lovell, D.P., G. Thomas, R. Dubow (1999), Issues related to the experimental design and subsequent statistical analysis of *in vivo* and *in vitro* Comet studies, *Teratogenesis Carcinogenesis Mutagenesis*, Vol. 19/2, pp. 109-19.

(59) Wiklund, S.J., E. Agurell (2003), Aspects of design and statistical analysis in the Comet assay, *Mutagenesis*, Vol. 18/2, pp. 167-75.

(60) Bright, J. et al. (2011), Recommendations on the statistical analysis of the Comet assay, *Pharmaceutical Statistics*, Vol. 10/6, pp. 485-93.

(61) Lovell, D.P., T. Omori (2008), Statistical issues in the use of the Comet assay, *Mutagenesis*, Vol. 23/3, pp. 171-82.

© OECD, (2016)

ANNEX 1

Definitions

Alkaline single cell gel electrophoresis: Sensitive technique for the detection of primary DNA damage at the level of individual cell/nucleus

Comet: The shape that nucleoids adopt after submitted to one electrophoretic field, due to its similarity to comets: the head is the nucleus and the tail is constituted by the DNA migrating out of the nucleus in the electric field.

A critical variable/parameter: This is a protocol variable for which a small change can have a large impact on the conclusion of the assay. Critical variables can be tissue-specific. Critical variables should not be altered, especially within a test, without consideration of how the alteration will alter an assay response, for example as indicated by the magnitude and variability in positive and negative controls. The test report should list alterations of critical variables made during the test or compared to the standard protocol for the laboratory and provide a justification for each alteration.

Tail intensity or % tail DNA: This corresponds to the intensity of the comet tail relative to the total intensity (head plus tail). It reflects the amount of DNA breakage, expressed as a percentage.

© OECD, (2016)

ANNEX 2

The Factorial Design for Identifying Sex Differences in the *in vivo* Comet Assay

The factorial design and its analysis

1. In this design, a minimum of 5 males and 5 females are tested at each concentration level resulting in a design using a minimum of 40 animals (20 males and 20 females, plus relevant positive controls.)

2. The design, which is one of the simpler factorial designs, is equivalent to a two-way analysis of variance with sex and concentration level as the main effects. The data can be analysed using many standard statistical software packages such as SPSS, SAS, STATA, Genstat as well as using R.

3. The analysis partitions the variability in the dataset into that between the sexes, between the concentrations and that related to the interaction between the sexes and the concentrations. Each of the terms is tested against an estimate of the variability between the replicate animals within the groups of animals of the same sex given the same concentration. Full details of the underlying methodology are available in many standard statistical textbooks (see references) and in the 'help' facilities provided with statistical packages.

4. The analysis proceeds by inspecting the sex x concentration interaction term in the ANOVA table[1]. In the absence of a significant interaction term the combined values across sexes or across concentration levels provide valid statistical tests between the levels based upon the pooled within group variability term of the ANOVA.

5. The analysis continues by partitioning the estimate of the between concentrations variability into contrasts which provide for a test for linear and quadratic contrasts of the responses across the concentration levels. When there is a significant sex x concentration interaction this term can also be partitioned into linear x sex and quadratic x sex interaction contrasts. These terms provide tests of whether the concentration responses are parallel for the two sexes or whether there is a differential response between the two sexes.

6. The estimate of the pooled within group variability can be used to provide pair-wise tests of the difference between means. These comparisons could be made between the means for the two sexes and between the means for the different concentration level such as for comparisons with the negative control levels. In those cases where there is a significant interaction comparisons can be made between the means of different concentrations within a sex or between the means of the sexes at the same concentration.

[1] Statisticians who take a modelling approach such as using General Linear Models (GLMs) may approach the analysis in a different but comparable way but will not necessarily derive the traditional ANOVA table which dates back to algorithmic approaches to calculating the statistics developed in a pre-computer age.

© OECD, (2016)

References

There are many statistical textbooks which discuss the theory, design, methodology, analysis and interpretation of factorial designs ranging from the simplest two factor analyses to the more complex forms used in Design of Experiment methodology. The following is a non-exhaustive list. Some books provide worked examples of comparable designs, in some cases with code for running the analyses using various software packages.

Box, G.E.P, Hunter, W.G. and Hunter, J.S. (1978). Statistics for Experimenters. An Introduction to Design, Data Analysis, and Model Building. New York: John Wiley & Sons.

Box G.E.P. & Draper, N.R. (1987) Empirical model-building and response surfaces. John Wiley & Sons Inc.

Doncaster, C.P. & Davey, A.J.H. (2007) Analysis of Variance and Covariance: How to choose and Construct Models for the Life Sciences. Cambridge University Press.

Mead, R. (1990) The Design of Experiments. Statistical principles for practical application. Cambridge University Press.

Montgomery D.C. (1997) Design and Analysis of Experiments. John Wiley & Sons Inc.
Winer, B.J. (1971) Statistical Principles in Experimental Design. McGraw Hill.

Wu, C.F.J & Hamada, M.S. (2009) Experiments: Planning, Analysis and Optimization. John Wiley & Sons Inc.

© OECD, (2016)

ANNEX 3

Current Limitations of the Assay

Due to the current status of knowledge, several limitations are associated with the *in vivo* comet assay. It is expected that these limitations will be reduced or more narrowly defined as there is more experience with application of the assay to answer safety issues in a regulatory context.

1. Some types of DNA damage may be short-lived, i.e. may be repaired too quickly to be observed 24 hours or more after the last dose. There is no identifiable list of the types of short-lived damages, nor of the substances which are likely to cause this type of damage, nor is it known over what time period this type of damage can be detected. The optimum sampling time(s) may also be substance- or route-specific and sampling times should be determined from kinetic data (for example the time, T_{max}, at which the peak plasma or tissue concentration is achieved), when such data are available. Most of the validation studies supporting this guideline specified necropsy 2 or 3 hours following administration of the final dose. Most studies in the published literature describe administration of the final dose between 2 and 6 hours prior to sacrifice. Therefore these experiences were used as the basis for the recommendation in the test guideline that, in the absence of data indicating otherwise, the final dose should be administered at a specified time point between 2 and 6 hours prior to necropsy.

2. There are no identifiable study data that examine the sensitivity of the test for the detection of short-lived DNA damage following administration in food or drinking water compared to administration by gavage. DNA damage has been detected following administration in feed and drinking water, but there are relatively few such reports compared to the much greater experience with gavage and i.p. administration. Thus the sensitivity of the assay may be reduced for substances which induce short-lived damage administered through feed or drinking water.

3. No inter-laboratory studies have been conducted in tissues other than liver and stomach, therefore no recommendation has been established for how to achieve a sensitive and reproducible response in tissues other than liver, such as expected positive and negative control ranges. For the liver, agreement on setting a lower limit to the negative control value also could not be reached.

4. Although there are several publications demonstrating the confounding effect of cytotoxicity *in vitro*, very little data have been published *in vivo* and therefore no single measure of cytotoxicity could be recommended. Histopathological changes such as inflammation, cell infiltration, apoptotic or necrotic changes have been associated with increases in DNA migration however, as demonstrated by the JaCVAM validation trial (OECD, 2014), these changes do not always result in positive comet findings and consequently no definitive list of histopathological changes that are always associated with increased DNA migration is available. Hedgehogs (or clouds, ghost cells) have previously been suggested as an indicator of cytotoxicity, however, the etiology of the hedgehogs is uncertain. Data exist which suggest that they can be caused by substance-related cytotoxicity, mechanical/enzyme-induced damage initiated during sample preparation (Guerard et al., 2014) and/or a more extreme effect of test chemical genotoxicity. Other data seem to show they are due to extensive, but perhaps repairable DNA damage (Lorenzo et al., 2013).

5. Tissues or cell nuclei have been successfully frozen for later analysis. This usually results in a measurable effect on the response to the vehicle and positive control (Recio at al., 2010; Recio at al., 2012; Jackson at al., 2013). If used, the laboratory should demonstrate competency in freezing

© OECD, (2016)

methodologies and confirm acceptable low ranges of % tail DNA in target tissues of vehicle treated animals, and that positive responses can still be detected. In the literature, the freezing of tissues has been described using different methods. However, currently there is no agreement on how to best freeze and thaw tissues, and how to assess whether a potentially altered response may affect the sensitivity of the test.

6. Recent work demonstrates that the list of critical variables is expected to continue to become shorter and the parameters for critical variables more precisely defined (Guerard et al., 2014).

References

Guerard, M., C. Marchand, U. Plappert-Helbig (2014), Influence of Experimental Conditions on Data Variability in the Liver Comet Assay, *Environmental and Molecular Mutagenesis*, Vol. 55/2, pp. 114-21.

Jackson, P. et al. (2013), Validation of use of frozen tissues in high-throughput comet assay with fully-automatic scoring, *Mutagenesis*, Vol. 28/6, pp. 699-707.

Lorenzo ,Y. et al. (2013), The comet assay, DNA damage, DNA repair and cytotoxicity: hedgehogs are not always dead, *Mutagenesis*, Vol. 28/4, pp. 427-32.

OECD (2014), *Reports of the JaCVAM initiative international pre-validation and validation studies of the in vivo rodent alkaline comet assay for the detection of genotoxic carcinogens*, Series on Testing and Assessment, Nos. 195 and 196, OECD Publishing, Paris.

Recio L, Hobbs C, Caspary W, Witt KL, (2010), Dose-response assessment of four genotoxic chemicals in a combined mouse and rat micronucleus (MN) and Comet assay protocol, *J. Toxicol. Sci.* 35:149-62.

Recio, L. et al. (2012), Comparison of Comet assay dose-response for ethyl methanesulfonate using freshly prepared versus cryopreserved tissues, *Environmental and Molecular Mutagenesis*, Vol. 53/2, pp. 101-13.

© OECD, (2016)