S2(R1) Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use

This draft guidance, when finalized, will represent the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

For questions regarding this draft document contact (CDER) David Jacobson-Kram 301-796-0175.

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED TRIPARTITE GUIDELINE

GUIDANCE ON GENOTOXICITY TESTING AND DATA INTERPRETATION FOR PHARMACEUTICALS INTENDED FOR HUMAN USE

DRAFT JANUARY 28TH 2008 VERSION 5.3

Recommended for Adoption
at Step X of the ICH Process
on
by the ICH Steering Committee

This Guideline has been developed by the appropriate ICH Expert Working Group and has been subject to consultation by the regulatory parties, in accordance with the ICH Process. At Step X of the Process the draft is recommended for adoption to the regulatory bodies of the European Union, Japan and USA.

ICH Harmonised Tripartite Guideline

Having reached *Step X* of the ICH Process at the ICH Steering Committee meeting on XXXXXX, this guideline is recommended for adoption to the three regulatory parties to ICH

Table of Contents

1. **INTRODUCTION**

1.1 Objectives of the Guideline

1.2 Background

1.3 Scope of the Guideline

1.4 General Principles

2. **THE STANDARD TEST BATTERY**

2.1 Rationale

2.2 Description of the two options for the standard battery

2.3 Modifications of the battery

2.3.1 Compounds from well characterized classes where genotoxicity is expected

2.3.2 Testing compounds that are toxic to bacteria

2.3.3 Compounds bearing structural alerts for genotoxic activity

2.3.4 Limitations to the use of *in vivo* tests

2.4 Detection of germ cell mutagens

3. **RECOMMENDATIONS FOR *IN VITRO* TESTS**

3.1 Test repetition and interpretation

3.2 Recommended protocol for bacterial mutation assays

3.2.1 Selection of top dose level

3.2.2 Study design/Test protocol

3.3 Recommended protocols for mammalian cell assays

3.3.1 Selection of top concentration

3.3.2 Study design/Test protocols

3.3.3 Positive Controls

4. **RECOMMENDATIONS FOR *IN VIVO* TESTS**

4.1 Tests for the detection of chromosome damage *in vivo*

4.2 Automated analysis of micronuclei

4.3 Other *in vivo* genotoxicity tests

4.4 Use of male/female rodents for *in vivo* genotoxicity tests

4.5 Use of multiple administrations in genotoxicity assays *in vivo* and integration

into toxicology studies
4.5.1 Sampling times
4.5.2 Number of animals analyzed
4.6 Route of administration
4.7 Dose Selection for *in vivo* assays
4.7.1 Short term studies
4.7.2 Multiple administration studies
4.7.3 Additional guidance on dose selection for multiple administration studies
4.8 Demonstration of target tissue exposure for negative *in vivo* test results
4.8.1 When an *in vitro* genotoxicity test is positive (or not done)
4.8.2 When *in vitro* genotoxicity tests are negative
4.9 Use of positive controls for *in vivo* studies
5. **GUIDANCE ON EVALUATION OF TEST RESULTS AND ON FOLLOW-UP TEST STRATEGIES**
5.1 Assessment of biological relevance
5.2 Evaluation of results obtained in *in vitro* tests
5.2.1 Evaluation of positive results obtained *in vitro* in a bacterial mutation assay
5.2.2 Evaluation of positive results obtained *in vitro* in a mammalian cell assay
5.2.3 Evaluation of *in vitro* negative results
5.3 Evaluation of results obtained from *in vivo* tests
5.4 Follow-up strategies for positive results
5.4.1 Follow-up to findings in *in vitro* in mammalian cell tests
5.4.1.1 Mechanistic/*in vivo* follow-up
5.4.1.2 Follow-up to an *in vitro* positive result that is dependent upon S-9
5.4.2 Follow-up to a positive *in vivo* micronucleus assay
5.5 Follow-up genotoxicity testing in relation to tumor findings in a carcinogenicity bioassay
6. **NOTES**
7. **GLOSSARY**
8. **REFERENCES**

DRAFT

GUIDANCE ON GENOTOXICITY TESTING AND DATA INTERPRETATION FOR PHARMACEUTICALS INTENDED FOR HUMAN USE

1. INTRODUCTION

1.1 Objectives of the Guideline

This guidance replaces and combines the ICH S2A and S2B guidelines. The purpose of the revision is to optimize the standard genetic toxicology battery for prediction of potential human risks, and to provide guidance on interpretation of results, with the ultimate goal of improving risk characterization for carcinogenic effects that have their basis in changes in the genetic material. The revised guidance describes internationally agreed upon standards for follow-up testing and interpretation of positive results *in vitro* and *in vivo* in the standard genetic toxicology battery, including assessment of non-relevant findings.

1.2 Background

Unless otherwise noted in this guidance, the recommendations from the latest OECD guidelines and the reports from the International Workshops on Genotoxicity Testing (IWGT) have been considered where relevant. The following notes for guidance should be applied in conjunction with other ICH guidances.

1.3 Scope of the Guideline

The primary focus of this guidance is testing of "small molecule" drug substances, and not biologics as defined in the ICH S6 guidance.

1.4 General Principles

Genotoxicity tests can be defined as *in vitro* and *in vivo* tests designed to detect compounds that induce genetic damage by various mechanisms. These tests enable hazard identification with respect to damage to DNA and its fixation. Fixation of damage to DNA in the form of gene mutations, larger scale chromosomal damage or recombination is generally considered to be essential for heritable effects and in the multi-step process of malignancy, a complex process in which genetic changes may play only a part. Numerical chromosome changes have also been associated with tumorigenesis and can indicate a potential for aneuploidy in germ cells. Compounds that are positive in tests that detect such kinds of damage have the potential to be human

carcinogens and/or mutagens. Because the relationship between exposure to particular chemicals and carcinogenesis is established for humans, whilst a similar relationship has been difficult to prove for heritable diseases, genotoxicity tests have been used mainly for the prediction of carcinogenicity. Nevertheless, because germ line mutations are clearly associated with human disease, the suspicion that a compound might induce heritable effects is considered to be just as serious as the suspicion that a compound might induce cancer. In addition, the outcome of genotoxicity tests can be valuable for the interpretation of carcinogenicity studies.

2. THE STANDARD TEST BATTERY FOR GENOTOXICITY

2.1 Rationale

Registration of pharmaceuticals requires a comprehensive assessment of their genotoxic potential. Extensive reviews have shown that many compounds that are mutagenic in the bacterial reverse mutation (Ames) test are rodent carcinogens. Addition of *in vitro* mammalian tests increases sensitivity and broadens the spectrum of genetic events detected, but also decreases the specificity of prediction; i.e., increases the incidence of positive results that do not correlate with rodent carcinogenicity. Nevertheless, a battery approach is still reasonable because no single test is capable of detecting all genotoxic mechanisms relevant in tumorigenesis.

The general features of a standard test battery are as follows:

i. Assessment of mutagenicity in a bacterial reverse mutation test. This test has been shown to detect relevant genetic changes and the majority of genotoxic rodent and human carcinogens.

ii. Genotoxicity should also be evaluated in mammalian cells *in vitro* and/or *in vivo*.

Several *in vitro* mammalian cell systems are widely used and can be considered sufficiently validated: The *in vitro* metaphase chromosome aberration assay, the *in vitro* micronucleus assay (note 1) and the mouse lymphoma L5178Y cell *tk* gene mutation assay. These three assays are currently considered equally appropriate and therefore interchangeable when used together with other genotoxicity tests in a standard battery for testing of pharmaceuticals, if the test protocols recommended in this guideline are used.

In vivo test(s) for genetic damage should usually be a part of the test battery to provide additional relevant factors (absorption, distribution metabolism, excretion) that can influence the genotoxic activity of a compound and permit the detection of some additional genotoxic agents (note 2). An *in vivo* test for chromosomal damage in rodent cells largely fulfills this need, either an analysis of micronuclei in erythrocytes in blood or bone marrow, or of chromosomal aberrations at metaphase in bone marrow cells (note 3). Lymphocytes cultured from treated animals can also be used for cytogenetic analysis, although experience with such analyses is less widespread.

In vitro and *in vivo* tests that measure chromosomal aberrations in metaphase cells can detect a wide spectrum of changes in chromosomal integrity. Breakage of chromatids or chromosomes can result in micronucleus formation if an acentric fragment is produced; therefore assays that detect either chromosomal aberrations or micronuclei are appropriate for detecting clastogens. Micronuclei can also result from lagging of one or more whole chromosome(s) at anaphase and thus micronucleus tests have the potential to detect some aneuploidy inducers. The mouse lymphoma cell mutation assay detects mutations in the *tk* gene that result from both gene mutations and changes in chromosome integrity. There is some evidence that the mouse lymphoma assay can also detect chromosome loss.

There are several additional *in vivo* assays that can be used in the battery or as follow-up tests to develop weight of evidence in assessing results of *in vitro* or *in vivo* assays (see below). Negative results in appropriate *in vivo* assays (usually two), with adequate justification for the endpoints measured, and demonstration of exposure (see section 4.8) is sufficient to demonstrate absence of genotoxic activity.

2.2 Description of the two options for the standard battery

The following two options for the standard battery are considered equally suitable:

<u>Option 1</u>

i. A test for gene mutation in bacteria.
ii. A cytogenetic test for chromosomal damage (the *in vitro* metaphase chromosome aberration test or *in vitro* micronucleus test), or an *in vitro* mouse lymphoma *tk* gene mutation assay.
iii. An *in vivo* test for genotoxicity, generally a test for chromosomal damage using

rodent hematopoietic cells, either for micronuclei or for chromosomal aberrations in metaphase cells.

Option 2

i. A test for gene mutation in bacteria.

ii. An *in vivo* assessment of genotoxicity with two tissues, usually an assay for micronuclei using rodent hematopoietic cells and a second *in vivo* assay.

Under both standard battery options, the *in vivo* genotoxicity assays can often be integrated into repeat-dose toxicity studies when the doses are sufficient (see section 4.7). Under Option 2, if dose/exposure is not appropriate, an acute *in vivo* study (incorporating two genotoxicity assays in one study where possible) should be performed to optimize dose selection based on exposure/toxicity (see sections 4.7.2 and 4.7.3), or Option 1, including an *in vitro* mammalian cell assay, should be followed.

For compounds that give negative results, the completion of either test battery, performed and evaluated in accordance with current recommendations, will usually provide sufficient assurance of the absence of genotoxic activity and no additional tests will be needed. Compounds that give positive results in the standard test battery may, depending on their therapeutic use, need to be tested more extensively (see Section 5).

The standard battery does not include a required independent test designed specifically to test for aneuploidy. However, information on numerical changes can be derived from the mammalian cell assays *in vitro* and from the micronucleus assays. Elements of the standard protocols that provide such information are elevations in the mitotic index, polyploidy induction and micronucleus evaluation. There is also experimental evidence that spindle poisons can be detected in the mouse lymphoma *tk* assay. The preferred *in vivo* cytogenetic test under Option 2 is the micronucleus assay, not a chromosome aberration assay, to include more direct capability for detection of chromosome loss (potential for aneuploidy).

There are several *in vivo* assays (note 4) that may be used as the second part of the *in vivo* assessment under option 2 (see section 4.3). The liver is typically the preferred tissue because of exposure and metabolizing capacity, but choice of *in vivo* tissue and assay should be based on factors such as any knowledge of the potential mechanism, of the metabolism *in vivo*, and of the exposed tissues thought to be relevant. The *in vivo* genotoxicity assays may be integrated into existing (repeat dose) toxicity

studies when the dose levels are justifiable (see section 4.7) and the protocols are compatible.

The suggested standard set of tests does not imply that other genotoxicity tests are generally considered inadequate or inappropriate. Additional tests can be used for further investigation of genotoxicity test results obtained in the standard battery (see sections 4.3 and 5). Alternative species, including non-rodents, can also be used if indicated, and if sufficiently validated.

Under extreme conditions in which one or more tests in the standard battery cannot be employed for technical reasons, alternative validated tests can serve as substitutes provided sufficient scientific justification is given to support the argument that a given standard battery test is not appropriate.

2.3 Modifications to the test battery

The following sections give situations where modification of the standard test battery may be advisable.

2.3.1 Compounds from well characterized classes

For compounds from well characterized classes where genotoxicity is expected, e.g., some quinolone antibiotics and some nucleoside analogues, the battery may be modified to characterize these appropriately in the tests/protocols known to respond to them. (See also note 8).

2.3.2 Testing compounds that are toxic to bacteria

In cases where compounds are highly toxic to bacteria (e.g., some antibiotics), the bacterial reverse mutation (Ames) test should still be carried out, because mutagenicity can occur at lower, less toxic concentrations. In such cases, any one of the *in vitro* mammalian cell assays should be done, i.e., Option 1 is followed.

2.3.3 Compounds bearing structural alerts for genotoxic activity

Structurally alerting compounds (Note 5) are usually detectable in the standard test battery since the majority of "structural alerts" are defined in relation to bacterial mutagenicity. A few chemical classes are known to be more easily detected in mammalian cell chromosome damage assays than bacterial mutation assays. Thus negative results in either test battery with a compound that has a structural alert is usually sufficient assurance of a lack of genotoxicity. However, for compounds bearing certain specific structural alerts modification to standard protocols can be

appropriate (Note 5). The choice of additional test(s) or protocol modification(s) depends on the chemical nature, the known reactivity and any metabolism data on the structurally alerting compound in question.

2.3.4 Limitations to the use of *in vivo* tests

There are compounds for which many *in vivo* tests (typically in bone marrow, blood or liver) do not provide additional useful information. These include compounds for which data on toxicokinetics or pharmacokinetics indicate that they are not systemically absorbed and therefore are not available to the target tissues. Examples of such compounds are some radioimaging agents, aluminum based antacids, some compounds given by inhalation, and some dermally or other topically applied pharmaceuticals. In cases where a modification of the route of administration does not provide sufficient target tissue exposure, and no suitable genotoxicity assay is available in the most exposed tissue, it may be appropriate to base the evaluation only on *in vitro* testing. In some cases evaluation of genotoxic effects at the site of contact may be warranted, although such assays have not yet been widely used (note 6).

2.4 Detection of germ cell mutagens

Results of comparative studies have shown that, in a qualitative sense, most germ cell mutagens are likely to be detected as genotoxic in somatic cell tests so that negative results of *in vivo* somatic cell genotoxicity tests generally indicate the absence of germ cell effects.

3. RECOMMENDATIONS FOR *IN VITRO* TESTS

3.1 Test repetition and interpretation

Reproducibility of experimental results is an essential component of research involving novel methods or unexpected findings; however, the routine testing of drugs with standard, widely used genotoxicity tests often does not need replication. These tests are sufficiently well characterized and have sufficient internal controls that repetition of a clearly positive or negative assay is not usually needed. Ideally it should be possible to declare test results clearly negative or clearly positive. However, test results sometimes do not fit the predetermined criteria for a positive or negative call and therefore are declared “equivocal”. The application of statistical methods can aid in data interpretation; however, adequate biological interpretation is of critical

importance. An equivocal test that is repeated may result in (i) a clearly positive outcome, and thus an overall positive result; (ii) a negative outcome, so that the result is not reproducible and overall negative, or (iii) another equivocal result, with a final conclusion that remains equivocal.

3.2 Recommended protocol for the bacterial mutation assays

Advice on the protocols is given in the OECD guideline (1997) and the IWGT report (Gatehouse et al, 1994).

3.2.1 Selection of top dose level

Maximum dose level

The maximum dose level recommended is 5000 µg/plate when not limited by solubility or cytotoxicity.

Limit of solubility

For bacterial cultures, precipitating doses are scored provided precipitate does not interfere with scoring, toxicity is not limiting, and the top concentration does not exceed 5000µg/plate. There is some evidence that dose-related genotoxic activity can be detected when testing certain compounds in the insoluble range in bacterial genotoxicity tests. On the other hand, heavy precipitates can interfere with scoring colonies or render the test compound unavailable to enter cells and interact with DNA. If no cytotoxicity is observed, then the lowest precipitating dose should be used as the top dose scored. If dose related cytotoxicity or mutagenicity is noted, irrespective of solubility, the top dose scored is based on cytotoxicity as described below.

Limit of cytotoxicity:

In the bacterial reverse mutation test, the doses scored should show evidence of significant toxicity, but without exceeding a top dose of 5000 µg/plate. Toxicity may be detected by a reduction in the number of revertants, and/or clearing or diminution of the background lawn.

3.2.2 Study design/Test protocol

The recommended set of bacterial strains (OECD) includes those that detect base substitution and frameshift mutations as follows: *Salmonella typhimurium* TA98; TA100; TA1535; either TA1537 or TA97 or TA97a; and either TA102 or *Escherichia coli* WP2 *uvrA* or *Escherichia coli* WP2 *uvrA* (pKM101).

One difference from the OECD and IWGT recommendations is that, based on

experience with testing pharmaceuticals, a single bacterial mutation (Ames) test is sufficient when it is clearly negative or positive, and carried out with a fully adequate protocol including all strains with and without metabolic activation, a suitable dose range that fulfills criteria for top dose selection, and appropriate positive and negative controls. Also, for testing pharmaceuticals, either the plate incorporation or the pre-incubation method is appropriate for this single experiment (note 7). Equivocal or weak positive results may indicate the need to repeat the test, possibly with a modified protocol such as appropriate spacing of dose levels.

3.3 Recommended protocols for the mammalian cell assays

Advice on the protocols is given in the OECD guidelines (1997) and the IWGT publications (Kirsch-Volders et al 2003; Moore et al 2006). Several differences from these recommendations are noted here for testing pharmaceuticals, notably for selection of the top concentration, related to the maximum concentration, cytotoxicity and solubility (see details below).

3.3.1 Selection of top concentration

<u>Maximum concentration</u>

The maximum top concentration recommended is 1 mM or 0.5 mg/ml, whichever is lower, when not limited by solubility or cytotoxicity (note 8).

<u>Limit of solubility</u>

When solubility is limiting, the maximum concentration if not limited by cytotoxicity, should be the lowest concentration at which minimal precipitate is visible in cultures, provided there is no interference with scoring. Evaluation of precipitation should be done by methods such as light microscopy, noting precipitate that persists, or appears during culture (by the end of treatment).

<u>Cytotoxicity</u>

It is not necessary to exceed a reduction of about 50% in cell growth (notes 9 and 10) for in vitro cytogenetic assays for metaphase chromosome aberrations or for micronuclei, or a reduction of about 80% in RTG (relative total growth) for the mouse lymphoma *tk* mutation assay (note 9).

3.3.2 Study design/Test protocols

For the cytogenetic evaluation of chromosomal damage in metaphase cells *in vitro*, the test protocol includes the conduct of tests with and without metabolic

activation, with appropriate positive and negative controls. Treatment with the test articles is for 3 to 6 hours with a sampling time approximately 1.5 normal cell cycles from the beginning of the treatment. A continuous treatment without metabolic activation up to the sampling time of approximately 1.5 normal cell cycles is needed in case of negative or equivocal results for both short treatments, with and without metabolic activation. The same principles apply to the *in vitro* micronucleus assay, except that the sampling time is typically 1.5 to 2 normal cell cycles from the beginning of treatment to allow cells to complete mitosis and enter the next interphase. For both *in vitro* cytogenetic assays, certain chemicals may be more readily detected by longer treatment, delayed sampling times or recovery periods, e.g., some nucleoside analogues and some nitrosamines. In the metaphase aberration assay, information on the ploidy status should be obtained by recording the incidence of polyploid (including endoreduplicated) metaphases as a percentage of the number of metaphase cells. An elevated mitotic index (MI) or an increased incidence of polyploid cells may give an indication of the potential of a compound to induce aneuploidy. For the mouse lymphoma *tk* assay, the test protocol includes the conduct of tests with and without metabolic activation, with appropriate positive and negative controls, where the treatment with the test article is for 3 to 4 hours. A continuous treatment without metabolic activation for approximately 24 hours is needed in case of a negative or equivocal result for both short treatments, with and without metabolic activation. An appropriate mouse lymphoma *tk* assay includes (i) the incorporation of positive controls that induce mainly small colonies, and (ii) colony sizing for positive controls, solvent controls and at least one positive test compound concentration (should any exist), including the culture that gave the greatest mutant frequency.

For mammalian cell assays *in vitro*, built-in confirmatory elements, such as those outlined above (e.g., different treatment lengths, tests with and without metabolic activation), are used. Following such testing, further confirmatory testing in the case of clearly negative or positive test results is not usually needed. Equivocal or weak positive results may require repeating tests, possibly with a modified protocol such as appropriate spacing of the test concentrations.

3.3.3 Positive controls

Concurrent positive controls are important, but *in vitro* mammalian cell tests

for genetic toxicity are sufficiently standardized that use of positive controls for chromosome aberration and MLA assays can be confined to a positive control with metabolic activation (provided it is done concurrently with the non-activated test) to demonstrate the activity of the metabolic activation system and the responsiveness of the test system.

4. RECOMMENDATIONS FOR *IN VIVO* TESTS

4.1 Tests for the detection of chromosome damage *in vivo*

Either the analysis of chromosomal aberrations or the measurement of micronucleated polychromatic erythrocytes in bone marrow cells *in vivo* is appropriate for the detection of clastogens. Both rats and mice are appropriate for use in the bone marrow micronucleus test. Micronuclei may also be measured in immature (e.g., polychromatic) erythrocytes in peripheral blood in the mouse, or in the newly formed reticulocytes in rat blood (note 3). Likewise, immature erythrocytes can be used from any other species which has shown an adequate sensitivity to detect clastogens/aneuploidy inducers in bone marrow or peripheral blood (note 3). Chromosomal aberrations can also be analyzed in peripheral lymphocytes cultured from treated rodents (note 11).

Note that when no *in vitro* mammalian cell assay is conducted, (Option 2), the micronucleus test *in vivo* is recommended, not the metaphase chromosome aberration assay, to include more direct capability for detection of chromosome loss (potential for aneuploidy).

4.2 Automated analysis of micronuclei

Systems for automated analysis (image analysis and flow cytometry) can be used if appropriately validated (OECD, 1997; Hayashi et al 2000; 2007).

4.3 Other *in vivo* genotoxicity tests

The same *in vivo* tests described as the second test in the standard battery (option 2) can be used as follow-up tests to develop weight of evidence in assessing results of *in vitro* or *in vivo* assays (notes 4 and 11). While the type of effect seen *in vitro* and any knowledge of the mechanism can help guide the choice of *in vivo* assay, investigation of chromosomal aberrations or of gene mutations in endogenous genes is not feasible with standard methods in most tissues; while mutation can be measured in transgenes in

rodents this entails prolonged treatment (e.g., 28 days) to allow for mutation expression/fixation, especially in tissues with little cell division. Thus the second *in vivo* assay will often evaluate a surrogate (DNA damage) endpoint. Assays with the most published experience and advice on protocols include the DNA strand break assays such as the single cell gel electrophoresis ("Comet") assay and alkaline elution assay, the i*n vivo* transgenic mouse mutation assays and DNA covalent binding assays, (all of which may be applied in many tissues, note 4), in addition to the liver unscheduled DNA synthesis (UDS) assay.

4.4 Use of male/female rodents in *in vivo* genotoxicity tests

If sex-specific drugs are to be tested, then the assay can be done in the appropriate sex. *In vivo* tests by the acute protocol may generally be carried out in only one sex (note 12). For acute tests both sexes should be considered only if any existing toxicity/metabolism data indicate a substantial sex difference in the species being used. Otherwise, males alone are appropriate for acute genotoxicity tests. When the genotoxicity test is integrated into a repeat-dose toxicology study in two sexes, samples can be collected from both sexes, but a single sex can be scored if there is no substantial sex difference evident in toxicity/metabolism. The dose levels for the sex(es) scored should meet the criteria for appropriate dose levels (sections 4.7.2 and 4.7.3).

Similar principles can be applied for other established *in vivo* genotoxicity tests.

4.5 Use of multiple administrations in genotoxicity assays *in vivo* and integration into toxicology studies

4.5.1 Sampling times

When micronucleus analysis is integrated into multi-week studies, sampling of blood or bone marrow can be done the day after the final administration (see recommendation for additional blood sampling time below).

When blood or bone marrow is used for micronucleus measurement in a multiweek study (e.g., 28 days), marked hematotoxicity may affect the ability to detect micronuclei, i.e., a dose that induces detectable increases in micronuclei after acute treatment may be too toxic to analyze after multiple treatments. It can be useful to obtain an additional sample blood on day 2 to 4 of dosing (Hamada et al, 2001); see section 4.7.3). The early sample can be used if needed to provide assurance that

clastogens and potential aneugens are detected (but see notes 13 and 17).

For other genotoxicity assays, sampling time is selected as appropriate for the endpoint measured; for example DNA damage/strand break measurements are usually made a few (e.g., 2-6) hours after the last administration.

In principle, studies of any length may be appropriate provided the top dose/exposure is adequate.

4.5.2 Number of animals analyzed

The number of animals analyzed is determined by current recommendations for the micronucleus assay (OECD) or other genotoxicity assays and generally does not include all the animals treated for a toxicology study. (Animals used for genotoxicity analyses should be randomly selected).

4.6 Route of administration

The route of administration is generally the expected clinical route, e.g., oral, intravenous or subcutaneous, but can be modified if needed to obtain systemic exposure, e.g., for topically applied compounds (see section 2.3.4).

4.7 Dose selection for *in vivo* assays

Typically three dose levels are used (Hayashi et al, 2005).

4.7.1 Short-term studies

For short term (usually 1 to 2 administrations) protocols, the top dose recommended for genotoxicity assays is a limit dose of 2000 mg/kg if this is tolerated, or maximum tolerated dose defined, for example for the micronucleus assay (OECD 474) as the dose producing signs of toxicity such that higher dose levels, based on the same dosing regimen, would be expected to produce lethality. (Similar recommendations have been made for the Comet assay [Hartmann et al, 2003] and transgenic mutation assay [Heddle et al, 2000]). Suppression of bone marrow red blood cell production may also be taken into account in dose selection. Lower doses are generally spaced at approximately two to three fold intervals below this.

4.7.2 Multiple administration studies

In the Option 1 battery, when the *in vitro* mammalian cell assay is negative (or "non-relevant positive" (see section 5), if the *in vivo* genotoxicity test is integrated into a multiple administration toxicology study, the doses are generally considered appropriate when the toxicology study meets the criteria for an adequate study to

support human clinical trials. However, when carrying out follow-up studies to address any indication of genotoxicity, or when using Option 2 with no *in vitro* mammalian cell assay, several factors should be evaluated to demonstrate that the top dose is appropriate for genotoxicity evaluation, as follows:

Recommendations for determining whether the top dose in a toxicology study (typically in rats) is appropriate for micronucleus analysis and for other genotoxicity evaluation (any one of the following):

i. Maximum feasible dose (MFD) based on physico-chemical properties of the drug in the vehicle (provided the MFD in that vehicle is similar to that achievable with acute administration; note 14).

ii. Limit dose of 1000 mg/kg for studies of 14 days or longer, if this is tolerated

iii. Exposure:

a. Plateau/saturation in exposure

b. Accumulation

Substantial reduction in exposure to parent drug with time (e.g., ≥ 50% reduction from initial exposure) would usually disqualify the study. If this is seen in one sex, generally the sex with reduced exposure would not be scored, unless there is enhanced exposure to a metabolite of interest.

iv Top dose is ≥ 50% of the top dose that would be used for acute administration, i.e., close to the minimum lethal dose, if such acute data are available for other reasons. (The top dose for acute administration micronucleus test is currently described in OECD guidance as the dose above which lethality would be expected; similar guidance is given [e.g. Hartmann et al, 2003] for other *in vivo* assays.)

Selection of a top dose based only on an exposure margin (multiple over clinical exposure) without toxicity is not considered sufficient justification.

If dose levels/exposure are not appropriate, acute *in vivo* assays should be performed to maximize exposure or obtain the appropriate toxicity range, (preferably conducting two genotoxicity assays in the same animals), or an *in vitro* mammalian cell assay should be done if not already completed.

4.7.3 Additional guidance on dose selection for multiple administration studies

Compounds that induce aneuploidy, such as spindle poisons, are typically

detectable in *in vivo* micronucleus assays in bone marrow or blood only within a narrow range of doses approaching toxic doses. This is also true for some clastogens. If toxicological data indicate severe toxicity to red blood cell lineage (e.g., marked suppression of PCEs or reticulocytes), doses scored should be spaced not more than about 2 fold below the top, cytotoxic dose. If suitable doses are not included in a multi-week study, additional data may be required to ensure detection of aneugens and some toxic clastogens; these could be derived from any one of the following:

a. 2 -4 day blood sampling from the multiweek study before substantial hematotoxicity developed
b. an *in vitro* mammalian cell micronucleus assay
c. An acute bone marrow micronucleus assay

4.8 Demonstration of target tissue exposure for negative *in vivo* test results

In vivo tests have an important role in genotoxicity test strategies. The value of *in vivo* results is directly related to the demonstration of adequate exposure of the target tissue to the test compound. This is especially true for negative *in vivo* test results when *in vitro* test(s) have shown convincing evidence of genotoxicity, or when no *in vitro* mammalian cell assay is used. Evidence of adequate exposure could include toxicity in the tissue in question, or toxicokinetic data.

4.8.1 When an *in vitro* genotoxicity test is positive (or not done)

Assessments of *in vivo* exposure should be made at the top dose or other relevant doses using the same species, strain and dosing route used in the genotoxicity assay. When genotoxicity is measured in toxicology assays, exposure information is generally available as part of the toxicology assessment.

Demonstration of *in vivo* exposure should be made by any of the following measurements:

i. Cytotoxicity
   a. For cytogenetic assays: By obtaining a significant change in the proportion of immature erythrocytes among total erythrocytes in the tissue used (bone marrow or blood), at the doses and sampling times used in the micronucleus test or by measuring a significant reduction in mitotic index for the chromosomal aberration assay.
   b. For other *in vivo* genotoxicity assays: Toxicity in the liver or tissue being

assessed, e.g., by histopathological evaluation or blood biochemistry toxicity indicators.

ii. Bioavailability

a. Measurement of drug related material either in blood or plasma. The bone marrow is a well perfused tissue and levels of drug related materials in blood or plasma are generally similar to those observed in bone marrow. Liver is expected to be exposed for drugs with systemic exposure regardless of the route of administration.

b. Direct measurement of drug-related material in target tissue, or autoradiographic assessment of tissue exposure.

If systemic exposure is similar to or lower than expected clinical exposure, alternative strategies may be needed such as (i) use of a different route of administration; (ii) use of a different species with higher exposure; (iii) use of a different tissue or assay (see section 2.3.4, "Limitations to the use of standard *in vivo* tests".

If adequate exposure cannot be achieved e.g., with compounds showing very poor target tissue availability, conventional *in vivo* genotoxicity tests may have little value.

4.8.2 When *in vitro* genotoxicity tests are negative

If *in vitro* tests do not show genotoxic potential, *in vivo* (systemic) exposure can be assessed by any of the methods above, or can be assumed from the results of standard absorption, distribution, metabolism and excretion (ADME) studies in rodents done for other purposes.

4.9 Use of positive controls for *in vivo* studies

For *in vivo* studies, it is not necessary to include concurrent treatments with positive controls in every study, after a laboratory has established competence in the use of the assay (note 15).

5. GUIDANCE ON EVALUATION OF TEST RESULTS AND ON FOLLOW-UP TEST STRATEGIES

Comparative trials have shown conclusively that each *in vitro* test system generates both false negative and false positive results in relation to predicting rodent

carcinogenicity. Genotoxicity test batteries (of *in vitro* and *in vivo* tests) detect carcinogens that are thought to act primarily via a mechanism involving direct genetic damage, such as the majority of known human carcinogens. Therefore, these batteries are not expected to detect non-genotoxic carcinogens. Experimental conditions, such as the limited capability of the *in vitro* metabolic activation systems, can lead to false negative results in *in vitro* tests. The test battery approach is designed to reduce the risk of false negative results for compounds with genotoxic potential, whereas a positive result in any assay for genotoxicity does not necessarily mean that the test compound poses a genotoxic/carcinogenic hazard to humans.

Although positive *in vitro* data may indicate intrinsic genotoxic properties of a drug, appropriate i*n vivo* data determine the biological significance of these *in vitro* signals in most cases. Also, because there are several indirect mechanisms of genotoxicity that operate only above certain concentrations, it is possible to establish a safe level (threshold) for classes of drugs with evidence for such mechanisms (see 5.2. below, Müller and Kasper, 2000; Scott et al, 1991; Thybaud et al 2007).

5.1 Assessment of biological relevance

The recommendations below assume that the test has been conducted using appropriate spacing of doses, levels of toxicity etc.

Small increases in apparent genotoxicity *in vitro* or *in vivo* should first be assessed for reproducibility and biological significance. Examples of results that are not considered biologically meaningful include:

i. Small increases that are statistically significant compared with the negative or solvent control values but are within the historical control range for the testing facility
ii. Weak/equivocal responses that are not reproducible

If any of the above conditions apply the weight of evidence indicates a lack of genotoxic potential, the test is considered negative or the findings not biologically relevant, and no further testing is required.

5.2 Evaluation of results obtained in *in vitro* tests

In evaluating positive results, especially for the microbial mutagenicity test, the purity of the test compound should be considered, to determine whether the positive result may be attributable to a contaminant.

5.2.1 Evaluation of positive results obtained *in vitro* in a bacterial mutation assay

There are some well characterized examples of artefactual increases in colonies that are not truly revertants. These may occur due to contamination with amino acids, (providing histidine for *Salmonella* strains or tryptophan for *Escherichia Coli* strains), so that the bacterial reversion assay is not suitable for testing a peptide that is likely to degrade. Certain cases exist where positive results in bacterial mutation assays may be shown not to indicate genotoxic potential *in vivo* in humans, for example when bacterial-specific metabolism occurs, such as activation by bacterial nitroreductases.

5.2.2 Evaluation of positive results obtained *in vitro* in mammalian cell assays

Recommendations for assessing weight of evidence and follow-up testing for positive genotoxicity results are discussed in IWGT reports (e.g., Thybaud et al 2007). In addition, the scientific literature gives a number of conditions that may lead to a positive *in vitro* result of questionable relevance. Therefore, any *in vitro* positive test result should be evaluated based on an assessment of the weight of evidence as indicated below. This list is not exhaustive, but is given as an aid to decision-making.

i. Conditions that do not occur *in vivo* (pH; osmolality; precipitates)

   Note that the 1 mM limit avoids increases in osmolality, and that if the test compound alters pH it is advisable to adjust pH to the normal pH of untreated cultures at the time of treatment.

ii. The effect occurs only at the most toxic concentrations.

   In the MLA increases at ≥80% reduction in RTG

   For *in vitro* cytogenetics assays when growth is suppressed by ≥50%

If any of the above conditions apply the weight of evidence indicates a lack of genotoxic potential and no additional testing beyond the standard battery (option 1) with one negative *in vivo* test would be needed.

5.2.3 Evaluation of *in vitro* negative results

For *in vitro* negative results further testing should be considered in special cases, such as (the examples given are not exhaustive, but are given as an aid to decision-making): The structure or known metabolism of the compound indicates that standard techniques for *in vitro* metabolic activation (e.g., rodent liver S9) may be inadequate; the structure or known activity of the compound indicates that the use of

other test methods/systems may be appropriate.

5.3 Evaluation of results obtained from *in vivo* tests

In vivo tests have the advantage of taking into account absorption, distribution and excretion, which are not factors in *in vitro* tests, but are potentially relevant to human use. In addition metabolism is likely to be more relevant *in vivo* compared to the systems normally used *in vitro*. If the *in vivo* and *in vitro* results do not agree, then the difference should be considered/explained on a case-by-case basis, e.g., difference in metabolism; rapid and efficient excretion of a compound may occur *in vivo*, etc.

In vivo genotoxicity tests also have the potential to give misleading positive results that do not indicate true genotoxicity. For example, increases in micronuclei can occur without administration of any genotoxic agent, due to disturbance in erythropoeisis (Tweats et al, 2007 I), DNA adduct data should be interpreted in the light of the known background level of endogenous adducts, and indirect, toxicity-related effects can influence the results of the DNA strand break assays (e.g., alkaline elution and Comet assays). Thus it is important to take into account all the toxicological and hematological findings when evaluating the genotoxicity data (note 17). Indirect effects related to toxicological changes may have a safety margin and may not to be clinically relevant.

5.4 Follow-up strategies for positive results

5.4.1 Follow-up to findings *in vitro* in mammalian cell tests

The following discussion assumes negative results in the Ames bacterial mutation assay.

5.4.1.1 Mechanistic/in vivo follow-up

To evaluate *in vitro* mammalian cell assay positive results for which there is insufficient weight of evidence to indicate lack of relevance, recommended follow-up for mammalian cell assays would be to provide experimental evidence, either by additional in vitro studies ***or*** by carrying out two appropriate *in vivo* assays, as follows:

i. Mechanistic information that contributes to a weight of evidence for a lack of relevant genotoxicity is often generated *in vitro*, for example evidence that a test compound that induces chromosome aberrations, or mutations in the MLA is not a DNA damaging agent (e.g., other negative mutation/DNA damage tests in addition to the Ames test; structural considerations), or evidence for an indirect/threshold mechanism not relevant *in vivo* (e.g., inhibition of DNA

synthesis, reactive oxygen species produced only at high concentrations, etc, (Galloway et al, 1998; Scott et al, 1991; Muller and Kasper, 2000). Similar studies can be used to follow up a positive result in the *in vitro* micronucleus assay, or in this case evidence can include a known mechanism that indicates chromosome loss/aneuploidy, or centromere staining experiments (note 18) that indicate chromosome loss.

If the above mechanistic information and weight of evidence supports the lack of relevant genotoxicity, only a single *in vivo* test is needed, with appropriate evidence of exposure, to establish the lack of genotoxic activity. This is typically a cytogenetic assay, and the micronucleus assay *in vivo* is needed when following up potential for chromosome loss.

Polyploidy is a common finding in chromosome aberration assays *in vitro*. While aneugens can induce polyploidy, polyploidy alone does not indicate aneugenic potential and may simply indicate cell cycle perturbation; it is also commonly associated with increasing cytotoxicity. If polyploidy, but no structural chromosome breakage, is seen in an *in vitro* assay, generally a negative *in vivo* micronucleus assay with assurance of appropriate exposure would provide sufficient assurance of lack of potential for aneuploidy induction.

<u>Or</u>

ii. Two appropriate *in vivo* assays are done, usually with different tissues, and with supporting demonstration of exposure.

In summary, if the results of the *in vitro* mammalian cell assay are positive and there is not sufficient weight of evidence or mechanistic information to rule out relevant genotoxic potential, two *in vivo* tests are required, with appropriate endpoints and in appropriate tissues (usually two different tissues), and with an emphasis on obtaining sufficient exposure in the *in vivo* models.

Negative results in appropriate *in vivo* assays, with adequate justification for the endpoints measured, and demonstration of exposure (see section 4.8.1) is sufficient to demonstrate absence of genotoxic activity.

5.4.1.2 Follow-up to an *in vitro* positive result that is dependent upon S-9

activation

When positive results are seen only in the presence of the S-9 activation system, it should first be verified that metabolic activation is responsible and not some other difference in conditions (e.g., low or no serum in the S-9 mix, compared with ≥10% serum in the non-activated incubations). The follow-up strategy is then aimed at determining the relevance of any reactive metabolites produced *in vitro* to conditions *in vivo*, and will generally focus on *in vivo* studies in liver (note 16).

5.4.2 Follow-up to a positive *in vivo* micronucleus assay

If there is an increase in micronuclei *in vivo*, all the toxicological data should be evaluated to determine whether a non-genotoxic effect may be the cause or a contributing factor (note 17). If non-specific effects of disturbed erythropoeisis or physiology (such as hypo/hyperthermia) are suspected, an *in vivo* assay for chromosome aberrations may be more appropriate. If a "real' increase is suspected, strategies would be needed to demonstrate whether the increase is due to chromosome loss or chromosome breakage (note 18). There is evidence that aneuploidy induction, e.g., with spindle poisons, follows a non-linear dose response. Thus, it may be possible to determine that there is a threshold exposure below which chromosome loss is not expected and to determine whether an appropriate safety margin exists compared with clinical exposure.

In conclusion, the assessment of the genotoxic potential of a compound should take into account the totality of the findings and acknowledge the intrinsic values and limitations of both *in vitro* and *in vivo* tests.

5.5 Follow-up genotoxicity testing in relation to tumor findings in a carcinogenicity bioassay

Additional genotoxicity testing in appropriate models may be conducted for compounds that were negative in the standard test battery but which have shown increases in tumors in carcinogenicity bioassay(s) with insufficient evidence to establish a non-genotoxic mechanism. To help understand the mode of action, additional testing can include modified conditions for metabolic activation in *in vitro* tests or can include *in vivo* tests measuring genetic damage in target organs of tumour induction, such as DNA strand break assays (e.g., comet or alkaline elution assays), liver UDS test, DNA covalent binding (e.g., by ^{32}P-postlabelling), mutation induction in transgenes, or

641 molecular characterization of genetic changes in tumor-related genes (Kasper et al,
642 2007).

6. NOTES

1. The *in vitro* micronucleus assay has been widely evaluated in international collaborative studies (Kirsch-Volders et all, 2003), is considered validated by ECVAM (Corvi et al, 2008), and an OECD guideline is in preparation.

2. There is a small but significant number of genotoxic carcinogens that are reliably detected by the bone marrow tests for chromosomal damage but have yielded negative/weak/conflicting results in the *in vitro* tests outlined in the standard battery options. Carcinogens such as procarbazine, hydroquinone, urethane and benzene fall into this category. Some other examples from a survey of companies are described by Tweats et al, 2007, II.

3. In principle, micronuclei in hematopoeitic cells may be evaluated in bone marrow from any species, and in blood from species that do not filter out circulating micronucleated erythrocytes in the spleen. In laboratory mice, micronuclei can be measured in polychromatic erythrocytes in blood, and mature (normochromatic) erythrocytes can be used when mice are treated continuously for about 4 weeks or more. Although rats rapidly remove micronucleated erythrocytes from the circulation, it has been established that micronucleus induction by a range of clastogens and aneugens can be detected in rat blood reticulocytes (Wakata et al, 1998; Hamada et al 2001). Rat blood may be used for micronucleus analysis provided methods are used to ensure analysis of the newly formed reticulocytes (Hayashi et al, 2007; MacGregor et al, 2006), and the sample size is sufficiently large to provide appropriate statistical sensitivity given the lower micronucleus levels in rat blood than in bone marrow (Kissling et al, 2007). Whichever method is chosen, bone marrow or blood, automated or manual analysis, each laboratory should determine the minimum sample size required to ensure that scoring error is maintained below the level of animal-to-animal variation.

Some experience is now available for micronucleus induction in the dog. One example where such alternative species might be useful would be in evaluation of a human metabolite that was not sufficiently represented in rodents but was formed in the dog.

4. The inclusion of a second *in vivo* assay in the battery is to provide assurance of lack of genotoxicity by use of a tissue that is well exposed to a drug and/or its metabolites; a small number of carcinogens that are considered genotoxic gave positive

675 results in a test in liver but were negative in a cytogenetics assay *in vivo* in bone marrow.
676 These examples likely reflect a lack of appropriate metabolic activity or lack of reactive
677 intermediates delivered to the hematopoietic cells of the bone marrow.
678 Assays for DNA strand breaks, DNA adducts, and mutation in transgenes have the
679 advantage that they can be applied in many tissues. Internationally agreed protocols
680 are not yet in place for all the *in vivo* assays, although considerable experience and
681 published data exist for DNA strand break assays (Comet and alkaline elution assays)
682 DNA adduct (covalent binding) measurements and transgenic rodent mutation assays, in
683 addition to the UDS assay. Because cytotoxicity induces DNA strand breakage,
684 careful cytotoxicity assessment is needed to avoid confounding the results of DNA
685 strand break assays. This has been well characterized for the alkaline elution assay
686 (Storer et al, 1996) but not yet fully validated for the Comet assay. In principle the
687 DNA strand break assays may be used in repeat-dose toxicology assays with appropriate
688 dose levels and sampling times.
689 Since liver of mature animals is not a highly mitotic tissue, often a non-
690 cytogenetic endpoint is used for the second assay, but with special protocols, or in
691 young rats (Suzuki et al 2005), micronucleus analysis in liver is possible, and detects
692 known genotoxic compounds.
693 5. Certain structurally alerting molecular entities are recognized as being causally
694 related to the carcinogenic and/or mutagenic potential of chemicals. Examples of
695 structural alerts include alkylating electrophilic centers, unstable epoxides, aromatic
696 amines, azo-structures, *N*-nitroso groups, and aromatic nitro-groups (Ashby and Paton
697 1994). For some classes of compounds with specific structural alerts, it is established
698 that specific protocol modifications/additional tests are important for optimum detection
699 of genotoxicity (e.g., molecules containing an azo-group, glycosides, compounds such
700 as nitroimidazoles requiring nitroreduction for activation, compounds such as
701 phenacetin requiring a different rodent S9 for metabolic activation).
702 6. There is some experience with *in vivo* assays for micronucleus induction in
703 skin, liver and colon (Hayashi et al 2007) and DNA damage assays in these tissues can
704 also be an appropriate substitute.
705 7. A few chemicals are more easily detectable either with plate-incorporation or
706 with pre-incubation methods though differences are typically quantitative rather than

707 qualitative (Gatehouse et al, IWGT, 1994). Experience in the pharmaceutical industry
708 where drugs have been tested in both protocols has not resulted in different results for
709 the two methods and in the IWGT report the examples of chemical classes listed as
710 more easily detectable in the pre-incubation protocol are generally not pharmaceuticals
711 and are positive in *in vivo* genotoxicity tests in liver. These include short chain
712 aliphatic nitrosamines; divalent metals; aldehydes (e.g., formaldehyde, crotonaldehyde);
713 azo dyes (e.g., butter yellow); pyrrolizidine alkaloids; allyl compounds
714 (Allylisothiocyanate, allyl chloride), and nitro (aromatic, aliphatic) compounds.
715 8. The rationale for a maximum concentration of 1 mM for *in vitro* mammalian
716 cell assays includes the following: The test battery includes the Ames test and an *in*
717 *vivo* assay. Viewing the battery as a whole means that it is not necessary to detect in
718 the mammalian cell assay every compound considered to be a genotoxic carcinogen.
719 There is a low likelihood of such compounds of concern (DNA damaging carcinogens)
720 that are not detected in Ames test or *in vivo* genotoxicity assay, but are detectable in an
721 *in vitro* mammalian assay only above 1 mM. Second, a limit of 1 mM maintains the
722 element of hazard identification, being higher than clinical exposures to known
723 pharmaceuticals, including those that concentrate in tissues (Goodman & Gilman's,
724 2001), and is also higher than the levels generally achievable in preclinical studies in
725 vivo. Certain drugs are known to require quite high clinical exposures, e.g., nucleoside
726 analogs and some antibiotics. While comparison of potency with existing drugs may
727 be of interest to sponsors, perhaps even above the 1 mM limit, it is ultimately the *in vivo*
728 tests that determine relevance for human safety.
729 9. Although some genotoxic carcinogens are not detectable in *in vitro*
730 genotoxicity assays unless the concentrations tested induce some degree of cytotoxicity,
731 particularly when measured by colony forming assays, DNA damaging agents are
732 generally detectable with only moderate levels of toxicity (e.g., 30% reduction in
733 growth measured at the time of sampling in the chromosome aberration assay,
734 Greenwood et al, 2004). As cytotoxicity increases, mechanisms other than direct DNA
735 damage by a compound or its metabolites can lead to ‘positive’ results that are related to
736 cytotoxicity and not genotoxicity. Such indirect induction of DNA damage secondary
737 to damage to non-DNA targets are more likely to occur above a certain concentration
738 threshold. The disruption of cellular processes is not expected to occur at lower,

739 pharmacologically relevant concentrations.
740 In cytogenetic assays, even weak clastogens that are known to be carcinogens are
741 positive without exceeding a 50% reduction in cell counts. On the other hand,
742 compounds that are not DNA damaging, mutagenic or carcinogenic can induce
743 chromosome breakage at toxic concentrations. For both *in vitro* cytogenetic assays,
744 the chromosome aberration assay and the *in vitro* micronucleus assay, a limit of about
745 50% growth reduction is appropriate.
746 For cytogenetic assays in cell lines, measurement of cell population growth over time
747 (by measuring the change in cell number during culture relative to control, e.g., by the
748 method referred to as population doubling (PD; note 10), has been shown to be a useful
749 measure of cytotoxicity, as it is known that cell numbers can underestimate toxicity.
750 For lymphocyte cultures, an inhibition of proliferation not exceeding about 50% is
751 considered sufficient; this can be measured by mitotic index (MI) for metaphase
752 aberration assays and by an index based on cytokinesis block for *in vitro* micronucleus
753 assays. In addition, for the *in vitro* micronucleus assay, since micronuclei are scored
754 in the interphase subsequent to a mitotic division, it is important to verify that cells have
755 progressed through the cell cycle. This can be done by use of cytochalasin B to allow
756 nuclear division but not cell division, so that micronuclei can be scored in binucleate
757 cells (the preferred method for lymphocytes). For cell lines other methods to
758 demonstrate cell proliferation, including cell population growth over time (PD) as
759 described above, may be used (Kirsch-Volders et al 2003).
760 For the mouse lymphoma assay, appropriate sensitivity is achieved by limiting the top
761 concentration to one with close to 20% Relative Total Growth (RTG) both for soft agar
762 and for microwell methods (IWGT). Reviews of published data using the current
763 criteria described by Moore et al (2006) found very few chemicals that were positive in
764 MLA only at concentrations with less than 20% RTG and that were rodent carcinogens,
765 and convincing evidence of genotoxic carcinogenesis for this category is lacking. The
766 consensus (Moore et al, 2006) is that caution is needed in interpreting results when
767 increases in mutation are seen only below 20% RTG, and a result would not be
768 considered positive if the increase in mutant fraction occurred only at ≤ 10% RTG.
769 In conclusion, caution is appropriate in interpreting positive results obtained as
770 reduction in growth/survival approaches or exceeds 50% for cytogenetics assays or 80%

771 for the mouse lymphoma assay. It is acknowledged that the evaluation of cells treated
772 at these levels of cytotoxicity/clonal survival may result in greater sensitivity, but bears
773 an increased risk of non-relevant positive results. The battery approach for
774 genotoxicity is designed to ensure appropriate sensitivity without the need to rely on
775 single *in vitro* mammalian cell tests at high cytotoxicity.
776 To obtain an appropriate toxicity range, a preliminary range-finding assay over a broad
777 range of concentrations is useful, but in the genotoxicity assay it is often critical to use
778 multiple concentrations that are spaced quite closely (less than two–fold dilutions).
779 Extra concentrations may be tested but not all need be evaluated for genotoxicity. It is
780 not intended that multiple experiments be carried out to reach exactly 50% reduction in
781 growth, for example, or exactly 80% reduction in RTG.
782 10. For *in vitro* cytogenetic assays it is appropriate to use a measure of relative cell
783 growth to assess toxicity, because cell counts can underestimate toxicity (Greenwood et
784 al, 2004). Using calculated population doublings to estimate the 50% growth
785 reduction level it was demonstrated that the frequency of positive results with
786 compounds that are not mutagenic or carcinogenic is reduced, while true DNA
787 damaging agents are reliably positive.
788 11. In certain cases it may be useful to examine chromosome aberrations at
789 metaphase in lymphocytes cultured from test animals after one or more administrations
790 of test compound, just as bone marrow metaphase cells may be used. Because some
791 lymphocytes are relatively long-lived, in principle there is the potential for
792 accumulation of un-repaired DNA damage *in vivo*, that would give rise to aberrations
793 when the cells are stimulated to divide *in vitro*. The *in vivo* lymphocyte assay may be
794 useful in following up indications of clastogenicity, but in general another tissue such as
795 liver is a more informative supplement to the micronucleus assay in hematopoeitic cells
796 because exposure to drug and metabolite(s) is often higher in liver.
797 12. Extensive studies of the activity of known clastogens in the acute mouse bone
798 marrow micronucleus test have shown that in general male mice are more sensitive than
799 female mice for micronucleus induction. Quantitative differences in micronucleus
800 induction have been identified between the sexes, but no qualitative differences have
801 been described. Where marked quantitative differences exist, there is invariably a
802 difference in toxicity between the sexes. Thus males alone can be appropriate for

803 acute in vivo micronucleus tests.
804 13. Caution is required if the toxicological study design includes additional blood
805 sampling, e.g., for measurement of exposure. Such bleeding could perturb the results
806 of micronucleus analysis since erythropoeisis stimulated by bleeding can lead to
807 increases in micronucleated erythrocytes.
808 14. For common vehicles like aqueous methyl cellulose this would usually be
809 appropriate, but for vehicles such as Tween 80, the volume that can be administered
810 could be as much as 30 fold lower than that given acutely.
811 15. For micronucleus (and other cytogenetic) assays, the purpose of the positive
812 control is to verify that the individuals scoring the slides can reliably detect increases in
813 micronuclei. This can be accomplished by use of samples from periodic studies of
814 small groups of positive control animals (one sex). For manual scoring such slides can
815 be included in coded slides scored from each study, or used for periodic demonstration
816 of ability of readers to recognize positive responses. Positive control slides should not
817 be obvious to readers based on their staining properties or micronucleus frequency.
818 For automated scoring, appropriate quality control samples should be used with each
819 assay.
820 For other *in vivo* genotoxicity assays, the purpose of positive controls is to
821 demonstrate reliable detection of an increase in DNA damage/mutagenicity using the
822 assay in the chosen species, tissue and protocol. After a laboratory has demonstrated
823 that it can consistently detect appropriate positive control compounds in multiple
824 independent experiments, it is no longer necessary to carry out concurrent controls with
825 every assay using that protocol, but controls can be tested periodically.
826 16. Standard induced S-9 mix has higher activation capacity than human S-9, and
827 lacks phase two detoxification capability unless specific cofactors are supplied. Also,
828 non-specific activation can occur *in vitro* with high test substrate concentrations (see
829 Kirkland et al, 2007). Genotoxicity testing with human S-9 or other human-relevant
830 activation systems can be helpful. Analysis of the metabolite profile in the
831 genotoxicity test incubations for comparison with known metabolite profiles in
832 preclinical species, (in uninduced microsomes or hepatocytes, or *in vivo*), or in
833 preparations from humans, can also help determine the relevance of test results (Ku et al,
834 2007), and follow-up studies will usually focus on *in vivo* testing in liver. A

835 compound that gives positive results *in vitro* with S-9 may not induce genotoxicity *in*
836 *vivo* because the metabolite is not formed, is formed in very small quantities, or is
837 metabolically detoxified or rapidly excreted, indicating a lack of risk *in vivo*.
838 17. Increases in micronuclei can occur without administration of any genotoxic
839 agent, due to disturbance in erythropoeisis (such as regenerative anemia; extramedullary
840 hematopoeisis), stress, hypo- and hyperthermia (reviewed by Tweats et al 2007I, IWGT).
841 In blood, changes in spleen function that affect clearance of micronucleated cells from
842 the blood are expected to lead to increases in circulating micronucleated red blood cells.
843 18. Determination of whether micronucleus induction is due primarily to
844 chromosome loss or to chromosome breakage could include staining micronuclei *in*
845 *vitro* or *in vivo* to determine whether centromeres are present. e.g., using fluorescent in
846 situ hybridization (FISH) with probes for DNA sequences in the centromeric region, or
847 a labeled antibody to kinetochore proteins. If the majority of induced micronuclei are
848 centromere positive, this suggests chromosome loss. (Note that even potent tubule
849 poisons like colchicine and vinblastine do not produce 100% kinetochore positive
850 micronuclei, but more typically 70 to 80%, but are accepted as primarily aneugens for
851 assessing risk). An alternative approach is to carry out an *in vitro* or *in vivo* assay for
852 metaphase structural aberrations; if negative this would infer that micronucleus
853 induction is related to chromosome loss.

7. GLOSSARY

Alkaline elution assay: see *DNA strand break assay*

Aneuploidy: numerical deviation of the modal number of chromosomes in a cell or organism.

Base substitution: the substitution of one or more base(s) for another in the nucleotide sequence. This may lead to an altered protein.

Cell proliferation: the ability of cells to divide and to form daughter cells.

Centromere/kinetochore: structures in chromosomes essential for association of sister chromatids and for attachment of spindle fibers that move daughter chromosomes to the poles and ensure inclusion in daughter nuclei

Clastogen: an agent that produces structural breakage of chromosomes, usually detectable by light microscopy.

Cloning efficiency: the efficiency of single cells to form clones. Usually measured after seeding low numbers of cells in a suitable environment.

Comet assay: see DNA strand break assay

Culture confluency: a quantification of the cell density in a culture by visual inspection

Cytogenetic evaluation: chromosome structure analysis in mitosis or meiosis by light microscopy, or micronucleus analysis

DNA adduct: product of covalent binding of a chemical to DNA

DNA repair: reconstitution of the original DNA sequence after DNA damage

DNA strand breaks: single or double strand scissions in the DNA

DNA strand break assay: alkaline treatment converts certain types of DNA lesions into strand breaks that can be detected by the alkaline elution technique, measuring migration rate through a filter, or by the single cell gel electrophoresis or Comet assay in which cells embedded in a thin layer of gel on a microscope slides are subjected to electric current, causing shorter pieces of DNA to migrate out of the nucleus into a "Comet tail". The extent of DNA migration is measured visually under the microscope on stained cells.

Frameshift mutation: a mutation (change in the genetic code) in which one base or two adjacent bases are added to (inserted in) or deleted from the nucleotide sequence of a gene. This may lead to an altered or truncated protein.

Gene mutation: a detectable permanent change within a single gene or its regulating

sequences. The changes may be point mutations, insertions, deletions.

Genetic endpoint: the precise type or class of genetic change investigated (e.g., gene mutations, chromosomal aberrations, DNA strand breaks, DNA repair, DNA adduct formation, etc).

Genotoxicity, genotoxicity: a broad term that refers to any deleterious change in the genetic material regardless of the mechanism by which the change is induced.

Micronucleus: particle in a cell that contains nuclear DNA; it might contain a whole chromosome(s) or a broken centric or acentric part(s) of chromosome(s).

Mitotic index: percentage of cells in the different stages of mitosis amongst the cells not in mitosis (interphase) in a preparation (slide).

Plasmid: genetic element additional to the normal bacterial genome. A plasmid might be inserted into the host chromosome or form an extra-chromosomal element.

Numerical chromosome changes: chromosome numbers different from the original haploid or diploid set of chromosomes; for cell lines, chromosome numbers different from the modal chromosome set

Point mutations: changes in the genetic code, usually confined to a single DNA base pair.

Polychromatic erythrocyte: an immature erythrocyte in an intermediate stage of development that still contains ribosomes and, as such, can be distinguished from mature normochromatic erythrocytes (lacking ribosomes) by stains selective for RNA.

Population doubling or culture growth: This can be calculated in different ways; one example of an appropriate formula is: Population doublings (PDs) = the log of the ratio of the final count (N) to the starting (baseline) count (Xo), divided by the log of 2. That is: $PD = [\log(N \div Xo)] \div \log 2$.

Polyploidy: Numerical deviation of the modal number of chromosomes in a cell, with approximately whole multiples of the haploid number. Endoreduplication is a morphological form of polyploidy in which chromosome pairs are associated at metaphase as “diplochromosomes”

Recombination: breakage and balanced or unbalanced rejoining of DNA

RTG (relative total growth): This measure of cytotoxicity takes the relative suspension growth (based on cell loss and cell growth from the beginning of treatment to the second day post-treatment) and multiplies it by the relative plating efficiency at the time

of cloning for mutant quantization.

Single Cel Gel Electrophoresis assay: Comet assay. *See DNA strand break assay*

Survival (in the context of mutagenicity testing): proportion of living cells among dead cells, usually determined by staining or colony counting methods after a certain treatment interval.

Transgene: an exogenous or foreign gene inserted into the host genome, either into somatic cells or germ line cells

Unscheduled DNA synthesis (UDS): DNA synthesis that occurs at some stage in the cell cycle other than S-phase in response to DNA damage. It is usually associated with DNA excision repair.

8. REFERENCES

Ashby, J., D. Paton, "The influence of chemical structure on the extent and sites of carcinogenesis for 522 rodent carcinogens and 55 different human carcinogen exposures," *Mutation Research* 286:3-74, 1994.

Corvi, R., S. Albertini, T. Hartung, S. Hoffmann, D. Maurici, S. Pfuhler, J. van Benthem, P. Vanparys, "ECVAM Retrospective Validation of the *in vitro* micronucleus test (MNT)", *Mutagenesis*, submitted 2007.

.

Gatehouse, D., S. Haworth, T. Cebula, E. Gocke, L. Kier, T. Matsushima, C. Melcion, T. Nohmi, T. Ohta, S. Venitt, E. Zeiger, "Report from the working group on bacterial mutation assays: International Workshop on Standardisation of Genotoxicity Test Procedures," *Mutation Research* 312: 217-233, 1994

Goodman & Gilman "The Pharmacological Basis of Therapeutics". J. G. Hardman, L E. Limbird, A. G. Gilman (Eds.), McGraw-Hill Professional, New York; 10th edition (August 13, 2001).

Greenwood, S.K., R.B. Hill, J.T. Sun, M.J. Armstrong, T.E. Johnson, J.P. Gara, S.M. Galloway, "Population Doubling: A simple and more accurate estimation of cell growth suppression in the *in vitro* assay for chromosomal aberrations that reduces irrelevant positive results," *Environmental and Molecular Mutagenesis*, 43:36-44, 2004.

Hamada S., S. Sutou, T. Morita, A. Wakata, S. Asanami, S. Hosoya, et al., "Evaluation of the rodent micronucleus assay by a 28-day treatment protocol: summary of the 13th collaborative study by the collaborative study group for the micronucleus test (CSGMT)/Environmental Mutagen Society of Japan (JEMS)–Mammalian Mutagenicity Study Group (MMS)," *Environmental and Molecular Mutagenesis*, 37:93-110, 2001.

Hartmann A., E. Agurell, C. Beevers, S. Brendler-Schwaab, B. Burlinson, P. Clay, A. Collins, A. Smith, G. Speit, V. Thybaud, R.R. Tice, "Recommendations for conducting the in vivo alkaline Comet assay", *Mutagenesis* 18:45-51, 2003.

Hayashi, M., J.T. MacGregor, D.G. Gatehouse, I. Adler, D.H. Blakey, S.D. Dertinger, G. Krishna, T. Morita, A. Russo, S. Sutou, "*In vivo* Rodent Erythrocyte Micronucleus Assay. II. Some Aspects of Protocol Design Including Repeated Treatments, Integration With Toxicity Testing, and Automated Scoring," *Environmental and Molecular Mutagenesis*, 35:234-252, 2000.

Hayashi, M., J.T. MacGregor, D.G. Gatehouse, D.H. Blakey, S.D. Dertinger, L. Abramsson-Zetterberg, G. Krishna, T. Morita, A. Russo, N. Asano, H. Suzuki, W. Ohyama, D. Gibson, "*In vivo* erythrocyte micronucleus assay. III. Validation and regulatory acceptance of automated scoring and the use of rat peripheral blood reticulocytes, with discussion of non-hematopoietic target cells and a single dose-level limit test," *Mutation Research*, 627:10-30, 2007.

Heddle, J.A., S. Dean, T. Nohmi, M. Boerrigter, D. Casciano, G..R Douglas, B.W. Glickman, N.J. Gorelick, J.C. Mirsalis, H-J Martus, T.R. Skopek, V. Thybaud, K.R. Tindall, N. Yajima, "In vivo transgenic mutation assays", *Environmental and Molecular Mutagenesis*, 35:253-259, 2000.

Kasper, P., Y. Uno, R. Mauthe, N. Asano, G. Douglas, E. Matthews, M. Moore, L. Müller, M. Nakajima, T. Singer, G. Speit, "Follow-up testing of rodent carcinogens not positive in the standard genotoxicity testing battery: IWGT workgroup report," *Mutation Research*, 627:106-116, 2007.

Kenelly, J.C., R. Waters, J. Ashby, P.A. Lefevre, B. Burlinson, D.J. Benford, S.W. Dean, I deG. Mitchell, "*In vivo* rat liver UDS assay. In: Supplementary Mutagenicity Tests, UKEMS Recommended Procedures, ed Kirkland D.J. and Fox M., Cambridge University Press, pp 52-77, 1993

Kirkland, D.J., S. Pfuhler, D. Tweats, M. Aardema, R. Corvi, F. Darroudi, A. Elhajouji, H. Glatt, P. Hastwell, M. Hayashi, P. Kasper, S. Kirchner, A. Lynch, D. Marzin, D. Maurici, J-R. Meunier, L. Müller, G. Nohynek, J. Parry, E. Parry, V. Thybaud, R. Tice, J. van Benthem, P. Vanparys, P. White, "How to reduce false positive results when undertaking *in vitro* genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of the ECVAM workshop," *Mutation Research*, 628:31-55, 2007.
.
Kirsch-Volders, M., T. Sofuni, M. Aardema, S. Albertini, D. Eastmond, M. Fenech, M. Ishidate, S. Kirchner, E. Lorge, T. Morita, H. Norppa, J. Surralles, A. Vanhauwaert, A. Wakata, "Report from the *in vitro* micronucleus assay working group", *Mutation Research*, 540:153-163, 2003.

Kissling, GE., S.D. Dertinger, M. Hayashi, J.T. MacGregor., "Sensitivity of the erythrocyte micronucleus assay: Dependence on number of cells scored and inter-animal variability", *Mutation Research* 634:235–240, 2007.

Ku, W.W., A. Bigger, G. Brambilla, H. Glatt, E. Gocke, P.J. Guzzie, A. Hakura, M. Honma, H-J. Martus, R.S. Obach, S. Roberts, "Strategy for genotoxicity testing-Metabolic considerations,"*Mutation Research*, 627:59-77, 2007.

MacGregor J.T., M.E. Bishop, J.P. McNamee, M. Hayashi, N. Asano, A. Wakata, M. Makajima, J. Saito, A. Aidoo, M.M. Moore, S.D. Dertinger, "Flow Cytometric analysis of micronuclei in peripheral blood reticulocytes: II. An efficient method of monitoring chromosomal damage in the rat", *Toxicological Sciences,* 94:92-107, 2006.

Moore, M. M., M. Honma, J. Clements, G. Bolcsfoldi, B. Burlinson, M. Cifone, J. Clarke, R. Delongchamp, R. Durward, M. Fellows, B. Gollapudi, S. Hou, P. Jenkinson, M. Lloyd, J. Majeska, B. Myhr, M. O'Donovan, T. Omori, C. Riach, R. San, L. F. Stankowski, Jr., A. K. Thakur, F. Van Goethem, S. Wakuri, I. Yoshimura, "Mouse Lymphoma Thymidine Kinase Gene Mutation Assay: Follow-up Meeting of the International Workshop on Genotoxicity Testing–Aberdeen, Scotland, 2003–Assay Acceptance Criteria, Positive Controls, and Data Evaluation," *Environmental and Molecular Mutagenesis*, 47:1-5, 2006.

Müller, L, P. Kasper, “Human biological relevance and the use of threshold arguments in regulatory genotoxicity assessment: Experience with pharmaceuticals.” *Mutation Research,* 464: 9-34, 2000.

OECD Guidelines for Genetic Toxicology (1997)

Scott, D., S.M. Galloway, R.R. Marshall, M. Ishidate Jr., D. Brusick, J. Ashby, B.C. Myhr, “Genotoxicity under extreme culture conditions. A report from ICPEMC Task Group 9”, *Mutation Research*, 257:147-204, 1991.

Storer, R.D., T.W. McKelvey, A.R. Kraynak, M.C. Elia, J.E. Barnum, L.S. Harmon, W.W. Nichols, J.G. DeLuca, "Revalidation of the *in vitro* alkaline elution/rat hepatocyte assay for DNA damage: improved criteria for assessment of cytotoxicity and genotoxicity and results for 81 compounds," *Mutation Research*, 368:59-101, 1996.

Suzuki, H., N. Ikeda, K. Kobayashi, Y. Terashima, Y. Shimada, T. Suzuki, T. Hagiwara, S.Hatakeyama, K.Nagaoka, J. Yoshida, Y. Saito, J. Tanaka, M. Hayashi, “Evaluation of liver and peripheral blood micronucleus assays with 9 chemicals using young rats A study by the Collaborative Study Group for the Micronucleus Test (CSGMT)/Japanese Environmental Mutagen Society (JEMS) - Mammalian Mutagenicity Study Group (MMS),” *Mutation Research,* 583: 133-145, 2005

Thybaud, V., M. Aardema, J. Clements, K. Dearfield, S. Galloway, M. Hayashi, D. Jacobson-Kram, D. Kirkland, J. T. MacGregor, D. Marzin, W. Ohyama, M. Schuler, H. Suzuki, E. Zeiger, "Strategy for genotoxicity testing: Hazard identification and risk assessment in relation to *in vitro* testing," *Mutation Research*, 627:41-58, 2007.

Tweats, D. J., D. Blakey, R. H. Heflich, A. Jacobs, S. D. Jacobsen, T. Morita, T. Nohmi, M. R. O'Donovan, Y. F. Sasaki, T. Sofuni, R. Tice, "Report of the IWGT working group on strategies and interpretation of regulatory *in vivo* tests. I. Increases in micronucleated bone marrow cells in rodents that do not indicate genotoxic hazards," *Mutation Research*, 627:78-91, 2007.

Tweats, D. J., D. Blakey, R. H. Heflich, A. Jacobs, S. D. Jacobsen, T. Morita, T. Nohmi, M. R. O'Donovan, Y. F. Sasaki, T. Sofuni, R. Tice, "Report of the IWGT working group on strategy/interpretation of regulatory *in vivo* tests. II. Identification of *in vivo*-only positive compounds in the bone marrow micronucleus test," *Mutation Research*, 627:92-105, 2007.

Wakata, A., Y. Miyamae, S. Sato, T. Suzuki, T. Morita, N. Asano, T. Awogi, K. Kondo, M. Hayashi, "Evaluation of the Rat Micronucleus Test with Bone Marrow and Peripheral Blood: Summary of the 9th Collaborative Study by CSGMT/JEMS • MMS," *Environmental and Molecular Mutagenesis*, 32:84-100, 1998.