

1.0 EXECUTIVE SUMMARY

Induced chromosomal and gene mutations cause genetic diseases, birth defects and other disease conditions and play a role in carcinogenesis. While it is widely accepted that *in vivo* mutation assays are more relevant to the human condition than *in vitro* assays, our ability to evaluate mutagenesis *in vivo* in a broad range of tissues has historically been quite limited. The development of transgenic rodent (TGR) mutation models has given us the ability to detect, quantify and sequence mutations in a range of somatic and germ cells.

This document provides a comprehensive review of the TGR mutation assay literature and assesses the potential use of these assays in a regulatory context. The information is arranged as follows.

1. *TGR mutagenicity models and their use for the analysis of gene and chromosomal mutation are fully described.* The TGR mutation assay is based on transgenic rats and mice that contain multiple copies of chromosomally integrated plasmid and phage shuttle vectors that harbour reporter genes for the detection of mutation. Mutagenic events arising in a rodent are scored by recovering the shuttle vector and analysing the phenotype of the reporter gene in a bacterial host. TGR gene mutation assays allow mutations induced in a genetically neutral transgene to be scored in any tissue of the rodent and therefore circumvent many of the existing limitations associated with the study of *in vivo* gene mutation. TGR models for which sufficient data are available to permit evaluation include Muta™Mouse, Big Blue® mouse and rat, *lacZ* plasmid mouse and *gpt* delta mouse and rat. Mutagenesis in the TGR models is normally assessed as a mutant frequency; if required, however, molecular analysis can provide additional information.
2. *The principles underlying current Organisation for Economic Co-operation and Development (OECD) tests for the assessment of genotoxicity in vitro and in vivo, as well as non-transgenic assays available for the assessment of gene mutation, are described.* OECD guidelines exist for a range of *in vitro* mutation assays that are capable of detecting both chromosomal and gene mutations. *In vivo* assays are required components of a thorough genetic toxicity testing programme. For somatic cells, OECD guidelines are currently available only for assays capable of assessing induced chromosomal mutation. In addition, there are non-transgenic assays that can be used for analysis of gene mutation; none of these have an OECD test guideline. Existing *in vivo* assays are limited by a range of different factors, including cost of the assay, the number of tissues in which genotoxicity may be measured, the state of understanding of the endpoint and the nature of the chemicals that will be detected.
3. *All available information pertaining to the conduct of TGR assays and important parameters of assay performance are tabulated and analysed.* As of December 2007, 228 agents have been evaluated using TGR assays. The majority of experimental records have assessed a subset of these chemicals, most of which are strong mutagens and carcinogens. Of the 154 agents whose carcinogenicity has been evaluated, 118 are carcinogens and 36 are non-carcinogens (including 13 chemicals routinely used as vehicle controls). These numbers represent significant increases over those in the

previously published version of this paper.¹ The following conclusions may be drawn from the existing TGR mutation data:

- The ability to use all routes of administration has been demonstrated. Experiments can be tailored to use the most relevant route of administration.
- The ability to examine mutation in virtually all tissues has been demonstrated. TGR assays have most commonly examined mutagenicity in the liver and bone marrow.
- The majority of the experiments have used shorter administration times than is currently recommended by the International Workshop on Genotoxicity Testing (IWGT). There are limited data available to assess the effects of longer sampling time, except with extremely short administration times. However, two recent studies on weak mutagens confirm the robustness of the recommended protocol.
- Although it is recognised that a number of factors may influence the tissue specificity of mutation, including cell turnover, deoxyribonucleic acid (DNA) repair, toxicokinetics and the nature of the genetic target, there are currently limited experimental data specific to transgenes that are available to inform the discussion.
- Limited data are available to evaluate the results of TGR assays in known target tissues for carcinogenicity. A case-by-case analysis of instances in which discrepancies are apparent suggests that in the majority of cases, factors such as non-genotoxic mechanism of action, inappropriate mode of administration or inadequate study design may account for the observed negative result in the tissue of interest.
- Qualitatively similar results have been obtained in the majority of experiments that have assayed different transgenes using similar experimental parameters.
- The spontaneous mutant frequency in most somatic tissues of transgenic rats and mice is five- to ten-fold higher than that observed in available endogenous loci using the same animals. Factors such as the age of the animal, the tissue and the animal model influence the absolute value of the spontaneous mutant frequency. In most somatic tissues, with the exception of the brain, there is an age-related increase in mutant frequency throughout the life of the animal. Most, but not all, studies suggest that the spontaneous mutant frequency in male germline tissues remains low and constant throughout the life of the animal.
- Multiple treatments of mutagen appear to increase mutant frequencies in neutral transgenes in an approximately additive manner. However, extremely long treatment times (12 weeks or longer) may produce an apparent increase in mutant frequency through clonal expansion, genomic instability in developing preneoplastic foci or tumours or oxidative damage of DNA resulting from chronic induction of cytochrome P-450 mono-oxygenases.
- The time required to reach the maximum mutant frequency is tissue specific and appears to be related to the turnover time of the cell population. The optimal sampling time differs according to tissue, with liver and bone marrow at opposite extremes among proliferating somatic tissues: in bone marrow, the mutant frequency appears to reach a maximum at extremely short sampling times and then decreases over 28 days following an acute treatment; in liver, the induced mutant frequency increases over the month following exposure, reaches a maximum and remains relatively constant thereafter. There are insufficient data

¹ In Lambert *et al.* (2005), 163 agents had been evaluated by July 2004; of the 103 agents whose carcinogenicity had been evaluated, 90 were carcinogens and 13 were non-carcinogens.

available for other tissues to support any conclusion regarding optimal sampling time.

- The results of studies carried out on a given chemical using similar experimental protocols suggest that the TGR assays show good qualitative reproducibility in both somatic and germ cells and quantitative reproducibility over a limited range of conditions and laboratories. The data are insufficient to allow conclusions to be drawn regarding the quantitative reproducibility of the assays over a wider range of protocol variations.
 - Although there exists a theoretical possibility that *ex vivo* and *in vitro* mutations may arise during the course of a TGR experiment, these types of mutations are expected to be extremely rare in a properly conducted experiment using the major TGR models. For positive selection systems, any such mutations will not be detected.
 - The weight of evidence suggests that transgenes and endogenous genes respond in approximately the same manner to mutagens in the few instances where direct comparisons are possible. Sensitivity is determined in large part by the spontaneous mutant frequency: the higher spontaneous mutant frequency in transgenes, compared with testable endogenous genes, reduces their sensitivity, especially when acute treatments are used. The sensitivity of transgenes can be enhanced by increasing the administration time.
 - Mutagens that induce deletions are likely to be detected more easily in certain endogenous genes than in transgenes as a result of phenotypic selection issues.
 - A very high proportion of the TGR experiments carried out to date have examined the activities of compounds that are known to be strong mutagens. A limited number of non-carcinogens have been evaluated with TGR assays, although the number has increased by almost 300% since publication of the first version of this review.
 - Molecular analysis of induced mutations in transgenic targets is possible and provides additional information in situations where high interindividual variation is observed and clonal expansion is suspected, when weak responses are obtained or when mechanistic information is desired. However, DNA sequence analysis of mutants is laborious and adds to the cost of the experiment; sequencing would not normally be required when testing drugs or chemicals for regulatory applications, particularly where a clear positive or negative result is obtained.
4. *The performance of TGR assays, both in isolation and as part of a battery of in vitro and in vivo short-term genotoxicity tests, in predicting carcinogenicity is described.* Analysis of the predictivity of TGR assays for carcinogenicity is hindered somewhat by the fact that TGR data are available for only a small number of non-carcinogens. Of the 118 carcinogens and 36 non-carcinogens that have been assessed using TGR assays, the following conclusions can be drawn regarding the predictivity and complementarity of TGR assays compared with a range of other OECD *in vitro* and *in vivo* genotoxicity tests:
- The TGR assay has high sensitivity and positive predictivity for carcinogenicity, meaning that most carcinogens have positive results in TGR assays and that there is a high probability that a chemical with a positive result in a TGR assay is a carcinogen.
 - As is the case with most genotoxicity assays, the TGR assay exhibits relatively low specificity and negative predictivity (compared with sensitivity and positive predictivity), meaning that relatively fewer non-carcinogens are negative in TGR

assays and that there is a lower probability that a chemical with a negative result in a TGR assay is a non-carcinogen.

- The positive and negative predictivities for the individual TGR and *Salmonella* assays were almost identical. Among the *in vivo* tests, the TGR assays were comparable with all except the comet assay, which had a slightly better combined predictivity. Among two-test combinations (*i.e.* a chemical could be predicted to be carcinogenic if the outcome of *either* short-term assay was positive), the TGR “or” cytogenetic assay gave the best combined positive predictivity, except for the TGR “or” comet assay combination, which was slightly better. Despite the lack of substantial increases in predictive values of the test batteries compared with the component assays alone, the test batteries had a much lower false-negative rate.
 - The TGR assay was usually positive for those carcinogens that were positive in both *Salmonella* and *in vitro* chromosomal aberration assays. In contrast, the *in vivo* micronucleus assay had a much higher false-negative rate for the same chemicals (0.22 vs. 0.15). If *in vivo* confirmation of positive results from both *Salmonella* and *in vitro* chromosomal aberration assays is warranted, the TGR assay is likely a better choice than the *in vivo* micronucleus assay.
 - The number of chemicals tested in the same assays is very small. Nevertheless, for chemicals having positive *Salmonella* and negative *in vitro* chromosomal aberration results (presumptive gene mutagens), selecting either the TGR assay or the *in vivo* micronucleus assay as the *in vivo* confirmation assay did not markedly affect the proportion of correct carcinogenicity predictions. The *in vivo* comet assay had a higher false-negative rate than the above two scenarios in which the *Salmonella* assay is positive and the *in vitro* chromosomal aberration assay is negative.
 - For chemicals having positive *in vitro* chromosomal aberration and negative *Salmonella* results (presumptive clastogens), selecting the *in vivo* micronucleus assay as the *in vivo* confirmation assay led to a higher proportion of correct carcinogenicity predictions than did selecting the TGR or *in vivo* comet assay. It should be noted that this observation is also based on data from only a few chemicals.
 - For those carcinogens with negative results in both *Salmonella* and *in vitro* chromosomal aberration assays, adding TGR assays to the test battery improved the overall predictivity over the *in vivo* micronucleus or the comet assay, since neither assay identified the carcinogens missed by the *in vitro* assays.
5. *A novel model for determining the performance of new in vivo gene mutation assays was introduced.* This model determines the ability of a new assay to accurately detect (*i.e.* positively predict) whether the assay will correctly identify *in vivo* gene mutagens. This method relies on the extensive DNA sequence data in the Transgenic Rodent Assay Information Database (TRAID) and is consistent with the OECD guidance document on validation of new test methods, which recommends that new tests be validated against the most biologically relevant endpoint. The analysis shows that the TGR assay has a near-perfect positive predictivity for gene mutation. The model is developed further to argue that a test battery designed to most optimally detect the endpoints of genotoxicity will, accordingly, be the most effective for detecting genotoxins that are carcinogenic.
6. *Recommendations are made regarding the experimental parameters for TGR assays and the use of TGR assays in a regulatory context.* Recommendations, based on

internationally harmonised criteria (IWGT), are made regarding the proper conduct of a TGR assay. These recommendations relate to accepted characteristics of a TGR mutation assay, treatment protocols and post-treatment sampling procedures. Of particular importance in optimising TGR protocols are two experimental variables – the administration time and the sampling time. Based on observations that mutations accumulate with each treatment, a repeated-dose regimen for a period of 28 days is strongly encouraged, with sampling at 3 days following the final treatment. If slowly proliferating tissues are of particular importance, then a longer sampling time may be more appropriate. Additional confidence in the recommended test protocol will be provided by research that examines the following:

- *The influence of the administration time on the observed mutant frequency for weak mutagens.* It has not conclusively been determined if data (especially negative results) from experiments using an administration time of less than 28 days should be discounted, if a 28-day treatment period is sufficiently long to permit the detection of weak mutagen-induced mutations in all tissues or if weak mutagens could in fact be detected using treatment times shorter than 28 days. Although more studies are required, two recent studies have shown that the recommended protocol, which uses a 28-day exposure period, is robust enough to detect weak mutagens.
- *The influence of the frequency of treatment on the observed mutant frequency.* The effects of weekly versus daily administrations on mutant frequency and on the ultimate conclusions of TGR experiments have not yet been thoroughly investigated.
- *The influence of sampling time following repeated administrations on the mutant frequency in both slowly and rapidly dividing tissue, particularly when examining weak mutagens.* At the current time, there are insufficient comparative data available for a range of tissues.

7. *Recommendations are made regarding how a TGR assay might be used within a new short-term test battery for assessing new compounds.* The test battery consists of various combinations of four assays – *Salmonella, in vitro* chromosomal aberration, *in vivo* micronucleus and TGR. This proposed strategy is based on the conclusions obtained from the predictivity analysis and the relative costs of the *in vivo* assays. An alternative approach to the above is described that shows how TGR assays could be used to resolve conflicts between *in vitro* and *in vivo* tests that are currently components of the standard genotoxicity test battery – *Salmonella, in vitro* chromosomal aberration and *in vivo* micronucleus. In situations where the standard test battery has been conducted and there are conflicting results – particularly in situations where the *Salmonella* assay has a positive result but the *in vivo* micronucleus assay is negative – the TGR assay may be conducted as an additional test to resolve the conflict. Recommended test strategies are based on an analysis of the existing data. Confidence in these recommendations would be enhanced by additional experimental data in the following areas:

- TGR data for additional non-carcinogens to increase the proportion of non-carcinogens in the dataset;
- additional testing to fill data gaps for chemicals having a known TGR assay but missing data from the *Salmonella, in vitro* chromosomal aberration or *in vivo* micronucleus assay;
- the testing of additional chemicals using an accepted test guideline for TGR mutation assays.

This extensive review fulfils the requirements for the preparation of an OECD Detailed Review Paper according to the OECD *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals* (OECD, 2006). Based on the extensive information and analyses in this review, there is sufficient evidence to support the recommendation that OECD undertake the development of a Test Guideline on Transgenic Rodent Gene Mutation Assays. Accordingly, it is recommended that OECD establish an Expert Working Group to develop such a Test Guideline and serve as an international forum for undertaking any additional research that would lead to the development of a fuller understanding of the variables surrounding the conduct of TGR mutation assays.