Biological Sciences in Space, Vol.19 No.4 (2005): 237-241

Mutation induction after low-dose carbon-ion beam irradiation of frozen human cultured cells

Yukihiro Umebayashi[1], Masamitsu Honma[2], Tomoko Abe[3], Hiromichi Ryuto[3], Hiromi Suzuki[4], Toru Shimazu[4], Noriaki Ishioka[5], Masaya Iwaki[1] and Fumio Yatagai[1,5]

[1] Advanced Development and Support Center, RIKEN Institute, Wako-shi, Saitama 351-0198, Japan
[2] Division of Genetics and Mutagenesis, National Institute of Health Sciences, Setagaya-ku, Tokyo 158-8501, Japan
[3] Accelerator Research Program, Frontier Research system, RIKEN Institute, Wako-shi, Saitama 351-0198, Japan
[4] Space Utilization Promotion Department, Japan Space Forum, Chiyoda-ku, Tokyo 100-0004, Japan
[5] Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, Tsukuba-shi, Ibaraki 305-8505, Japan

Abstract To study the genetic influence of low-dose ionizing radiation at the chromosomal level, frozen human lymphoblastoid TK6 cells were exposed to a 10 cGy dose delivered by a carbon-ion (24.5 ± 2.0 keV/μm) beam. Mutation assays were performed within a few days or after about one month of preservation at –80°C following irradiation. The results showed an increase in mutation frequencies at the *thymidine kinase (TK)* gene locus of 1.6-fold (2.5×10^{-6} to 3.9×10^{-6}) and 2.1-fold (2.5×10^{-6} to 5.3×10^{-6}), respectively. The relative distributions of the observed mutation classes were not changed by the radiation exposure in either assay. Multilocus analysis using two *TK* locus markers and eleven microsatellite loci spanning chromosome 17 were used for the analysis of TK mutants exhibiting a loss of heterozygosity (LOH). An interesting characteristic was detected using this system; interstitial deletion patterns were observed in hemizygous LOH mutants, which were specific for radiation exposure, and considered to be the result of end-joining repair of carbon-ion-induced DNA double-strand breaks. The relative increase in TK mutation frequency of the irradiated cells after the longer preservation at –80°C is probably due to the lower cell-viability compared to the unirradiated level. These results clearly demonstrate that this type of analysis can be used for the detection of low-dose ionizing radiation effects in frozen cells.

Key words: loss of heterozygosity (LOH), thymidine kinase (TK), frozen cells, carbon-ion beam

Introduction

The genetic influences of low-dose or low-dose rate ionizing radiation at the cellular level are important. From the viewpoint of human radiation risk, this kind of information could contribute to the estimation of such risks for persons working in special environmental conditions such as air line crews, radiation workers in the medical and industrial fields, and workers in nuclear power plants. Genetic alterations such as gene mutations, chromosome aberrations, and micronuclei induction are frequently used as indicators of radiation-exposure effects. These indicators were useful in recent studies concerning the bystander effect and the adaptive response, where low-dose or low-dose rate effects were critical parameters of radiation exposure (Bonner, 2003; Hei *et al.*, 2004; Mitchell *et al.*, 2004; Shao *et al.*, 2004). Since genetic consequences after radiation exposure must reflect cellular responses to radiation, these types of studies can also provide important insights for understanding the mechanisms of repair and mutagenesis after low-dose exposure.

For convenience in the molecular analysis of mutations located in a specific gene as well as effects at the chromosome level, a new methodology using human lymphoblastoid TK6 cells has been developed which has been successful in detecting low-dose radiation effects (Morimoto *et al.*, 2002a; Morimoto *et al.*, 2002b). Irradiation with 10 cGy of accelerated carbon-ions increased the mutation frequency at the heterozygous *thymidine kinase (TK)* locus from 5.7×10^{-6} to 17.9×10^{-6}, approximately a 3-fold increase over the spontaneous background level (Morimoto *et al.*, 2002b). This increase was due to the enhancement of TK mutants exhibiting a hemizygous-type loss of heterozygosity (LOH). Surprisingly, the radiation-specific LOH patterns were observed in more than half of these hemizygous-type LOH events. Such LOH mutants can be considered to be the result of end-joining repair of DNA double-strand breaks (DSBs). It also appears that DSBs are more efficiently formed by carbon-ion irradiation than by the same low-dose (10 cGy) X-ray irradiation, resulting in the relevant characteristics of LOH events described previously (Morimoto *et al.*, 2002a; Morimoto *et al.*, 2002b).

Article ID: 051904032
Original paper
Received Dec 26, 2005
Accepted Feb 13, 2006

Address for correspondence: Dr. Fumio Yatagai
Advanced Development and Support Center,
RIKEN Institute, Wako-shi, Saitama 351-0198, Japan
E-mail: yatagai@postman.riken.go.jp

To extend this line of study, the work described here was designed to see if radiation-specific LOH events in TK6 could be detected after the low-dose irradiation of frozen cells. If the detection of such events is possible after the thawing and growth of irradiated frozen-cells, this would suggest that DSBs are formed in the chromosomes of frozen cells leading to the induction of LOH events. If so, the LOH analysis system described here might be sensitive enough to detect genetic alterations caused by DSBs, which are less efficiently produced under the frozen conditions which permit almost no observation of indirect radiation effects mediated by free radicals. This could also provide the basis for a system to obtain information about the genetic influences of space environmental radiation by using a frozen cell system in future experiments at International Space Station. The work described here reports the success of an attempt to detect LOH events using frozen cells exposed to low dose irradiation.

Materials and Methods

A new methodology for LOH analysis using human lymphoblastoid TK6 cells has been established by Honma's group (Honma *et al.*, 2000) and successfully applied for the detection of LOH events after this kind of low-dose radiation effect (Morimoto *et al.*, 2002a; Morimoto *et al.*, 2002b). The methodology used here for the detection of TK-deficient mutants, the cell line used, incubation medium and conditions has already been reported in the above papers. Briefly, TK6 cells were incubated in RPMI1640 medium supplemented with HAT to eliminate pre-existing TK^- mutants. The cells (2×10^6) were suspended in 1 ml of fresh medium containing 10% DMSO and frozen in the cryogenic vials of diameter 10 mm at –80°C. The vials were exposed to 10 cGy of a carbon-ion beam (135 MeV/u) produced by the RIKEN Ring Cyclotron. During irradiation, the vials were cooled with a dry ice and ethanol bath. The cell irradiation condition was exactly same as the previous report (Morimoto *et al.*, 2002b) except the sample container as described above. The LET of carbon-ion at the sample site can be estimated to be 24.5 ± 2.0 keV/μm. Then the cells were stored at –80°C until the assays were performed. TK^- mutant clones were isolated on the basis of their resistance to 4 μg/ml trifluorothymidine (TFT). The early TK^- mutant clones (EM) and the late mutant clones (LM) were selected after two and four weeks incubation, respectively. The selected TK^- mutants were first analyzed to determine the loss of *TK* heterozygosity (LOH) with a PCR analysis of the exon 4 and 7 regions in the *TK* locus. This PCR analysis also can distinguish two types of LOH, homozygous and hemizygous, in which the functional *TK* allele is replaced by the mutated TK^- allele or lost, respectively. These two types of LOH, hemizygous and homozygous, can be considered to be the result of end-joining repair or homologous recombination repair of DNA double-strand breaks (DSBs), respectively. For additional determination of the extent of deleted or substituted portions of the chromosome in the LOH mutants, 11 microsatellites regions (D17S588, D17S1784, D17S785, D17S789, D17S802, D17S807, D17S928, D17S932, D17S1299, D17S1566 and THRA) were analyzed on chromosome 17 using multiple PCR reactions as described previously (Yatagai *et al.*, 2004).

Results

Table 1 represents the average values of the plating efficiency (PE) in three independent experiments. The assays which were performed within a few days after irradiation, showed almost no difference in the PE values between the two samples; unirradiated and irradiated. The assays performed after one month of post-irradiation storage showed a higher PE value for the unirradiated sample when compared to that for the 10 cGy irradiated sample, but this difference was not statistically significant. These measurements demonstrated that this level of low-dose irradiation delivered to frozen cells produced little effect on cell viability. The average values of the TK^- mutation frequency (MF) in three independent

Table 1 Measurements of plating efficiencies (PE) and thimidine kinase (TK) mutation frequencies (MF)

Sample		PE (%)	EM* ($\times 10^{-6}$)	LM* ($\times 10^{-6}$)	Total MF ($\times 10^{-6}$)
After several days of storage at -80°C	Unirradiated	52 ± 5.2 (60, 53, 42)	0.61 ± 0.31 (0, 1.0, 0.83)	1.9 ± 0.45 (2.1, 1.0, 2.5)	2.5 ± 0.42 (2.1, 2.0, 3.3)
	10 cGy Irradiation	53 ± 4.6 (62, 47, 50)	1.2 ± 0.29 (0.69, 1.7, 1.1)	2.8 ± 0.71 (1.7, 2.5, 4.1)	3.9 ± 0.82 (2.4, 4.2, 5.2)
	T-test**	0.77	0.27	0.3	0.13
After about one month of storage at -80°C	Unirradiated	73 ± 9.1 (78, 85, 55)	0.42 ± 0.15 (0.22, 0.31, 0.72)	2.0 ± 0.47 (1.1, 2.4, 2.6)	2.5 ± 0.56 (1.4, 2.7, 3.3)
	10 cGy Irradiation	60 ± 7.2 (53, 74, 52)	1.4 ± 0.4 (1.8, 1.0, 1.4)	3.9 ± 0.87 (2.2, 5.0, 4.6)	5.3 ± 0.67 (4.0, 6.0, 6.0)
	T-test**	0.18	0.081	0.049	0.006

* EM and LM represent the TK^- mutation frequencies for early mutants (EM) and late mutants (LM), respectively (see text). **P-values were calculated with the T-test.

Table 2 Classification of TK⁻ mutants using PCR analysis of the *TK* gene

Mutation Class		After several days of storage at -80°C				After about one month of storage at -80°C			
		Unirradiated		10cGy Irradiation		Unirradiated		10cGy Irradiation	
		Number (EM, LM)	MF* ($\times 10^{-6}$)	Number (EM, LM)	MF* ($\times 10^{-6}$)	Number (EM, LM)	MF* ($\times 10^{-6}$)	Number (EM, LM)	MF* ($\times 10^{-6}$)
Non LOH		6 (4, 2)	0.6 [24%]	15 (9, 6)	1.1 [28%]	7 (6, 1)	0.6 [25%]	14 (12, 2)	1.4 [26%]
LOH	Homo-zygous	8 (0, 8)	0.8 [32%]	15 (0, 15)	1.1 [28%]	10 (0, 10)	0.9 [36%]	21 (0, 21)	2.1 [40%]
	Hemi-zygous	11 (4, 7)	1.1 [44%]	24 (13, 11)	1.7 [44%]	11 (3, 8)	1.0 [39%]	18 (7, 11)	1.8 [34%]
Total		25 (8, 17)	2.5 [100%]	54 (22, 32)	3.9 [100%]	28 (9, 19)	2.5 [100%]	53 (19, 34)	5.3 [100%]

* The relative percentage of the MFs among the individual sample groups are shown in the brackets: [ ].

experiments are also shown in Table 1. In both assays, performed within several days of each other, and about one month after irradiation, total TK⁻ MFs (EM+LM) showed an increase after the 10 cGy carbon-ion irradiation of 1.6-fold (2.5×10^{-6} to 3.9×10^{-6}) and 2.1-fold (2.5×10^{-6} to 5.3×10^{-6}), respectively. Such increases were not statistically significant in the early assay (several days after irradiation), but were significant for the one month post-irradiation sample. The relative increase in TK mutation frequency after longer preservation of irradiated cells at –80°C is probably due to the relatively lower viability of irradiated cells (PE=60%) compared to the unirradiated case (PE=73%), but this increase does not mask the radiation exposure effect as described below. The 2.0- and 3.3-fold enhancements in the *TK*⁻ mutation frequencies for EM in the cases of a few days and one month, respectively, were also statistically not significant. However, the characteristic features in the classification of TK mutants and the characterization of LOH mutants were observed in both cases as described below.

The mutation classes were determined by PCR analyses for the exon 4 and 7 regions in the *TK* gene locus of the TK mutants. The distribution of mutation

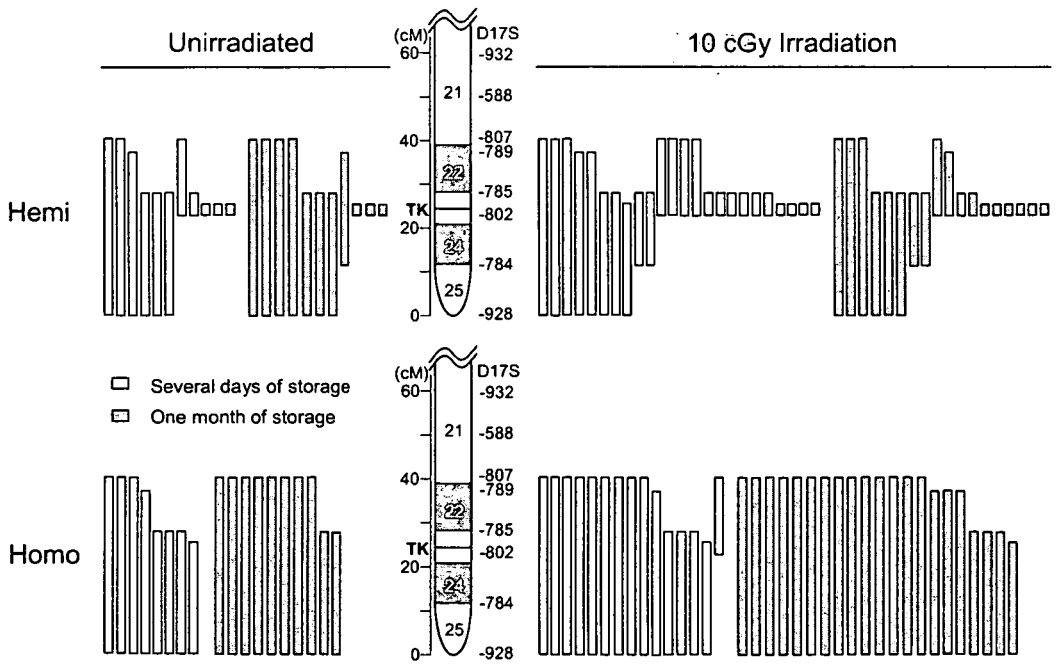

Fig. 1. Extent of hemizygous and homozygous LOH events on chromosome 17 in TK-deficient mutants obtained after a 10 cGy exposure to carbon-ions.

Table 3 Detailed characteristics of LOH-type events defined by chromosome mapping

Type of LOH		After several days of storage				After about one month of storage			
		Unirradiated		10 cGy Irradiation		Unirradiated		10cGy Irradiation	
		Number	(EM, LM)	Number	(EM, LM)	Number	(EM, LM)	Number	(EM, LM)
Homo-zygous	Type 1	8	(0, 8)	14	(0, 14)	10	(0, 10)	9	(0, 9)
	Type 2	0	(0, 0)	1	(0, 1)	0	(0, 0)	0	(0, 0)
Hemi-zygous	Type 1	6	(0, 6)	8	(0, 8)	7	(0, 7)	2	(0, 2)
	Type 2	5	(4, 1)	16	(13, 3)	4	(3, 1)	12	(7, 5)
	[Large]*	[2	(1, 1)]	[12	(10, 2)]	[1	(0, 1)]	[6	(2, 4)]

* Larger size deletion event extending past the region delimited by TK and D17S802 (see text)

classes for the analyzed mutants are shown in Table 2, together with the corresponding mutation frequencies (MF). The mutation class "non-LOH" means that the two exon regions, 4 and 7 in *TK* gene locus in the *TK*$^-$ mutant can be amplified by PCR in a manner similar to the same regions in the parental TK heterozygous cell. Taking into consideration both, the results of previous analyses and the principle of the mutation analysis used here, most of non-LOH mutants could be point mutations, although the DNA base sequences in the *TK* gene of these mutants were not determined in this study. As described above, homozygous and hemizygous types of LOH mutations represent cases where the functional *TK* allele is replaced by a mutated *TK*$^-$ allele or is lost, respectively. The corresponding MFs in all of these three classes increased after the radiation exposure, and the variation of these increased rates among these classes were not large after storage periods of a few days or one month. As a result, no specific change in the distribution of mutation classes was detected after radiation exposure and storage at –80°C for up to one month after irradiation.

As a result of chromosome mapping analyses, the characteristic changes in the fine structure of the LOH patterns in the analyzed *TK*$^-$ mutants were observed after low-dose carbon-ion exposure. A type 1 LOH event is a so-called terminal event, and defined as a pattern in which the deleted or exchanged chromosome segment extends to the telomere marker (D17S928). In a type 2 event this exchanged segment does not reach the marker. A "large" type 2 hemizygous event was defined as one where changes are found which extend past the limited region or neighborhood of the deletion (*TK* and its nearest marker D17S802) (Table 3). Type 2 hemizygous events (interstitial deletions), especially ones with a large size deletion, were more frequently observed after radiation exposure. For a better depiction of the extent of LOH mutants, the deleted or replaced regions of chromosome 17 in each LOH mutant are shown as bars in Fig.1. This figure more clearly represents the above characteristic estimated from Table 3. In addition, this characteristic was essentially the same in the assays performed at both times, after about a few days and about one-month of storage at –80°C after irradiation.

Discussion

As can be seen in Table 1, with the level of low-dose irradiation used here, 10 cGy of radiation delivered to frozen cells, there was little effect on cell viability. This result showing almost no toxic effect on cell viability appears reasonable in view of previous experimental results with carbon-ion irradiation using the same dose under normal incubation conditions with the cells in suspension culture (Morimoto *et al.*, 2002b). The mutation frequency for the unirradiated control cells measured in this study is lower than the level of background mutations reported in previous experiments (5.7×10^{-6}). This is probably due to differences in the procedures using HAT medium; in particular, a shorter incubation period for the cells was used in the earlier studies. The increase in the TK$^-$ mutation frequency was not statistically significant in the assays performed after a few days of storage in a –80°C freezer following the irradiation. Although the TK mutation frequency increased after irradiation, the relative distributions of mutation classes were not changed in assays performed after a few days or after about one month of storage. However, the molecular nature of the LOH events revealed an interesting characteristic which appears to be specific for mutations found after radiation exposure. The occurrences of type 2 hemizygous LOH events (interstitial deletions) were clearly enhanced by the radiation exposure in both assays (Table 3). These enhancements become more obvious if the comparison is restricted to the large size deletions. This "large" event was frequently observed following the same 10 cGy dose delivered by a carbon-ion beam under suspension culture conditions in previous experiments (Morimoto *et al.*, 2002b). This event might reflect the effect of carbon-ion exposure, because DSBs created by carbon-ion traversal through the cell results in this type 2 hemizygous event via a non-homologous end-joining repair (NHEJ) process for DSBs. This 10 cGy irradiation can be estimated to provide a 1.1 ion traversal on the average through the nuclear region of a cell, and approximately 10 DSBs per cell (Taucher-Scholz *et al.*, 1999; Ward, 1988). This means that the proportion of cell having at least, a single DSB on chromosome17 is less than about one fifth of

the total cells. However, this kind of estimation reflects the high efficiency of DSB repair and lower production of mutational events even after taking into consideration that the DSBs occurring around the restricted *TK* gene locus and its neighboring region lead to LOH events more efficiently than those in other region. In fact, we have already succeeded to suggest the high efficiency of DSB repair through NHEJ using this cell line (Honma *et al.*, 2003). Regardless of radiation exposure, type 1 homozygous or hemizygous LOH events were observed as LM. It should also be emphasized that this characteristic event can be detected after about a one-month preservation of cells at –80°C. In other words, radiation effects are probably not erased during a long preservation period while frozen. Interestingly, only a single type 2-event for homozygous LOH was found among in the mutants irradiated with 10 cGy (Table 3 and Fig.1). This event might reflect the effects of radiation exposure, because this specific event was also observed at a very low frequency after low-dose rate γ-ray irradiation (data, not shown).

The present experiments were performed within the low-dose range using the accelerated charged particles. An interesting question is whether the observed increase in the TK mutation frequency after this kind of low-dose radiation exposure might reflect the bystander effect, because a large proportion, close to 33% of cells are considered to be not hit by the carbon-ion due to the poisson distribution of average number of nuclear hit 1.1. Such possibility seems to be not high, because we recently obtained the following unpublished data: the TK mutation frequency of unirradiated cells were found to be rather decreased if they are cultured by the medium used for the carbon-ion beam irradiation (2Gy) of cells. The present observations also suggest that this mutational analysis can be used for estimation of the influence of low-dose IR, especially heavy ions. Chromosome damage caused by the low-dose and/or low-dose rate radiation exposure can accumulate in frozen cells. This is one of the reasons why this frozen-cell system is radiation-sensitive, although indirect radiation-effects, for example mediated by free radicals, seem to be limited in this system. Astronauts will be exposed to space radiation, and it is necessary to estimate the risk of such exposures. Heavy ions are known to be included in the important components of space radiation. Taking this into consideration together with the possible long-term preservation of irradiated cells as discussed above, it should be possible to detect low-dose exposure effects by using this system with the –80°C storage of irradiated frozen cells. In fact, preparations are under way for space studies which are being planned to analyze mutations found after the recovery of cells from a long (about 3 months) preservation period in the International Space Station (ISS).

Acknowledgements

This study was partially supported by the Budget for Nuclear Research of the Ministry of Education, Culture, Sports, Science and Technology, and was reviewed by the Atomic Energy Commission.

References

Bonner, W.M. (2003) Low-dose radiation: thresholds, bystander effects, and adaptive responses. *Proc. Natl. Acad. Sci. U.S.A.*, **100(9)**, 4973-4975.

Hei, T.K., Persaud, R., Zhou, H. and Suzuki, M. (2004) Genotoxicity in the eyes of bystander cells. *Mutat. Res.*, **568(1)**, 111-120.

Honma, M., Izumi, M., Sakuraba, M., Tadokoro, S., Sakamoto, H., Wang, W., Yatagai, F. and Hayashi, M. (2003) Deletion, rearrangement, and gene conversion; genetic consequences of chromosomal double-strand breaks in human cells. *Environ. Mol. Mutagen.*, **42(4)**, 288-298.

Honma, M., Momose, M., Tanabe, H., Sakamoto, H., Yu, Y., Little, J.B., Sofuni, T. and Hayashi, M. (2000) Requirement of wild-type p53 protein for maintenance of chromosomal integrity. *Mol. Carcinog.*, **28(4)**, 203-214.

Mitchell, S.A., Marino, S.A., Brenner, D.J. and Hall, E.J. (2004) Bystander effect and adaptive response in C3H 10T(1/2) cells. *Int. J. Radiat. Biol.*, **80(7)**, 465-472.

Morimoto, S., Kato, T., Honma, M., Hayashi, M., Hanaoka, F. and Yatagai, F. (2002a) Detection of Genetic Alterations Induced by Low-Dose X Rays: Analysis of Loss of Heterozygosity for *TK* Mutation in Human Lymphoblastoid Cells. *Radiat. Res.*, **157(5)**, 533-538.

Morimoto, S., Honma, M. and Yatagai, F. (2002b) Sensitive detection of LOH events in a human cell line after C-ion beam exposure. *J. Radiat. Res. (Tokyo)*, **43 (Suppl.)**, S163-167.

Shao, C., Folkard, M., Michael, B.D. and Prise, K.M. (2004) Targeted cytoplasmic irradiation induces bystander responses. *Proc. Natl. Acad. Sci. U.S.A.*, **101(37)**, 13495-13500.

Taucher-Scholz, G. and Kraft, G. (1999) Influence of radiation quality on the yield of DNA strand breaks in SV40 DNA irradiated in solution. *Radiat. Res.*, **151(5)**, 595-604.

Yatagai, F., Morimoto, S., Kato, T. and Honma, M. (2004) Further characterization of loss of heterozygosity enhanced by p53 abrogation in human lymphoblastoid TK6 cells: disappearance of endpoint hotspots. *Mutat. Res.*, **560**, 133-145.

Ward, J.F. (1988) DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Prog. Nucleic. Acid. Res. Mol. Biol.*, **35**, 95-125.

Available online at www.sciencedirect.com
ScienceDirect
Mutation Research 627 (2007) 31–35
Genetic Toxicology and Environmental Mutagenesis
www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

Fourth International Workgroup on Genotoxicity testing: Results of the in vivo Comet assay workgroup

Brian Burlinson[a,*], Raymond R. Tice[b], Günter Speit[c], Eva Agurell[d], Susanne Y. Brendler-Schwaab[e], Andrew R. Collins[f], Patricia Escobar[g], Masamitsu Honma[h], Tirukalikundram S. Kumaravel[i], Madoka Nakajima[j], Yu F. Sasaki[k], Veronique Thybaud[l], Yoshifumi Uno[m], Marie Vasquez[n], Andreas Hartmann[o]

[a] *Huntingdon Life Sciences, Cellular & Molecular Toxicology, Woolley Road, Alconbury, Huntington, Cambs PE28 4HS, UK*
[b] *National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA*
[c] *Universitätsklinikum, Abteilung Humangenetik, Ulm, Germany*
[d] *Medical Products Agency, Uppsala, Sweden*
[e] *Federal Institute for Drugs and Medical Devices, Bonn, Germany*
[f] *Department of Nutrition, University of Oslo, Norway*
[g] *BioReliance, Invitrogen Bioservices, Rockville, MD, USA*
[h] *National Institute of Health Sciences, Tokyo, Japan*
[i] *Covance Laboratories Ltd., Harrogate, UK*
[j] *Bio Safety Research Center, Foods, Drugs and Pesticides, Shizuoka, Japan*
[k] *Hachinohe National College of Technology, Hachinohe, Japan*
[l] *Sanofi-Aventis, Vitry Sur Seine, France*
[m] *Mitsubishi Chemical Safety Institute Ltd., Shiba, Minato-ku, Tokyo, Japan*
[n] *Helix3 Inc., North Carolina, USA*
[o] *Safety Profiling & Assessment, Novartis Pharma AG, Basel, Switzerland*

Received 28 February 2006; received in revised form 31 July 2006; accepted 14 August 2006
Available online 22 November 2006

Abstract

As part of the Fourth International Workshop on Genotoxicity Testing (IWGT), held 9–10 September 2005 in San Francisco, California, an expert working group on the Comet assay was convened to review and discuss some of the procedures and methods recommended in previous documents. Particular attention was directed at the in vivo rodent, alkaline (pH >13) version of the assay. The aim was to review those protocol areas which were unclear or which required more detail in order to produce a standardized protocol with maximum acceptability by international regulatory agencies. The areas covered were: number of dose levels required, cell isolation techniques, measures of cytotoxicity, scoring of comets (i.e., manually or by image analysis), and the need for historical negative/positive control data. It was decided that a single limit dose was not sufficient although the required number of dose levels was not stipulated. The method of isolating cells was thought not to have a qualitative effect on the assay but more data were needed before a conclusion could be drawn. Concurrent measures of cytotoxicity were required with histopathological examination of tissues for necrosis or apoptosis as the "Gold Standard". As for analysing the comets, the consensus was that image analysis was preferred

* Corresponding author. Tel.: +44 1480 893235.
E-mail address: burlinsb@ukorg.huntingdon.com (B. Burlinson).

1383-5718/$ – see front matter
doi:10.1016/j.mrgentox.2006.08.011

but not required. Finally, the minimal number of studies required to generate a historical positive or negative control database was not defined; rather the emphasis was placed on demonstrating the stability of the negative/positive control data. It was also agreed that a minimum reporting standard would be developed which would be consistent with OECD in vivo genotoxicity test method guidelines.

Keywords: Single cell gel assay; Comet assay; DNA damage; Genotoxicity; Alkaline electrophoresis

1. Introduction

The Comet assay, also referred to as the single cell gel electrophoresis (SCG or SCGE) assay, is a rapid, visual, and quantitative technique for measuring DNA damage in eukaryote cells [1–7]. Under alkaline (pH >13) conditions, the assay can detect single and double-stranded breaks, incomplete repair sites, alkali labile sites, and also possibly both DNA–protein and DNA–DNA cross-links, in virtually any eukaryotic cell population that can be obtained as a single cell suspension.

As the Comet assay has gained in popularity as a standard laboratory technique for evaluating DNA damage and/or repair, the question of how it can be applied within the current regulatory strategy of genotoxicity testing has become a matter of debate [8]. The primary focus of interest has been on the alkaline (pH >13) version, as it is applied to *in vivo* genotoxicity testing strategies [6,8–11]. This is especially important now that acceptance of the *in vivo* Comet assay by regulatory agencies in a number of countries is growing, with some already citing it as an acceptable second test [12,13]. Part of the reason for this acceptance has been the development of a standard protocol and acceptance criteria for the assay through the IWGT working parties [6] and international Comet assay workshops [10]. The purpose of this meeting was to review the procedures and methods recommended in previous guidance documents [6,10], with particular attention being given to the *in vivo* rodent alkaline (pH >13) assay.

Prior to the actual IWGT session, the members of the working group were assigned to different subgroups with each subgroup responsible for reviewing a particular topic. At the IWGT meeting, the subgroups presented their conclusions and recommendations to the complete working group for consideration and discussion, with input from the audience. This report provides an overview of the topics discussed and the consensus reached by the working group with regard to the *in vivo* rodent alkaline (pH >13) Comet assay (hereafter designated as the *in vivo* Comet assay).

2. Discussion topics and recommendations

2.1. Multiple dose levels versus limit dose

For this topic, the discussions focused on the number of dose levels to be used in the *in vivo* Comet assay, especially for cases where there is no evidence of animal toxicity. For example, as stated in the Organisation for Economic Development and Co-operation (OECD) test guideline 474 (rodent bone marrow micronucleus test), a chemical which shows no sign of toxicity up to the limit dose of 2 g/kg need only be tested at that dose [14].

The consensus of the working group was that a single dose level was not sufficient even for substances that could be tested at the limit dose of 2 g/kg. The reasoning behind this consensus was that there were not yet sufficient data to conclude that downturns in the dose response curve (i.e., a bell shaped dose response curve) would not occur for some substances due, for example, to altered bioavailability at higher dose levels. The 'downturn phenomenon', was also a matter of discussion among the members of the IWGT *in vivo* micronucleus (MN) group [15], where this phenomenon has been shown to occur in some MN studies although the underlying mechanism(s) have yet not been identified. In such cases, positive responses occurred at the second highest dose level. Therefore, it was concluded that the use of a single dose level could lead to problems in data interpretation. There was also the feeling that positive responses at multiple dose levels could reinforce the biological relevance of the result.

2.2. Cell isolation process

The background behind this discussion point was the disparate *in vivo* rodent Comet assay data sets published about *ortho*-phenyl phenol. When tested by Sasaki et al. [16], *ortho*-phenyl phenol was positive in the mouse using stomach, liver, kidneys, lung, urinary bladder as target organs. However, when tested by Bomhard et al. [17] in the same species, *ortho*-phenyl phenol was negative in the tissues investigated. One possible explanation for the difference in results was how the tissues were pro-

cessed. In Sasaki et al. [16], isolated nuclei were used, whereas in Bomhard et al. [17], isolated whole cells were used. Although there was much discussion on this subject along with data from two groups which showed that the method of tissue processing (i.e., isolated cells versus isolated nuclei) did not have a qualitative effect on the comet response, it was decided that more data were needed before a conclusion could be made and that any international validation study should consider both processing methods.

2.3. Concurrent measures of cytotoxicity

Cell death is a process that leads to DNA degradation. Thus, all test methods that evaluate primary DNA damage, including the Comet assay, have the potential to detect agents that are cytotoxic rather than genotoxic. However, since DNA damage in the Comet assay is assessed at the level of the individual cell, it is possible in some cases to identify dead or dying cells by their specific image. Under alkaline conditions, necrotic or apoptotic cells can result in comets with small or non-existent head and large diffuse tails [18] as observed in *in vitro* studies following treatment with cytotoxic, non-genotoxic compounds [19–21]. However, such images may not be uniquely diagnostic for apoptosis or necrosis since they may also be detected after treatment with high doses of radiation or high concentrations of strong mutagens [22]. For the *in vivo* Comet assay, only limited data are available to establish whether cytotoxicity results in increased DNA migration in tissues of experimental animals. It was discussed that migration levels detected at the time of sampling are dependent on the tissue and the slope of the dose response for a particular tested compound. For some chemicals, despite the presence of necrosis or apoptosis in target organs such as kidneys [23], testes [24], and liver or duodenum [11], an increase in DNA migration was not observed. In contrast, enhanced DNA migration was seen in cells isolated from the livers of mice dosed with carbon tetrachloride under conditions that also resulted in necrosis, as determined from a histopathological examination [25]. It is also possible that at cytotoxic doses, a decrease in DNA migration may be detected due to the loss of heavily damaged or dying cells during sample processing and/or electrophoresis.

There was consensus on the need to include measures of cytotoxicity and to address the possible effects of cytotoxicity in comet data interpretation. The suggested methods included: a dye exclusion test for membrane integrity and metabolic competency [26] and determining the frequency of cells with low molecular weight DNA using the neutral diffusion assay [6,27]. The "Gold Standard" for assessing levels of necrosis and apoptosis when an *in vivo* Comet assay gave positive results was concluded to be histopathology. It was pointed out that there was a need to standardize ways to present histopathological findings.

2.4. Image analysis (IA) or manual scoring

A variety of commercial and freeware IA systems are available for assessing DNA migration in individual cells. In addition, manual scoring can also be used to determine the length of DNA migration, the percentage of cells with and without migration, or the proportion of comets that can be "binned" into various migration categories (generally one of five, from undamaged to maximally damaged depending upon the tail length) [28]. However, a limitation of this categorization method may be a potential inability to take into account the density or shape of tails which can include short but dense tails and long but sparse tails depending on the effects of compounds tested. With IA systems, the most common parameters analyzed are the percentage DNA in the tail (% tail DNA), tail moment, and tail length and/or image length (referring to nucleus plus migrated DNA). The percentage DNA in the tail is generally defined as the fraction of DNA in the tail divided by the amount of DNA in the cell multiplied by 100, while the tail length is the distance from the middle or the estimated perimeter of the comet head to the last visible signal in the tail. There are several measures of tail moment. The one most commonly used, called the Olive tail moment, is the product of the amount of DNA in the tail and the mean distance of migration in the tail [29]. It is important to note that some parameters (e.g., tail moment) may be calculated differently among IA systems and this can lead to quantitative differences, which can be problematic when comparing inter-laboratory data.

The consensus was that IA is preferred but not required. Heavily damaged cells exhibiting a specific microscopic image (commonly referred to as hedgehogs) consisting of small or non-existent head and large and diffuse tails [18] potentially represent dead or dying cells and may be excluded from data collection. However, determining their frequency may be useful for data interpretation. If IA is used, then % tail DNA appeared to be the most linearly related to dose and the easiest to intuitively understand [30]. However, there was no consensus that this IA measure of DNA migration must be the only one used. If some measure of tail moment is used, than % tail DNA and tail length data must be provided also. Data on the distribution of migration among

cells should also be presented. This is accomplished by sorting cells within "bins" based on the metric used to evaluate DNA migration and presenting the data as the percentage of cells within each bin.

2.5. Historical negative/positive control data

The minimal number of studies needed was not defined but enough studies need to be conducted to demonstrate the stability of the negative/positive control data. Criteria for determining the acceptability of new studies, based on historical control data, should be developed for each tissue by each laboratory. There was discussion on the background responses for negative controls and there was a consensus that negative controls should exhibit measurable DNA migration. However, there was no consensus as to how much mean DNA migration was needed among the control cells. It was recognized that the ability to detect chemicals that predominantly induce DNA cross-linking, damage that reduces the ability of the DNA to migrate, depends on the extent of average DNA migration in the control cells. Investigators who are attempting to detect such chemicals will need to demonstrate the adequacy of their *in vivo* Comet assay protocol for this purpose.

2.6. Minimal reporting standards

It was agreed that to ensure that all studies can be independently evaluated, a minimum reporting standard for regulatory submissions and publications will be developed. This standard will be consistent with OECD *in vivo* genetic toxicology test method guidelines. Previous publications have covered some aspects of protocol design and reporting [10,31].

2.7. Conclusions

In recent years, the *in vivo* Comet assay has become increasingly used for regulatory purposes and acceptance of the test method by regulatory agencies is growing (reviewed in [8]). However, several issues on study design and on data analysis and assessment that required further investigation remain and it was these issues that were discussed by the IWGT working group. In addition to guidance provided in previous published guidelines [6,10], consensus among the participants of the working group was reached with regards to the selection of the number of dose levels, the need to include concurrent measures of cytotoxicity in the studies, the adequacy of manual scoring, and the need to develop historical control data. Consensus was also reached on the need for an international validation study to stringently evaluate the reliability and accuracy of the *in vivo* Comet assay (as well as *in vitro* versions). This validation study would compare, among other protocol issues, test results obtained using isolated nuclei versus isolated whole cells from various tissues.

Acknowledgements

During the preparation of this report, the Japanese Center for the Validation of Alternative Methods (JaCVAM) announced that they were forming a study management team including participants from the European Centre for the Validation of Alternative Methods (ECVAM), the US Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), US National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), and the Mammalian Mutagenicity Study Group/Japanese Environmental Mutagen Society (MMS/JEMS) to conduct an international Comet assay validation study. This validation study is scheduled to start in late 2006 and will focus initially on the *in vivo* Comet assay, to be followed by the validation of various *in vitro* Comet assays. ECVAM has also implemented an initiative to evaluate the validity of the *in vitro* Comet assay.

References

[1] O. Ostling, K.J. Johanson, Microelectrophoretic study of radiation induced DNA damages in individual mammalian cells, Biochem. Biophys. Res. Commun. 123 (1984) 291–298.
[2] N.P. Singh, M.T. McCoy, R.R. Tice, E.L. Schneider, A simple technique for quantitation of low levels of DNA damage in individual cells, Exp. Cell Res. 175 (1988) 184–191.
[3] R.R. Tice, P.W. Andrews, O. Hirai, N.P. Singh, The single cell gel (SCG) assay: an electrophoretic technique for the detection of DNA damage in individual cells, in: C.R. Witmer, R.R. Snyder, D.J. Jollow, G.F. Kalf, J.J. Kocsis, I.G. Sipes (Eds.), Biological Reactive Intermediates. IV. Molecular and Cellular Effects and Their Impact on Human Health, Plenum Press, New York, 1991, pp. 157–164.
[4] V.J. McKelvey-Martin, M.H. Green, P. Schmezer, B.L. Pool-Zobel, M.P. De Meo, A. Collins, The single cell gel electrophoresis assay (comet assay): a European review, Mutat. Res. 288 (1993) 47–63.
[5] D.W. Fairbairn, P.L. Olive, K.L. O'Neill, The Comet assay: a comprehensive review, Mutat. Res. 339 (1995) 37–59.
[6] R.R. Tice, E. Agurell, D. Anderson, B. Burlinson, A. Hartmann, H. Kobayashi, Y. Miyamae, E. Rojas, J.-C. Ryu, Y.F. Sasaki, The single cell gel/comet assay: Guidelines for in vitro and in vivo genetic toxicology testing, Environ. Mol. Mutagen. 35 (2000) 206–221.
[7] G. Speit, A. Hartmann, The comet assay—a sensitive genotoxicity test for the detection of DNA damage and repair, in: D.S.

Henderson (Ed.), Methods in Molecular Biology, vol. 113, DNA Repair Protocols: Eucaryotic Systems, 2nd ed., Humana Press Inc., Totowa, NJ, 2006, pp. 275–286.
[8] S.Y. Brendler-Schwaab, A. Hartmann, S. Pfuhler, G. Speit, The in vivo comet assay: use and status in genotoxicity testing, Mutagenesis 20 (2005) 245–254.
[9] Y.F. Sasaki, K. Sekihashi, F. Izumiyama, E. Nishidate, A. Saga, K. Ishida, S. Tsuda, The comet assay with multiple mouse organs: comparison of comet assay results and carcinogenicity with 208 chemicals selected from the IARC monographs and US NTP Carcinogenicity Database, Crit. Rev. Toxicol. 30 (2000) 629–799.
[10] A. Hartmann, E. Agurell, C. Beevers, S. Brendler-Schwaab, B. Burlinson, P. Clay, A.R. Collins, A. Smith, G. Speit, V. Thybaud, R.R. Tice, Recommendations for conducting the in vivo alkaline comet assay, Mutagenesis 18 (2003) 45–51.
[11] A. Hartmann, M. Schumacher, U. Plappert-Helbig, P. Lowe, W. Suter, L. Mueller, Use of the alkaline in vivo comet assay for mechanistic genotoxicity investigations, Mutagenesis 19 (2004) 51–59.
[12] COM (United Kingdom Committee on Mutagenicity of Chemicals in Food, Consumer Products, and the Environment), Guidance on a Strategy for Testing of Chemicals for Mutagenicity, COM, 2000.
[13] FDA, Guidance for Industry and Review Staff Recommended Approaches to Integration of Genetic Toxicology Study Results, FDA, 2006, http://www.fda.gov/cder/guidance/index.htm.
[14] OECD, Mammalian erythrocyte micronucleus test OECD Guidelines for the Testing of Chemicals, vol. 474, OECD, Paris, 1997 (test guideline).
[15] M. Hayashi, 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, Mutat. Res., this issue.
[16] Y.F. Sasaki, A. Saga, M. Akasaka, K. Yoshida, E. Nishidate, Y.Q. Su, N. Matsusaka, S. Tsuda, In vivo genotoxicity of orthophenylphenol, biphenyl, and thiabendazole detected in multiple mouse organs by the alkaline single cell gel electrophoresis assay, Mutat. Res. 395 (1997) 189–198.
[17] E.M. Bomhard, S.Y. Brendler-Schwaab, A. Freyberger, B.A. Herbold, K.H. Leser, M. Richter, *O*-Phenylphenol and its sodium and potassium salts: a toxicological assessment, Crit. Rev. Toxicol. 32 (2002) 551–626.
[18] P.L. Olive, J.P. Banath, Sizing highly fragmented DNA in individual apoptotic cells using the comet assay and a DNA crosslinking agent, Exp. Cell. Res. 221 (1995) 19–26.
[19] L. Henderson, A. Wolfreys, J. Fedyk, C. Bourner, S. Windebank, The ability of the comet assay to discriminate between gentoxins and cytotoxins, Mutagenesis 13 (1998) 89–94.
[20] A. Hartmann, E. Kiskinis, A. Fjaellman, W. Suter, Influence of cytotoxicity and compound precipitation on test results in the alkaline comet assay, Mutat. Res. 497 (2001) 199–212.
[21] M. Kiffe, P. Christen, P. Arni, Characterization of cytotoxic and genotoxic effects of different compounds in CHO K5 cells with the comet assay (single-cell gel electrophoresis assay), Mutat. Res. 537 (2003) 151–168.
[22] M.S. Rundell, E.D. Wagner, M.J. Plewa, The comet assay: genotoxic damage or nuclear fragmentation? Environ. Molec. Mutagen. 42 (2003) 61–67.
[23] T. Mensing, P. Welge, B. Voss, L.M. Fels, H.H. Fricke, T. Bruning, M. Wilhelm, Renal toxicity after chronic inhalation exposure of rats to trichloroethylene, Toxicol. Lett. 128 (2002) 243–247.
[24] G. Scassellati-Sforzolini, R. Pasquini, M. Moretti, M. Villarini, C. Fatigoni, P. Dolara, S. Monarca, G. Caderni, F. Kuchenmeister, P. Schmezer, B.L. Pool-Zobel, In vivo studies on genotoxicity of pure and commercial linuron, Mutat. Res. 390 (1997) 207–221.
[25] Y.F. Sasaki, A. Saga, M. Akasaka, S. Ishibashi, K. Yoshida, Y.Q. Su, N. Matsusaka, S. Tsuda, Detection of in vivo genotoxicity of haloalkanes and haloalkenes carcinogenic to rodents by the alkaline single cell gel electrophoresis (Comet) assay in multiple mouse organs, Mutat. Res. 419 (1998) 13–20.
[26] G.H.S. Strauss, Non-random cell killing in cryopreservation: implications for performance of the battery of leukocyte tests (BLT). I. Toxic and immunotoxic effects, Mutat. Res. 252 (1991) 1–15.
[27] M. Vasquez, R.R. Tice, Detecting genotoxic activity against high molecular weight DNA using the alkaline single cell gel (SCG) assay, Environ. Molec. Mutagen. 29 (1997) 53.
[28] A.R. Collins, A. Ai-guo, S.J. Duthie, The kinetics of repair of oxidative DNA damage (strand breaks and oxidised pyrimidines) in human cells, Mutat. Res. 336 (1995) 69–77.
[29] P.L. Olive, J.P. Banath, R.E. Durand, Heterogeneity in radiation-induced DNA damage and repair in tumor and normal cells measured using the "Comet" assay, Radiat. Res. 122 (1990) 86–94.
[30] T.S. Kumaravel, A.N. Jha, Reliable Comet assay measurements for detecting DNA damage induced by ionising radiation and chemicals, Mutat. Res. 605 (2006) 7–16.
[31] S.J. Wiklund, E. Agurell, Aspects of design and statistical analysis in the comet assay, Mutagenesis 18 (2003) 167–175.

Available online at www.sciencedirect.com
ScienceDirect
Mutation Research 627 (2007) 36–40
Genetic Toxicology and Environmental Mutagenesis
www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

Mouse lymphoma thymidine kinase gene mutation assay: Meeting of the International Workshop on Genotoxicity Testing, San Francisco, 2005, recommendations for 24-h treatment

Martha M. Moore[a,*], Masamitsu Honma[b], Julie Clements[c], George Bolcsfoldi[d], Brian Burlinson[e], Maria Cifone[f], Jane Clarke[g], Philip Clay[h], Rupa Doppalapudi[i], Michael Fellows[j], Bhaskar Gollapudi[k], Saimei Hou[d], Peter Jenkinson[l], Wolfgang Muster[m], Kamala Pant[g], Darren A. Kidd[c], Elisabeth Lorge[n], Melvyn Lloyd[c], Brian Myhr[o], Michael O'Donovan[j], Colin Riach[p], Leon F. Stankowski, Jr.[q], Ajit K. Thakur[f], Freddy Van Goethem[r]

[a] *National Center for Toxicological Research, Food and Drug Administration, USA*
[b] *National Institute of Health Sciences, Division of Genetics & Mutagenesis, Tokyo, Japan*
[c] *Covance Laboratories, Ltd., Harrogate, North Yorkshire, United Kingdom*
[d] *Safety Assessment, AstraZeneca R&D, Sodertalje, Sweden*
[e] *Huntingdon Life Sciences, Huntingdon, United Kingdom*
[f] *Covance Laboratories Inc., Vienna, VA, USA*
[g] *BioReliance Invitrogen Bioservices, Rockville, MD, USA*
[h] *Servier Group, Drug Safety Assessment, F-45403 Orleans-Gidy, France*
[i] *SRI International, Menlo Park, CA, USA*
[j] *Safety Assessment, AstraZeneca R&D, Alderley Park, Macclesfield, United Kingdom*
[k] *The Dow Chemical Company, TERC, Midland, MI, USA*
[l] *Safepharm Laboratories Ltd., Shardlow, Derbyshire, United Kingdom*
[m] *F. Hoffmann-La Roche Ltd., Basel, Switzerland*
[n] *Syngenta CTL, Alderley Park, Macclesfield SK10 4TJ, United Kingdom*
[o] *Genotox Consulting, Bethesda, MD, USA*
[p] *Charles River Laboratories, Tranent, Edinburgh, United Kingdom*
[q] *Johnson and Johnson Pharmaceutical Research & Development L.L.C., Raritan, NJ, USA*
[r] *Johnson & Johnson Pharmaceutical Research and Development, Beerse, Belgium*

Received 28 February 2006; received in revised form 31 July 2006; accepted 14 August 2006
Available online 6 December 2006

Abstract

The Mouse Lymphoma Assay (MLA) Workgroup of the International Workshop on Genotoxicity Testing (IWGT), comprised of experts from Japan, Europe and the United States, met on September 9, 2005, in San Francisco, CA, USA. This meeting of the MLA Workgroup was devoted to reaching a consensus on issues involved with 24-h treatment. Recommendations were made concerning the acceptable values for the negative/solvent control (mutant frequency, cloning efficiency and suspension growth) and the criteria

* Corresponding author at: National Center for Toxicological Research, Food and Drug Administration, NCTR/DGRT, HFT-120, 3900 NCTR Road, Jefferson, AR 72079, USA. Tel.: +1 870 543 7050; fax: +1 870 543 7393.
E-mail address: Martha.Moore@fda.hhs.gov (M.M. Moore).

1383-5718/$ – see front matter
doi:10.1016/j.mrgentox.2006.08.013

to define an acceptable positive control response. Consensus was also reached concerning the use of the global evaluation factor (GEF) and appropriate statistical trend analysis to define positive and negative responses for the 24-h treatment. The Workgroup agreed to continue their support of the International Committee on Harmonization (ICH) recommendation that the MLA assay should include a 24-h treatment (without S-9) in those situations where the short treatment (3–4 h) gives negative results.

Keywords: Mouse lymphoma assay; In vitro mutation; Thymidine kinase

1. Introduction

The Mouse Lymphoma Assay (MLA) Workgroup of the International Workshop on Genotoxicity Testing (IWGT), comprised of experts from Japan, Europe and the United States, met on September 9, 2005, in San Francisco, CA, USA. This meeting of the MLA Workgroup was devoted to reaching a consensus on issues related to the use of 24-h treatment.

The first meeting of the MLA Workgroup was held as a part of the International Workshop on Genotoxicity Testing Procedures in Washington, DC, in the spring of 1999. Since that time, the Workgroup has been working to address three main issues of importance to the assay. These include: (1) the conduct of a data-based analysis and a final recommendation for using the relative total growth (RTG) as the appropriate measure for cytotoxicity; (2) the criteria for data acceptance (based on the negative/vehicle and positive controls) and a new method [the global evaluation factor (GEF)] for data evaluation; (3) the issues related to the International Conference on Harmonization of technical requirements for registration of pharmaceuticals for human use (ICH) recommended use of a 24-h treatment time (including the ability of the assay to detect aneugens). This is the 5th meeting of the Workgroup in which consensus has been reached and reported. The previous four meetings are reported in Moore et al. [1–4].

2. 24-h treatment

Following the ICH recommendation requiring that the MLA be conducted using a 24-h treatment (without S-9) in situations where the short treatment (3–4 h) was negative, laboratories have conducted such experiments. With the goal of determining the approximate frequency at which chemicals require a 24-h treatment to express their mutagenic potential, and to make recommendations concerning the conduct of the 24-h treatment, the Workgroup solicited data from laboratories conducting both the agar and microwell versions of the assay.

Laboratories were asked to evaluate data obtained since 2002 (some earlier data were submitted) and to base a positive result on the assay evaluation criteria in force for each participating laboratory at the time of the performance of the assay. They were asked to identify: (i) compounds uniquely positive following 24-h treatment in the absence of S-9, (ii) compounds positive following short (3- or 4-h) treatment times in the absence of S-9, but negative following 24-h treatment in the absence of S-9, and (iii) compounds either known or suspected to be aneugens.

An estimated 990 data sets (compounds) were reviewed by the individual laboratories to identify compounds fitting into one of these three categories. The majority of assays (approximately 900) were performed using the microwell method; approximately 90 assays used the agar method. Of these approximately 990 tests, 71 (7%) were positive, as assessed by the individual laboratory. It should be noted that the nine laboratories that submitted data only provided the actual data for these 71 test agents. These data were compiled, analyzed and summarized by three members of the Workgroup and the summary information was used in the deliberations of the entire Workgroup. It should be noted that only five data sets were submitted for the third category (known or suspected aneugens). This small number was considered insufficient to permit meaningful analysis and although the data was compiled, there was no discussion of this category by the Workgroup.

2.1. Category 1 responses

A total of 56 data sets (54 in the microwell assay and 2 in the agar assay) were uniquely positive following 24-h treatment in the absence of S-9. The negative controls for each of the 56 data sets were evaluated to determine whether they met all of the revised assay acceptance criteria agreed in the 3rd Workgroup meeting, held in Plymouth in 2002 [3] and the acceptable range for solvent control mutant frequencies (MF) agreed in the 4th Workgroup meeting held in Aberdeen in 2003 [4]. The application of these acceptance criteria eliminated 19 data sets. See Fig. 1 for a breakdown as to the causes of the unacceptable experiments.

A number of the remaining 56 data sets showed only a very small induced MF (IMF). In fact, the maximum

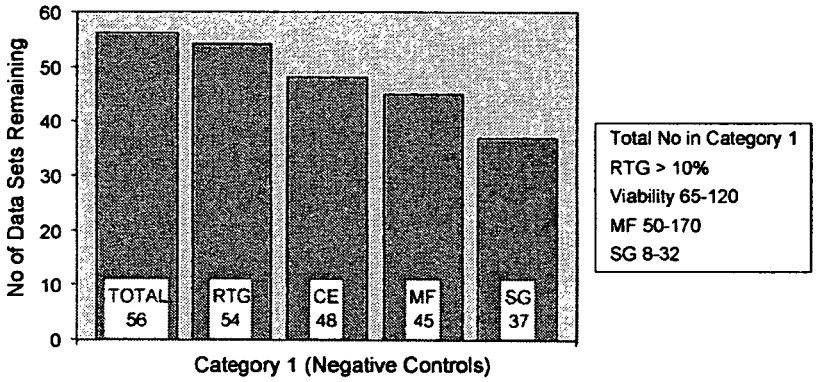

Fig. 1. Column graph demonstrating how many Category 1 (unique 24-h positive) data sets would be excluded as the acceptance criteria [4] are applied sequentially left to right. Note that SG as described in this instance were the values recommended for the short term treatment. They were applied to the 2 day expression period and excluded the 24-h treatment.

increase in MF observed at any data point (compared to concurrent controls) did not exceed the GEF in 26 out of 56 data sets (Fig. 2). Therefore, these data sets did not meet the new criteria required for a positive response, as agreed in the 4th Workgroup meeting in Aberdeen in 2003 [4]. Positive responses are defined as those that exceed the GEF and show statistically positive dose response trends (see discussion below concerning data evaluation).

After applying both the acceptance criteria for the negative controls and the GEF, the number of unique 24-h positive compounds was reduced to 18.

2.2. Category 2 responses

Only 10 data sets (9 in the microwell assay and 1 in the agar assay) were positive following the short (3- or 4-h) treatment in the absence of S-9, but apparently negative following 24-h treatment in the absence of S-9. Application of the acceptance criteria to the negative controls in these data sets reduced this number to 7 and the additional application of the GEF to the data further reduced the number to 4 (Fig. 3). The Workgroup made no additional recommendations concerning the 3–4 h treatment, based on this data.

2.3. Acceptance criteria for negative/vehicle controls (24-h treatment)

Previously, following an extensive evaluation of negative/vehicle control data from a number of laboratories using the short (3- or 4-h) treatment time, the Workgroup reached consensus on the acceptance criteria for individual experiments based upon several negative/vehicle control parameters [4]. With the exception of the suspension growth (SG) parameter, the Workgroup recommended that the same criteria be applied to the 24-h treatment experiments. Because the 24-h treatment includes 3 rather than 2 days of suspension growth, the acceptance criteria for the 24-h SG was revised to 32–180. The theoretical optimum suspension growth is about 5-fold per day, or 125 over the 3-day

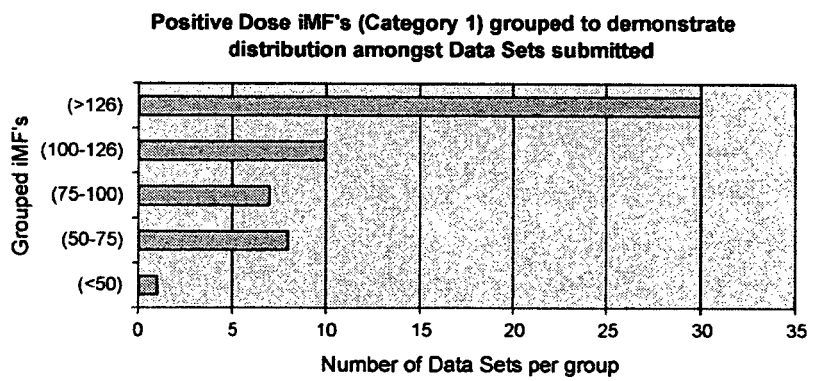

Fig. 2. Bar chart demonstrating breakdown of the 56 Category 1 (unique 24-h positive) studies in terms of the induced MF of highest positive data point.

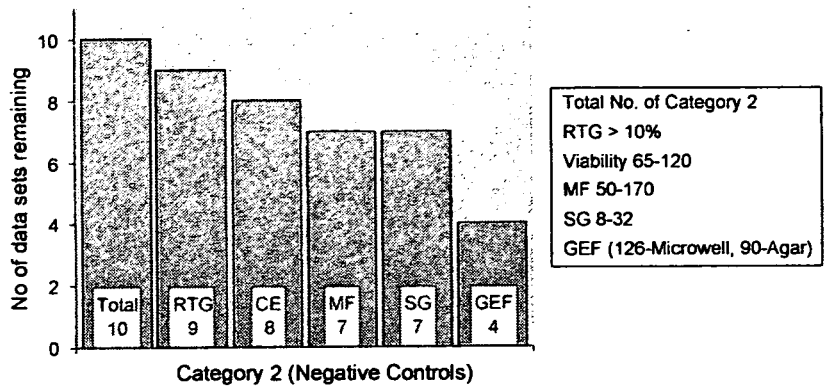

Fig. 3. Column graph demonstrating the number of Category 2 (unique 3- or 4-h positive) data sets that would be excluded as stated acceptance criteria [4] are applied in sequence of left to right.

period. However, there is variability in growth rates and the Workgroup would be very concerned with suspension growth less than 32. The high limit of 180 allows for reasonable errors in cell counting and dilution. As with the short treatment, the acceptance criteria for the background MF are specific to the agar or microwell version of the assay. For both methods the cloning efficiency (CE) referred to in the criteria is the absolute CE obtained at the time of mutant selection. The suspension growth of the negative/vehicle control refers to the growth during both the 24-h treatment and the 2-day expression period following treatment. It is defined as the fold-increase of the cell number during this 3-day period. The SG is calculated by the treatment period fold-increase multiplied by both the expression day 1 and expression day 2 fold-increases in cell number.

The acceptance criteria for the negative/vehicle control parameters for the soft agar and microwell methods of performing the MLA using 24-h treatment are now as follows:

Agar method	Microwell method
MF: 35–140 × 10^{-6}	MF: 50–170 × 10^{-6}
CE: 65–120%	CE: 65–120%
SG: 32–180	SG: 32–180

2.4. Acceptance criteria for positive controls (24-h treatment)

As with the short treatment time experiments, positive control cultures should be included in every 24-h treatment MLA experiment. To assess the adequacy of detection of both small and large colony mutants in the 24-h treatment, the Workgroup agreed that it is appropriate to apply the same acceptance criteria developed for the short treatment time. There are two equally acceptable approaches to assuring an adequate positive control response. (1) The laboratory should use a dose of a mutagenic chemical that yields an absolute increase in total MF that is an increase above the spontaneous background MF [an induced MF] of at least 300 × 10^{-6}. The small colony MF should account for at least 40% of that IMF. For instance, in a culture showing an IMF of 300 × 10^{-6}, the small colony IMF should be at least 120 × 10^{-6}. (2) The second approach requires the use of a dose of a chemical that increases the small colony MF at least 150 × 10^{-6} above that seen in the concurrent negative/vehicle control (a small colony IMF of 150 × 10^{-6}).

In addition, the upper limit of cytotoxicity observed in the positive control culture should have a relative total growth (RTG) that is greater than 10% [2]. The Workgroup recognizes that some laboratories prefer to use more than one dose of their positive control and/or to use a dose that gives a small increase in MF. For these laboratories, it is sufficient if only a single dose of the positive control meets the acceptance criteria.

2.5. Data evaluation

Once the criteria for experimental acceptance have been satisfied, the data from each individual experiment can be evaluated to determine whether the response is positive, negative or equivocal. The Workgroup agreed that data generated using 24-h treatment should be evaluated using the same method previously developed by the Workgroup for use with the short treatment times. A brief summary of the previous analyses conducted by the Workgroup and the rationale for developing the new method using the GEF in conjunction with appropriate statistical analysis to ascertain the presence of a dose-related positive trend is included in the Plymouth and Aberdeen Meeting Reports [3,4]. It should

be noted that the GEF approach takes into account previous guidance documents (i.e. FDA Redbook [http://www.cfsan.fda.gov/~redbook/red-toca.html] and OECD [5], which states that biological relevance should be a major factor in data evaluation.

The GEF evaluation method requires that the IMF exceeds a value based on the global distribution of the background MF for each method (agar or microwell). This value, the GEF, was established by the Workgroup, based on short treatment experiments, to be 126 for the microwell version of the assay and 90 for the agar version. The GEF is applied as follows: if the negative/vehicle control MF in a microwell experiment is 100×10^{-6}, then one of the treatment groups must have a MF of at least 100 + 126 (the microwell GEF) = 226×10^{-6} in order to meet the GEF criterion for a positive call. An appropriate statistical trend test should be applied to determine whether there was a positive dose-related increase.

A test agent response in an experiment is positive if both the IMF for any treatment meets or exceeds the GEF and a positive trend test is obtained. A test agent response is clearly negative if both the trend analysis and the GEF are negative. Situations where either (but not both) the GEF or statistical analysis is positive should be evaluated on a case-by-case basis. It should be noted that it is generally advisable to conduct one or more additional experiments to better define the assay response (particularly in the 30–10% RTG cytotoxicity range).

For more detail on the Workgroup recommendations on the steps for proper assay evaluation, the reader is referred to the summaries of the New Orleans, Plymouth and Aberdeen meetings [2–4]. All of these recommendations are equally applicable to the short treatment and the 24-h treatment.

3. Conclusions

From this analysis, it is clear that only a very small percentage of chemicals (less than 2%) are uniquely positive at 24 h, and an even smaller percentage appear to be uniquely positive at short (3- or 4-h) treatment times. The low numbers of unique 3- or 4-h positive results may be attributable to the regulatory guideline requirements that positive results observed following 3- or 4-h treatments do not need to be further evaluated at 24 h. In some of the unique 24-h treatment cases, the longer treatment time provides for the effective treatment of a higher dose of compound. This is particularly true when a chemical's insolubility prevents testing to adequate toxicity in the short treatment time. There are also some situations in which the maximum recommended concentration (5000 μg/ml) was not mutagenic and was insufficiently toxic in the short treatment, but was mutagenic following a 24-h treatment. There is also evidence that some (but not all) aneugens require longer treatment time [6].

Based upon all the available data, the Workgroup agreed to continue its support of the ICH recommendation that 24-h treatment be used when the short treatment time is negative or equivocal.

References

[1] M.M. Moore, M. Honma, J. Clements, T. Awogi, G. Bolcsfoldi, J. Cole, B. Gollapudi, K. Harrington-Brock, A. Mitchell, W. Muster, B. Myhr, M. O'Donovan, M.-C. Ouldelhkim, R. San, H. Shimada, L. Stankowski Jr., The mouse lymphoma thymidine kinase locus (Tk) gene mutation assay: International Workshop on Genotoxicity Test Procedures (IWGTP) Workgroup Report, Environ. Mol. Mutagen. 35 (2000) 185–190.

[2] M.M. Moore, M. Honma, J. Clements, K. Harrington-Brock, T. Awogi, G. Bolcsfoldi, M. Cifone, D. Collard, M. Fellows, K. Flanders, B. Gollapudi, P. Jenkinson, P. Kirby, S. Kirchner, J. Kraycer, S. McEnaney, W. Muster, B. Myhr, M. O'Donovan, J. Oliver, M.-C. Ouldelhkim, K. Pant, R. Preston, C. Riach, R. San, H. Shimada, L. Stankowski Jr., Mouse lymphoma thymidine kinase gene mutation assay: Follow-up International Workshop on Genotoxicity Test Procedures, New Orleans, Louisiana, April 2000, Environ. Mol. Mutagen. 40 (2002) 292–299.

[3] M.M. Moore, M. Honma, J. Clements, G. Bolcsfoldi, M. Cifone, R. Delongchamp, M. Fellows, B. Gollapudi, P. Jenkinson, P. Kirby, S. Kirchner, W. Muster, B. Myhr, M. O'Donovan, M.-C. Ouldelhkim, K. Pant, R. Preston, C. Riach, R. San, L.F. Stankowski Jr., A. Thakur, S. Wakuri, I. Yoshimura, Mouse lymphoma thymidine kinase locus gene mutation assay: International Workshop (Plymouth England) on Genotoxicity Test procedures Workgroup Report, Mut. Res. 540 (2003) 127–140.

[4] M.M. Moore, 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. Thakur, F. Van Goethem, S. Wakuri, I. Yoshimura, Mouse lymphoma thymidine kinase gene mutation assay: Follow-up Meeting of the International Workshop on Genotoxicity Tests – Aberdeen, Scotland, 2003 – Assay Acceptance Criteria, Positive Controls and Data Evaluation, Environ. Mol. Mutagen. 47 (2006) 1–5.

[5] OECD Guideline 476, In Vitro Mammalian Cell Mutation Test, Adopted July 21, 1997.

[6] M. Honma, L.Z. Zhang, H. Sakamoto, M. Ozaki, K. Takeshita, M. Momose, M. Hayashi, T. Sofuni, The need for long-term treatment in the mouse lymphoma assay, Mutagenesis 14 (1999) 23–29.

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Mutation Research 627 (2007) 59–77
Genetic Toxicology and Environmental Mutagenesis
www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

Strategy for genotoxicity testing—Metabolic considerations

Warren W. Ku[a,*], Anita Bigger[b], Giovanni Brambilla[c], Hansruedi Glatt[d], Elmar Gocke[e], Peggy J. Guzzie[a], Atsushi Hakura[f], Masamitsu Honma[g], Hans-Joerg Martus[h], R. Scott Obach[i], Stanley Roberts[j]

[a] *Pfizer Global Research and Development, Drug Safety Research and Development, Groton, CT 06340, USA*
[b] *U.S. Food and Drug Administration, CDER, Division of Antiviral Products, Silver Spring, MD 20993, United States*
[c] *Department of Internal Medicine, Division of Clinical Pharmacology and Toxicology, University of Genoa, I-16132 Genoa, Italy*
[d] *German Institute of Human Nutrition Potsdam-Rehbrücke, Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany*
[e] *F. Hoffmann-La Roche Ltd., Preclinical Safety, CH-4070 Basel, Switzerland*
[f] *Eisai Co., Ltd., Drug Safety Research Laboratories, Kakamigahara, GIFU 501-6195, Japan*
[g] *National Institute of Health Sciences, Division of Genetics & Mutagenesis, Tokyo, Setagaya-Ku 158-8501, Japan*
[h] *Novartis Pharma AG, Exploratory Development, Genetic Toxicology and Safety Pharmacology, 4002 Basel, Switzerland*
[i] *Pfizer Global Research and Development, Pharmacokinetics, Dynamics and Metabolism, Groton, CT 06340, USA*
[j] *Abbott Laboratories, Drug Metabolism and Pharmacokinetics, Abbott Park, IL 60064, USA*

Received 28 February 2006; received in revised form 31 July 2006; accepted 14 August 2006
Available online 1 December 2006

Abstract

The report from the 2002 International Workshop on Genotoxicity Tests (IWGT) Strategy Expert Group emphasized metabolic considerations as an important area to address in developing a common strategy for genotoxicity testing. A working group convened at the 2005 4th IWGT to discuss this area further and propose practical strategy recommendations. To propose a strategy, the working group reviewed: (1) the current status and deficiencies, including examples of carcinogens "missed" in genotoxicity testing, established shortcomings of the standard *in vitro* induced S9 activation system and drug metabolite case examples; (2) the current status of possible remedies, including alternative S9 sources, other external metabolism systems or genetically engineered test systems; (3) any existing positions or guidance. The working group established consensus principles to guide strategy development. Thus, a human metabolite of interest should be represented in genotoxicity and carcinogenicity testing, including evaluation of alternative genotoxicity *in vitro* metabolic activation or test systems, and the selection of a carcinogenicity test species showing appropriate biotransformation. Appropriate action triggers need to be defined based on the extent of human exposure, considering any structural knowledge of the metabolite, and when genotoxicity is observed upon *in vitro* testing in the presence of metabolic activation. These triggers also need to be considered in defining the timing of human pharmaceutical ADME assessments. The working group proposed two strategies to consider; a more proactive approach, which emphasizes early metabolism predictions to drive appropriate hazard assessment; and a retroactive approach to manage safety risks of a unique or "major" metabolite once identified and quantitated from human clinical ADME studies. In both strategies, the assessment of the genotoxic potential of a metabolite could include the use of an alternative or optimized *in vitro* metabolic activation system, or direct testing of an isolated or synthesized metabolite. The working group also identified specific areas where more data or experiences need to be gained to reach consensus. These included defining a discrete exposure action trigger for safety assessment and when direct testing of a metabolite of interest is warranted versus the use of an alternative *in vitro* activation system, a universal recommendation for the timing of human ADME studies for drug candidates and the positioning of metabolite structural knowledge (through *in silico* systems, literature, expert analysis) in supporting metabolite safety qualification. Lastly, the working group outlined future considerations for refining the initially proposed

* Corresponding author. Tel.: +1 860 441 3969; fax: +1 860 715 1251.
E-mail address: warren.w.ku@pfizer.com (W.W. Ku).

1383-5718/$ – see front matter
doi:10.1016/j.mrgentox.2006.10.004

strategies. These included the need for further evaluation of the current *in vitro* genotoxicity testing protocols that can potentially perturb or reduce the level of metabolic activity (potential alterations in metabolism associated with both the use of some solvents to solubilize test chemicals and testing to the guidance limit dose), and proposing broader evaluations of alternative metabolic activation sources or engineered test systems to further challenge the suitability of (or replace) the current induced liver S9 activation source.

Keywords: Genotoxicity; Metabolism; Testing; Strategy

1. Introduction

The report on initial activities of the IWGT Strategy Expert Group [1] emphasized metabolic considerations as an important area to address in developing a common strategy for genotoxicity testing and risk assessment. For human risk assessment, differences in metabolism attributed to the conditions of standard *in vitro* test systems or animal species pose challenges for establishing human relevance. This can be important in cases where the ultimate toxic chemical entity may be a metabolite of the parent chemical agent, whether it applies to pharmaceuticals, pesticides or industrial chemicals. Current genotoxicity testing guidelines incorporate a battery of *in vitro* and *in vivo* tests to identify hazards [2,3]. To address the potential role of metabolism, induced rat liver S9 has been adopted for *in vitro* genotoxicity tests as an exogenous activation system for detecting promutagens [4,5]. In evaluating test results, current *in vitro* test systems can generate both false negative and false positive results in relation to predicting rodent carcinogenicity. These potential discrepancies could be attributed to metabolic considerations, such as the limited capability of *in vitro* exogenous activation systems [6]. These discrepancies may be reduced by considering the use of alternative activation and/or testing approaches based on knowledge of certain structural features or chemical classes (potential chemical reactivity or substrate preference for a specific enzyme not represented or inactive in the standard S9 activation system). *In vivo* genotoxicity tests, typically conducted in rodent species (rat or mouse), take into consideration route of administration, toxicokinetics and absorption–distribution–metabolism–excretion (ADME) in the assessment of relevance and risk of any *in vitro* findings. These same rodent species are also employed for later assessment of carcinogenic potential. The ability of both *in vitro* and *in vivo* test systems to model human metabolites of interest can be quite variable and therefore limit current capabilities to adequately conduct human genotoxicity and carcinogenicity risk assessment in certain cases.

An independent working group convened at the 4th International Workshop on Genotoxicity Tests (IWGT) to discuss the issues regarding metabolic considerations and propose practical strategy recommendations. To arrive at a proposed strategy, the working group reviewed the following: (1) the current status and deficiencies, including examples of carcinogens "missed" in genotoxicity testing, established shortcomings of the standard *in vitro* induced S9 activation system and drug metabolite case examples; (2) the current status of possible remedies, including alternative S9 sources, other external metabolism systems or genetically engineered test systems; (3) any existing and related recommendations, guidance or strategies [7–10]. Based upon this review, an initial working strategy was proposed along with working group consensus and non-consensus areas.

2. Key learnings from assessing historical and current information

2.1. Current status and deficiencies

2.1.1. Representative examples of carcinogens "missed" in standard genotoxicity testing due to inadequate metabolism

The standard test battery (ICH S2B) for genotoxic potential was designed to avoid the risk of false negative results for compounds with genotoxic potential. However, it cannot be taken for granted that this risk is completely eliminated when compounds give negative results in all the battery assays. Brambilla and Martelli [11] have shown that there are some chemicals, classified by the International Agency for Research on Cancer (IARC) as probable (Group 2A) or possible (Group 2B) human carcinogens, that tested consistently negative in the standard battery but were positive in other non-standard genotoxicity assays. These chemicals often induced DNA damage and repair synthesis or the formation of DNA adducts in the target organs of their carcinogenic activity, as well as in primary cultures of cells of the target organ from animal and, in some cases, human donors. These chemicals include certain polychlorinated biphenyls, trichloroethylene, tetrachloroethylene, *o*-toluidine, carbon tetrachloride, cyproterone acetate and its structural analogues chlormadinone acetate and megestrol acetate, *p*-dichlorobenzene, hexachlorobenzene, hexachloroethane, nitrilotriacetic acid and its salts,

and ochratoxin A. The failure of the standard test battery to detect some genotoxic carcinogens is attributable to several causes, but the principal ones are the following: *in vitro*, the artificial metabolic activity of the liver S9-mix, and the different biotransformation of chemicals in cells of different tissues and from different species; *in vivo*, the pharmacokinetics of the test compound and its possible species-, sex- and tissue-specificity.

Recently, further investigation [12 and unpublished data] by this research group demonstrated that four rat thyroid carcinogens (propylthiouracil, methimazole, *N,N'*-diethylthiourea and ethylenethiourea) (see IARC Monograph 79, 2001) that gave negative or contradictory responses in the standard battery are thyroid-specific DNA-damaging agents. All four compounds induce DNA single-strand breaks and alkali-labile sites in primary cultures of human thyroid cells. Following oral administration to rats, these compounds induced DNA fragmentation in the thyroid with no evidence of DNA lesions being induced in liver and kidneys. This organ specificity may be due to a difference in metabolism in the thyroid versus liver and kidney and/or to the occurrence in the thyroid of a higher concentration or a more persistent presence of these compounds.

2.1.2. A historical introduction: 2-acetylaminofluorene

2-Acetylaminofluorene (2-AAF) was developed, but never used, as an insecticide, as early animal studies indicated high carcinogenic activity [13]. Instead it became a key model compound in experimental carcinogenesis. It is active over wide tissue and species ranges, often with marked sex-dependence. In a feeding study in rats, it induced high incidences of liver tumors in males, but none in females (where mammary tumors were very common), and moderate incidences of neoplasias in the ear duct and the small intestine in both sexes [14]. In the 1960s, James and Elizabeth Miller pioneered the concept that carcinogens are either electrophiles as such or are metabolically activated to electrophiles, which covalently bind to cellular macromolecules [15]. At this time, DeBaun et al. [16] and King et al. [17] discovered that *N*-hydroxy-2-AAF, a major phase-1 metabolite of 2-AAF, extensively bound to proteins and nucleic acids in the presence of a cytosolic preparation from rat liver fortified with a system generating 3'-phosphoadenosine-5'-phosphosulfate (PAPS), the cofactor required by sulfotransferases (SULTs). This finding marked a milestone, as for the first time the activation pathway of a carcinogen to an electrophilic metabolite was elucidated. Later it was shown that the SULT inhibitor pentachlorophenol drastically decreases the formation of DNA adducts and the induction of tumors by 2-AAF and *N*-hydroxy-2-AAF in liver of rats and mice [18–21]. Likewise, a genetic defect leading to reduced hepatic biosynthesis of PAPS strongly protected infant mice against the hepatocarcinogenicity of *N*-hydroxy-2-AAF [20]. These various findings clearly demonstrate that a sulfo conjugate is the principal ultimate carcinogen of 2-AAF in mouse and rat liver.

The detection of the bioactivation of carcinogens revolutionized genetic toxicology. External activating systems, usually in the form of liver S9, were incorporated in all kinds of *in vitro* test systems. Ironically, standard S9 does not provide any SULT activity—the key factor of the hallmark findings with *N*-hydroxy-2-AAF described in the preceding paragraph. The reason for this deficiency is lack of the cofactor PAPS. Nevertheless, 2-AAF is mutagenic in the Ames test in the presence of liver S9 [22], although its activity is very modest taking into account the exceptional carcinogenic activity of 2-AAF and that a major target tissue, the liver of the male rat, is the source of standard S9. Surprisingly, the mutagenicity of *N*-hydroxy-2-AAF in the presence of hepatic cytosol was decreased upon addition of PAPS, whereas its covalent binding to free DNA was enhanced by the same modification [23]. Likewise, chemically synthesized *N*-sulfooxy-AAF, the putative ultimate carcinogen of 2-AAF, showed negligible mutagenic activity in *Salmonella* [24].

2-AAF, tested up to a concentration of 300 μM, was negative in a gene (*hprt*) mutation test in V79 cells in the presence of liver S9, whereas is was clearly mutagenic even at 0.1 μM in V79-derived cells engineered for co-expression of rat cytochrome P450 (CYP) 1A2 and rat SULT1C1 [25]. Both enzymes were required for the activation. SULT1C1 is expressed with high tissue and sex specificity in the liver of male rats [26,27], in agreement with the sex-specific major target tissue in male rats [14] (the activation pathways leading to tumors in other tissues, such as the mammary gland, have not been elucidated; they may involve other SULT forms or be independent of SULT).

Subsequently, *N*-hydroxy-2-AAF was tested in *Salmonella* strains engineered for expression of various individual human SULTs. A strong activation was observed in the presence of SULT1A2 [28]. This finding implies that *N*-sulfooxy-AAF generated within the target cell is mutagenic to bacteria. Thus, the negligible activity of externally added or generated *N*-sulfooxy-AAF has to be due to insufficient penetration. The half-life of *N*-sulfooxy-AAF in water at 37 °C amounts to ~4 s [29]. A similar half-life has been reported for aflatoxin B_1 exo-8,9-epoxide, the active metabolite of aflatoxin B_1 [30].

Since bioactivated aflatoxin B_1 is readily detected in the Ames test [22], other properties of *N*-sulfooxy-AAF may be important for its poor membrane permeability. The most plausible factor is the negative charge (combined with the relatively short lifespan).

Key learnings from the findings with 2-AAF are:

- Non-CYP enzymes may be critically involved in the activation of carcinogens and mutagens.
- These non-CYP enzymes may not be active in standard target cells (in the case of 2-AAF due to lack of the critical enzyme, a SULT) or S9 preparations (due to lack of the corresponding cofactor, PAPS) [31–33].
- Supplementation of the cofactor may not be a reliable remedy, as some active metabolites (especially phase-2 metabolites) may not permeate into the target cells. In fact, external activation may lead to decreased mutagenicity as shown for *N*-hydroxy-2-AAF [23] and thus preclude detoxification. The more appropriate remedies are the use of metabolically competent target cells.

The SULT form required for hepatocarcinogenicity of 2-AAF (SULT1C1) is expressed with high tissue selectivity and is rapidly lost when hepatocytes are taken into culture. Liu et al. [34] cultured primary rat hepatocytes under 12 different conditions. Even under the best conditions, the level of SUL1C1 mRNA fell to approximately 20% and <5% of its initial values within 24 and 96 h, respectively. Under certain conditions, the SULT1C1 mRNA was already undetectable after 24 h. For such expression-sensitive enzymes, the use of genetically engineered cells appears to be the most practical approach to date.

2.1.3. Issues and potential shortcomings of the induced rat liver S9 fraction in in vitro genotoxicity tests

In standard *in vitro* genotoxicity testing, an activation system is included with the purpose of generating electrophilic metabolites that can react with macromolecules including nucleic acids. This has been a hallmark of the Ames bacterial mutagenicity test, and its utility was apparent when some of the first environmental mutagens were characterized in this system. It was a natural extension of the pioneering work of the Millers who demonstrated bioactivation of xenobiotics to reactive metabolites [35]. Many mutagens require oxidative metabolism to reactive species before demonstrating mutagenicity.

The induced rat liver S9 fraction is, in essence, employed as a 'metabolite factory' in the Ames and other *in vitro* genotoxicity tests. In principle, other systems capable of generating reactive metabolites could also be used (see below). However, the regulatory standard presently requires the induced rat liver system, which has been used for decades. Its initial choice was logical; the levels of several CYP enzymes are elevated after induction, in particular the CYP1A subfamily of enzymes (CYP1A1 and 1A2), which are efficient catalysts of the bioactivation of polycyclic aromatic hydrocarbons and azaarenes, aromatic amines and aflatoxins. These types of compounds were some of the first and best understood mutagens, and the Aroclor 1254-induced rat S9 fraction effectively allowed their identification as mutagens. Its choice was also logical in that it provided a reliable, robust and readily available bioactivation system at a time when human-derived systems were rare or unavailable. Also, a rodent system can be more easily standardized than an exogenous human-derived system that normally would rely on human tissue samples, which are subject to significant biological variation.

However, to identify human hazards, the relevance of the induced rat liver S9 fraction can be called into question. It is now known that the rat and human CYP enzymes can differ in their substrate selectivities and reactions catalyzed. While the CYP1A family is reasonably well conserved between rats and humans with regard to structure and function, other CYP enzymes that are more frequently involved in drug metabolism are more divergent. In particular, the CYP2 family (e.g. CYP2C, 2D) demonstrates vast differences in activity between rat and human. Furthermore, upon induction with Aroclor 1254, an unnatural complement of CYP enzymes is generated: the CYP1A and 2B enzymes are very markedly elevated; others such as CYP3A are affected in a minor way, while others (e.g. CYP2C11) may decrease (Table 1).

Table 1
Complement of rat liver cytochrome P450 (CYP) enzymes in uninduced and Aroclor-1254 induced states

Enzyme	Cytochrome P450 (nmol/mg microsomal protein)	
	Untreated	Aroclor treated
CYP1A1	0.04	1.45
CYP1A2	<0.03	1.23
CYP2B1	0.03	1.29
CYP2B2	0.07	1.46
CYP2C6	0.36	0.36
CYP2C11	1.20	0.27
CYP2D1	0.15	0.15
CYP3A	0.39	0.77

From Ref. [73].

Little has been done to compare the metabolite profiles generated in Aroclor-induced rat liver S9 to those in human liver S9 for drug-like molecules. In a preliminary unreported study, the metabolism of ten common drugs spanning a range of structural types and CYP substrates in Aroclor-1254 induced rat and pooled human liver S9 fractions was examined. In this study, it was demonstrated that in a qualitative sense, metabolite profiles in the two systems did not differ substantially—for the most part metabolites generated in one system were observed in the other. However, relative quantities of metabolites did differ between the two systems and in a few cases, metabolites observed in one system were not observed in the other. In addition, there is recent evidence demonstrating a large variation in the Ames test results between human and induced rat S9 [36–43]. For example, according to a recent study, certain human S9 preparations detected 25% of compounds with higher sensitivity than induced rat S9 although the remaining 75% was detected at higher levels with the rat enzymes. The chemicals that showed stronger mutagenic activity in the human than in the rat S9 systems included some aromatic amines, polycyclic azaarenes, and nitrosamines, which constitute important classes of mutagenic substances. In addition, a large variation (inter-individual diversity) in the Ames test results was found between multiple human S9 samples that varied by chemical. The Ames test data with 18 different human S9 samples indicated that although the mutagenicity of 2-amino-3-methylimidazo[4,5-*f*]quinoline (IQ) and benzo[*a*]pyrene (BP) in the presence of a high activity human S9 was equal to that observed in the presence of induced rat S9, their mutagenicity in the presence of pooled human S9 was much lower than that observed in the presence of induced rat S9 (IQ: 0.1-fold and BP: 0.05-fold). In contrast, the mutagenicity of dimethylnitrosamine in the presence of either high activity human S9 or pooled human S9 was eight-fold higher than that found in the presence of induced rat S9.

The above examples call into question the relevance of a rodent *in vitro* system for evaluating human metabolites. The possibility exists that the induced rat system may generate mutagenic metabolites of no relevance, or worse still may not generate a mutagenic metabolite that would be generated by human enzymes. Considerations for replacing or complementing the rat *in vitro* system with a human system are highlighted particularly when we need to assess the mutagenicity of unique or major human metabolites. Identification of cases in which discordance between metabolite profiles in rat liver S9 fraction and corresponding systems from other species can be readily made. Over the past decade, profiling and identification of metabolites in various *in vitro* systems (e.g. liver microsomes, S9 fractions, hepatocytes) from human and laboratory animal species has become a standard part of the contribution of Drug Metabolism and Pharmacokinetics scientists to drug discovery and development processes. Such metabolite profile data can be leveraged in the design of appropriate *in vitro* genotoxicity tests—comparison of rat and human profiles can be made, and if markedly different, the rat system could be replaced.

Based on these situations stated above, consideration of replacing the rat *in vitro* system with a human system has been highlighted particularly when we need to assess the mutagenicity of human unique or major metabolites. These findings, therefore, suggest the utility of human S9 in mutagenicity testing systems for human hazard identification.

In addition to the potential shortcomings of the use of an *in vitro* system from a non-human species, it must also be considered that the use of an S9 system with NADPH (for CYP and flavin-containing monooxygenase activities) also represents an incomplete picture of the metabolism that can occur *in vivo*. Some enzymes that can sequester proximate mutagens, such as UDP-glucuronosyltransferases (UGTs), or inactivate electrophilic metabolites, such as glutathione transferases (GSTs), are not active in the NADPH supplemented S9 system because other cofactors and additives would be needed. In particular, sulfotransferases, which through the transfer of a sulfuryl group, can sequester some proximate mutagens but are well known to activate others are also not active in the standard testing system because the necessary cofactor, PAPS, is not added [31,32].

2.1.4. Issue of poisoning of enzymes by organic solvents

The use of solvents to deliver the test chemicals to the *in vitro* genotoxicity tests is a common practice. Regulatory requirements are such that new compounds should be tested to 10 mM (or 5 mg/ml) or 5 mg/plate in the Ames test or to the limit of solubility, whichever is lower [2]. In order to achieve such high concentrations for lipophilic molecules such as drugs and environmental chemicals, solvents are used in the preparation of concentrated stock solutions. Due to its nearly universal capability to dissolve organic chemicals, aqueous miscibility, and its very low reactivity, dimethylsulfoxide (DMSO) has become a routinely used vehicle for the genotoxicity tests. For instance, in the Ames test (preincubation version) DMSO volumes of 100 and 200 μl per incubation mix (700 or 800 μl) have been recommended