Food and Chemical Toxicology 106 (2017) 609–615

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevier.com/locate/foodchemtox

The current limitations of *in vitro* genotoxicity testing and their relevance to the *in vivo* situation

Fabrice Nesslany [a, b, *]

[a] Genetic Toxicology Department, Institut Pasteur de Lille, 1 rue du Pr. Calmette, 59019, Lille, France
[b] EA 4483, Université Lille 2, 59000, Lille, France

ARTICLE INFO

Article history:
Received 13 July 2016
Received in revised form
12 August 2016
Accepted 30 August 2016
Available online 31 August 2016

Keywords:
In vitro and in vivo genotoxicity
Limitations
False positive
False negative

ABSTRACT

The standard regulatory core battery of genotoxicity tests generally includes 2 or 3 validated tests with at least one *in vitro* test in bacteria and one *in vitro* test on cell cultures. However, limitations in in vitro genotoxicity testing may exist at many levels. The knowledge of the underlying mechanisms of genotoxicity is particularly useful to assess the level of relevance for the *in vivo* situation. In order to avoid wrong conclusions regarding the actual genotoxicity status of any test substance, it appears very important to be aware of the various origins of related bias leading to 'false positives and negatives' by using *in vitro* methods. Among these, mention may be made on the metabolic activation system, experimental (extreme) conditions, specificities of the test systems implemented, cell type used etc. The knowledge of the actual 'limits' of the *in vitro* test systems used is clearly an advantage and may contribute to avoid some pitfalls in order to better assess the level of relevance for the *in vivo* situation.

1. Introduction

From a public health point of view, one of the main objectives of toxicology is to eliminate hazardous substances from our environment. In particular, genotoxicity assessment theoretically is aimed at eliminating genotoxic carcinogens as soon as possible. Regarding scientific and mechanistic aspects, many genetic events may lead to genotoxicity meaning that the testing strategy should cover all the possible mechanisms of genotoxicity, *i.e.* gene mutation and chromosomal aberrations including both clastogenic and aneugenic effects.

Therefore, in the regulatory side, whatever the domain of application of the test substances, knowing that no single test is able to display all events possibly leading to genotoxicity and/or mutagenicity, a battery of tests is required in order to assess gene point mutation and/or chromosomal aberrations (structural & numerical) induction. Primary DNA damage (*e.g.*, Comet assay) may be specifically assessed but usually in a second intention. The current set should include tests with high sensitivity (to theoretically avoid « false negative ») and specificity (to theoretically avoid « false positive » and to reach a mutagenicity-carcinogenicity relationship as high as possible).

Consequently, the regulatory core battery of tests usually includes 2 or 3 validated tests (for which guidelines are available) with at least one *in vitro* test in bacteria and one *in vitro* test on cell cultures. For instance, the current strategy for food and feed safety assessment (EFSA, 2011) or for human drugs (ICH S2-R1 option 1) recommends performing both the Ames and an *in vitro* test on mammalian cells, *e.g.*, the micronucleus test. Even if there is an evident complementarity between these assays (use of prokaryotic cells *vs* eukaryotic cells; gene mutations *vs* chromosomal aberrations), it remains theoretical and limitations (*i.e.*, deficiencies, lack of relevance, inadequacies, bias, wrong interpretation …) in *in vitro* genotoxicity testing may exist at many levels. The goal of this paper is not to deplore the lack of systematic *in vivo* testing but instead to keep in mind that misleading predictive results using *in vitro* systems exist. Subsequently, the description of the various origins of related limitations to avoid false positives and negatives mainly by describing the underlying mechanisms of genotoxicity is particularly useful to assess the level of relevance for the *in vivo* situation.

2. Metabolic consideration: one of the major sources of deficiencies?

Knowing that some substances may lead to genotoxic entities after metabolization, there is a need for testing both the whole

* Genetic Toxicology Department, Institut Pasteur de Lille, 1 rue du Pr. Calmette, 59019, Lille, France.
E-mail address: fabrice.nesslany@pasteur-lille.fr.

http://dx.doi.org/10.1016/j.fct.2016.08.035
0278-6915/

substance and its metabolite(s) meaning that the *in vitro* test systems should include the use of metabolic activation, either endo- or exogenous which is a major source of deficiencies. Indeed, the level of metabolic efficacy is directly managed by the metabolic activation system itself depending on its degree of organization (*e.g.*, whole organs, slices of organs, enzymatically competent primary cultured cells or cell lines, subcellular fractions …).

Regarding the most currently used exogenous metabolic activation system in *in vitro* assays (namely S9) it may directly influence results depending on different parameters:

- The inducer used (if any) = "*inducer effect*". As largely published, the metabolic induction leads to strong modulations in terms of expression of CYP450. As demonstrated by Guengerich et al. (1982), individual forms of CYP-450 can be induced by different compounds (e.g., phenobarbital, 5,6-benzoflavone, pregnenolone-16 alpha-carbonitrile, isosafrole, or the polychlorinated biphenyl mixture Aroclor 1254) and a single compound can lower the level of one form of CYP-450 while inducing one or more other forms of CYP-450. For instance, induction factors (ratio of level of expression of untreated rats/inducer-treated rats) range from 0.23 (CYP450 2C11) to up to 41 (CYP450 1A2) after induction with Aroclor 1254 (Table 1 hereafter). These results underline that a metabolic pattern could be 'preferred' to the detriment of another (preferential activation).

Furthermore, it is noteworthy that the normal level of expression of these CYP450 in human liver may be deeply different from the ones of rodent induced liver (up to the total absence of expression) which already raises the question of the transposition to human.

- When induction is performed, parameters such as the origin in terms of organ (*e.g.*, liver, kidney, intestine …, *e.g.*, Obrecht-Pflumio et al., 1999; Nishimuta et al., 2013) and/or species (rat, mouse, dog, human …, *e.g.*, Beaune et al., 1985; Neis et al., 1986; Nishimuta et al., 2013; Cox et al., 2016) but also age and sex of animals (*e.g.*, Imaoka et al., 1991; Kamataki et al., 1983) directly influence the qualitative and quantitative rate of metabolization, and thus possibly the mutagenic response.
- In the same way, the final concentration (*e.g.*, percentage of S9 in final concentration) also directly impacts the global metabolic capacity which may result in different responses (Rees et al., 1989).

Therefore, the metabolic efficacy depends on different parameters of the system used which representativity and transposition to human are not always guaranteed.

Many attempts to overcome these limits were historically proposed, *e.g.*, use of genetically engineered cells to express human or rat CYP-450, use of human S9 … but it appeared that these 'solutions' were finally not more reliable. As a matter of fact, trials on pro-mutagenic reference compounds (*i.e.*, mutagens only after metabolization), performed using S9 from different species displayed a species-dependent mutagenic effect. In particular, comparison of results using either human or rat S9 clearly shows that S9 from human origin is not always the most « effective ». For instance, Beaune et al. (1985) investigated the mutagenic effects on different well-known genotoxins (Aflatoxin B1, 2-amino-anthracene, 2-amino-fluorene, 3-methyl-cholanthrene, and cigarette smoke condensate) with the Ames test by using either human or rat liver S9 (Aroclor-treated or untreated). For each concentration of product, the number of revertants was measured as a function of the quantity of S9 added to the medium.

Aflatoxin B1, 3-methylcholanthrene and cigarette-smoke condensate were much less mutagenic when activated with human-liver S9 than with untreated rat-liver while S9 treatment with Aroclor increased the mutagenic potency of rat-liver S9. With human S9, aflatoxin B1 remained mutagenic whereas 3-methylcholanthrene and cigarette-smoke condensate were not. Under these experimental conditions, 2-amino-fluorene was activated to a similar degree by human-liver S9 and by untreated rat-liver S9. On the contrary, 2-aminoanthracene was much more mutagenic after activation by human-liver S9 than by rat-liver S9. This single example clearly demonstrated that human S9 is not always the most enzymatically efficient.

In the same way, a recent review concluded that the metabolic activity of induced rat S9 was found to be higher than human S9, and linked to high mutagenic potency results (Cox et al., 2016). Indeed, human S9 often yields significantly lower Salmonella mutagenic potency values, especially for polycyclic aromatic hydrocarbons, aflatoxin B1 and heterocyclic amines (~3- to 350-fold). Conversely, assessment with human S9 activation yields higher potency for aromatic amines (~2- to 50-fold). Similar trends were observed in experimentally generated mammalian micronucleus cell assays, however human S9 elicited potent cytotoxicity in L5178Y, CHO and TK6 cell lines. Due to the potential for reduced sensitivity and the absence of a link between enzyme activity levels and mutagenic potency, human liver S9 is not recommended for use alone in *in vitro* genotoxicity screening assays; however, human S9 may be extremely useful in follow-up tests, especially in the case of chemicals with species-specific metabolic differences, such as aromatic amines.

As a conclusion, from the moment when a metabolic activation system is needed, the 'ideal' situation does not exist. In return, the knowledge of the actual 'limits' of the (exogenous) metabolic activation system used is clearly an advantage and may contribute

Table 1
Differences in CYP450 expression depending on induction and on species (adapted from Guengerich et al., 1982).

Enzyme CYP450	CYP nmol/mg protein		Induction factor[a] (rat)	Expression level in human liver
	Untreated rat	Aroclor- 1254 treated rat		
1A1	0.04	1.45	36	**0**
1A2	<0.03	1.23	>41	**0**/+
2B1	0.03	1.29	**43**	+
2B2	0.07	1.46	21	
2C6	0.36	0.36	1	++
2C11	1.20	0.27	**0.23**	
2D1	0.15	0.15	1	**0**/+
3A2	0.39	0.77	2	+++

In bold, the minimum and the maximum induction factors.
[a] Aroclor- 1254 treated Rat/Untreated.