Review

ALTERNATIVE MODELS FOR TOXICITY TESTING OF XENOBIOTICS*

Jasna KNIEWALD, Ivana KMETIČ, Višnja GAURINA-SRČEK, and Zlatko KNIEWALD

Faculty of Food Technology and Biotechnology, University of Zagreb, Zagreb, Croatia

Received in October 2004

The alternatives to whole-animal testing include endpoint assays, cell and tissue cultures, use of tissue slices, toxicokinetic modelling, and structure-activity relationships and databases. The use of *in vitro* systems (subcellular fractions, cell lines, primary cell cultures, tissue slices, organ cultures, etc.) as research tools in toxicology is widespread. In the past few years, the apoptosis phenomena were followed by very precise intracellular changes where, through programmed cell death, a cell can be removed from a population. The *in vitro* systems are ideally suited for investigations of the molecular, cellular and physiological mechanisms of chemically induced toxicity, which cannot readily be studied *in vivo* for known target organ and target species toxicity studies and for answering specific questions about toxic effects. The main justification for developing *in vitro* toxicity tests is that they will make toxicology a more scientifically based practice. It is increasingly apparent that the development and incorporation of stepwise testing strategies, combining experimental data from a range of alternative methods (physicochemical techniques, quantitative structure-activity relationships - QSAR, metabolic and kinetic modelling and *in vitro* tests), provide the most advanced way to predict toxicity, reducing at the same time the number of laboratory animals used for testing.

KEY WORDS*: alternative methods, biomarkers,* in vitro *methods, predicting toxic hazard, validation*

Technical improvements in tissue culture and development of the Ames test, which uses bacteria to detect mutagens, challenge the view that animal testing is the only option in toxicity testing. In addition, there has been a growing recognition of the limitations of certain standard *in vivo* testing procedures. The explosion of knowledge in molecular biology has also significantly affected toxicology. Moreover, the costs of assessing potential health effects of some 200,000 substances per year that are newly identified or synthesized necessitate alternatives to animal testing. It has been estimated that the cost of testing of a single substance using whole animals is frequently in excess of $2 million. In addition, the *in vitro* testing provides the researcher with considerably more control of the variables than the whole-animal testing. However, new tools for toxicity testing must be looked on as adjuncts to traditional testing methods. Any testing method has inherent difficulties; when using whole animals, data must be extrapolated from one species to another, and when using cell or tissue culture, data must be extrapolated to the whole animal.

The new testing methods that more accurately assess hazards, are less expensive and are able to determine toxicity more rapidly. It is the pursuit of these goals that accounts for the great progress in the alternative testing techniques. At the same time, certain basic tests, such as the Ames test, continue to be the workhorses for specific areas of toxicology.

The alternatives to whole-animal testing include endpoint assays, cell and tissue cultures, use of tissue slices, toxicokinetic modelling, and structure-activity relationships and databases. Some examples of non-whole animal methods include the use of bacteria

* Partly presented at the 3rd Croatian Congress of Toxicology, Plitvice, Croatia, 26-29 May 2004

and yeast to assess mutagenicity, chicken embryos to assess teratogenicity and fixed enzyme systems to screen for biological effects. The main questions concerning the use of alternatives are:

- How do we extrapolate from an *in vitro* system to an *in vivo* system (i.e., How do we relate effects in single cells to complex interactions in whole animals)?
- How do we use available *in vitro* and *in vivo* data to design better experimental approaches?
- How do we predict potential biological effects from the chemical structure of a substance?

The concordance between results from alternative tests and those from mammals is an important issue in protecting public safety. It is also important to note that the research using cells, tissue cultures or non-mammalian systems is conducted not only as an alternative to using mammals, but because a given alternative system provides the best answers to the question under study. The *in vitro* studies also allow researchers to understand the discrete steps in a specific sequence of events, which is difficult to do in whole animals.

One reason why cell and tissue culture systems are valuable in toxicity testing is that they can be observed with a light microscope while various components of the system are manipulated. For instance, one can observe the beating of cultured heart cells and note those effects of adding various chemicals to the culture medium. The human oral fibroblasts are used for testing dental materials, and the cell mats have been used for screening human tumours for sensitivity to anticancer drugs.

The use of *in vitro* systems (subcellular fractions, cell lines, primary cell cultures, tissue slices, organ cultures, etc.) as research tools in toxicology is widespread (1). In the past few years the apoptosis phenomena are followed by very precise intracellular changes, where through programmed cell death, a cell can be removed from a population. This regulated physiological process is followed by DNA fragmentation, nuclear blebbing, and cell shrinkage or by a number of marker enzymes. The molecular pathways that regulate apoptosis provide a possible target for oxidative damage, suggesting a role for transcriptional regulation of genes implicated in cell survival. Cells and tissues are subjected to oxidative stress when the concentration of reactive oxygen species (ROS) exceeds the antioxidant capability of the cell. The ROS levels can be modulated by changes in endogenous cytochrome P450s or exogenously following the administration of drugs, hormones, xenobiotics or toxicants.

In recent years there has also been a move toward standardizing the various techniques and procedures available, with the concomitant development and evaluation of numerous *in vitro* tests which have been proposed as potential replacements for various animal tests currently required by regulatory authorities (2, 3). There are four main factors driving the development of *in vitro* toxicology (4):

- the need for relatively simple, inexpensive and efficient systems for testing large number of chemicals for which the toxicological data are required,
- public and legislative pressures to reduce animal experimentation,
- the need for a better understanding of the mechanisms of chemical-induced toxicity, in order to improve the scientific basis of current risk assessment procedures, and
- the desire to use human cells and tissues whenever possible, for example to undertake inter-species comparisons of xenobiotic metabolism and toxicity.

VALIDATION OF TOXICITY TEST PROCEDURES

Validation is the process by which the reliability and relevance of a test method are evaluated for the purpose of supporting a specific use. The approaches and methods conform to scientific principles of objectivity and appropriate experimental design.

A test is considered validated when its performance characteristics, advantages and limitations have been adequately determined for a specific purpose. The measurement of a test reliability and relevance are independent stages in the validation of the test method, and both are required. Reliability is an objective measure of an intra- and interlaboratory reproducibility of the method. If the test is not sufficiently reliable, it cannot be used for its intended purpose. Alternatively, if the test is not relevant, or is of a questionable relevance to the biological effect of interest, or if it is not an appropriate measure of the effect, its reliability is academic. The relevance of the test may be linked to the mechanism of the toxic effect it measures and to its proposed uses. The measures of relevance of the test include the calculated

operational characteristics (e.g., sensitivity, specificity, etc) or statistically derived correlation coefficients and determinations of the mechanistic association of the measured effects with the toxic events of interest.

There are no optimum or minimum levels of reproducibility or association with the event of interest that must be reached for the test to be considered "validated". The conditions under which the test will be used and the purposes to which its results will be applied will determine the levels of reliability and relevance that are needed.

The new tests can be designed either as substitutes to replace, or to be interchangeable with, currently accepted tests, or as tests that have no correlate with currently used tests or endpoints. "Definitive tests" provide data that are used to measure toxic effects or unequivocally identify the hazardous substances and asses the risks posed by exposure to them. "Screening methods" are generally used to make preliminary hazard decisions (i.e., identify potential adverse effects), or to select chemicals or set priorities for other, more definitive tests. They often provide only a qualitative or semi-quantitative response and are generally not designed to serve as definitive tests. In contrast, "adjunct tests" are used to increase information base and/or aid the interpretation of results from other, definitive methods. They are not used in isolation or as substitutes for definitive methods, but they often support the relevance of the definitive test method by providing information related to the mechanism of toxicity. For example, a test showing that relevant metabolic pathways are similar in the test system and in the species of interest supports the use of information from the definitive test system for hazard and risk assessments. These tests can be developed for newly identified endpoints or effects, or they can be used to replace existing adjunct methods (5).

Because the data from a substitute test will be used in lieu of a currently used test, its adoption requires evidence from validation studies that the use of the method will provide a comparable or better level of protection of human health or of the environment than the current methods or approaches. Often, new tests are developed that identify or provide data about toxicological effects not addressed by existing methods. These new tests should be based on specific biological mechanisms or enpoints related to an effect of concern. The test itself often will help define the effect. When a new method is designed to measure an effect that is newly discovered or not well defined, there usually is no benchmark against which the usefulness or effectiveness of such a method can be judged.

Methods may be designed to stand alone, or as components of tiers, batteries, or hierarchical testing strategies, e.g. stepwise sequence of tests from simple to more complex ones. The process is leading to the scientific validation of stand-alone tests, or tests that are used only as a component of a test.

Validation is a scientific process designed to characterize the operational characteristics, advantages, and limitations of a test method, and to demonstrate its reliability and relevance. The designation of a test as "validated" or "not validated" for a specific purpose is not irrevocable; subsequent data and experience with the test can lead to a loss or affirmation of its validation status. Also, the test method could be considered validated for a specific use, but not for other uses. The criteria for validation of the test method are the function of the purpose for which the test method will be used. For example, the mechanisms of some effects are known or are relatively straightforward (e.g., skin corrosivity, estrogen receptor binding), while others (e.g., carcinogenicity, developmental toxicity) are complex and multi-faceted, or not well understood. The validation of tests for these different types of effects requires different approaches (6).

Critical to the validation process are the standardized test protocol and the ability of component laboratories to perform the test. "Prevalidation" is the process by which testing laboratories are selected and demonstrate competence in performing the testing procedures, and during which the test protocols are standardized. If a protocol cannot be standardized or reproduced using known chemicals, it cannot be validated.

A prerequisite for the performance and evaluation of any validation study is a formal protocol or procedural manual that can be readily understood and followed by individuals in other laboratories, and by administrative and scientific review personnel. This protocol should clearly state the purpose of the test. All test responses, regardless of whether they are in humans, animals, or cultured cells, contain a certain level of between-animal or between-culture variability, which may or may not be defined, that must be considered when evaluating the performance of the candidate method. Generally, the variability of *in vivo* methods is greater than that of *in vitro*

methods because of the wider degree of genetic and physiological diversity among whole animals.

When evaluating a new method, there must be a sufficient number of chemicals to demonstrate the test's performance within a chemical class or among a range of chemical classes or products and among substances of different reactivities. Other considerations are the cost and complexity of the method; for example, an *in vitro* test to identify oestrogen receptor-binding chemicals would require less time and resources to perform than an *in vivo* rodent reproductive test. These cost and time considerations would determine the numbers of chemicals that could reasonably be tested in a validation exercise.

In summary, the specific goals of the validation study and the hypotheses to be tested must be clearly defined. The test method must be shown to be reproducible and understandable in the context of science and, for substitute tests, the procedure should offer an advantage over the currently accepted procedures. Because tests can be designed and used for different purposes (e.g., as substitutes or screens) by different organizations, and with varying categories of substances, the test validation process should be highly flexible and adapted to the specific test and its proposed use. Despite this need for flexibility, all the various factors that make up a validation process must be included.

ALTERNATIVE METHODS DATABASES

It is a fundamental requirement that any scientist who intends to start a new project is well-informed about the current status of the proposed field of investigation. The animals should only be used in a study if such use is justifiable after all possible alternatives have been identified and found to be inadequate. Ready access to relevant information on the proposed topic of study can prevent unnecessary duplication of work, thereby contributing to a reduction in animal use.

The bibliographic and factual databases can be a powerful information tool, and the database searches are an essential part of a stepwise approach to toxicity testing. The field of toxicology is, in fact, an area of continuous investigation with respect to the identification of new directions in database development for the storage and dissemination of toxicological information.

There are many web sites that mention the alternatives only briefly as part of the scope of laboratory animal care. However, few are devoted exclusively to the alternative concept. For the researcher, educator or student in search of information on the alternatives, wandering from site to site for the best full-text information can be a challenge. The list of some web sites that follows emphasizes the alternatives and the sites that provide unique information developed by the host institution (7-9).

- *Alternatives to Skin Irritation Testing in Animals*, Toxicology Consulting Services, Plainsboro, USA; *http://www.invitroderm.com*. This non-profit web site is developed to display information about alternatives to skin irritation testing in animals.
- *Altweb*, John Hopkins Center for Alternatives to Animal Testing (CAAT), Baltimore, USA; *http://caat.jhsph.edu*. This is very extensive site for news, information, discussion and resources from the field of alternatives to animal testing.
- *Animal Welfare Information Center (AWIC*), Beltsville, USA; *http://www.nal.usda.gov/awic*. The AWIC was established in 1985 as part of an amendment to the Animal Welfare Act. The AWIC's mission is to provide training about more humane animal care and use for the researchers, who use animals, and to provide information for improving methods of the experimentation on animals that can reduce or replace the use of animals or minimize pain or distress to animals.
- *Center for Alternatives to Animal Testing (CAAT*), Baltimore, USA; *http://caat.jhsph.edu*. The CAAT was founded in 1981 as a resource of alternatives and a granting organization. It has created and maintained the Altweb site.
- *European Centre for Validation of Alternative Methods (ECVAM)*, Ispra, Italy; *http://ecvam.jrc.cec.eu.int/index.htm*. The ECVAM was established in 1993, and was created to play a leading role at the European level in the independent evaluation of the relevance and reliability of tests and test strategies for specific purposes.
- *Fund for the Replacement of Animals in Medical Experiments (FRAME),* Nottingham, The United Kingdom; *http://www.frame-uk.demon.co.uk*. The FRAME is founded in 1969, and is

charitable trust that advocates the 3 Rs approach (reducement, replacement, refinement) by encouraging realistic consideration of ethical and scientific issues. The FRAME web site provides an original section on searching for alternatives.

- *Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM),* Research Triangle Park, USA; *http://iccvam.niehs.nih.gov*. The ICCVAM coordinates development, validation, acceptance, and harmonization of alternative toxicological test methods to improve toxicity characterization, increase savings in time and cost, and implement the 3 Rs. The web site includes full texts of current ICCVAM reports, meeting proceedings, information about testing method development and methods under review.
- *Information on Alternative Databases*; *http://oslovet.veths.no/databasesintro.htlm*. This collection links to more than 20 databases.

ALTERNATIVE APPROACHES FOR PREDICTING TOXIC HAZARD

The scientific and regulatory acceptance of alternative methods are of importance to promote the 3 Rs (reduce, refine or replace) on the use of laboratory animals. One of the main priorities was the implementation of procedures, which would enable it to become well-informed about the state-of-the-art of non-animal test development and validation, and the potential for the possible incorporation of alternative tests into regulatory procedures (10).

Quantitative Structure-Activity Relationships (QSAR) might be used in the design, evaluation and validation of *in vitro* tests and in the selection of appropriate test chemicals for validation studies. The principles underlying the development of QSARs are based on the premise that the properties of a chemical are implicit in its molecular structure. Therefore it follows that, if the mechanism for the activity of a group of chemicals can be elucidated, and relevant parameters can be measured or calculated, then, theoretically, a QSAR can be established. For the QSAR to be valid and reliable, the activities of all the chemicals covered must be elicited by a mechanism, which is both common and relevant. Attempts to derive QSARs for data sets where this is not the case have not always been successful. The same principles need to be applied to the development of *in vitro* toxicity tests. However, in many cases, such principles are overlooked. As a result, some alternative tests predict endpoints which are different from those which they claim to predict (because the wrong mechanism has been identified) or they cannot predict endpoints accurately for all classes of chemicals (because a common mechanism is lacking).

In terms of developing integrated approaches to the use of alternative methods for the prediction of toxic hazard, it is important to consider how the QSARs, computer modelling techniques (for example, physiologically based pharmacokinetic /PBPK/ models) and *in vitro* methods might be used in combination (11).

The QSAR is a model which relates the biological activities of a series of similar compounds to one or more (physico)chemical or structural properties of the compounds. In this definition, "similar" means having the same mechanism of action, but not necessarily having a related chemical structure. However, it is often difficult (if not impossible) to determine the mechanism of action, whereas it is usually less difficult to establish chemical similarity. Hence, the QSARs are generally developed for sets of chemically similar compounds (congeneric series), hoping that they will also have the same mechanism of action. Any compounds which do not act by the same mechanism are likely to fit the correlation only poorly and to appear as "outliers".

When a chemical is administered to an organism, two events must occur for a biological response to be triggered. Firstly, the compound has to be transported to the site of action (the "receptor") and, secondly, it must interact with the target in an appropriate manner. The interaction with the target (receptor) is governed largely by two factors: the size and shape of the xenobiotic, which will control how well the molecule fits the receptor site and the nature and relative positions of appropriate functional groups on the molecule, which will affect the type and strength of the interaction with complementary groups on the receptor. Many physicochemical properties can be used to model receptor interaction (12).

It is important to note that, in effect, the QSAR is concerned with the *change* in biological activity brought about by a change in chemical structure. But, there are a number of caveats to the use of the QSARs, therefore they have to be borne in mind,

otherwise the application of the QSARs may not be valid. Some of them are:

- The QSARs can be applied currently only to pure compounds. There are yet no firm guidelines for the use of mixtures in this respect.
- The set of compounds used to derive the QSAR (*training set*) should be selected from knowledge, or assumption, of a common mechanism of action.
- If possible, the QSAR parameters should be selected on the basis of mechanistic considerations. Alternatively, they should be amenable to mechanistic interpretation.
- For comparative purposes, concentrations or doses must be in molar, not weight, units.
- Each QSAR should be validated by investigating its predictive ability using a different set of compounds (*test set*), which could cover the same ranges of parameter space.
- The QSAR must not be applied outside of its domain of validity (i.e. outside of the parameter space covered by the training set).

Future requirements for the development of *in vitro* tests include more attention being focused on the elucidation of mechanisms of toxicity *in vivo*, the use of the QSARs in the rational selection of test chemicals for the development of integrated testing strategies which combine information on toxicological effects from different sources, including human studies where appropriate (13).

What are the ways of achieving a better integration of the QSAR and *in vitro* methods? The *in vitro* tests fall into three categories: *empirical* – those for which no clear mechanistic basis can be identified; *mechanistic* – those with a clear mechanistic basis; and *analogous* – those in which the *in vivo* test system is essentially reproduced *in vitro*. The QSAR methods may help to identify the mechanisms operating in *in vitro* assays. These mechanisms should be defined in operational terms, i.e. in terms of the level of organisation (molecular, cellular, tissue, organ, whole organism) at which they are working in the animal. From this information it should be possible to judge the "mechanistic relevance" of the *in vitro* test.

The results from appropriate *in vitro* tests may also be used as molecular descriptors to develop the QSAR equations. For example, the *in vitro* cytotoxicity data can be used to model a toxic endpoint at the level of the whole organ. Although data from *in vitro* tests can be used to construct QSARs, it is also important to use existing animal data where available; the predictions obtained from such QSARs could then be used to improve the *in vitro* tests *via* definition of relevant mechanisms. A single QSAR or *in vitro* test cannot be used directly to extrapolate from one level of organisation to another. However, the families of QSARs or *in vitro* tests may be used to define rate-limiting parameters, which govern effects at different levels; for example, to relate the inhibition of acetylcholinesterase activity to lethality in the whole organism.

Current attempts to validate *in vitro* methods as alternatives to animal tests are based on the premise that *in vitro* potency can be related directly to *in vivo* potency. However, it is of fundamental importance to recognize the general inapplicability of this premise, since the relationships between *in vitro* and *in vivo* potencies can be expressed only in terms of probabilities. The practical significance of this for the validation of alternative methods is that, in the absence of a clear mechanistic relationship between *in vitro* and *in vivo* potencies, such validation must proceed in two stages:

- Stage 1: establishment of a general statistical relationship between the responses in the *in vitro* and *in vivo* tests. This stage must be performed using a wide range of chemicals, with different physicochemical properties, acting *via* different mechanisms of toxicity.
- Stage 2: establishment of a specific relationship between the *in vitro* and *in vivo* responses using separate models for each chemical class, which is defined by a common mechanism of toxicity.

It follows that, where a clear mechanistic relationship between *in vitro* and *in vivo* potencies has already been established, validation can begin in stage 2.

Generally, it is recommended that (14):

- The development of integrated testing strategies for predicting the systemic toxicity of chemicals should be given greater priority, due to animal welfare and economic considerations.
- The most appropriate statistical techniques must be used to analyze the data, taking into account whether the data are continuous or discrete.
- Depending on the claims made for the applicability of an *in vitro* test, a range as wide as possible of relevant chemicals must be selected to assess its utility. The known boundary conditions in test performance/interpretation must be taken

into account during the selection of the test chemicals wherever possible.
- The development and validation of expert systems for predicting toxicity and metabolism should be encouraged.

CURRENT ALTERNATIVES FOR REPRODUCTIVE TOXICITY

There has been a great deal of research and debate during the past 20 years concerning the use of alternatives to living mammals for testing the potential reproductive toxicities of chemical and physical agents and mixtures.

Reproduction is a continuous cycle, but for purposes of toxicity testing it is broadly divided into pregnancy in females (during which period a pre-natal and post-natal developmental toxicity may be induced) and the remainder of the cycle in both males and females (during which period fertility may be impaired). The majority of research into the development of alternative tests has concentrated specifically on teratogenicity, which, in turn, is one aspect of developmental toxicity. This is, no doubt, because of the potential complexity of, and multiple targets for, adverse effects on fertility (which include effects on sexual behaviour, spermatogenesis, oogenesis, hormonal activity or physiological response) could interfere with the capacity to fertilise, fertilisation itself, or the development of the zygote up to and including implantation.

Due to the complexity of the reproductive cycle, from gamete maturation to the implantation of the early embryo into the maternal uterus, and because of the lack of validated alternative tests for most steps in the cycle, testing in living animals is the only option currently available for assessing the possible effects of chemicals on reproduction. It is recommended that one or more of the following procedures are included: for males, detailed histological examination of the testicles and epididymides, semen analysis (15) and/or flow cytometric analysis of spermatogenic cell types; for females, detailed histological examination of the ovary and vaginal cytology (16).

In testing female fertility, alternative approaches still exist. Some aspects of female reproductive function can be modelled *in vitro*, and several cellular components of the female reproductive organs can be maintained in culture. Ovarian somatic cells can be maintained in culture, and any adverse effects can be assessed by examining cell morphology, and by determining cell viability (17) and hormonal responsiveness.

According to the assessments of the male fertility, several testicular cell types can be maintained in culture, either alone or in combination. These include Sertoli-germ cell co-cultures, Sertoli cell-enriched cultures, germ cell-enriched cultures, Leydig cell cultures, and Leydig-Sertoli cell co-cultures. All of these systems have been used successfully to study specific features of testicular toxicity. Primary cultures of testicular cells retain many of the differentiated characteristics of their *in vivo* counterparts, but they are inherently variable and generally have only a limited lifespan, which necessitates the frequent preparation of fresh cultures. The ability to study individual cell populations from a heterogeneous target organ such as the testis is a powerful tool for probing mechanisms of toxicity (18).

The mammalian embryos can be maintained in culture for short periods through the phase from fertilization to the end of organogenesis. Screening systems using mouse and rat embryos have been proposed, and the culture of rabbit embryos has recently been optimised. At the end of the culture period, a number of endpoints can be measured, including: effects on the development of the visceral yolk sac vascularisation and circulation; effects on haematopoiesis, embryonic growth and differentiation; dysmorphogenic effects. The interpretation of the results obtained takes into account the adverse effects on yolk sac development, embryonic growth and differentiation, as well as specifically on dysmorphogenesis.

However, the system has clear limitations (18):
- It is relatively complex, covers only a part of organogenesis, and requires a high level of technical skill.
- The test can be costly, and it uses mammalian tissue and serum.
- Whether it is justified or not, its use as a screening tool should be evaluated by including it in a comparative trial with other simpler *in vitro* systems. It is likely that these different systems will provide complementary information and selection of the appropriate test for a particular application should be made on a case study basis.

BIOMARKERS IN TOXICITY TESTING

The term "biomarker" has grown in popularity in recent years, and so has its meaning. It ranges from exposure measurements and biological indices, which support a mechanistic postulate, to clinical markers with diagnostic implications. Some meanings of the term are given below:

- "Biological markers are indicators signalling events in biological systems or samples" and "Biological markers are measurements of body fluids, cells or tissues that indicate, in biological or cellular terms, the presence and magnitude of toxicants or of host responses" (19).
- "Biomarkers are cellular, biochemical, or molecular alternations, which are measurable in biological media such as human tissues, cells or fluids" (20).
- "A biomarker is a xenobiotically induced variation in cellular or biochemical components or processes, structures, or functions, that is measurable in a biological system or sample" (21).
- "Biomarkers are parameters that putatively represent some step along the causal pathway between exposure and effect" (22).
- "A biomarker is a parameter, which can be evaluated quantitatively, semi-quantitatively or qualitatively, and which provides information on exposure to a xenobiotic, or on the actual or potential effects of that exposure in an individual or in a group" (23) etc.

The meaning for which the term biomarker is used clearly depends upon the context and this is reflected most clearly in the parameters of the database which is used as the basis for any search.

The biomarker of *exposure* could be predictive of risk, if sufficient information were available on the dose-response relationship. The biomarkers of *effect* cover a range of measurements which may be indicative of exposure to a particular agent, although their specificity is generally lower than that of biomarkers of exposure. Biomarkers of *susceptibility* most notably include genetic susceptibility.

The validation of biomarkers is not directly comparable with the validation of *in vitro* tests, but must follow the basic principle of demonstrating reproducibility, reliability and fitness-for-purpose. The actual process will vary, depending on the type of biomarker and its intended use. Although biomarkers currently cannot replace or reduce the use of animals in toxicity testing, future technological developments show great promise for making this possible.

The *in vitro* techniques may play an important role in the development of biomarkers of effect, as the greater control which is possible in *in vitro* studies facilitates the investigation of mechanisms of toxicity, and therefore the identification of key events. The mechanistic understanding of the progression of a toxic effect is a necessary prerequisite for the use of biomarker approaches in the prediction of toxicity. It is possible to envisage a natural progression, in which a potential biomarker is identified during *in vitro* mechanistic studies. The value tested in animal models is then incorporated in *in vivo* testing protocols (refinement of animal testing), and is subsequently validated for use in *in vitro* toxicity tests (leading to reduction and replacement of animal testing).

At present, it is impossible to reduce, refine and replace all animal experiments by applying biomarkers in environmental bioassays for toxicological evaluation. The acquisition of knowledge, for example, as a result of genomic and proteomic technologies, should make this feasible in the future (24).

CONCLUSIONS

The *in vitro* systems are ideally suited for investigations of the molecular, cellular and physiological mechanisms of chemically induced toxicity, which cannot readily be studied *in vivo* for known target organ and target species toxicity studies and for answering specific questions about toxic effects. The main justification for developing *in vitro* toxicity tests is that they will make toxicology a more scientifically-based practice. The understanding of the mechanisms by which chemicals cause cell and tissue damage, and the reasons for the increased susceptibility of certain species, individuals or tissues to particular chemicals, will markedly improve our ability to predict possible consequences of human and/or environmental exposure to them. Perhaps the greatest advantage of *in vitro* toxicity tests is that human cells and tissues can be used, thereby obviating the need to extrapolate data from laboratory animals to man.

The greatest progress in the use of *in vitro* test systems has been achieved in the area of local ocular and dermal toxicity and target organ toxicity, especially in testing for hepatotoxicity, nephrotoxicity

and neurotoxicity. Despite numerous comparisons of *in vitro* cytotoxicity data with rodent LD_{50} values, there has been little progress in the use of *in vitro* tests for predicting overall systemic toxicity, for which the need to incorporate data on biokinetics in particular is a critical issue.

It is increasingly apparent that the development and incorporation of stepwise testing strategies, combining experimental data from a range of alternative methods (physicochemical techniques, quantitative structure-activity relationships - QSAR, metabolic and kinetic modelling and *in vitro* tests), provide the most progressive way to predict toxicity, reducing at the same time the number of laboratory animals used for testing purposes.

REFERENCES

1. Watson RR, editor. *In vitro* methods of toxicology. Boca Raton (FL): CRC Press; 1992.
2. O'Hare S, Atterwill CK, editors. *In vitro* toxicity testing protocols. Totowa (NJ): Humana Press Inc.; 1995.
3. Gad SC. Alternatives to *in vivo* studies in toxicology. In: Ballantyne B, Marrs T, Turner P, editors. General and Applied Toxicology. Volume 1. Basingstoke (UK): Macmillan; 1993. p. 179-206.
4. Garle MJ, Fentern JH, Fry JR. *In vitro* cytotoxicity tests for the prediction of acute toxicity. Toxicol In Vitro 1994;8:1303-12.
5. Walum E, Clemedson C, Ekwall B. Principles for the validation of *in vitro* toxicology test methods. Toxicol In Vitro 1994;8:807-12.
6. Balls M, Blaauboer BJ, Fentem JH, Bruner L, Combes RD, Ekwall B, Fielder RJ, Guillouzo A, Lewis RW, Lovell DP, Reinhardt CA, Repetto G, Sladowski D, Spielman H, Zucco F. Practical aspects of the validation of toxicity test procedures. Altern Lab Anim 1995;23:129-47.
7. Kreger M. Alternative methods databases: Internet resources that are more than links to other sites. AWIC Bulletin 2000 [cited June 2004]. Available from: http://www.nal.usda.gov/awic/newsletters/
8. Janusch A, van der Kamp MDO, Bottrill K, Grune B, Anderson DC, Ekwall B, Howald M, Kolar R, Kuiper JD, Larson J, Loprieno G, Sauer UG, Smith AJ, van der Valk JBF. Current status and future developments of databases on alternative methods. Altern Lab Anim 1997;25:411-22.
9. Todd M, Garthoff B, Bernhardt W, Bode G, Farr S, Kolar R, Lumley C, Pieters M, Rogiers V, Sauer U, Spielman H, Straus J, Warngard L. ATLA 1998;26:13-20.
10. Anon. ECVAM News&Views. Altern Lab Anim 1994;22:7-11.
11. Cooper-Hannan R, Harbell JW, Coecke S, Balls M, Bowe G, Cervinka M, Clothier R, Hermann F, Klahm LK, de Lange J, Liebsch M, Vanparys P. The principles of good laboratory practice: Application to *in vitro* toxicology studies. Altern Lab Anim 1999;27:539-77.
12. Testing MFW, Baumans V, Combes RD, Halder M, Hendriksen CFM, Howard BR, Lovell DP, Moore GJ, Overend P, Wilson MS. Reducing the use of laboratory animals in biomedical research: problems and possible solutions. Altern Lab Anim 1998;26:283-301.
13. Lumley CE, van Cauteren H. Harmonisation of international toxicity testing guidelines for pharmaceuticals: contribution to refinement and reduction in animal use. EBRA Bulletin November 1997. p. 4-9.
14. Balls M, Goldberg AM, Fentem JH, Broadhead CL, Burch RL, Festing MFW, Frazier JM, Hendriksen CFM, Jennings M, van der Kamp MDO, Morton DB, Rowan AN, Russell C, Russell WMS, Spielman H, Stephens ML, Stokes WS, Straughan DW, Yager JD, Zurlo J, van Zutphen BFM. The three Rs: the way forward. The report and recommendations of ECVAM workshop 11. Altern Lab Anim 1995;23:838-66.
15. Kniewald J, Jakominić M, Tomljenović A, Šimić B, Romac P, Vranešić Đ, Kniewald Z. Disorders of male reproductive tract under the influence of atrazine. J Appl Toxicol 2000;20:61-8.
16. Anon. OECD Guidelines for testing of chemicals, Guideline 422: Combined repeated dose toxicity study with the reproduction/developmental toxicity screening test. Paris: OECD; 1993.
17. Kmetič I, Gaurina-Srček V, Šimić B, Kniewald Z, Kniewald J. Cytotoxic effects of organochlorine insecticides in Baby Hamster Kidney (BHK 21 C13) cell line. In: Kniewald Z et al., editors. Current Studies of Biotechnology. Volume 3 - Food. Zagreb: Croatian Society of Biotechnology and Medicinska naklada; 2003. p.177-84.
18. Brown NA, Spielmann H, Bechter R, Flint OP, Freeman SJ, Jelinek RJ, Koch E, Nau H, Newall DR, Palmer AK, Renault J-Y, Repetto MF, Vogel R, Wiger R. Screening chemicals for reproductive toxicity: the current alternatives. Altern Lab Anim 1995;23:868-82.
19. Nuclear Regulatory Commission. Biological markers in environmental health research. Environ Health Perspect 1987;74:3-9.
20. Hulka BS, Wilcosky T. Biological markers in epidemiologic research. Arch Environ Health 1989;43:83-9.
21. Willet WC, The use of biomarkers in nutritional epidemiology. In: Kok FJ, van't Veer P, editors. Biomarkers of dietary exposure. London (UK): Smith-Gordon; 1991. p.9-14.
22. Greenwald P, Witkin KM, Malone WF, Byar DP, Freedman LS, Stern HR. The study of markers of biological effects in cancer prevention research trials. Int J Cancer 1992;52:189-96.

23. Bottrill K. The use of biomarkers as alternatives to current animal tests on food chemicals. Altern Lab Anim 1998;26:421-80.
24. Benford DJ, Hanley AB, Bottrill K, Oehlschlager S, Balls M, Branca F, Castegnaro JJ, Descontes J, Hemminiki K, Lindsay D, Schilter B. Biomarkers as predictive tools in toxicity testing. Altern Lab Anim 2000;28:119-31.

Sažetak

ALTERNATIVNI MODELI ISPITIVANJA TOKSIČNOSTI KSENOBIOTIKA

Alternativni sustavi ispitivanja toksičnosti na životinjama obuhvaćaju postupke koji uključuju stanice i kulture, uporabu dijelova tkiva, toksikokinetičko modeliranje, kao i odnose aktivnosti prema strukturi ksenobiotika. Uporaba *in vitro* sustava (staničnih frakcija, staničnih linija, primarnih staničnih kultura, dijelova tkiva, kulture organa itd.) ima široku primjenu u toksikološkim istraživanjima. U posljednjih nekoliko godina fenomen apoptoze može se pratiti prema vrlo preciznim unutarstaničnim promjenama te utvrđena programirana smrt stanice iskoristiti u uklanjanju takvih stanica iz populacije. *In vitro* sustavi idealno se rabe u istraživanjima molekularnih, staničnih i fizioloških mehanizama toksičnosti izazvanih kemikalijama, dok *in vivo* studije toksičnosti ne mogu dati te odgovore. Glavna potvrda razvoja i praćenja toksičnosti spojeva u *in vitro* testovima jest znanstvena utemeljenost toksikologije. Razvoj i uvođenje strategija testiranja koje primjenjuju kombinaciju eksperimentalnih podataka iz niza alternativnih metoda (fizikalno-kemijske tehnike, QSAR tehnike, metaboličko i kinetičko modeliranje) omogućuje najefikasniju procjenu toksičnosti, kao i istodobno smanjenje broja laboratorijskih životinja potrebnih u postupcima testiranja toksičnosti spojeva.

KLJUČNE RIJEČI*: alternativne metode, biomarkeri,* in vitro *metode, predviđanje toksičkih rizika, provjera valjanosti, testiranje toksičnosti*

REQUESTS FOR REPRINTS:

Professor Jasna Kniewald, Ph. D.
Faculty of Food Technology and Biotechnology, Zagreb University
Pierottijeva 6, HR-10000 Zagreb, Croatia
E-mail: *jkniewal@pbf.hr*