High information content technologies in support of read-across in chemical risk assessment

Technical Report No. 109

Brussels, December 2010

ISSN-0773-8072-109 (print)
ISSN-2079-1526-109 (online)

ECETOC TECHNICAL REPORT No. 109

© Copyright – ECETOC AISBL
European Centre for Ecotoxicology and Toxicology of Chemicals
4 Avenue E. Van Nieuwenhuyse (Bte 6), B-1160 Brussels, Belgium.

All rights reserved. No part of this publication may be reproduced, copied, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the copyright holder. Applications to reproduce, store, copy or translate should be made to the Secretary General. ECETOC welcomes such applications. Reference to the document, its title and summary may be copied or abstracted in data retrieval systems without subsequent reference.

The content of this document has been prepared and reviewed by experts on behalf of ECETOC with all possible care and from the available scientific information. It is provided for information only. ECETOC cannot accept any responsibility or liability and does not provide a warranty for any use or interpretation of the material contained in the publication.

High information content technologies in support of read-across in chemical risk assessment

CONTENTS

SUMMARY 1

1. INTRODUCTION 2
1.1 Terms of reference 3

2. READ-ACROSS APPROACHES 5

3. DEVELOPING AND TESTING HYPOTHESES BASED ON ANALOGUE SEARCHES 6

4. READ-ACROSS BASED ON A HYPOTHESIS OF COMMON BIOLOGICAL ACTIVITY AS A COMPARABLE COMPOUND 9
4.1 Read-across based on a hypothesis of similar biological activity 9
4.1.1 Toxicogenomics for phthalates 9
4.1.2 High-throughput assays for oestrogens 10
4.1.3 Examples for reactive chemistry 10

5. CONCLUSIONS AND RECOMMENDATIONS 11
5.1 Recommendations 11
5.1.1 Better analogue identification and hypothesis generation 11
5.1.2 Demonstrations of the use of high-information content data to support read-across hypotheses 12

BIBLIOGRAPHY 13

APPENDIX A: PARTIAL LIST OF DATABASES AND RESEARCH TOOLS TO SUPPORT READ-ACROSS 23

APPENDIX B: METABOLOMICS AS A USEFUL TOOL FOR READ ACROSS IN CHEMICAL RISK ASSESSMENT 37

APPENDIX C: READ-ACROSS IN THE ASSESSMENT OF SKIN SENSITISATION 53

APPENDIX D: REQUEST FOR RESEARCH PROPOSALS 70

MEMBERS OF THE TASK FORCE 72

MEMBERS OF THE SCIENTIFIC COMMITTEE 73

SUMMARY

Read-across exploits information on structurally related (similar) analogues to derive hypotheses about the activity of the new chemical and hence predict its toxicity without experimental testing. Large existing databases on traditional toxicological endpoints and mechanisms of action are available that can be searched by data mining and cheminformatics tools (a selection is presented in the report). In addition, high-information-content techniques such as ‛omics (toxicogenomics and metabolomics in particular) can be utilised to generate and test these hypotheses, notably about the mechanism of action. Examples are given in the report for phthalates, oestrogens and skin sensitisers.

There is scope for improvement of the heuristics of analogue identification and hypothesis generation. Furthermore, real examples of using high-information-content data are needed to support read-across, e.g. to provide a biology based rationale for chemical grouping.

1. INTRODUCTION

Regulatory toxicology testing first became formalised in the late 1930s, following the sulphanilamide tragedy. The goal of toxicity testing is to generate information that identifies hazard, estimates the exposure level at which risk is negligible, and enables risk management measures to be developed, thus ensuring that public health is appropriately protected from the adverse effects of exposures to environmental agents. Toxicity testing is aimed at identifying chemicals with the potential to produce human diseases, such as birth defects, cancer, and organ toxicity. A battery of tests was established that screen for toxicity to most organ systems, covering numerous life stages, and some biological functions (e.g. reproductive function, some aspects of nervous system function). Current approaches to toxicity testing still rely on observing adverse biologic responses in animals. Such studies are costly, resource intensive and can use large numbers of animals, hence only a small proportion of chemicals out of all of those in commerce have been evaluated using these approaches. The information from these studies is integrated with human data (if any) as well as estimates of human exposure.

Despite the expense over the past 50 years of using animal testing, the field of applied toxicology has generated a substantial amount of data on adverse effects for thousands of chemicals, and toxicology research has identified a large number of mechanisms of toxicity. This large data set, coupled with the computational capability to manage and query the data in novel ways (e.g. searching for structural features of chemicals) has brought us to the point where we can potentially approach toxicity testing needs in a different way. Instead of considering each new chemical as an unknown entity, we can instead use the information about the toxicity of related materials to generate hypotheses about the nature of the toxicity of the new chemical, and to design an assessment strategy (which may involve experimentation) that tests the hypotheses. Very often the similarities with existing analogues are sufficiently robust that little if any experimental work is needed to support hypotheses about similar activity. (For example, when mechanism of action is known, targeted biochemical or *in vitro* assays may provide all the information necessary to support read-across.)

This is, in effect, the basis of chemical grouping approaches, expert rule-driven structure activity relationship and some quantitative structure activity relationship (QSAR) approaches where the underlying hypothesis is that similar chemicals will have similar toxicity profiles. The application of this structure activity method to predict toxicity is known as read-across. Read-across is an approach used to predict toxicity endpoint information for an untested chemical (the target compound) by using existing data from the same endpoint from another chemical(s) which is considered to be ‘similar’ on the basis of molecular structure, chemical properties and/or other activities. Read-across is therefore a data gap filling exercise based upon existing test results from closely related chemicals (analogues) which together form a common group or chemical category.

The challenge lies in characterising the context of similarity in terms of the pertinent chemical characteristics. Data mining and cheminformatics techniques on existing toxicity data provide one approach in generating hypotheses for read-across. Emerging technologies such as toxicogenomics, proteomics or metabolomics allow investigators to evaluate globally the changes in gene or protein or endogenous metabolite expression in target cells or tissues or even in remote tissues like blood that can be easily sampled. These approaches can be tremendously useful adjuncts both in generating hypotheses for read-across, and in testing those hypotheses. These technologies can be considered to be high information content technologies, in that a single experiment generates tens of thousands of data points that are highly relevant for understanding the biological response to an external perturbation. High-throughput assays (assays that are automated so that hundreds of chemicals can be evaluated at one time) can also be considered to be high-information content assays, especially when a series of such assays is run.

With the advent of legislation such as REACH in Europe and proposed revisions to TSCA in the US, the generation of toxicity data and the use of that data for risk assessment purposes are undergoing a major change. Systematic read-across supported by SAR and QSAR approaches are expected to be applied more routinely. Data mining/cheminformatics approaches for analogue identification and hypothesis generation are potentially powerful tools to facilitate read-across but these rely on a reasonable volume of data. There have been a number of initiatives, both public and private, to compile and store toxicology data in a manner that facilitates the efficient search and retrieval of data and chemical structure information. Several relational toxicology databases now exist, many of which are freely accessible. These databases contain publicly available data from the published and grey literature and are replete with data entries for endpoints of toxicity that are routinely considered such as genotoxicity, carcinogenicity, reproductive toxicity and repeated-dose toxicity. In addition to traditional toxicology information, these databases access the results of high-throughput assays, e.g. those that are being conducted by the US National Centre for Chemical Genomics. These data sets, together with supplementary information about physical chemistry properties, provide a rich foundation for hypothesis generation about the potential biological effects of new chemicals.

1.1 Terms of reference

With these considerations in mind, the Long-range Research Initiative (LRI) of the European Chemical Council (CEFIC) requested ECETOC to develop and position a suitable request for research proposals (RfP) on “High information-content technologies for read-across in chemical risk assessment”. Consequently the ECETOC Scientific Committee established a Task Force with the following terms of reference.

- Review how to employ high-throughput screening techniques (e.g. 'omics) in the application of existing *in silico* methods (e.g. enzyme kinetics modelling, systems biology) used for performing read-across in chemical hazard identification/characterisation.
- Demonstrate their use in several practical case studies, highlighting their potential strengths and limitations.
- Review how to evaluate the relative robustness of the results and adequacy of the read-across approach.
- Identify gaps in knowledge and priorities for research.
- Present a synopsis (report) and recommend suitable topics (Request for Proposals) to LRI.

The RfP should aim at a mechanistic concept using 'omics tools and *in silico* methods, unlike the empirical ToxCast approach followed by the US EPA.

2. READ-ACROSS APPROACHES

Use of chemical categories and analogue read-across is well recognised in the regulatory arena as useful approaches to address data gaps. The means of deriving categories or selecting analogues for read-across has typically been carried out in an *ad-hoc* fashion whereby selection of category members or analogues has been on the basis of common functional group/chemical class. Many chemical categories that have been developed and published, for example, under the High Production Volume programmes in the US or under OECD, have contained small numbers of very closely related analogues often where there is a clear obvious trend, e.g. in carbon chain length. Such categories have likely been manually assembled.

Advancement in computational tools enables a systematic means of developing and evaluating new chemical categories for read-across. Available tools can be helpful in both identifying potential category members and assessing their suitability in a fast and systematic manner – thus they can be helpful in establishing the category rationale as well as its scope (applicability domain).

The specific computational approach used will depend on the availability of information upfront. For example, the tools and techniques that will be used when starting from a large inventory of substances which need to be subdivided into smaller groups (top-down approach) will be different than starting from the premise of addressing data gaps for a single (target) substance (bottom-up approach). The top-down approach is most useful for chemical prioritisation where a pragmatic means of rank ordering chemical substances is required.

For the purposes of this report, only this bottom-up approach will be discussed. The bottom-up approach considers a single substance and sets about identifying structurally related analogues that could be potentially used for read-across purposes. A workflow describing this bottom-up approach together with a illustrative flowchart is described in both the REACH technical guidance on chemical categories and read-across and the OECD guidance on chemical categories.

The most common approach for formulating a chemical grouping for read-across is through a systematic search of potential analogues through using available analogue searching tools. The starting point for this bottom-up approach is either for a (‘seed’) chemical or a small set of (‘seed’) chemicals. There are a number of tools both freely and commercially available to facilitate this type of searching. A description of a few of these tools is provided in Appendix A.

3. DEVELOPING AND TESTING HYPOTHESES BASED ON ANALOGUE SEARCHES

The nature of these hypotheses is, in general, that the new compound is sufficiently similar to chemicals that have already been tested, and that it can be inferred that it behaves similarly in a biological system, or that it will be metabolised to a compound for which the toxicity is known. These analogies can be refined by considering the physical chemistry of the new compound, its potential to react with endogenous molecules, and other characteristics such as pharmacokinetics. These axes of similarity have been conveniently encoded in tools such as the OECD QSAR Toolbox and in publications (e.g. Wu *et al*, 2010). When the hypothesis involves metabolism, the steps for testing it are relatively straightforward. Techniques for evaluating metabolism *in vivo* or *in vitro* have been well worked out, although this may also be an area in which computational approaches can and do play a role, both in identifying previously generated information on the metabolism of similar chemicals, or in predicting metabolic reactions. TIMES (Mekenyan *et al*, 2004) and METEOR (Greene *et al*, 1999) aim to predict likely metabolic pathways (Appendix A). Neither has reached the level of sophistication necessary to predict quantitative aspects of metabolism such as Michaelis-Menten kinetics. There is also a need to predict the hierarchy of possible metabolic pathways (i.e. it doesn't matter that a chemical could theoretically undergo a hundred different reactions if it is rapidly metabolised by the most avid reaction), and to couple metabolic predictions with simple pharmacokinetic considerations to enhance the value of read-across based on metabolism predictions.

The other possibility that the new compound acts in the same way as its analogues can be a little more difficult to address. If mechanism of action is known or can be inferred for the analogues, then it may be possible to evaluate the toxicological similarity of the new compound by a simple *in vitro* or biochemical assay. However, in many cases there is insufficient information about mechanism of action of the analogues. To address this issue, we will be able to rely on increasingly available data and methods in the area of toxicogenomics, which comprehensively evaluates all possible biological responses in a target tissue at the level of gene expression; proteomics, which evaluates most biological responses at the level of protein expression; and metabolomics, which evaluates changes in the profiles of large numbers of small molecules, usually in a clinically accessible fluid like blood plasma. US EPA is considering a mechanistic approach to ecological risk assessment as it introduces the concept of Adverse Outcome Pathways (AOPs) (Ankley *et al*, 2010). In this scheme, risk assessment is based on the mechanistic event(s) that are critical for an adverse effect. Compounds that act via a common mechanism/pathway would be expected to produce comparable effects. Therefore, the focus of hazard identification shifts from outcome to molecular mechanism, a shift that accommodates the read-across approach. The OECD is also adopting this AOP concept as part of the ongoing development of the OECD QSAR Toolbox.

Research in toxicogenomics has been underway for almost ten years. It has become clear that, in virtually all instances, toxicity is accompanied by changes in gene expression, and these changes, when taken as a whole, are indicative of the mode of action of the toxicant. Considerable work has been done to identify gene expression signatures for various modes of action (MoA) (e.g. Naciff *et al*, 2002; Fielden *et al*, 2008). Although this work has not been applied yet to read-across, it is clear that the potential to support read-across exists. One particularly useful illustration of the power of this technology is a study by Liu *et al* (2005) in which they evaluated gene expression in foetal rat testis after maternal treatment with four phthalates that affect male reproductive development, and three phthalates that do not, despite being very close in chemical structure.

Agencies that deal with the regulation of commodity chemicals are starting to appreciate the power of toxicogenomics, and to consider using it in aspects of hazard identification and risk assessment. For example, the US EPA has just finalised a report in which they analysed the utility of toxicogenomics data in refining its risk assessment for dibutylphthalate. While the report concentrated on the utility of microarray data in characterising the MoA of dibutylphthalate, it also acknowledges the potential use of transcript data for other purposes (Makris *et al*, 2010). Predictive toxicology by read-across is the most obvious potential application.

Along with those conclusions are a list of data needs, many of which are consistent with the scientific discoveries and technology development that will need to occur to support read-across. These will include the development of data sets on analogous chemicals to ensure comparability across chemicals that produce similar adverse effects, statistical tools for analysis of microarray data, and informatics tools that facilitate the identification of affected molecular pathways and support the construction of *in silico* models. It will also be important to develop cell-based methods to complement whole-animal approaches and to facilitate interspecies comparisons.

The literature on the effect of chemicals on protein expression is not as well developed as the toxicogenomics literature. The main reason for this is that the methodology for proteomics is more specialised and difficult to standardise as compared to toxicogenomics. The most widely used approaches for large-scale protein identification are 2-dimensional gel electrophoresis and mass spectrometry. Both are powerful tools, but are limited in their ability to be standardised, and neither is capable of identifying all possible protein changes. Still, it is reasonable to assume that a global analysis of protein expression would be as predictive of toxic mechanism as an analysis of mRNA expression.

Metabolite profiling (metabolomics) is a technique which allows for in depth analysis of changes in endogenous metabolites that result from and are indicative of toxicity. In this context, metabolites are defined as small endogenous compounds such as carbohydrates, amino acids,

nucleic acids or fatty acids and their derivates that are present as the result of intermediary metabolism (Lindon *et al*, 2004). The use of sensitive LC-MS and GC-MS techniques to analyse the metabolome in plasma of rats in regulatory studies, such as OECD guideline 407 (OECD, 1995), offers the possibility to obtain significant additional information on toxicological activities in such studies. No additional animals are needed for metabolome analysis, which is a clear contribution to animal welfare in terms of refinement. Data generated from new studies can be compared to a database containing the metabolite profiles of reference compounds of known toxicity. Since the basis for the metabolite profiling as described here is an *in vivo* rat study in which classical endpoints like clinical and histopathology can also be included, the method is not limited to the comparison with the data base itself. From the classical as well as the metabolome parameters, unexpected effects can be observed and compounds can be directly compared to assess their similarity in terms of toxicological / biological effects. A detailed description of the methodology can be found in Van Ravenzwaay *et al* (2007, 2010a).

It may also be possible to use high-throughput assay systems as support for analogue-based hypotheses. The most obvious cases are those instances in which the mechanism of action is well characterised (e.g. the inhibition of a specific enzyme). As batteries of automated high-throughput screening tests, such as ToxCast, become more widely available, it may be possible to use this approach to identify putative mechanisms of action instead of using an omics approach. The amount of data that would need to be generated would probably be the same, but in one instance the data stream would emanate from a single assay that evaluates thousands of endpoints (i.e. microarrays) and in the other from hundreds or thousands of assays that each evaluates only single endpoints.

4. READ-ACROSS BASED ON A HYPOTHESIS OF COMMON BIOLOGICAL ACTIVITY AS A COMPARABLE COMPOUND

There are numerous examples in which a new chemical is a variation on existing chemicals. For some of these, the new chemical can be expected to be metabolised to a chemical with existing safety data. Some trivial examples of this are compounds with a lengthy aliphatic chain. A triglyceride with 18-carbon chains would be expected to have comparable toxicity to one with 16-carbon chains, as they would have similar physical chemistry and both chains are likely to be substrates for β-oxidation. A propyl ester and a butyl ester of benzoic acid would both be hydrolysed rapidly, yielding daughter compounds of known toxicity.

There are many other metabolic reactions that can be predicted based on what is known about xenobiotic and intermediary metabolism. These include both bio-activation and inactivating reactions.

4.1 Read-across based on a hypothesis of similar biological activity

Examples of the use of high-information content data to evaluate the toxicity of analogues.

4.1.1 Toxicogenomics for phthalates

Phthalates are a large family of compounds that are di-esters of phthalic acid. Esters that are between 4 and 10 carbons have been shown to adversely affect male reproductive system development in rats, while lower molecular weight phthalate esters or terephthalates have been shown to have no effect on male reproductive system development, even at extremely high dose levels. These conclusions are based on decades of research on phthalates using traditional toxicology techniques. However, it has recently been shown that the same pattern of activity can be demonstrated with an evaluation of gene expression (Liu *et al*, 2005). In this study, pregnant rats were administered one of several phthalate esters during pregnancy. mRNA was obtained from the testes of male foetuses and evaluated for changes in gene expression versus controls at the same developmental stage. Phthalates that have been shown to affect male reproductive development (e.g. dibutylphthalate and diethylhexylphthalate) produced a characteristic pattern of gene expression, whereas inactive analogues (e.g. diethylphthalate) produced a pattern of gene expression equivalent to controls.

The metabolome is generally reflective of mode of action of toxicity and is believed to be connected to phenotypic effects. Therewith it provides robust data in the context of toxicological evaluation, especially in the case that a database for comparison against the metabolite profiles of reference compounds of known toxicity is available. For example, dibutylphthalate and

diethylhexylphthalate induce a practically identical metabolome in the rat because they induce toxicity with the same MoA (Van Ravenzwaay *et al,* 2010b). On the other hand, compounds like 2-acetylaminofluorene and 4-acetylaminofluorene, despite being structurally very similar at first glance, can be clearly discriminated by metabolomics. A more detailed description of these examples is presented in Appendix B.

4.1.2 High-throughput assays for oestrogens

Due to intense interest in endocrine-active compounds, both in pharmacology and toxicology, a great deal of information has emerged on the chemistry, binding affinity, and biological action (agonist, partial agonist and antagonist) of oestrogen receptors. There are also a number of assay systems for oestrogenic activity, many of which can be run in a high-throughput mode. By high-throughput, we mean assays that can be automated such that hundreds or thousands of chemicals can be evaluated in a period of hours. (For the purpose of evaluating single chemicals versus closely related analogues, there is little need for assay systems to be run in high-throughput mode, however.) An example of the use of oestrogen receptor assays to evaluate the biological activity of closely related analogues is the parabens and *p*-hydroxybenzoic acid. Routledge *et al* (1998) demonstrated weak *in vitro* activity for butyl-, propyl-, ethyl- and methyl-paraben, with diminishing activity as chain length decreased. The root compound, *p*-hydroxybenzoic acid, had no activity.

Another potential use of high-throughput assays comes from ToxCast. The ToxCast battery includes a number of assays that evaluate key molecular controllers of angiogenesis or vasculogenesis. A series of pesticides have been tested in these assays and there appears to be a good correlation (albeit with significant species-specificity) between agents that affect certain aspects of embryonic development and a positive response in several of these assays (Judson *et al*, 2010). It may be possible to use this kind of approach to support read-across, but much more research needs to be done to determine whether the assays have strong mechanistic predictivity.

4.1.3 Examples for reactive chemistry

Skin sensitisation is an endpoint that is dependent on reaction of a hapten with an endogenous peptide. There is a fairly large database on *in vivo* testing for skin sensitisation that supports the development of chemical class categorisation, and even predictions of potency based on reactivity. The case for read-across for this endpoint is developed in Appendix C.

5. CONCLUSIONS AND RECOMMENDATIONS

Traditional toxicity testing methods provide a good basis for assessing hazard and risk of chemicals. However, they are time-consuming and resource-intensive. With the advent of legislation such as REACH that mandates the development of comprehensive toxicity information on tens of thousands of chemicals, it is clear that the traditional methods cannot keep pace with the demand for information. Read-across approaches provide a practical alternative for the development of information on chemicals that are structurally related to already tested compounds. Read-across methods are supported by the large database of toxicology studies that have been conducted over the past 50+ years. These data have been compiled into formats that can be searched by chemical structure/ substructure. Heuristics that identify appropriate structure search strategies, and which can be used to classify the suitability of analogues, have been developed for some aspects of toxicity, most notably toxicity that is dependent on reactive chemistry. Some rules have also been developed for receptor affinity (e.g. the oestrogen receptor), although the development of heuristics for weak interactions such as receptor binding and enzyme inhibition is an area for further research.

In many cases, the outcome of a structure search is sufficiently robust that it can be viewed as definitive. However, in other cases the result can be viewed as a hypothesis that requires further testing. Toxicogenomics, metabolomics, and other high-information-content technologies provide an opportunity to test such hypotheses rapidly and completely. The use of high-information-content methods has the potential to minimise animal use because it may be possible with some technologies (especially toxicogenomics and high-throughput assays) to use *in vitro* models. We have presented a few examples of such an approach here, but this type of research is sparse at present. More is needed to provide convincing evidence that this approach is valid. The generation of such information will serve not only to legitimise the approach, but will also add considerable value to the toxicity databases in which analogue searches are done and which currently are limited mostly to traditional toxicity endpoint data.

5.1 Recommendations

5.1.1 Better analogue identification and hypothesis generation

Although there are a large number of databases and programs available to support read-across, there is a significant need to develop better heuristics for appropriate analogue identification and hypothesis generation. Wu *et al* (2010) provide a good start on codifying the heuristics of analogue identification, especially for reactive chemistry. More research along this line, especially to identify analogues and predict toxicity for chemistry that acts via weak interactions, is needed.

5.1.2 Demonstrations of the use of high-information content data to support read-across hypotheses

There are still very few examples of the use of high-information content data to support read-across. Importantly, there is not enough data to support a consensus on the best and most economical ways to test read-across hypotheses. Existing methods for read-across in chemical risk assessment still need some optimisation, to give real examples to put them in a broader scientific context and last but not least, regulatory acceptance (e.g. REACH). A specific focus for generation of such high-information content data in the context of REACH should be to evaluate the potential of such data to support chemical grouping approaches based on biological response in addition to chemical structure. Beyond this, there is a great need for the development of assay systems, especially *in vitro* assay systems that are robust enough to produce conclusive data to support or reject the hypotheses based on analogue identification.

One promising approach in toxicogenomics is connectivity mapping, first described by Lamb *et al* (2006). This approach is attractive because it uses a small number of cell types that express a large number of small molecule receptors, and does not require the use of large number of microarrays that are needed in traditional toxicogenomics experiments. However, it is not clear whether this approach will be useful for analogue identification.

Concerning metabolomics, databases have been populated with metabolome data from huge numbers of reference compounds of known toxicity (Van Ravenzwaay *et al*, 2010a). Metabolomics is a promising approach since the information can be obtained from minimal invasively collected samples such as blood. The applicability of ‘omics sciences to enhance the quality of chemical grouping has been discussed at an international workshop in 2010 (ECETOC, 2010). It was concluded that these technologies could help to provide a better (biology based) rationale for chemical grouping under the REACH legislation.

These recommendations are formulated as a request for research proposals in Appendix D.

BIBLIOGRAPHY

Ankley GT, Bennett RS, Erickson, RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrano JA, Tiege JE, Villeneuve DL. 2010. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem* 29:730-741.

Aptula AO, Roberts DW. 2006. Mechanistic applicability domains for non-animal-based prediction of toxicological end points: General principles and application to reactive toxicity. C*hem Res Toxicol* 19:1097-1105.

Aptula AO, Patlewicz G, Roberts DW. 2005. Skin sensitization: Reaction mechanistic applicability domains for structure–activity relationships. *Chem Res Toxicol* 18:1420-1426.

Aptula AO, Patlewicz G, Roberts DW, Schultz TW. 2006. Non-enzymatic glutathione reactivity and *in vitro* toxicity: A non-animal approach to skin sensitization. *Toxicol in Vitro* 20:239-247.

Ashikaga T, Sakaguchi H, Nukada Y, Kosak N, Sono S, Nishiyama N, Itagaki H. 2008. Database of h-CLAT (cell-based skin sensitization test) for clarification of applicability domain. Abstracts of the 45th Congress of the European Societies of Toxicology. *Toxicol Lett* 180S:S95.

Basketter DA, Gerberick GF, Kimber I, Loveless SE. 1996. The local lymph node assay: A viable alternative to currently accepted skin sensitization tests. *Food Chem Toxicol* 34:985-997.

Benigni R, Bossa C, Jeliazkova N, Netzeva T, Worth A. 2008. The Benigni / Bossa rule base for mutagenicity and carcinogenicity – A module of Toxtree. European Commission report EUR 23241 EN. European Commission, Joint Research Centre, Institute for Health and Consumer Protection, Ispra, Italy.

Bitsch A, Hadjiolov N, Klöhn PC, Bergmann O, Zwirner-Baier I, Neumann HG. 2000. Dose response of early effects related to tumor promotion of 2-acetylaminofluorene. *Toxicol Sci* 55:44-51.

Bitsch A, Jacobi S, Melber C, Wahnschaffe U, Simetska N, Mangelsdorf I. 2006. REPDOSE: A database on repeated dose toxicity studies of commercial chemicals. A multifunctional tool. *Regul Toxicol Pharmacol* 46:202-210.

Böhme A, Thaens D, Paschke A, Schüürmann G. 2009. Kinetic glutathione chemo-assay to quantify thiol reactivity of organic electrophiles − Application to α,β-unsaturated ketones, acrylates, and propiolates. *Chem Res Toxicol* 22:742-750.

Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, Schlatter J, Seed J, Vickers C. 2008. IPCS Framework for analysing the relevance of a non-cancer mode of action for humans. *Crit Rev Toxicol* 38:87-96.

Buehler EV. 1965. Delayed contact hypersensitivity in the guinea pig. *Arch Dermatol* 91:171-177.

Cramer GM, Ford RA, Hall RL. 1978. Estimation of toxic hazard − A decision tree approach. *Fd Cosmet Toxicol* 16:255-276.

David RM, Moore MR, Cifone MA, Finney DC, Guest D. 1999. Chronic peroxisome proliferation and hepatomegaly associated with the hepatocellular tumorigenesis of di(2-ethylhexyl)phthalate and the effects of recovery. *Toxicol Sci* 50:195-205.

Dinkova-Kostova AT, Holtzclaw WD, Kensler TW. 2005. The role of Keap1 in cellular protective responses. *Chem Res Toxicol* 18:1779-1791.

Dos Santo GG, Reinders J, Ouwehand K, Rustemeyer T, Scheper RJ, Gibbs S. 2009. Progress on the development of human *in vitro* dendritic cell based assays for assessment of the sensitizing potential of a compound. *Toxicol Appl Pharmacol* 236:372-382.

ECETOC. 2010. 'Omics in (eco)toxicology: Case studies and risk assessment, 22-23 February 2010, Málaga, Spain. Workshop report no. 19. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium.

Enoch SJ, Madden JC, Cronin MT. 2008a. Identification of mechanisms of toxic action for skin sensitisation using a SMARTS pattern based approach. *SAR QSAR Environ Res* 19:555-578.

Enoch SJ, Cronin MTD, Schultz TW, Madden JC. 2008b. Quantitative and mechanistic read across for predicting the skin sensitization potential of alkenes acting via Michael addition. *Chem Res Toxicol* 21:513-520.

Fielden MR, Nie A, McMillian M, Elangbam CS, Trela BA, Yang Y, Dunn RT, Dragan Y, Fransson-Stehen R, Bogdanffy M, Adams SP, Foster WR, Chen SJ, Rossi P, Kasper P, Jacobson-Kram D, Tatsuoka KS, Wier PJ, Gollub J, Halbert DN, Roter A, Young JK, Sina JF, Marlowe J, Martus HJ, Aubrecht J, Olaharski AJ, Roome N, Nioi P, Pardo I, Snyder R, Perry R, Lord P, Mattes W, Car BD. 2008. Interlaboratory evaluation of genomic signatures for predicting carcinogenicity in the rat. *Toxicol Sci* 103:28-34.

Flaks B. 1972. Early changes in the fine structure of rat hepatocytes, induced by the non-carcinogen 4-acetylaminofluorene. *Chem Biol Interact* 5:127-137.

Flaks B. 1973. The effect of prolonged dietary administration of the non-carcinogen, 4-acetylaminofluorene on the liver of the rat. An electron microscope study. *Chem Biol Interact* 7:151-64.

Freidig AP, Verhaar HJM, Hermens JLM. 1999. Quantitative structure property relationships for the chemical reactivity of acrylates and methacrylates. *Environ Toxicol Chem* 18:1133-1139.

Gerberick GF, Vassallo JD, Bailey RE, Chaney JG, Morrall SW, Lepoittevin JP. 2004. Development of a peptide reactivity assay for screening contact allergens. *Toxicol Sci* 81:332-343.

Gerberick GF, Ryan CA, Kern PS, Schlatter H, Dearman RJ, Kimber I, Patlewicz GY, Basketter DA. 2005. Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. *Dermatitis* 16:157-202.

Gerberick GF, Vassallo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin JP. 2007. Quantification of chemical peptide reactivity for screening contact allergens: A classification tree model approach. *Toxicol Sci* 97:417-427.

Gerberick GF, Aleksic M, Basketter D, Casati S, Karlberg AT, Kern P, Kimber I, Lepoittevin JP, Natsch A, Ovigne JM, Rovida C, Sakaguchi H, Schultz T. 2008. Chemical reactivity measurement and the predictive identification of skin sensitisers. *ATLA* 36:215-242.

Gerner I, Liebsch M, Spielmann H. 2005. Assessment of the eye irritating properties of chemicals by applying alternatives to the Draize rabbit eye test: The use of QSARs and *in vitro* tests for the classification of eye irritation. *ATLA* 33:215-237.

Gray TJB, Rowland IR, Foster PMD, Gangolli SD. 1982. Species difference in the testicular toxicity of phthalate esters. *Toxicol Lett* 11:141-147.

Gray TJB, Lake BG, Beamand JA, Foster JR, Gangolli SD. 1983. Peroxisomal effects of phthalate esters in primary cultures of rat hepatocytes. *Toxicology* 28:167-180.

Greene N, Judson PN, Langowski JJ, Marchant CA. 1999. Knowledge-based Expert Systems for Toxicity and Metabolism Prediction: DEREK, StAR, and METEOR. *SAR QSAR Environ Res* 10:299-314.

Hooyberghs J, Schoeters E, Lambrechts N, Nelissen I, Witters H, Schoeters G, Van Den Heuvel R. 2008. A cell-based *in vitro* alternative to identify skin sensitizers by gene expression. *Toxicol Appl Pharmacol* 231:103-111.

HSDB. 2010a. Hazardous Substances Data Bank entry for 2-AAF (53-96-3). National Library of Medicine, toxicology data network. National Institutes of Health, Bethesda, Maryland, USA. http://toxnet.nlm.nih.gov/cgi-bin/sis/search.

HSDB. 2010b. Hazardous Substances Data Bank entry for 4-AAF (28322-02-3). National Library of Medicine, toxicology data network. National Institutes of Health, Bethesda, Maryland, USA. http://toxnet.nlm.nih.gov/cgi-bin/sis/search.

Jacobs A. 1997. Understanding organic reaction mechanisms. Cambridge University Press, Cambridge, United Kingdom.

Jansen EHJM, Van den Ham WA, De Fluiter P, Van Leeuwen FXR. 1993. Toxicological investigation of dibutylphthalate in rats. Report no. 618902013. National Institute of Public Health and Environmental Protection (RIVM), Bilthoven, Netherlands.

Jowsey IR, Basketter DA, Westmoreland C, Kimber I. 2006. A future approach to measuring relative skin sensitising potency: A proposal. *J Appl Toxicol* 26:341-350.

Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, Reif DM, Rotroff DM, Shah I, Richard AM, Dix DJ. 2010. *In vitro* screening of environmental chemicals for targeted testing prioritization: The ToxCast project. *Environ Health Perspect* 118:485-492.

Karlberg AT, Bergström MA, Börje A, Luthman K, Nilsson JL. 2008. Allergic contact dermatitis – Formation, structural requirements, and reactivity of skin sensitizers. *Chem Res Toxicol* 21:53-69.

Kern PS, Gerberick GF, Ryan CA, Kimber I, Aptula A, Basketter DA. 2010. Local lymph node data for the evaluation of skin sensitization alternatives: A second compilation. *Dermatitis* 21:8-32.

Kimber I, Griffin AC, Jones K. 1986. The influence of chemical carcinogens on natural killer cell function in rats. A comparison of 2-acetylaminofluorene with 4-acetylaminofluorene. *Cancer Lett* 30:41-48.

Kimber I, Dearman, RJ, Basketter DA, Ryan CA, Gerberick GF. 2002a. The local lymph node assay: Past, present and future. *Contact Dermatitis* 47:315-328.

Kimber I, Basketter DA, Gerberick GF, Dearman RJ. 2002b. Allergic contact dermatitis. *Int Immunopharmacol* 2:201-211.

Koleva YK, Madden JC, Cronin MTD. 2008. Formation of categories from structure-activity relationships to allow read-across for risk assessment: Toxicity of α,β-unsaturated carbonyl compounds. *Chem Res Toxicol* 21:2300-2312.

Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR. 2006. The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313:1929-1935.

Lambrechts N, Vanheel H, Nelissen I, Witters H, Van Den Heuvel R, Van Tendeloo V, Schoeters G, Hooyberghs J. 2010. Assessment of chemical skin sensitizing potency by an *in vitro* assay based on human dendritic cells. *Toxicol Sci* 116:122-129.

Lepoittevin JP. 2006. Metabolism versus chemical transformation or pro-versus pre-haptens? *Contact Dermatitis* 54:73-74.

Lewis DFV, Ioannides C, Parke DV. 1994. Molecular modelling of cytochrome CYP1A1: A putative access channel explains differences in induction potency between the isomers benzo(*a*)pyrene and benzo(*e*)pyrene, and 2- and 4-acetylaminofluorene. *Toxicol Lett* 71:235-243.

Lindon JC, Holmes E, Nicholson JK. 2004. Toxicological applications of magnetic resonance. *Prog Nucl Magn Reson Spectrosc* 45:109-143.

Liu K, Lehmann KP, Sar M, Young SS, Gaido KW. 2005. Gene expression profiling following in utero exposure to phthalate esters reveals new gene targets in the etiology of testicular dysgenesis. *Biol Reprod* 73:180-192.

Makris SL, Euling SY, Gray LE, Benson R, Foster PM. 2010. Use of genomic data in risk assessment case study: I. Evaluation of the dibutylphthalate male reproductive development data set. *Toxicol Appl Pharmacol* Sep 16 2010 [Epub ahead of print].

Maurer T, Arthur A, Bentley P. 1994. Guinea-pig contact sensitization assays. *Toxicology* 93:47-54.

Mekenyan OG, Dimitrov SD, Pavlov TS, Veith GD. 2004. A systematic approach to simulating metabolism in computational toxicology. I. The TIMES heuristic modelling framework. *Curr Pharm Des* 10:1273-1293.

Naciff JM, Jump ML, Torontali SM, Carr GJ, Tiesman JP, Overmann GJ, Daston GP. 2002. Gene expression profile induced by 17α-ethinyl estradiol, bisphenol A, and genistein in the developing female reproductive system of the rat. *Toxicol Sci* 68:184-199.

Natsch A. 2010. The Nrf2-Keap1-ARE toxicity pathway as a cellular sensor for skin sensitizers – Functional relevance and a hypothesis on innate reactions to skin sensitizers. *Toxicol Sci* 113:284-292.

Natsch A, Emter R. 2008. Skin sensitizers induce antioxidant response element dependent genes: Application to the *in vitro* testing of the sensitization potential of chemicals. *Toxicol Sci* 102:110-119.

Natsch A, Emter R, Ellis G. 2009. Filling the concept with data: Integrating data from different *in vitro* and *in silico* assays on skin sensitizers to explore the battery approach for animal-free skin sensitization testing. *Toxicol Sci* 107:106-121.

Natsch A, Gfeller H, Rothaupt M, Ellis G. 2007. Utility and limitations of a peptide reactivity assay to predict fragrance allergens *in vitro*. *Toxicol in Vitro* 21:1220-1226.

Neumann HG. 2007. Aromatic amines in experimental cancer research: tissue-specific effects, an old problem and new solutions. *Crit Rev Toxicol* 37:211-136.

NTP. 1995. Technical Report on toxicity studies of dibutylphthalate (CAS No. 84-74-2) administered in feed to F344/N rats and B6C3F$_1$ mice. Marsman DS. National Toxicology Program, Research Triangle Park, NC, USA. *Toxic Rep Ser* 30.

NTP. 2007a. Testing status 2-acetylaminofluorene (2-AAF) (53-96-3): http://ntp.niehs.nih.gov/index.cfm?objectid=BCB379E3-123F-7908-7BCDB93DD9677E62. National Toxicology Program, Research Triangle Park, NC, USA.

NTP. 2007b. Testing status 4-acetylaminofluorene (4-AAF) (28322-02-3): http://ntp.niehs.nih.gov/index.cfm?objectid=BCDF130A-123F-7908-7B2BA957CBFE66DD. National Toxicology Program, Research Triangle Park, NC, USA.

OECD. 1995. Repeated dose 28-day oral toxicity study in rodents. Guidelines for testing of chemicals No. 407. Organisation for Economic Co-operation and Development, Paris, France.

OECD. 2002. Skin sensitisation: Local lymph node assay. Guidelines for testing of chemicals No. 429. Organisation for Economic Co-operation and Development, Paris, France.

OECD. 2007. Guidance on grouping of chemicals. OECD Environment Health and Safety Publications, Series on testing and assessment No. 80: Report ENV/JM/MONO(2007)28. Organisation for Economic Co-operation and Development, Paris, France.

Oishi S, Hiraga K. 1980. Testicular atrophy induced by phthalic acid esters: Effect on testosterone and zinc concentrations. *Toxicol Appl Pharmacol* 53:35-41.

Python F, Goebel C, Aeby P. 2007. Assessment of the U937 cell line for the detection of contact allergens. *Toxicol Appl Pharmacol* 220:113-124.

Roberts DW, Aptula AO. 2008. Determinants of skin sensitisation potential. *J Appl Toxicol* 28:377-387.

Roberts DW, Natsch A. 2009. High throughput kinetic profiling approach for covalent binding to peptides: Application to skin sensitization potency of Michael acceptor electrophiles. *Chem Res Toxicol* 22:592-603.

Roberts DW, Aptula AO, Patlewicz G. 2007a. Electrophilic chemistry related to skin sensitization. Reaction mechanistic applicability domain classification for a published data set of 106 chemicals tested in the mouse local lymph node assay. *Chem Res Toxicol* 20:44-60.

Roberts DW, Patlewicz G, Kern PS, Gerberick GF, Kimber I, Dearman RJ, Ryan CA, Basketter DA, Aptula AO. 2007b. Mechanistic applicability domain classification of a local lymph node assay dataset for skin sensitization. *Chem Res Toxicol* 20:1019-1030.

Roberts DW, Aptula AO, Patlewicz G, Pease C. 2008. Chemical reactivity indices and mechanism-based read-across for non-animal based assessment of skin sensitisation potential. *J Appl Toxicol* 28:443-454.

Roberts DW, Schultz TW, Wolf EM, Aptula AO. 2010. Experimental reactivity parameters for toxicity modeling: Application to aquatic toxicity of Sn2 electrophiles to *Tetrahymena pyriformis*. *Chem Res Toxicol* 23:228-234.

Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L. 2000. Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *Plant J*. 23:131-142.

Routledge EJ, Parker J, Odum J, Ashby J, Sumpter JP. 1998. Some alkyl hydroxyl benzoate preservatives (parabens) are estrogenic. *Toxicol Appl Pharmacol* 153:12-19.

Sakaguchi H, Ashikaga T, Miyazawa M, Yoshida Y, Ito Y, Yoneyama K, Hirota M, Itagaki H, Toyoda H, Suzuki H. 2006. Development of an *in vitro* skin sensitization test using human cell lines: Human cell line activation test (h-CLAT). *Toxicol in Vitro* 20:774-784.

Sakaguchi H, Ashikaga T, Miyazawa M, Kosaka N, Ito Y, Yoneyama K, Sono S, Itagaki H, Toyoda H, Suzuki H. 2009. The relationship between CD86/CD54 expression and THP-1 cell viability in an *in vitro* skin sensitization test – human cell line activation test (h-CLAT). *Cell Biol Toxicol* 25:109-126.

Sanderson DM, Earnshaw CG. 1991. Computer prediction of possible toxic action from chemical structure; The DEREK System. *Hum Exp Toxicol* 10:261-273.

Schilling K. 1992. Confidential Report from BASF, Department of Toxicology. Study of the oral toxicity of dibutylphthalate in Wistar rats. Administration via the diet over 3 months. Unpublished report, project No. 31S0449/89020. (As quoted by Van Ravenzwaay *et al*, 2010a).

Schultz TW. 2010. Adverse outcome pathways: A way of linking chemical structure to *in vivo* toxicological hazards. In Cronin MTD, Madden JC, eds. *In silico toxicology: Principles and applications*. Royal Chemical Society, London, UK, pp 346-371.

Schultz TW, Yarbrough JW, Johnson EL. 2005. Structure−activity relationships for glutathione reactivity of carbonyl-containing compounds. *SAR QSAR Environ Res* 16:313-322.

Schultz TW, Carlson RE, Cronin MTD, Hermens JLM, Johnson R, O'Brien PJ, Roberts DW, Siraki A, Wallace KD, Veith GD. 2006. A conceptual framework for predicting toxicity of reactive chemicals: Models for soft electrophilicity. *SAR QSAR Environ Res* 17:413-428.

Schultz TW, Yarbrough JW, Hunter RS, Aptula AO. 2007. Verification of the structural alerts for Michael acceptors. *Chem Res Toxicol* 20:1359-1363.

Schultz TW, Rogers K, Aptula AO. 2009. Read-across to rank skin sensitization potential: Subcategories for the Michael acceptor domain. *Contact Dermatitis* 60:21-31.

Schwöbel JAH, Koleva YK, Bajot F, Enoch SJ, Hewitt M, Madden JC, Roberts DW, Schultz TW, Cronin MTD. 2010. Measurement and estimation of electrophilic reactivity for predictive toxicology. *Chem Res Toxicol* (in press).

Srivastava SP, Srivastava S, Saxena DK, Chandra SV, Seth P. 1990. Testicular effects of di-*n*-butylphthalate (DBP); biochemical and histopathological alterations. *Arch Toxicol* 64:148-152.

Van Ravenzwaay B, Coelho-Palermo Cunha G, Leibold E, Looser R, Mellert W, Prokoudine A, Walk T, Wiemer J. 2007. The use of metabolomics for the discovery of new biomarkers of effect. *Toxicol Lett* 172:21-28.

Van Ravenzwaay B, Coelho-Palermo Cunha G, Strauss V, Wiemer J, Leibold E, Kamp H, Walk T, Mellert W, Looser R, Prokoudine A, Fabian E, Krennrich G, Herold M. 2010a. The individual and combined metabolite profiles (metabolomics) of dibutylphthalate and di(2-ethylhexyl)phthalate following a 28-day dietary exposure in rats. *Toxicol Lett* 198:159-170.

Van Ravenzwaay B, Cunha GC, Fabian E, Herold M, Kamp H, Krennrich G, Krotzky A, Leibold E, Looser R, Mellert W, Prokoudine A, Strauss V, Trethewey R, Walk T, Wiemer J. 2010b. Chapter 8: The use of metabolomics in cancer research. In Cho WCS, ed, *An omics perspective on cancer research*. Springer, Dordrecht, Netherlands, pp 141-166.

Vandebriel RJ, Van Loveren H. 2010. Non-animal sensitisation testing: State-of-the-art. *Crit Rev Toxicol* 40:389-404.

Verhaar HJM, Mulder W, Hermens JLM. 1995. QSARs for ecotoxicity. An overview of structure-activity relationships for environmental endpoints. Part 1: General outline and procedure. In Hermens JLM, ed. Report on QSAR for predicting fate and effect of chemicals in the environment. Prepared within the framework of the project "QSAR for prediction of fate and effects of chemicals in the environment", an international project of the Environment; Technologies RTD programme (DGXII/D-1) of the European Commission under contract number EV5V-CT92-0211.

Verhaar HJM, Van Leeuwen CJ, Hermens JLM. 1992. Classifying environmental pollutants. 1. Structure-activity relationships for prediction of aquatic toxicity. *Chemosphere* 25:471-491.

Walker JD, Gerner I, Hulzebos E, Schlegel K. 2005. The skin irritation corrosion rules estimation tool (SICRET). *QSAR Comb Sci* 24:378-384.

Weltzien HU, Corsini E, Gibbs S, Lindstedt M, Borrebaeck C, Budde P, Schulz-Knappe P, Thierse HJ, Martin SF, Roggen EL. 2009. Safe cosmetics without animal testing? Contributions of the EU Project Sens-it-iv. *J Verbr Lebensm* 4, Suppl 2:S41-S48.

Williams GM, Iatropoulos MJ, Jeffrey AM. 2004. Thresholds for the effects of 2-acetylaminofluorene in rat liver. *Toxicol Pathol* 32 Suppl 2:85-91.

Wondrouch D, Böhme A, Thaens D, Ost N, Schüürmann G. 2010. Local electrophilicity predicts the toxicity-relevant reactivity of Michael acceptors. *J Phys Chem Lett* 1:1605-1610.

Wong HL, Liebler DC. 2008. Mitochondrial protein targets of thiol-reactive electrophiles. *Chem Res Toxicol* 21:796-804.

Wu S, Blackburn K, Amburgey J, Jaworska J, Federle T. 2010. A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogues for SAR-based toxicological assessments. *Reg Tox Pharmacol* 56:67-81.

Yamada H, Yamahara A, Yasuda S, Abe M, Oguri K, Fukushima S, Ikeda-Wada S. 2002. Dansyl chloride derivatization of methamphetamine: A method with advantages for screening and analysis of methamphetamine in urine. *J Anal Toxicol* 26:17-22.

Yarbrough JW, Schultz TW. 2007. Abiotic sulfhydryl reactivity: A predictor of aquatic toxicity for carbonyl-containing α,β-unsaturated compounds. *Chem Res Toxicol* 20:558-562.

APPENDIX A: PARTIAL LIST OF DATABASES AND RESEARCH TOOLS TO SUPPORT READ-ACROSS

A.1 Databases with structural search capabilities

A.1.1 AIM

The Analog Identification Methodology (AIM) was developed by the U.S. Environmental Protection Agency (US EPA) to facilitate read-across and chemical grouping by identifying chemical analogues that have existing test data publicly available. AIM is a web-based, computerised tool (cannot be downloaded) that identifies chemical analogues based on structure. The tool also provides the user with pointers or links to publicly available experimental data on the closely related chemical(s) identified.

AIM identifies chemical analogues from a default database that currently contains 31,031 compounds that have some type of toxicity data publicly available. AIM can be searched on the basis of structure, SMILES, CAS number or chemical name. It employs a fragment-based search method to identify analogous compounds using a set of 645 pre-defined fragments and correction factors, and a 'three-pass' searching strategy to locate structures through defined rules and allowable substitution patterns for different types of structural features. For each chemical entered, AIM attempts to identify at least seven analogues using a 'three-pass' method in which analogues with the greatest similarity are identified first. If fewer than seven analogues are found another pass will find chemicals with some allowable differences. Again if fewer than seven analogues are found, another pass will be made to locate analogues with greater differences. If seven analogues are not found with the substitutions allowed in the 3rd pass, AIM will present the analogues found (sometimes only one or two). If data are available on the chemical under study, these data will be presented first and labelled 'Exact Chemical Match', and are followed by those analogues that appear most frequently in the indexed data sources.

AIM was developed to identify analogues for neutral organic compounds and not for metals, inorganic substances, and organic salts. Chemical classes (not an exhaustive list) for which AIM is not expected to produce reliable results include: Chemicals with unknown or variable composition; Mixtures; and High molecular weight compounds.

The tool provides a simple means of identifying analogues that have some kind of toxicity data available, but it does not categorise or rank the analogues returned. This approach leaves it up to the end user to determine when a specific analogue is suitable for a specific assessment, as the determination of what structure is 'appropriate' can vary depending on the endpoint assessed.

The available test data is accessed in the form of hyperlink pointers. The test data is not structured in any way and cannot be downloaded into Excel or other tools for analyses. Some

hyperlinks point to a general webpage, e.g. IUCLID homepage or RTECS homepage; hence the end-user may need the appropriate licences to be able to extract available information. Other links take the user directly to the data source. Thus, the pointer informs that there is a record for the chemical, but does not always indicate the specific type of data available. A summary table of the analogues and which databases are associated with test data can be copied and pasted into Excel.

AIM allows users to rapidly categorise multiple chemicals, focus available resources, facilitate read-across, and streamline assessment exercises.

Further information about AIM is available from the US EPA website at http://www.epa.gov/oppt/sf/tools/aim.htm

A.1.2 ChemIDPlus

ChemIDPlus provides access to structure and nomenclature information for the identification of chemical substances cited in the National Library of Medicine (NLM) databases. The database contains over 370,000 chemical records, of which 200,000 include chemical structures. A ChemIDplus Lite version is also available for NAME and RN searching without the need for plugins or applets.

A user may enter compound identifiers such as Chemical Name, CAS Registry Number, Molecular Formula, Classification Code, Locator Code, and Structure or Substructure. New searchable features include search and display by Toxicity indicators such as Median Lethal Dose (LD50), by Physical/Chemical Properties such as LogP, and by Molecular Weight.

The Advanced search page is available at
http://chem.sis.nlm.nih.gov/chemidplus/chemidheavy.jsp

There are various means of searching for structures using ChemIDPlus to identify analogues. These are as follows:

The Substructure Search will look for the drawn (or transferred) structure embedded in the structures of other compounds. This is useful for finding a set of compounds that share a common ‘Sub Structure’ which may cause good or bad biological activity.

The Similarity Search will look for other compounds with structural features that are similar to those of the drawn (or transferred) structure. The default is to find compounds that are 80% similar. This similarity percentage may be changed to be between 50% and 100%.

The similarity index is based on the Tanimoto index. This compares one compound with another on the basis of fingerprints to arrive at a number between zero and one. An index of 0 signifies that the 2 compounds have nothing in common whereas a value of 1 reveals them to be identical.

The index is defined as follows: C/(A+B-C), where A is number of bits on in molecule A, B is number of bits on in molecule B, and C is number of bits in common between A and B.

The Exact (parent only) will look for the structure that is drawn (or transferred) as a complete entity, with all the structure's atoms and bonds identical in the retrieved compound.

The Flex (parent, salts, mixture) will look for the structure that is drawn (or transferred), with all bonds identical in the retrieved compound including stereo-chemical and tautomeric bonds. Flex will also find salts, mixtures, hydrates, and polymers of the parent that is drawn.

The Flexplus (parent, all variations) will look for the structure that is drawn (or transferred), including compounds that have stereo-chemical and tautomeric variations in their bonds. It will also find salts, mixtures, hydrates, and polymers of the parent that is drawn, plus compounds containing metal atoms bonded to the parent. Flexplus may find compounds with different biological activity from the parent because of differences in stereochemistry and metal bonding.

A typical search will give rise to one or more record which can be viewed one at a time as separate webpages. Results cannot be downloaded or exported into Excel.

A.1.3 Leadscope

Leadscope is a software tool developed and commercialised by Leadscope Inc. (http://www.leadscope.com). It possesses a unique chemical hierarchy containing over 27,000 chemical fingerprints which represent functional groups, chemical groupings and pharmacophores. The software can be purchased with various toxicity databases and/or known drugs databases. The toxicity database contains integrated information on over 160,000 chemical structures from multiple sources including the FDA PAFA Database, the US National Toxicology Program (NTP), RTECS, and the DSSTox Carcinogenicity Potency Database (CPDB). The database covers a range of endpoints including acute and multiple dose studies, such as subchronic liver, carcinogenicity, genetic toxicity, reproductive and irritation. Additional databases have been collated as part of a CRADA agreement with the US FDA (CDER and CFSAN groups within the FDA). These include compiled experimental data from genetic toxicity studies, carcinogenicity studies, chronic/subchronic studies (< 700 substances), acute toxicity data (< 1,000 substances), genetic toxicity and carcinogenicity. Two SAR ready databases for genetic toxicity and carcinogenicity are also available which have been structured

using a specific controlled vocabulary known as ToxML. The controlled vocabulary means that very specific searches can be constructed to extract out very detailed study information. ToxML is an XML database standard based on toxicity controlled vocabulary for use in database standardisation. These databases may be queried by structure (such as substructure or similarity), type of study, toxic effect, species, sex, dosage, duration and route of exposure. Results can be viewed and exported in convenient formats, such as Excel files or as structure data files (sdf).

A.1.4 Danish (Q)SAR Database

The Danish Environmental Protection Agency (EPA) constructed a database of (Q)SAR predictions made by some 70 models for about 166,000 organic chemicals for a wide range of different endpoints. An internet-accessible version of this database is available from the JRC website (http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=DDB). Different types of searching are possible including structure (substructure/exact match) searching, ID (CAS number, name) searching and parameter (endpoint) searching. The (Q)SAR models encompass endpoints for physicochemical properties, fate, eco-toxicity, absorption, metabolism and toxicity. Searches result in a hit list of webpages that can be viewed one at a time. No export functionality exists. The Danish database has since been incorporated into the OECD QSAR Toolbox.

A.1.5 ChemSpider

ChemSpider is a chemistry search engine. It has been built with the intention of aggregating and indexing chemical structures and their associated information into a single searchable repository, and making it freely available to the community. The simple search screen allows searching based on Systematic Name, Synonym, Trade Name, SMILES string or InChI string. Sub-Structure searching is also possible. Searches can be constructed to extract certain property information such as calculated/experimental physico-chemical data, chemical synthesis information, patent information, and supplier information. More information can be found at http://www.chemspider.com.

A.1.6 SciFinder

SciFinder is a commercially available research tool providing access the world's largest collection of biochemical, chemical, chemical engineering, medical, and other related information (http://www.cas.org/SCIFINDER). It provides a means of using a single source to obtain scientific information in journals and patent literature from around the world. It is possible to

explore the database by chemical name, structure, substructure, biological sequence and reaction, as well as by research topic, author and company.

A.1.7 OECD QSAR Toolbox

The OECD Application Toolbox is a software application intended to be used by governments, chemical industry and other stakeholders in filling gaps in (eco)toxicity data needed for assessing the hazards of chemicals. The Toolbox incorporates information and tools from various sources into a logical workflow. Crucial to this workflow is grouping chemicals into chemical categories. The Toolbox is designed to facilitate the development, evaluation and documentation of chemical categories. The main feature of the Toolbox is in its ability to identify relevant structural characteristics and potential mechanisms or modes of action (MoAs) of a target chemical and the search and retrieval of related analogues with those same structural / MoA features. The OECD Toolbox house a range of different databases covering mammalian, e-fate, physico-chemical and ecotoxicity endpoints in addition to a number of regulatory inventories. Workflows are structured to group related (source) chemicals to facilitate data gap filling for a target chemical. Data-gap filling can be performed by qualitative read-across, trend analysis or through the application of external QSARs. Further information is available at http://www.oecd.org/document/54/0,3343,en_2649_34379_42923638_1_1_1_1,00.html

A.1.8 AMBIT

AMBIT is a chemo-informatics data management tool developed by Ideaconsult Ltd (Sofia, Bulgaria) with funding from the CEFIC LRI project: Building blocks for a future (Q)SAR ((Quantitative) Structure Activity Relationship) decision support system.

The AMBIT software is available both as an online and stand-alone application at http://ambit.sourceforge.net/

The AMBIT system comprises a database and functional modules allowing a variety of flexible searches and mining of the data stored in the database. The AMBIT database stores chemical structures, their identifiers such as CAS, Einecs, Inchi numbers; attributes such as molecular descriptors, experimental data together with test descriptions and literature references. The database can also store QSAR models. In addition the software can generate 2-dimensional and 3-dimensional molecular descriptors. Search options include searching by name, CAS number, SMILES, substructures and structure-based similarity, and by descriptor ranges. It can also apply grouping approaches based on mechanistic understanding, such as the Verhaar classification

scheme. The suite of software tools includes a module for QSAR applicability domain assessment, known as Ambit Discovery.

A.1.9 RepDose database on repeated dose toxicity

New regulations for chemicals, biocides and cosmetics require thorough and careful data mining for SAR approaches and further prioritisation for integrated testing. With this in mind, the Fraunhofer ITEM has developed a relational database 'RepDose' on experimental NOEL/LOEL values for various repeated dose toxicity endpoints. While CEFIC-LRI has supported the database development since its inception by ECETOC, RepDose currently contains quality-assured data on subacute to chronic toxicity from nearly 2,300 studies on 655 chemicals. The data sources include well-known review series like WHO-EHC and German MAK, and original study reports (e.g. NTP). So far, the focus has been on oral and inhalation studies in rats or mice; harmonised glossaries describing targets/organs and observed effects were created for RepDose. The content and structure of the database RepDose provide a sound basis to perform analyses on repeated dose toxicity. The philosophy and structure are explained in a paper by Bitsch *et al* (2006).

RepDose can be accessed on the internet using Ambit structure descriptors (http://www.fraunhofer-repdose.de). The RepDose website includes a tool with standard queries that are expected to be useful for the chemical industry in preparing dossiers and classifying compounds under REACH, GHS, and beyond. The database can be used for the evaluation of categories and the development of integrated testing strategies which are useful for the REACH process as well as for the refinement of toxicological concepts such as TTC.

A companion database 'FeDTex' on fertility and developmental endpoints is under construction. FeDTex will enable the evaluation of reproductive toxicity data, in particular NOELs/LOELs on fertility and developmental effects, of presently 100 chemicals.

A.1.10 Toxmatch

Toxmatch is a chemical similarity tool, developed by Ideaconsult under a JRC contract that encodes a range of different similarity measures which can be used to facilitate the development of generic and endpoint-specific categories. The tool includes a functionality to facilitate read-across, as well as to compare chemicals of interest with existing categories. Toxmatch includes several endpoint datasets as examples including those for skin sensitisation, skin irritation, aquatic toxicity and bioaccumulation to assist in the development of endpoint specific chemical categories. Datasets can be imported into Toxmatch to facilitate the identification of related

analogues. See http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXMATCH for more information.

A.2 Databases with physical chemistry, Lipinski rule, or other relevant data

A.2.1 Syracuse Research Corporation databases

Syracuse Research Corporation (SRC) offers free online searches of a number of databases as a public service. For example available free demos and databases are the Environmental Fate Database (EFDB, Bibliographic and experimental values data files on environmental fate and physical/chemical properties) and Physical Properties Database (PHYSPROP, this on-line interactive demo retrieves data from the PHYSPROP database, containing chemical structures, names and physical properties for over 25,000 compounds) (http://www.syrres.com/what-we-do/product.aspx?id=133). The commercial version of these databases is also available for purchase. The PHYSPROP database is also available within the OECD Toolbox and Chemspider described previously.

A.2.2 ACToR

ACToR (http://actor.epa.gov/actor/faces/ACToRHome.jsp) is a collection of databases collated or developed by the US EPA National Center for Computational Toxicology (NCCT). It aggregates data from over 500 public sources on over 500,000 environmental chemicals. It is searchable by chemical name, other identifiers, and by chemical structure. The data includes chemical structure, physico-chemical values, *in vitro* assay data and *in vivo* toxicology data. Chemicals include, but are not limited to, high and medium production volume industrial chemicals, pesticides (active and inert ingredients), and potential ground and drinking water contaminants.

A.3 Heuristics for determining analogue suitability

A.3.1 Toxtree

Toxtree, developed by Ideaconsult Ltd, is a freely available application from the JRC website (http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE) which is able to estimate different types of toxic hazard by applying structural rules. Options include applying the Cramer decision tree, the Verhaar scheme, the BfR rules and SICRET rules and the Benigni-Bossa rules for mutagenicity/carcinogenicity, the START rules for persistence/biodegradation potential and the skin sensitisation alerts.

The Cramer classification scheme (tree) is probably the best known approach for structuring chemicals in order to make estimations of the so-called Threshold of Toxicological Concern (TTC; Cramer *et al*, 1978). The tree relies primarily on chemical structures and estimates of total human intake to establish priorities for testing. The procedure uses recognised pathways for metabolic deactivation and activation, toxicity data and the presence of a substance as a component of traditional foods or as an endogenous metabolite. Substances are classified into one of three classes:

- Class 1 contains substances of simple chemical structure with known metabolic pathways and innocuous end products which suggest a low order of oral toxicity.
- Class 2 contains substances that are intermediate. They possess structures that are less innocuous than those in Class 1 but they do not contain structural features that are suggestive of toxicity like those in Class 3.
- Class 3 contains substances with structures that permit no strong initial impression of safety and may even suggest a significant toxicity.

The Verhaar scheme is a widely used scheme for determining the MoA of chemicals that display aquatic toxicity. It divides chemicals into four groups: Non-polar narcotics, polar narcotics, reactive chemicals and specifically-acting chemicals (Verhaar *et al*, 1992, 1995).

The BfR and SICRET rules predict skin irritation and corrosion on the basis of physicochemical exclusion rules and structural alert inclusion rules (Walker *et al,* 2005).

The Benigni-Bossa rule base provides users with a number of models aimed at predicting the carcinogenicity and mutagenicity of chemicals, based on the knowledge of their structure. The main tool is a list of Structural Alerts (SAs) for carcinogenicity. The present list of SAs refers mainly to the knowledge on the action mechanisms of genotoxic carcinogenicity (thus they apply also to the mutagenic activity in bacteria), but includes also a number of SAs flagging potential non-genotoxic carcinogens. The software also includes QSAR models for: 1) the mutagenic activity of aromatic amines in the *Salmonella typhimurium* TA100 strain (Ames test); 2) the carcinogenic activity of the aromatic amines in rodents (summary activity from rats and mice); 3) the mutagenic activity of α-, β-unsaturated aldehydes in the *Salmonella typhimurium* TA100 strain (Ames test).

The START rule base estimates potential biodegradability or environmental persistence, by using a series of SAs in combination with a decision tree. There are 32 SAs included in START; 23 SAs refer to environmental persistent chemicals mechanisms of action while 9 SAs refer to readily biodegradable chemicals. The SAs were derived from the Guidance Manual for the Categorisation of Organic and Inorganic Substances on Canada's Domestic Substances List, which was assembled from the Existing Substances Branch of Environment Canada.

The skin sensitisation alerts were those developed by Aptula and Roberts (2006) and subsequently encoded as SMARTS patterns by Enoch *et al* (2008a).

A.3.2 TIMES

The Tissue MEtabolism Simulator (TIMES) is a commercially available system that aims to produce plausible biotransformation pathways from a query molecule by using rules developed from a comprehensive library of biotransformations (Mekenyan *et al,* 2004). The system was developed by the Laboratory of Mathematical Chemistry (LMC, Bulgaria). The generation of metabolites by TIMES can be limited to the most likely ones or can be extended to include less likely ones. The developers have also integrated reactivity models for various macromolecular interactions, for example for mutagenicity and sensitisation, to simulate the generation of reactive metabolites by specific metabolising systems, such as S9.

A.3.3 Derek for Windows

The Derek for Windows (DfW) system is a knowledge-based expert system created with knowledge of structure-toxicity relationships and an emphasis on the need to understand mechanisms of action and metabolism. It is marketed and developed by LHASA (Logic and Heuristics Applied to Synthetic Analysis) Limited, a not-for-profit company and educational charity (http://www.lhasalimited.org/index.php).

Within DfW, there are over 504 alerts covering a wide range of toxicological endpoints. An alert consists of a toxicophore (a substructure known or thought to be responsible for the toxicity) and is associated with literature references, comments and examples. The DfW knowledge base covers a broad range of toxicological endpoints, but its main strengths lie in the areas of mutagenicity, carcinogenicity and skin sensitisation. All the rules in DfW are based either on hypotheses relating to mechanisms of action of a chemical class or on observed empirical relationships. Information used in the development of rules includes published data and suggestions from toxicological experts in industry, regulatory bodies and academia. The toxicity predictions are the result of two processes. The program first checks whether any alerts in the knowledge base match toxicophores in the query structure. The reasoning engine then assesses the likelihood of a structure being toxic. There are 9 levels of confidence: certain, probable, plausible, equivocal, doubted, improbably, impossible, open, contradicted. The reasoning model considers the following information: a) the toxicological endpoint; b) the alerts that match toxicophores in the query structure; c) the physicochemical property values calculated for the query structure; and d) the presence of an exact match between the query structure and a supporting example within the knowledge base.

A further application of DfW is its integration with the Meteor system to enable predictions of toxicity for both parent and metabolites. DfW is marketed under the name Derek Nexus.

A.3.4 OncoLogic

The Cancer Expert System or OncoLogic is an expert system that assesses the potential of chemicals to cause cancer. OncoLogic was developed under a cooperative agreement between EPA's Office of Pollution Prevention and Toxics (OPPT) and LogiChem, Inc. It predicts the potential carcinogenicity of chemicals by applying the rules of SAR analysis and incorporating what is known about the mechanisms of action and human epidemiological studies. OncoLogic has the ability to reveal its line of reasoning just as human experts can. After supplying the appropriate information about the structure of the compound, an assessment of the potential carcinogenicity and the scientific line of reasoning used to arrive at the assessment outcome are produced. This information provides a detailed justification of a chemical cancer causing potential. The Cancer Expert System is comprised of four subsystems that evaluate fibres, metals, polymers, and organic chemicals of diverse chemical structures. The OncoLogic Cancer Expert System was previously distributed exclusively by LogiChem, Inc. but is now available for download from the US EPA website (http://www.epa.gov/oppt/sf/pubs/oncologic.htm).

The above databases and tools are summarised in Table A.1.

Table A.1: Databases for approaches to read-across

Database	Import / output	Endpoints	Chemical domain	License type	Predominant utility	Main limitations
Databases with structural search capabilities						
AIM http://www.epa.gov/oppt/sf/tools/aim.htm	Search by CAS, Name, SMILES and structure (drawn by applet)/excel table	Various – all endpoints that are covered in databases such as RTECS, TSCATS, DssTox	Neutral organics (Metal complexes, UVCBS, mixtures, inorganics and ionisable substances are excluded)	Free to use	To identify potential analogues with experimental data	Not possible to direct export the experimental data
ChemIDPlus	Search by: Chemical name, CAS registry number, molecular formula, classification code, locator code, and structure or substructure, limited number of endpoints	Not really covered. Literature for the identified analogues can be found via a TOXNET search over several toxicological databases with the respective CAS or chemical name of analogue	Broad, but works best for organics	Free to use	To identify structurally similar chemicals or those with similar substructures with self-defined % of similarity	No direct access to experimental data, no direct export of results, one by one search needed
Leadscope	Sdf and mol files (exact/substructure/similarity) /excel; sdf; rtf	Mammalian toxicity endpoints (extent of endpoints covered and level of detail that can be exported will depend on which toxicity databases have been licensed)	Predominantly discrete organic chemicals	Commercial, Annual License fee for Leadscope software + databases	To identify potential analogues with experimental data, to apply data mining approaches to elicit new hypothesis for structure alert generation or model development	Limited to mammalian toxicity endpoints; cannot handle more than one result for a given endpoint

Table A.1: Databases for approaches to read-across (cont'd)

Database	Import / output	Endpoints	Chemical domain	License type	Predominant utility	Main limitations
Databases with structural search capabilities						
Danish (Q)SAR Database http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=DDB	Search by CAS, Name and structure (similarity/substructure/exact) (drawn by applet)/html page for viewing	Various – physico-chemical, mammalian, e-fate, ecotoxicity	Predominantly discrete organic chemicals	Free to use	To identify potential analogues in regulatory inventories with estimated data	No experimental data available. No export of data possible
Chemspider http://www.chemspider.com	Search by Name, InChI, SMILES, Structure (Exact/Substructure)	Chemical information (spectral data, Pubchem data), predicted properties (ACD Labs)	Predominantly discrete organic chemicals	Free to use	To identify 2-dimensional structures from chemical/trade name, SMILES	No experimental data available
SciFinder http://www.cas.org/SCIFINDER	Search: Chemical name, CAS number, reaction, structure, substructure	Scientific literature	Unlimited	Commercial annual license fee	Identify chemical analogues with data	No direct access to experimental data, literature search tool with a number of search algorithms
OECD Application Toolbox http://www.oecd.org/document/54/0,3343,en_2649_34379_42923638_1_1_1_1,00.html	Import as Sdf, mol, SMILES, CAS, Name/output as smi, text, specified report formats (pdf, rtf)	Various – physico-chemical, mammalian tox, e-fate, ecotoxicity	Predominantly organic chemicals	Free to use	To develop endpoint specific categories	Limited by experimental data that is contained within the Toolbox; Read-across/Trend analysis is limited to simple linear regression
AMBIT (source and executable to download) http://ambit.sourceforge.net	Sdf and mol files (exact/substructure/similarity) CSV, SMILES	Any endpoint from user imported dataset	ECHA preregistration list and several other databases	Free, Open source (LGPL)	To run online descriptor calculations, building QSAR models and applying QSAR models	No accounting for metabolism

Table A.1: Databases for approaches to read-across (cont'd)

Database	Import / output	Endpoints	Chemical domain	License type	Predominant utility	Main limitations
RepDose http://www.fraunhofer-repdose.de	Chemical name, CAS, functional groups, linked to AMBIT for structural similarity search	Repeated dose toxicity	Organic chemicals with repeated dose data	Free, registration needed	Identify common target organs, relate structures to effects, find range of NOAELs and LOAELs for groups. Details of experimental data available	Limited to repeated dose toxicity. Limited number of predefined queries
Toxmatch http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXMATCH	Text files (txt, csv), smi, Sdf	User defined – starter demo datasets addressing BCF, Aquatic fish toxicity, Skin irritation and sensitisation are provided	Discrete organics	Free to use	To perform many to one/many to many read-across based on in built similarity indices based on experimental data	Limited by access to available datasets although user is free to import other datasets for evaluation
Databases with physical chemistry, Lipinski rule, or other relevant data						
Syracuse Research Corporation databases	Search by substructure, name, fragment, or any physical properties	Contains chemical structures, names, and physical properties for over 41,000 chemicals	Predominantly organic chemicals	Commercial; The online interactive demo retrieves data from the PhysProp database	To identify chemicals with physico-chemical property data	Some of the values are estimated rather than experimental
ACToR http://actor.epa.gov/actor/faces/ACToRHome.jsp	Search by name, CAS, structure	Contains data from 500 public sources on over 500,000 chemicals	HPV, MPV chemicals, pesticides, ground/drinking water contaminants	Free to use	To identify chemicals with available data (physico-chemical, ecotoxicological, mammalian toxicity)	Limited by experimental data that is available. No export/ download possible

Table A.1: Databases for approaches to read-across (cont'd)

Database	Import / output	Endpoints	Chemical domain	License type	Predominant utility	Main limitations
Heuristics for determining analogue suitability						
Toxtree	Text files (txt, csv), smi, Sdf	Rule bases include – BfR rule base (skin/eye irritation), Cramer structural classes (TTC), Verhaar MoA scheme (aquatic toxicity), Benigni-Bossa alerts (muta-/carcino-genicity), Sensitisation alerts	Discrete organics	Free to use	To categorise substances by their Cramer classes, to assign MoA, to identify structure alerts	No accounting for metabolism
TIMES	Import as Sdf, mol, SMILES output as specified report (txt)	Various – skin sensitisation, mutagenicity, carcinogenicity	Discrete organics	Commercial	Estimate toxicity endpoints	Some data used in the training set are proprietary
Derek for Windows (Derek Nexus)	Import as mol, or ISIS draw output as specified report formats (txt, rtf, xls)	Various toxicological endpoints, including: Carcinogenicity Mutagenicity Genotoxicity Skin Sensitisation Teratogenicity Irritancy Respiratory Sensitisation Hepatotoxicity	Discrete organics	Commercial	Estimate toxicity endpoints	No applicability domain is defined. No accounting for metabolism
OncoLogic http://www.epa.gov/oppt/sf/pubs/oncologic.htm	No specific input format – all must be performed on the fly – search/assess on the basis of name, CAS or structure/dat text files	Carcinogenicity	All substances – organics, polymers, organo-metallic	Free to use	To assign a level of concern for carcinogenicity (low, marginal, moderate, high) based on structural alert information and route of exposure	Relies on user expertise in terms of selecting the right chemical functional groups

APPENDIX B: METABOLOMICS AS A USEFUL TOOL FOR READ ACROSS IN CHEMICAL RISK ASSESSMENT

Many chemical categories of structurally similar compounds have been developed and are accepted for read-across (e.g. in OECD HPV). Categories have been manually assembled on a case by case basis on the decisions of experts. Such categories are often based mainly on structural similarity. However, the adverse outcome of exposure to a chemical in reality is caused by a pathway of mechanistic chemical-biological interactions. Interactions of the chemical with a biological system at the molecular level that subsequently cause biological effects at the subcellular, cellular, tissue, organ and whole animal levels of observation is called toxicological mechanism (or mode) of action (MoA). The toxicological MoA is linked to the structure of the molecule and therefore structurally similar molecules often have the same toxicological MoA. However, structural similarity does not guarantee that the toxicological MoA of two compounds is identical. This is demonstrated in the following two case studies by two pairs of practically structurally identical chemicals that trigger the same toxicological MoA in one example (Section B.2.1, peroxisome proliferation) and two different toxicological MoA in the other example (Section B.2.2, liver carcinogenicity).

Metabolome analysis (metabolomics) is able to provide MoA information and can be applied to typical 28 day repeated dose toxicity guideline studies. The metabolome is considered to be all of the measurable changes in small molecule metabolites in a particular tissue or biological fluid, in response to an external perturbation. Toxicants produce characteristic changes in the metabolome, which can be regarded as an instantaneous 'snapshot' of the physiology of the treated organism. Based on defined metabolite patterns the toxicological MoA of a chemical with limited information can be predicted. This technique allows read-across for complex toxicological endpoints. Metabolome analysis combines advantages of *in vivo* and *in silico* strategies by extracting considerably more information from a short *in vivo* study. In contrast to pure *in silico* methods, this procedure is open to unknown effects and new toxicological MoA. In parallel, the method can be used to explain experimental results from a mechanistic point of view and to assist design of intelligent testing strategies.

B.1 The method

Metabolite profiling (metabolomics) is a technique which allows for in depth analysis of changes in endogenous metabolites as a result of exogenous insults like toxicity. In this context, metabolites are defined as small endogenous compounds such as carbohydrates, amino acids, nucleic acids or fatty acids and their derivatives that are involved in intermediary metabolism (Lindon *et al*, 2004). The use of sensitive LC-MS and GC-MS techniques to analyse the metabolome in plasma of rats in regulatory studies, such as OECD guideline 407 (OECD, 1995), offers the possibility to obtain significant additional information on toxicological MoA.

No additional animals are needed for metabolome analysis, a clear contribution to animal welfare in terms of refinement. Moreover, the metabolome is believed to be closely related to histopathological effects observed in the study, thereby providing robust data on biological responses at the biochemical level. This information is especially valuable when a database for comparison against the metabolite profiles of reference compounds of known toxicity is available. Since the basis for the metabolite profiling as described here is an in vivo rat study in which classical endpoints like clinical and histopathology can also be included, the method is not limited to the comparison with the data base itself. From the classical as well as the metabolome parameters, unexpected effects can be observed and compounds can be directly compared to assess their similarity in terms of toxicological / biological effects (Figure B.1). A detailed description of the methodology can be found in Van Ravenzwaay et al, 2007 and 2010a.

Figure B.1: Advantages of metabolomics

- Complex toxicological end points like repeated dose toxicity studies are covered.
- Combining advantages of *in vitro* and *in silico* strategies by extracting the maximum data from a short *in vivo* study.
- The procedure is open to unknown effects and mechanisms as it uses a short *in vivo* study as starting point. There are no limitations like validity area as for pure *in silico* methods.
- Because the method is conducted in parallel it can be used:
    - To explain experimental results from a mechanistic point of view;
    - to assist design of intelligent testing strategies.

Briefly, the studies were performed as follows:

For each dose group, Wistar rats (5/sex) were either fed with a diet containing the individual test substance or a combination thereof. The doses were chosen based on the symptoms and no observed effect levels as described in the literature or in-house studies. The dosed rats were compared to untreated maintenance diet controls. Blood samples were taken from the retro-orbital sinus in all rats under isoflurane anaesthesia (1.0 ml K-EDTA blood on study days 7, 14 and 28) after a fasting period of 16 to 20 hours. The EDTA plasma samples were covered with nitrogen and frozen at -80°C until metabolite profiling was performed. For mass spectrometry-based metabolite profiling analysis, EDTA plasma samples were extracted to yield in a polar and non-polar fraction. Gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-MS/MS (LC-MS/MS) were used for broad profiling; solid phase extraction-LC-MS/MS (SPE-LC-MS/MS) was applied for the determination of catecholamine and steroid hormone levels. For GC-MS analysis, the non-polar fraction was treated with methanol under acidic conditions to yield the fatty acid methyl esters derived from both free fatty acids and

hydrolysed complex lipids. The non-polar and polar fractions were further derivatised with O-methyl-hydroxylamine hydrochloride and pyridine to convert the oxo-groups to O-methyl-oximes and subsequently with a silylating agent before analysis (Roessner *et al*, 2000). HPLC was performed by gradient elution using methanol/water/formic acid on reversed phase separation columns. Mass spectrometric detection technology was applied which allows target and high sensitivity multiple reaction monitoring (MRM) profiling in parallel to a full screen analysis (WO2003073464). For GC-MS and LC-MS/MS profiling, data were normalised to the median of reference samples which were derived from a pool formed from aliquots of all samples to account for inter- and intra-instrumental variation. Steroids hormones, catecholamines and their metabolites were measured by online SPE-LC-MS/MS (Yamada *et al*, 2002).

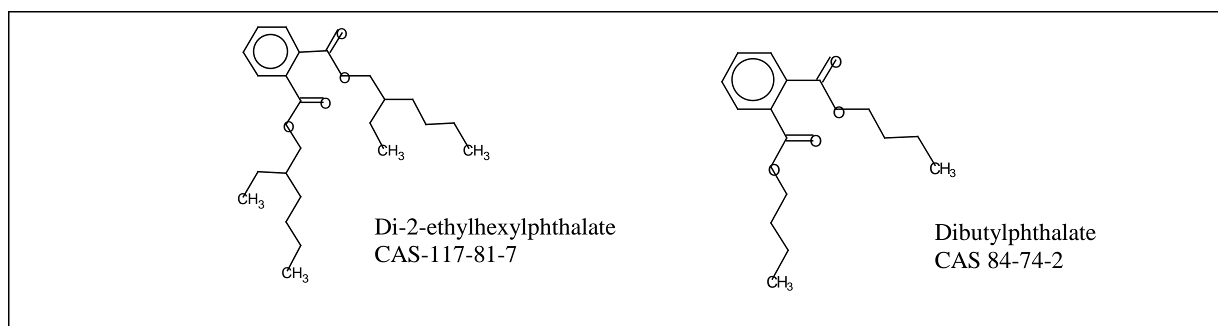
B.2 Practical examples

The following two case studies describe a metabolomic analysis of two pairs of nearly structural identically chemicals. In one example both compounds trigger the same toxicological MoA (Section B.2.1, peroxisome proliferation) and in the other example the compounds trigger two different toxicological MoA (Section B.2.2, liver carcinogenicity). These examples are chosen to demonstrate that structurally similar molecules often, but not always, have the same toxicological MoA.

B.2.1 Practical example DEHP and DBP (same toxicological MoA)

Two structurally related phthalate esters, dibutylphthalate (DBP) and di-2-ethylhexylphthalate (DEHP) (Figure B.2) were administered to rats over 28 days in order to provoke changes in the plasma metabolome. For both compounds, the toxicity induced in the rat is well known. Furthermore, literature and studies performed by BASF clearly demonstrate common toxic effects induced. The comparison of the metabolite changes for the two compounds was used to evaluate whether they had comparable effects on the physiology of treated rats.

Figure B.2: Chemical structure of dibutylphthalate (DBP) and di-2-ethylhexylphthalate (DEHP)



Available in vivo data

The subacute or subchronic repeated administration of DBP and DEHP result in rat liver hepatomegaly with peroxisome proliferation (David *et al,* 1999; Jansen *et al*, 1993; Schilling, 1992; NTP, 1995). In male animals spermatogenesis was negatively affected and atrophy of the seminiferous tubules was observed (Gray *et al*, 1982, 1983; Oishi and Hiraga 1980; Srivastava *et al*, 1990). In reproductive toxicity studies, infertility was seen in the male offspring. Particular effects observed in studies with DBP and DEHP are presented in Table B.1.

Table B.1: effects observed in studies with DBP and DEHP

Compound	Target organs	Effects observed in toxicity studies
DBP	• Liver • reproductive organs • kidney • testis • hematopoietic system • lung • adrenal gland	• ↓ body weight; • ↑ organ weights: liver, lung, kidney, adrenal glands, brain • ↓ testes weights; bilateral testes atrophy; Zn concentration in testes and testosterone ↓; number of epididymal spermatozoa ↓; degeneration of spermatogenic cells; germinal epithelial degeneration with alteration in sperm counts and infertility; hypo-spermatogenesis; antiandrogen • CPY P-450 ↓ (lung) • PAL-CoA activity in hepatocytes ↑; peroxisome proliferation; 11- and 12-hydroxylase ↑; lipid deposition in hepatocytes; hepatomegaly; necrosis and degeneration of liver cells; succinate- and pyruvate dehydrogenase activity of liver mitochondria ↓; • white blood cells ↓; erythrocytes, haemoglobin (males), haematocrit (males) ↓; neutrophils (males), nucleated erythrocytes, blood platelets ↑ (males) • ↑ blood parameters: urea nitrogen, glucose, albumin (males), γ-globulin, AP, bile acids, ALT, AST • ↓ blood parameters: cholesterol, triglycerides, T3, Zn
DEHP	• Liver • reproductive organs • kidney • testis • ovary • spleen • urogenital tract	• ↑ organ weights: liver, kidney • ↓ organ weights: ovary, testes • mononuclear cell infiltration in liver, kidney, heart; mitotic figures in liver ↑; focal loss of glycogen in liver; focal dilation of renal tubules; centrilobular eosinophilia ↑; proliferation of the smooth endoplasmatic reticulum; number of lipid droplets ↑; cyanide-insensitive palmitoyl-CoA oxidation ↑; α-glycerophosphate dehydrogenase ↑; carnitine acetyl transferase ↑; catalase activity ↑; CYP 450 ↑; glucose-6-phosphatase ↓; non-enzymatic reducing agents ↓; peroxisome proliferation; epithelial cells of the proximal tubules ↑; cytoplasmic basophilia ↓; hepatocellular adenoma and carcinoma • bilateral testicular atrophy with interstitial cell hyperplasia; spermatogenesis ↓ • erythrocytes ↑ (males); haematocrit (females), MCV, MCH, MCHC ↓; urine porphyrin ↑ • ↑ blood parameters: non-esterified fatty acids, total protein, albumin, AP • ↓ blood parameters: cholesterol, triglycerides, phospholipids; T4, Zn

Results of Meta Map Tox

At 7,000 ppm, a total of 47 metabolite levels out of 238 were changed in male rats following DBP treatment for 28 days. Most of these were decreased (41) and only 6 were increased. The profile (i.e. number and size of metabolite level changes) was more pronounced at day 14 and 28 compared to day 7 (Table 2a). In female rats dosed with 7,000 ppm DBP only 12 (out of 238) metabolites were consistently changed, 6 of these were increased, 6 decreased (Table 2b).

At 3,000 ppm, a total of 65 metabolite levels out of 238 were changed in male rats following DEHP treatment for 28 days, only 8 being increased and the majority (57) decreased. The profile appeared to be equally strong at all three time points. At 3,000 ppm, a total of 18 metabolites out of 238 were changed in female rats following DEHP treatment for 28 days, 13 being increased and 5 decreased (Tables 3a and 3b).

Table B.2a: Metabolite level changes in male Wistar rats treated with 150, 1,000 and 7,000 ppm DBP for 4 weeks

Metabolite levels were measured at study days 7, 14 and 28. Metabolites are mentioned with at least at 2 of 3 study days increased (red marked) or decreased levels (yellow marked; WELCH t-test, p-value ≤ 0.05) in one of the three mentioned dose groups. Figures represent relative changes of the median metabolite levels compared to controls.

	DBP 7000 ppm			DBP 1000 ppm			DBP 150 ppm		
Metabolite	day 7	day 14	day 28	day 7	day 14	day 28	day 7	day 14	day 28
14-Methylhexadecanoic acid	0.92	0.84	0.81	1.06	0.85	0.94	1.11	0.92	1.03
16-Methylheptadecanoic acid	0.68	0.56	0.49	0.88	0.80	0.96	0.98	1.01	1.07
17-Methyloctadecanoic acid	0.81	0.51	0.54	0.90	0.79	0.89	1.22	0.91	0.98
3-Hydroxybutyrate	2.41	2.17	1.94	1.24	1.18	0.97	0.95	1.02	0.91
3-O-Methylsphingosine [1]	0.81	0.69	0.80	0.88	0.68	0.72	1.00	0.86	0.93
Alanine	0.69	0.73	0.98	0.94	0.88	1.17	1.18	0.88	0.82
alpha-Tocopherol	0.72	0.78	0.82	0.98	0.80	0.86	1.09	0.97	0.96
Arachidonic acid (C20:cis[5,8,11,14]4)	0.75	0.70	0.72	1.02	0.86	0.97	0.96	0.98	1.03
Cholesterol	0.95	0.72	0.82	0.92	0.89	0.94	1.03	1.03	0.99
Coenzyme Q10	0.67	0.51	0.60	0.66	0.69	0.70	0.94	0.88	1.07
dihomo-gamma-Linolenic acid (C20:cis[8,11,14]3)	1.51	1.88	2.11	1.00	1.14	1.33	0.87	1.09	1.05
Docosahexaenoic acid (C22:cis[4,7,10,13,16,19]6)	0.88	0.71	0.56	1.08	0.90	0.83	1.14	1.05	1.01
Dodecanol	0.95	0.72	0.65	1.15	0.77	0.86	0.86	0.91	0.93
Elaidic acid (C18:trans[9]1)	0.93	0.76	0.77	1.01	0.96	0.98	0.91	0.88	0.99
Glycerol phosphate, lipid fraction	0.87	0.81	0.75	1.03	0.86	0.80	0.96	0.83	0.91
Heptadecanoic acid (C17:0)	0.78	0.62	0.76	0.86	0.82	0.85	0.97	0.85	1.13
Isoleucine	0.87	0.81	0.96	1.00	0.95	1.19	0.98	0.97	0.93
Leucine	0.84	0.86	1.02	0.96	0.92	1.22	0.95	0.97	0.88
Linoleic acid (C18:cis[9,12]2)	0.93	0.72	0.69	1.02	0.87	0.90	0.81	0.72	0.79
Linolenic acid (C18:cis[9,12,15]3)	0.93	0.68	0.57	1.19	1.05	1.18	0.71	0.70	0.91
Lysine	1.07	1.15	1.33	0.95	0.88	0.90	0.89	0.92	0.81
Lysophosphatidylcholine (C:17) [2]	0.70	0.67	0.73	0.80	0.88	0.92	0.96	0.89	0.88
Methionine	0.76	0.74	0.83	0.91	0.86	1.02	1.08	0.94	0.91
myo-Inositol, lipid fraction	0.83	0.82	0.82	0.93	0.86	0.93	0.90	0.97	0.84
Myristic acid (C14:0)	0.72	0.71	0.52	0.94	0.86	0.90	0.58	0.86	0.90
Pantothenic acid	1.19	1.65	1.66	1.19	1.15	1.21	1.17	1.08	0.97
Phosphatidylcholine (C16:0, C20:4) [1]	1.00	1.01	1.03	0.97	1.01	1.01	0.97	1.00	1.05
Phosphatidylcholine (C18:0, C22:6) [1]	0.90	0.89	0.79	0.94	0.92	0.88	0.95	0.96	0.94
Phosphatidylcholine No 02 [1]	0.81	0.73	0.71	1.03	1.05	0.96	1.05	0.89	0.87
Proline	0.79	0.83	0.86	0.93	0.95	1.02	0.94	0.93	0.92
Serine	0.75	0.86	0.89	0.90	1.01	1.07	0.86	0.98	0.88
Stearic acid (C18:0)	0.76	0.69	0.67	0.91	0.85	0.87	0.91	0.99	0.96
Threonine	0.75	0.76	0.82	0.92	0.99	0.93	0.97	0.89	0.90
trans-4-Hydroxyproline	0.84	0.88	0.87	1.07	0.91	0.90	0.95	0.99	0.90
Triacylglyceride (C18:1, C18:2) [1]	0.64	0.56	0.49	1.03	0.83	0.73	0.62	0.81	0.95
Unknown lipid (28000470)	0.78	0.63	0.58	0.93	0.85	0.66	0.94	0.85	0.76
Unknown lipid (28000473)	0.61	0.37	0.45	0.88	0.75	0.87	0.81	0.79	0.77
Unknown lipid (68000015)	0.88	0.78	0.66	0.94	0.82	0.75	1.02	1.03	0.91
Unknown lipid (68000017)	0.82	0.71	0.83	0.85	0.78	0.80	0.96	0.92	0.94
Unknown lipid (68000020)	0.76	0.57	0.61	0.89	0.80	0.77	0.93	0.80	0.95
Unknown lipid (68000028)	0.44	0.66	0.28	1.24	0.88	1.16	0.68	0.93	0.97
Unknown lipid (68000038)	0.61	0.67	0.61	0.94	0.90	1.00	1.13	0.96	1.04
Unknown lipid (68000047)	0.91	0.83	0.78	0.95	0.89	0.90	0.99	1.01	1.02
Unknown lipid (68000053)	0.87	0.92	0.89	0.96	0.99	0.92	0.98	1.09	1.03
Unknown lipid (69000059)	0.56	0.51	0.38	1.04	0.92	0.97	0.86	0.74	0.86
Unknown polar (58000136)	1.41	1.40	1.14	1.10	0.97	0.87	0.99	1.04	1.15
Valine	0.88	0.85	0.97	1.01	0.94	1.15	0.99	0.94	0.85

[1] Structure annotation is based on strong analytical evidence (e.g. combinations chromatography, mass spectrometry, chemical reactions, deuterium-labelling, database and literature search and comparison to similar/homologue/isomeric reference compounds)

[2] Metabolite exhibits identical qualitative analytical characteristics (chromatography and mass spectrometry) compared to a metabolite with footnote 1). Further structural and analytical investigations of this metabolite are still pending.

Table B.2b: Metabolite level changes in female Wistar rats treated with 150, 1,000 and 7,000 ppm for 4 weeks

Metabolite levels were measured at study days 7, 14 and 28. Metabolites are mentioned with at least at 2 of 3 study days increased (red marked) or decreased levels (yellow marked; WELCH t-test, p-value ≤ 0.05) in one of the three mentioned dose groups. Figures represent relative changes of the median metabolite levels compared to controls (NA = not analysed).

	DBP 7000 ppm			DBP 1000 ppm			DBP 150 ppm		
Metabolite	day 7	day 14	day 28	day 7	day 14	day 28	day 7	day 14	day 28
3-Hydroxybutyrate	1.87	2.05	2.31	1.24	1.50	1.41	0.93	0.98	0.95
Alanine	0.75	0.82	1.00	0.75	0.79	1.07	0.83	0.84	1.17
Arginine	0.84	0.76	0.92	0.79	0.92	0.93	0.98	0.93	1.09
Cortisol	NA	0.26	3.41	1.42	2.46	0.95	0.94	0.05	2.29
Cysteine	0.87	0.89	1.00	0.85	0.83	0.95	0.96	1.01	0.98
Glutamate	1.17	1.33	0.88	0.98	1.18	1.03	1.15	1.25	1.09
Glycine	1.09	1.04	1.13	1.02	1.01	1.09	1.09	1.05	1.14
Methionine	0.83	0.77	0.97	0.79	0.89	0.89	1.10	1.04	1.13
Pantothenic acid	1.20	1.81	1.66	1.03	1.06	1.40	1.07	1.25	1.38
Phosphatidylcholine (C18:0, C18:2) [1)]	0.97	1.00	0.96	0.98	0.96	1.00	1.00	1.02	1.00
Phosphatidylcholine (C18:2,C20:4) [1)]	1.10	1.06	1.03	1.02	1.06	1.07	1.04	1.04	1.00
Tyrosine	0.80	0.75	0.99	0.96	0.86	1.02	1.12	1.16	1.31

[1)] Structure annotation is based on strong analytical evidence (e.g. combinations chromatography, mass spectrometry, chemical reactions, deuterium-labelling, database and literature search and comparison to similar/homologue/isomeric reference compounds)

Table B.3a: Metabolite level changes in male Wistar rats treated with 3,000 ppm DEHP for 4 weeks

Metabolite levels were measured at study days 7, 14 and 28. Metabolites are mentioned with at least at 2 of 3 study days increased (red marked) or decreased levels (yellow marked; WELCH t-test, p-value ≤ 0.05) in the mentioned dose group. Figures represent relative changes of the median metabolite levels compared to controls.

Metabolite	DEHP 3000 ppm		
	day 7	day 14	day 28
14-Methylhexadecanoic acid	0.75	0.69	0.73
16-Methylheptadecanoic acid	0.48	0.41	0.41
17-Methyloctadecanoic acid	0.44	0.35	0.45
3-O-Methylsphingosine [1]	0.82	0.67	0.97
Arachidonic acid (C20:cis[5,8,11,14]4)	0.67	0.61	0.82
Behenic acid (C22:0)	0.74	0.74	0.82
Cholesterol	0.77	0.78	0.82
Choline plasmalogen (C18, C20:4) [1]	0.87	0.82	0.98
Coenzyme Q10	0.60	0.50	0.55
Creatine	1.29	1.39	1.54
Diacylglyceride (C18:1, C18:2) [2]	0.71	0.52	0.69
Docosahexaenoic acid (C22:cis[4,7,10,13,16,19]6)	0.55	0.48	0.53
Dodecanol	0.77	0.70	0.80
Eicosanoic acid (C20:0)	0.65	0.63	0.70
Elaidic acid (C18:trans[9]1)	0.80	0.63	0.80
Galactose, lipid fraction	0.72	0.77	0.75
Glycerol, lipid fraction	0.50	0.49	0.58
Heptadecanoic acid (C17:0)	0.69	0.61	0.66
Leucine	1.04	1.14	1.25
Linoleic acid (C18:cis[9,12]2)	0.62	0.47	0.58
Linolenic acid (C18:cis[9,1...	0.54	0.44	0.47
Lysophosphatidylcholine (C16:0) [2]	1.05	1.15	1.21
Lysophosphatidylcholine (C17:0) [2]	0.67	0.75	0.66
Lysophosphatidylcholine (C18:0) [1]	0.87	0.95	0.93
myo-Inositol, lipid fraction	0.75	0.71	0.79
myo-Inositol-2-phosphate, lipid fraction	0.67	0.52	0.89
Myristic acid (C14:0)	0.45	0.47	0.56
Nervonic acid (C24:cis[15]1)	0.55	0.64	0.71
Ornithine	1.10	1.18	1.48
Palmitoleic acid (C16:cis[9]1)	0.58	0.45	0.69
Pantothenic acid	1.40	1.80	1.61
Phenylalanine	1.04	1.23	1.14
Phosphate, lipid fraction	0.88	0.73	1.06
Phosphatidylcholine (C16:1, C18:2) [1]	0.60	0.53	0.54
Phosphatidylcholine (C18:0, C18:1) [1]	0.79	0.81	0.86
Phosphatidylcholine (C18:0, C22:6) [1]	0.73	0.70	0.69
Phosphatidylcholine (C18:2, C20:4) [1]	0.87	0.87	0.88
Phosphatidylcholine No 02 [1]	0.69	0.69	0.63
Proline	0.84	0.90	0.90
Serotonin (5-HT)	0.25	0.77	0.12
Stearic acid (C18:0)	0.63	0.60	0.67
Triacylglyceride (C16:0, C16:1) [1]	0.51	0.58	0.71
Triacylglyceride (C16:0, C18:2) [1]	0.81	0.56	0.62
Triacylglyceride (C18:1, C18:2) [1]	0.44	0.41	0.43
Triacylglyceride (C18:2, C18:2) [1]	0.61	0.42	0.49
Triacylglyceride (C18:2, C18:3) [1]	0.32	0.27	0.35
Unknown lipid (28000470)	0.74	0.52	0.55
Unknown lipid (28000473)	0.42	0.36	0.33
Unknown lipid (28000493)	0.59	0.48	0.69
Unknown lipid (68000009)	0.70	0.70	0.86
Unknown lipid (68000015)	0.81	0.79	0.79
Unknown lipid (68000017)	0.66	0.61	0.73
Unknown lipid (68000020)	0.54	0.49	0.64
Unknown lipid (68000028)	0.31	0.35	0.20
Unknown lipid (68000033)	0.82	0.77	0.79
Unknown lipid (68000034)	0.43	0.53	0.49
Unknown lipid (68000038)	0.55	0.52	0.55
Unknown lipid (68000044)	0.50	0.42	0.41
Unknown lipid (68000053)	0.81	0.84	0.88
Unknown lipid (68000056)	0.48	0.26	0.34
Unknown lipid (68000057)	0.81	0.54	0.71
Unknown lipid (68000058)	0.47	0.51	0.48
Unknown lipid (68000059)	0.36	0.29	0.34
Unknown lipid (68000060)	1.19	1.60	1.30
Valine	1.07	1.20	1.31

[1] Structure annotation is based on strong analytical evidence (e.g. combinations chromatography, mass spectrometry, chemical reactions, deuterium-labelling, database and literature search and comparison to similar/homologue/isomeric reference compounds).

[2] Metabolite exhibits identical qualitative analytical characteristics (chromatography and mass spectrometry) compared to a metabolite with footnote 1). Further structural and analytical investigations of this metabolite are still pending.

Table B.3b: Metabolite level changes in female Wistar rats treated with 3,000 ppm DEHP for 4 weeks

Metabolite levels were measured at study days 7, 14 and 28. Metabolites are mentioned with at least at 2 of 3 study days increased (red marked) or decreased levels (yellow marked; WELCH t-test, p-value ≤ 0.05) in the mentioned dose group. Figures represent relative changes of the median metabolite levels compared to controls.

Metabolite	DEHP 3000 ppm		
	day 7	day 14	day 28
3-Hydroxybutyrate	1.69	1.91	1.60
3-O-Methylsphingosine [1)]	1.69	1.74	1.44
5-O-Methylsphingosine [1)]	1.37	1.82	1.35
Arginine	0.73	0.84	0.87
Coenzyme Q10	1.11	1.63	1.32
Coenzyme Q9	1.37	1.91	1.77
Lignoceric acid (C24:0)	1.67	1.87	1.61
Normetanephrine	0.93	0.82	0.69
Palmitic acid (C16:0)	1.20	1.44	1.27
Pantothenic acid	1.50	1.50	1.65
Sphingomyelin (d18:1, C16:0) [1)]	1.13	1.37	1.55
Sphingomyelin No 02 [2)]	1.03	1.27	1.21
Triacylglyceride (C18:1, C18:2) [1)]	0.74	1.42	1.55
Tyrosine	0.86	0.81	0.92
Unknown lipid (68000021)	1.36	1.16	1.41
Unknown lipid (68000033)	0.88	0.81	0.91
Unknown lipid (68000038)	0.79	0.83	0.81
Unknown lipid (68000060)	1.84	1.60	1.68

[1)] Structure annotation is based on strong analytical evidence (e.g. combinations chromatography, mass spectrometry, chemical reactions, deuterium-labelling, database and literature search and comparison to similar/homologue/isomeric reference compounds).

[2)] Metabolite exhibits identical qualitative analytical characteristics (chromatography and mass spectrometry) compared to a metabolite with footnote 1). Further structural and analytical investigations of this metabolite are still pending.

From the above-mentioned tables, it can be seen that DBP and DEHP differ in some metabolites, but share the majority of metabolite changes. The metabolite changes common to both compounds reflect their ability to induce similar biological responses in the metabolome and can therewith be considered as specific for the toxicological activity of these two phthalate esters. In fact, one can use a larger number of reference compounds as a basis to derive specific patterns for toxicity, e.g. peroxisome proliferation (Van Ravenzwaay *et al*, 2010b).

Using statistical correlation analysis on the whole metabolite profile, the similarity of the two compounds is also evident. Both phthalate are clearly grouped together. Comparing DEHP (= rank 1) against > 600 metabolomic profiles in the database ranks DBP on rank 5. However, one has to consider that rank 2 to 4 are other phthalates and mixtures of DEHP and DBP that are in the MetaMapTox-database. The correlation coefficient of the pattern of DEHP and DBP based on Pearson is 0.39 (Table 4). Comparable correlation coefficients are also obtained by comparing DBP and DEHP with other peroxisome proliferators such as Clofibrate, Fenofibrate in a pair-wise comparison against all substances of the MetaMapTox-database. Based on this, it is obvious that the metabolomic pattern of peroxisome proliferators as published by Van Ravenzwaay *et al* (2010b) identifies DEHP and DBP as peroxisome proliferators.

Discussion

Based on studies with numerous reference compounds of known toxicity, specific metabolite changes (patterns) have been detected representing toxicological modes of action using a metabolomics approach developed by BASF (Van Ravenzwaay *et al*, 2010b). One of the patterns for liver toxicity, peroxisome proliferation, matches quite well with the metabolite changes induced by DBP and DEHP.

The phthalate esters used in this study, DBP and DEHP, are known to share toxicological features. Both are peroxisome proliferators showing the subsequent effects on the liver including liver weight gain, hepatocellular damage and increased cyanide insensitive palmitoyl-Co-A oxidation. Structurally, DBP and DEHP share the same chemical core. Moreover, toxicity to the testis is well established accompanied by disturbed spermatogenesis, secondary endocrine effects and male infertility.

Table B.4: Pair-wise comparison of all metabolites DEHP versus DBP

Pair wise comparison is a measure of metabolomic similarity after exposure to different compounds. It can be used to define how similar the toxicological MoA of two different compounds probably is. This test can be done without any further knowledge of the compounds.

Pearson			**Spearman**			**Norm vector product**		
r	**p**	**rank**	**r**	**p**	**rank**	**r**	**p**	**rank**
0.389	0	5	0.339	0	15	0.3	7.99 E-15	18

The difference between the two structures is in the aliphatic chain of the alcohols, with which the phthalic acid is esterified. In case of DBP, the linear *n*-butanol is used, in case of DEHP, 2-Ethylhexanol is the alcoholic moiety. The structural similarity DBP and DEHP is high and can be calculated using different algorithms comparing structural moieties.

This example shows that metabolite profiling or metabolomics is a technique, which can be used two evaluate the biological effects caused by test substances in a minimal invasive manner. The toxic effects induced by DBP and DEHP in the rat studies were comparable. This was reflected by changes in specific metabolites, which have been published as being indicative for peroxisome proliferators. As can be seen from Tables B.2a,b and B.3a,b, not all metabolites changed under treatment with one of the two compounds were in a similar manner affected by the respective other compound, however, DBP and DEHP shared numerous changes. From these changes, a conclusion on the similarity of the two compounds is possible as has been seen with a statistical

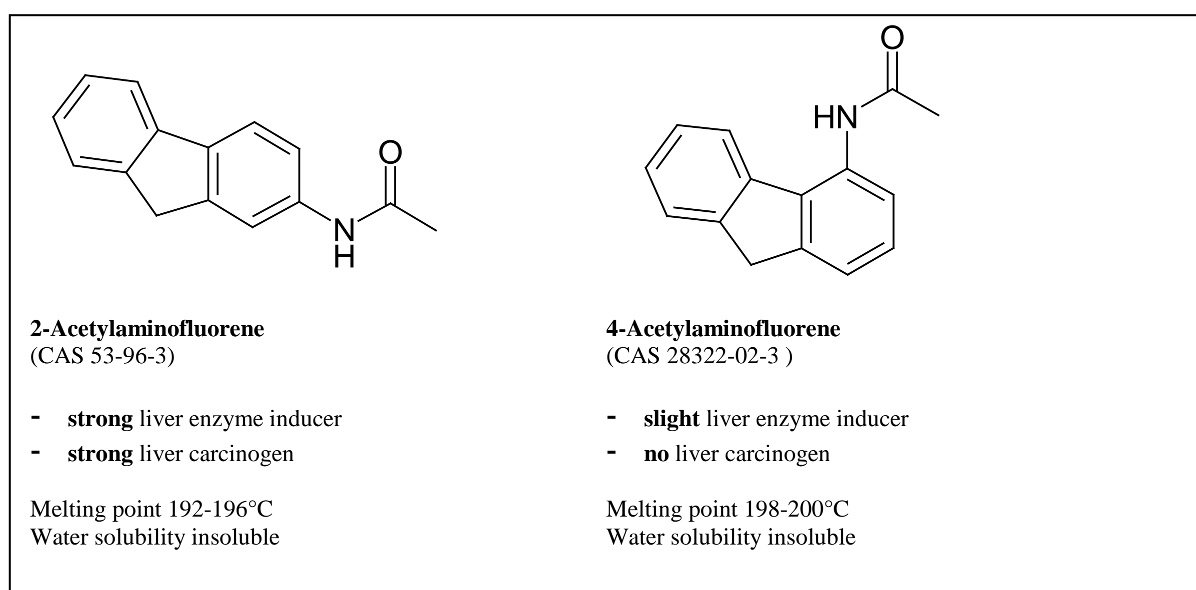
correlation analysis. Moreover, the changes induced by DBP and DEHP were sufficient to identify the target organs as well as the toxicological MoA (as described in Van Ravenzwaay *et al*, 2010b).

Taken together, DBP and DEHP are an example of a successful application of an 'omics technology (in this case metabolomics) to enable read-across (as suggested by REACH) based on structural as well as biological activity (QBAR).

B.2.2 Practical example 2-AAF and 4-AAF (different toxicological MoA)

Two structurally related compounds, 2-acetylaminofluorene (2-AAF) and 4-acetylaminofluorene (4-AAF) (Figure B.3) were administered to rats over 28 days in order to provoke changes in the plasma metabolome. The study design was the same as for the study on DEHP and DBP. The toxicity induced by 2-AAF and 4-AAF in the rat is well known. Furthermore, literature and this study performed by BASF clearly demonstrate that despite similar structure different toxic effects are induced. The comparison of the metabolite changes for the two compounds was used to evaluate the similarity of their effects on the physiology in treated rats.

Figure B.3: Chemical structure of 2-acetylaminofluorene (2-AAF) and 4-acetylaminofluorene (4-AAF)



Available in vivo data

2-AAF is a known mutagen. It is used as positive control in mutagenicity testing such as the Ames test and UDS test. Clastogenic activity has been proven *in vivo*. Oral administration of 2-AAF at doses of 6.25 to 200 mg/kgbw induced chromosome aberrations (35%), SCEs at up to 0.9 per chromosome and micronuclei (2.5%). Several studies have proven carcinogenic activity of this compound (bladder, liver, testis Zymbal gland) (NTP, 2007a; HSDB, 2010a).

2-AAF is a potent inducer of CYP 1A1 and CYP 1A2 acting via the Ah receptor. Liver tumours are seen in less than a year. The mechanism of 2-AAF carcinogenicity is based on N-hydroxylation to the arylhydroxamic acid followed by enzymatic sulphonation or acetylation to sulphoxyfluorenylacetamide, acetylfluorenylacetamide, respectively. Both can be regarded as esters of the hydroxamic acid and are considered to be the ultimate carcinogen (Lewis *et al,* 1994; Bitsch *et al,* 2000; Williams *et al,* 2004; Neumann, 2007).

Beyond this, 2-AAF displays immunosuppressant activities as well. 2-AAF induces a significant depression of native and interferon activated natural killer cells from rats exposed to 2-AAF for 7 or 13 days (Kimber *et al,* 1986).

4-Acetylaminofluorene (4-AAF) was found to be negative for chromosomal aberrations, however positive results were obtained in a sister chromatid exchange assay. 4-AAF was positive in the Ames test and several mouse lymphoma L5178Y TK+/- assays gave positive as well as negative results. Two dominant lethal test as well as an UDS Test were negative. 4-AAF is mitogen, however, in contrast to 2-AAF it is not supposed to be a liver carcinogen as in limited cancer studies (~ 1 year) no liver tumours were observed (NTP, 2007b; HSDB, 2010b).

A study that directly compared the influence of 2-AAF and 4-AAF on natural killer cells revealed that 2-AAF but not 4-AAF induces a significant depression of native and interferon activated natural killer cells from rats exposed for 7 or 13 days (Kimber *et al*, 1986). 4-AAF is not potent inducer of CYP 1A1 and CYP 1A2 (Lewis *et al*, 1994).

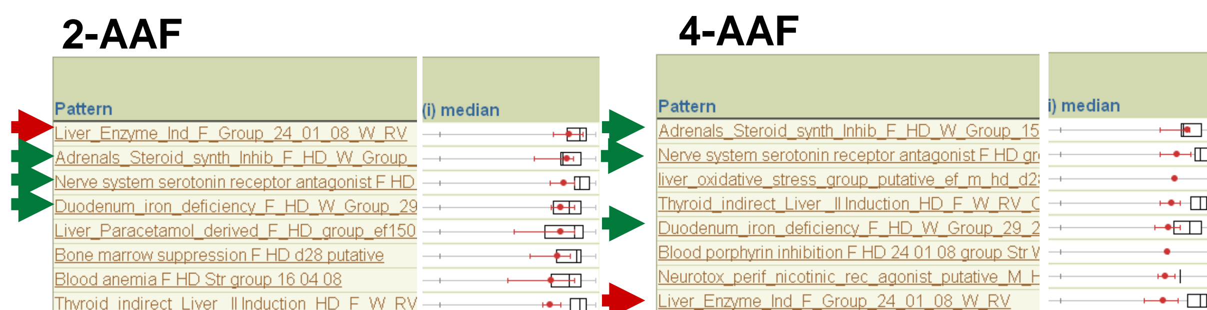
4-AAF caused proliferation of agranular endoplasmic reticulum and glycogen depletion in hepatocytes in an oral feeding study (0.05% of diet) in male rats. No other changes were observed after 3 to 4 weeks of treatment, apart from minor degree of intracellular lipid accumulation. These effects produced by 4-acetylaminofluorene were markedly different from that produced by the carcinogenic isomer. The same results were obtained when 4-AAF was fed to male leads rats at concentration of diet for up to 10 months, producing conspicuous morphological alterations in pancreatic granular endoplasmic reticulum together with mitochondrial damage and focal cytoplasmic degradation (NTP, 2007b; Flaks 1972, 1973).

Results of Meta Map Tox

Metabolome analysis of the high dose female rats is presented in the following. To elicit information on the toxicological MoA a first step in metabolome analysis is to compare the pattern of the compound induced changes of the metabolome with the metabolome pattern of reference compounds from the MetaMapTox data base representing a known MoA of toxicity (pattern ranking). As can be seen in Figure B.4, there are similarities in the pattern ranking of 2-AAF and 4-AAF. One prominent example is the pattern concerning adrenal steroid synthesis inhibition. The pattern of induced changes in the metabolome of both compounds is showing a high similarity to a pattern caused by this toxicological MoA. On the other hand, a pattern representing induction of certain liver enzymes is matched only by 2-AAF. Matching of 4-AAF is not high enough to consider this pattern as the prominent toxicological MoA for 4-AAF (Figure B.4).

Figure B.4: Screenshot of MetaMapTox-database

Comparison of 2-AAF and 4-AAF with metabolomic pattern of known toxicological MoA. A metabolomic pattern is defined by up or down regulation of a defined set of metabolites induced by reference substances of a known toxicological MoA.



That 4-AAF does not fit to the liver enzyme induction pattern can be also seen, when taking a closer look to the metabolomic pattern of liver enzyme inducers (compare 4-AAF and treatment). There is almost no significant induction of metabolites that generally rise after exposure to liver enzyme inducers incl. 2-AAF (Figure B.5).

Figure B.5: Screenshot of MetaMapTox-database

Metabolome pattern of 2-AAF and 4-AAF in comparison to reference substances causing liver enzyme induction (toxicological MoA). Numbers given are factors of up regulation of relevant metabolites defining the toxicological MoA in high dose females on day 7, 14 and 28.

Select	Direction	Anchor	Metabolite	2-Acetylaminofluorene (MOA65)			4-Acetylaminofluorene (MOA64)			Treatment 086 (MOA11)			Treatment 337 (MOA3)			Treatment 427 (MOA11)			Treatment 362 (MOA12)			Treatment 077 (MOA4)		
				fh7	fh14	fh28	fh7	fh14	fh28	fh7	fh14	fh28	fh7	fh14	fh28	fh7	fh14	fh28	fh7	fh14	fh28	fh7	fh14	fh28
	up		Metabolite 105	1.62	1.97	1.47	1.19	1.26	1.16	1.31	1.21	1.26	1.21	1.77	1.42	1.22	1.53	1.38	2.2	3.17	2.03	1.36	1.52	1.43
	up		Metabolite 106	1.29	1.37	1.42	1.21	1.16	1.25	1.23	1.19	1.08	1.41	1.67	1.29	1.17	1.72	1.53	2.32	3.1	2.53	1.73	1.61	1.79
	up		Metabolite 107	1.37	1.45	1.38	1.16	1.24	1.27	1.33	1.35	1.27	1.37	1.72	1.19	1.33	1.94	1.5	2.38	2.75	2.05	1.83	1.81	2.0
	up		Metabolite 110	1.16	1.19	1.15	1.24	1.23	1.25	1.71	1.61	1.56	1.77	1.95	1.34	0.95	1.45	1.32	1.56	1.64	1.58	1.92	1.91	2.03
	up		Metabolite 112	1.18	1.2	1.14	1.22	1.22	1.28	1.76	1.6	1.53	1.91	2.2	1.53	1.05	1.6	1.49	1.92	2.18	2.01	2.08	1.94	2.29
	up		Metabolite 116	1.16	1.36	1.32	1.17	1.56	1.48	1.62	1.41	1.36	2.53	3.04	2.48	1.33	1.98	1.64	1.99	2.96	2.84	2.68	2.22	2.33
	up		Metabolite 143	1.4	1.35	1.28	1.35	1.34	1.07	1.5	1.44	1.16	1.54	1.98	1.28	1.1	1.4	1.22	2.0	1.8	1.64	2.08	1.92	2.51
	up		Metabolite 151	1.07	1.22	1.24	1.12	1.19	1.14	2.26	1.68	2.11	2.36	2.1	1.78	1.57	1.81	1.9	2.34	1.87	2.2	2.77	2.92	3.09
	up		Metabolite 185	1.25	1.17	1.12	1.11	1.11	1.31	1.19	1.25	1.18	1.27	1.46	1.27	1.02	1.33	1.26	1.48	1.53	1.42	1.52	1.79	1.76
	up		Metabolite 356	1.16	1.5	1.68	1.36	1.39	1.42	1.3	1.74	1.49	1.41	1.79	1.17	1.26	2.0	1.85	2.72	4.92	3.75	2.66	2.96	3.16
	up		Metabolite 372	1.56	1.26	1.34	1.34	1.32	1.28	2.22	1.81	2.19	2.31	1.96	1.56	1.51	1.84	1.74	2.38	1.93	2.11	3.77	3.66	3.5
	up		Metabolite 804	1.09	1.1	1.2	1.02	1.01	1.05	1.22	1.36	1.41	1.14	1.25	0.9	1.17	1.35	1.12	1.53	1.57	1.18	1.94	1.74	1.5
	up		Metabolite 816	1.55	1.8	1.36	1.43	1.5	1.38	1.39	1.19	1.53	2.21	2.83	3.11	0.94	1.77	1.68	2.67	4.23	3.49	2.17	1.68	1.85
	up		Metabolite 819	2.43	3.66	1.96	1.31	1.69	2.06	1.61	1.59	1.51	1.63	3.02	1.79	1.85	2.47	1.96	5.08	5.07	4.07	1.98	1.84	1.97
	up		Metabolite 825	1.14	1.12	1.08	1.1	1.13	1.09	1.42	1.38	1.25	1.4	1.32	1.31	1.3	1.35	1.2	1.54	1.54	1.65	1.4	1.14	1.49
	up		Metabolite 844	1.07	1.13	1.23	1.1	1.06	1.03	1.48	1.71	1.48	1.16	1.34	1.39	1.36	1.43	1.28	2.08	1.27	1.23	1.72	1.82	1.3
	up		Metabolite 847	4.0	5.81	2.96	1.79	2.32	2.39	1.09	1.83	0.85	1.02	2.42	1.86	1.64	3.54	2.44	4.4	6.0	3.02	0.84	0.91	1.56
	up		Metabolite 848	2.77	4.38	2.16	1.81	1.84	2.16	1.5	1.58	1.93	1.44	2.79	1.91	1.63	2.37	2.23	5.25	7.19	4.37	1.75	1.65	1.5

Another way to use the metabolome data is to screen for similar substances by pair wise comparison of induced changes to the concentration of all analysed metabolites of a compound to all compounds in that database. All compounds are ranked. Ranking can be done based on different correlation coefficients. A metabolomic pattern of a substance (rank = 1) is compared against > 600 metabolomic profiles in the MetaMapTox-database.

Despite the structure of 2-AAF and 4-AAF a appearing to be very similar 2-AAF and 4-AAF appears to be completely different substances according to the ranking of the metabolomic pattern (Table B.5a). This result is even more prominent when compared to the pair wise comparison of DBP and DEHP (Section B.2.1). DBP and DEHP are ranked together (Table B.5b). One has to consider that substances ranked between DBP and DEHP are also phthalates, mixtures of DBP and DEHP, respectively.

Table B.5a: Pair wise comparison of all metabolites analysed (2-AAF versus 4-AAF)

Pair wise comparison is a measure of metabolomic similarity after exposure to different compounds. It can be used to define how similar the toxicological MoA of two different compounds probably is. This test can be done without any further knowledge of the compounds. Pair wise comparison of all metabolites DEHP versus DBP (Table B.4) is presented again in Table B.5b for a direct comparison of both results.

Pearson			Spearman			Norm vector product		
r	**p**	**rank**	**r**	**p**	**rank**	**r**	**p**	**rank**
0.131	5.33E4	336	0.143	1.51E-4	326	-0.065	0.097	548

Table B.5b: Pair wise comparison of all metabolites DEHP versus DBP

Pearson			Spearman			Norm vector product		
r	**p**	**rank**	**r**	**p**	**rank**	**r**	**p**	**rank**
0.389	0	5	0.339	0	15	0.3	7.99 E-15	18

Discussion

These examples show that metabolite profiling or metabolomics can be used to evaluate the biology in a minimally invasive manner. The toxic effects induced by 2-AAF and 4-AAF in the rat studies are different. This was reflected by changes in specific metabolites.

The adverse outcome of exposure to a chemical in reality is caused by a pathway of mechanistic chemical-biological interactions called the toxicological MoA. The toxicological MoA is linked to the structure of the molecule and therefore structurally similar molecules often have the same toxicological MoA. However, structural similarity does not guarantee that the toxicological MoA of two compounds is identical. This has been demonstrated in this case study. Two structurally near-identical chemicals such as 2-AAF and 4-AAF can trigger two different toxicological MoA. Metabolome analysis is able to clearly discriminate 2-AAF and 4-AAF and to show that read-across is not valid in this case.

B.3 Conclusion on metabolomics as a tool for read-across

Metabolic profiling can give an instantaneous 'snapshot' of the physiology of a living organism. Based on defined metabolite patterns, the toxicological MoA of an unknown chemical can be predicted. Metabolome analysis is able to clearly discriminate substances (no read-across possible = read-across not valid) such as 2-AAF and 4-AAF or group them together as DEHP and DBP (read-across possible = read-across valid).

Taken together, the two case studies are an example for the successful application of an ‘omics technology (in this case metabolomics) to support a decision on validity of a read-across (as suggested by REACH) based on biological activity (QBAR).

B.4 Glossary

Metabolomics, or metabonomics, is the study of a total population of low molecular weight compounds (metabolome).

Metabolome represents the collection of all metabolites in a biological organism.

Metabolites are directly responsible for the phenotype (the observable characteristics) of a biological sample and thus provide a direct link as to what causes various biological states, such as disease states or toxicity.

APPENDIX C: READ-ACROSS IN THE ASSESSMENT OF SKIN SENSITISATION

C.1 Introduction

While reducing and replacing *in vivo* testing for skin sensitisation are long standing goals, at this time there is no single or set of alternative methods which enjoys universal acceptance, especially for replacement. There is a concerted effort to shift hazard assessment for health endpoints, such as skin sensitisation, from one based on animal testing to one based on alternative methods (Weltzein *et al*, 2009; Vanderbriel and Van Loveren, 2010). Among those methods, which show promise in reducing *in vivo* testing is the structure-activity method known as read-across.

The reasons that chemical category-based read-across is gaining acceptance in predicting skin sensitisation include and are in general agreement: 1) that any substance that covalently bonds to proteins has the potential to be a sensitiser, 2) that the adverse outcome pathway leading to sensitisation is well known, 3) on the mechanistic understanding of the molecular initiating event(s), and 4) on the array of organic reaction, which are points of divergence in the cascades of events leading to sensitisation. Armed with the above information, a plausible chemical category can be formed which identifies properties that are common to the members of the category (e.g. skin sensitiser by the Michael addition chemical reaction with thiol groups). When coupled with read-across, this means that it is possible to and reliable to extend measured data to similar untested compounds and help in not only deciding if additional testing is required but also the disposition and extent of any testing.

Qualitative read-across for skin sensitisation (i.e. the presence or absence of sensitisation) based on common 2-dimensional molecular structure between analogues is well accepted, especially for chemical reactions that are well represented within *in vivo* databases. Evidence is presented which shows that quantitative read-across for skin sensitisation is only valid if reading-across within a particular subcategory (e.g. acrylates versus methacrylates) for a specified chemical reaction (e.g. Michael addition). Further evidence is presented suggesting that the use of iso-reactive groups may provide a means of expanding read-across to other subcategories.

Read-across predictions for skin sensitisation are: 1) generally more acceptable if they predict a positive versus negative effect, and 2) strengthened if a mechanistic basis and/or common metabolism can be demonstrated, such as through an adverse outcome pathway. Confidence in read-across predictions for skin sensitisation are: 1) enhanced when experimental data for structural analogues bracket the predicted value for the target chemical (e.g. data for methyl acrylate and butyl acrylate is used to predict ethyl or propyl acrylate), 2) increased as the number of analogues within the chemical category increase (i.e. reading-across from many to one), and 3) improved when supplemented by data from relevant *in vitro* and *in chemico* endpoints (i.e. increased weight of evidence). Confidence in read-across from a single analogue

is improved if it can be demonstrated that the tested analogue is likely to be more toxic than the target chemical (e.g. measured data for ethyl priopiolate is used to predict the effect of ethyl acrylate).

C.1.1 Conclusion

The key to read-across data gap filling is having adequate category data and category rationale to support the prediction. While qualitative read-across for skin sensitisation is currently possible, quantitative read-across, which requires shared potency and the use of subcategories within a given reaction category, is currently not always possible. At the present time, the read-across technique to data gap filling in the assessment of *in vivo* skin sensitisation is sharply limited by the depth and breadth of the available data sets, which were never designed for such an exercise. Currently the best way around this data limitation is to use *in chemico* reactivity as a surrogate for forming the subcategory for quantitative *in vivo* read-across.

C.2 In vivo testing for skin sensitisation

Historically, methods for identifying skin sensitisers have relied on *in vivo* endpoints that make known the onset of allergic contact dermatitis or contact hypersensitivity. Among these *in vivo* methods are the occluded patch test (Buehler, 1965), the Magnusson-Kligman guinea-pig maximisation test (Maurer *et al*, 1994), and the murine local lymph node assay (LLNA) (Basketter *et al*, 1996; Kimber *et al*, 2002a). The LLNA is a well-established and extensively evaluated method. It has undergone validation and has an OECD approved test guideline (OECD, 2002). As the gold-standard for *in vivo* sensitisation, the LLNA assay allows for a quantitative assessment of skin sensitisation potency through the use of the EC_3 value and broader grouping of sensitisation as extreme, strong, moderate, weak, and non-sensitisers. Since the development and validation of the LLNA, it has been primarily used by industry in proprietary product development. The net result is that publicly available databases (see Gerberick *et al*, 2005; Kern *et al*, 2010) are limited in both scope and number. Complicating the issue of predicating skin sensitisation is the observation that, while covalent protein-binding is considered a common property for skin sensitising agents, not all protein binders are skin sensitisers. Realising the importance of chemical reactivity, efforts have been made to divide the LLNA data into reactive domains (Roberts *et al*, 2007a,b). However, even after categorising the LLNA data into separate reaction, it is apparent that sensitisation can be present for some, but not all members of a category. For example: Even within a common reactive mechanism (e.g. Michael addition), compounds may not show a consistent trend when sensitising potency is considered. This has led to proposing that it can be helpful to divide the chemical category into subcategories or more ‘similar’ restrictive analogous chemical sets (Schultz *et al*, 2009).

C.3 Read-across within structural-based categories

Since the 1930s, there has been a growing agreement that the main potency determining step in skin sensitisation is the formation of a stable covalent association with carrier protein or formation of the hapten-protein complex (Roberts and Aptula, 2008). Consequently, the molecular initiating event (covalent perturbation of dermal proteins) at the proper molecular site of action (selected nucleophilic sites in dermal proteins) is predicated on the probability of a particular protein-binding reaction. This reaction is also a point of divergence along the pathway, as different reactions, each with a separate chemical space, can serve as the molecular initiating event leading to skin sensitisation.

In read-across for skin sensitisation, it is extremely beneficial to establish whether the target compound is electrophilic, and secondly, if it is an electrophilic compound, what reaction(s) it is associated with (Roberts *et al*, 2008; Schultz *et al*, 2009). Much of the information relating to electrophilic reactions is obtained from classic organic chemistry and is amenable to computational coding. Such computer-assisted tools are used in the formation of chemical categories and form the basis of the protein binding profiler within the OECD QSAR Toolbox.

In the case of qualitative read-across for skin sensitisation, the presence or absence of an activity for the untested chemical of interest is inferred from the presence or absence of the same activity for the tested analogue(s). Commonality in 2-dimensional molecular structure between members of a category leads to the elucidation of structural alerts. Such structural alerts are then used to identify other category members and define the molecular structural limits of the chemical space of the category. In qualitative read-across, the analogous category is often defined solely on the reactive mechanism. For example, the category of chemicals which act by Michael type nucleophilic addition to thiol is defined by the structural alert for a "polarised α,β-unsaturated group" (Schultz *et al,* 2007).

In the case of quantitative read-across, a known parameter value for tested analogue(s) is used to estimate the value for the untested chemical, with the assumption that the potency of the effect of interest is shared by both the tested and untested analogue. Meeting the assumption of shared potency in skin sensitisation requires the use of subcategories within a given mechanistic category (Enoch *et al*, 2008a,b; Koleva *et al*, 2008; Schultz *et al*, 2009). This is consistent with the goal of moving assessment of skin sensitisation from one of qualification to one of quantification. In the past, skin sensitisers (or non-sensitisers) have been grouped together using expert judgment, which typically reflected a common feature exhibited by the chemicals (e.g. polarised α,β-unsaturated structural fragment).

Table C.1 lists in order of potency 44 Michael acceptors, which have been experimentally evaluated in the LLNA. An initial examination of Table C.1 reveals that LLNA potency for

Michael acceptors vary from extreme sensitisers such as *p*-benzoquinone, to non-sensitisers such as 2-hydroxypropyl methacrylate. Thus, qualitative read-across is of no value in assessing Michael acceptors.

A further examination of Table C.1 reveals several structural groups with more than one member having experimental EC_3 values. These included the 2-alkenal, α-carbon-substituted-2 alkenals, acrylates, methacrylates, 2-alkynoates, and chalcones. The 2-alkenal include compounds 15, 17, 18, 21, and 23; the EC_3 values for these five compounds vary from 2.5 to 5.5. Applying the read-across from the most potent category member gives an estimated EC_3 value of 2.5.

The α-carbon-substituted-2 alkenals include compounds 22, 30, 32, 33 and 44; the EC_3 values for these five compounds vary from 4.5 (strongly moderate sensitiser) to no observed response (non-sensitiser), with compounds 30, 32, and 33 all having an EC_3 value of 11. The LLNA data for this category makes the read-across less clear. Applying the read-across from the most potent category member gives an estimated EC_3 value of 4.5 and the classification of strongly moderate sensitiser. Whereas, applying the read-across from the most common potent gives an estimated EC_3 value of 11 and the classification of weakly moderate sensitiser.

The acrylates include compounds 9, 29, 31, 36, and 37; the EC_3 values for these five compounds vary from 1.4 to 28. This range of values is sharply influenced by the high potency of 2-hydroxyethyl acrylate. Applying the read-across from the most potent category member gives an estimated EC_3 value of 1.4 (strongly moderate sensitiser); applying the read-across from compounds 29 and 31 gives an estimated EC_3 value of 11 and the classification of weakly moderate sensitiser, while applying the read-across from compounds 36 and 37 gives an estimated EC_3 value of 20 and the classification of weak sensitiser. Among the simple alkyl acrylates, the lower molecular weight ones are less potent than the higher weight ones.

The methacrylates include compounds 38, 40, and 43; the EC_3 values for these three compounds vary from 28 (weak sensitiser) to no observed response (non-sensitiser). Applying the read-across from the most potent category member gives an estimated EC_3 value of 28 (weak sensitiser).

The chalcones include compounds 1, 3, 5, 42; the EC_3 values for these four compounds vary from 0.002 (extreme sensitiser) to no observed response (non-sensitiser). This range of values is sharply influenced by the lack of sensitisation of 2',4-Dihydroxychalone. Since this is the only tested chalcone, with substitution on both rings, it is logical to assume this structural feature accounts for the marked difference in observed potency for these compounds. Thus, read-across should only be applied to chalcones that have no substitution on the prime-ring. Applying the read-across from the most potent category member gives an estimated EC_3 value of 0.002 (extreme sensitiser).

Table C.1:. LLNA (EC_3) and GSH reactivity (RC_{50}) data for selected Michael acceptors

ID no.	CAS no.	Chemical name	Reaction sub-domain	EC_3 (%)	RC_{50} (mM)	Kinetics (s^{-1} M^{-1})
1	Not known	4'-Hydroxychalcone	Extreme	0.002		
2	106-51-4	*p*-Benzoquinone	Extreme	0.010	0.046	> 200
3	1482-74-2	2',3',4'-Trihydroxychalcone	Strong	0.11		
4	111-12-6	Methyl 2-octynoate	Strong	0.45	0.28	
5	Not known	2',4'-Dihydroxychalcone	Strong	0.56		
6	1210-39-5	β- Phenyl cinnamaldehyde	Strong	0.60		
7	55-55-0	Metol	Strong	0.80		
8	35691-65-7	Dibromodicyanobutane	Strong	0.90	0.027	> 200
9	818-61-1	2-Hydroxyethyl acrylate	Strongly moderate	1.4	0.27	0.5
10	357650-26-1	5,6,7-Trimethyl-(2Z)- 2,5-octadien-4-one	Strongly moderate	1.6		
11	1337-81-1	Vinyl pyridine	Strongly moderate	1.6	0.33 [a]	
12	29043-97-8	5,5-Dimethyl-3-methylenedihydro-2(3H)-furanone	Strongly moderate	1.8		
13	224031-70-3	1-Spiro[4.5]dec-7-en-7-yl-4-penten-1-one	Strongly moderate	2.2		
14	111-80-8	Methyl 2-nonynoate	Strongly moderate	2.5	0.27	0.2
15	3913-71-1	*trans*-2-Decenal	Strongly moderate	2.5	0.17	
16	56973-85-4	α-Dynascone	Strongly moderate	3.0		
17	104-55-2	Cinnamic aldehyde	Strongly moderate	3.0	0.87 [b]	3.2 [b]
18	142-83-6	*trans,trans*-2,4-Hexadienal	Strongly moderate	3.5	1.5	
19	154750-20-6	3-Bromomethyl-5,5-dimethyl-dihydro-2(3H)-furanone	Strongly moderate	3.6		
20	122-57-6	Benzylidene acetone (4-Phenyl-3-buten-2-one)	Strongly moderate	3.7	3.5	0.01
21	5910-85-0	2,4-Heptadienal	Strongly moderate	4.0	1.2 [c]	
22	101-39-3	α-Methyl cinnamic aldehyde	Strongly moderate	4.5	25	0.003
23	6728-26-3	*trans*-2-Hexenal	Weakly moderate	5.5	0.42	0.3
24	141-05-9	Diethyl maleate	Weakly moderate	5.8	1.80	0.02

Table C.1:. LLNA (EC_3) and GSH reactivity (RC_{50}) data for selected Michael acceptors (cont'd)

ID no.	CAS no.	Chemical name	Reaction sub-domain	$EC_{3\,(\%)}$	$RC_{50\,(mN)}$	Kinetics (s-1 M-1)
25	23726-92-3	β-Damascone	Weakly moderate	6.7		
26	2111-75-3	Perillaldehyde	Weakly moderate	8.1	5.50	
27	104-27-8	1-(*p*-Methoxyphenyl)-1-penten-3-one	Weakly moderate	9.3		
28	Not known	Linalool aldehyde	Weakly moderate	9.5		
29	103-11-7	2-Ethylhexyl acrylate	Weakly moderate	10	0.44	
30	122-40-7	α-Amylcinnamic aldehyde	Weakly moderate	11	NRAS	
31	141-32-2	Butyl acrylate	Weakly moderate	11	0.77	
32	101-86-0	α-Hexyl cinnamic aldehyde	Weakly moderate	11	NRAS	
33	7492-44-6	α-Butyl cinnamic aldehyde	Weakly moderate	11	NRAS	
34	6485-40-1	R-Carvone	Weakly moderate	13	NRAS	
35	103-41-3	Benzyl cinnamate	Weak	18	NRAS	
36	96-33-3	Methyl acrylate	Weak	20	0.48	0.4
37	140-88-5	Ethyl acrylate	Weak	28	0.51	0.7
38	97-90-5	Ethylene glycol dimethacrylate	Weak	28	5.8	0.008
39	1565-94-2	2,2-*bis*-[4-(2-Hydroxy-3-methacryloxypropoxy)-phenyl)]-propane (*bis*-GMA)	Weak	45		
40	80-62-6	Methyl methacrylate	Weak	90	35	0.003
41	75-35-4	Vinylidene dichloride	Ns	NS		
42	13323-66-5	2',4-Dihydroxychalcone	Ns	NS		
43	923-26-2	2-Hydroxypropyl methacrylate	Weak	NS	25	0.007
44	497-03-0	*trans*-2-Methyl-2-butenal	Weak	NS	12	

NS, non-sensitiser; NRAS, not reactive at saturation
[a] for 4-Vinyl pyridine 100-43-6
[b] for *trans*-Cinnaminic aldehyde 14371-10-9
[c] for *trans,trans*-Heptadienal 4313-03-5

With noted disparities in experimental values, these 5 dataset are robust enough to allow for read-across to untested members of their respective categories; however, the data is not in all cases consistent enough to do the read-across with confidence (e.g. the acrylates).

The 2-alkynoates (compounds 4 and 14) are an example of the dilemma of having only two tested analogues with markedly different values. Applying the read-across from the most potent category member gives an estimated EC_3 value of 0.45 (strong sensitiser).

C.3.1 Conclusion

Currently, the direct use of the read-across technique to fill data gaps for the assessment of *in vivo* skin sensitisation is sharply limited by the depth and breadth of the available data sets including the LLNA, which were never designed for such an exercise.

C.4 Read-across within iso-reactive categories

Measurements and estimations of electro(nucleo)philic reactivity have recently been reviewed (Schwöbel *et al*, 2010). Computational or *in silico* techniques to predict chemical reactivity have been developed; however, they vary in complexity from the relatively simple approach of forming chemical categories from 2-dimensional structural alerts (i.e. SARs for qualitative identification of chemical sub-structures with the potential of being reactive) to QSAR models (i.e. quantitative prediction of relative reactivity).

Reactivity with biological molecules includes a spectrum of conjugation and substitution where an electron-rich compound interacts with electron-deficient ones (Jacobs, 1997). In contrast to receptor-mediated chemical interactions (e.g. oestrogen-receptor binding), electrophilic compounds are not specific in regards to their molecular targets. Wong and Liebler (2008), in their examination of mitochondrial proteins from cells treated with two different electrophilic compounds, observed that adducts were formed with more than 800 proteins. Moreover, some chemicals, such as acrolein, are able to react with several different nucleophilic substituents. Therefore, the identification of the specific target protein is not critical to predicting skin sensitisation. It should be recognised that reactivity measured alone does not necessarily reflect the specific chemical reaction, as many reactions target the same chemical substituent.

Nucleophilic sites related to skin sensitisation are in order of increasing hardness; the thiol group of cysteine and glutathione, sulphur atoms of methionine, primary amino groups of lysine and arginine, and secondary amino groups of histidine. Soft electrophilic interactions involving the thiol group can be modelled with small molecules as well as cysteine,

acetylcysteine, or peptides with a cysteine residue (Natsch, *et al*, 2007). Glutathione (GSH, *L*-γ-glutamyl-*L*-cysteinyl-glycine) is one of the most widely used model nucleophiles in reactivity assays. It is the most prevalent cellular thiol and the most abundant low molecular weight peptide in eukaryotic cells (Aptula *et al*, 2006).

Reaction rates can only be compared in the context of the same reaction domain, for example: Michael addition reactions measured by glutathione depletion. This also holds for other mechanisms, such as nucleophilic substitution, aromatic nucleophilic substitution, Schiff base formation or acylation (Aptula *et al*, 2005; Aptula and Roberts, 2006).

While a variety of *in chemico* protocols have been developed (see Schwöbel *et al*, 2010), many are based on GSH as the model nucleophile, which is usually dissolved in an aqueous phosphate buffer solution. Typically after a defined reaction time, the concentration of free thiol groups is measured. The fraction of free thiol is usually quantified by UV/VIS spectroscopy at 412 nm absorption after reaction with the chromophore 5,5'- dithio-*bis*-(2-nitrobenzoic acid) (DTNB). Good qualitative relationships between GSH reactivity of Michael type acceptors have been demonstrated. For example: Although structurally very similar, the different toxicity of methacrylate, crotonate, and acrylate can be explained by their different in GSH reactivity (Wondrouch *et al*, 2010). Methacrylates are very slowly reactive, while crotonates are moderately reactive, and acrylates are highly reactive (Yarbrough and Schultz, 2007).

The importance of reaction chemistry for sensitisation indicates that identification of the reaction-limited chemical spaces is critical to select the correct chemicals for read-across. Systematic databases for reaction-specific chemical spaces are being developed (Roberts and Natsch, 2009). *In chemico,* databases reporting measurements of reactive potency currently exist for Michael acceptors (Yarbrough and Schultz, 2007; Böhme *et al*, 2009; Roberts and Natsch, 2009) and halo-sp3C S_N2 substitution (Roberts *et al*, 2010). GSH reactivity databases for other reactions, including disulphide exchange, and reactions with epoxides and nitroso-containing compounds, are currently under development.

While the use of model nucleophiles containing primary amino ($–NH_2$) or secondary amino groups ($–NHCH_3$), such as in the amino acids lysine and histidine, respectively, are less well documented with the principle of measuring relative reactivity being the same as for thiol. For example, *n*-butylamine is used as a surrogate for primary amines such as lysine. Currently, amine-based reactivity databases are scattered and fragmentary.

Gerberick *et al* (2004) have put forth a battery of *in chemico* assays. They utilised four different depletion assays (GSH, cysteine, and lysine, or histidine containing synthetic heptapeptides (Gerberick *et al,* 2004, 2007). Since both soft and hard interactions are captured by this approach, it is a more robust assessment of reactivity which then could be related to toxicological

pathways, including the one proposed for skin sensitisation. While this battery is a good compromise between accuracy, time and effort, it should be noted that the assays does not consider the underlying reaction domain of the test compounds.

Michael addition includes a number of subcategories where chemical reactivity is consistent within the subcategory, but varies between subcategories (Schultz *et al*, 2005, 2007; Roberts and Natsch, 2009). An example of the difference in reactive potency between subcategories is seen by comparing the RC_{50} values in Table C.1. Such examples point to the possibility of basing read-across for LLNA on iso-reactive groups which incorporate *in chemico* reactive potency similarities into forming categories for *in vivo* read-across. This has the added value of using the existing *in vivo* data to predict skin sensitisation for a greater number of chemicals, while at the same time assuring mechanistic plausibility. Using *in chemico* reactivity data also may provide a greater likelihood that the estimated sensitisation value will be accurate.

One way that reactivity data can be used to expand read-across is by providing supplemental data on data poor subcategories. Only one quinone, *p*-benzoquinone, has been tested in the LLNA (Table C.1). However, ten quinones have been evaluated for *in chemico* GSH reactivity (Table C.2). These data support the read-across from *p*-benzoquinones to other quinones having at least one hydrogen-substituted C-atom.

Table C.2: GSH reactivity (RC_{50}) for selected quinones

Name	CAS number	RC_{50} (mM)
1,4-Benzoquinone	106-51-4	0.047
Methyl-1,4-benzoquinone	553-97-9	0.046
tert-Butyl-1,4-benzoquinone	3602-55-9	0.050
Chloro-1,4-benzoquinone	695-99-8	0.025
2,6-Dimethyl-1,4-benzoquinone	527-61-7	0.085
2,6-Dimethoxy-1,4-benzoquinone	530-55-2	0.021
1,2-Naphthoquinone	524-42-5	0.023
1,4-Naphthoquinone	130-15-4	0.028
2-Methyl-1,4-naphthoquinone	58-27-5	0.057
2,3-Dibromo-1,4-naphthoquinone	13243-65-7	0.025

Another way that reactivity data can be used to expand read-across is to provide supplemental data for subcategories where the *in vivo* data varies. As noted above, LLNA for acrylates exhibits significant potency differences. Eighteen acrylates have been evaluated for *in chemico* GSH reactivity (Table C.3).

Table C.3: GSH reactivity (RC_{50}) for selected acrylates

ID number	Name	CAS number	RC_{50} (mM)
1	Methyl acrylate	96-33-3	0.48
2	Ethyl acrylate	140-88-5	0.51
3	*n*-Propyl acrylate	925-60-0	0.85
4	*n*-Butyl acrylate	141-32-2	0.79
5	Isobutyl acrylate	106-63-8	0.48
6	*tert*-Butyl acrylate	1663-39-4	1.2
7	*n*-Pentylacrylate	2998-23-4	0.80
8	Isoamyl acrylate	4245-35-6	0.70
9	*n*-Hexylacrylate	2499-95-8	0.82
10	2-Ethylhexyl acrylate	103-11-7	0.44
11	Allyl acrylate	999-55-3	0.40
12	Propargyl acrylate	10477-47-1	0.24
13	2-Hydroxyethyl acrylate	818-61-1	0.27
14	Benzyl acrylate	2495-35-4	0.25
15	*cis*-Ethyl-3-iodoacrylate	31930-36-6	0.70
16	*cis*-Ethyl-(β-cyano)acrylate	40594-97-6	0.10
17	Phenyl acrylate	937-41-7	0.02
18	Ethyl 2-(bromomethyl)acrylate	17435-72-2	0.01

Within the alkyl acrylates (compounds 1 – 10), the smaller methyl and ethyl acrylates are more reactive that the large analogues. Moreover, the highly branched *tert*-butyl derivative (compound 6) is less reactive than its less branched analogues (compounds 4 and 5), while the 2-position substituted derivative (compounds 5 and 10) are more reactive than straight chained analogues. Moreover, acrylates substituted with other substituents (e.g. hydroxyl, cyano, and aromatic moieties) are more reactive than simple alkyl derivatives. These observations support the *in vivo* data, where 2-hydroxyethyl acrylate is more potent than the alkyl acrylates. Furthermore, it leads support to the hypothesis that the EC_3 values for the low molecular weight acrylates may be a result of abiotic loss due to evaporation from the mouse ear during the exposure phase of the LLNA.

A third way in which reactivity data can be used to expand read-across is to provide data for forming iso-reactive subcategories where *in vivo* data is sparse. Expanding on *p*-benzoquinones and the quinones in Table C.2, we see that the RC_{50} value in all cases is less than 0.1 mM. Employing this value as a cut off, several other groups of Michael acceptors can be estimated via read-across to be extreme sensitisers. These include nitrovinyl compounds, vinyl sulphones and sulphonate, propiolates as well as maleimides, acetylenedicarboxylates, and other Michael acceptors with two polar activating groups (Table C.4).

Such use of read-across in skin sensitisation has less contention for extreme sensitisers than for moderate or weak sensitisers. At the very least, such a use of read-across would prioritise chemicals for further testing.

Table C.4: GSH reactivity (RC_{50}) for other extreme sensitisers

Group	Name	CAS number	RC_{50} (mM)
1	1-Nitro-1-cyclohexene	2562-37-0	0.028
1	*trans*-*b*-Nitrostyrene	5153-67-3	0.079
1	*trans*-4-Methyl-*b*-nitrostyrene	5153-68-4	0.075
1	*trans*-3-Methoxy-*b*-nitrostyrene	3179-09-7	0.045
1	4-Fluoro-b-nitrostyrene	706-0801	0.054
1	*trans*-4-Chloro-*b*-nitrostyrene	706-07-0	0.066
1	*trans*-4-Bromo-β-nitrostyrene	5153-71-9	0.067
2	Divinyl sulfone	77-77-0	0.050
2	Phenyl vinyl sulfonate	1562-34-1	0.058
3	Ethyl propiolate	623-47-2	0.091
3	Methyl propiolate	922-67-8	0.097
4	N-Methylmaleimide	930-88-1	0.059
4	N-Ethylmaleimide	128-53-0	0.054
4	N-Propylmaleimade	21746-40-7	0.051
5	Dimethyl acetylenedicarboxylate	762-42-5	0.030
5	Diethyl acetylenedicarboxylate	762-21-0	0.025
5	Diethyl ethylidenemalonate	1462-12-0	0.067
5	4-Cyclopentene-1,3-dione	930-60-9	0.075
5	Fumaronitrile	764-42-1	0.073

C.4.1 Conclusion

Since the main potency determining step in skin sensitisation is the formation of a stable covalent association with carrier protein or formation of the hapten-protein complex, chemical reactivity is considered essential for establishing the rational and data selection in read-across for assessment of skin sensitisation. Since both reactions specificity and rate-related reactive potency are molecular structure-dependent, both *in silico* and *in chemico* methods can be used to establish plausibility for the read-across as well as expand its use.

C.5 Metabolism

Another ancillary event in identifying protein-binding is metabolism and/or abiotic transformation. *In vivo*, the keratinocyte is the primary site of metabolism. While *in silico*, methods for identifying reactive metabolites exist (e.g. Times-SS); their current predictivity varies depending on the reaction being simulated. Oxidation of simple compounds, such as polyphenols and unsaturated alcohols, is well documented and well modelled *in silico*. However, metabolism of more structurally complex chemicals, especially ones with the potential for multiple reactions, are less well understood and predictivity is hampered by not knowing which metabolites are the more likely ones to be formed.

C.5.1 Conclusion

It is clear that metabolism and its simulation can have a significant impact on the selection of read-across analogues. If metabolism is suspected, selection of analogues for read-across should be done in the most conservative manner possible.

C.6 The adverse outcome pathway for skin sensitisation

An adverse outcome pathway strategy based on protein binding reactions can be developed to establish and test mechanistic plausibility for the *in vivo* hazard of skin sensitisation. Skin sensitisation is Type IV contact allergy and as such can be conceptually difficult due to it biological complexity (i.e. the requirement of multiple cellular and organ interactions) and large number of individual chemical reactions that can initiate the biological response set. There is, however, general agreement that any substance that covalently bonds to proteins has the potential to be a sensitiser (Gerberick *et al*, 2008). Yet, in order to determine if a chemical will be an *in vivo* sensitiser, other issues such as skin penetration (i.e. bioavailability), epidermal dendritic cell activation, and T-cell proliferation, may have to be considered (Jowsey *et al*, 2006). While current knowledge of dermal immunogenicity is sufficiently developed so inclusion of other factors can be proposed with pathway justification, often our ability to quantify these aspects of immunotoxicity is lacking.

It is well documented that a chemical must undergo a number of interactions before it induces skin sensitisation (Kimber *et al*, 2002b). These interactions are characteristically divided into two phases. The first is the induction phase, which entails the chemical or allergen penetrating the outer epidermis of the skin and forming a complex with a dermal protein (i.e. a hapten-protein complex). This complex is then processed by the epidermal dendritic cells (i.e. Langerhans cells), which subsequently mature and migrate out of the skin to the local lymph nodes. In the

lymph nodes, the complex is presented and recognised by naive T-cells leading to cell proliferation and the production of allergen chemical-specific memory T-cells, some of which re-circulate throughout the body.

The second phase is the elicitation or challenge phase. This phase occurs following a subsequent contact with the allergen. Again, the hapten-protein complex is formed and subsequently taken up by epidermal dendritic cells, as well as other cell types. The circulating allergen-specific, activated memory T-cells are triggered to secrete specific cytokines which induce the release of inflammatory cytokines and mobilisation of other T-cells, as well as other inflammatory cells from the circulating blood. These cells migrate to the epidermis of the skin and induce the distinguishing local inflammatory response of red rash, blisters and welts, and itchy and burning skin.

Since both the fundamental chemical basis of protein binding and the chronology of the biological events leading to skin sensitisation are understood, there is general agreement on the generic events along the adverse outcome pathway leading to skin sensitisation (Karlberg *et al*, 2008; Schultz, 2010). This pathway can be summarised as:

1. The chemical must be bioavailable (i.e. dermal penetration).
2. A chemical must be: a) a direct-acting electrophilic compound; b) be converted from a non-reactive compound (pro-electrophilic) to a reactive metabolite via metabolism, or c) be converted form a non-reactive compound (pre-electrophilic) to a reactive derivative via an abiotic process, typically auto-oxidation (Lepoittevin, 2006).
3. Molecular sites of action are nucleophilic sites in proteins (e.g. cysteine and lysine residues) in the dermis (Roberts *et al*, 2008).
4. The molecular initiating event is the covalent perturbation of dermal proteins, which is irreversible (Schultz *et al*, 2006).
5. Biochemical pathways affected by the chemical's action on the molecular targets are incompletely known, but often include mitogen-activated protein signalling pathway and the oxidative stress response pathway.
6. Cellular-level outcomes are incompletely known, but include immune recognition of chemical allergens by Langerhans cells (specialised dermal dendritic cells) and dendritic cells responses in the form of expression of specific cell surface markers and cytokines, which are typically taken as evidence of dendritic cell maturation.
7. Organ-level responses (lymph node) include:
    a. Dendritic cell migration to the lymph node;
    b. formation and proliferation of allergen-specific memory T-cells.
8. Target organ is the skin with the requirement of intact local lymph nodes; the target tissues are the immune cells.
9. Key physiological response is acquisition of sensitivity (i.e. the ability to recognise hapten-specific hapten-protein complexes).

10. Key organism response is dermal inflammation upon receiving the chemical challenge in the elicitation phase.
11. Overall effect on the mammals is allergic contact dermatitis in humans or its rodent equivalent contact hypersensitivity.

Events along an adverse outcome pathway are either chemical structure-related and thus amenable to structure-activity analyses and read-across, or structure-independent events reflecting the cascade of biological events which follow the molecular initiating event. The latter events are important in confirming the outcome pathway is being followed at the different levels of biological organisation; they may not be a critical event required to predict potency.

To support an event as being critical to an adverse outcome pathway, there needs to be experimental data which characterise and measure the event. Critical events are ones that effect potency or are points of divergence along the biological responses leading to the adverse outcome. In relation to skin sensitisation, there are currently a limited number of such events which can be used to build the rationale for the read-across. All of these events are associated with the induction phase of sensitisation, which makes sense since LLNA only reports the induction phase of sensitisation.

C.7 In vitro data as weight of evidence for read-across

The International Program on Chemical Safety has played a central role in harmonising global approaches to chemical risk assessment. Building on earlier work, this program recently released a framework applicable to non-cancer endpoints (Boobis *et al*, 2008). The central focus of the framework is on the hypothesise MoA, which is comprised of the ‘key events’ causally related to the toxic effect, identified using an approach similar to the Bradford Hill criteria. First qualitative, and then quantitative comparisons, are made between data from experimental models (*in vitro* and *in vivo*) and the human situation. The output is a clear statement of the confidence, analysis, interpretations, and conclusions, and is intended to be transparent evaluation of all available data. Using the knowledge of and information available on selected events along the adverse outcome pathway, strategy can add confidence to the results from read-across data gap filling in the assessment of skin sensitisation. This knowledge/information can further support the mechanistic plausibility of the hypothesised protein binding reaction and provide a weight of evidence to the prediction.

C.7.1 Biochemical pathways related to skin sensitisation

Biochemical or cellular pathways affected by the action of reactive chemicals on molecular targets are incompletely known. However, there is evidence that mitogen-activated protein signalling pathways are particularly important to skin sensitisation. In particular, the biochemical pathways involving extracellular signal-regulating kinases, the c-Jun N-terminal kinases and the p38 kinases, have been shown to be activated upon exposure to protein-binding chemicals. These studies are very preliminary and based on only a few compounds in particular 1-chloro-2,4-dinitrobenzene and nickel dichloride (see Vanderbriel and Van Loveren, 2010 for further discussion).

Among the recent investigations on *in vitro* assays based on cellular pathways which show promise as an alternative for detect protein reactivity of chemicals acting via covalent binding to thiol groups, is one based on the stimulation of the DNA antioxidant-response element (ARE), also known as electrophilic response element in the reporter cell line AREc32, which contains an eight-fold repeat of the rat GSTA2 ARE-sequence upstream of a luciferase reporter gene in the human breast cancer cell line MCF7. Natsch and Emter (2008) described this Keap1-Nrf2-ARE assay. Briefly, the signalling pathway involves the repressor protein Keap1 (Kelch-like ECH-associates protein 1) and the transcription factor Nrf2 (nuclear factor-erythroid 2-related factor 2), which binds to the antioxidant response element (ARE) in the promoter of many phase II detoxification genes. The Keap1-Nrf2-ARE assay is a potential cellular marker because Keap1 is thiol-rich sensor protein which has been shown to be covalently modified by electrophilic compounds that leads to activation of ARE-dependent genes (Dinkova-Kostova *et al*, 2005). Natsch *et al* (2009) show the Keap1-Nrf2-ARE assay responds in a quantitative fashion, which is related to LLNA potency (i.e. strong, moderate, and weak).

Recently, Natsch (2010) discussed Keap1-Nrf2-ARE as a specific cellular toxicity pathway that is relevant to skin sensitisation and notes that the pathway appears to be selective for reactions that preferentially target thiol moieties. It appears to be a logical complement to *in chemico* measurements of reactivity (Gerberick *et al*, 2008) in measuring a selective biological responses key to the adverse outcome pathway.

Currently the Keap1-Nrf2-ARE assay offer a straightforward means of measuring electrophilic reactivity at the cellular level of organisation in a high-throughput fashion. As such, it is the proof of concept of an *in vitro* assay, which is an alternative method related to the skin sensitisation pathway. It is amenable to building databases required to address issues, such as it being a point of divergence along the adverse outcome pathway by reacting differently to different reactions, reacting preferentially with the thiol targets, and by its measured endpoint the EC1.5 being quantitative in nature (e.g. different values for acrylates, crotonates, and methacrylates).

C.7.2 Dendritic cell recognition and activation:

In vitro assays for sensitisation using cell or tissue cultures are particular to event 5 of the adverse outcome pathway - immune recognition of chemical allergens by dendritic cells (Dos Santo *et al*, 2009). Typically, such assays have, as their endpoint, a single event associated with the stimulation of dendritic cells. While measurement of the expression of certain surface markers (e.g. MHC class I and class II molecules), adhesion molecules (e.g. CD54), co-stimulatory molecules (e.g. CD86) or measurement of secretion of specific cytokines (e.g. IL-1β) are an area of active research (Vanderbriel and Van Loveren, 2010), no assays have yet been robust evaluated with chemicals representing different chemical reactions or the different thiol and amino targets. The assay, which is closest to meeting these criteria, is VITOSEN a human CD34+ progenitor-derived dendritic cell (CD34-DC) assay (Hooyberghs *et al,* 2008). Briefly, this *in vitro* assay uses CD34+ cells isolated from human cord blood and is based on the differential expression of the cAMP-responsive element modulator (CREM) and monocyte chemotactic protein-1 receptor (CCR2), which is highly discriminatory between chemical skin sensitisers and non-sensitisers after six hours of exposure (Hooyberghs *et al,* 2008). More recently, Lambrechts *et al* (2010) expanded the use of the CD34-DC assay to include the IC_{20} endpoint (i.e. the concentration of a compound, which causes 20% cell damage in CD34-DC). More of interest is that the data for 15 chemicals (representing several reactions, which target thiol) the IC20, which is independent from the gene expression changes, linearly correlates with LLNA EC_3.

Two dendritic cell-like systems should be mentioned, as they have drawn particular attention. The human cell line activation test (h-CLAT) measures CD54 and CD86 expression by the pro-monocytic THP-1 cell line (Ashikaga *et al*, 2006; Sakaguchi *et al*, 2006, 2009). In contrast, the myeloid U937 skin sensitisation test (MUSST) (see Python *et al,* 2007) measures CD86 expression.

C.7.4 Conclusion

The adverse outcome pathway establishes both chemical and biological relevance for use in selected *in vitro* methods and data used as weight of evidence in support of a read-across. Biochemical or cellular pathways, especially mitogen-activated protein signalling pathways, are particularly important to skin sensitisation. While the Keap1-Nrf2-ARE assay is the best example of a cellular pathway/test system, others may be applicable to skin sensitisation. As they are currently data-limited, it is unclear if dendritic cell recognition/activation modulates sensitisation potency or only reflects the cascade of biological events.

C.8 Consequences of the read-across in skin sensitisation:

To shift the skin sensitisation assessment paradigm to one based on non-animal test methods is complicated by the fact that the alternative method(s) must be shown to be relevant to the *in vivo* endpoint they are expected to replace. Read-across, since it is based on *in vivo* data, side steps this issue of relevancy. However, adapting chemical reactivity into the read-across for skin sensitisation will require accepting an additional layer of uncertainty; namely, that which associates with the basis for determining the final category or subcategory for read-across.

It is accepted by many that the uncertainty of an LLNA EC_3 value is between ½ and 2 times of the report value. This uncertainty appears to have little effect on extreme or strong sensitisers. However, correctly classifying moderate and weak sensitisers is, at present, difficult and often requires judgment calls.

Currently, we are not sure what the uncertainty is with placing a chemical into a reaction-based category, especially a potency subcategory. While placing a slow reacting chemical into a fast reacting category will over estimate its hazard by reading across from more potency chemicals, placing a fast reacting chemical into a slow reacting category will under estimate its hazard by reading across from more potency chemicals.

C.8.1 Conclusion

Qualitative read-across for skin sensitisation based on common 2-dimensional structural alerts is a well-accepted practice, especially for reactions such as Michael addition that are well represented within *in vivo* databases. Evidence shows that quantitative read-across for skin sensitisation is only valid if performed within a particular subcategory (e.g. acrylates versus methacrylates) and within a specified chemical reaction (e.g. Michael addition). Further evidence is presented suggesting that the use of iso-reactive groups may provide a means of expanding read-across to other subcategories. Read-across appears to be a good way of filling data gaps for extreme and strong sensitisers, especially when supplemented with reactivity data. However, it currently appears to be less useful in filling data gaps for moderate and weak sensitisers. While both *in silico* and *in chemico* data help in identifying non-sensitisers, read-across predictions for skin sensitisation are generally more acceptable if they predict a positive rather than a negative effect.

APPENDIX D: REQUEST FOR RESEARCH PROPOSALS

D.1 Title: Approaches for read-across in chemical risk assessment

D.2 Background

Traditional toxicity testing methods are resource-intensive. With the advent of legislation such as REACH that mandates the development of comprehensive toxicity information for large numbers of chemicals, it is clear that current methods cannot keep pace with the demand for information. Read-across approaches provide a practical alternative for the development of information on chemicals that are structurally related to already tested compounds. Read-across methods are supported by the large database of toxicology studies that have been conducted over the past 50+ years. These data have been compiled into formats that can be searched by chemical structure / substructure. Heuristics that identify appropriate structure search strategies, and which can be used to classify the suitability of analogues, have been developed for some aspects of toxicity, most notably toxicity that is dependent upon reactive chemistry. Some rules have also been developed for receptor affinity (e.g. the oestrogens receptor), although the development of heuristics for weak interactions such as receptor binding and enzyme inhibition is an area for further research.

In many cases, the outcome of a structure search is sufficiently robust that it can be viewed as definitive. However, in other cases the result can be viewed as a hypothesis that requires further testing. Toxicogenomics, metabolomics, and other high-information-content technologies provide an opportunity to test such hypotheses rapidly and completely with a minimum of animals. They also provide the potential to rapidly populate toxicity databases with information that will both augment the traditional toxicology information and facilitate the discovery of chemical classes based on biological activity in addition to chemical structure.

D.3 Objective

The objective is to improve the process of read-across by the development of new heuristics that support the identification of appropriate analogues, the development of high-information-content approaches to test the results of read-across and to augment databases with enough information that chemical classes can be identified by similar biological activity, not just chemical structure.

D.4 Scope and deliverables

1. Development of transparent expert-based rules for analogue search, based on existing knowledge of chemistry, xenobiotic metabolism and toxicology. Ideally, these rules will extend beyond reactive chemistry to encompass weak interactions such as receptor binding and enzyme inhibition.
2. Development of a critical mass of high-information-content data to support the testing of hypotheses generated by chemical analogue identification. This includes the incorporation of existing data (in public and private databases) in a format that is easily transferable (platform independent).
3. Elucidating the value and limitations of various high-information content data streams in supporting read-across. This includes a critical evaluation of the predictive value of these technologies versus traditional toxicity results.
4. Providing a process for using high-information content data to improve expert-rules and to incorporate high-info-content data into chemical databases to support better hypothesis generation. This includes the incorporation of existing data (in public and private databases).
5. Each point will be illustrated by examples that cover several relevant modes of action and endpoints.
6. Multi-disciplinary approach including chemistry, toxicology, biotechnology and informatics.

D.5 Costs, duration, timing

A budget of 300,000 Euros per year for 3 years is needed. The project might start in 2011.

D.6 Reference

ECETOC. 2009. Advanced technologies in read-across for chemical risk assessment. Technical Report No. 109. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium. (This report).

MEMBERS OF THE TASK FORCE

George Daston (Chairman)	Procter & Gamble USA - Cincinnati
Nora Aptula	Unilever UK - Sharnbrook, Bedford
Frank Bringezu	Merck D - Darmstadt
Jens-Olaf Eichler	BASF D - Ludwigshafen
Sylvia Jacobi	Albemarle Europe B - Louvain-La-Neuve
Grace Patlewicz	DuPont Haskell Global Centers for Health & Environmental Sciences USA - Newark
Nick Sturgess	Syngenta UK - Jealott's Hill, Bracknell
Henk Vrijhof (Scientific Secretary)	ECETOC B - Brussels

Acknowledgements

Dr Watze de Wolf (European Chemicals Agency, Helsinki, Finland) provided personal comments during peer review.

The Task Force is indebted to Dr Terry W. Schultz (The Schultz Group, Knoxville, Tennessee, USA) for preparing, and to CEFIC-LRI for funding, a review of current methods used for read-across in the assessment of skin sensitisation (Appendix C).

MEMBERS OF THE SCIENTIFIC COMMITTEE

(Peer review committee)

F. Lewis (Chairman) Head of Environmental Safety - EAME	Syngenta UK - Jealott's Hill, Bracknell
D. Owen [a] (Vice Chairman) Regulatory and Science Issues Manager	Shell UK - London
R. Bars Team Leader, Toxicology Research	Bayer CropScience F - Sophia Antipolis
P. Calow Professor, Dept. of Environmental, Social and Spatial Change	Roskilde University DK - Copenhagen
D. Farrar Occupational Health Business Manager	Ineos Chlor UK - Runcorn
A. Flückiger Head of Corporate Health Protection	F. Hoffmann - La Roche CH - Basel
H. Greim Director, Institute of Toxicology and Environmental Hygiene	Technical University München D - München
G. Malinverno Governmental & Regulatory Affairs – EU and Italian Manager	Solvay B - Brussels / I - Milano
S. Marshall Environmental Science Leader	Unilever SEAC UK - Bedford
C. Money Industrial Hygiene Adviser, Europe	ExxonMobil B - Brussels
M. Pemberton Global Product Integrity Manager	Lucite UK - Billingham

[a] Responsible for primary peer review.

MEMBERS OF THE SCIENTIFIC COMMITTEE (cont'd)

C. Rodriguez Principal Toxicologist, Corporate Central Product Safety	Procter and Gamble B - Strombeek-Bever
D. Salvito Vice president, Environmental Sciences	RIFM USA - Woodcliff Lake
G. Swaen Epidemiologist, Epidemiology, Health Services	Dow NL - Terneuzen
J. Tolls Director Environmental Safety Assessment	Henkel D - Düsseldorf
S. van der Vies Professor of Biochemistry	Vrije Universiteit Amsterdam NL - Amsterdam
B. van Ravenzwaay Senior Vice President - Experimental Toxicology	BASF D - Ludwigshafen
H.-J. Wiegand Product Stewardship, Corporate Environment, Safety, Health, Quality	Evonik D - Essen

ECETOC PUBLISHED REPORTS

Monographs

No.	Title
No. 1	Good Laboratory Practice (Published October 1979)
No. 2	A Contribution to Strategy for Identification and Control of Occupational Carcinogens (Published September 1980)
No. 3	Risk Assessment of Occupational Chemical Carcinogens (Published May 1985)
No. 4	Hepatocarcinogenesis in Laboratory Rodents: Relevance for Man (Published October 1982)
No. 5	Identification and Assessment of the Effects of Chemicals on Reproduction and Development (Reproductive Toxicology) (Published December 1983)
No. 6	Acute Toxicity Tests, LD_{50} (LC_{50}) Determinations and Alternatives (Published May 1985)
No. 7	Recommendations for the Harmonisation of International Guidelines for Toxicity Studies (Published December 1985)
No. 8	Structure-Activity Relationships in Toxicology and Ecotoxicology: An Assessment (Summary) (Published June 1986)
No. 9	Assessment of Mutagenicity of Industrial and Plant Protection Chemicals (Published June 1987)
No. 10	Identification of Immunotoxic Effects of Chemicals and Assessment of their Relevance to Man (Published August 1987)
No. 11	Eye Irritation Testing (Published June 1988)
No. 12	Alternative Approaches for the Assessment of Reproductive Toxicity (with emphasis on embryotoxicity/teratogenicity) (Published November 1989)
No. 13	DNA and Protein Adducts: Evaluation of their Use in Exposure Monitoring and Risk Assessment (Published October 1989)
No. 14	Skin Sensitisation Testing (Published March 1990)
No. 15	Skin Irritation (Published July 1990)
No. 16	Early Indicators of Non-Genotoxic Carcinogenesis (Published June 1991)
No. 17	Hepatic Peroxisome Proliferation (Published May 1992)
No. 18	Evaluation of the Neurotoxic Potential of Chemicals (Published September 1992)
No. 19	Respiratory Allergy (Published August 1993)
No. 20	Percutaneous Absorption (Published August 1993)
No. 21	Immunotoxicity: Hazard Identification and Risk Characterisation (Published September 1994)
No. 22	Evaluation of Chemicals for Oculotoxicity (Published November 1994)
No. 23	Receptor Mediated Mechanisms in Chemical Carcinogenesis (Published December 1995)
No. 24	Risk Assessment for Carcinogens (Published July 1996)
No. 25	Practical Concepts for Dose Selection in Chronic Toxicity and Carcinogenicity Studies in Rodents (Published February 1996)
No. 26	Aquatic Toxicity Testing of Sparingly Soluble Volatile and Unstable Substances (Published September 1996)
No. 27	Aneuploidy (Published August 1997)
No. 28	Dose-response and threshold-mediated mechanisms in mutagenesis - Mutation Research Special Issue (Published January 2000)
No. 29	Skin Sensitisation Testing for the Purpose of Hazard Identification and Risk Assessment (Published September 2000)
No. 30	Genetic Susceptibility to Environmental Toxicants (Published October 2001) Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, Volume 482, Issues 1-2, Pages 1-115 http://www.sciencedirect.com/science/journal/00275107
No. 31	Guidance on Evaluation of Reproductive Toxicity Data (Published February 2002)
No. 32	Use of Human Data in Hazard Classification for Irritation and Sensitisation (Published July 2002)

No. 33 Application of Physiological - Toxicokinetic Modelling to Health Hazard Assessment of Chemical Substances (Published February 2003)
Toxicology Letters, Volume 138, Issues 1-2
http://www.sciencedirect.com/science/journal/03784274

No. 34 Toxicogenomics in Genetic Toxicology and Hazard Determination (Published August 2005)
Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, Volume 575, Issues 1-2
http://www.sciencedirect.com/science/journal/00275107

No. 35 Biomarkers and molecular epidemiology (Published August 2006)
Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, Volume 600, Issues 1-2
http://www.sciencedirect.com/science/journal/00275107

No. 36 Environmental Genotoxins in Children and Adults (Published August 2006)
Mutation Research/Genetic Toxicology and Environmental Mutagenesis, Volume 608, Issue 2
http://www.sciencedirect.com/science/journal/13835718

No. 37 Biomarkers in Children and Adults (Published July 2007)
Toxicology Letters, Volume 172, Nos. 1-2
http://www.sciencedirect.com/science/journal/03784274

No. 38 Toxicity of Engineered Nanomaterials (published May 2009)
Toxicology Letters, Volume 186, Issue 3
http://www.sciencedirect.com/science/journal/03784274

Technical Reports

No.	Title
No. 1	Assessment of Data on the Effects of Formaldehyde on Humans (Published January 1979) (Updated by TR No. 6)
No. 2	The Mutagenic and Carcinogenic Potential of Formaldehyde (Published May 1981)
No. 3	Assessment of Test Methods for Photodegradation of Chemicals in the Environment (Published August 1981)
No. 4	The Toxicology of Ethylene Glycol Monoalkyl Ethers and its Relevance to Man (Published June 1982) (Updated by TR No. 17)
No. 5	Toxicity of Ethylene Oxide and its Relevance to Man (Published September 1982)
No. 6	Formaldehyde Toxicology: An Up-Dating of ECETOC Technical Reports 1 and 2 (Published September 1982)
No. 7	Experimental Assessment of the Phototransformation of Chemicals in the Atmosphere (Published September 1983)
No. 8	Biodegradation Testing: An Assessment of the Present Status (Published November 1983)
No. 9	Assessment of Reverse-Phase Chromatographic Methods for Determining Partition Coefficients (Published December 1983)
No. 10	Considerations Regarding the Extrapolation of Biological Data in Deriving Occupational Exposure Limits (Published February 1984)
No. 11	Ethylene Oxide Toxicology and its Relevance to Man: An Up-Dating of ECETOC Technical Report No. 5 (Published March 1984)
No. 12	The Phototransformation of Chemicals in Water: Results of a Ring-Test (Published June 1984)
No. 13	The EEC 6th Amendment: A Guide to Risk Evaluation for Effects on the Environment (Published March 1984)
No. 14	The EEC 6th Amendment: A Guide to Risk Evaluation for Effects on Human Health (Published March 1984)
No. 15	The Use of Physical-Chemical Properties in the 6th Amendment and their Required Precision, Accuracy and Limiting Values (Published June 1984)
No. 16	A Review of Recent Literature on the Toxicology of Benzene (Published December 1984)
No. 17	The Toxicology of Glycol Ethers and its Relevance to Man: An Up-Dating of ECETOC Technical Report No. 4) (Published April 1985) (Updated by TR No. 64)
No. 18	Harmonisation of Ready Biodegradability Tests (Published April 1985)
No. 19	An Assessment of Occurrence and Effects of Dialkyl-o-Phthalates in the Environment (Published May 1985)
No. 20	Biodegradation Tests for Poorly-Soluble Compounds (Published February 1986)
No. 21	Guide to the Classification of Carcinogens, Mutagens, and Teratogens under the 6th Amendment (Published February 1986)
No. 22	Classification of Dangerous Substances and Pesticides in the EEC Directives. A Proposed Revision of Criteria for Inhalational Toxicity (Published January 1987)
No. 23	Evaluation of the Toxicity of Substances to be Assessed for Biodegradability (Published November 1986)
No. 24	The EEC 6th Amendment: Prolonged Fish Toxicity Tests (Published October 1986)
No. 25	Evaluation of Fish Tainting (Published January 1987)
No. 26	The Assessment of Carcinogenic Hazard for Human Beings exposed to Methylene Chloride (Published January 1987)
No. 27	Nitrate and Drinking Water (Published January 1988)
No. 28	Evaluation of Anaerobic Biodegradation (Published June 1988)
No. 29	Concentrations of Industrial Organic Chemicals Measured in the Environment: The Influence of Physico-Chemical Properties, Tonnage and Use Patterns (Published June 1988)
No. 30	Existing Chemicals: Literature Reviews and Evaluations (Fifth Edition) (No longer available) (Published May 1994)
No. 31	The Mutagenicity and Carcinogenicity of Vinyl Chloride: A Historical Review and Assessment (Published July 1988)
No. 32	Methylene Chloride (Dichloromethane): Human Risk Assessment Using Experimental Animal Data (Published May 1988)

No. 33 Nickel and Nickel Compounds: Review of Toxicology and Epidemiology with Special Reference to Carcinogenesis (Published February 1989)
No. 34 Methylene Chloride (Dichloromethane): An Overview of Experimental Work Investigating Species Differences in Carcinogenicity and their Relevance to Man (Published March 1989)
No. 35 Fate, Behaviour and Toxicity of Organic Chemicals Associated with Sediments (Published January 1990)
No. 36 Biomonitoring of Industrial Effluents (Published April 1990)
No. 37 Tetrachlorethylene: Assessment of Human Carcinogenic Hazard (Published May 1990)
No. 38 A Guide to the Classification of Preparations Containing Carcinogens, Mutagens and Teratogens (Published July 1990)
No. 39 Hazard Assessment of Floating Chemicals After an Accidental Spill at Sea (Published July 1990)
No. 40 Hazard Assessment of Chemical Contaminants in Soil (Published April 1992)
No. 41 Human Exposure to N-Nitrosamines, their Effects and a Risk Assessment for N-Nitrosodiethanolamine in Personal Care Products (Published August 1990)
No. 42 Critical Evaluation of Methods for the Determination of N-Nitrosamines in Personal Care and Household Products (Published February 1991)
No. 43 Emergency Exposure Indices for Industrial Chemicals (Published March 1991)
No. 44 Biodegradation Kinetics (Published September 1991)
No. 45 Nickel, Cobalt and Chromium in Consumer Products: Allergic Contact Dermatitis (Published March 1992)
No. 46 EC 7th Amendment: Role of Mammalian Toxicokinetic and Metabolic Studies in the Toxicological Assessment of Industrial Chemicals (Published May 1992)
No. 47 EC 7th Amendment "Toxic to Reproduction": Guidance on Classification (Published August 1992)
No. 48 Eye Irritation: Reference Chemicals Data Bank (Second Edition) (Published June 1998)
No. 49 Exposure of Man to Dioxins: A Perspective on Industrial Waste Incineration (Published December 1992)
No. 50 Estimating Environmental Concentrations of Chemicals using Fate and Exposure Models (Published November 1992)
No. 51 Environmental Hazard Assessment of Substances (Published January 1993)
No. 52 Styrene Toxicology Investigation on the Potential for Carcinogenicity (Published August 1992)
No. 53 DHTDMAC: Aquatic and Terrestrial Hazard Assessment (CAS No. 61789-80-8) (Published February 1993)
No. 54 Assessment of the Biodegradation of Chemicals in the Marine Environment (Published August 1993)
No. 55 Pulmonary Toxicity of Polyalkylene Glycols (Published December 1997)
No. 56 Aquatic Toxicity Data Evaluation (Published December 1993)
No. 57 Polypropylene Production and Colorectal Cancer (Published February 1994)
No. 58 Assessment of Non-Occupational Exposure to Chemicals (Published May 1994)
No. 59 Testing for Worker Protection (Published April 1994)
No. 60 Trichloroethylene: Assessment of Human Carcinogenic Hazard (Published May 1994)
No. 61 Environmental Exposure Assessment (Published September 1994)
No. 62 Ammonia Emissions to Air in Western Europe (Published July 1994)
No. 63 Reproductive and General Toxicology of some Inorganic Borates and Risk Assessment for Human Beings (Published February 1995)
No. 64 The Toxicology of Glycol Ethers and its Relevance to Man (Published August 1995)
No. 65 Formaldehyde and Human Cancer Risks (Published May 1995)
No. 66 Skin Irritation and Corrosion: Reference Chemicals Data Bank (Published March 1995)
No. 67 The Role of Bioaccumulation in Environmental Risk Assessment: The Aquatic Environment and Related Food Webs (Published October 1995)
No. 68 Assessment Factors in Human Health Risk Assessment (Published August 1995) (Updated by TR No. 86)
No. 69 Toxicology of Man-Made Organic Fibres (Published April 1996)
No. 70 Chronic Neurotoxicity of Solvents (Published February 1996)
No. 71 Inventory of Critical Reviews on Chemicals (Only available to ECETOC members) (Published August 1996)

No. 72 Methyl *tert*-Butyl Ether (MTBE) Health Risk Characterisation (Published June 1997)
No. 73 The Value of Aquatic Model Ecosystem Studies in Ecotoxicology (Published December 1997)
No. 74 QSARs in the Assessment of the Environmental Fate and Effects of Chemicals (Published June 1998)
No. 75 Organophosphorus Pesticides and Long-term Effects on the Nervous System (Published December 1998)
No. 76 Monitoring and Modelling of Industrial Organic Chemicals, with Particular Reference to Aquatic Risk Assessment (Published January 1999)
No. 77 Skin and Respiratory Sensitisers: Reference Chemicals Data Bank (Published August 1999)
No. 78 Skin Sensitisation Testing: Methodological Considerations (Published December 1999)
No. 79 Exposure Factors Sourcebook for European Populations (with Focus on UK Data) (Published June 2001)
No. 80 Aquatic Toxicity of Mixtures (Published July 2001)
No. 81 Human Acute Intoxication from Monochloroacetic Acid: Proposals for Therapy (Published November 2001)
No. 82 Risk Assessment in Marine Environments (Published December 2001)
No. 83 The Use of T25 Estimates and Alternative Methods in the Regulatory Risk Assessment of Non-threshold Carcinogens in the European Union (Published December 2002)
No. 84 Scientific Principles for Soil Hazard Assessment of Substances (Published July 2002)
No. 85 Recognition of, and Differentiation between, Adverse and Non-adverse Effects in Toxicology Studies (Published December 2002)
No. 86 Derivation of Assessment Factors for Human Health Risk Assessment (Published February 2003)
No. 87 Contact Sensitisation: Classification According to Potency (Published April 2003)
No. 88 Environmental Risk Assessment of Difficult Substances (Published June 2003)
No. 89 (Q)SARS: Evaluation of the Commercially Available Software for Human Health and Environmental Endpoints with Respect to Chemical Management Applications (Published September 2003)
No. 90 Persistence of Chemicals in the Environment (Published October 2003)
No. 91 Aquatic Hazard Assessment II (Published November 2003)
No. 92 Soil and Sediment Risk Assessment (Published December 2004)
No. 93 Targeted Risk Assessment (Published December 2004)
No. 94 Whole Effluent Assessment (Published December 2004)
No. 95 The Toxicology of Glycol Ethers and its Relevance to Man (Fourth Edition) Volume I and Volume II Substance Profiles (Published February 2005)
No. 96 Trends in Children's Health and the Role of Chemicals: State of the Science Review (Published June 2005)
No. 97 Alternative Testing Approaches in Environmental Safety Assessment (Published December 2005)
No. 98 Risk Assessment of PBT Chemicals (Published December 2005)
No. 99 Toxicological Modes of Action: Relevance for Human Risk Assessment (Published July 2006)
No. 100 Contribution to the Methodology for the Development of Acute Exposure Threshold Levels in Case of Accidental Chemical Release (Published July 2006)
No. 101 Guidance for Setting Occupational Exposure Limits: Emphasis on Data-Poor Substances (Published October 2006)
No. 102 Intelligent Testing Strategies in Ecotoxicology: Mode of Action Approach for Specifically Acting Chemicals (Published December 2007)
No. 103 Toxicity of Possible Impurities and By-products in Fluorocarbon Products (Published December 2008)
No. 104 Framework for the Integration of Human and Animal Data in Chemical Risk Assessment (Published January 2009)
No. 105 Evaluation of Cardiac Sensitisation Test Methods (Published October 2009)
No. 106 Guidance on Identifying Endocrine Disrupting Effects (Published June 2009)
No. 107 Addendum to ECETOC Targeted Risk assessment report No. 93 (Published December 2009)
No. 108 Collation of Existing Marine Biodegradation Data and its Use in Environmental Risk Assessment (Published December 2009)
No. 109 High information content technologies in support of read-across in chemical risk assessment (Published December 2010)
No. 110 Guidance on Assessment Factors to Derive a DNEL (Published November 2010)

Joint Assessment of Commodity Chemicals (JACC) Reports

No.	Title
No. 1	Melamine (Published February 1983)
No. 2	1,4-Dioxane (Published February 1983)
No. 3	Methyl Ethyl Ketone (Published February 1983)
No. 4	Methylene Chloride (Published January 1984)
No. 5	Vinylidene Chloride (Published August 1985)
No. 6	Xylenes (Published June 1986)
No. 7	Ethylbenzene (Published August 1986)
No. 8	Methyl Isobutyl Ketone (Published May 1987)
No. 9	Chlorodifluoromethane (Published October 1989)
No. 10	Isophorone (Published September 1989)
No. 11	1,2-Dichloro-1,1-difluoroethane (HFA-132b) (Published May 1990)
No. 12	1-Chloro-1,2,2,2-tetrafluoroethane (HFA-124) (Published May 1990) (Updated by JACC No. 25)
No. 13	1,1-Dichloro-2,2,2-trifluoroethane (HFA-123) (Published May 1990) (Updated by JACC No. 33)
No. 14	1-Chloro-2,2,2-trifluoromethane (HFA-133a) (Published August 1990)
No. 15	1-Fluoro 1,1-dichloroethane (HFA-141) (Published August 1990) (Updated by JACC No. 29)
No. 16	Dichlorofluoromethane (HCFC-21) (Published August 1990)
No. 17	1-Chloro-1,1-difluoroethane (HFA-142b) (Published August 1990)
No. 18	Vinyl Acetate (Published February 1991)
No. 19	Dicyclopentadiene (CAS: 77-73-6) (Published July 1991)
No. 20	Tris-/Bis-/Mono-(2 ethylhexyl) phosphate (Published May 1992)
No. 21	Tris-(2-butoxyethyl)-phosphate (CAS: 78-51-3) (Published March 1992)
No. 22	Hydrogen Peroxide (CAS: 7722-84-1) (Published January 1993)
No. 23	Polycarboxylate Polymers as Used in Detergents (Published November 1993)
No. 24	Pentafluoroethane (HFC-125) (CAS: 354-33-6) (Published May 1994)
No. 25	1-Chloro-1,2,2,2-tetrafluoroethane (HCFC 124) (CAS No. 2837-89-0) (Second Edition) (Published July 1994) (Updated by JACC 46)
No. 26	Linear Polydimethylsiloxanes (CAS No. 63148-62-9) (Published September 1994)
No. 27	*n*-Butyl Acrylate (CAS No. 141-32-2) (Published August 1994)
No. 28	Ethyl Acrylate (CAS No. 140-88-5) (Published September 1994)
No. 29	1,1-Dichloro-1-fluoroethane (HCFC-141b) (CAS No. 1717-00-6) (Published December 1994)
No. 30	Methyl Methacrylate (CAS No. 80-62-6) (Published February 1995)
No. 31	1,1,1,2-Tetrafluoroethane (HFC-134a) (CAS No. 811-97-2) (Published February 1995) (Updated by JACC No. 50)
No. 32	Difluoromethane (HFC-32) (CAS No. 75-10-5) (Published May 1995) (Updated by JACC No. 54)
No. 33	1,1-Dichloro-2,2,2-trifluoroethane (HCFC-123) (CAS No. 306-83-2) (Published February 1996) (Updated by JACC No. 47)
No. 34	Acrylic Acid (CAS No. 79-10-7) (Published September 1995)
No. 35	Methacrylic Acid (CAS No. 79-41-4) (Published May 1996)
No. 36	*n*-Butyl Methacrylate; Isobutyl Methacrylate (CAS No. 97-88-1) (CAS No. 97-86-9) (Published December 1996)
No. 37	Methyl Acrylate (CAS No. 96-33-3) (Published September 1998)
No. 38	Monochloroacetic Acid (CAS No. 79-11-8) and its Sodium Salt (CAS No. 3926-62-3) (Published June 1999)
No. 39	Tetrachloroethylene (CAS No. 127-18-4) (Published December 1999)
No. 40	Peracetic Acid (CAS No. 79-21-0) and its Equilibrium Solutions (Published January 2001)

No. 41 *n*-Butanol (CAS No. 71-36-3) (Published March 2004)
No. 42 Tetrafluoroethylene (CAS No. 116-14-3) (Published December 2003)
No. 43 *sec*-Butanol (CAS No. 78-92-2) (Published December 2004)
No. 44 1, 1, 1, 3, 3-Pentafluoropropane (HFC-245fa) (Published June 2004)
No. 45 1, 1-Difluoroethane (HFC-152a) (CAS No. 75-37-6) (Published September 2004)
No. 46 1-Chloro-1,2,2,2-tetrafluoroethane (HCFC 124) CAS No. 2837-89-0 (Third Edition) (Published November 2004)
No. 47 1,1-Dichloro-2,2,2-trifluoroethane (HCFC-123) CAS No. 306-83-2 (Third Edition) (Published May 2005)
No. 48 Hexafluoropropylene (HFP) CAS No. 116-15-4 (Published September 2005)
No. 49 Vinylidene Fluoride CAS No. 75-38-7 (Published November 2005)
No. 50 1,1,1,2-Tetrafluoroethane (HFC-134a) (CAS No. 811-97-2) (Second Edition) (Published January 2006)
No. 51 Synthetic Amorphous Silica (CAS No. 7631-86-9) (Published September 2006)
No. 52 Trifluoroethane (HFC-143a) CAS No. 420-46-2 (Published October 2006)
No. 53 Cyanides of Hydrogen, Sodium and Potassium, and Acetone Cyanohydrin (CAS No. 74-90-8, 143-33-9,151-50-8 and 75-86-5) (Published September 2007)
No. 54 Difluoromethane (HFC-32) CAS No. 75-10-5 (Second Edition) (Published June 2008)

Special Reports

No.	Title
No. 8	HAZCHEM; A Mathematical Model for Use in Risk Assessment of Substances (Published October 1994)
No. 9	Styrene Criteria Document (Published June 1995)
No. 10	Hydrogen Peroxide OEL Criteria Document (CAS No. 7722-84-1) (Published July 1996)
No. 11	Ecotoxicology of some Inorganic Borates (Published March 1997)
No. 12	1,3-Butadiene OEL Criteria Document (Second Edition) (CAS No. 106-99-0) (Published January 1997)
No. 13	Occupational Exposure Limits for Hydrocarbon Solvents (Published August 1997)
No. 14	*n*-Butyl Methacrylate and Isobutyl Methacrylate OEL Criteria Document (Published May 1998)
No. 15	Examination of a Proposed Skin Notation Strategy (Published September 1998)
No. 16	GREAT-ER User Manual (Published March 1999)
No. 17	Risk Assessment Report for Existing Substances Methyl *tertiary*-Butyl Ether (Published December 2003)

Documents

No.	Title
No. 32	Environmental Oestrogens: Male Reproduction and Reproductive Development (Published January 1996)
No. 33	Environmental Oestrogens: A Compendium of Test Methods (Published July 1996)
No. 34	The Challenge Posed by Endocrine-disrupting Chemicals (Published February 1996)
No. 35	Exposure Assessment in the Context of the EU Technical Guidance Documents on Risk Assessment of Substances (Published May 1997)
No. 36	Comments on OECD Draft Detailed Review Paper: Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals (Published August 1997)
No. 37	EC Classification of Eye Irritancy (Published December 1997)
No. 38	Wildlife and Endocrine Disrupters: Requirements for Hazard Identification (Published January 1998)
No. 39	Screening and Testing Methods for Ecotoxicological Effects of Potential Endocrine Disrupters: Response to the EDSTAC Recommendations and a Proposed Alternative Approach (Published January 1999)
No. 40	Comments on Recommendation from Scientific Committee on Occupational Exposure Limits for 1,3-Butadiene (Published October 2000)
No. 41	Persistent Organic Pollutants (POPs) Response to UNEP/INC/CEG-I Annex 1 (Published January 2000)
No. 42	Genomics, Transcript Profiling, Proteomics and Metabonomics (GTPM). An Introduction (Published April 2001)
No. 43	Contact Sensitisation: Classification According to Potency. A Commentary (Published July 2003)
No. 44	Guidance for the Interpretation of Biomonitoring Data (Published November 2005)
No. 45	Triggering and Waiving Criteria for the Extended One-Generation Reproduction Toxicity Study (Published March 2008)
No. 46	Potency Values from the Local Lymph Node Assay: Application to Classification, Labelling and Risk Assessment (Published December2008)

Workshop Reports

No.	Title
No. 1	Workshop on Availability, Interpretation and Use of Environmental Monitoring Data. 20-21 March 2003, Brussels (Published December 2003)
No. 2	Strategy Report on Challenges, Opportunities and Research needs arising from the Definition, Assessment and Management of Ecological Quality Status as required by the EU Water Framework Directive based on the workshop EQS and WFD versus PNEC and REACH - are they doing the job? 27-28 November 2003, Budapest (Published March 2004)
No. 3	Workshop on the Use of Human Data in Risk Assessment. 23-24 February 2004, Cardiff (Published November 2004)
No. 4	Influence of Maternal Toxicity in Studies on Developmental Toxicity. 2 March 2004, Berlin (Published October 2004)
No. 5	Workshop on Alternative Testing Approaches in Environmental Risk Assessment. 7-9 July 2004, Paris (Published December 2004)
No. 6	Workshop on Chemical Pollution, Respiratory Allergy and Asthma. 16-17 June 2005, Leuven (Published December 2005)
No. 7	Workshop on Testing Strategies to Establish the Safety of Nanomaterials. 7-8 November 2005, Barcelona (Published August 2006)
No. 8	Workshop on Societal Aspects of Nanotechnology. 7-8 November 2005, Barcelona (Published October 2006)
No. 9	Workshop on the Refinement of Mutagenicity/Genotoxicity Testing. 23-24 April 2007, Malta (Published September 2007)
No. 10	Workshop on Biodegradation and Persistence. 26-27 June 2007, Holmes Chapel (Published September 2007)
No. 11	Workshop on the Application of 'Omics in Toxicology and Ecotoxicology: Case Studies and Risk Assessment. 6-7 December 2007, Malaga (Published July 2008)
No. 12	Workshop on Triggering and Waiving Criteria for the Extended One-Generation Reproduction Toxicity Study. 14-15 April 2008, Barza d'Ispra (Published August 2008)
No. 13	Counting the Costs and Benefits of Chemical Controls: Role of Environmental Risk Assessment in Socio-Economic Analysis 4 June 2008, Brussels (Published September 2008)
No. 14	Use of Markers for Improved Retrospective Exposure Assessment in Epidemiology Studies. 24-25 June 2008, Brussels (Published February 2009)
No. 15	The Probabilistic Approaches for Marine Hazard Assessment. 18-19 June 2008, Oslo (Published June 2009)
No. 16	Guidance on interpreting endocrine disrupting effects. 29-30 June 2009, Barcelona (Published October 2009)
No. 17	Significance of Bound Residues in Environmental Risk Assessment. 14-15 October 2009, Brussels (Published December 2009)
No. 18	The Enhancement of the Scientific Process and Transparency of Observational Epidemiology Studies. 24-25 September 2009, London (Published December 2009)
No. 19	'Omics in (Eco)toxicology: Case Studies and Risk Assessment. 22-23 February 2010, Málaga (Published June 2010)

All ECETOC reports can be downloaded from http://www.ecetoc.org/publications