REPORT OF THE EXPERT CONSULTATION ON SCIENTIFIC AND REGULATORY EVALUATION OF ORGANIC CHEMISTRY MECHANISM-BASED STRUCTURAL ALERTS FOR THE IDENTIFICATION OF PROTEIN-BINDING CHEMICALS

Series on Testing and Assessment
No. 139

20 October 2010, Oxford, United Kingdom
REPORT OF THE EXPERT CONSULTATION ON SCIENTIFIC AND REGULATORY EVALUATION OF ORGANIC CHEMISTRY MECHANISM-BASED STRUCTURAL ALERTS FOR THE IDENTIFICATION OF PROTEIN BINDING CHEMICALS
Also published in the Series on Testing and Assessment:


No. 9, Guidance Document for the Conduct of Studies of Occupational Exposure to Pesticides during Agricultural Application (1997)


No. 12, Detailed Review Document on Classification Systems for Germ Cell Mutagenicity in OECD Member Countries (1998)

No. 13, Detailed Review Document on Classification Systems for Sensitising Substances in OECD Member Countries 1998)


No. 15, Detailed Review Document on Classification Systems for Reproductive Toxicity in OECD Member Countries (1998)

No. 16, Detailed Review Document on Classification Systems for Skin Irritation/Corrosion in OECD Member Countries (1998)
No. 17,  Environmental Exposure Assessment Strategies for Existing Industrial Chemicals in OECD Member Countries (1999)


No. 19,  Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (1999)


No. 25,  Detailed Review Document on Hazard Classification Systems for Specifics Target Organ Systemic Toxicity Repeated Exposure in OECD Member Countries (2001)

No. 26,  Revised Analysis of Responses Received from Member Countries to the Questionnaire on Regulatory Acute Toxicity Data Needs (2001)

No. 27,  Guidance Document on the Use of the Harmonised System for the Classification of Chemicals which are Hazardous for the Aquatic Environment (2001)


No. 29,  Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (2001)

No. 30,  Detailed Review Document on Hazard Classification Systems for Mixtures (2001)


No. 32,  Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (2000)
No. 33, Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures (2001)


No. 35, Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies (2002)


No. 38, Detailed Background Review of the Uterotrophic Assay Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Assay (2003)

No. 39, Guidance Document on Acute Inhalation Toxicity Testing (in preparation)


No. 41, Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which in Contact with Water Release Toxic Gases (2003)


No. 43, Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment (2008)

No. 44, Description of Selected Key Generic Terms Used in Chemical Hazard/Risk Assessment (2003)


No. 48,  New Chemical Assessment Comparisons and Implications for Work Sharing (2004)


No. 51,  Approaches to Exposure Assessment in OECD Member Countries: Report from the Policy Dialogue on Exposure Assessment in June 2005 (2006)

No. 52,  Comparison of emission estimation methods used in Pollutant Release and Transfer Registers (PRTRs) and Emission Scenario Documents (ESDs): Case study of pulp and paper and textile sectors (2006)


No. 57,  Detailed Review Paper on Thyroid Hormone Disruption Assays (2006)


No. 60,  Report of the Initial Work Towards the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1A) (2006)
No. 61,  Report of the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1B) (2006)


No. 64,  Guidance Document on Overview of Residue Chemistry Studies (2006)


No. 67,  Additional data supporting the Test Guideline on the Uterotrophic Bioassay in rodents (2007)

No. 68,  Summary Report of the Uterotrophic Bioassay Peer Review Panel, including Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the follow up of this report (2006)


No. 70,  Report on the Preparation of GHS Implementation by the OECD Countries (2007)

No. 71,  Guidance Document on the Uterotrophic Bioassay - Procedure to Test for Antioestrogenicity (2007)


No. 79, Validation Report of the Full Life-cycle Test with the Harpacticoid Copepods Nitocra Spinipes and Amphiascus Tenuiremis and the Calanoid Copepod Acartia Tonsa - Phase 1 (2007)

No. 80, Guidance on Grouping of Chemicals (2007)


No. 82, Guidance Document on Amphibian Thyroid Histology (2007)


No. 84, Report on the Workshop on the Application of the GHS Classification Criteria to HPV Chemicals, 5-6 July Bern Switzerland (2007)


No.88,  Workshop on Integrated Approaches to Testing and Assessment (2008)

No.89,  Retrospective Performance Assessment of the Test Guideline 426 on Developmental Neurotoxicity (2008)


No.95,  Detailed Review Paper on Fish Life-Cycle Tests (2008)


No.102,  Guidance Document for using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping of Chemicals (2009)
No. 103, Detailed Review Paper on Transgenic Rodent Mutation Assays (2009)

No. 104, Performance Assessment: Comparison of 403 and CxT Protocols via Simulation and for Selected Real Data Sets (2009)


No. 107, Preservative treated wood to the environment for wood held in storage after treatment and for wooden commodities that are not cover and are not in contact with ground. (2009)


No. 109, Literature review on the 21-Day Fish Assay and the Fish Short-Term Reproduction Assay (2009)

No. 110, Report of the validation peer review for the weanling Hershberger Bioassay and agreement of the working of national coordinators of the test guidelines programme on the follow-up of this report (2009)


No. 114, Performance Assessment of Different Cytotoxic and Cytostatic Measures for the In Vitro Micronucleus Test (MNVIT): Summary of results in the collaborative trial (2010)


No. 119, *Classification and Labelling of chemicals according to the UN Globally Harmonized System: Outcome of the Analysis of Classification of Selected Chemicals listed in Annex III of the Rotterdam Convention (2010)*

No. 120, *Explanatory Background Document to the OECD Draft Test Guideline on in vitro Skin Irritation Testing (2010)*

No. 121, *Detailed review paper (DRP) on Molluscs life-cycle Toxicity Testing (2010)*

No. 122, *Guidance Document on the determination of the Toxicity of a Test Chemical to the Dung Beetle Aphodius Constans (2010)*

No. 123, *Guidance Document on the Diagnosis of Endocrinerelated Histopathology in Fish Gonads (2010)*

No. 124, *Guidance for the Derivation of an Acute Reference Dose (2010)*

No. 125, *Guidance Document on Histopathology for Inhalation Toxicity Studies, Supporting TG 412 (Subacute Inhalation Toxicity: 28-Day) and TG 413 (Subchronic Inhalation Toxicity: 90-Day) (2010)*

No. 126, *Short Guidance on the Threshold approach for Acute Fish Toxicity (2010)*

No. 127, *Peer review report of the validation of the 21-day androgenised female stickleback screening assay (2010)*

No. 128, *Validation Report of the 21-day Androgenised Female Stickleback Screening Assay (2010)*


No. 133, *Peer Review Report for the H295R Cell-Based Assay for Steroidogenesis (2010)*


© OECD 2011
Applications for permission to reproduce or translate all or part of this material should be made to: Head of Publications Service, RIGHTS@oecd.org, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France
ABOUT THE OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 34 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD’s work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD’s workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in ten different series: Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides and Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and Safety of Manufactured Nanomaterials. More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD’s World Wide Web site (www.oecd.org/ehs/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. UNDP is an observer. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.
This publication is available electronically, at no charge.

For this and many other Environment, Health and Safety publications, consult the OECD’s World Wide Web site (www.oecd.org/ehs/)

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France

Fax: (33-1) 44 30 61 80

E-mail: ehscont@oecd.org
FOREWORD

This document is a report of the expert consultation held on 20 October 2010 with the aim to evaluate a set of structural alerts for estimating covalent binding of chemicals with proteins. This consultation was held based on a key recommendation from the OECD Workshop on Structural Alerts for the OECD (Q)SAR Application Toolbox held in May 2008 [see ENV/JM/MONO(2009)4]. The resulting set of alerts will be implemented in version 2.0 of the OECD (Q)SAR Toolbox.

This document is published under the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology of the OECD.
TABLE OF CONTENTS

ABOUT THE OECD .................................................................................................................................... 14
FOREWORD................................................................................................................................................ 16
REPORT OF THE EXPERT CONSULTATION ON SCIENTIFIC AND REGULATORY EVALUATION OF ORGANIC CHEMISTRY MECHANISM-BASED STRUCTURAL ALERTS FOR THE IDENTIFICATION OF PROTEIN BINDING CHEMICALS ................................................ 18
  Background ............................................................................................................................................... 18
  Workshop .................................................................................................................................................. 19
  Scope and Objectives ................................................................................................................................ 19
  Preparation for the Expert Consultation .................................................................................................... 19
  Preparatory Work by Experts .................................................................................................................... 20
  Proceedings of the Expert Consultation .................................................................................................... 21
  Outcome of the Expert Consultation ......................................................................................................... 21

ANNEX 1: AGENDA OF THE EXPERT CONSULTATION ON SCIENTIFIC AND REGULATORY EVALUATION OF ORGANIC CHEMISTRY MECHANISM-BASED STRUCTURAL ALERTS FOR THE IDENTIFICATION OF PROTEIN BINDING CHEMICALS ....... 23

ANNEX 2: LIST OF PARTICIPANTS............................................................................................................. 25

ANNEX 3: A REPORT ON RE-EVALUATION OF STRUCTURAL ALERTS FOR THE BINDING OF MOLECULES TO PROTEINS AND THE DEVELOPMENT OF A COMPREHENSIVE PROFILER OF ALERTS ............................................................................................... 30

ANNEX 4: SUPPLEMENTARY INFORMATION TO THE REPORT .............................................................. 59

ANNEX 5: PRESENTATION BY LIVERPOOL JOHN MOORES UNIVERSITY .............................................. 60

ANNEX 6: REVIEW REPORT BY DR. WOO ............................................................................................... 82

ANNEX 7: REVIEW REPORT BY DRS. VANDERBRIEL AND RORIJE ..................................................... 89

ANNEX 8: REVIEW REPORT BY DR. APTULA ......................................................................................... 95
REPORT OF THE EXPERT CONSULTATION ON SCIENTIFIC AND REGULATORY EVALUATION OF ORGANIC CHEMISTRY MECHANISM-BASED STRUCTURAL ALERTS FOR THE IDENTIFICATION OF PROTEIN BINDING CHEMICALS

Background

1. The OECD QSAR Toolbox has six work modules, which are used in a work flow with the goal of filling data gaps through the use of the chemical category. The Toolbox modules include: 1) Chemical Input; 2) Profiling; 3) Endpoints; 4) Category Definition; 5) Filling Data Gaps, and 6) Report. To build a category or to perform a simple analogue approach, the user goes through these modules sequentially. However, Category Definition and Profiling are the critical steps in the workflow. While the Toolbox provides many ways to set up a category, defining a category based on similar mechanisms or modes of action of its members is the most appropriate.

2. The careful use of expert judgment to define the boundaries of a chemical category is crucial to the reliable application of the Toolbox to estimate properties of untested chemicals. Formal definitions of which chemicals should be included in a category and conversely which chemicals should be excluded (i.e. well defined applicability domain) are essential for reliable estimates of missing values. The expert judgment for forming the category should be described in a transparent manner so that the category can be evaluated by others.

3. Experience from using the Toolbox has shown a common mechanism of action to be a critical factor in deciding what chemicals would be expected to be members of a category. Variations in chemical structure can affect both toxicokinetics (uptake and bioavailability) and toxicodynamics (e.g. interactions with receptors and enzymes). Two-D structural alerts (SAs) in the form of mechanistic profilers have proven to be useful in identifying a chemical category for filling data gaps. For example, having an amino group substituted on an aromatic system is relevant to enzymatic transformation to the hydroxylamine derivative and the hazard endpoint carcinogenicity.

4. The protein binding categorization scheme (Protein-binding profiler) used in the Toolbox (Version 1.1) is based on the model developed by the Laboratory of Mathematical Chemistry (LMC) and donated to the proof of concept version of the Toolbox. The scheme includes more than 30 categories with each category being defined by SAs that are a necessary condition for a chemical to covalently interact with proteins either at the thiol or primary amine sites. Definition of these alerts was justified by their interaction mechanisms with proteins, found in the literature, especially in the area of skin sensitization. This classification scheme is particularly relevant for skin sensitization and excess acute aquatic toxicity.

5. The OECD QSAR Toolbox is a stand-alone system intended to facilitate the formation of chemical categories and filling data gaps. The first version of the Toolbox released in March 2008 is already helpful to member countries and other stakeholders in forming categories and using existing data to fill data gaps. Phase 2 of the development of the Toolbox started in November 2008 and the aim is to ensure that the categories approach works uniformly for all discrete organic chemicals and for all regulatory endpoints. The 42nd Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology agreed that the main work item in the phase 2 project will be to gather and maintain additional categorisation methods [ENV/JM(2008)7].
Workshop

6. The expert consultation was held on 20 October 2010 hosted by the United Kingdom in Oxford. The agenda is outlined in Annex 1.

7. The consultation was attended by experts from Australia, Canada, Denmark, Germany, Japan, Netherlands, Poland, Spain, Sweden, the United Kingdom, the United States, the European Commission, BIAC, ICAPO. The list of the participants is attached to this document as Annex 2. The expert consultation was chaired by the OECD Secretariat.

Scope and Objectives

8. The stated scope of the expert consultation was to evaluate a revised set of SAs for estimating covalent protein binding.

9. The objectives of the expert consultation were to:

- get an overview of the revised organic chemistry-based mechanistic SAs for identifying protein binding chemicals, the literature on which they are based, and how they may be used in the OECD QSAR Toolbox;

- get a review from experts of the proposed SAs for protein binding;

- propose with the example of Michael addition reaction domain as a proof of concept how quantification of reactivity in the form of isoreactive groups may be used in the OECD QSAR Toolbox;

- get a recommendation from the experts on the proposed structural alerts and accompanying documentation and isoreactive groups, in particular if they should be implemented in the Toolbox.

10. The importance of alert-based expert systems (so-called profilers) in the Toolbox is to allow for the formation of toxicologically meaningful categories. Such a category means that all the chemicals falling within it can be assessed when only a few members are tested. This enables transparent and defensible categories to be formed. Version 1.1 of the Toolbox only contains a relatively small number of profilers many of which are incomplete. Incorporation of new and better profilers is seen as being essential to add new functionalities to the Toolbox. The better the profiler, the better and more precise the category. It is important to note that in the Toolbox profilers are not to be used to predict adverse effects. Rather, the profilers are used to group chemicals to allow for read-across using existing experimental results.

Preparation for the Expert Consultation

11. A scientific re-evaluation of known SAs for covalent protein binding was undertaken by Liverpool John Moores University (LJMU) in coordination with LMC as part of the development of a new profiler for the OECD QSAR Toolbox, financed by the European Chemicals Agency. The focus of this work was SAs for protein binding based on organic chemical reactions, especially ones that target thiol and primary amine groups. The work plan identified six issues:

   A. Identification of the scientific literature detailing SAs for protein binding.
   B. Analysis of the identified alerts and the rationalization of the associated mechanistic chemistry.
   C. Identification of mitigating factors that may alter protein binding.
D. Construction of clear and concise documentation related to each alert including name and pictorial representation of the alert, as well as sections detailing the mechanistic chemistry that leads the alert being able to covalently bind to proteins, any mitigating factors that should be considered as part of the alert, and references that support the mechanistic chemistry information.

E. Associated confidence in the suggested alerts.

F. Proposal of a method(s) for the implementation of protein binding-related reactive potency (glutathione RC50 for Michael addition) into the Toolbox workflow.

12. In preparations for the expert consultation LJMU prepared a consultation document entitled: “Re-evaluation of Structural Alerts for the Binding of Molecules to Protein and the Development of a Comprehensive Profiler of Alerts”, which is reported in Annex 3. The consultation document has a number of supplementary information, which is compiled in Annex 4. In addition, the LJMU presented an overview of the alerts, which is reported in Annex 5.

Preparatory Work by Experts

13 Expert reviewers were selected by the Secretariat. These were, Dr Yintak Woo from the US-EPA, Drs Rob Vanderbriel and Emiel Rorije from the Netherlands, and Dr Nora Aptula from Unilever and representing BIAC. The reviewers were provided with the consultation documents.

14. The reviewers were asked to comment on a series of queries drafted by the OECD Secretariat. These are as follows.

**Query 1.** Please comment on the completeness of literature reviewed. Please indicate any additional literature, which you feel would further clarify or support SAs for protein binding.

**Query 2.** Please comment on the adequacy and completeness of the SAs for forming categories based on mechanisms of protein binding.

**Query 3.** Please comment on the adequacy of the mitigating factors, affecting either toxicokinetics or toxicodynamics, which alter protein binding ability.

**Query 4.** Please comment on documentation associated with each alert. In particular is the rationalization complete yet easy to follow.

**Query 5.** Please comment on the associated confidence noted for each alert, especially for those alerts where you feel the confidence may be misstated.

**Query 6.** Please comment on the scientific rationale and clarity of the proposed implementation of reactive potency into the Toolbox workflow.

**Query 7.** Please make any further suggestions for improvements in presenting the SAs and their underlying rationale.

15. The review report of Dr. Woo is reported in Annex 6.

16. The review report of Drs. Vanderbriel and Rorije is reported in Annex 7.

17. The review report of Dr. Aptula is reported in Annex 8.
Proceedings of the Expert Consultation

18. The Expert Consultation was conducted as described in the agenda reported in Annex 1. Briefly, the LJMU presented an overview of the SAs, which was followed by a demonstration by LMC on how reactive potency may be integrated into the workflow of the Toolbox. These presentations were followed by the reviews of Drs. Woo, Vanderbriel and Rorije, and Aptula. LJMU then provided clarification and response to the reviews. The clarifications were followed by a general discussion by all participants.

Outcome of the Expert Consultation

19. Summary responses to the queries asked of the reviewers and agreed upon by the expert consultation are as follows:

1) Regarding the completeness of the literature search the meeting agreed with the reviewers that the contractors did a good job of identifying the majority of the currently available relevant literature related to covalent protein binding and their mechanistic interpretation, especially in the context of direct-acting reactions that target thiols and primary amines.

2) Regarding the adequacy and completeness of the SAs for forming categories based on mechanism of covalent protein binding the meeting agreed with the reviewers that in general the new SAs presented by the contractors were a significant improvement over the SAs listed in version 1.1 of the Toolbox. Noted exceptions were the missing alerts referred to by Drs Woo and Aptula. The meeting furthermore agreed with the suggestions from the contractor that these new alerts be added to their proposed series of SAs.

3) Regarding the adequacy of the mitigating factors, the meeting agreed with the reviewers that the covalent protein binding profiler with its basis on chemical mechanism should project binding in the broader or more generic sense. The meeting furthermore agreed that mitigating factors are important and their usage should be considered further (e.g. in a subcategorization profiler). However, this can only be done with regard to steric and electronic factors, as biological factors will be endpoint specific.

4) Regarding the documentation of the SAs the meeting agreed with the reviewers that the documentation as presented is easy to follow, and consistent with other profilers in the Toolbox, but that it may need to be improved for some SAs (e.g. adding pictorial representation of the mitigating factors). The meeting agreed that this information would be a welcome addition to the Toolbox. The meeting furthermore agreed with the suggestions from the contractor to review the documentation for each SA and update the documentation as suggested.

5) Regarding the confidence noted for each SA the meeting agreed with the reviewers that the confidence is related to our knowledge of organic chemistry and not reported biological effects and this needs to be stated.

6) Regarding the proposed methods of incorporating reactive potency for protein binding reactions into the Toolbox the meeting agreed with the reviewers that the two examples demonstrated for the Michael Addition mechanism of binding was an improvement in the Toolbox and a start at moving from qualitative read-across to quantitative read-across. It was further noted that the availability of additional reactivity potency data was the major limitation to universal implementation of the methods.

7) Regarding further suggestions for improvements the meeting agreed with the reviewers that there is a need for additional chemical reactivity data to if possible, reflect the total chemical space of each SA. The meeting further agreed with the reviewers that there is a need for an evaluation of the performance of the SAs, especially in the context of skin sensitization and excess acute aquatic
toxicity. The contractors have agreed to conduct and report on such an evaluation. Additionally, the meeting agreed that a strategy of using the different chemical reactivity-based profilers in the Toolbox needs to be developed. The OECD Secretariat agreed to coordinate development of such a strategy.

20. The Conclusions and Recommendations from the expert consultation are as follows:

1) While not all covalent protein binding SAs have been identified the new proposed chemical mechanism-based protein binding profiler is an improvement over the profiler in Toolbox version 1.1.

2) The OECD is encouraged to expand this work to include the recommendations of the reviewers.

3) The SAs are useful to build chemical categories.

4) The reactivity potency data for Michael acceptors is adequate to be used as a descriptor in trend analysis and in the development of a subcategorisation profiler.

5) It is recommended that the SAs be implemented in the OECD QSAR Toolbox.

6) It was recommended that the Michael acceptor sub-profiler be implemented in the OECD QSAR Toolbox.

7) Further recommendations include:
   - after implementation the contractors should evaluate the performance of the alerts with regard to skin sensitization and excess acute aquatic toxicity
   - other reaction-specific reactivity databases should be identified.
   - additional guidance is necessary on how to improve the confidence in the read-across approach for biological endpoints, using the reactivity-based profilers and the relevant experimental data in the Toolbox databases.
ANNEX 1:
AGENDA OF THE EXPERT CONSULTATION ON SCIENTIFIC AND REGULATORY EVALUATION OF ORGANIC CHEMISTRY MECHANISM-BASED STRUCTURAL ALERTS FOR THE IDENTIFICATION OF PROTEIN BINDING CHEMICALS

Randolph Hotel, Oxford, UK

20 October 2010

The meeting starts at 08h30 and closes at 17h30.

PRELIMINARY DRAFT AGENDA (11 August 2010)
The meeting starts at 08h30 and closes at 17h30.

08h30 1 Opening and the adoption of the agenda (10min)
The meeting will be opened by the OECD Secretariat. The Secretariat will explain the purpose of the Expert Consultation and housekeeping items. The Secretariat will also confirm that the participants have all meeting documents. The meeting participants will briefly introduce themselves to the meeting (Tour de Table). The Consultation participants will be asked to approve the agenda, and discuss changes in meeting papers and scheduling of the agenda items if necessary.

08h40 2 Background information (10min)
The Secretariat will explain the history and rational for the project leading to this Expert Consultation. The Consultation participants will be invited to take note of this activity.

08h50 2 Overview of the revised protein-binding profiler (75min)
The OECD Secretariat will ask Drs Mark Cronin and Steve Enoch of Liverpool John Moores University to present a detailed overview of the Organic Chemistry Mechanism-Based Structural Alerts for the Identification of Protein Binding Chemicals and Dr. Ovanes Mekenyan of the Laboratory of Mathematical Chemisity to present the proposal for Incorporation of Quantification of in chemico Reactivity in to the Toolbox. The consultation participants will be invited to take note of this activity and ask questions as appropriate.

10h00 Coffee Break (30min)

10h35 3 First review of the revised protein binding profiler (30min)
The Secretariat will ask the first reviewer to present their review of the alerts and the proposed use of reactivity described by Drs Cronin, Enoch and Mekenyan. The consultation participants will be invited to take note of this review and ask questions as appropriate.

11h05 4 Second review of revised protein binding profiler (30min)
The Secretariat will ask the second reviewer to present their review of the alerts and the proposed use of reactivity described by Drs Cronin, Enoch and Mekenyan. The consultation participants will be invited to take note of this review and ask questions as appropriate.

11h35 **Third review of the of the revised protein binding profiler (30min)**
The Secretariat will ask the third reviewer to present their review of the alerts and the proposed use of reactivity described by Drs Cronin, Enoch and Mekenyan. The consultation participants will be invited to take note of this review and ask questions as appropriate.

12h05 **Lunch Break (85min)**

13h30 **Clarification and Responses to the Review’s Comments (30min)**
The Secretariat will ask Drs Cronin, Enoch and Mekenyan to provide clarification of the alerts and the proposed use of reactive, and respond to the reviews. The consultation participants will be invited to take note of this activity.

14h00 **Discussion of the Organic Chemistry Mechanism-Based Structural Alerts for the Identification of Protein binding Chemicals (60min)**
The OECD Secretariat will invite the participants to discuss the alerts and comment as appropriate.

15h00 **Discussion of the proposal to Incorporation of Quantification of in chemico Reactivity in to the Toolbox (45min)**
The OECD Secretariat will invite the participants to discuss the proposed use of isoreactive groups and comment as appropriate.

15h45 **Coffee Break (30min)**

16h15 **Initial Finding of the Consultation (45min)**
The OECD Secretariat will present the initial findings of the consultation. The consultation participants will be invited to provide comments as appropriate.

17h00 **Any Other Issues Relating to the Consultation (30min)**
The Secretariat will consider any other issues related to the consultation raised by the participants.

17h30 **Meeting adjourns**
ANNEX 2: LIST OF PARTICIPANTS

**Australia/Australie**

Julija FILIPOVSKA  
Senior Regulatory Scientist  
NICNAS

**Canada**

Mark BONNELL  
Senior Science Advisor  
Environment Canada

Sunil KULKARNI  
Scientific Evaluator, Risk Assessment  
Risk Assessment Bureau  
Health Canada

**Czech Republic/République Tchèque**

Marian RUCKI  
Expert  
National Institute of Public Health

Milon TICHY  
Expert  
Centre of Occupational Health  
National Institute of Public Health

**Denmark/Danemark**

Jay NIELASMA  
Department of Toxicology and Risk Assessment  
Danish Institute for Food and Veterinary Research

**Germany/Allemagne**

Frauke STOCK  
Chemicals section  
Federal Environment Agency (Umweltbundesamt)

Monika BATKE  
Expert  
Chemical Risk Assessment  
Fraunhofer Institute of Toxicology and Experimental Medicine (ITEM), Chemical Risk Assessment
Italy/Italie

Romualdo BENIGNI
Experimental and Computational Carcinogenesis

Japan/Japon

Yumi AKAHORI
Research Scientist
Chemicals Assessment and Research Center
Chemicals Evaluation and Research Institute, Japan (CERI)

Ayako FURUHAMA
NIES Special Researcher
Research Center for Environmental Risk
National Institute for Environmental Studies

Yuki SAKURATANI
Chief
Safety Assessment Division, Chemical Management Center
National Institute of Technology and Evaluation

Jun YAMADA
Technical Adviser
Safety Assessment Division
National Institute of Technology and Evaluation

Netherlands/Pays-Bas

Emiel RORIJE
SEC
RIVM-National Institute for Public Health and the Environment

Rob VANDEBRIEL
Laboratory for Health Protection Research
National Institute for Public Health & the Environment

Poland/Pologne

Jacek CIESLA
Expert of the Ministry of Health, Chief Specialist, IT Manager
Dept for Risk Assessment
Bureau for Chemical Substances and Preparations
Spain/Espagne

Juan José IZQUIERDO
Conseiller
Environment Department
INIA

Sweden/Suède

Jörgen HENRIKSSON
Senior Scientific Officer
Swedish Chemicals Agency (KEMI)

United Kingdom/Royaume-Uni

Ian DOYLE
Senior Chemicals Assessment Scientist
Chemicals Assessment Unit
Environment Agency

United States/États-Unis

Kelly BEAN
US Environmental Protection Agency

Yintak WOO
Senior Toxicologist
US Environmental Protection Agency

EU/UE

Tatiana NETZEVA
Senior Scientific Officer
Unit C2 - Registration
European Chemicals Agency -ECHA

Doris HIRMANN
Scientific Officer
Unit for Scientific IT Tools
European Chemicals Agency

Business and Industry Advisory Committee (BIAC)/Comité consultatif économique et industriel (BIAC)

Katsutoshi KINOSHITA
MITSUI CHEMICAL, INC.
Shinji KOTACHI  
Chemical Control Department  
Japan Chemical Industry Association

Grace PATLEWICZ  
DuPont

International Council on Animal Protection in OECD Programmes

Chad SANDUSKY  
Director of Toxicology and Research  
Physicians Committee for Responsible Medicine

Alexander TROPSHA  
Director  
Laboratory for Molecular Modeling

OECD/OCDE

Mark CRONIN  
School of Pharmacy and Chemistry  
Liverpool John Moores University

Bob DIDERICH  
Principal Administrator  
ENV/EHS  
OECD

Sabcho DIMITROV  
Laboratory of Mathematical Chemistry (LMC)  
University "Prof. Assen Zlatarov"

Steven ENOCH  
Post-Doctoral Research Fellow  
Liverpool John Moores University

Philip JUDSON  
Scientific Advisor  
Lhasa Limited

Ovanes MEKENYAN  
Laboratory of Mathematical Chemistry (LMC),  
Bourgas University

Michihiro OI  
Administrator, (Q)SARs and Exposure Assessment  
ENV/EHS  
OECD
Todor PAVLOV  
Laboratory of Mathematical Chemistry (LMC), Bourgas University

Terry SCHULTZ  
Administrator, (Q)SARs  
ENV/EHS  
OECD

Dave WATSON  
Chief Executive Officer  
Lhasa Limited
ANNEX 3: A REPORT ON RE-EVALUATION OF STRUCTURAL ALERTS FOR THE BINDING OF MOLECULES TO PROTEINS AND THE DEVELOPMENT OF A COMPREHENSIVE PROFILER OF ALERTS

Steven J. Enoch and Mark T. D. Cronin

School of Pharmacy and Chemistry
Byrom Street
Liverpool John Moores University
Liverpool
L3 3AF
England
s.j.enoch@ljmu.ac.uk, m.t.cronin@ljmu.ac.uk
Contents

1.0 Introduction

2.0 Aims of the investigation

3.0 Existing approaches for modelling endpoints where protein binding is the key molecular initiating event

4.0 Electrophilic reaction chemistry

5.0 Structural alert literature sources

5.1 Skin sensitisation – OECD Toolbox compilation

5.2 Skin sensitisation – Roberts and Aptula compilation

5.3 Skin sensitisation – Zinke et al compilation

5.4 Skin sensitisation – Gerner et al compilation

5.5 Respiratory sensitisation – Enoch et al compilation

5.6 Skin irritation and corrosion – Hulzebos et al compilation

5.7 Excess acute fish toxicity – Verhaar et al compilation

5.8 Excess acute fish toxicity – Hermens compilation

5.9 Excess acute fish toxicity – von der Ohe et al compilation

5.10 Excess acute fish toxicity – Nendza and Muller compilation

5.11 In chemico data from glutathione depletion assays – Schultz et al compilation

6.0 Mechanistic alerts

6.1 Development of structural alerts for The Updated Profiler

6.2 Mechanistic domains and alerts

6.3 Mitigating factors

6.4 The meta data associated with the structural alerts

7.0 Potency sub-categorisation

7.1 Skin sensitisation: cinnamic aldehyde
7.2 Acute aquatic toxicity: 4-methyl-2-pentenal

8.0 Conclusions

9.0 Supplementary material

10.0 References
1.0 Introduction

A number of toxicological endpoints relating to human health have been suggested to be related to the ability of a chemical to form covalent adducts with proteins. The major endpoints include: skin sensitisation, respiratory sensitisation, skin and eye irritation, and hepatotoxicity (1-5). In addition, covalent binding to a protein can result in a chemical showing excess acute toxicity above narcosis in both aquatic and terrestrial species. This means the chemical is more toxic than would be predicted from baseline toxicity (non-polar narcotic chemicals form the so-called baseline level of toxicity in these species, such chemicals do not bind directly to proteins) (2, 6, 7). The ability of a chemical to form a protein-adduct has also been suggested to be play a role in other endpoints such as carcinogenesis (via chromosomal aberration (8)) and reproductive toxicity (9). Through the formation of a covalent protein-chemical adduct, these endpoints are considered to share a common molecular initiating event (10). However, despite the common initiating event, the resulting biological pathway that causes toxicity is different (although in the case of skin and respiratory sensitisation the biological responses are closely related (11)).

Humans come into contact with a wide range of industrial chemicals that are capable of forming covalent protein adducts through deliberate, accidental and occupational exposures. In addition, these chemicals are frequently disposed in the environment. There is therefore a desire to assess the toxic potential of such chemicals to both humans and environmental species. There are a number of methods to determine the toxicity of such chemicals spanning a number endpoints, both in vitro and in vivo. However, there is an increasing appreciation of the costs (financial and animal use) associated with such testing, particularly with regard to their regulatory assessment (12-14). Thus, alternatives are being sought to rapidly screen compounds for toxicity (15). These techniques are also extremely useful in the efficient development of new compounds and to determine the potential effects of existing compounds.

Key techniques to develop non-test methods to assess toxicological endpoints related to covalent protein adduct formation include the in silico approaches (16). Such methods attempt to relate the chemical and / or structural properties of a molecule to its activity. The techniques include the development of (quantitative) structure-activity relationships ((Q)SARs) and the formation of categories to facilitate read-across (17-20). The possibility of relating chemical structure to (for example) skin sensitisation activity has been explored for several decades (21). As well as developing “traditional” QSARs (i.e. statistical techniques relating activity to molecular descriptors), there has been a keen interest in developing so-called “structural alerts”. These are essentially molecular fragments that are known, or thought, to be related to the toxic effect in question. The philosophy of such an approach is clear and elegant: it allows structural fragments associated with endpoints such as (for example) skin sensitisation to be defined and related to a mechanism of action. Structural alerts have been used to identify compounds that may potentially elicit skin sensitisation and acute toxicity to aquatic species elevated above narcosis (2, 22-24). An added bonus has been their more recent use to develop chemical categories or groupings which may assist in the filling of data gaps for regulatory purposes. Such collections, or compilations, of structural alerts can be thought of as profilers for this property and, if associated with a defined chemical grouping, are easily coded into a computational format for further application. As yet there have been no global approaches to unify these alerts on the basis of mechanistic reaction chemistry.
2.0 Aims of the investigation

This report details the recent efforts to update the protein binding profiler within OECD (Q)SAR Application Toolbox V1.1.01 (referred to throughout as „The Toolbox”). The alerts for the new direct-acting covalent protein binding profiler are referred to throughout as „The Updated Profiler”. The aims of this analysis can be summarised as follows:

- To review the current scientific knowledge relating to structural alerts for toxicological endpoints for which covalent protein binding is the key molecular initiating event.

- To review structural alerts related to direct acting covalent protein binding. Alerts for pre / pro (chemicals requiring abiotic and metabolic activation respectively) electrophiles were not included; this being due to the metabolic simulators within The Toolbox being designed to deal with the conversion of such chemicals into direct acting electrophiles.

- To map the existing structural alerts in terms of their relationships with mechanistic organic chemistry (i.e. identify alerts from the published compilations related to direct acting covalent protein binding).

- To undertake an analysis of the underlying mechanistic chemistry for each alert.

- To compile a new and complete (at the time of writing) set of direct acting covalent protein binding structural alerts.

- To develop a „proof of concept” approach for the use of in chemico data for the development of potency sub-categories within the Michael addition mechanistic domain.

It is important to state that no attempt has been made to evaluate or validate the newly suggested covalent binding protein alert compilation against toxicological data (although the mechanisms associated with the alerts have been reported in at least one peer-reviewed literature source to be capable of directly covalently binding to proteins).
3.0 Existing approaches for modelling endpoints where protein binding is the key molecular initiating event

A number of toxicological endpoints require the formation of a covalent protein-chemical adduct as the molecular initiating event. The mechanisms of these endpoints have been studied extensively and a number of predictive models and approaches have been developed. The resulting models have ranged from statistically-derived global modelling approaches, through to so-called expert systems and the definition of mechanistically derived structural alerts. In addition, a number of local (quantitative) structure-activity relationship ((Q)SAR) models have been published. For example, for skin sensitisation the local lymph node assay potency for congeneric series of chemicals has been modelled (for a review of global and local modelling approaches for skin sensitisation see (20)).

The varying approaches to modelling toxicity can be considered as a continuous spectrum with increasing mechanistic interpretability and transparency (Figure 3.1). Approaches can be considered according to their mechanistic relevance i.e. can the model be interpreted with regard to the mechanism of action to which it relates; transparency i.e. can the user of the structural alert quickly and rationally obtain the mechanistic relevance? Whilst these are by no means the only criteria for evaluating an in silico model, they are key to the development of meaningful structural alerts and their compilation into a profiler suitable for category formation. The following very general conclusions can be drawn from the analysis of the modelling approaches (a full evaluation of individual approaches according the OECD Principles for the Validation of (Q)SARs is outside the scope of this report).

- Statistically derived global models are usually the least mechanistically interpretable. However, in some cases, they can be used to derive new mechanistic information about the causes of toxicity through post rationalisation of the model.

- Mechanistic category formation is fully transparent and mechanistically relevant. It is based on current scientific knowledge regarding the ability of chemicals to cause toxicity. It is within a mechanistic category that a local QSAR can be utilised to predict a biological endpoint of interest, for example skin sensitisation or excess acute toxicity.

- Knowledge based expert systems fall somewhere in-between these two extremes with software being constructed using either a mixture of mechanistic alerts and statistical models.

![Figure 3.1: Schematic showing the relationship between modelling approach and mechanistic interpretability and transparency](image-url)
A number of existing approaches for modelling skin sensitisation, respiratory sensitisation and excess acute aquatic toxicity and other relevant endpoints are summarised in Table 3.1. The advantages and disadvantages of the three methodologies (global statistical models, expert systems, and mechanistic categories) are summarised in Table 3.2.

<table>
<thead>
<tr>
<th>Software</th>
<th>Modelling approach</th>
<th>Further information</th>
<th>Example endpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAESAR</td>
<td>Global statistical</td>
<td><a href="http://www.caesar-project.eu/software/index.htm">http://www.caesar-project.eu/software/index.htm</a></td>
<td>Skin sensitisation</td>
</tr>
<tr>
<td>Topkat</td>
<td>Global statistical</td>
<td><a href="http://accelrys.com/products/discovery-studio/toxicology/">http://accelrys.com/products/discovery-studio/toxicology/</a></td>
<td>Skin sensitisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aquatic toxicity</td>
</tr>
<tr>
<td>MultiCase</td>
<td>Global statistical</td>
<td><a href="http://www.multicase.com/">http://www.multicase.com/</a></td>
<td>Skin sensitisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Respiratory sensitisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aquatic toxicity</td>
</tr>
<tr>
<td>Various QSARs</td>
<td>Local statistical</td>
<td>For a review of such models for skin sensitisation and acute aquatic toxicity see references (6, 20)</td>
<td>Any for which data are available</td>
</tr>
<tr>
<td>Hazard Expert Pro</td>
<td>Expert system</td>
<td><a href="http://www.compudrug.com/">http://www.compudrug.com/</a></td>
<td>Irritation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sensitisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Immunotoxicity</td>
</tr>
<tr>
<td>DEREK for Windows</td>
<td>Expert system</td>
<td><a href="http://www.lhasalimited.org">http://www.lhasalimited.org</a></td>
<td>Skin sensitisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Respiratory irritation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hepatotoxicity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Skin / eye irritation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chromosomal aberration</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aquatic toxicity</td>
</tr>
<tr>
<td>OECD QSAR Application Toolbox</td>
<td>Mechanistic categories / local QSARs</td>
<td><a href="http://www.oecd.org/document/23/0,3343,en_2649_34379_339570_15_1_1_1_1_1_00.html">http://www.oecd.org/document/23/0,3343,en_2649_34379_339570_15_1_1_1_1_1_00.html</a></td>
<td>Any for which data are available</td>
</tr>
</tbody>
</table>

Table 3.1: Non-exhaustive summary of previous efforts in the modelling of skin sensitisation, respiratory sensitisation and excess aquatic toxicity
<table>
<thead>
<tr>
<th>Modelling approach</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global statistical</td>
<td>Useful when mechanism of action is unknown.</td>
<td>Can be difficult to interpret the models.</td>
</tr>
<tr>
<td></td>
<td>Good for screening large numbers of chemicals.</td>
<td>Models are frequently not transparent to the end-user.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Applicability domain can be difficult to define.</td>
</tr>
<tr>
<td>Knowledge based expert system</td>
<td>Often mechanistically based.</td>
<td>Applicability domain difficult to define.</td>
</tr>
<tr>
<td></td>
<td>Good for screening large numbers of chemicals.</td>
<td>Not always transparent or mechanistically based.</td>
</tr>
<tr>
<td>Mechanistic alerts</td>
<td>Derived from knowledge of the underlying mechanism of action.</td>
<td>Each alert has a limited (although well defined) applicability domain.</td>
</tr>
<tr>
<td></td>
<td>Useful for defining chemical categories.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interpretable and transparent.</td>
<td></td>
</tr>
<tr>
<td>Local QSARs</td>
<td>Mostly utilise mechanistic knowledge.</td>
<td>Limited, although well defined applicability domain.</td>
</tr>
<tr>
<td></td>
<td>Interpretable and transparent.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Well defined applicability domain (which can be derived from a structural alert)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Advantages and disadvantages of the differing approaches to modelling endpoints where covalent protein binding is the molecular initiating event.

In contrast to the *in silico* models described in Table 3.1 that predict an individual biological effect, The Toolbox uses a category approach that aims to group chemicals around so-called molecular initiating events. One such event, applicable to skin sensitisation, respiratory sensitisation and excess aquatic toxicity is covalent protein binding. The presence of a common molecular initiating event within a series of chemicals (often highlighted by a common structural feature) allows the end-user of The Toolbox to develop mechanistic categories within which (Q)SAR and trend analysis can be performed. Such analyses are aimed at allowing toxicological data gaps to be filled for a chemical (or a series of chemicals), thus reducing animal usage to comply with, for example, the REACH legislation (25). The Toolbox contains a number of profilers that enable chemicals to be grouped into categories using mechanistic information (other methods such as structural similarity are also encoded). The current covalent protein binding profiler within The Toolbox contains structural alerts derived from the TIMES software (26).
4.0 Electrophilic reaction chemistry

It is important to realise that one of the biological mechanisms of action that all of the approaches described in Table 3.1 are trying to predict, implicitly or explicitly is the ability of a chemical to bind covalently to a protein. If one considers the mechanistically based models specifically, then an understanding of how the alerts relate to the underlying chemistry is extremely important. It is has been known for several decades that in order for a chemical to exhibit toxicity across a number of endpoints (for example, skin sensitisation) then one of the key mechanisms is that it must be (or must be metabolised to) an electrophile (16).

Electrophilic chemistry is well understood and defined in the simplest terms as involving an electron rich nucleophilic centre on a protein chain attacking an electron deficient electrophilic centre on an exogenous chemical. This results in the formation of a new chemical (covalent) bond, and in the case of protein-chemical interactions, an adduct. This adduct formation is considered to be an initiation process for a range toxicity endpoints as noted above.

As an example of the role of electrophilic chemistry as a molecular initiating event, acrolein (Figure 4.1) is considered as it is a well-known skin sensitiser. The underlying chemistry of this event can be understood by considering the chemical structure and likely reactivity of acrolein. In acrolein, the terminal carbon atom (the so-called β-carbon) of the alkene of acrolein is electron deficient due to the presence of the carbonyl group (which contains an electronegative oxygen atom that „pulls” electron density towards itself). This makes this β-carbon electrophilic and thus susceptible to nucleophilic attack (which means in this case attack by a nucleophile such as a protein) and hence capable of forming a covalent bond. The fragment (based on the mechanism) required to make the β-carbon electrophilic as a structural alert can thus be defined. The chemistry and related structure is defined in Figure 4.2.

![Figure 4.1: Schematic of a nucleophile attacking the electrophile acrolein resulting in a protein-chemical adduct (Nu = biological nucleophile)](image)

Whilst an analysis of the structure of reactive chemicals allows for the definition of electrophilic fragments, little information is known about the exact source or nature of the biological nucleophile. For the majority of toxicological endpoints for which protein binding is important, the biological nucleophile is assumed to be a cysteine, lysine or serine. It is likely that, in reality, reactive electrophilic chemicals react with all three nucleophiles, with the exact extent of adduct formation to each being dependent on the relative hardness / softness of the electrophile and nucleophile (the so-
called hard-soft acid-base theory). The inability to identify the exact biological nucleophile is less important than the information regarding the electrophile, as hard-soft acid base theory states that a soft electrophile will have a relative preference for a soft nucleophile and that a hard electrophile will have a relative preference for a hard nucleophile. Thus, for a series of chemicals assigned to the same category within a mechanistic domain, the relative rates of reactivity between each electrophile and (unknown) nucleophile will remain the same. For example, if chemical A is more reactive than chemical B when the nucleophile is cysteine then it will also be more reactive then chemical B when the nucleophile is lysine or serine (assuming both chemicals are within the same mechanistic domain). However, it is important to realise that the actual rate of reactivity of a chemical to differing nucleophiles does vary (i.e. depending how soft the chemical is will determine how fast it reacts with cysteine, lysine or serine). Since the formation of a protein-chemical adduct is the molecular initiating event for a range of toxic endpoints; information regarding the rate of formation for such adducts (obtained either experimentally or by calculations) is useful in establishing (quantitative) structure-activity relationships for the prediction of toxicity within mechanistic domains. This is the basis for grouping chemicals into mechanistic categories to allow for read across and local (Q)SAR developments.
5.0 Literature sources for structural alert for protein binding

A number of publications were identified as sources of structural alerts for covalent protein binding. These publications were studies relating to several toxicological endpoints for which covalent protein binding is the key molecular initiating event, these being: skin and respiratory sensitisation, skin irritation, skin corrosion and excess acute aquatic toxicity. In addition, sources of in chemico data were also retrieved, these data relate to experimental determination using glutathione as a nucleophile (a recent review details the wide range of available in chemico data (41)). In addition, evidence for the ability of direct acting covalent DNA binding structural alerts (compiled previously (27)) to also bind covalently to proteins was sought. The set of structural alerts for covalent protein binding in v1.1.01 of The Toolbox are those detailed by Dimitrov et al (26). The literature relating to protein binding structural alerts identified:

5.1 Skin sensitisation – OECD Toolbox compilation

The current protein binding profiler within v1.1.01 of The Toolbox consists of 38 structural alerts. These alerts were developed from an analysis of skin sensitisation data by a consortium of industrial partners and academics. This analysis led to the development of the TIMES-SS system from which the structural alerts within the profiler were taken (26, 28).

5.2 Skin sensitisation – Roberts and Aptula compilation

A number of related studies have illustrated the ability of organic reaction chemistry to develop clear, mechanistically-based, categories for the ability of a chemical to cause skin sensitisation (16, 29-31). The mechanistic chemistry within these studies was developed from an analysis of historical local lymph node assay data and builds on a number of earlier mechanistic studies by Ashby, Payne and Barrett (23, 32, 33). The mechanistic chemistry encapsulated within these studies have been developed into a compilation of structural alerts (34). Analysis of this compilation shows it to contain 28 direct acting covalent binding structural alerts covering five mechanistic domains (acylation, Michael addition, Schiff base formation, SNAr, and SN2).

5.3 Skin sensitisation – Zinke et al compilation

A recent literature review detailed an analysis of the Derek for Windows skin sensitisation rulebase (22, 24). This study utilised an external dataset to evaluate the alerting groups present in the Derek for Windows rulebase, with the authors making suggestions and comments regarding the utility of each of the structural alerts. Inspection of the data within this study revealed 29 direct acting covalent protein binding alerts (24).

5.4 Skin sensitisation – Gerner et al compilation

A recent analysis of skin sensitisation data resulted in the development of 15 alerts for electrophiles (22). The structural alerts were developed from an analysis of proprietary data contained within a regulatory database. All but one of the resulting alerts presented in the study were accompanied by a mechanistic rationale.
5.5 Respiratory sensitisation - Enoch et al compilation

A recent analysis of human respiratory sensitisation data allowed 12 structural alerts related to direct acting covalent protein binding to be developed (3, 35). In contrast to skin sensitisation data (for which sulphur and nitrogen containing amino acid side chains can act as the biological nucleophile) the alerts presented within this work are likely to be related to covalent binding nitrogen containing side chains only. This difference is due to the likelihood of thiol groups to be oxidised to disulphide bridges within the highly oxidising environment of the lung.

5.6 Skin irritation and corrosion – Hulzebos et al compilation

Hulzebos et al presented an analysis of structural features and physico-chemical properties related to a chemical’s ability to cause skin irritation or corrosion (5). In this analysis a dataset of 99 corrosive and 118 irritating chemicals was investigated. A number of potential mechanisms of action were identified, including the ability of a chemical to bind covalently to skin proteins. 41 structural alerting groups were identified, analysis of which revealed 22 to be related to direct acting covalent protein binding.

5.7 Excess acute fish toxicity – Verhaar et al compilation

Verhaar et al published a very well utilised decision tree approach aimed at classifying chemicals into modes of action for acute aquatic fish toxicity (2). Twenty two electrophilic fragments were identified as part of the rulebase for the identification of „Class 3 type compounds – unspecific reactivity”. Analysis of these fragments reveals 21 of them to have a clear mechanism related to covalent protein binding.

5.8 Excess acute fish toxicity – Hermens compilation

Hermens investigated the importance of electrophiles in acute aquatic fish toxicity (36). Thirty six direct acting covalent protein binding structural alerts were identified within this study.

5.9 Excess acute fish toxicity – von der Ohe et al compilation

von der Ohe reported the findings of a study in which a classification model aimed at discriminating chemicals exhibiting excess aquatic toxicity was developed using structural alerts (37). The study identified 13 structural alerts, of which seven are related to a chemical’s ability to react as direct covalent electrophile with proteins.

5.10 Excess acute fish toxicity – Nendza and Muller compilation

Nendza and Muller described a QSAR study in which a number of models were used to discriminate aquatic toxicants by mode of action using sub-structure indicators (38). The study identified 40
molecular sub-structures (structural alerts) that were useful in the modelling process, of these 29 are capable of direct acting covalent protein binding.

5.11 *In chemico* data from glutathione depletion assays – Schultz *et al* compilation

Schultz *et al* have described at least two studies highlighting the use of an *in chemico* glutathione depletion assay to measure chemical reactivity and help define mechanistic domains (19, 39). These data have been used to help define reactivity sub-categories within the Michael acceptor mechanistic domain. This work is a proof-of-concept approach to such sub-categorisation.

6.0 Mechanistic alerts

Given the range of alert compilations available in the literature, it is clear that a method for the development of an updated set of direct acting covalent protein binding alerts is required. To achieve this, this study investigated the mechanistic domains suggested by work of Aptula and Roberts (16) as the central premise for the development of the new profiler. These domains being defined as:

- Acylation
- Aromatic nucleophilic substitution (SNAr)
- Bimolecular nucleophilic substitution (SN2)
- Michael addition
- Schiff base formation
- Unimolecular nucleophilic substitution (SN1)

In this approach chemicals are grouped into common mechanistic domains which encompass a number of so-called traditional structural alerts. For example, one would group all polarised alkenes into the Michael addition domain, rather than having individual alerting groups for acroleins, acrylates, methacrylates etc. This can be thought of a developing mechanism based structural alerts (or mechanistic alerts) rather than structural alerts based on a congeneric series of chemicals. Importantly, both approaches ensure that the resulting structural alerting groups are defined in terms of a common mechanism of action, with the former being more use in a category building system such as The Toolbox and the latter more use in a predictive system such as Derek for Windows.

6.1 Development of structural alerts for The Updated Profiler

The mechanistic domains reported by Roberts and co-workers were used as the central basis for the development of The Updated Profiler (16, 29-31). In addition, the structural alerts identified for these mechanistic domains (34) were further grouped into sub-domains (referred to throughout as a mechanistic alert). Mechanistic alerts were created on the basis of a common reactive centre (the site of attack by a biological nucleophile) being activated by a number of substituents. For example, an alkene acting as a Michael acceptor due to the influence of a polarising moiety formed a mechanistic alert. A separate mechanistic alert was formed for alkynes polarised by the same set of substituents, despite the fact that both sets of chemicals act as Michael acceptors. This is due to the fact that chemicals within these two mechanistic alerts have differing reactive centres (an alkene carbon atom
versus an alkyne carbon atom). A number of structural alerting groups make up each mechanistic alert, for example the polarised alkene mechanistic alert consists of 14 structural alerts. This approach allows one to profile at either the mechanistic alert level or the structural alert level. The schematic shown in Figure 6.1 details the levels within this approach.

Figure 6.1: Levels of information within mechanistic domains

Given this approach, the development of an updated set of structural alerts for direct acting covalent protein binding was undertaken using the following procedure for each alert (for all of the alerts within the literature compilations outlined in Section 5.0):

1. The mechanistic chemistry of the structural alert was investigated and assigned to one of six mechanistic domains, these being: acylation, Michael addition, Schiff base formation, SNAr, SN1, and SN2.

2. An attempt was then made to assign the structural alert to an existing mechanistic alert within the assigned mechanistic domain. If the structural alert could be assigned to an existing mechanistic alert, then an inspection of the structural features of the structural alert was undertaken to investigate whether it expanded the existing chemistry space. If the structural alert added additional structural features to the mechanistic alert then it was included. The structural alert was able to add structural information to an existing mechanistic alert in one of two ways:
   a. Additional structural information about an existing structural alert within the mechanistic alert.
   b. The definition of a completely new structural alert within the mechanistic alert.

3. If the structural alert could not be assigned to an existing mechanistic alert, then a new mechanistic alert within the assigned mechanistic domain was created. The structural alert was then assigned to the newly created mechanistic alert.

6.2 Mechanistic domains and alerts

The analysis detailed above resulted in the development of a set of mechanistic domains sub-divided into mechanistic alerts. Each mechanistic alert is made up of a group of structural alerts related by a common reaction site. The results of this analysis can be summarised as shown in Table 6.1. The mechanistic alerts and the associated structural alerts can be further detailed as shown in Table 6.2. It
is worth noting the absence of any mechanistic alerts for the Sn1 domain. This is due to The Updated Profiler only covering direct acting electrophiles. In general chemicals acting via an Sn1 mechanism need to be metabolically activated to a reactive electrophilic fragment (for example a carbenium ion). In addition, the majority of Sn1 reactive chemicals also react via the equivalent Sn2 mechanism and thus are covered by this domain in The Updated Profiler.

<table>
<thead>
<tr>
<th>Mechanistic Domain</th>
<th>Number of Mechanistic Alerts</th>
<th>Number of Structural Alerts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acylation</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Michael addition</td>
<td>4</td>
<td>31</td>
</tr>
<tr>
<td>Schiff base formation</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>SN2</td>
<td>8</td>
<td>39</td>
</tr>
<tr>
<td>SNAr</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 6.1: Summary of the mechanistic domains, mechanistic alerts and structural alerts defined as a result of the analysis detailed in Section 6.1 for The Updated Profiler
<table>
<thead>
<tr>
<th>Mechanistic Domain</th>
<th>Mechanistic Alert</th>
<th>Structural Alert</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acylation</td>
<td>Direct acylation</td>
<td>Acyl halides&lt;br&gt;Acetates&lt;br&gt;Anhydrides&lt;br&gt;Azlactones&lt;br&gt;Sulphonyl halides&lt;br&gt;Phosphonic acid halides&lt;br&gt;Dialkyl carbamoylhalides</td>
</tr>
<tr>
<td></td>
<td>Ring opening acylation</td>
<td>β-Lactones&lt;br&gt;Cyclopropenones</td>
</tr>
<tr>
<td></td>
<td>Isocyanates and related chemicals</td>
<td>Thiocyanates&lt;br&gt;Isocyanates&lt;br&gt;Isothiocyanates&lt;br&gt;Dithiocarbonimidic acid esters&lt;br&gt;Carbodiimides&lt;br&gt;Ketenes</td>
</tr>
<tr>
<td>Michael addition</td>
<td>Polarised alkenes and related chemicals</td>
<td>Polarised alkene - aldehydes&lt;br&gt;Polarised alkene - ketones&lt;br&gt;Polarised alkene - esters&lt;br&gt;Polarised alkene - amides&lt;br&gt;Polarised alkene - nitros&lt;br&gt;Polarised alkene - cyano&lt;br&gt;Polarised alkene - sulfonate&lt;br&gt;Polarised alkene - sulfone&lt;br&gt;Polarised alkene - sulfinyl&lt;br&gt;Polarised alkene - pyridines&lt;br&gt;Polarised alkene - pyrazines&lt;br&gt;Polarised alkene - pyrimidines&lt;br&gt;Polarised alkene - triazines&lt;br&gt;Azocarbonamides</td>
</tr>
<tr>
<td></td>
<td>Polarised alkynes</td>
<td>Polarised alkyne - aldehydes&lt;br&gt;Polarised alkyne - ketones&lt;br&gt;Polarised alkyne - esters&lt;br&gt;Polarised alkyne - amides&lt;br&gt;Polarised alkyne - nitros&lt;br&gt;Polarised alkyne - cyano&lt;br&gt;Polarised alkyne - sulfonate&lt;br&gt;Polarised alkyne - sulfone&lt;br&gt;Polarised alkyne - sulfinyl&lt;br&gt;Polarised alkyne - pyridine&lt;br&gt;Polarised alkyne - pyrazine&lt;br&gt;Polarised alkyne - pyrimidine&lt;br&gt;Polarised alkyne - triazine</td>
</tr>
</tbody>
</table>
| Quinones and quinone-type chemicals | Benzoquinones  
|                                     | Quinone-methides  
|                                     | Pyranones  
| Acid imides | Acid imides  
| Schiff base |  
| SN2 |  
| SN2 reaction at a sp\(^{3}\) carbon atom | Alkyl halides  
|                                     | Sulfates  
|                                     | Sulfonates  
|                                     | Allyl acetates  
|                                     | Nitrosoureas (carbon)  
|                                     | \(\alpha\)-Halocarbonyls  
|                                     | Phosphates  
|                                     | Thiophosphates  
|                                     | Phosphonates  
|                                     | \(\alpha\)-Halo ethers  
|                                     | \(\beta\)-Halo ethers  
|                                     | Alkyl diazo  
|                                     | \(\alpha\)-Haloalkenes  
|                                     | \(\alpha\)-Haloalkynes  
|                                     | \(\alpha\)-Halobenzyls  
| Epoxides and related chemicals | Epoxides  
|                                     | Aziridines  
|                                     | Sulfuranes  
| Ring Opening SN2 Reaction | \(\beta\)-Lactones  
| SN2 reaction at a nitrogen atom | Nitrosoureas (nitrogen)  
|                                     | N-Acetoxy-N-acetyl-phenyl  
|                                     | N-Acyloxy-N-alkoxyamides  
| SN2 reaction at a sulphur atom | Isothiazol-3-ones (sulphur)  
|                                     | Aromatic sulphonic acids  
|                                     | Thiocyanates  
|                                     | Thiols  
|                                     | Disulfides  
|                                     | Thiosulfonates  
|                                     | Sulfoxides of disulfides  
|                                     | Sulphenyl halides  
| SN2 reaction at a halo atom | N-Chloro-sulphonamides  
|                                     | N-Haloimides  
| SN2 reaction at a sp\(^{3}\) carbon | Polarised alkene with a  
|                                     | halogen leaving group  
|                                     | Polarised alkene with a  
|                                     | sulfonate leaving group  
|                                     | Polarised alkene with a sulfate  
|                                     | leaving group  
|                                     | Polarised alkene with a  
|                                     | phosphonate leaving group  

46
Polarised alkene with a phosphate leaving group

<table>
<thead>
<tr>
<th>Episulfonium ion formation</th>
<th>Mustards 1,2-Dihaloalkanes</th>
</tr>
</thead>
<tbody>
<tr>
<td>S_{\text{NAr}}</td>
<td>S_{\text{NAr}}</td>
</tr>
<tr>
<td>Activated halo-benzenes</td>
<td>Activated halo-pyridines</td>
</tr>
<tr>
<td></td>
<td>Halo-pyrimidines</td>
</tr>
<tr>
<td></td>
<td>Halo-triazines</td>
</tr>
</tbody>
</table>

Table 6.2: Detailed breakdown of the mechanistic domains, mechanistic alerts and structural alerts resulting from the analysis presented in section 6.1 for The Updated Profiler.

6.3 Mitigating factors

The presence of mitigating factors associated with a structural alert was also defined within the mechanistic chemistry analysis, where such data existed in the literature. In the current analysis, a mitigating factor is only considered as part of the alert if it has been shown to completely abolish covalent protein binding activity. This is in keeping with other definitions of the mitigating factors and ensures that such factors are transparent and (importantly) easy to encode computationally into the final alert. Mitigating factors can be (approximately) divided into three classes: steric, electronic, and detoxifying. Previous analysis of the mechanistic chemistry associated with covalent DNA binding has shown all three of these mitigating factors to be important factors (27). In contrast, analysis of the mitigating factors identified during the analysis of structural alerts for direct acting covalent protein binding revealed only steric and electronic factors to be significant. The lack of detoxifying mitigating factors is perhaps unsurprising given the current analysis did not investigate chemistry associated with metabolic and oxidatively activated alerting groups.

6.4 The meta data associated with the structural alerts

The mechanistic chemistry information contained within the meta data files is designed to be as transparent and simple as possible. The concept is to help the user understand the chemistry (and associated mitigating factors) associated with the structural alerts identified within each of the mechanistic alerts. The meta data are not designed to highlight any examples of toxicological data associated with a category as this information can be accessed by running the profiler through the databases contained within The Toolbox. The meta data detail the structural alert, the mechanism (or mechanisms) involved in covalent protein binding, known mitigating factors, and the literature sources from which the mechanistic information has been drawn. The meta data contain the following information (under the headings shown):

- Mechanistic alert: Defines the name of the mechanistic alert. A short description is given.
- Structural alert: Defines the structural features.
- Mechanism: Defines the electrophilic reaction chemistry.
• Category mitigating factors: Defines the structural features that remove reactivity for each of the structural alerts that make up the mechanistic alert.

• References: Details the literature supporting the mechanistic chemistry for each of the structural alerts within the mechanistic alert.

Importantly the meta data do not contain any of the following information:

• Information about the types of protein adducts formed.

• Results of toxicological testing (these can be found by running the profiler though databases within The Toolbox).

The resulting meta data for each structural alert have been compiled into mechanistic domain specific meta data files (see supplementary information). These files contain all the meta data for alerts associated with a given mechanistic domain organised into mechanistic alerts as detailed in Section 6.1. Importantly, if an alert had potential mechanisms across more than a single mechanistic domain, then it was assigned to mechanistic alerts within both domains.

7.0 Potency sub-categorisation

In addition to the description of the mechanistic chemistry associated with direct acting covalent protein binding, the current work also investigated the use of experimental reactivity data (so called in chemico data). The aim of this was to identify sub-categories within mechanistic domains based on experimental potency data (analogous to the rate of reaction between electrophile and nucleophile). The rate of reaction between an electrophile and biological nucleophile should give an indication of the chemicals likely toxicity, with a faster rate constant indicating a more toxic chemical (40). In the current work plan the in chemico data for the Michael addition mechanistic domain were utilised as a proof of concept (the remaining domains will be addressed in due course given the approval of the approach by The Toolbox management group). A recent review article details a wide range of in chemico reactivity data and is likely to be of use in the future for potency based sub-categorisation (41).

Experimentally determined in chemico data were retrieved from a literature source for the recently developed glutathione depletion assay (42). The in chemico assay is short-term, static, concentration-response protocol with the tripeptide glutathione as the model nucleophile. The basic assay mixes the test chemical with a solution of GSH of known concentration. The reaction is stopped after 120 minutes and the concentration of GSH remaining is determined. From a knowledge of the concentration of GSH at the start and after 120 minutes, the relative depletion of GSH can be determined. The assay is performed at a number of different concentrations of the test chemical such that the 50% reactive concentration (termed the RC50 value), which measures the concentration required to complete half the reaction within a fixed time as compared to a control, can be determined. The concentration of free GSH is determined after 120 minutes by reacting it with another chemical, which produces a definite colour and this is recorded. This value is proportional to the rate constant of the reaction between the electrophile and glutathione. A number of studies have shown the ability of data generated using this assay to model the toxicity of electrophilic chemicals to Tetrahymena
In addition, a recent study highlighted the use of such data to sub-categorise skin sensitisers allowing read-across predictions to be made (19). Analysis of these data allowed a number of potency driven sub-categories to be developed based on the suggested RCso ranges as shown in Table 7.1. It is worth noting that the „Suspect” category is used for chemicals which are found to be reactive in the assay but for which contamination is likely to be an issue. A detailed overview of the structural alerts that make up each of the sub-categories can be found in the supporting documentation accompanying this report and the following publications (19, 39, 42).

<table>
<thead>
<tr>
<th>Sub-category</th>
<th>RC50 range (mM)</th>
<th>Number of structural alerts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extremely reactive</td>
<td>&lt;0.050 - 0.099</td>
<td>23</td>
</tr>
<tr>
<td>Highly reactive</td>
<td>0.10 - 0.99</td>
<td>19</td>
</tr>
<tr>
<td>Moderately reactive</td>
<td>1.0 - 15</td>
<td>3</td>
</tr>
<tr>
<td>Slightly reactive</td>
<td>16 - 70</td>
<td>4</td>
</tr>
<tr>
<td>Suspect</td>
<td>70.1 - 135</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7.1: In chemico potency from the Schultz et al GSH depletion assay based sub-categorisation for the Michael mechanistic domain and the number of structural alerts for each potency sub-category

7.1 Skin sensitisation: cinnamic aldehyde

As an example of the use of potency based sub-categories within mechanistic domains, consider the skin sensitisation potential and assigned reactive sub-category based on in chemico data for the small series of chemicals tested in Table 7.2 (all skin sensitising data taken from (45)). Consider, for example, if one was trying to predict the skin sensitising potential for cinnamic aldehyde. Inspection of the in chemico data for cinnamic aldehyde shows it to be classified as highly reactive. One can use this reactivity information in conjunction with the in chemico and in vivo data for the other three chemicals to make a read-across prediction for the in vivo skin sensitising potential of cinnamic aldehyde. This analysis suggests that cinnamic aldehyde should have a similar skin sensitising potential to trans-2-decenal and trans-2-hexenal (but lower than that for benzoquinone). Thus one can predict (correctly) cinnamic aldehyde to be a moderate skin sensitiser.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>EC3 (%wt)</th>
<th>Sensitising class</th>
<th>Potency sub-category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzoquinone</td>
<td>0.0099</td>
<td>Extreme</td>
<td>Extremely reactive</td>
</tr>
<tr>
<td>Trans-2-decenal</td>
<td>2.5</td>
<td>Moderate</td>
<td>Highly reactive</td>
</tr>
<tr>
<td>Cinnamic aldehyde</td>
<td>3.0</td>
<td>Moderate</td>
<td>Highly reactive</td>
</tr>
<tr>
<td>Trans-2-hexenal</td>
<td>5.5</td>
<td>Moderate</td>
<td>Highly reactive</td>
</tr>
</tbody>
</table>

Table 7.2: Skin sensitisation (local lymph node assay) and reactive sub-categories for a small series of chemicals within the Michael acceptor mechanistic domain

7.2 Acute aquatic toxicity: 4-methyl-2-pentenal

The glutathione reactivity data in The Toolbox can be used in two ways to develop categories suitable for filling the data gap that is present for aquatic toxicity as measured in *Tetrahymena pyriformis* for 4-methyl-2-pentenal (Figure 7.1). Since the endpoint is aquatic toxicity one can use the Protein Binding by OASIS profiler to investigate the potential mechanisms of action for 4-methyl-2-pentenal.
This profiling results in two potential mechanisms, Schiff base formation and Michael addition. Extracting *Tetrahymena pyriformis* toxicity data (IGC50) from the *Aquatic OASIS* database indicates there to be a data gap for 4-methyl-2-pentenal (Figure 7.2).

![Structure of 4-methyl-2-pentenal](image)

Figure 7.1: Structure of 4-methyl-2-pentenal (structural alert for „polarised alkene – aldehyde“ shown in red)
Figure 7.2: Mechanistic profiling (shown in red text) and toxicity data gap (highlighted by the grey box) for 4-methyl-2-pentenal
Category formation and data gap filling using glutathione RC50 values as a descriptor in trend analysis

The mechanistic analysis indicates that 4-methyl-2-pentenal can act via either a Michael addition mechanism or by Schiff base formation. It is likely that the Michael addition mechanism will dominate in this chemical due to the steric accessibility at the β-carbon atom (the site nucleophilic attack in a Michael addition reaction) combined with the deactivating effect of the conjugated alkene upon the carbonyl carbon (the site of nucleophilic attack in a Schiff base mechanism).

One can develop an initial mechanistic chemical category based on a protein binding mechanism of Michael addition consisting of 29 chemicals (aquatic toxicity data extracted from the Aquatic toxicity OASIS database). Inspection of this category shows that 23 of these chemicals have Tetrahymena pyriformis toxicity data associated with them. In addition, seven chemicals also have glutathione reactivity data (RC50 values). Therefore, one can use these RC50 values within the Michael addition glutathione reactivity database as descriptors in a trend analysis to predict the toxicity of 4-methyl-2-pentenal to Tetrahymena pyriformis. There are six chemicals for which data exist for both glutathione reactivity and toxicity to Tetrahymena pyriformis. This leads to a trend analysis that predicts an IGC50 value for 4-methyl-2-pentenal of 1.32 x 10^-4 mol/l (Figure 7.3).

![Figure 7.3: Trend analysis within a mechanistic category using RC50 as a descriptor allowing the toxicity of 4-methyl-2-pentenal to Tetrahymena pyriformis to be predicted](image-url)
Sub-categorisation using Michael addition potency profiler

An alternative way that the glutathione data can be utilised is by sub-categorising the initial mechanistic category using the Michael addition potency profiler. This profiler has been designed by analysing the glutathione RC₅₀ values within the database allowing so-called iso-reactive categories of chemicals to be identified. Chemicals within an iso-reactive category can be considered to be equally reactive and thus for aquatic toxicity hydrophobicity should be the controlling factor for toxicity.

Utilising the Michael addition potency profiler to sub-categorise the 29 chemical mechanistic category developed above results in an iso-reactive category of nine chemicals. The key structural feature for the iso-reactive category is an alkene polarised by an aldehyde with an alkyl chain (in which the carbon atom attached to the alkene must be sp³ hybridised) or a benzene ring at the β-carbon atom (Figure 7.4). A trend analysis upon this category using LogKₐw as the descriptor can predict the toxicity to *Tetrahymena pyriformis* of 4-methyl-2-pentenal. This results in a predicted IGC₅₀ of 1.65 x 10⁻⁴ mol/l (Figure 7.5). The predicted value is in keeping with that predicted using the RC₅₀ data as a descriptor and thus one has confidence that the two approaches lead to complimentary data gap filling results.

![Structural domain for the iso-reactive category for 4-methyl-2-pentenal.](image)

Figure 7.4: Structural domain for the iso-reactive category for 4-methyl-2-pentenal.
Figure 7.5: Trend analysis within an iso-reactive category using LogKow as a descriptor allowing the toxicity of 4-methyl-2-pentenal to Tetrahymena pyriformis to be predicted (shown as the red data point). Note that the relationship is not really inverse as the Y axis would normally be plotted as “– log(IGC50)” rather than „log(IGC50)”.

8.0 Conclusions

This report has detailed the development of The Updated Profiler which has been based upon previously published lists of structural alerts. These alert compilations have been analysed in order to place the information contained within the literature alerts into a mechanistic chemistry framework. It is this mechanistic chemistry that will be used as the basis for chemical category formation when utilising the new profiler, and thus the associated meta data in the OECD (Q)SAR Application Toolbox. The structural alerts identified within each of the five mechanistic domains have been organised into mechanistic alerts based on the presence of common reactivity sites. This has resulted in the development of 17 mechanistic alerts covering 92 structural alerts. These data are supported by mechanistic chemistry and references to the scientific literature. In addition, a framework for the use of in chemico reactivity data allowing the development of potency based sub-categories has been developed for the Michael addition mechanistic domain.
9.0 Supplementary material
Several additional documents detailing the mechanistic domains (the meta data) and the suggested sub-categories developed using *in chemico* data accompany this report. These documents are as follows:

- Potency sub-category data: Michael acceptors – potency categorisation.pdf
- Acylation meta data: AC.pdf
- Michael addition meta data: MA.pdf
- Schiff base meta data: SB.pdf
- Bimolecular aliphatic nucleophilic substitution meta data: SN2.pdf
- Aromatic nucleophilic substitution meta data: SNAR.pdf
10.0 References


ANNEX 4: SUPPLEMENTARY INFORMATION TO THE REPORT

- Potency sub-category data: Michael acceptors – potency categorisation.pdf
- Acylation meta data: AC.pdf
- Michael addition meta data: MA.pdf
- Schiff base meta data: SB.pdf
- Bimolecular aliphatic nucleophilic substitution meta data: SN2.pdf
- Aromatic nucleophilic substitution meta data: SNAR.pdf

<To be attached later>
Re-Evaluation of Structural Alerts for the Binding of Molecules to Proteins and the Development of a Comprehensive Profiler of Alerts (Deliverable D5.10)

Steven Enoch and Mark Cronin
School of Pharmacy and Chemistry
Liverpool John Moores University
England

Importance of Protein Binding

- Protein binding is a key molecular initiating event in:
  - Skin and respiratory sensitisation
  - Elevated acute aquatic and terrestrial toxicity
  - Protein binding leads to downstream biological effects
    - Immunological responses
    - Disruption of cellular proteins
Biological Nucleophiles

• Cysteine, lysine and serine can act as nucleophiles

• Evidenced in the types of mechanistic chemistry observed in differing biological endpoints

• Skin sensitisation: cysteine and lysine

• Respiratory sensitisation: lysine

• Role of hard-soft acid-base theory

Molecular Initiating Event (MIE)
Aims of Deliverable 5.10

• Review currently published structural alert compilations for direct acting protein binding

• Perform a mechanistic chemistry analysis of these alerts to generate the ‘meta data’

• Compile a new set of alerts for the OECD Toolbox for direct acting electrophiles

Aims of Deliverable 5.10

• Illustrate how in chemico reactivity can be utilised within the Michael addition domain

• Develop Michael addition potency profiler

• Use reactivity data as a descriptor

• N.B. No validation was to be undertaken of the alerts or potency profiler
QSAR Modelling Approaches

Global statistical models → Expert systems → Mechanistic categories

Increasing mechanistic interpretability and transparency

MIE Based Chemical Categories

- Knowledge of protein binding is useful to group chemicals

- Skin sensitisation
  - Group chemicals by an electrophilic mechanism
  - Data gaps can then be filled using read across

- Aquatic toxicity
  - Identify electrophilic chemicals
  - Local (Q)SARs and read across
  - Non-electrophilic chemicals are likely to act via narcosis
Electrophilic Reaction Chemistry

- Chemicals capable of protein binding chemicals are electrophiles (or can be metabolised into electrophiles)
- Proteins contain nucleophilic centres
- Thus, principles derived from organic chemistry are essential in mechanistic category formation
- Structural alerts can be utilised to relate the chemistry to protein binding

Electrophilic Reaction Chemistry

\[
\text{Mechanism?} \quad \text{Protein binding} \\
\text{Electrophilic carbon atom} \\
\text{Electronegative carbonyl polarises the alkene}
\]
Electrophilic Reaction Chemistry

Mechanistic Structural Alerts

- Mechanistic chemistry enables transparent category formation
- QSAR / read-across more likely to be successful within mechanistic domains
- Such analyses in keeping with OECD Principles for the Validation of (Q)SARs
- Profilers can be implemented into the current OECD Toolbox software architecture
Structural Alerts: The Toolbox

- 38 protein binding alerts currently in The Toolbox
- Developed from an analysis of skin sensitisation data
- Original rule base developed during a collaboration between industry and LMC
- This collaboration produced the TIMES-SS system

Structural Alerts: Acute Aquatic Toxicity

- Several alert compilations from acute fish toxicity
  - 21 alerts from Verhaar et al (1992)
  - 36 alerts from Hermens et al (1990)
  - 29 alerts from Nendza et al (2007)
- Analysis of *Daphnia magna* toxicity
  - 7 alerts from von der Ohe et al (2005)
Structural Alerts: Skin and Respiratory Sensitisation / Irritation

- Analysis of skin sensitisation expert systems
  - 29 alerts from Zinke et al (2002)

- Analysis of respiratory sensitisation data
  - 12 alerts from Enoch et al (2010)

- Analysis of skin irritation and corrosion data
  - 22 alerts from Hulzebos et al (2005)

Structural Alerts: Roberts and Aptula

- 28 direct acting structural alerts for electrophiles

- Derived from expert analysis of skin sensitisation data from the local lymph node assay

- Defines the electrophilic mechanistic domains
  (Michael addition, acylation, $S_N1$, $S_N2$, Schiff base, $S_{NAr}$)

- Alerts assigned to domains based on this study
Mechanistic and Structural Alerts

- Groups of related structural alerts are grouped into so-called mechanistic alert.

- Alerts grouped by a common reaction site within a mechanistic domain into a mechanistic alert.

- For example, structural alerts for acroleins, acrylates and methacrylates grouped into ‘polarised alkenes’ mechanistic alert.

Mechanistic and Structural Alerts

[Diagrams showing examples of polarised alkenes and polarised alkynes]
Mechanistic Domains, Mechanistic Alerts and Structural Alerts

• Mechanistic domain
  • Mechanistic alert 1
    • Structural alert 1
    • Structural alert 2
    • ...
  • Mechanistic alert 2
    • Structural alert 1
    • Structural alert 2
    • ...

Mechanistic Domains, Mechanistic Alerts and Structural Alerts

• Mechanistic domain: Michael addition
  • Mechanistic alert 1: Polarised alkenes & related
    • Structural alert 1: Polarised alkene - aldehydes
    • Structural alert 2: Polarised alkene - ketones
    • ...
  • Mechanistic alert 2: Quinones
    • Structural alert 1: Benzoquinones
    • Structural alert 2: Quinone-methides
    • ...
Methodology – Documented Information

• Establish a clear mechanism(s) for each literature alert
• Document the mechanism schematically
• Assign the structural alert a mechanistic alert and thus to a mechanistic domain
  • Michael acceptor
  • Schiff base formation
  • $S_NAr$
  • $S_N^1$
  • $S_N^2$

Methodology – Information Not Documented

• Potential protein adducts

• Toxicological data associated with each alert / mechanism

• Documented information is provided to assist the user understand the chemistry behind the mechanism and thus the potential category
Alert Mapping

1. Can the alert be assigned to a mechanistic domain?

   - Yes
   - No

2. Can the alert be assigned to an existing mechanistic alert?

   - Yes
   - No

3. Does the alert expand any of the established structural alerts?

   - Yes
     - Alert becomes a new structural alert within a new mechanistic alert
   - No
     - Alert becomes a new structural alert

Example: azocarbonamide

1. Can the alert be assigned to a mechanistic domain?

   - Yes

2. Can the alert be assigned to an existing mechanistic alert?

   - Yes

3. Does the alert expand any of the established structural alerts?

   - No

Alert defines a new structural alert
Example: cinnamic aldehyde

1. Can the alert be assigned to a mechanistic domain? 
- Yes

2. Can the alert be assigned to an existing mechanistic alert? 
- Yes

3. Does the alert expand any of the established structural alerts? 
- Yes

Alert Mapping Overview

- Structural alerts are grouped based on common reaction sites resulting in a mechanistic alert
- For example, β-carbon atom in polarised alkenes
- Exceptions do occur: for example azocarbonamides are mechanistically analogous to polarised alkenes
Alert Summary

<table>
<thead>
<tr>
<th>Mechanistic domain</th>
<th>Mechanistic alert</th>
<th>Structural alerts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acylation</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Michael addition</td>
<td>4</td>
<td>31</td>
</tr>
<tr>
<td>Schiff base formers</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>$S_{N1}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$S_{N2}$</td>
<td>8</td>
<td>39</td>
</tr>
<tr>
<td>$S_{N}Ar$</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

Michael Domain Summary

<table>
<thead>
<tr>
<th>Mechanistic domain</th>
<th>Mechanistic alert</th>
<th>Structural alert</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael addition</td>
<td>Polarised alkenes and related chemicals</td>
<td>Polarised alkene - aldehyde</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Polarised alkene - ketone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Azocarbonamides</td>
</tr>
<tr>
<td>Polarised alkynes</td>
<td>Polarised alkyne – aldehyde</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Polarised alkyne - triazines</td>
</tr>
<tr>
<td>Quinones and quinone-type chemicals</td>
<td>Benzoquinones</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quinone-methides</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pyranones</td>
<td></td>
</tr>
<tr>
<td>Acid imides</td>
<td>Acid imides</td>
<td></td>
</tr>
</tbody>
</table>
Mitigating Factors

• 6 mechanistic categories have mitigating factors associated with them

• Three types of mitigating factor identified:

  • Steric hindrance at the reaction site
  
  • Electronic deactivation of the reaction site
  
  • Biological preference for a given nucleophile

<table>
<thead>
<tr>
<th>Mitigating factor</th>
<th>Example chemical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steric</td>
<td><img src="image" alt="Steric" /></td>
</tr>
<tr>
<td>Electronic</td>
<td><img src="image" alt="Electronic" /></td>
</tr>
<tr>
<td>Biological</td>
<td><img src="image" alt="Biological" /></td>
</tr>
</tbody>
</table>
The ‘Meta Data’

• Structural alert: the structural features that define the chemicals within the class

• Mechanism: the electrophilic reaction chemistry

• Mitigating factors: structural features that remove activity

• References: literature supporting the mechanism and the mitigating factors

The ‘Meta Data’

• The ‘meta data’ doesn't include:
  
  – Information about the types of protein adducts formed

  – Results of toxicological testing
Category Formation using the Updated Protein Binding Profiler

Protein Binding Profiling

4-methyl-2-pentenal

Mechanistic alert

Structural alert
### Category Formation: Mechanistic Alert Level

#### Structure

- **X** = aldehyde, ketone, ester, amide, nitro, cyano

#### Predicted log(IGC50) = $6.16 \times 10^{-4}$ mol/l
Category Formation: Structural Alert Level
Category Formation: Structural Alert Level

Predicted log(IGC50) = 1.66 x 10^{-4} \text{ mol/l}
Mechanistic and Structural Alert Categories

- Mechanistic alert level profiling:
  - Covers a broader range of chemical classes
  - Useful when you have reactivity data
  - Predicted log(IGC50) = 6.16 x 10^{-4} mol/l

- Structural alert level profiling:
  - Covers a single chemical class (aldehydes)
  - Can use hydrophobicity as a descriptor
  - Structural domain important
  - Predicted log(IGC50) = 1.66 x 10^{-4} mol/l

Conclusions

- D5.10 aimed to develop a comprehensive listing of structural alerts based on current literature knowledge

- This has resulted in 17 mechanistic alerts covering 92 structural alerts supported by mechanistic chemistry

- Two levels of profiling are available:
  - Mechanistic alert level
  - Structural alert level
Future Work

• Coding of the alerts into The Toolbox

• Input of meta data information

• Checking of functionality of alerts

• Evaluation of performance of the alerts
Overview of the evaluation

- The “Updated profiler” for covalent protein binding can significantly enhance the mechanistic support of the QSAR toolbox and categorizing/predictive capability
- The mechanistic classification can help to achieve structural association with mechanistic backing
- Coverage extensive and close to being exhaustive
- GSH reactivity great in chemico approach for potency; role of SA stability?, applicable to other mech. domains?
- Mitigating factors need refinement/expansion for some SAs to avoid over-sensitivity and lower specificity
- Confidence evaluation of SA not clear?
- Future consideration: will the metabolic profiler be able to cover all pre/pro-protein binders? SA and protein adduct stability? Relative potency ranking?
Additional databases for consideration

• Isothiazolinone allergy compounds (expand SA coverage)
• Hemoglobin binding/adduct
  – Extensive database
  – Structural diversity
  – Quantitative data available (HBI)
  – (Q)SAR studies available
  – Human biomonitoring data
  – Hematotoxicity predictor
  – Nongenotoxic mechanism of splenic tumorigenesis

Additional SAs for consideration

• Quinoneimine and quinonediimine (the N analogs of quinone)
• Sultone (the S analog of β-propiolactone)
• Acyl glucuronides
  - metabolites of various pharma
  - intramolecular rearrangement to protein binder
Some thoughts on Mitigating Factors

- Current version may need some refinement
- Handling of F as leaving group in some SAs needs correction
- Handling of dibasic/tribasic esters as alkylators needs correction
- Reactivity (e.g., size of R as alkylator?)
- Stability (e.g., short-lived acylators or \(\alpha\)-halo-ethers/thio-ethers portal of entry concerns only?)
- Toxicodynamics issue: adduct stability
- Enhancing factors for higher concern?
Comments on Mechanistic Domains
Mechanistic Domain: $S_{N\text{Ar}}$

• Looks good
• Can perhaps mention F a better leaving group than other halogens in activated aromatic ring (if ranking is needed)

Comments on Mech. Domain: Acylation

• Fast hydrolysis expected for most SAs; concern may be limited to portal of entry
• Possible typos: “azide” should be “cyano”
• Phosphonic acid halides: R cannot be H (too unstable), X not defined
• Dialkylcarbamoyl halides: R cannot be H
• Sultone should be added to the list of SA
• Thiocyanate/isocyanate/isothiocyanate: R cannot be H
Comments on Mech. Domain:
Michael Addition

• SA polarized alkene-cyano has wrong structure in formula and reaction mechanism (-NO₂ should be –CN)
• SA azocarbonamide not a polarized alkene
• Polarized alkynes: R should be H only
• Add quinoneimine and quinonediimine as SA to quinone-like chemicals

Comments on Mech. Domain:
Schiff Base Formers

• Looks good
• Should perhaps make a separate SA for R = α,β-unsaturated (including β-disubstitution)
• Can adduct stability be a toxicodynamic issue?
Comments on Mech. Domain: S_N2

- X cannot be F in several SAs (e.g., alkyl halide, β-haloether, etc) due to strength of C-F bond
- Dibasic (e.g., sulfate) and tribasic (e.g., phosphate/thiophosphate) esters must be fully esterified to be alkylators. R cannot be H.
- Alkylating esters not reactive if R is too big
- SA phosphate X = O only; SA thiophosphate X = S
- α-Haloethers/thioethers hydrolyze in seconds; portal of entry concern only
- α-Haloethers/thioethers and β-haloethers/thioethers vastly different in potency
- Isothiazol-3-one/isothiazolin-3-one SA should expand coverage to allow halogen/benzo at 4,5-positions

Potency subcategorization by GSH depletion

- Interesting *in chemico* approach with great potential
- “Proof of concept”, details?
- Role of SA stability
- Applicability to other mechanistic domains
- Data on isothiazolones/isothiazolinones?
- Alternative ranking approaches
Some thoughts on Confidence

- Currently not clearly discussed/provided?
- Degree of confidence (probable/well established vs. possible/hypothetical)
- Screening for hazard vs. assessing potential risk
- Tolerance for false positive/negative
- Supportive evidence/documentation (e.g., actual protein adduct data, representative protein adduct-related toxicity endpoint data, SAR)
- Regulatory acceptance

Suggestions for future consideration

- Can metabolic profilers correctly predict pre/protein binders?
- Consideration of relative ranking of SA reactivity and SA stability
- Impact of protein adduct stability
- Systematic supportive evidence for confidence
- Biological significance and predictive values of protein adducts
ANNEX 7: REVIEW REPORT BY DRS. VANDERBRIEL AND RORIJE

RIVM/Rob Vandebriel, Emiel Rorije
Bilthoven, The Netherlands,
8-10-2010

The remarks below should not necessarily be seen as criticism on the report, which we think is a very nice piece of work, complete and well organized. It should be read as discussion points on what might be needed to include this information in a useful way in the QSAR Toolbox.

1. We are not aware of additional literature which would extend, clarify or support SAs for protein binding.
2. The SAs for forming categories based on mechanisms of protein binding seem to be adequate and complete. See 7a (validation) and 7b (pre- and prohaptns) for discussion.
3. The mitigating factors which alter the chemical protein binding ability seem to be adequate. It is unclear whether these affect toxicokinetics and toxicodynamics, and whether other factors might have an influence on the toxicokinetics and dynamics of protein binding. In this respect a (small) discussion on the role of factors influencing bioavailability / skin permeability e.g. governed by lipophilicity, octanol-water partition coefficient, water solubility, molecular size/weight etc. would be expected (either in the report, or better in the Toolbox). There have been several investigations which conclude that bioavailability does not play a significant role in distinguishing sensitizers from non-sensitizers. However, e.g DerekfW provides a standard text with every skin sensitization alert highlighting possible mitigating action of physicochemical properties of a substance. In addition, see discussion point on detoxification (as a mitigating factor) below.
4. The rationalization seems complete and is easy to follow. Specifically the hierarchy going from mechanistic domain to mechanistic alert into structural alerts is very well implemented and makes it easier to see the logic behind different structural alerts.
5. Differences in confidence associated with each alert are not as such present in the overview. The same goes for confidence associated with the mitigating factors. Actually the confidence (performance) of the alerts should be evaluated using the Toolbox workflow, with data on specific toxicological endpoints. This should indicate whether an alert has a high confidence for that specific toxicological endpoint. We think it is not useful to have confidence indications for the protein binding alerts.
6. Distinguishing between different potency classes for reactivity, within the set of structural alerts, seems a scientifically sound approach. Proof of the concept is e.g. the trend analysis as shown in figure 7.5. If within an iso-reactive category the toxicity is governed solely by partitioning behavior, this indicates that all substances within this category are actually equally toxic (expressed in internal concentrations). This is actually the whole concept behind the EPA ECOSAR models for predicting fish toxicity. However, this bottom-up approach based on reactivity potency classes might allow combining several different ECOSAR classes into one “reactivity potency” class.
7. a. As a next step in the process, validation for specific toxicological endpoints is required. Potentially missing SAs and wrong predictions/bad alerts should present themselves when applying the protein binding profile for predicting e.g. skin sensitization. A suggestion would be to use the ICCVAM LLNA dataset (466 compounds) and run the toolbox. That is also the way to come up with (endpoint specific) confidence indications (q.5) for each alert.
b. The (eventual) inclusion of a separate profiler which deals with known pre- and prohaptns should be considered, instead of relying fully on the metabolic simulators, and the user actually applying the metabolic simulators as standard.
c. For the Michael acceptor potency compilation, although being a proof-of-principle, data from e.g. Gerberick et al. and Natsch et al. might also be included, or at least could be mentioned as having been considered.

Additional discussion points:

**Detoxification**
Detoxifying mitigating factors are absent in this overview, since pre- and prohaptens were excluded from the investigation. However, (empirical) observations that certain substances/alerts will in general be quickly metabolized, leading to effective detoxification for a certain toxicological endpoint, should be included somewhere in the QSAR Toolbox, possibly as part of/addition to the metabolic simulator(s).
1. Completeness of literature

- We are not aware of additional literature which would extend, clarify or support SAs for protein binding.

- Interesting article from Mekenyan et al. discussing the overlap between DNA binding alerts and Protein binding alerts.
  - Do we want profiles related to (toxicological) endpoints?
  - Do we want one reactivity profile, with reactive potency determining which toxicological endpoint will be affected?
  - currently a “mix” in the Toolbox
2. Adequacy and Completeness of the SAs

- The SAs for forming categories based on mechanisms of protein binding seem to be adequate and complete.
  - Validation needed to find missing or bad alerts?
  - A separate profiler for pre- and prohaptens?

3. Adequacy of the mitigating factors

- The mitigating factors which alter the chemical protein binding ability seem to be adequate.

- not discussed:
  - Physico-chemical factors influencing toxicokinetics & dynamics
  - Detoxification (pre-haptens)?:
    > (empirical) observations that certain substances/alerts will in general be quickly metabolized, leading to effective detoxification for a certain toxicological endpoint, should be included somewhere in the QSAR Toolbox, possibly as part of/in addition to the metabolic simulator(s).
4. Documentation associated with each alert.
   ● The rationalization seems complete and is easy to follow.
   ● Hierarchy from mechanistic domain to mechanistic alert into structural alerts is very clear and instructive.

5. Associated confidence noted for each alert
   ● No Confidence (performance) indications given in the report
   ● Confidence for protein binding is not possible?
   ● Confidence of the alerts should be evaluated (by the user) using the Toolbox workflow, using data on specific toxicological endpoints (Guidance).
6. Implementation of reactive potency

- The concept seems scientifically sound, and useful! as a next step in the evaluation of toxicity of substances

- EPA ECOSAR models for predicting fish toxicity are based on the same concept – but empirical.

- Should reactive potency be dealt with before covering pre- and prohaphtens?

7. Further suggestions for improvements (recap)

improvements:
- Validation
- Pre- and prohaphtens (separate pre/prohapten profiler? detoxification as part of metabolic simulator?)
- Physico-chemical factors influencing toxicokinetics & dynamics
- More data on peptide reactivity data is available for the reactive potency approach (Natsch, Gerberick).

discussion:
- One reactivity profile vs. different toxicological endpoint profiles?
- Should reactive potency be dealt with before covering pre- and prohaphtens?
Evaluation of the Structural Alerts for Protein-Binding in the Context of the OECD (Q)SAR Toolbox

Nora Aptula
20th October 2010

- I have been asked to comment on the completeness of literature reviewed and to indicate any additional literature, which I feel would further clarify or support structural alerts for protein binding.

- While the literature review is quite thorough, it may be augmented by an examination of:
  - Fabjan & Hulzebos, Tox. in Vitro, 22 (2008), 468
  - Bergstrom et al., J. Med. Chem. 51 (2008), 2541
  - Kern et al., Dermatitis 21 (2010), 8
I have been asked to comment on the adequacy and completeness of the structural alerts for forming categories based on mechanisms of protein binding.

While the proposed alerts are an improvement on those in version 1 of the Toolbox, they may be improved by:

Modification to the Schiff base domain (Mechanistic alert: Direct acting Schiff base formers; Structural alert: Mono-carbonyls) to include:
- $\alpha,\beta$-unsaturated aldehydes that have the $\beta$-C disubstituted
- heterocyclic aldehydes

Similarly, the alerts may be improved by:

Modification to the Michael acceptor domain to include:
- new structural alert for oximes under the “polarised alkene mechanistic alert”
I have been asked to comment on the adequacy of the mitigating factors, affecting either toxicokinetics or toxicodynamics which alter protein binding ability.

Since mitigating factors are typically identified through testing, it is likely that for well tested reactions such as Michael Addition these factors are complete and therefore adequate.

As more reactivity data for different reaction becomes available – new mitigating factors might be defined. This information will need to be updated into the Toolbox.

When mitigating factors were observed:

- They are reported after the corresponding structural alert
- They are clearly stated in written
- But no pictorial representation is given – is it needed?
I have been asked to comment on documentation associated with each structural alert. In particular if the rationalization was complete yet easy to follow

All cited references are peer-reviewed papers in the scientific literature

They are clearly referenced after each proposed reaction mechanism and the mechanism is pictorially represented

The rational for each alert is clear, comprehensive and easy to follow, and should be of great value to the non-chemist

I have been asked to comment on the associated confidence noted for each structural alert, especially for those alerts where I feel the confidence may be misstated

All alerts have been extracted from published literature and all are based on hazard or reactivity data so they can be stated with a high degree of confidence

However, there is no “numeric confidence” stated in the documentation – is this necessary?
I have been asked to comment on the scientific rationale and clarity of the proposed implementation of reactive potency into the Toolbox workflow.

While only limited to the Michael acceptor domain and only based on a single model nucleophile - glutathione, the RC50 reactivity data is proposed to be used in two very clear and transparent ways: 1) to develop potency-based sub-categories for read-across and 2) as an independent variable in QSAR modelling.

Both of these approaches fit within the category-based data-gap filling mission of the Toolbox.

However, since Unilever is donating the GSH RC50 data for the implementation of reactive potency into the Toolbox,

I am not without bias.
Summary

- The new series of structural alerts for protein binding is an improvement of the alerts in version 1 of the Toolbox

- All identified alerts have an associated mechanistic chemistry rational

- They are comprehensive and represent the currently availability hazard data

- The meta-data for each alert is clear, comprehensive and easy to follow

Summary

- The proposed ways of implementing chemical reactivity in the form of GSH RC50 values is consistent with the workflow of the Toolbox

- However, reactivity data for all the structural alerts and with model nucleophiles (not just GSH) representing different regions of the hard-soft electrophilic spectrum will be needed to complete this task