QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

User manual

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

Document history

Version	Comment
Version 1.0	January 2011: Manual based on QSAR Toolbox version 2.0

Issue date: January 2011

Language: English

If you have questions or comments that relate to this document, please send them to ehscont@oecd.org or visit the QSAR Toolbox discussion forum at https://community.oecd.org/community/toolbox_forum

Table of Contents

How to use this guidance document 8

1 Genetic toxicity and carcinogenicity 10

1.1 *In vitro* endpoints 10

1.2 In vivo tests 11

2 Primary profilers relevant to genetic toxicity and/or carcinogenicity 13

2.2 DNA binding by OASIS 17

2.3 Protein binding by OASIS 18

2.4 Mutagenicity & carcinogenicity alerts by Benigni/Bossa 19

2.5 Micronucleus alerts by Benigni/Bossa 20

2.6 OncoLogic primary classification 20

3 Database relevant to genetic toxicology and carcinogenicity 22

4 Profiling results: What they tell us about a grouping strategy 24

4.1 Mutagenicity 24

4.2 Chromosomal aberration (*in vivo* and *in vitro*) 28

4.3 Genotoxic carcinogenicity 30

4.4 General conclusions regarding the outcome of profiling strategy results 32

5 Secondary profilers relevant to genetic toxicology and carcinogenicity endpoints 35

5.1 Defining the structural domain of a chemical category 37

6 Initial category formation and subsequent sub-categorisation with primary profilers relevant to genetic toxicity and/or genotoxic carcinogenicity 41

6.1 Mutagenicity 42

6.2 Chromosomal aberration 43

6.3 Genotoxic carcinogenicity 45

6.4 General conclusions regarding the sub-categorisation with the primary profilers 47

7 Profiling examples which result in the ability to fill data gaps through read across 48

7.1 Profiling and data gap filling for Ames mutagencity for 3-methoxyaniline 48

7.1.1 Initial profiling using the primary profilers 49

7.1.2 Initial category formation and sub-categorisation using the primary profilers 50

7.1.3 Empiric sub-categorisation using the secondary profilers 51

7.1.4 Data-gap filling via read across 51

7.2 Profiling and data gap filling for a chemical with more than a single covalent mechanism, 3-aminobenzaldhyde, for Ames mutagenicity 53

7.2.1 Initial category formation and sub-categorisation using the primary profilers - Nitrenium ion formation 54

7.2.2 Empiric sub-categorisation using the secondary profilers 57

7.2.3 Data gap filling using read across for 3-aminobenzaldehyde for the nitrenium ion category 59

7.2.4 Initial category formation and sub-categorisation using the primary profilers - Schiff base formation due to the presence of a mono-aldehyde 61

7.2.5 Empiric sub-categorisation using the secondary profilers 63

7.2.6 Data gap filling using read across for 3-aminobenzaldehyde for the nitrenium ion category 66

7.2.7 Summary of category formation for 3-aminobenzaldehdye 66

8 General approach for the development of chemical categories for genetic toxicity and genotoxic carcinogenicity 68

9 References 70

Table of Tables

Table 2.1 Primary profilers for genetic toxicity and/or genotoxic carcinogenicity available in the OECD QSAR Toolbox V2.0 15

Table 3.1 Summary of databases available within the OECD QSAR Toolbox V2.0 that are relevant to genetic toxicity and/or carcinogenicity 22

Table 4.1 Summary of the applicable primary profilers for the Ames mutagenicity, chromosomal aberration and genotoxic carcinogenicity available in the OECD QSAR Toolbox V2.0. 34

Table 5.1 Secondary profilers relevant to genetic toxicity and genotoxic carcinogenicity endpoints available in the OECD QSAR Toolbox V2.0. 35

Table of Figures

Figure 2.1 Screenshot of the list of the primary profilers available in the OECD QSAR Toolbox V2.0. Those appropriate for genetic toxicity and genotoxic carcinogenicity are checked. 14

Figure 2.2 Schematic representation of how structural alerts are assigned to mechanistic alerts in the 'DNA binding by OECD' profiler. 17

Figure 3.1 Screenshot of the list of the databases available within the OECD QSAR Toolbox V2.0 that are relevant to genetic toxicity and/or carcinogenicity (checked) 23

Figure 4.1 Result of primary profiling for 4-aminobiphenyl using a combination of the mechanistic and endpoint specific profilers available for mutagenicity in the OECD QSAR Toolbox V2.0. 25

Figure 4.2 Result of primary profiling for quinone using a combination of the mechanistic and endpoint specific profilers for available for mutagenicity in the OECD QSAR Toolbox V2.0. 26

Figure 4.3 Result of primary profiling for 1,3-benzodioxole using a combination of the mechanistic and endpoint specific profilers for available for mutagenicity in the OECD QSAR Toolbox V2.0. 27

Figure 4.4 Result of primary profiling for 3-methyl-2-butenal using a combination of the mechanistic and endpoint specific profilers for available for mutagenicity in the OECD QSAR Toolbox V2.0. 28

Figure 4.5 Result of primary profiling for ethylene oxide using a combination of the mechanistic and endpoint specific profilers for available for chromosomal aberration in the OECD QSAR Toolbox V2.0. 29

Figure 4.6 Result of primary profiling for sulphur mustard using a combination of the mechanistic and endpoint specific profilers for available for chromosomal aberration in the OECD QSAR Toolbox V2.0. 30

Figure 4.7 Result of primary profiling for N-nitrosodimethylamine using a combination of the mechanistic and endpoint specific profilers for available for genotoxic carcinogenicity in the OECD QSAR Toolbox V2.0. 31

Figure 4.8 Result of primary profiling for trichlorotoluene using a combination of the mechanistic and endpoint specific profilers for available for genotoxic carcinogenicity in the OECD QSAR Toolbox V2.0. 32

Figure 5.1 Standard sub-categorisation using the organic functional group profiler (chemicals with functional groups in grey would be removed). 39

Figure 5.2 Inclusion of additional functional groups when using the organic functional group profiler to sub-categorise (chemicals with functional groups in grey would be removed). Notice the inclusion of ether, methyl and methylene compared to Figure 5.1. 40

Figure 6.1 Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OECD' profiler when sub-categorising the initial category formed for the target chemical ethanal. 43

Figure 6.2 Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OECD' profiler when sub-categorising the initial category formed for the target chemical ethylene oxide. 45

Figure 6.3 Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OASIS' profiler when sub-categorising the initial category formed for the target chemical N-nitrosodimethylamine. 47

Figure 7.1 Data gap present (shown in grey) for Ames mutagenicity data for 3-methoxyaniline. 49

Figure 7.2 Profiling results from the initial battery of profilers applicable to Ames mutagenicity for 3-methoxyaniline. 50

Figure 7.3 Read across prediction made for the Ames mutagenicity for 3-methoxyaniline in the TA 1537 strain of Salmonella in the presence of the S9 liver fraction. 52

Figure 7.4 Mechanistic profiling for the target chemical 3-amino-benzaldehyde using the 'DNA binding by OECD' profiler in the OECD QSAR Toolbox V2.0. 54

Figure 7.5 Initial mechanistic profiling – mechanisms in highlighted are removed from the categorisation process by selecting the down arrow. 55

Figure 7.6 Selection of Schiff base mechanisms when sub-categorising using the 'DNA binding by OECD' profiler. The mechanisms relating to Schiff base formation are circled in red and should be selected when performing the sub-categorisation process. 56

Figure 7.7 Data matrix for the five member category formed for nitrenium ion formation.58

Figure 7.8 Data matrix including methyl substituents in the structural domain. 59

Figure 7.9 Additional in vitro and in vivo chromosomal aberration data showing genotoxicity for aniline. 60

Figure 7.10 Initial mechanistic profiling – highlighted mechanisms are removed from the categorisation process by selecting the down arrow. 62

Figure 7.11 Selection of nitrenium ion formation mechanisms when sub-categorising using the 'DNA binding by OECD' profiler. The mechanisms relating to nitrenium ion formation are circled in red and should be selected when performing the sub-categorisation process.63

Figure 7.12 Data matrix formed after the initial secondary profiling. This results in only a single analogue being identified. 65

Figure 7.13 Data matrix formed after extending the structural domain to include chemicals with methyl groups. 65

Figure 7.14 matrix formed after eliminating non-aromatic chemicals. 66

Figure 8.1 General scheme for category formation for genetic toxicity and genotoxic carcinogenicity. 69

How to use this guidance document

This document is intended to offer guidance on how to form robust chemical categories suitable for data gap filling for genetic toxicity and genotoxic carcinogenicity in the OECD QSAR Toolbox. The examples (sections 4.0 – 7.0) used in this document are intended to illustrate one approach to the formation of chemical categories with the OECD QSAR Toolbox. It is important to realise that no effort has been made to validate the read across predictions (by analysing additional data sources not available in V2 of the OECD QSAR Toolbox) that are made for these chemicals. The term genetic toxicity and genotoxic carcinogenicity will refer to the following endpoints in this document:

- Ames mutagenicity
- *In vitro* and *in vivo* chromosomal aberration
- Genotoxic carcinogenicity

The document is split into several sections, these being:

Sections 1.0 – 3.0: Introductory material about the profilers and databases available for genetic toxicity and genotoxic carcinogenicity. These sections include additional references to the literature that provide more in depth background material.

Sections 4.0 – 7.0: Worked examples for profiling target chemicals and how to use this information to form chemical categories for each of the endpoints. These sections are intended to be used as examples that the reader can follow as illustrations of some of the recommended strategies. A first-time reader of this document should spend time ensuring that they can use the profilers and form the suggested categories as instructed. A good working knowledge of the OECD QSAR Toolbox is required for these examples to be of benefit and thus users who do not consider themselves as experts should consult the appropriate guidance documents available from the OECD (*1*).

Section 8.0: Summary of a general strategy that can be used to generate chemical categories suitable for read across for a user's own data. This summary was used to generate the example categories covered in this guidance document. The worked examples in sections 4.0 – 7.0 should be undertaken first before attempting to use this summary information.

1 Genetic toxicity and carcinogenicity

This guidance document aims to outline strategies for grouping chemicals into chemical categories for genetic toxicity and genotoxic carcinogenicity. The effects covered in this document are:

- *In vitro* gene mutation (Ames mutagenicity)
- *In vitro* chromosomal aberration
- *In vivo* chromosomal aberration (micronucleus assay)
- Genotoxic carcinogenicity

These endpoints are assumed to all share a common molecular initiating event (*2*) in the covalent binding of a chemical to DNA (*3-8*). Other molecular initiating events exist, especially for carcinogenicity, such as protein binding, non-covalent interactions with protein receptors, intercalation with DNA and the formation of free radicals (for details of these mechanisms see reference (*8*)). Mechanistic information for so-called non-genotoxic carcinogenicity is sparse and currently The QSAR Toolbox lacks such profilers. Thus, in this document, the focus will only be on genotoxicity.

1.1 *In vitro* endpoints

This guidance document covers the two *in vitro* effects that are typically required for the regulatory assessment of a chemical's mutagenicity. The first of these is the Ames mutagenicity assay (OECD guideline 471). This is specifically designed to assess the ability of a chemical to cause point mutations in the DNA of the bacterium Salmonella typhimurium (*9, 10*). A number of bacterial strains (TA1537, TA1535, TA100, TA98 and TA97) have been engineered to detect differing classes of mutagenic chemicals. The basic test only detects direct acting mutagens (those chemicals able to interact with DNA without the need for metabolic activation). The inclusion of an S9 mix of rodent liver

enzymes is designed to assess those chemicals requiring bioactivation in order to be mutagenic. Typically, chemicals are assayed with and without the presence of the S9 mix. The assay results are reported in a binary fashion, with a positive result in any of the bacterial strains confirming mutagenic potential.

The second in vitro test discussed in this document is the micronucleus assay for chromosomal aberration (OECD guideline 473). The *in vitro* micronucleus assay involves the use of rodent or human cell lines to detect chemicals that can bind to chromosomes resulting in the production of micronuclei (small membrane bound fragments of DNA) during nuclear division. The production of micronuclei is an indicator of damage to the chromosome and/or the ability of the cell to divide correctly (*11*). As with the Ames assay, the micronucleus test is conducted with and without the inclusion of S9 mix of rodent liver enzymes. The in vitro micronucleus assay results are also reported in a binary fashion, with a positive result indicating genotoxic potential.

1.2 In vivo tests

In addition to the two *in vitro* tests described in section 1.1, the use of data from two *in vivo* tests will also be discussed in this guidance document. The first is the *in vivo* micronucleus assay carried out in rodents (OECD guideline 474). As with the *in vitro* micronucleus assay, this assay is specifically designed to detect chemicals capable of causing the production of micronuclei during nuclear division. The *in vivo* assay has several advantages over the *in vitro* micronucleus assay in that metabolism, pharmacokinetics and potential DNA repair processes are all taken into account. A positive result is confirmed by the presence of micronuclei in either extracted bone marrow or blood samples from the animal. This indicates that the chemical is able to produce micronuclei (under the test conditions) in rodents and, thus, be potentially genotoxic in humans. A negative result indicates (under the test conditions) that in rodents the chemical is not able to produce micronuclei.

Experimental assays relating to the assessment of genotoxic carcinogenicity are also covered by this guidance document (OECD guideline 451). Data are available from a number of experimental protocols within the OECD QSAR Toolbox, with the general approach being to investigate the ability of a chemical administered orally to rodents to cause cancerous lesions in a variety of tissues. Two common results are reported for carcinogenicity studies; namely histopathology and TD_{50}. The results of histopathology examinations are reported as a positive/negative, with a positive result indicating the presence of a tumour in at least one of the tissues examined. The second result, TD50 is the dose-rate in mg/kg body wt/day which, if administered chronically for the standard lifespan of the species, will halve the probability of remaining tumourless throughout that period (*12*). Importantly, in addition to covering metabolism, pharmacokinetics and potential DNA repair mechanisms, carcinogenicity assays also cover both genotoxic and non-genotoxic mechanisms of action.

2 Primary profilers relevant to genetic toxicity and/or carcinogenicity

The primary profilers relevant to genetic toxicity and/or carcinogenicity can be divided into two types. The first of type of primary profilers have been developed from a purely mechanistic standpoint and are related to the underlying chemistry governing the ability of a chemical to bind covalently to DNA. A screenshot of the list of primary profilers available in version 2.0 of the OECD QSAR Toolbox is shown in Figure 2.1. These profilers define fragments (so-called structural alerts) within chemicals that have been shown to be associated with a given reaction mechanism known to be important in covalent binding to DNA. Importantly, there are not necessarily toxicological data associated with these structural alerts - with a number of the structural alert being derived from the results of *in chemico* testing (defined as any data taken from chemistry based studies (*13*)). The second type of primary profilers have been developed using toxicological data for a given assay. These profilers define a series of structural alerts that are associated with toxicity for a given assay, i.e. the structural alert has been identified as being with associated with a positive result in either a genotoxicity or carcinogenicity assay. The available profilers in the OECD QSAR Toolbox, along with how they have been developed, are summarised in Table 2.1 (and profilers relevant to genotoxicity are shown checked in Figure 2.1).

Figure 2.1 Screenshot of the list of the primary profilers available in the OECD QSAR Toolbox V2.0. Those appropriate for genetic toxicity and genotoxic carcinogenicity are checked.

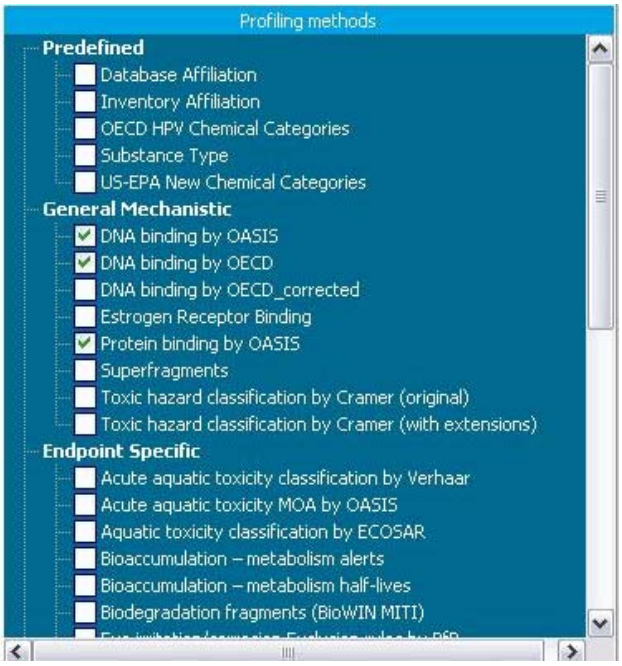


Table 2.1 Primary profilers for genetic toxicity and/or genotoxic carcinogenicity available in the OECD QSAR Toolbox V2.0

Profiler name	Type	Applicable endpoint(s)	Number of alerts	Data source(s) that the structural alerts have been derived from	See section
DNA binding by OECD	General mechanistic	Covalent DNA binding	94	Ames test Carcinogenicity Idiosyncratic drug toxicity *In chemico*	2.1
DNA binding by OASIS	General mechanistic	Covalent DNA binding	41	Ames test	2.2
Protein binding by OASIS	General mechanistic	Covalent protein binding	67	Skin sensitisation	2.3
Micronucleus alerts by Benigni/Bossa	Endpoint specific	*In vivo* chromosomal aberration (micronucleus assay)	35	Chromosomal aberration	2.4
Mutagenicity/Carcinogenicity alerts by Benigni/Bossa	Endpoint specific	Genotoxicity (mutagenicity/carcinogenicity) Non-genotoxic carcinogenicity	37	Ames test Carcinogenicity	2.5
OncoLogic primary classification	Endpoint specific	Carcinogenicity	48	Carcinogenicity	2.6

2.1 DNA binding by OECD

The 'DNA binding by OECD' profiler is based solely on structural alerts for the electrophilic reaction chemistry associated with covalent DNA binding. A detailed review of the electrophilic reaction chemistry covered by this profiler is given by Enoch and Cronin (2010) (*5*). In the simplest terms, the applicable electrophilic reaction chemistry can be defined as the formation of a new chemical bond between a DNA base containing a nucleophilic centre [an area of the molecule with a (partial) negative charge, typically a lone pair of electrons on a nitrogen or oxygen atom] and a exogenous chemical containing an electrophilic centre [an area of a molecule with a (partial) positive charge]. Thus, the profiler details a range of structural alerts that contain electrophilic centres or those that can be metabolically activated to electrophiles. Importantly, the data used to develop this profiler were not necessarily from regulatory toxicological endpoints (such as the Ames test). Instead a range of data sources covering experiments in which a chemical had been shown to bind covalently to one of the nucleobases of DNA. For example, mutagenicity and carcinogenicity data, *in chemico* data and data from idiosyncratic drug toxicity studies were considered to develop these structural alerts.

The structural alerts were assigned to one of six mechanistic domains based on the definitions developed by Aptula and Roberts (*14*). These domains are acylation, unimolecular nucleophilic substitution (S_N1), bimolecular nucleophilic substitution (S_N2), Michael addition, Schiff base formation and radical formation. Within each of these mechanistic domains each of the structural alerts was assigned to a mechanistic alert based on the nature of the reactive centre within the electrophile (for example the formation of a common electrophilic metabolite such as a nitrenium ion). Thus, all of the structural alerts within a given mechanistic domain that have been shown to be reactive by a common feature were assigned to the same mechanistic alert. Taking the nitrenium

ion mechanistic alert as an example, there are 16 structural alerts assigned to it ranging from aromatic amines to heterocyclic amines (Figure 2.2).

Figure 2.2 Schematic representation of how structural alerts are assigned to mechanistic alerts in the 'DNA binding by OECD' profiler.

1. Mechanistic domain (for example S_N1)
    - Mechanistic alert 1 (Nitrenium ion formation)
        - Structural alert 1 (Aromatic amines)
        - Structural alert 2 (Protected aromatic amines)
        - ….
        - Structural alert 16 (Heterocyclic amines)

The use of structural and mechanistic alerts within a mechanistic domain allows for a two tired approach to profiling depending on the exact requirements and the availability of experimental data (mechanistic alert level and structural alert level). As one progresses from a mechanistic domain into a mechanistic alert and then to a structural alert then the resulting chemical category becomes more focussed containing significantly closer analogous to the target chemical of interest.

2.2 DNA binding by OASIS

This profiler is a mechanistic profiler developed from an analysis of Ames mutagenicity data. The structural alerts within this profiler are detailed by Mekenyan et al (2004) and Serafimova et al (2007) (*15, 16*). It contains a number of structural alerts that have been shown to be related to established electrophilic reaction chemistry known to be important in covalent DNA binding. The chemical categories developed using this profiler can be considered similar to those obtained at the structural alert level using the 'DNA binding by OECD' profiler. The electrophilic reaction chemistry for each structural alerting group is detailed within the profiler's meta data. In addition, the meta data included in

the OECD QSAR Toolbox detail examples of DNA adducts for each structural alerting group with references to toxicological data sources. This is useful information that can help to support a mechanistic analysis.

Clearly there is an overlap between the 'DNA binding by OASIS' profiler and the 'DNA binding by OECD' profiler as Ames data have been used in the development of both profilers. It is important to realise that the 'DNA binding by OASIS' profiler is focussed on well established electrophilic mechanisms that have been have been shown to lead to DNA adducts in the Ames assay. In contrast to the 'DNA binding by OECD' profiler, only structural alerts that have supporting toxicological data are included in the profiler. Thus, one can consider this profiler as an intermediate step between the mechanistic chemistry based 'DNA binding by OECD' profiler and the endpoint specific profilers (discussed in sections 2.4 – 2.6).

2.3 Protein binding by OASIS

It has been suggested that the ability of a chemical to bind covalently to proteins is important in chromosomal aberration (*17*). Thus, when developing categories for chromosomal aberration, mechanistic information regarding both covalent DNA and protein binding is useful. The OECD QSAR Toolbox contains a mechanistic profiler for protein binding that has been developed from an analysis of skin sensitisation data (for a review of the method see (*18*)). Obviously skin sensitisation data are not the same as chromosomal aberration data. However, there is considerable overlap in the underlying mechanistic chemistry, thus the applicability of this profiler to chromosomal aberration. This profiler has been developed in the same manner as the 'DNA binding by OASIS' profiler in that it contains a number of structural alerts related to established electrophilic reaction chemistry. The difference being that it documents chemistry related to covalent interactions with proteins that involve sulphur, nitrogen and/or oxygen acting as the

biological nucleophiles. The meta data detail both information about the electrophilic reaction chemistry and potential protein adducts. In addition, the meta data also includes references to toxicological data sources that document potential protein adducts that can be formed upon the formation of a covalent bond between a protein and an exogenous chemical.

2.4 Mutagenicity & carcinogenicity alerts by Benigni/Bossa

This profiler has been developed by the Instituto Superiore di Sanita, Rome, Italy (*3*) as part of the development of the ToxTree software (*19*). The profiler is specific to *in vitro* mutagenicity (Ames test) and *in vivo* carcinogenicity (in rodents). The profiler has been developed using a structural alert approach allowing chemical categories to be formed based on the presence of a given structural alert (for example the presence of an aromatic amine). Most of the alerts are refined with modulating factors. This approach is similar to the structural alert profiling method discussed for the 'DNA binding by OECD' profiler. Importantly, the majority of the structural alerting groups defined in this profiler have a genotoxic mechanistic basis related to interaction with DNA. In addition to structural alerts for genotoxicity, alerts are available for a few non-genotoxic carcinogenic mechanisms (for example thioureas). Documentation on both the biological mechanisms and the electrophilic reaction chemistry is provided for this profiler within The Toolbox. The structural alerting groups covered by this profiler are derived directly from the mechanistic knowledge accumulated during decades of carcinogenicity and mutagenicity research, and were validated against the database of experimental results available in the public domain (e.g., the ISSCAN database on chemical carcinogens (*20*)). This approach has the advantage that an alert present in this profiler has definitive toxicological data associated with it i.e. a chemical containing the alert has been identified previously that is either mutagenic or carcinogenic. This set of structural alerts

was originally developed as a standalone tool for predictive toxicology; however, in the context of the OECD QSAR Toolbox it is used as a piece of information contributing enabling chemicals to be placed into categories.

2.5 Micronucleus alerts by Benigni/Bossa

The micronucleus profiler is an endpoint specific profiler relating to the *in vivo* micronucleus assay conducted in rodents that was initially developed using the genotoxic structural alert compilation developed by Benigni and Bossa for mutagenicity and carcinogenicity (*3*). These structural alerts were used to screen a database of over 700 chemicals that were tested in the *in vivo* micronucleus assay. An analysis of the results identified a number of additional structural alerts associated with activity in the micronucleus assay (for a detailed review of the analysis see (*4*)). The majority of the structural alerts developed for this profiler have a clear mechanistic basis in terms of their electrophilic reaction chemistry. However, a number of the additional alerts identified from the screening of the database do not appear to have a clear mechanistic rationale. As with the Mutagenicity & Carcinogenicity profiler by the same authors, supporting mechanistic chemistry is available in support of the alerts within this profiler (within The Toolbox). Again, within the OECD QSAR Toolbox the identification of a structural alert within a target chemical is intended to allow for category formation.

2.6 OncoLogic primary classification

The OncoLogic primary classification was developed to be used specifically to address carcinogenicity potential. The profiler was developed by the Laboratory of Mathematical Chemistry (LMC) solely to mimic the structural classes of known/potential carcinogens covered in version 7.0 of the United States Environmental Protection Agency's (US EPA) OncoLogic Cancer Expert System for predicting carcinogenic potential (*21*). No attempts were made to incorporate the additional expert system rule base from the OncoLogic

software into the OECD QSAR Toolbox. Since the additional enhancing/mitigating rules of OncoLogic are not part of the Toolbox profiler, users should note that classification of a query chemical in an OncoLogic class does not automatically mean that the chemical will be a carcinogen. As with the two Benigni/Bossa profilers, this means that for each structural alerting group in the profiler, there is at least a single reported incidence of a chemical causing carcinogenicity. However, the same caveat also applies in that the presence of a structural alerting group within a query chemical does not necessarily indicate that the query chemical will be a carcinogen. In addition, no supporting mechanistic chemistry is available for this profiler (within the OECD QSAR Toolbox).

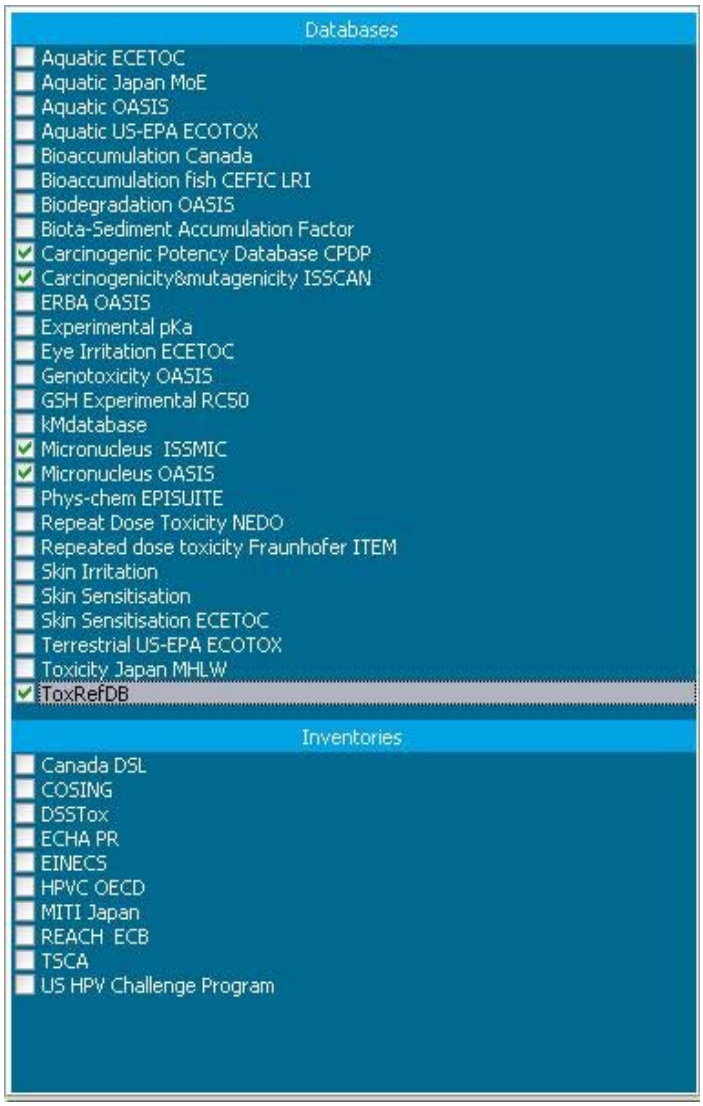
3 Database relevant to genetic toxicology and carcinogenicity

The OECD QSAR Toolbox contains a number of databases that are relevant to genetic toxicology and carcinogenicity. These have been donated by various organisations and cover a range of regulatory endpoints and species. These databases are summarised as shown in Table 3.1. A screenshot of the listing of databases in the OECD QSAR Toolbox with relevant ones highlighted is shown in Figure 3.1.

Table 3.1 Summary of databases available within the OECD QSAR Toolbox V2.0 that are relevant to genetic toxicity and/or carcinogenicity

Database	Number chemicals	Endpoint	Species
Carcinogenic Potency DataBase CPDB	1778	Carcinogenicity (TD_{50})	*In vivo*: Rats, mice, hamsters, dogs and non human primates
Carcinogenicity&Mutagenicity ISSCAN V3a	1153	Carcinogenicity (TD_{50}) Ames mutagenicity (positive/negative)	*In vivo*: Rats and mice *In vitro: Salmonella typhimurium*
Genotoxicity OASIS	7500	Gene mutation (positive/negative) Chromosomal aberration (positive/negative)	*In vitro*: *Salmonella typhimurium* *In vitro*: Chinese hamster lung cells *In vitro*: T-lymphoma cell lines
Micronucleus ISSMIC V2a	151	Chromosomal aberration (positive/negative)	*In vivo*: Rats and mice
Micronucleus OASIS	577	Chromosomal aberration	*In vivo*

Figure 3.1 Screenshot of the list of the databases available within the OECD QSAR Toolbox V2.0 that are relevant to genetic toxicity and/or carcinogenicity (checked)



4 Profiling results: What they tell us about a grouping strategy

The QSAR Toolbox aims to group chemicals into categories on the basis of a common molecular initiating event. This allows data gaps to be filled via trend analysis and/or read across. As outlined in section 1.0, one common molecular initiating event for genetic toxicity and genotoxic carcinogenicity is often the ability of a chemical to bind covalently to DNA. Thus, when using The QSAR Toolbox to group chemicals into categories for the endpoints outlined in sections 1.1 and 1.2, the following general steps for a target chemical for which a data gap exists would be undertaken.

1. Profile the target chemical for potential mechanism of action related to the molecular initiating event
2. Use the result of this mechanistic profiling to select chemicals analogues from endpoint specific toxicological databases
3. Define the mechanistic and structural domain of the resulting chemical category
4. Fill data gaps using trend analysis and/or read across

The following sections (4.1 – 4.4) outline how one should interpret a number of differing results that can occur when one profiles the target chemical of interest (step 1 above) for the endpoints outlined in sections 1.1 and 1.2. The remaining steps (2 – 4) are dealt with subsequent sections of this guidance document.

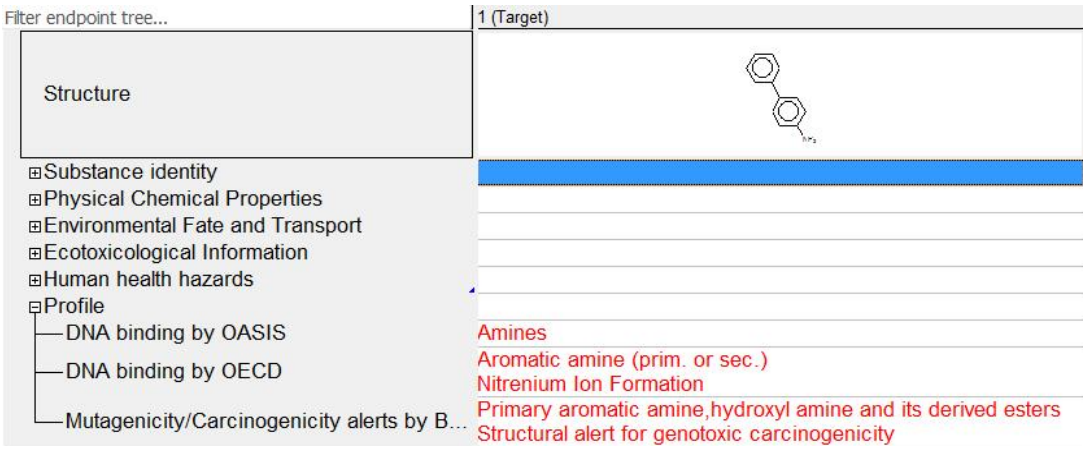
4.1 Mutagenicity

There are three primary profilers that should be used when developing chemical categories for mutagenicity as measured in the Ames assay. In terms of chemical mechanism, both the 'DNA binding by OECD' and 'DNA binding by OASIS' profilers should be applied. Both of these profilers detail knowledge related to mechanisms for covalent DNA binding which is (together with DNA intercalation) a key molecular initiating event

for Ames mutagenicity. A single endpoint specific profiler (i.e. 'mutagenicity/carcinogenicity alerts by Benigni/Bossa' profiler) should also be used in conjunction with the two mechanistic profilers. This profiler details knowledge regarding known structural features that causes mutagenicity in the Ames assay. In profiling a target chemical, one is ideally hoping for consistency from the three initial profilers i.e. that the same mechanism is suggested by the two mechanistic profilers and that a complimentary structural feature is identified by the endpoint specific profiler.

To illustrate the use of the three profilers, consider if one was trying to develop a chemical category that would allow data gaps to be filled for Ames mutagenicity for the target chemical aniline (Figure 4.1). The result of profiling using the two mechanistic profilers suggests a single covalent DNA binding mechanism based on oxidation of the aromatic amine moiety to a reactive nitrenium ion. This mechanistic analysis is supported by the presence of a structural alert for aromatic amines within the endpoint specific profiler (Figure 4.1 shows a screenshot of the profiling results). These complimentary profiling results, for 4-aminobiphenyl, would lead to a good degree of confidence in the resulting chemical category.

Figure 4.1 Result of primary profiling for 4-aminobiphenyl using a combination of the mechanistic and endpoint specific profilers available for mutagenicity in the OECD QSAR Toolbox V2.0.

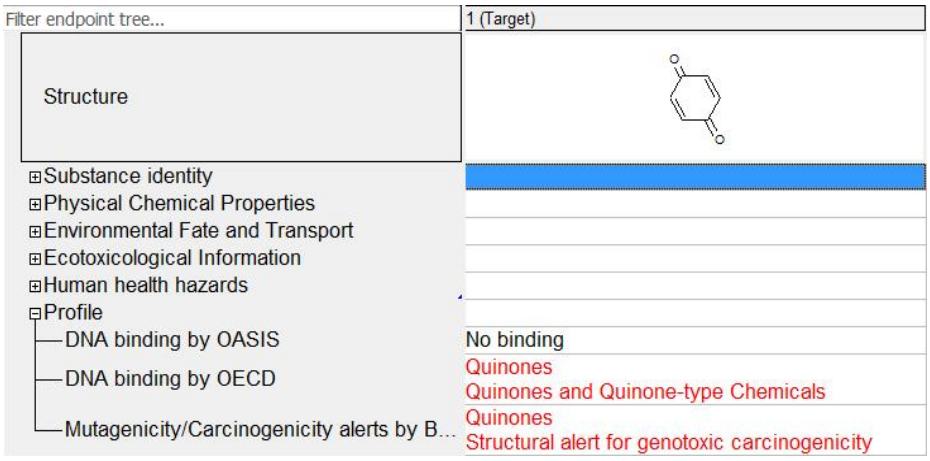


Clearly not every target chemical is going to produce the ideal profiling results that occur when one investigates aniline. In many cases the profiling results will not be complimentary and may even be conflicting in nature. Consider the following three profiling examples:

(i) Quinone

Repeating the above profiling process for the target chemical, quinone results in a conflict between the two mechanistic profilers, with the 'DNA binding by OECD' indicating that the quinone has an electrophilic mechanism, whilst in contrast the 'DNA binding by OASIS' suggests no binding potential. However, inspection of the endpoint profiling results supports the mechanistic assignment that the quinone is able to bind covalently to DNA (Figure 4.2). In this example, one would probably have sufficient confidence that quinone (and quinone-type) chemicals are capable of covalent binding to DNA.

Figure 4.2 Result of primary profiling for quinone using a combination of the mechanistic and endpoint specific profilers for available for mutagenicity in the OECD QSAR Toolbox V2.0.

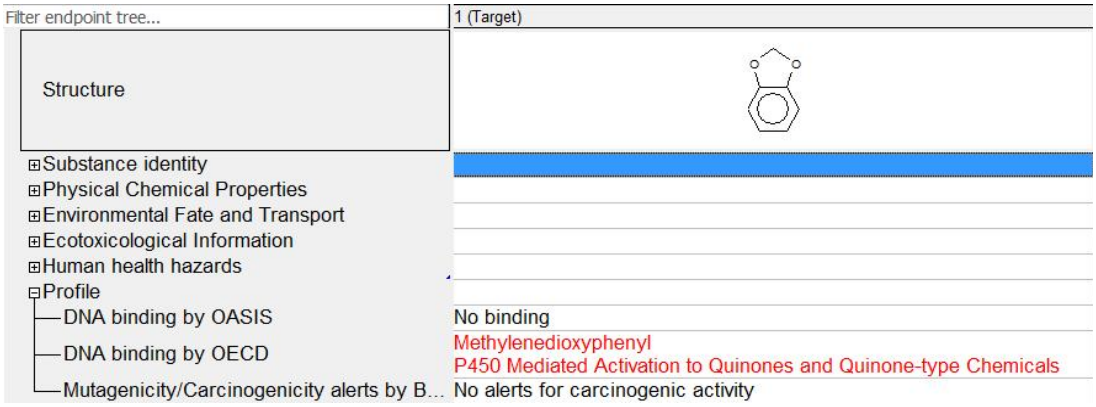


(ii)1,3-Benzodioxole

The profiling results for the chemical 1,3-benzodioxole further illustrate the potential for the conflict between the profilers. In this case, only the 'DNA binding by OECD' profiler

suggests a possibility of covalent DNA binding (Figure 4.3). It is worth noting that the development of this profiler utilised additional data sources (over and above toxicological data) to attempt to develop a broad rule base for covalent DNA binding, thus it might be expected that it contains additional mechanistic categories for which there are little or no supporting toxicological data. However, the lack of support from the other two profilers should make one cautious about the resulting chemical category. This does not mean that one should discard the results for the potential for covalent DNA binding and assume that the chemical cannot bind to DNA and cause mutagenicity. Instead, one should adopt the precautionary principle and use the mechanistic information presented by the single mechanistic profiler to build a chemical category. Subsequent analysis of the mutagenic activity of the category members then allows one to make a decision about the likely toxicity of target chemical via the suggested mechanism.

Figure 4.3 Result of primary profiling for 1,3-benzodioxole using a combination of the mechanistic and endpoint specific profilers for available for mutagenicity in the OECD QSAR Toolbox V2.0.

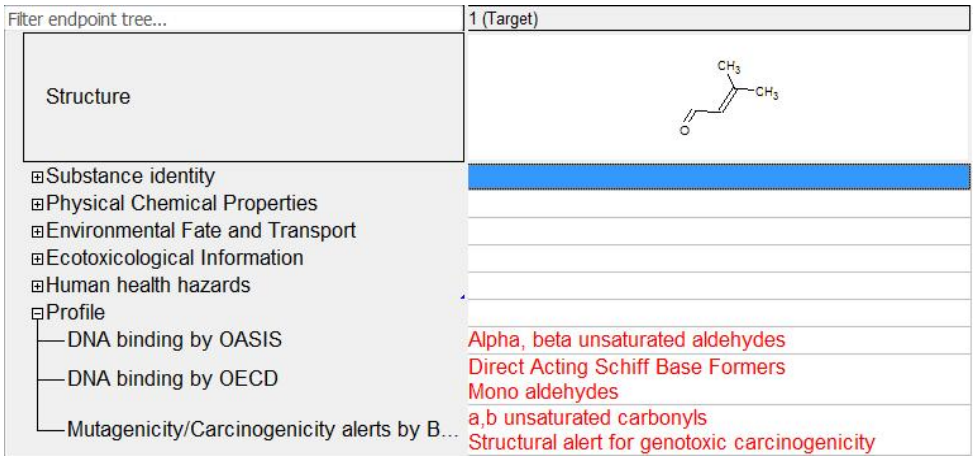


(iii) 3-Methyl-2-butenal

A different situation arises upon inspection of the results for the profiling of 3-methyl-2-butenal (Figure 4.4). The two mechanistic profilers conflict one another for this chemical; the 'DNA binding by OECD' profiler indicating a Schiff base mechanism and the 'DNA

binding by OASIS' profiler suggests Michael addition due to the presence of an α, β-unsaturated aldehyde moiety. The endpoint profiler confirms that one should expect the presence of this functionality to result in toxicity and thus supports the category premise. One approach to dealing with profiling results such as this is to build two chemical categories, one for each of the suggested mechanisms. One can then use the data in each of these chemical categories to assess the likelihood of the target chemical to be toxic via each of the mechanisms. One would then fill any data gaps using both categories and, taking the precautionary principle into account, would then take the worst case toxicological predictions.

Figure 4.4 Result of primary profiling for 3-methyl-2-butenal using a combination of the mechanistic and endpoint specific profilers for available for mutagenicity in the OECD QSAR Toolbox V2.0.



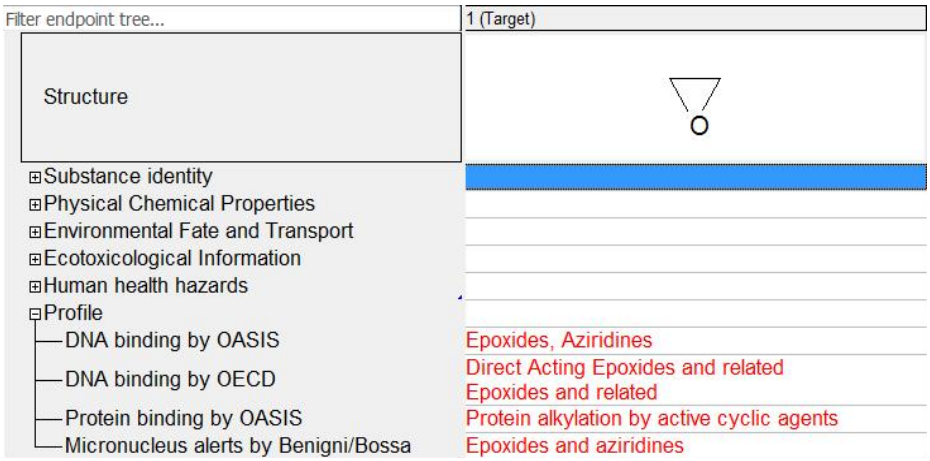
4.2 Chromosomal aberration (*in vivo* and *in vitro*)

A similar profiling analysis can be carried out for chromosomal aberration. It is important to note that, from a profiling point of view, one can treat *in vivo* and *in vitro* chromosomal aberration as the same endpoint. This is because the current version of the OECD QSAR Toolbox contains only a single endpoint specific profiler developed using results from *in vitro* and *in vivo* assays. In contrast to mutagenicity and genotoxic carcinogenicity, the applicable mechanistic profilers for chromosomal aberration include

those related to both covalent DNA and protein binding. Thus, the relevant profilers are 'DNA binding by OECD', 'DNA binding by OASIS', 'protein binding by OASIS' and 'micronucleus alerts by Benigni/Bossa'. As with mutagenicity, the ideal scenario is for these four profilers to reveal complimentary results, with the three mechanistic profilers showing the same mechanism supported by the presence of a structural alert known to be related to the suggested mechanism.

Profiling ethylene oxide shows convergence between the profilers (Figure 4.5). Although the protein binding alert fired for this compound does not refer to "epoxides" directly, examination of the mechanistic explanation of the alert reveals that all of the alerts highlight the same mechanism.

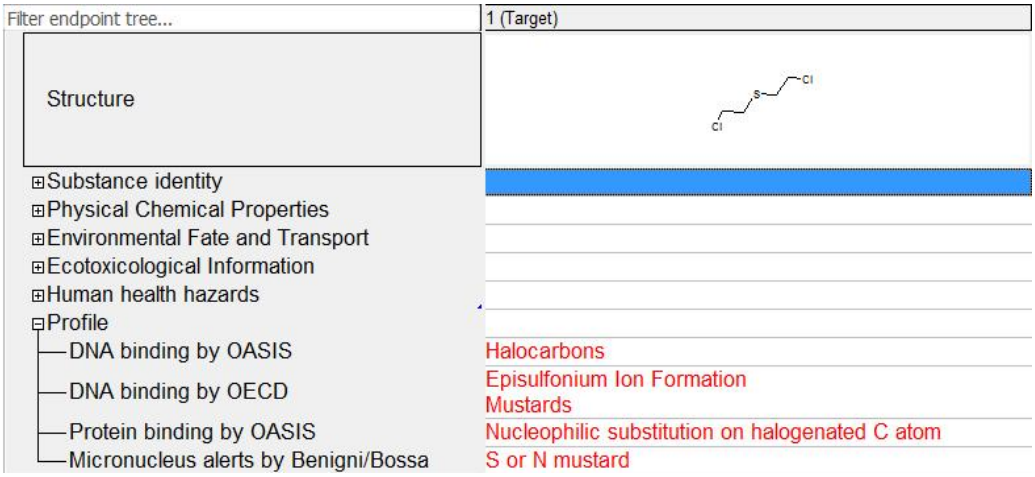
Figure 4.5 Result of primary profiling for ethylene oxide using a combination of the mechanistic and endpoint specific profilers for available for chromosomal aberration in the OECD QSAR Toolbox V2.0.



It is possible that the profiling results from the DNA and protein binding profilers may differ. In cases where the evidence for a potential covalent DNA binding mechanism is well supported by the endpoint specific profiler then one can remain fairly confident in the mechanistic analysis and thus the resulting category. This is supported by a recent analysis that showed that structural alerts for covalent DNA binding (and thus the

electrophilic mechanisms) are more correlated to activity in the chromosomal aberration assay than structural alerts for covalent protein binding (*4*). Consider the example of sulphur mustard profiled by the three primary profilers for chromosomal aberration (Figure 4.6). The 'DNA binding by OECD' profiler and the 'micronucleus alerts by Benigni/Bossa' profiler both highlight the potential for concern because of the mustard group, whereas the protein binding profiler highlights the potential for nucleophilic substitution because of the haloalkane groups.

Figure 4.6 Result of primary profiling for sulphur mustard using a combination of the mechanistic and endpoint specific profilers for available for chromosomal aberration in the OECD QSAR Toolbox V2.0.



Conflicting outcomes from mechanistic profiling where one of the mechanistic profilers suggests a mechanism that is not supported by the endpoint specific profiler should be treated with the most caution. This is not to say the suggested mechanism is incorrect, rather that no further information is (currently) available within the OECD QSAR Toolbox (see the examples discussed for mutagenicity, Figures 4.1 – 4.4).

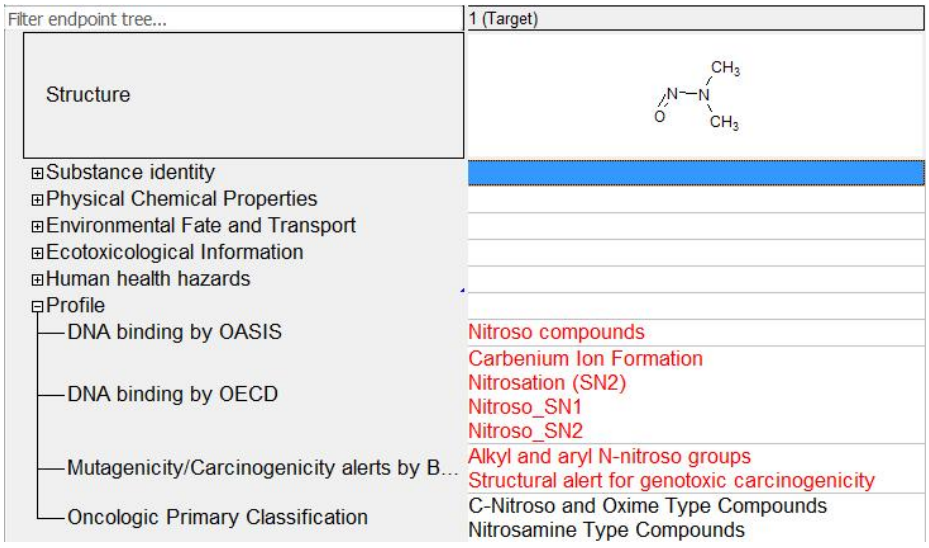
4.3 Genotoxic carcinogenicity

The final profiling strategy discussed in this section is related to genotoxic carcinogenicity. There are two mechanistic and two endpoint specific profilers applicable

to this endpoint; 'DNA binding by OECD', 'DNA binding by OASIS', 'mutagenicity/carcinogenicity by Benigni/Bossa' and 'OncoLogic primary classification' respectively. As in the previous endpoints discussed, the ideal scenario when profiling a target chemical is that the four profilers complement one another with the two mechanistic profilers identifying the same mechanism and are supported by the identification of the same structural alert (related to the identified mechanism).

Profiling N-nitrosodimethylamine shows this ideal situation (Figure 4.7). Several possible mechanisms are proposed by the profilers. However, a concurrent theme of the presence of the nitrosation mechanism is supported by all four of the primary profilers for genotoxic carcinogenicity.

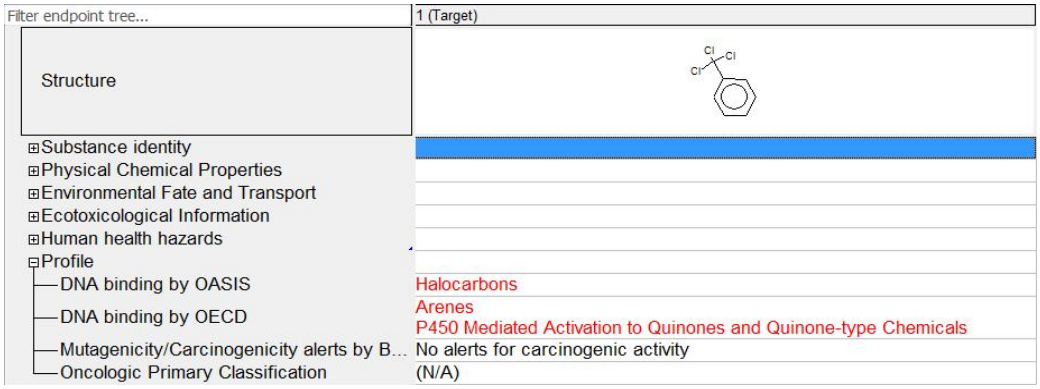
Figure 4.7 Result of primary profiling for N-nitrosodimethylamine using a combination of the mechanistic and endpoint specific profilers for available for genotoxic carcinogenicity in the OECD QSAR Toolbox V2.0.



Given the presence of four primary profilers for this endpoint, there is a high likelihood that the results of the profilers will not converge. An example of such a result is for trichlorotoluene (Figure 4.8). As with the other endpoints discussed in this guidance document, one should treat conflicting profiling results with caution. As before, this is not

to imply the suggested mechanism is not important only that further analysis is recommended before one can be confident in the suggested category (for example, building multiple chemical categories for each mechanism as was discussed in section 4.1 for mutagenicity).

Figure 4.8 Result of primary profiling for trichlorotoluene using a combination of the mechanistic and endpoint specific profilers for available for genotoxic carcinogenicity in the OECD QSAR Toolbox V2.0.



4.4 General conclusions regarding the outcome of profiling strategy results

The above examples of the profiling strategies highlight the fact that the initial battery of profilers should not be given equal weightings. In fact when dealing with genetic toxicity and genotoxic carcinogenicity, one should ideally hope to base any chemical category around a mechanism suggested by one of the DNA binding mechanistic profilers. In addition, one should remember that the 'DNA binding by OECD' profiler is the broader of these two profilers and contains significantly more mechanistic classes that have been shown to be capable of binding to DNA covalently. One can consider the chemistry in this profiler as the primary source of mechanistic data for endpoints in which covalent DNA binding is the molecular initiating event. The second mechanistic profiler is narrower in scope as it was developed from the analysis of mutagenicity data.

For chromosomal aberration, one should also take into account the mechanistic information related to protein binding. As discussed, if the two mechanistic profilers (three in the case of chromosomal aberration) agree then the confidence in the resulting chemical category should be high. The endpoint specific profilers have been developed using toxicological data and thus are of greatest use in confirming the results of one of the mechanistic profilers via the identification of a complimentary structural alert within the target chemical. Such a confirmation can be thought of an indication that data exist that shows that the chemical with the specified moiety is toxic for the endpoint of interest. However, it should be noted that this does not mean that the target chemical being profiled is itself toxic – this decision must be based on the trend analysis or a read across prediction using the resulting members of the chemical category. Finally, one should try and avoid using the endpoint specific profilers for mutagenicity and genotoxic carcinogenicity in isolation for the development of a chemical category. Significant weighting should be placed on the presence of well defined mechanistic chemistry associated with DNA binding (complimented by protein binding for chromosomal aberration). The applicable primary profilers for each of the three endpoints (for the purposes of profiling taking *in vitro* and *in vivo* chromosomal aberration as a single endpoint) are summarised in Table 4.1.

Table 4.1 Summary of the applicable primary profilers for the Ames mutagenicity, chromosomal aberration and genotoxic carcinogenicity available in the OECD QSAR Toolbox V2.0.

Endpoint	Primary profilers
Ames mutagenicity	DNA binding by OECD DNA binding by OASIS Mutagenicity/Carcinogenicity alerts by Benigni/Bossa
Chromosomal aberration	DNA binding by OECD DNA binding by OASIS Micronucleus alerts by Benigni/Bossa
Genotoxic carcinogenicity	DNA binding by OECD DNA binding by OASIS Mutagenicity/Carcinogenicity alerts by Benigni/Bossa OncoLogic primary classification

5 Secondary profilers relevant to genetic toxicology and carcinogenicity endpoints

In addition to the primary profilers, a number of secondary profilers are also of use in category formation for genetic toxicity and genotoxic carcinogenicity endpoints. These profilers are summarised in Table 5.1. In contrast to the initial battery of profilers which are used in combination with one another, the secondary profilers are best utilised individually to help sub-categorise a chemical category. Such sub-categorisation is often needed to refine the structural domain of a chemical category allowing transparent structure-activity relationships to be developed. However, it is important to state that these secondary profilers are based on various measures of chemical similarity or the presence of functional groups. Thus, they should be used cautiously in order to ensure that sub-categorisation is carried out in a logical and transparent manner. It is not the intention to delete or exclude structural analogues from a category for unknown reasons.

Table 5.1 Secondary profilers relevant to genetic toxicity and genotoxic carcinogenicity endpoints available in the OECD QSAR Toolbox V2.0.

Profiler name	Type	Number of alerts
Organic functional groups	Empiric	227
Organic functional groups (nested)	Empiric	227
Organic functional groups (US EPA)	Empiric	467
Chemical elements	Empiric	29
Superfragments	General mechanistic	NA
Liver metabolism simulator	Metabolism	NA

The most commonly utilised and useful secondary profilers are the organic functional groups and chemical elements profilers. These profilers allow the user to develop sub-categories based on the presence or absence of common organic functional groups such as carbonyl, nitro or many others. In addition, the chemical elements profiler allows sub-

categories to be developed based on the presence or absence of chemical elements. A combination of one of the organic functional group profilers and the chemical elements profiler can provide useful sub-categories depending on the makeup of the chemical category. The choice of which of the three organic functional group profilers to use is largely dependent on the data within the category one wishes to sub-categorise. However, as a general approach, it is advised to use the organic functional group profiler as it relates to well established organic functional groups and thus is the most interpretable. The remaining two organic functional group profilers should be used in cases where the organic functional group profiler does not provide a satisfactory sub-category. The superfragments profiler is also based on chemical fragments – specifically those used in the ClogP calculations for predicting hydrophobicity. In addition to the organic functional group profiler, the chemical elements profiler is also a useful secondary profiler. This profiler encodes the chemical elements within a molecule allowing the user to exclude a given element or sets of elements. This would become useful during the fine-tuning of a chemical category as it allows the user to restrict the category members to those whose elements are the same as are present in the target chemical.

The final, relevant, secondary profiler is the liver metabolism simulator (*22*). This profiler encodes a number of rules governing the biotransformation in the liver of chemicals into potentially electrophilic metabolites. It is worth noting that the primary profilers (discussed in sections 2.1 – 2.6) already include a number these transformations in their rule bases. This is particularly true of the two mechanistic profilers. Thus, the use of the liver metabolism simulator should be limited when investigating genetic toxicity and genotoxic carcinogenicity endpoints. However, its use should not be excluded if the initial

profiling using either the 'DNA binding by OECD' or 'DNA binding by OASIS' profilers does not indicate the presence of an electrophilic mechanism.

5.1 Defining the structural domain of a chemical category

One of the key functions of the secondary profilers is in the definition of the structural domain of the chemical category. It is important that chemicals containing (significantly) different elements and functional groups are removed from the category. Typically this is achieved using a combination of the organic functional group and chemical elements profilers (see Table 3 in section 5.0). Ideally, the category resulting from the primary profiling should contain only chemicals with the same elements and functional groups as the target chemical (those identified in the 'target menu' of the sub-categorisation window). However, this is not always possible and using such a tight structural domain may result in the elimination of too many analogues from the category. In these instances, one can include more functional groups (by selecting them 'by hand' in the 'analogues menu' in the sub-categorisation window) to increase the number of analogues in the resulting category. A useful approach to ensure that the structural domain of the chemical category is suitable for subsequent data gap filling is as follows (in usage order):

1. Profile the endpoint specific category using the organic functional group profiler removing all chemicals that contain functional groups not present in the target chemical (Figure 5.1 – only functional groups that are within the target chemical listed in the 'target window' are included). Inspect the resulting chemical category (achieved by selecting the 'Remove' button shown in Figure 5.1) – if it contains sufficient analogues (that one considers) suitable to fill the data gaps of interest then no further sub-categorisation is required (the absolute minimum for read across is a category containing the target chemical and a single analogue as this would allow for one-to-one read across. However, ideally one would like a category in which trend analysis and/or read across predictions could be made on a many-to-one basis. Thus, one wants a category containing at least two or three analogues if possible. For general guidance on grouping, chemical category formation and read across see (*23*)).

2. If step 1 results in insufficient chemicals considered suitable for data gap filling, then re-profile the endpoint specific category using the organic functional group profiler. However, instead of removing all chemicals as before, additional simple non-ionisable organic functional groups not present in the target chemical should be included (Figure 5.2 – note the inclusion of ether, methyl and methylene). This increases the likelihood that there will be sufficient chemicals included in the resulting chemical category to allow for data gap filling.

When profiling for organic functional groups 'by hand' (as in step 2 above), it is extremely important to visually inspect the types (i.e. the chemical structures and associated functional groups) of chemicals that one is eliminating. The chemicals that will be eliminated can be visualised by right clicking on the 'sub-categorisation' window and selecting 'display selected'. One approach when sub-categorising organic functional groups in this way is to try to produce a chemical category in which a suitable descriptor (for example hydrophobicity or chemical reactivity) is responsible for (the majority of) the trends in toxicity. Doing so will ensure that any subsequent predictions made by read across or trend analysis are as transparent and interpretable as possible.

Figure 5.1 Standard sub-categorisation using the organic functional group profiler (chemicals with functional groups in grey would be removed).

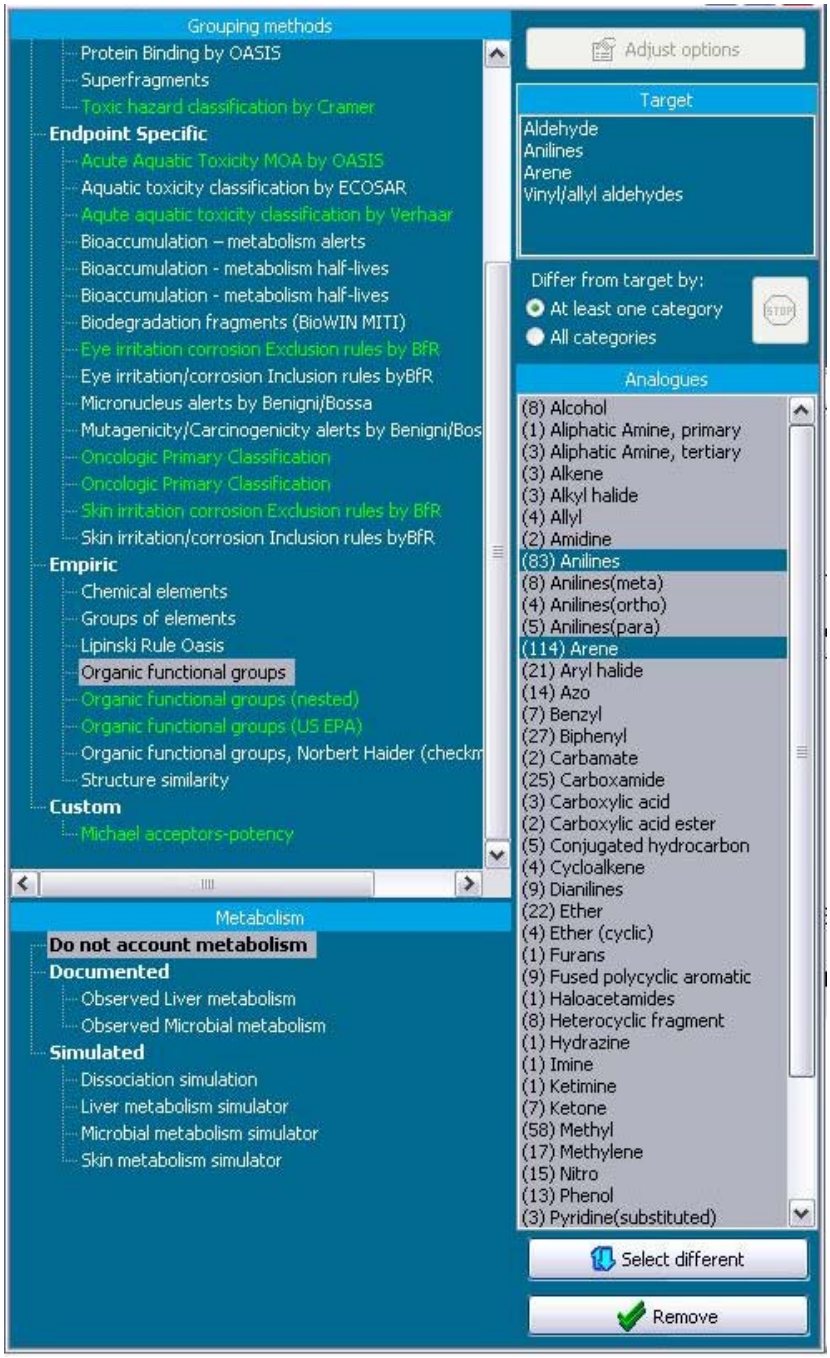
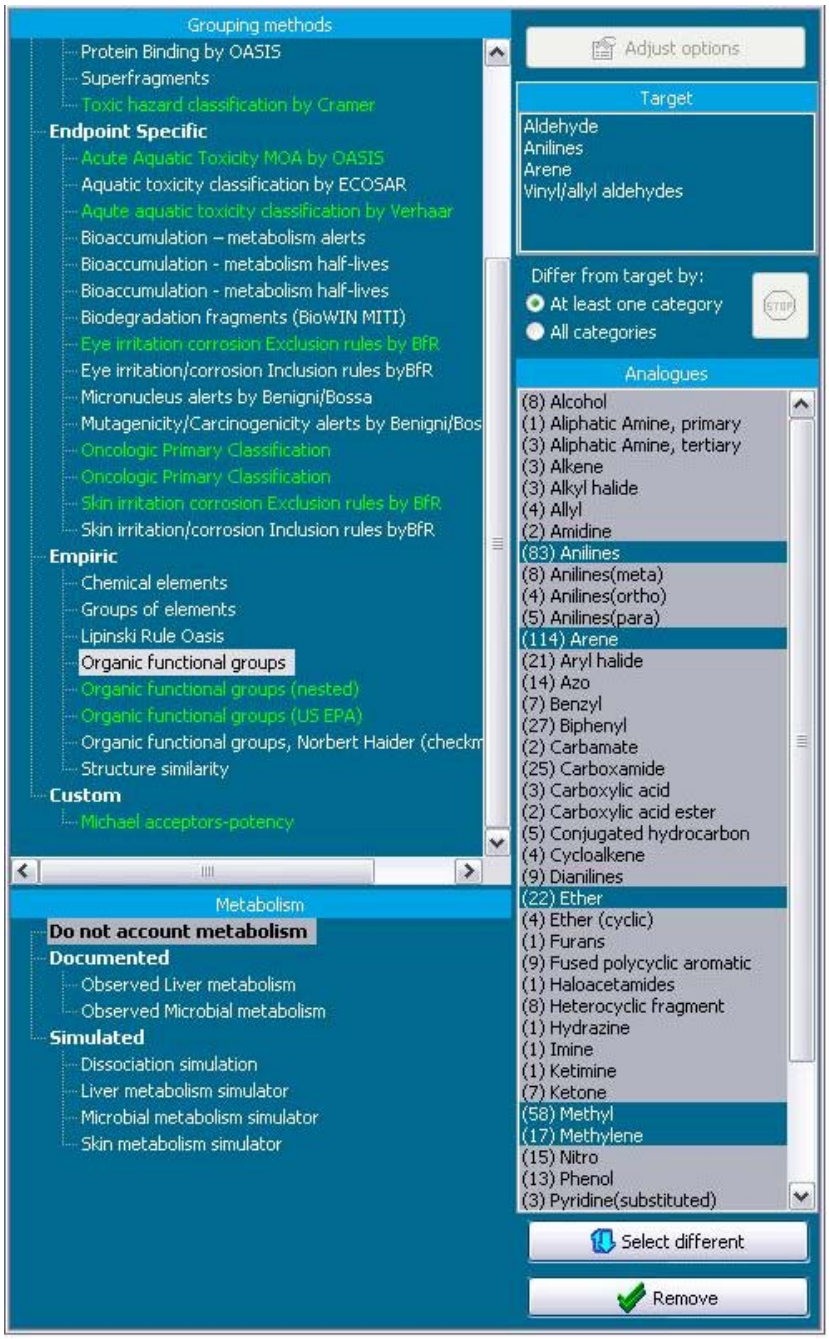


Figure 5.2 Inclusion of additional functional groups when using the organic functional group profiler to sub-categorise (chemicals with functional groups in grey would be removed). Notice the inclusion of ether, methyl and methylene compared to Figure 5.1.



6 Initial category formation and subsequent sub-categorisation with primary profilers relevant to genetic toxicity and/or genotoxic carcinogenicity

The formation of the initial chemical category is carried out by profiling the relevant databases (see Section 3) to genetic toxicity and genotoxic carcinogenicity using one of the two mechanistic profilers ('DNA binding by OECD' or 'DNA binding by OASIS' profilers). The profiling and formation of the initial category is carried out as follows:

1. Profile the target chemical using the relevant primary profilers as outlined in sections 4.1 – 4.3.

2. Using the results of one of the mechanistic profilers, profile the relevant databases to the endpoint of interest for chemical analogues acting via the same mechanism as the target chemical. It is important to take into account the discussion presented in sections 4.1 – 4.4 in terms of how to form a chemical category for a target chemical with more than a single mechanism of action (or profiling results in which conflicting mechanisms are identified).

3. The resulting category is termed the 'initial category'.

It is frequently necessary to perform a sub-categorisation of the initial category using one or more of the primary profilers relevant to genetic toxicity and genotoxic carcinogenicity (see sections 2.1 – 2.6). This is to ensure that the chemical category relates to a single mechanism of action. Sub-categorisation of a category is carried out as follows:

1. Profile the initial category with the primary profiler that was used to develop the initial category. This profiling will identify the mechanisms (if using a mechanistic profiler) or structural alerting groups (if using an endpoint specific profiler) that are present in the initial category. These mechanisms (or structural alerting groups) are displayed in the sub-categorisation window.

2. Eliminate analogues from the initial category that contain additional mechanisms (or structural alerting groups if using an endpoint specific profiler). Ensure that the 'differ from target by' option in the sub-categorisation window is set to 'at least one category'.

3. Repeat steps 1 and 2 for the other relevant primary profilers. The order in which the primary profilers are used to perform the additional sub-categorisations is not important.

Specific examples will now be discussed to show how the above process works for each of the three endpoints.

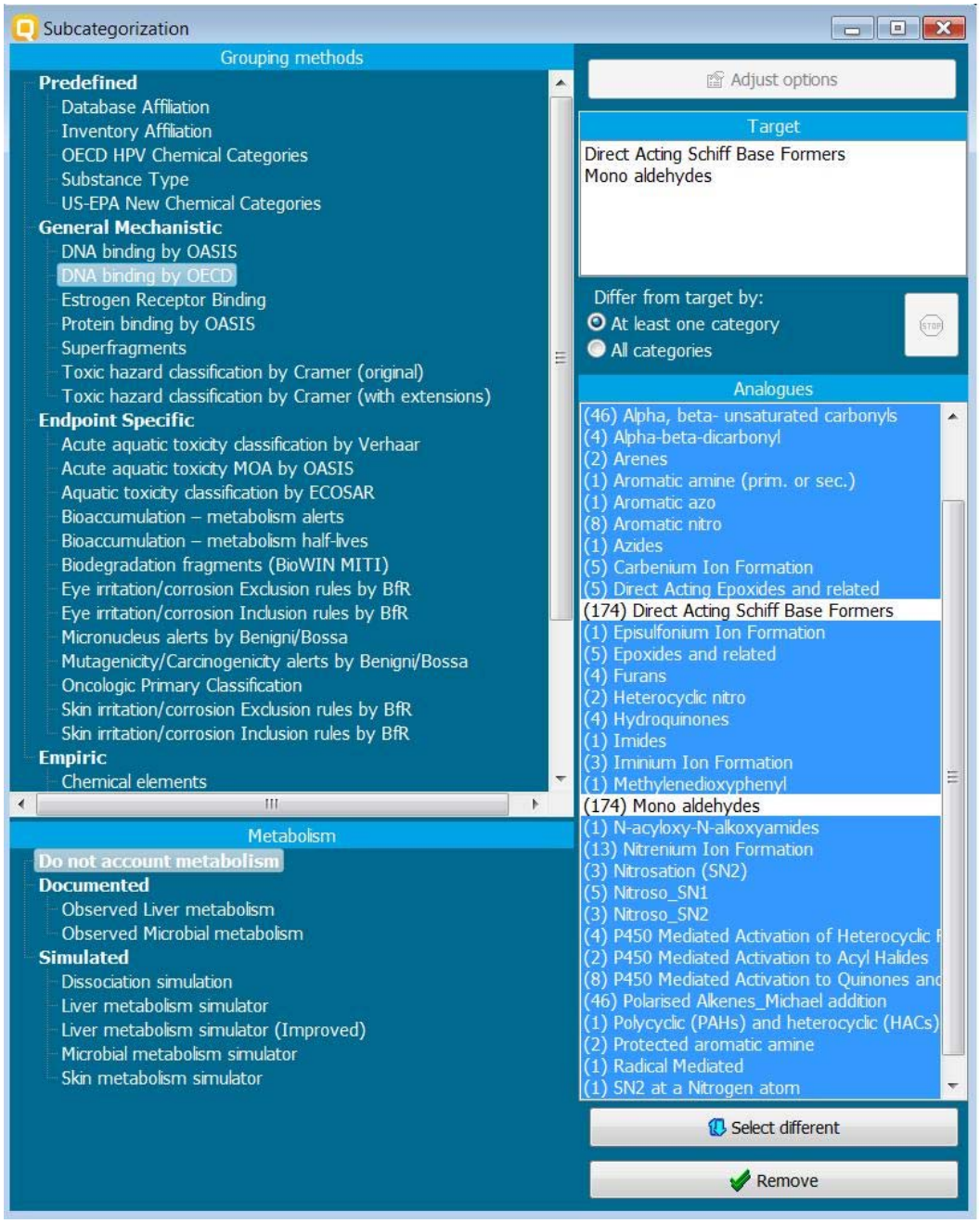
6.1 Mutagenicity

This discussion relates to ethanal for the development of a chemical category suitable for data gap filling for Ames mutagenicity. The results of profiling using the 'DNA binding by OECD', 'DNA binding by OASIS' and 'mutagenicity/carcinogenicity alerts by Benigni/Bossa' profilers indicate that ethanal is a potential Schiff base former. This information can be used to develop an initial category using the 'DNA binding by OECD' profiler to select category analogues from the two applicable databases available in the OECD QSAR Toolbox (Carcinogenicity & Mutagenicity ISSCAN and Genotoxicity OASIS databases). Doing so results in an initial category consisting of 175 chemicals.

The following sub-categorisations are then required to ensure the category contains analogues acting via a single mechanism of action:

- Sub-categorisation of the initial category of 175 chemicals using the 'DNA binding by OECD' profiler (the profiler that was used to develop the initial category). Figure 6.1 highlights the additional electrophilic mechanisms that are present in the initial category. This sub-categorisation results in a category of 86 chemicals.
- Sub-categorisation of the category of 86 chemicals using the 'DNA binding by OASIS' profiler. This sub-categorisation results in a category of 19 chemicals.
- Sub-categorisation of the category of 19 chemicals using the 'mutagenicity/ carcinogenicity alerts by Benigni/Bossa' profiler. This sub-categorisation results in a category of 18 chemicals.

Figure 6.1 Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OECD' profiler when sub-categorising the initial category formed for the target chemical ethanal.



6.2 Chromosomal aberration

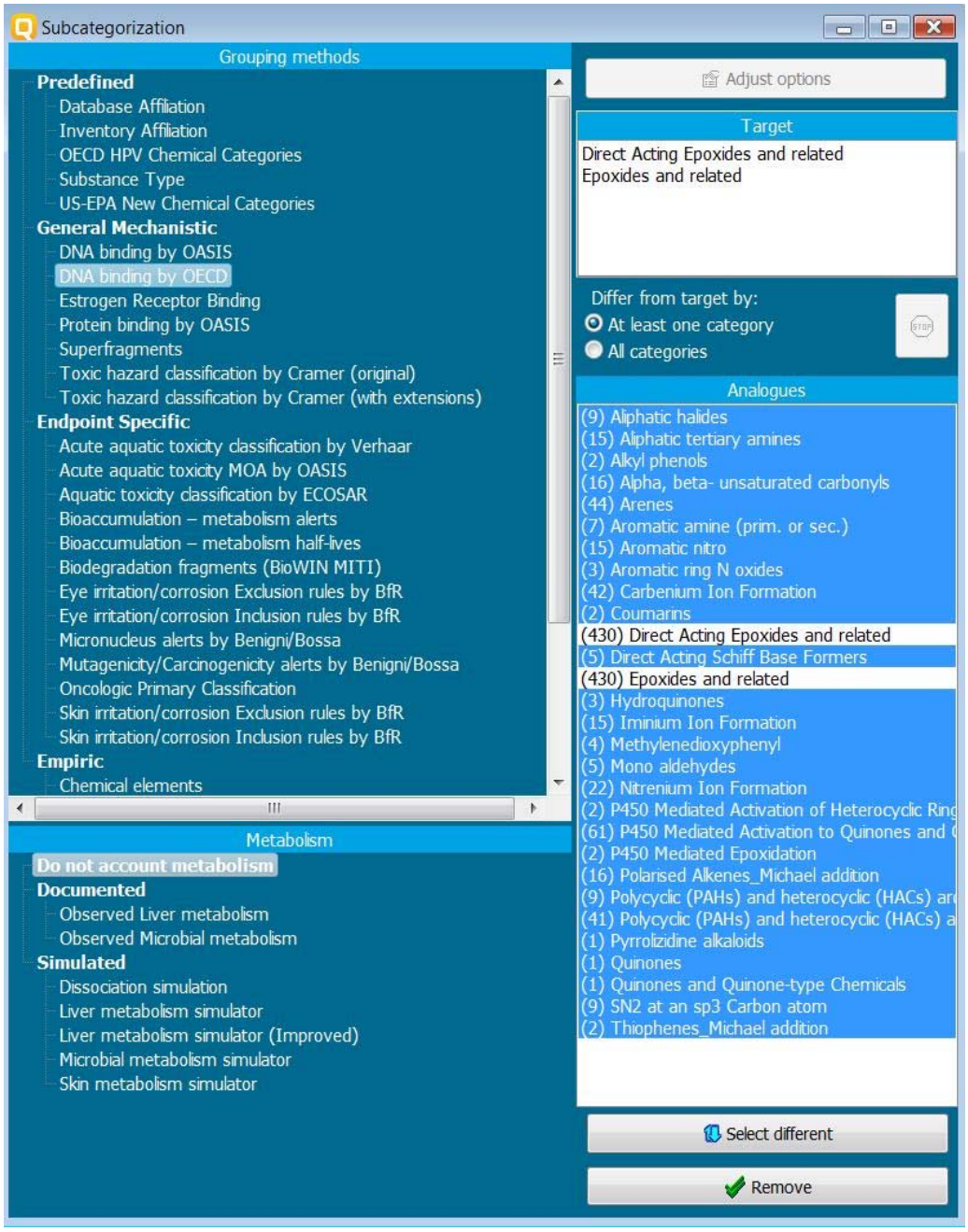
An analogous category formation and sub-categorisation process using the primary profilers can also be carried out for chromosomal aberration using ethylene oxide as the

target chemical. Four primary profilers are relevant for this endpoint (both *in vitro* and *in vivo*) these being: 'DNA binding by OECD', 'DNA binding by OASIS', 'protein binding by OASIS' and 'micronucleus alerts by Benigni/Bossa'. The four primary profilers are in agreement that the presence of the epoxide moiety in ethylene oxide is the most likely route to toxicity. This mechanistic information can be used to select analogues from the three relevant databases (genotoxicity OASIS, micronucleus ISSMIC and micronucleus OASIS) to create an initial category using the 'DNA binding by OECD' profiler. This results in an initial category of 431 chemicals.

The following sub-categorisations are then required to ensure the category contains analogues acting via a single mechanism of action:

- Sub-categorisation of the initial category of 431 chemicals using the 'DNA binding by OECD' profiler (the profiler that was used to develop the initial category). Figure 6.2 highlights the additional electrophilic mechanisms that are present in the initial category. This sub-categorisation results in a category of 270 chemicals.
- Sub-categorisation of the category of 270 chemicals using the 'DNA binding by OASIS' profiler. This sub-categorisation results in a category of 234 chemicals.
- Sub-categorisation of the category of 234 chemicals using the 'protein binding by OASIS' profiler. This sub0categorisation results in a category of 224 chemicals.
- Sub-categorisation of the category of 224 chemicals using the 'micronucleus alerts by Benigni/Bossa' profiler. This sub-categorisation results in a category of 123 chemicals.

Figure 6.2 Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OECD' profiler when sub-categorising the initial category formed for the target chemical ethylene oxide.



6.3 Genotoxic carcinogenicity

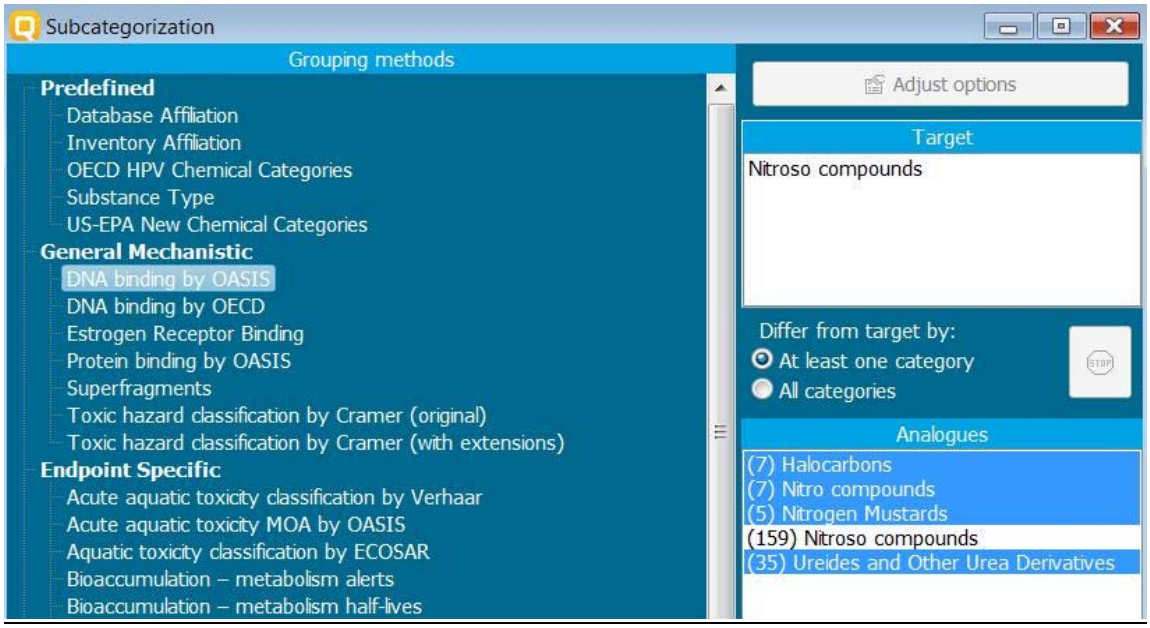
The formation of an initial category and then the subsequent sub-categorisation process for genotoxic carcinogenicity can be achieved using four primary profilers, these being:

'DNA binding by OECD', 'DNA binding by OASIS', 'mutagenicity/carcinogenicity alerts by Benigni/Bossa' and 'OncoLogic primary classification' profilers. Using N-nitrosodimethylamine to illustrate this point, an initial category can be created by selecting analogues from the Carcinogenic Potency DataBase CPDB and Carcinogenicity & Mutagenicity ISSCAN database using the 'DNA binding by OASIS' profiler based on the presence of the nitrosation mechanism (and is supported by the other three profilers). An initial chemical category can be developed consisting of 160 chemicals.

The following sub-categorisations are then required to ensure the category contains analogues acting via a single mechanism of action:

- Sub-categorisation of the initial category of 160 chemicals using the 'DNA binding by OASIS' profiler (the profiler that was used to develop the initial category). Figure 6.3 highlights the additional electrophilic mechanisms that are present in the initial category. This sub-categorisation results in a category of 118 chemicals.
- Sub-categorisation of the category of 118 chemicals using the 'DNA binding by OECD' profiler. This sub-categorisation results in a category of 92 chemicals.
- Sub-categorisation of the category of 92 chemicals using the 'mutagenicity/carcinogenicity alerts by Benigni/Bossa' profiler. This sub-categorisation results in a category of 91 chemicals.
- Sub-categorisation of the category of 91 chemicals using the 'OncoLogic primary classification' profiler. This sub-categorisation results in a category of 69 chemicals.

Figure 6.3 Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OASIS' profiler when sub-categorising the initial category formed for the target chemical N-nitrosodimethylamine.



6.4 General conclusions regarding the sub-categorisation with the primary profilers

The examples discussed for the endpoints covered by this guidance document highlight the importance of performing a sub-categorisation with the applicable primary profilers. Such sub-categorisations are important to ensure that the resulting categories consist of chemicals acting via a single mechanism of action related to covalent DNA binding. In addition, the sub-categorisations carried out using the endpoint specific primary profilers ensure that only analogues that contain the same structural alerts present in the target chemical are included in the category. This sub-categorisation should be carried out in a stepwise manner starting with the mechanistic profilers and then using the relevant endpoint specific profilers to refine the chemical category further. This will ensure a single mechanism of action is applicable for all category members, which is the main aim of this sub-categorisation process. This approach to sub-categorisation is in keeping with the OECD Guidance on the Grouping of Chemicals (*23*)).

7 Profiling examples which result in the ability to fill data gaps through read across

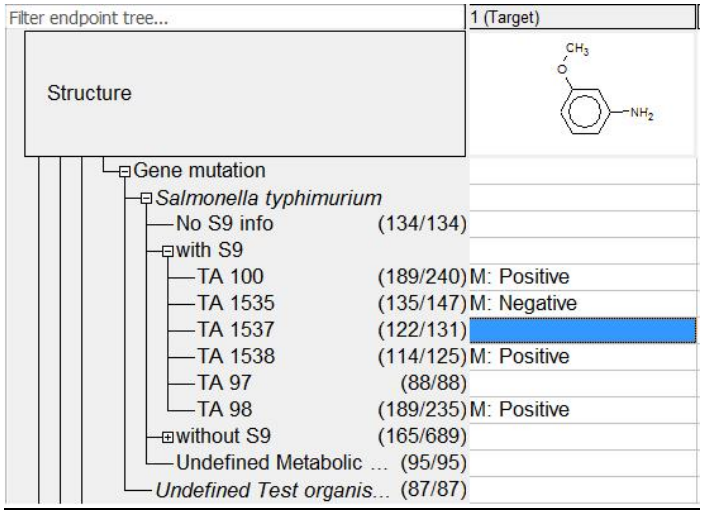
This section details how to use the primary and secondary profilers to build a chemical category for a chemical with a single electrophilic mechanism and for a chemical with two possible electrophilic mechanisms. It is important to realise that the chemicals that have been chosen to illustrate a suggested method for the formation of chemical categories. No attempt has been made to validate the resulting read across predictions by investigating additional data sources not available in V2 of the OECD QSAR Toolbox.

7.1 Profiling and data gap filling for Ames mutagencity for 3-methoxyaniline

This section will outline how to profile 3-methoxyaniline in order to build chemical categories suitable for read across enabling data gaps to be filled for the Ames mutagenicity assay. The following example assumes the user is familiar with the workflow of the OECD QSAR Toolbox. Thus, multiple steps and keystrokes in the workflows are omitted with only key screenshots being included. All of the profiling steps detailed should be carried out with the 'differ from target by' option set to 'at least one category' unless otherwise stated.

The following section details the steps that are required in order to fill the data gap that is present for the TA-1537 strain of *Salmonella* in the Ames mutagenicity assay (Figure 7.1). The intended usage is that one can follow the instructions presented in a stepwise manner allowing an identical category to be produced. It is important that the user is familiar with the general approach to category formation within the OECD QSAR Toolbox (those users who are not familiar should see the guidance material (*1*)).

Figure 7.1 Data gap present (shown in grey) for Ames mutagenicity data for 3-methoxyaniline.

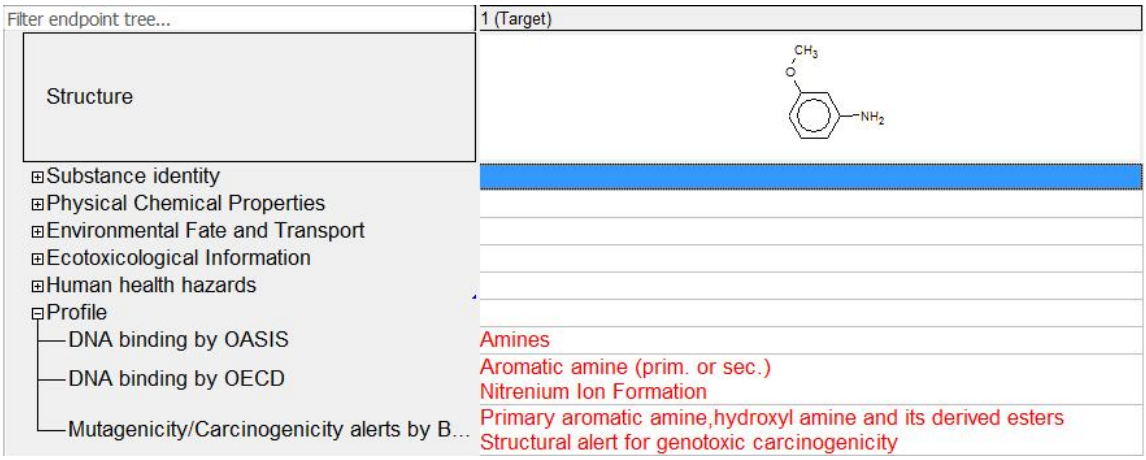


7.1.1 Initial profiling using the primary profilers

The first step in the development of a chemical category for 3-methoxyaniline is to profile it using the primary profilers applicable to Ames mutagenicity (Table 4.1 in section 4.4). The profiling results for 3-methoxyaniline show both mechanistic profilers are in agreement; the presence of an aromatic amine moiety is key structural feature and hence the most likely mechanism for covalent DNA binding (Figure 7.2). Inspection of the meta data accompanying the 'DNA binding by OECD' profiler shows that aromatic amines can be metabolised into reactive nitrenium ions capable of binding to DNA. This mechanism is supported by meta data accompanying 'DNA binding by OASIS' which details a number of possible DNA adducts formed due to this mechanism. Given the agreement between the two mechanistic profilers, one can be confident that the most likely mechanism that could lead to covalent DNA binding is as a result of nitrenium ion formation due to the presence of an aromatic amine moiety. This mechanistic information is reinforced by the identification of a structural alert for aromatic amines when profiling using the 'mutagenicity/carcinogenicity alerts by Benigni/Bossa' profiler. It is important to re-state that the presence of the structural alert does not indicate that 3-

methoxyaniline is mutagenic (the OECD QSAR Toolbox is not an expert system in this sense). Instead, the presence of such an alert is an indication that chemicals containing an aromatic amine moiety have been identified as being mutagenic in at least one strain of *Salmonella* in the Ames assay.

Figure 7.2 Profiling results from the initial battery of profilers applicable to Ames mutagenicity for 3-methoxyaniline.



7.1.2 Initial category formation and sub-categorisation using the primary profilers

The initial profiling results indicated that the most likely molecular initiating event for covalent DNA binding for 3-methoxyaniline was nitrenium ion formation due to the presence of an aromatic amine. This information was used to retrieve mechanistic analogues from the applicable databases within the OECD QSAR Toolbox. The relevant databases to search are:

- Carcinogenicity & Mutagenicity ISSCAN
- Genotoxicity OASIS

Profiling these two databases with the 'DNA binding by OECD' profiler creates an initial category of 810 chemicals. As discussed in sections 6.0 and 6.1 this initial category requires sub-categorisation in order to ensure that it contains only analogues that via the

same mechanism as 3-methoxyaniline. Thus, the following sub-categorisations are required:

1. Sub-categorisation of the initial category of 810 chemicals using the 'DNA binding by OECD' profiler (the profiler that was used to develop the initial category). This sub-categorisation results in a category of 472 chemicals.

2. Sub-categorisation of the category of 472 chemicals using the 'DNA binding by OASIS' profiler. This sub-categorisation results in a category of 177 chemicals.

3. Sub-categorisation of the category of 177 chemicals using the 'mutagenicity/ carcinogenicity alerts by Benigni/Bossa' profiler. This sub-categorisation results in a category of 152 chemicals.

7.1.3 Empiric sub-categorisation using the secondary profilers

The final stage in the development of a robust chemical category suitable for data gap filling is to ensure that the structural domain is well defined. One method to do this involves sub-categorising using a combination of the empiric profilers (removing all chemicals from the category that contain elements and functional groups not present in 3-methoxyaniline). This sub-categorisation process is analogous to that carried using the primary profilers. The following sub-categorisations should be carried out on the category of 152 chemicals generated in section 7.2:

1. Sub-categorisation of the category of 152 chemicals using the 'organic functional group' profiler. This sub-categorisation results in a category of 22 chemicals.

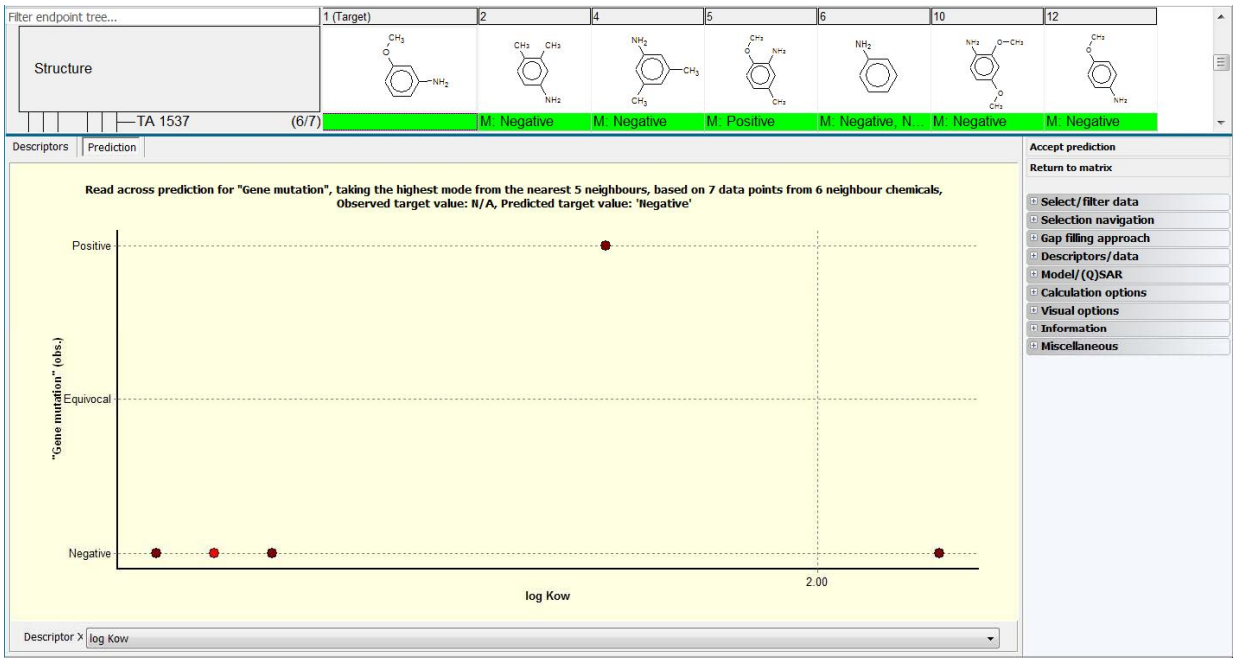
2. Sub-categorisation of the category of 22 chemicals using the 'chemical elements' profiler. This sub-categorisation results in a category of 19 chemicals.

7.1.4 Data-gap filling via read across

The sub-categorisation carried using the primary and secondary profilers results in a category that has a well defined mechanistic (defined as a result of the sub-categorisation in section 7.1.2) and structural (defined in section 7.1.2) domains. This category can now be used to fill the data gap that is present for 3-methoxyaniline in the TA 1537 strain of *Salmonella* in the presence of the S9 liver fraction. Inspection of the nearest five chemicals (in this example using hydrophobicity) in the category shows four

of them to be negative in TA 1537 strain and one to be positive (Figure 7.3). It is interesting to note the positive result for 2-methoxy-5-methyl-aniline (chemical 5 in Figure 7.3). It is possible that this chemical might be a false positive in the TA 1537 strain of *Salmonella*. This possibility is supported by the negative TA 1537 result for the related chemical, 2,4-dimethoxyaniline (chemical 10 in Figure 7.3). The identification of potential outlying chemicals is one of the advantages of the category approach. The weight of evidence based read across prediction based on the five analogues shown in Figure 7.3 suggest that 3-methoxyaniline would be negative if tested in the TA 1537 strain of *Salmonella*.

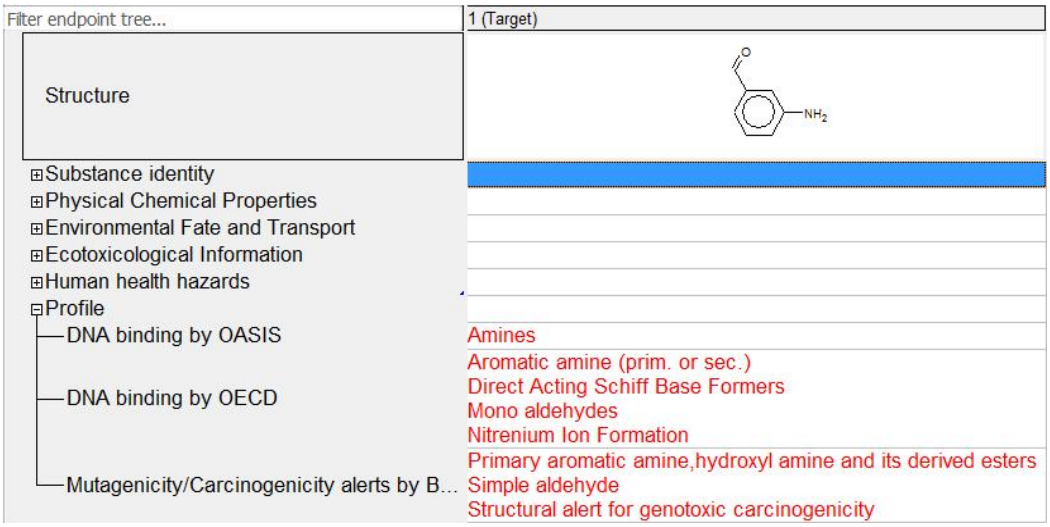
Figure 7.3 Read across prediction made for the Ames mutagenicity for 3-methoxyaniline in the TA 1537 strain of Salmonella in the presence of the S9 liver fraction.



7.2 Profiling and data gap filling for a chemical with more than a single covalent mechanism, 3-aminobenzaldhyde, for Ames mutagenicity

In the previous example, the target chemical contained only a single reactive site susceptible to covalent bond formation. It is sometimes the case that a target chemical may contain more than a single reactive site and thus more than one possible covalent mechanism of action is possible. As an example, consider the profiling results for 3-aminobenzaldehyde (Figure 7.4). The 'DNA binding' by OECD profiler suggests two possible mechanism of action, nitrenium ion formation due to the aromatic amine and Schiff base formation due to the mono-aldehyde. In contrast, the 'DNA binding by OASIS' profiler indicates a single mechanism due to the aromatic amine. This example investigates the Ames mutagenicity of this chemical (although it should be remembered that it is the approach to dealing with chemicals of this type that is important and not the actually read across prediction). Endpoint profiling using the 'mutagenicity/carcinogenicity alerts by Benigni/Bossa' profiler adds weight to the possibility of two mechanisms of action leading to mutagenicity for this chemical due to the presence of two corresponding structural alerts. It is important to recall that a robust and defensible chemical category should encapsulate an area of chemical space relating to a single mechanism of action, with a well defined structural domain. Thus, for chemicals such as 3-aminobenzaldehyde, one approach is to build two separate categories, one for each mechanism of action. The data gap can then be filled by (for example) read across using both categories. Using the precautionary principle, the predicted worst case scenario would be taken i.e. the more toxic. The following sections cover the development of categories for Ames mutagenicity for the two mechanisms identified for 3-aminobenzaldehyde.

Figure 7.4 Mechanistic profiling for the target chemical 3-amino-benzaldehyde using the 'DNA binding by OECD' profiler in the OECD QSAR Toolbox V2.0.



7.2.1 Initial category formation and sub-categorisation using the primary profilers - Nitrenium ion formation

An initial category can be created by profiling the Carcinogenicity & Mutagenicity ISSCAN and genotoxic OASIS databases using the DNA binding by OECD profiler. It is important to de-select the additional Schiff base mechanism before profiling the two databases for analogues. This ensures that the focus of the initial category is towards nitrenium ion formation due to an aromatic amine. This is achieved by de-selecting the Schiff base mechanisms as shown in Figure 7.5 (select the mechanisms as shown and then click the down arrow). This creates an initial category of 811 chemicals.

As with the previous example this initial category requires a series of sub-categorisations to remove chemicals with additional mechanisms and structural alerts. The following sub-categorisations are required:

1. Sub-categorisation of the initial category of 811 chemicals using the 'DNA binding by OECD' profiler (the profiler that was used to develop the initial category). Importantly, the Schiff base mechanisms must be selected in this sub-categorisation (see Figure 7.6). This sub-categorisation results in a category of 474 chemicals.
2. Sub-categorisation of the category of 474 chemicals using the 'DNA binding by OASIS' profiler. This sub-categorisation results in a category of 179 chemicals.

3. Sub-categorisation of the category of 179 chemicals using the 'mutagenicity/ carcinogenicity alerts by Benigni/Bossa' profiler. This sub-categorisation results in a category of 154 chemicals.

Figure 7.5 Initial mechanistic profiling – mechanisms in highlighted are removed from the categorisation process by selecting the down arrow.

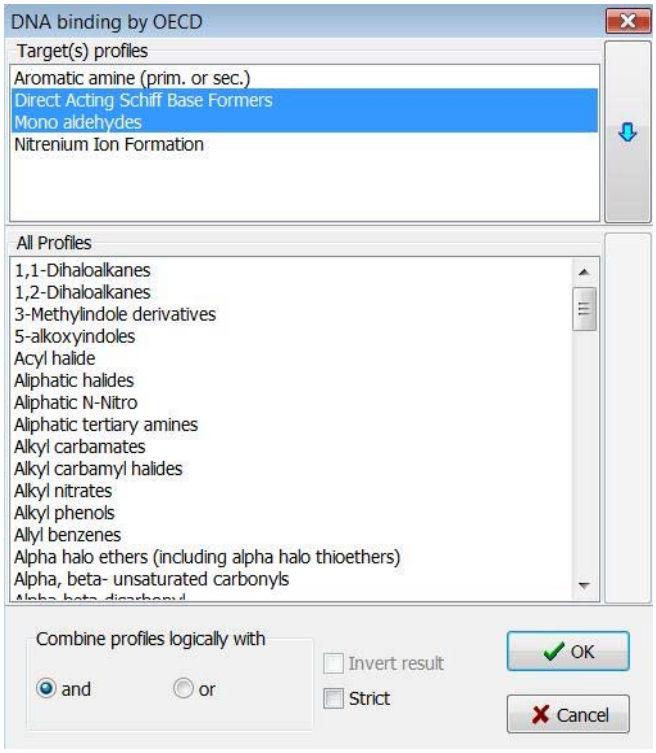
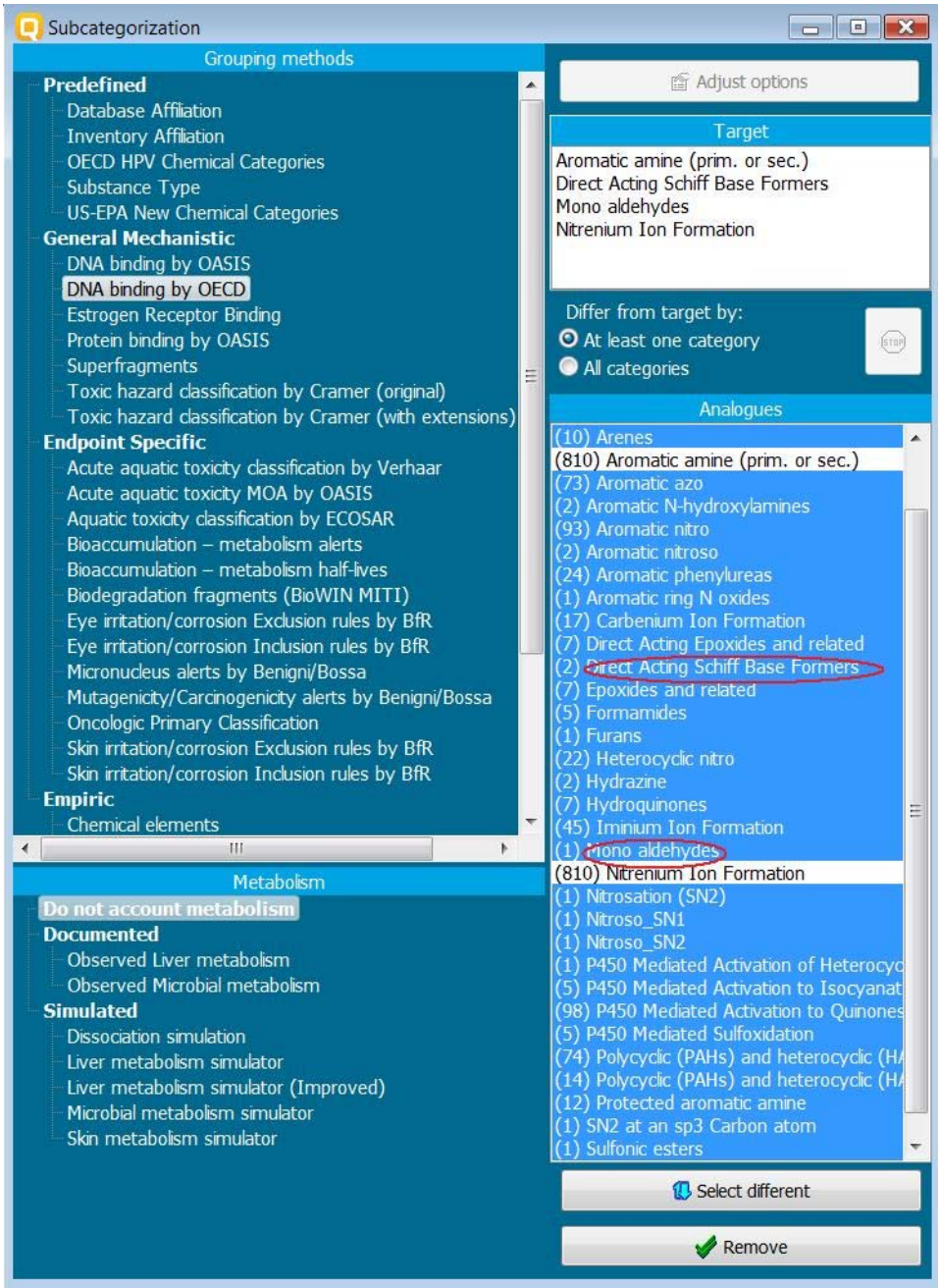


Figure 7.6 Selection of Schiff base mechanisms when sub-categorising using the 'DNA binding by OECD' profiler. The mechanisms relating to Schiff base formation are circled in red and should be selected when performing the sub-categorisation process.



7.2.2 Empiric sub-categorisation using the secondary profilers

The final steps in the sub-categorisation process are the use of the empiric profilers to define the structural domain of the category in terms of the target chemical. Two sub-categorisations of the category of 154 chemicals generated in section 8.1 are required:

1. Sub-categorisation of the category of 154 chemicals using the 'organic functional group' profiler. This sub-categorisation results in a category of 7 chemicals.
2. Sub-categorisation of the category of seven chemicals using the 'chemical elements' profiler. This sub-categorisation results in a category of six chemicals.

Inspection of the data matrix for the five member category shows it to contain for data the TA 100 strain of *Salmonella* for only a single chemical (Figure 7.7). This is a less than ideal situation as any read across prediction would be made based on a one-to-one basis rather than many-to-one. In cases such as this it is useful to return to the sub-categorisation using the 'organic functional group' profiler and include additional simple functional groups (see section 5.1 for a more detailed discussion). Repeating the sub-categorisation (steps 1 and 2 above) including methyl groups within the structural domain results in a category of 13 chemicals. This category has data for three chemicals in TA 100 strain of *Salmonella* (Figure 7.8).

Figure 7.7 Data matrix for the five member category formed for nitrenium ion formation.

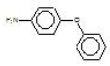
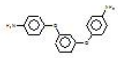
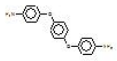
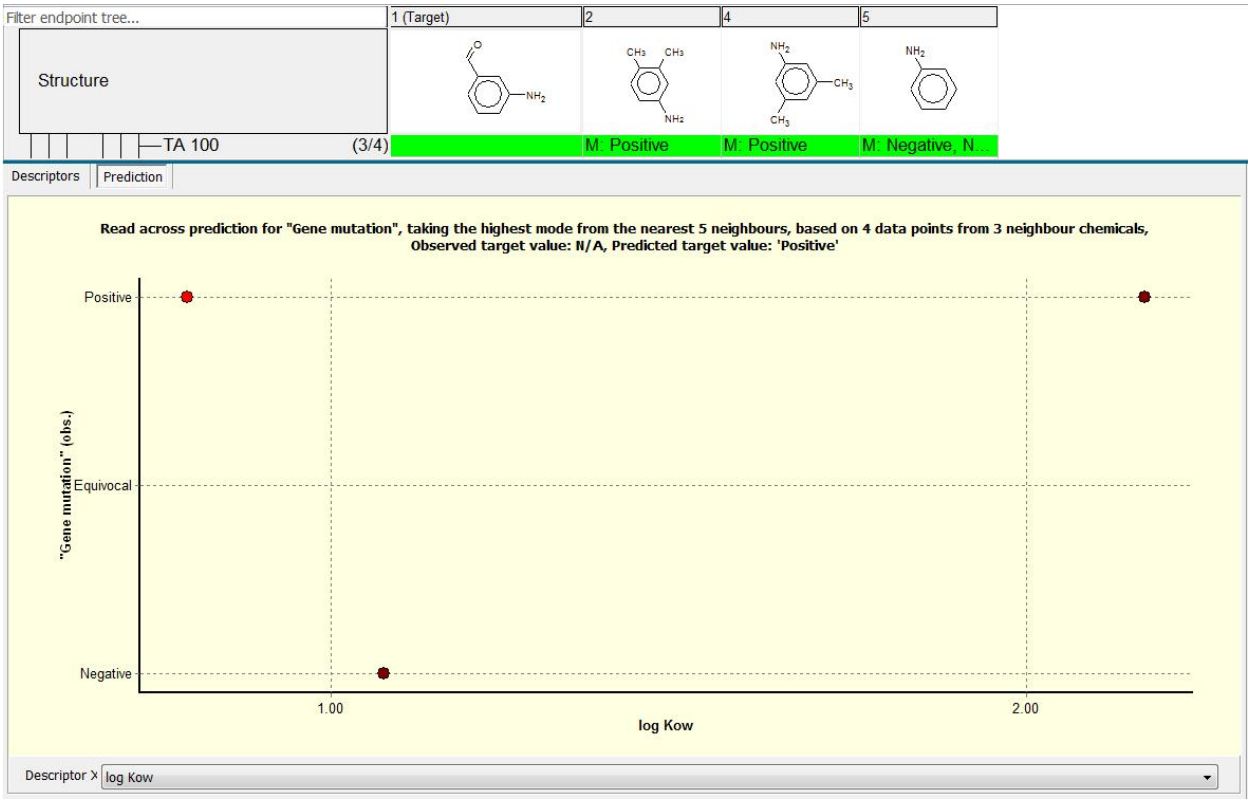
Filter endpoint tree...	1 (Target)	2	3	4	5
Structure					
⊞Substance identity					
⊞Physical Chemical Properties					
⊞Environmental Fate and Transport					
⊞Ecotoxicological Information					
⊟Human health hazards					
—Acute Toxicity					
—⊞Carcinogenicity					
—⊞Developmental toxicity / teratogenicity					
—⊟Genetic Toxicity					
—⊟in vitro					
—⊟bacterial reverse mutation assay (e....					
—⊟Gene mutation					
—⊟*Salmonella typhimurium*					
—No S9 info					
—⊟with S9					
—TA 100 (1/2)		M: Negative, N...			
—TA 1535 (1/2)		M: Negative, N...			
—TA 1537 (1/2)		M: Negative, N...			
—TA 1538 (1/1)		M: Negative			
—TA 97 (1/1)		M: Negative			
—TA 98 (1/2)		M: Negative, N...			
—⊞without S9 (1/10)		M: Negative, N...			
—Undefined Metabolic activation					
—*Undefined Test organisms (spe...*					

Figure 7.8 Data matrix including methyl substituents in the structural domain.



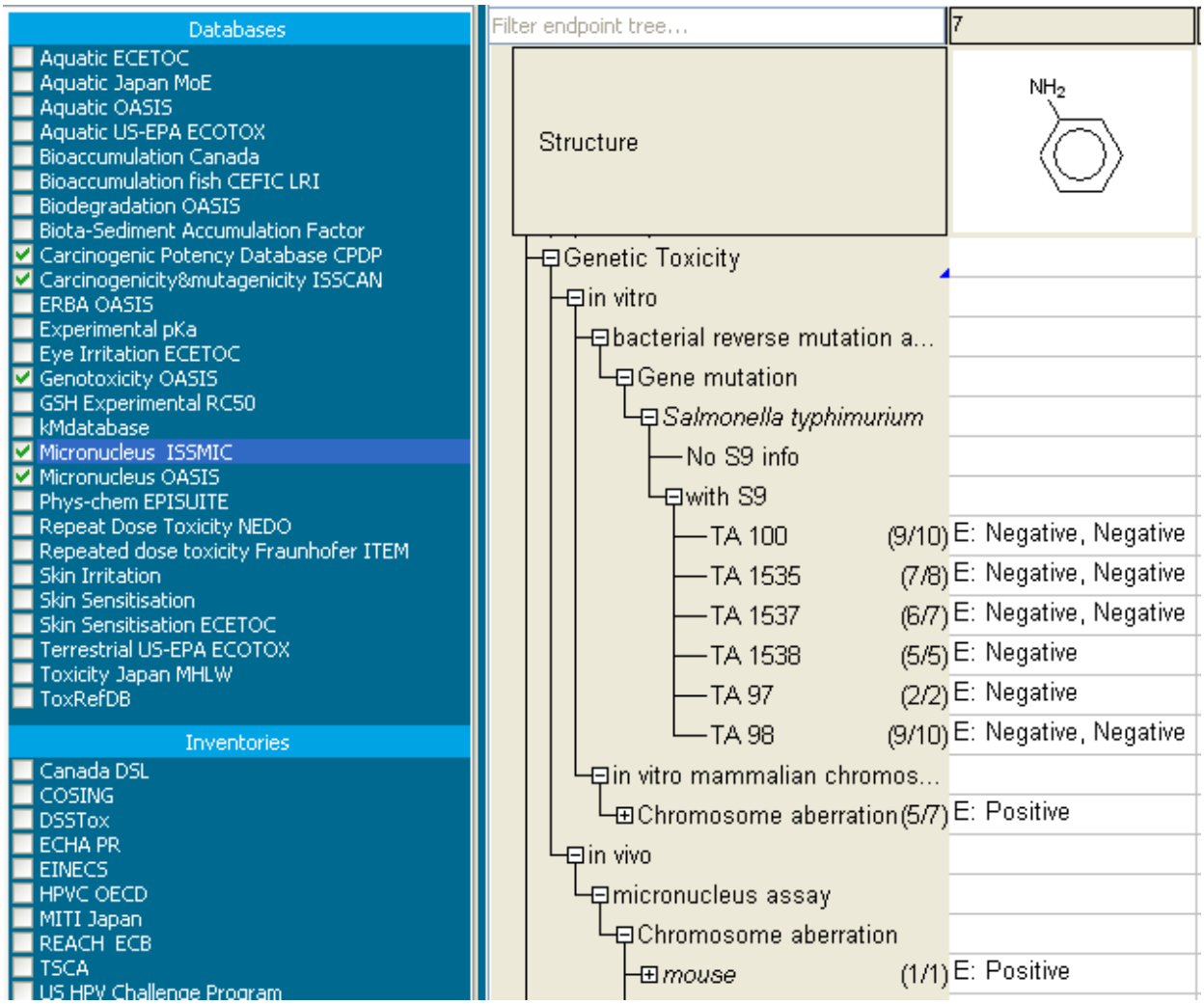
7.2.3 Data gap filling using read across for 3-aminobenzaldehyde for the nitrenium ion category

The three analogues identified in section 8.2 can now be used to predict the toxicity of 3-aminobenzaldehdye via read across. This read across prediction is for toxicity due to the presence of the amine function group that can bind to DNA via the formation of a nitrenium ion. The data matrix (Figure 7.8) indicates that two of the chemicals are positive in the TA 100 strain of *Salmonella* and the third was negative. This results in a positive read across prediction for 3-aminobenzaldehyde based on the available weight of evidence.

It is interesting to note the negative test data in the TA 100 strain of *Salmonella* for aniline (chemical 5 in Figure 7.8). This appears to suggest that this chemical is an outlier compared to the other two category members. It is likely that such conflicting data will

occur in numerous categories developed for genotoxicity and thus an analysis of toxicological data in categories formed for related endpoints will help resolve such conflicts and add weight to the read across predictions. In this example, inspection of data within the OECD QSAR Toolbox for *in vitro* and *in vivo* chromosomal aberration data suggests that aniline is genotoxic (Figure 7.9). This helps to add to confidence to the read across predictions made for 3-aminobenzaldehyde.

Figure 7.9 Additional in vitro and in vivo chromosomal aberration data showing genotoxicity for aniline.



7.2.4 Initial category formation and sub-categorisation using the primary profilers - Schiff base formation due to the presence of a mono-aldehyde

The same process can be undertaken for the second possible mechanism present in 3-aminobenzaldehyde. Profiling the Carcinogenicity & Mutagenicity ISSCAN and genotoxic OASIS databases for analogues capable of covalently binding to DNA via Schiff base formation using the 'DNA binding by OECD' profiler (deselecting the nitrenium ion formation mechanisms – Figure 7.10) results in an initial category of 176 chemicals.

This initial category requires a series of sub-categorisations to remove chemicals with additional mechanisms and structural alerts. The following two sub-categorisations are required (note that the sub-categorisation process is not carried out using the 'DNA binding by OASIS' profiler as this profiler did not identify a Schiff base mechanism in the initial profiling):

1. Sub-categorisation of the initial category of 176 chemicals using the 'DNA binding by OECD' profiler (the profiler that was used to develop the initial category). Importantly, the nitrenium ion formation mechanisms must be selected in this sub-categorisation (see Figure 7.11). This sub-categorisation results in a category of 87chemicals.

2. Sub-categorisation of the category of 87 chemicals using the 'mutagenicity/ carcinogenicity alerts by Benigni/Bossa' profiler. This sub-categorisation results in a category of 57 chemicals.

Figure 7.10 Initial mechanistic profiling – highlighted mechanisms are removed from the categorisation process by selecting the down arrow.

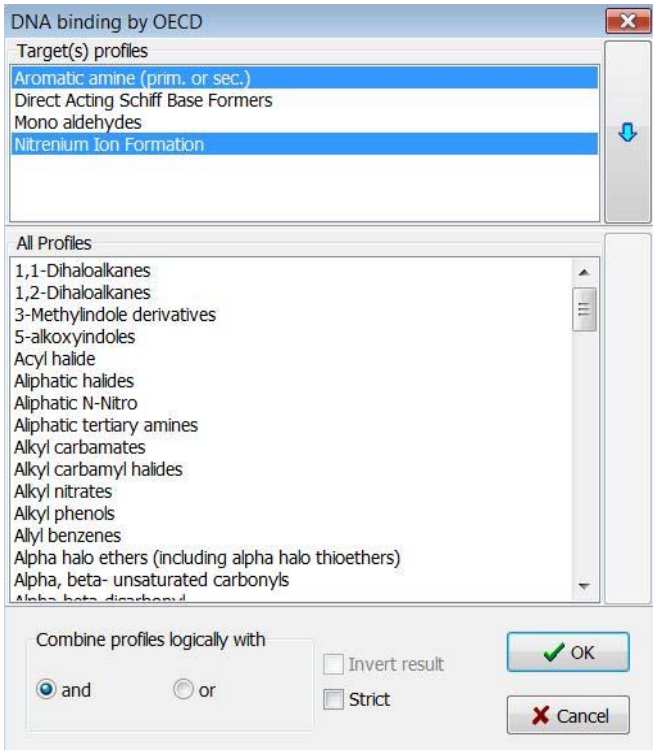
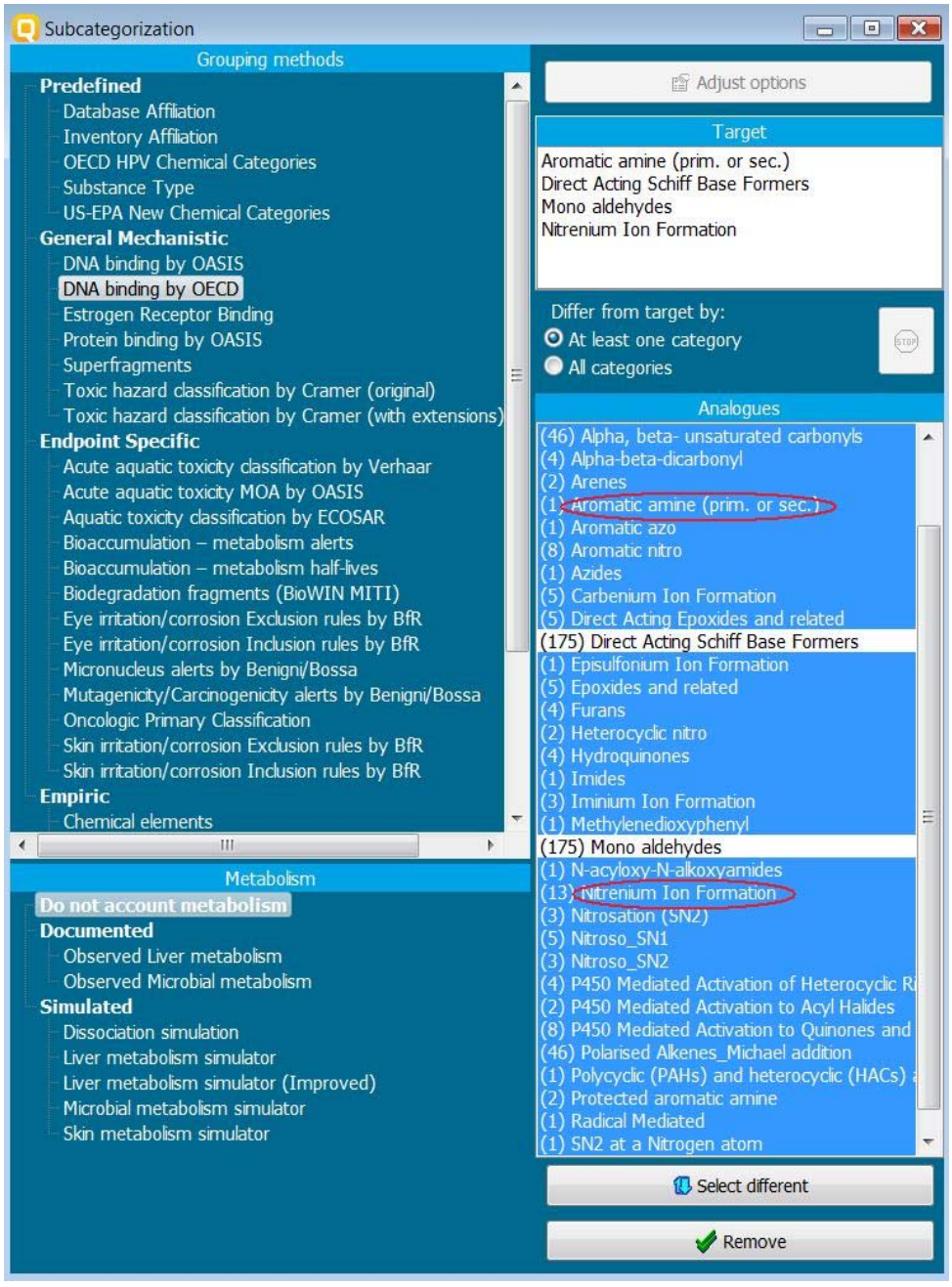


Figure 7.11 Selection of nitrenium ion formation mechanisms when sub-categorising using the 'DNA binding by OECD' profiler. The mechanisms relating to nitrenium ion formation are circled in red and should be selected when performing the sub-categorisation process.



7.2.5 Empiric sub-categorisation using the secondary profilers

The final steps in the sub-categorisation process are the use of the empiric profilers in order to define the structural domain of the category in terms of the target chemical.

Two sub-categorisations of the category of 57 chemicals generated in section 7.2.4 are required:

1. Sub-categorisation of the category of 57 chemicals using the 'organic functional group' profiler. This sub-categorisation results in a category of two chemicals.
2. Sub-categorisation of the category of two chemicals using the 'chemical elements' profiler. This sub-categorisation does not identify any chemicals for removal and thus the two chemicals remain.

Inspection of the data matrix for the category containing two chemicals shows it to contain data for the TA 100 strain of *Salmonella* for only a single chemical (Figure 7.12). As previously, this is a less than ideal situation as any read across prediction would be made based on a one-to-one basis rather than many-to-one. In cases such as this it is useful to return to the sub-categorisation using the 'organic functional group' profiler and include additional simple functional groups (see section 5.1 for a more detailed discussion). Repeating the sub-categorisation (steps 1 and 2 above) including methyl groups within the structural domain results in a category of six chemicals. This category has data for four chemicals for the *Salmonella* TA 100 strain. However, the resulting category consists of aromatic and non-aromatic aldehydes, given that the target chemical is an aromatic aldehyde it is a good idea to remove the non-aromatic aldehydes from the category. This is done on a chemical by chemical basis in the trend analysis window (Figures 7.13 and 7.14 show the before and after categories). The inclusion of the additional analogue results in a category of three chemicals.

Figure 7.12 Data matrix formed after the initial secondary profiling. This results in only a single analogue being identified.

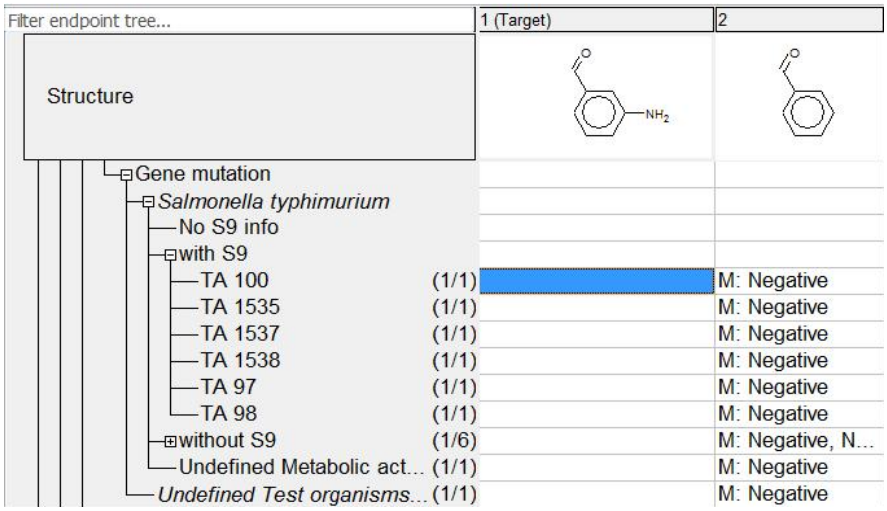


Figure 7.13 Data matrix formed after extending the structural domain to include chemicals with methyl groups.

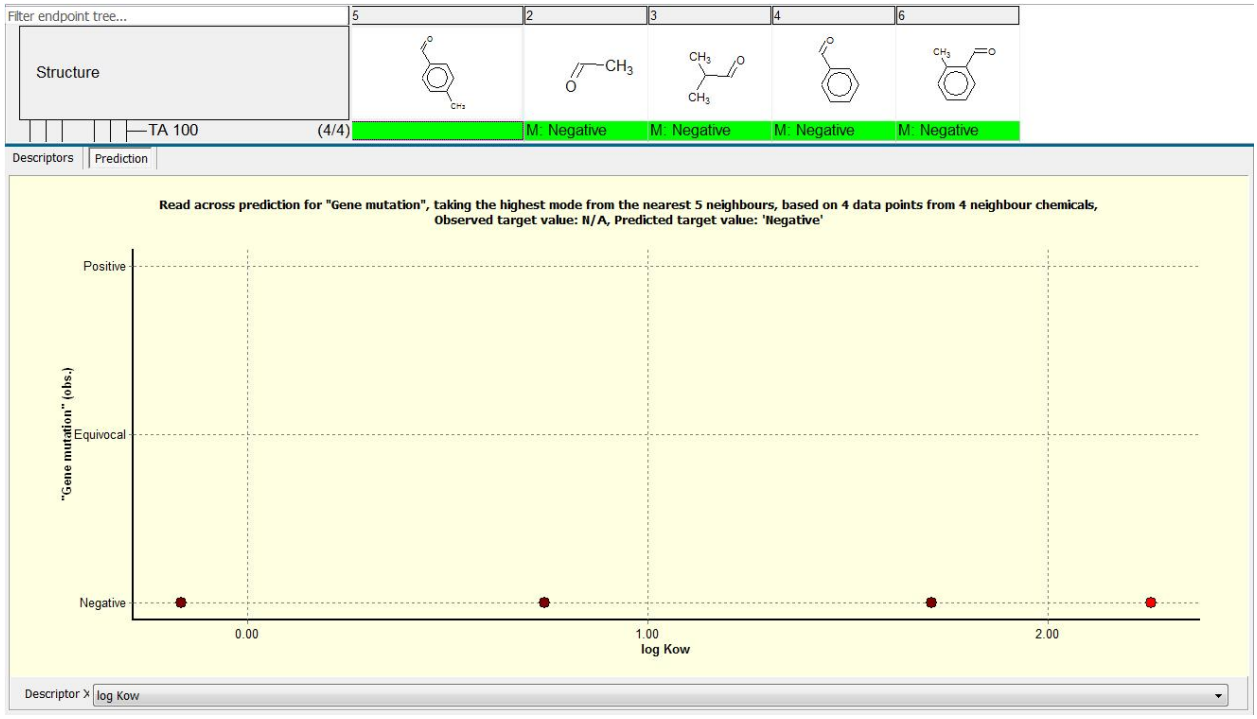
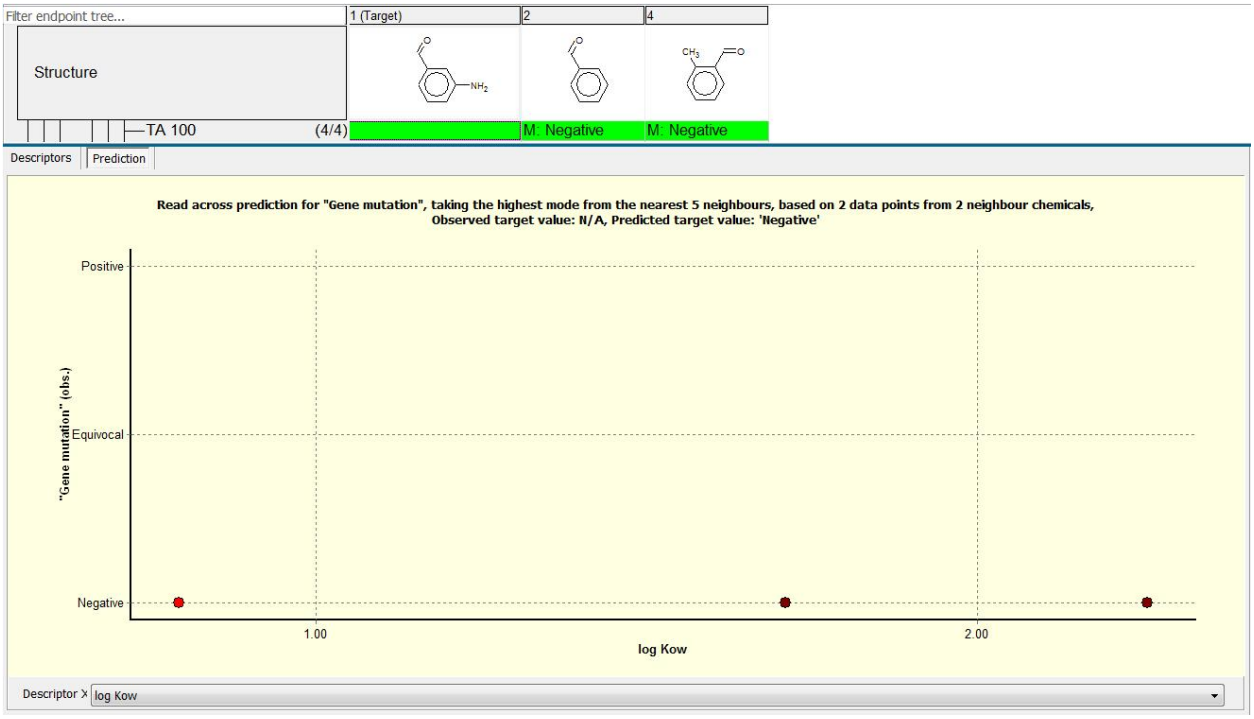


Figure 7.14 matrix formed after eliminating non-aromatic chemicals.



7.2.6 Data gap filling using read across for 3-aminobenzaldehyde for the nitrenium ion category

The two categories developed in section 8.5 allow two read across predictions to be made. The first is a one-to-one prediction that predicts 3-aminobenzaldehyde to be negative in *Salmonella* TA 100. This prediction is supported by the second category containing two analogues. This category allows a two-to-one prediction to be made, which is also negative.

7.2.7 Summary of category formation for 3-aminobenzaldehdye

The above category formation for 3-aminobenzaldehdye illustrates how multiple categories can be formed to fill data gaps for chemicals with more than a single mechanism for covalent DNA binding. In the case of 3-aminobenzaldehdye the categories formed suggest that genotoxicity in the Ames assay is likely via the formation of a nitrenium ion due to the presence of the aromatic amine moiety. This assessment of

Ames mutagenicity in *Salmonella* TA 100 is made on a worst case scenario basis. It is also important to note the use of additional structural analogues to increase the confidence in the initially small categories (those formed after the first definition of the structural domain). The ability to define the structural domain is also extremely important in the formation of a robust chemical category suitable for read across.

8 General approach for the development of chemical categories for genetic toxicity and genotoxic carcinogenicity

The following outline can be considered a good general approach for the development of chemical categories for genetic toxicity and genotoxic carcinogenicity. These instructions are summarised in a flow chart (Figure 8.1).

1. Profile the target chemical using the two mechanistic profilers; 'DNA binding by OECD' and 'DNA binding by OASIS'. Consistency between these two profilers is a good indication of a robust mechanism.

2. Sub-categorisation using a combination of the mechanistic profilers to eliminate chemicals that contain additional potential covalent mechanism of action. Bear in mind that a robust category is applicable to a single mechanism of action.

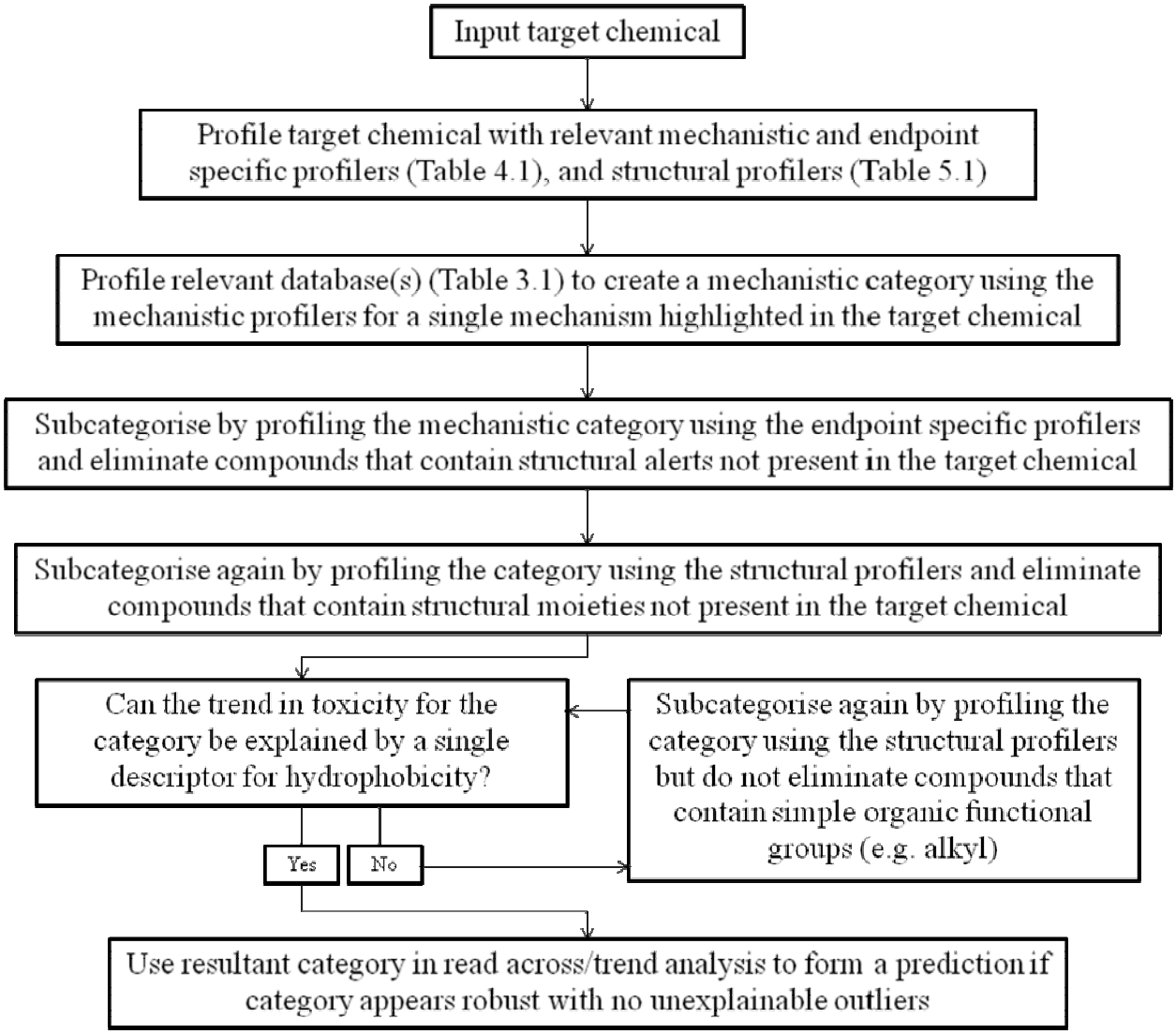
3. Sub-categorisation of the category using an appropriate endpoint specific profiler. These profilers identify known structural alerts related to endpoint specific toxicological data. Eliminate any chemicals that contain additional structural alerts that do not occur in the target chemical. In a robust chemical category the endpoint profiling will identify a structural feature in keeping with the mechanistic profiling results.

4. Sub-categorisation using the secondary profilers in order to define the structural domain. One should use a combination of the empiric profilers (it is recommended to use the organic functional group and chemical elements profilers in the majority of cases) to restrict the structural domain of the category so that it is similar to that of the target chemical. The guiding principle should be towards the descriptor that one will use in any subsequent read across or trend analysis. This helps keep any predictions made using read across or trend analysis as transparent as possible. It is worth recalling that sometimes this profiling step requires the inclusion of analogues containing simple organic functional groups that are not present in the target chemical.

5. Always ensure that the data used in any read across or trend analysis predictions are quality checked and that unusual or outlying data within a category are investigated before use. Please remember that the OECD is not responsible for the quality of the data within the OECD QSAR Toolbox.

6. Create the appropriate reporting format in the OECD QSAR Toolbox (see guidance available from (*1*)).

Figure 8.1 General scheme for category formation for genetic toxicity and genotoxic carcinogenicity.



9 References

(1) OECD. (2010) The OECD (Q)SAR Application Toolbox Guidance Documents available from http://www.oecd.org/document/54/0,3343,en_2649_34379_42923638_1_1_1_1,00.html#Guidance_Documents_and_Training_Materials_for_Using_the_Toolbox.

(2) Schultz, T. W. (2010) Adverse outcome pathways: A way of linking chemical structure to in vivo toxicological hazards, In In silico Toxicology: Principles and Applications (Cronin, M. T. D. and Madden, J. C., Eds.), Royal Society of Chemistry, Cambridge, UK.

(3) Benigni, R. and Bossa, C. (2008) Structure alerts for carcinogenicity, and the Salmonella assay system: A novel insight through the chemical relational databases technology. Mutation Research 659, 248-261.

(4) Benigni, R., Bossa, C., Tcheremenskaia, O. and Worth, A. (2009) Development of structural alerts for the in vivo micronucleus assay in rodents. http://ecb.jrc.ec.europa.eu/DOCUMENTS/QSAR/EUR_23844_EN.pdf.

(5) Enoch, S. J. and Cronin, M. T. D. (2010) A review of the electrophilic reaction chemistry involved in covalent DNA binding. Critical Reviews in Toxicology 40, 728-748.

(6) Ashby, J. and Tennant, R. W. (1988) Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. Mutation Research 204, 17-115.

(7) Ashby, J. and Tennant, R. W. (1991) Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemical tested by the US NTP. Mutation Research 257, 229-306.

(8) Woo, Y.-T. and Lai, D. Y. (2003) Mechanisms of action of chemical carcinogens and their role in structure-activity relationships (SAR) analysis and risk assessment, In Quantitative Structure-Activity Relationship (QSAR) Models of Mutagens and Carcinogens (Benigni, R., Ed.), CRC Press LLC.

(9) Ames, B. N., Garry, B. and Herzenberg, L. A. (1960) The genetic control of the enzymes of histidine biosysnthesis in Salmonella typhimurium. Journal of General Microbiology 22, 369-378.

(10) Mortelmans, K. and Zeiger, E. (2000) The Ames Salmonella/microsome mutagenicity assay. Mutation Research 455, 29-60.

(11) Kirsh-Volders, M., Elhajouji, A., Cundari, E. and Van Hummelen, P. (1997) The in vitro micronucleus test: a multi-endpoint assay to detect simultaneously mitotic delay, apoptosis, chromosome breakage, chromosome

loss and non-disjunction. Mutation Research 392, 19-30.

(12) Peto, R., Pike, M. C., Bernstein, L., Gold, L. S. and Ames, B. N. (1984) The TD_{50}: A proposed general convention for the numerical description of the carcinogenic potency of chemicals in chronic-exposure animal experiments. Environmental Health Perspectives 58, 1-8.

(13) Schultz, T. W., Carlson, R. E., Cronin, M. T. D., Hermens, J. L. M., Johnson, R., O'Brien, P. J., Roberts, D. W., Siraki, A., Wallace, K. B. and Veith, G. D. (2006) A conceptual framework for predicting the toxicity of reactive chemicals: modeling soft electrophilicity. SAR and QSAR in Environmental Research 17, 413-428.

(14) Aptula, A. O. and Roberts, D. W. (2006) Mechanistic applicability domains for nonanimal-based prediction of toxicological end points: General principles and application to reactive toxicity. Chemical Research in Toxicology 19, 1097-1105.

(15) Mekenyan, O., Dimitrov, S., Serafimova, R., Thompson, E., Kotov, S., Dimitrova, N. and Walker, J. D. (2004) Identification of the structural requirements for mutagenicity by incorporating molecular flexibility and metabolic activation of chemicals I: TA100 Model. Chemical Research in Toxicology 17, 753-766.

(16) Serafimova, R., Todorov, M., Pavolo, T., Kotov, S., Jacob, E., Aptula, A. and Mekenyan, O. (2007) Identification of the structural requirements for mutagencitiy, by incorporating molecular flexibility and metabolic activation of chemicals. II. General Ames mutagenicity model. Chemical Research in Toxicology 20, 662-676.

(17) Mekenyan, O., Todorov, M., Serafimova, R., Stoeva, S., Aptula, A., Finking, R. and Jacob, E. (2007) Identifying the structural requirements for chromosomal aberration by incorporating molecular flexibility and metabolic activation of chemicals. Chemical Research in Toxicology 20, 1927-1941.

(18) Dimitrov, S. D., Low, L. K., Patlewicz, G. Y., Kern, P. S., Dimitrova, G. D., Comber, M. H. I., Phillips, R. D., Niemela, J., Bailey, P. T. and Mekenyan, O. G. (2005) Skin sensitization: Modeling based on skin metabolism simulation and formation of protein conjugates. International Journal of Toxicology 24, 189-204.

(19) JRC. (2009) ToxTree. (available from: http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE).

(20) ISSCAN. (2009) ISSCAN database. (available from: http://www.iss.it/ampp/dati/cont.php%3Fid=233%26lang=1%26tipo=7).

(21) Woo, Y.-T. and Lai, D. Y. (2005) A mechanism-based expert system for predicting the carcinogenic potential of chemicals In Predictive Toxicology (Helma, C., Ed.) pp 385-413, Taylor and Francis, Boca Raton, Florida.

(22) Mekenyan, O., Dimitrov, S., Dimitrova, N., Dimitrova, G., Pavlov, T., Chankov, G., Kotov, S., Vasilev, K. and Vasilev, R. (2006) Metabolic activation of chemicals: in-silico simulation. SAR and QSAR in Environmental Research 17, 107-120.

(23) OECD. (2007) Guidance on Grouping of Chemicals. (available from: http://www.oecd.org/LongAbstract/0,3425,en_2649_34379_39404999_119829_1_1_37407,00.html).

OECD
2, rue André Pascal
75775 Paris Cedex 16
France
Tel.: +33 1 45 24 82 00
Fax: +33 1 45 24 85 00
ehscont@oecd.org