

OECD QSAR Toolbox v.4.2

Manipulation of data matrix and manual transferring of data to the target outside data gap filling module

Outlook

- **Background**
- Objectives
- Specific aim
- Manipulation of data matrix
- Example

Background

- This is a step-by-step presentation designed to introduce the user to the newly created functionalities for manipulation of data matrix.
- A simple example of read-across assessment where the new functionalities are applied is shown.

Outlook

- Background
- **Objectives**
- Specific aim
- Manipulation of data matrix
- Example

Objectives

This presentation demonstrates a number of functionalities of the Toolbox:

- Rearrangement of columns of data matrix;
- Filtering of parameters/experimental data or profiling results of the analogues within the category;
- Hide/show gap-filling chart while in Data gap filling module;
- Transferring of data to the target chemical outside Data gap filling module.

Outlook

- Background
- Objectives
- **Specific aim**
- Manipulation of data matrix
- Example

Aim

- To introduce and make the user familiar with:
 - Manipulation of the data matrix;
 - Filtering data matrix with respect to parameters, experimental data which to appear for the selected analogues;
 - Hide/Show the gap filling chart while the user is in the Gap filling module for better screening and analysis of data on data matrix;
 - Transfer of data to target chemical outside Data gap filling.

Outlook

- Background
- Objectives
- Specific aim
- **Manipulation of data matrix**
- Example

Data matrix Overview

- The data matrix window has three main parts:
 - Area with the Endpoint tree (1)
 - Area with the selected chemicals (2) and
 - Area with data (experimental, predicted) (3)

The screenshot displays the QSAR Toolbox 4.2 interface. The 'Data Matrix' window is active, showing a list of 12 chemicals (labeled 1-12) and their corresponding data for various endpoints. The endpoints are grouped into categories like Aquatic Toxicity, Environmental Fate and Transport, and Human Health Hazards. The data matrix table is as follows:

Endpoint	1	2	3	4	5	6	7	8	9	10	11	12
AW SW		M: 152 mg/L	M: 145 mg/L	M: 31.7 mg/L	M: 112 mg/L	M: 3.9 mg/L	M: 1.5 mg/L	M: 7.96 mg/L	M: 8.2 mg/L	M: 937 mg/L	M: 191 mg/L	M: 5.01 mg/L
Growth	105/123											
Growth Inhibition	12/24											
Immobilisation	15/15											
Intoxication	1/1											
Mortality	81/131	M: 17.8 mg/L	M: 10.5 mg/L	M: 6.62 mg/L	M: 7.77 mg/L							
Physiology	21/28					M: 1.66 mg/L	M: 14.6 mg/L	M: 5.64 mg/L				
Reproduction	10/24								M: 3.19 mg/L		M: 20 mg/L	M: 5.01 mg/L

Manipulation of data matrix Implementation in Toolbox

The functionality allows the user manually to manipulate the matrix via:

- Filtering the parameters and/or experimental data and/or profiling results which appear on the data matrix for the selected analogues;
- Reordering the columns with analogues in the category in order to more effectively analyze the data between the target and analogues
- Transferring of data (experimental/predicted) from analogues to the target chemical outside the gap filling module
- Hide/Show the data matrix once the user is in the stage of Data gap filling module

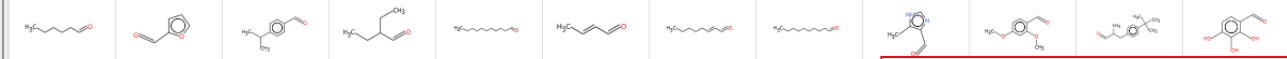
Illustration of the functionalities are shown on the next few slides.

Manipulation of data matrix Implementation in Toolbox

Filtering the data matrix

Filter endpoint tree... 1 [target] 2 3 4 5 6 7 8 9 10 11 12

Structure



Data matrix before filtering

Structure info	1 [target]	2	3	4	5	6	7	8	9	10	11	12
2D												
(Q) Acidic pKa (Chemaxon)	15.6	No value	No value	16.9	15.6	16.6	17.7	15.6	11.5	No value	15.8	7.22
(Q) Basic pKa (Chemaxon)	-6.94	-4.11	-7.1	-7.02	-6.94	-4.09	-4.2	-6.94	5.64	-4.56	-7.02	-6.57
BAF	0.81 log(L/kg bd...	0.05 log(L/kg bd...	2.09 log(L/kg bd...	0.76 log(L/kg bd...	2.6 log(L/kg bdwt)	0.1 log(L/kg bdwt)	1.52 log(L/kg bd...	2.36 log(L/kg bd...	0.03 log(L/kg bd...	0.91 log(L/kg bd...	2.93 log(L/kg bd...	0.16 log(L/kg bd...
BAF (lower trophic)	0.637 log(L/kg b...	0.034 log(L/kg b...	1.91 log(L/kg b...	0.593 log(L/kg b...	2.67 log(L/kg bd...	0.066 log(L/kg b...	1.33 log(L/kg bd...	2.34 log(L/kg bd...	0.022 log(L/kg b...	0.717 log(L/kg b...	2.95 log(L/kg bd...	0.2 log(L/kg bdwt)
BAF (mid trophic)	0.677 log(L/kg b...	0.037 log(L/kg b...	1.95 log(L/kg bd...	0.632 log(L/kg b...	2.65 log(L/kg bd...	0.073 log(L/kg b...	1.38 log(L/kg bd...	2.35 log(L/kg bd...	0.024 log(L/kg b...	0.761 log(L/kg b...	2.94 log(L/kg bd...	0.195 log(L/kg b...
BAF (upper trophic)	0.807 log(L/kg b...	0.047 log(L/kg b...	2.09 log(L/kg bd...	0.757 log(L/kg b...	2.6 log(L/kg bdwt)	0.099 log(L/kg b...	1.52 log(L/kg bd...	2.36 log(L/kg bd...	0.026 log(L/kg b...	0.905 log(L/kg b...	2.93 log(L/kg bd...	0.161 log(L/kg b...
BAF (upper trophic, biotransformation...)	0.873 log(L/kg b...	0.069 log(L/kg b...	2.27 log(L/kg bd...	0.825 log(L/kg b...	3.77 log(L/kg bd...	0.122 log(L/kg b...	1.63 log(L/kg bd...	3.03 log(L/kg bd...	0.05 log(L/kg bd...	0.956 log(L/kg b...	3.95 log(L/kg bd...	0.612 log(L/kg b...
BCF	0.84 log(L/kg bd...	0.5 log(L/kg bdwt)	1.76 log(L/kg bd...	0.81 log(L/kg bd...	1.1 log(L/kg bdwt)	0.5 log(L/kg bdwt)	1.36 log(L/kg bd...	2.15 log(L/kg bd...	0.5 log(L/kg bdwt)	0.9 log(L/kg bdwt)	2.54 log(L/kg bd...	0.64 log(L/kg bd...
BCF (lower trophic)	0.637 log(L/kg b...	0.034 log(L/kg b...	1.91 log(L/kg bd...	0.593 log(L/kg b...	2.65 log(L/kg bd...	0.066 log(L/kg b...	1.33 log(L/kg bd...	2.33 log(L/kg bd...	0.022 log(L/kg b...	0.717 log(L/kg b...	2.91 log(L/kg bd...	0.2 log(L/kg bdwt)
BCF (mid trophic)	0.677 log(L/kg b...	0.037 log(L/kg b...	1.95 log(L/kg bd...	0.632 log(L/kg b...	2.64 log(L/kg bd...	0.073 log(L/kg b...	1.38 log(L/kg bd...	2.35 log(L/kg bd...	0.024 log(L/kg b...	0.761 log(L/kg b...	2.92 log(L/kg bd...	0.195 log(L/kg b...

Physical Chemical Properties	1	2	3	4	5	6	7	8	9	10	11	12
Environmental Fate and Transport												
Ecotoxicological Information												
Aquatic Toxicity												
Growth	105/123	M: 152 mg/L	M: 145 mg/L	M: 31.7 mg/L	M: 112 mg/L	M: 3.9 mg/L	M: 1.5 mg/L	M: 7.96 mg/L	M: 8.2 mg/L	M: 937 mg/L	M: 191 mg/L	M: 77.2 mg/L
Growth Inhibition	12/24											
Immobilisation	15/15											
Intoxication	1/1											
Mortality	81/131	M: 17.8 mg/L	M: 10.5 mg/L	M: 6.62 mg/L	M: 7.77 mg/L				M: 3.19 mg/L		M: 20 mg/L	M: 5.01 mg/L
Physiology	21/28											
Reproduction	10/24					M: 1.66 mg/L	M: 14.6 mg/L	M: 5.64 mg/L				
Sediment toxicity												
Terrestrial Toxicity												
Human Health Hazards												
Profile												
Predefined												
US-EPA New Chemical Categories		Aldehydes (Acut...	Aldehydes (Acut...	Aldehydes (Acut...	Aldehydes (Acut...	Aldehydes (Acut...	Aldehydes (Acut...	Aldehydes (Acut...	Aldehydes (Acut...	Aldehydes (Acut...	Aldehydes (Acut...	Aldehydes (Acut...
Endpoint Specific												
Acute aquatic toxicity classification by...		Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...
Acute aquatic toxicity MOA by OASIS		Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes
Aquatic toxicity classification by ECOS...		Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)	Vinyl/Allyl Aldeh...	Vinyl/Allyl Aldeh...	Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)
Empiric												
Organic functional groups		Aldehyde	Aldehyde	Aldehyde	Aldehyde	Aldehyde	Alkene	Alkene	Aldehyde	Aldehyde	Aldehyde	Aldehyde
Organic functional groups (nested)		Aldehyde	Aldehyde	Aldehyde	Aldehyde	Aldehyde	Allyl	Allyl	Aldehyde	Aldehyde	Aldehyde	Aldehyde
Organic functional groups (US EPA)		Aldehyde, alipha...	Aldehyde, aroma...	Aldehyde, aroma...	Aldehyde, alipha...	Aldehyde, alipha...	Aliphatic Carbon...	Aliphatic Carbon...	Aldehyde, alipha...	Aldehyde, aroma...	Aldehyde, aroma...	Aldehyde, alipha...
Organic functional groups, Norbert Ha...		Aldehyde	Aldehyde	Aldehyde	Aldehyde	Aldehyde	No functional gr...	No functional gr...	Aldehyde	Aldehyde	Aldehyde	Aldehyde

Manipulation of data matrix Implementation in Toolbox

Filtering the data matrix

The screenshot displays the QSAR Toolbox interface. On the left, the 'Filter endpoint tree...' dialog is open, showing a tree view of endpoints. A red circle highlights the dropdown menu, which currently shows 'target'. A red arrow points from this dropdown to the 'Select' dialog box in the center. The 'Select' dialog box contains a tree view of endpoints with checkboxes for selection. The 'OK' button is highlighted with a red box. On the right, a red box contains the text 'Filter data matrix by select/unselect the corresponding checkboxes'. At the bottom right, another red box contains the text 'Click OK button to confirm the selection'.

Filter data matrix by
select/unselect the
corresponding checkboxes

Click OK button to
confirm the selection

Manipulation of data matrix Implementation in Toolbox

Filtering the data matrix

Filter endpoint tree... 1 [target] 2 3 4 5 6 7

Structure	1 [target]	2	3	4	5	6	7
Structure							
Structure info							
Parameters							
2D							
BCF	0.84 log(L/kg bd...	0.5 log(L/kg bdwt)	1.76 log(L/kg bd...	0.81 log(L/kg bd...	1.1 log(L/kg bdwt)	0.5 log(L/kg bdwt)	1.36 log(L/kg
Ecotoxicological Information							
Aquatic Toxicity	AW SW						
Growth							
IGC50							
48 h							
Protozoa							
Ciliophora							
Ciliatea							
Tetrahymena pyriformis...	97/97	M: 152 mg/L	M: 145 mg/L	M: 31.7 mg/L	M: 112 mg/L	M: 3.9 mg/L	M: 14 mg/L
Profile							
Endpoint Specific							
Acute aquatic toxicity classification by Verha...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unsp...
Acute aquatic toxicity MOA by OASIS	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Reactive unspeci...	Aldehydes
Aquatic toxicity classification by ECOSAR	Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)	Vinyl/Allyl Aldehy...	Vinyl/Allyl Alc...

Select

Select All Unselect All

- Structure info
- Parameters
 - 2D
 - 3D
- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
 - Aquatic Toxicity
 - Growth
 - EC50
 - IGC50
 - 48 h
 - Protozoa
 - Ciliophora
 - Ciliatea
 - Tetrahymena pyriformis
 - LOEC
 - NOEC
 - Growth Inhibition
 - Immobilisation
 - Intoxication
 - Mortality
 - Physiology
 - Reproduction
 - Sediment toxicity
 - Terrestrial Toxicity
- Human Health Hazards
- Profile
 - Predefined
 - Endpoint Specific
 - Acute aquatic toxicity classification by Verhaar (Modified)
 - Acute aquatic toxicity MOA by OASIS
 - Aquatic toxicity classification by ECOSAR
 - Empiric

OK Cancel

Data matrix after filtering includes the selected items (data/parameters/profilers) only

Manipulation of data matrix Implementation in Toolbox

Reordering of the analogues in the category

QSAR Toolbox 4.2 [Document 1]

QSAR TOOLBOX

Input Profiling Data Category definition Data Gap Filling Report










Profiling Custom profile

Apply View New Delete

Documents

document 1
CAS: 66251
Aldehydes (Acute toxicity) (US-EPA New Chemi

Filter endpoint tree... 1 [target] 5 3 4 6 7 8 9 10

Structure	1 [target]	5	3	4	6	7	8	9	10	
Structure										
Structure info										
Parameters										
Ecotoxicological Information	97/97	M: 152 mg/L	M: 14 mg/L	M: 31.7 mg/L	M: 112 mg/L	M: 3.9 mg/L	M: 7.96 mg/L	M: 8.2 mg/L	M: 937 mg/L	M: 191 mg/L
Profile										
Endpoint Specific										
Acute aquatic toxicity classification by...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	Class 3 (unspecif...	
Acute aquatic toxicity MOA by OASIS	Aldehydes	Reactive unspeci...	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	
Aquatic toxicity classification by ECOS...	Aldehydes (Mono)	Vinyl/Allyl Aldehy...	Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)	Vinyl/Allyl Aldehy...	Aldehydes (Mono)	Aldehydes (Mono)	Aldehydes (Mono)	

Profiling methods

Options

Select All Unselect All Invert

- Skin irritation/corrosion Exclusion rules
- Skin irritation/corrosion Inclusion rules
- Empiric
 - Chemical elements
 - Groups of elements
 - Lipinski Rule Oasis
 - Organic functional groups
 - Organic functional groups (nested)
 - Organic functional groups (US EPA)
 - Organic functional groups, Norbert Ha
 - Structure similarity
 - Tautomers unstable


Metabolism/Transformations

Options

Select All Unselect All Invert

- Documented
 - Observed Mammalian metabolism
 - Observed Microbial metabolism
 - Observed Rat In vivo metabolism
 - Observed rat liver metabolism with qu
 - Observed Rat Liver S9 metabolism











Left-click on the chemical, hold it and drag it next to the target



Manipulation of data matrix Implementation in Toolbox

Transferring of data (exp./predicted) from analogues to the target

Filter endpoint tree...

	Parent chemical...	metabolite #1	metabolite #2	metabolite #3	metabolite #4	metabolite #5	metabolite #6	metabolite #7	metabolite #8	metabolite #9
Structure										
Structure info	Methamphetamine	Generated observed rat in vivo metabolites								
Parameters										
Physical Chemical Properties										
Environmental Fate and Transport										
Ecotoxicological Information										
Human Health Hazards										
Acute Toxicity										
Bioaccumulation										
Carcinogenicity	2/6							M: Negative	M: Negative	
Developmental Toxicity / Teratogenicity										
Genetic Toxicity										
in Vitro	3/49	M: Negative						M: Equivocal	M: Negative	
in Vivo		R: Positive						M: Positive M: Positive M: Positive R: Positive		
Micronucleus Assay	2/5									
Immunotoxicity										
Irritation / Corrosion										
Neurotoxicity										
Photoinduced toxicity										
Repeated Dose Toxicity										
Sensitisation	AW SW AOP									
ToxCast										
Toxicity to Reproduction										
Toxicokinetics, Metabolism and Distributi...										

Manipulation of data matrix Implementation in Toolbox

Transferring of data (exp./predicted) from analogues to the target

Filter endpoint tree... Parent chemical... metabolite #1 metabolite #2 metabolite #3 metabolite #4 metabolite #5 metabolite #6 metabolite #7 metabolite #8 metabolite #9

Structure

Methamphetamine

Structure info
Parameters
Physical Chemical Properties
Environmental Fate and Transport
Ecotoxicological Information
Human Health Hazards
Acute Toxicity
Bioaccumulation
Carcinogenicity 2/6
Developmental Toxicity / Teratogenicity
Genetic Toxicity
in Vitro 3/49 M: Negative
in Vivo 1/4 R: Positive
Micronucleus Assay
Immunotoxicity
Irritation / Corrosion
Neurotoxicity
Photoinduced toxicity
Repeated Dose Toxicity
Sensitisation
ToxCast
Toxicity to Reproduction
Toxicokinetics, Metabolism and Distributi...

Select data points

- R: Positive
- M: Positive
- M: Positive
- M: Positive

1 M: Negative M: Negative
2 M: Equivocal M: Negative
3 M: Positive
4 M: Positive
5 M: Positive
6 R: Positive

Explain
Delete prediction
Explain prediction
Transfer to target
Set AOP target
Use for AOP
Copy

Selection of measured or prediction data to transfer to the target

Manipulation of data matrix Implementation in Toolbox

Hide/Show the chart while in data gap filling module

Filter endpoint tree...

- Structure
- 2D
 - Boiling point
 - log Kow
 - Molecular Weight
 - Vapor Pressure (Antoine method)
 - Water Solubility
- Physical Chemical Properties
- Human Health Hazards
 - Carcinogenicity
 - Genetic Toxicity
 - in Vitro
 - Bacterial Reverse Mutation Assay (e.g....
 - Gene mutation
 - Salmonella typhimurium
 - No S9 Info
 - With S9
 - Without S9

1 [target]	2	4	11	12	15	16
363 °C	368 °C	379 °C	352 °C	352 °C	383 °C	387 °C
4.65	5.06	5.55	4.15	4.15	2.69	3.15
223 Da	237 Da	251 Da	209 Da	209 Da	210 Da	220 Da
3.1E-06 mm Hg	2.22E-06 mm Hg	7.94E-07 mm Hg	8.41E-06 mm Hg	8.41E-06 mm Hg	4.89E-07 mm Hg	3.59E-07 mm Hg
3.56 mg/L	1.32 mg/L	0.42 mg/L	11.1 mg/L	11.1 mg/L	195 mg/L	37.7 mg/L
Data matrix						
3/3				M: Positive	M: Positive	M: Positive
7/29 M: Positive	M: Positive	M: Positive	M: Positive	M: Positive	M: Negative	M: Positive
7/20 M: Negative	M: Negative	M: Negative	M: Negative	M: Negative	M: Negative	M: Negative

**Click the minimize button to
hide the gap filling chart
(see next slide)**

Descriptors

Prediction

Read-across prediction for Gene mutation, based on 5 values
Observed: Positive (x4); Predicted: Positive

Active descriptor X log Kow

Select / filter data

Subcategorize

Mark chemicals by WS

Mark chemicals by descriptor value

Filter points by test conditions

Mark focused chemical

Mark focused points

Remove marked data

Clear existing marks

✓ Accept prediction

Manipulation of data matrix Implementation in Toolbox

Hide/Show the chart while in data gap filling module

The screenshot displays the QSAR Toolbox interface. On the left is a 'Filter endpoint tree...' sidebar with categories like Structure, Physical Chemical Properties, Human Health Hazards, and Profile. The main area shows a data matrix with columns numbered 1 to 16 and rows for various properties. A red box highlights the text 'Data matrix' in the center of the table. A callout box on the right points to the minimize button in the window's title bar, with the text 'Click the minimize button to restore the gap filling chart'.

Endpoint	1 [target]	2	4	11	12	15	16
Structure							
Boiling point	363 °C	368 °C	379 °C	352 °C	352 °C	383 °C	387 °C
log Kow	4.65	5.06	5.55	4.15	4.15	2.69	3.15
Molecular Weight	223 Da	237 Da	251 Da	209 Da	209 Da	210 Da	220 Da
Vapor Pressure (Antoine method)	3.1E-06 mm Hg	2.22E-06 mm Hg	7.94E-07 mm Hg	8.41E-06 mm Hg	8.41E-06 mm Hg	4.89E-07 mm Hg	3.59E-07 mm Hg
Water Solubility	3.56 mg/L	1.32 mg/L	0.42 mg/L	11.1 mg/L	11.1 mg/L	195 mg/L	37.7 mg/L
Data matrix							
in Vitro							
Bacterial Reverse Mutation Assay (e.g...)							
Gene mutation							
Salmonella typhimurium							
No S9 Info	3/3				M: Positive	M: Positive	M: Positive
With S9	7/29	M: Positive	M: Positive	M: Positive	M: Positive	M: Negative	M: Positive
Without S9	7/29	M: Negative	M: Negative	M: Negative	M: Negative	M: Negative	M: Negative
Undefined Test organisms...	1/2					M: Positive	
Mammalian Cell Gene Mutation A...	1/4					M: Inadequate	
in Vivo	1/2					M: Positive	
Profile							
Predefined							
Substance type	Discrete chemical	Discrete chemical	Discrete chemical	Discrete chemical	Discrete chemical	Discrete chemical	Discrete chemical
General Mechanistic							
Endpoint Specific							
Empiric							
Metabolism/Transformations							
Rat liver S9 metabolism simulator	12 metabolite(s)	8 metabolite(s)	16 metabolite(s)	9 metabolite(s)	9 metabolite(s)	5 metabolite(s)	6 metabolite(s)
Predefined							
General Mechanistic							
Endpoint Specific							
DNA alerts for AMES by OASIS	1 x Radical >> R... 1 x SN1 >> Nucl...	1 x AN2 1 x AN2 >> Carb...	1 x Radical >> R... 1 x SN1 >> Nucl...	1 x AN2 1 x AN2 >> Carb...	1 x AN2 1 x AN2 >> Carb...	1 x AN2 1 x AN2 >> Carb...	1 x AN2 1 x AN2 >> Carb...

Outlook

- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- **Example**

Example

- In this example we will predict *AMES mutagenicity* of Benzamide, N-((ethylnitrosoamino)methyl) [CAS# 59665-03-1], which will be the “target” chemical
- Collect data and profiling results for the target according to the suitable profilers
- Generate *in vitro* rat liver metabolites of the target chemical
- Make read-across prediction for the target by transferring observed data of the preliminary generated metabolites outside gap filling module

Workflow

- **The Toolbox workflow include six modules :**
 - Input
 - Profiling
 - Data
 - Category Definition
 - Data Gap Filling
 - Report
- **In this example we will use only the first three modules in order to fulfil the aims of the example.**

Outlook

- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- Example
 - **Input**

Input Overview

- This module provides the user with several means of entering the chemical of interest or the target chemical.
- Since all subsequent functions are based on chemical structure, the goal here is to make sure the molecular structure assigned to the target chemical is the correct one.

Input

Entering a target chemical by CAS#

The screenshot shows the QSAR Toolbox 4.2 interface. The 'Input' button in the top toolbar is highlighted with a red box and labeled '1'. The 'Search by CAS #' dialog box is open, showing the CAS# '59665031' entered in the search field and the 'OK' button highlighted with a red box and labeled '3'. The results table below the dialog box shows the following information:

Search by CAS #		
59665031 Search		
Select All Unselect All Invert Selection Selected 1 of 1		
1	CAS	59665-03-1
	SMILES	CCN(CNC(=O)c1ccccc1)N=O
	CS Relation	Low
<input checked="" type="checkbox"/>	Substance	Mono constituent
	Composition	
	Name	Benzamide, N-((ethylnitroso... N-Ethyl-N-(benzoylaminomet...

Chemical structure image of Benzamide, N-((ethylnitroso... N-Ethyl-N-(benzoylaminomet... is shown next to the search results.

1. Click **CAS#**
2. Enter **CAS# 59665-03-1**
3. Click **OK**

Outlook

- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- Example
 - Input
 - **Define target endpoint**

Define target endpoint

1. Once you are in the **Input** module click **Define** button

2. Select **Genetic Toxicity** node part of **Human Health Hazard**

3. Select the specific data from the pop-up menus as shown

4. Click **Finish**

5. The row related to the defined target endpoint is getting yellow highlighted

Outlook

- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- The exercise
 - Input
 - Define target endpoint
 - **Profiling**

Profiling Overview

- “Profiling” refers to the electronic process of retrieving relevant information on the target compound, other than environmental fate, ecotoxicity and toxicity data, which are stored in the Toolbox database.
- Available information includes likely mechanism(s) of action, as well as observed or simulated metabolites.
- Based on the “profilers’ relevancy” the most suitable ones are getting colour highlighted (see next slide)*
- For the purpose of our example only suitable profilers in combination with metabolism simulator are used

*For more details regarding relevancy of profilers see ppt: *Example for predicting skin sensitization taking into account alert performance*

Profiling

Profiling the target chemical

The screenshot displays the QSAR Toolbox 4.2 interface. The top toolbar features several icons, with the 'Profiling' icon (1) highlighted in a red box. Below the toolbar, the left sidebar contains two sections: 'Profiling methods' (2) and 'Metabolism/Transformation' (2). Both sections have a 'Suitable' checkbox checked, and several specific methods and simulators are also checked. The central panel, 'Filter endpoint tree...', shows a tree view of endpoints, with 'Non-covalent interaction' (4) and 'Endpoint Specific' (5) highlighted. The right panel displays the results for the selected endpoint, showing 'No alert found' (5) and '8 metabolite(s)' (6) produced by the 'Rat liver S9 metabolism simulator'. The results for the metabolites are listed below.

1. Go to **Profiling**; 2. Check the suitable profiles and simulators (rat liver in this case); 3. Click **Apply**; 4. *No alert* is found in the target structure based on endpoint-specific *DNA alerts for AMES by OASIS* profiler; 5. Eight metabolites are produced for the target after applying the *in vitro* rat liver metabolism simulator; 6. Structural alerts for interaction with DNA are found in 3 of the metabolites.

Profiling

Explain of profiling results

The screenshot displays the 'Profiling results' window. On the left is a 'Filter endpoint tree...' with categories like Structure, Parameters, Physical Chemical Properties, Environmental Fate and Transport, Ecotoxicological Information, Human Health Hazards, and Profile. The main area shows a list of results for '1 [target]'. A context menu is open over the results, with 'Explain' highlighted (labeled '1'). Another context menu is open over the '3 x N-Nitroso Compounds' alert, with 'Display chemicals' highlighted (labeled '2'). A separate window titled 'File' shows three chemical structures, each labeled 'No CAS number'.

1. **Right-click** over the results and select **Explain** for more details. 2. Right-click on the identified alert and select **Display chemicals** to see generated metabolites where active alert has been found.

Outlook

- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- Example
 - Input
 - Define target endpoint
 - Profiling
 - **Data**

Data Overview

- “Data” refers to the electronic process of retrieving the environmental fate, ecotoxicity and toxicity data that are stored in the Toolbox.
- Data gathering can be executed in a global fashion (i.e., collecting all data for all endpoints) or on a more narrowly defined basis (e.g., collecting data for a single or limited number of endpoints).
- Once the endpoint is selected, the databases, which contain such type of data are highlighted in green (see next slide).

Data Collect data

1. Go to **Data** module

2. Select the green highlighted databases without ECHA CHEM*

3. Click **Gather**

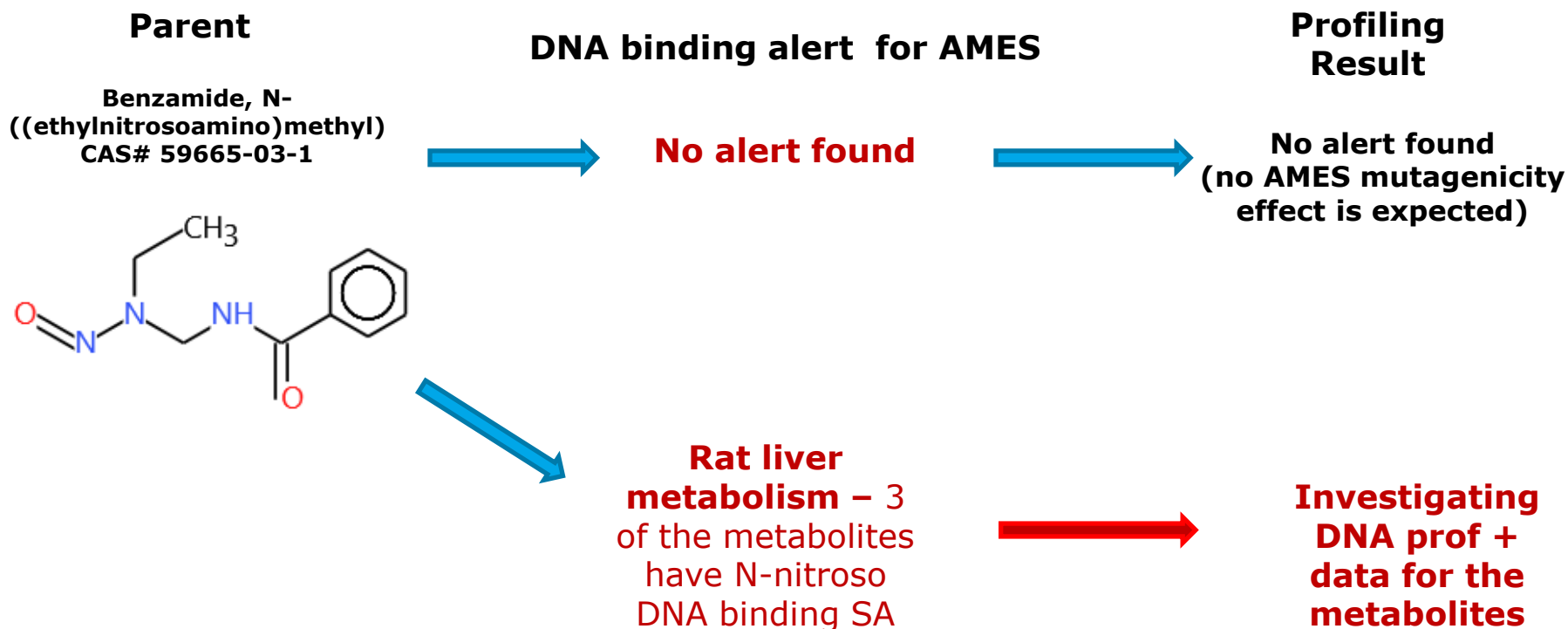
4. No data has been found for target chemical

*ECHA Chem is excluded from this selection due to absence of experimental data related to the target endpoint either for the target chemical or for its analogues as well (based on preliminary investigation).

Recap

- The first module, which introduces the target chemical, ensure correctness of the structure.
- The second module (Profiling) shows that there is no DNA binding alert for target chemical itself, but structural alerts responsible for DNA interaction have been found in the generated rat liver metabolites. The latter determines the forthcoming actions of the workflow.
- In the third module (Data), you have found that the target chemical has no data associated with the target endpoint
- The study continues with investigating profiling results and data of generated rat liver metabolites of target chemical (see next slides).

Recap



Outlook

- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- Example
 - Input
 - Define target endpoint
 - Profiling
 - Data
 - **Simulation of rat liver S9 metabolism**

Handling of rat liver S9 metabolism of target chemical

- Metabolizing the target chemical by Rat liver S9 metabolism simulator
- The simulation of metabolism of target chemical is accomplished in section **Input**
- The generated metabolites appear in tree like form (see next slide)

Handling of rat liver S9 metabolism of target chemical

Multiplication of target chemical

The screenshot shows the QSAR Toolbox 4.2 interface. The 'Input' module is selected in the menu bar. In the document tree, the target chemical 'Rat liver S9 m' is right-clicked, and the 'Multiplication' option is chosen. The 'Metabolism/Transformations' sub-menu is open, and 'Rat liver S9 metabolism simulator' is selected. The main workspace displays the parent chemical and eight generated metabolites (metabolite #1 to #8) in a grid-like structure.

1. Go to **Input** module
2. Click on the level with **CAS #** of the target chemical and right-click on it, then
2. Select **Multiplication-Metabolism/Transformations /Rat liver S9 metabolism simulator**
4. Generated metabolites appear in tree like form and also are aligned next to the target

Next actions are focused on investigating the profilers of the generated metabolites and collecting data for them

Outlook

- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- Example
 - Input
 - Define target endpoint
 - Profiling
 - Data
 - Generation of rat liver S9 metabolism
 - **Profiling and collecting data for metabolites**

Handling of rat liver S9 metabolism of target chemical

Profiling the package of metabolites

4

1

2

3

5

Profiling results related to the most suitable profilers of package parent and metabolites

Endpoint	Parent chemical...	metabolite #1	metabolite #2	metabolite #3	metabolite #4	metabolite #5	metabolite #6	metabolite #7	metabolite #8
General Mechanistic	DNA binding by OASIS	Non-covalent int., SN1	No alert found	SN1	Non-covalent int., SN2	No alert found	SN1	Non-covalent int., No alert found	No alert found
	DNA binding by OECD	Non-covalent int., SN1	No alert found	SN2	Schiff base form., SN1	No alert found	SN1	Non-covalent int., No alert found	No alert found
Endpoint Specific	DNA alerts for Ames by OASIS	No alert found	No alert found	SN1	SN1	No alert found	SN1	No alert found	No alert found
	in vitro mutagenicity (Ames test) alert...	Alkyl and aryl N-...	No alert found	Alkyl and aryl N-...	Alkyl and aryl N-...	Simple aldehyde	Alkyl and aryl N-...	No alert found	Aliphatic azo an..., Aliphatic azo an...

1. Go to **Profiling**; 2. Check suitable profilers related to the target endpoint; 3. Unselect Rat liver S9 metabolism simulator 4. Click **Apply**; 5. The profiling results appears on data matrix.

Handling of rat liver S9 metabolism of target chemical

Collect data for metabolites

1. Go to **Data** module;

2. The databases related to the defined target endpoint are already selected;

3. Click **Gather**;

4. The data for the parent and metabolites appears on data matrix.

Gathered data for parent and metabolites

Endpoint	Parent chemical...	metabolite #1	metabolite #2	metabolite #3	metabolite #4	metabolite #5	metabolite #6	metabolite #7	metabolite #8
Gene mutation 4/82		M: Negative	M: Negative			M: Positive	M: Negative		
in Vitro Mammalian C... 3/6		M: Negative	M: Negative			M: Negative			
Mammalian Cell Gene... 1/1						M: Positive			
in Vivo 1/2						M: Negative			

Outlook

- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- Example
 - Input
 - Define target endpoint
 - Profiling
 - Data
 - Generation of rat liver S9 metabolism
 - Profiling and collecting data for metabolites
 - **Transferring data to the target outside data gap filling module**

Transferring experimental data of metabolite to the target outside data gap filling

Hold (1) and drag the chemical next to the target. The profiling results (2) and the data (3) are available for the parent and metabolites (see next slide).

Filter endpoint tree...	Parent chemical...	metabolite #1	metabolite #2	metabolite #3	metabolite #4	metabolite #5	metabolite #6	metabolite #7	metabolite #8
Structure									
Carcinogenicity									
Developmental Toxicity / Teratogenicity									
Genetic Toxicity									
in Vitro									
Bacterial Reverse Mutation Assay (e.g....									
Salmonella typhimurium									
Gene mutation	4/82	M: Negative			M: Negative	M: Positive	M: Negative		
in Vitro Mammalian Chromosome...	3/6	M: Negative	M: Negative		M: Negative				
Mammalian Cell Gene Mutation A...	1/1				M: Positive				
in Vivo	1/2				M: Negative				
Immunotoxicity									
Irritation / Corrosion									
Neurotoxicity									
Photoinduced toxicity									
Repeated Dose Toxicity									
Sensitisation	AW SWAOP								
ToxCast									
Toxicity to Reproduction									
Toxicokinetics, Metabolism and Distributi...									
Profile									
General Mechanistic									
Endpoint Specific									
DNA alerts for AMES by OASIS	No alert found	No alert found	SN1 >> Nucleop... SN1 >> Nucleop... SN1 >> Nucleop... SN1 >> Nucleop...	SN1 >> Nucleop...	No alert found	SN1 >> Nucleop... SN1 >> Nucleop... SN1 >> Nucleop...	No alert found	No alert found	No alert found
Metabolism/Transformations									
Rat liver S9 metabolism simulator	8 metabolite(s)								
General Mechanistic									
Endpoint Specific									

2 Positive data (-S9)

The structural alert for interaction with DNA for metabolite #5 (2) coincide with the identified positive experimental data of the metabolite as parent (3)

Structural alerts for interaction with DNA identified in the generated metabolites

Transferring experimental data of metabolite to the target outside data gap filling

1. Right-click the cell with observed data of the metabolite #5

2. Select **Transfer to target**

3. Select which data points to be transferred to the target

4. Click **OK** button

The positive data points of the metabolite #5 will be transferred to the target chemical. In case of conflicting data points (positive/negative) then a read-across should be obtained based on worst case-data points

Transferring experimental data of metabolite to the target outside data gap filling

QSAR Toolbox 4.2 [Document 3]

QSAR TOOLBOX

Input Profiling Data Category definition Data Gap Filling Report

Profiling Custom profile

Apply View New Delete

Documents

- Document 1
- Document 2
- Document 3
 - # CAS: 59665031
 - Rat liver S9 metabolism simulator
 - metabolite #1
 - metabolite #2
 - metabolite #3
 - metabolite #4
 - metabolite #5
 - metabolite #6
 - metabolite #7
 - metabolite #8

Profiling methods

Options

Select All Unselect All Invert

- Suitable
 - DNA alerts for AMES by OASIS
 - DNA binding by OASIS
 - DNA binding by OECD
 - in vitro mutagenicity (Ames test) alert
- Plausible
 - Aquatic toxicity classification by ECOS
 - Chemical elements
 - DNA alerts for CA and MNT by OASIS
 - Groups of elements
 - in vivo mutagenicity (Micronucleus) alert
 - Lindecker Rule based

Metabolism/Transformations

Options

Select All Unselect All Invert

- Suitable
 - Rat liver S9 metabolism simulator
- Plausible
 - Dissociation simulator
 - Hydrolysis simulator (neutral)
 - in vivo Rat metabolism simulator
- Unclassified

Filter endpoint tree...

- Structure
- Structure info
- Parameters
- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
 - Acute Toxicity
 - Bioaccumulation
 - Carcinogenicity
 - Developmental Toxicity / Teratogenicity
 - Genetic Toxicity
 - in Vitro
 - Bacterial Reverse Mutation Assay (e.g. ...)
 - Salmonella typhimurium
 - Gene mutation 5/83
 - M: Positive
 - M: Positive
 - in Vitro Mammalian Chromosome...
 - Mammalian Cell Gene Mutation A... 1/1
 - in Vivo 1/2
 - Immunotoxicity
 - Irritation / Corrosion
 - Neurotoxicity
 - Photoinduced toxicity
 - Repeated Dose Toxicity
 - Sensitisation AW SW AOP
 - ToxCast
 - Toxicity to Reproduction
 - Toxicokinetics, Metabolism and Distributi...
- Profile

Parent chemical...	metabolite #5	metabolite #1	metabolite #2	metabolite #3	metabolite #4	metabolite #6	metabolite #7	metabolite #8
		R: Positive						
		M: Positive	M: Negative			M: Negative	M: Negative	
		M: Positive	M: Negative			M: Negative	M: Negative	
			M: Negative			M: Negative	M: Negative	
			M: Negative			M: Negative		
						M: Positive		
						M: Negative		

Read-across prediction based on positive observed data of the metabolite appeared for the target chemical

Congratulation

- You have now been introduced to the Data matrix manipulation options;
- You have now been introduced to the transfer of read-across prediction to the target chemical outside gap filling module.
- Note proficiency comes with practice.