QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

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

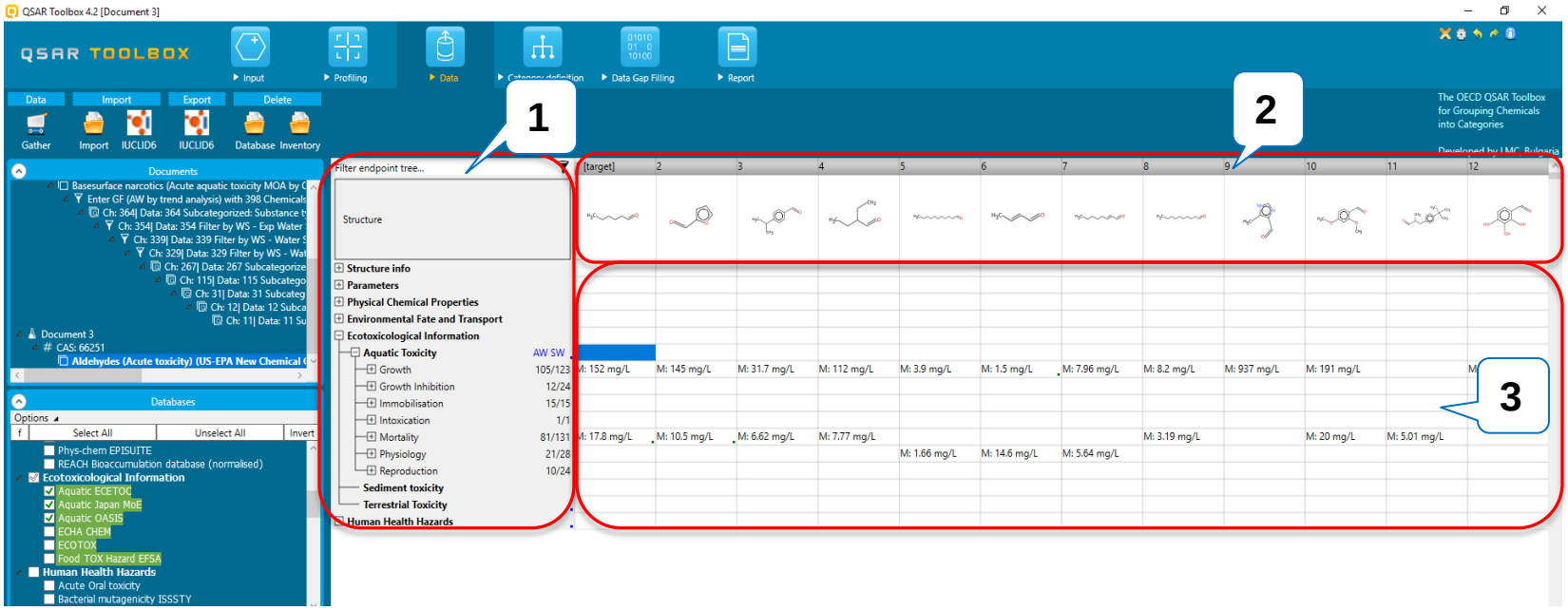
- To introduce and make the user familiar with:
  - Manipulation of the data matrix;
  - Filtering data matrix with respect to parameters, experimental data which to appear for the selected analogues;
  - Hide/Show the gap filling chart while the user is in the Gap filling module for better screening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



Manipulation of data matrix
Implementation in Toolbox

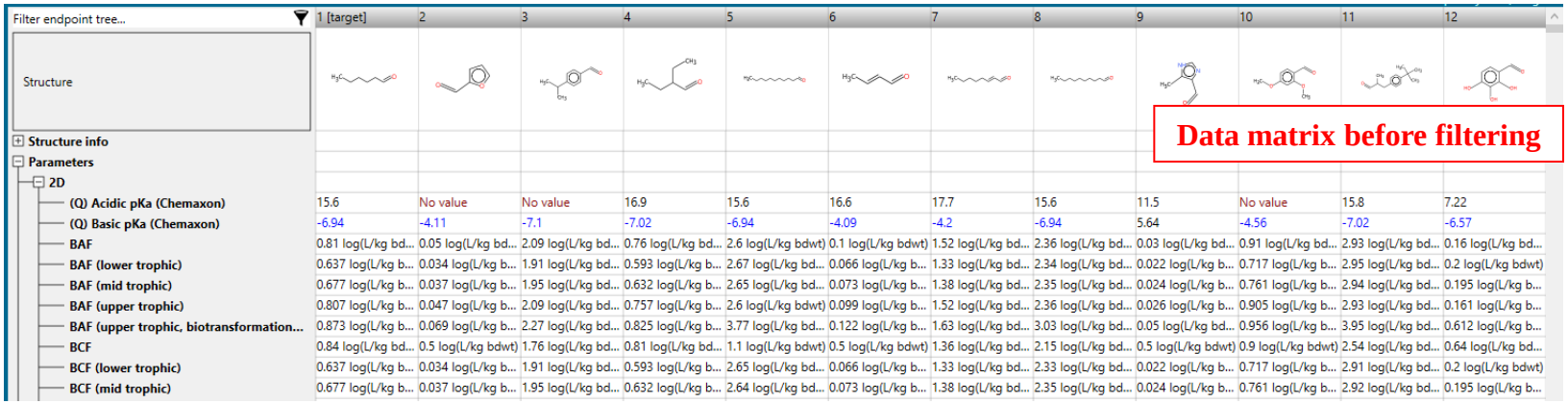
The functionality allows the user manually to manipulate the matrix via:

- 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



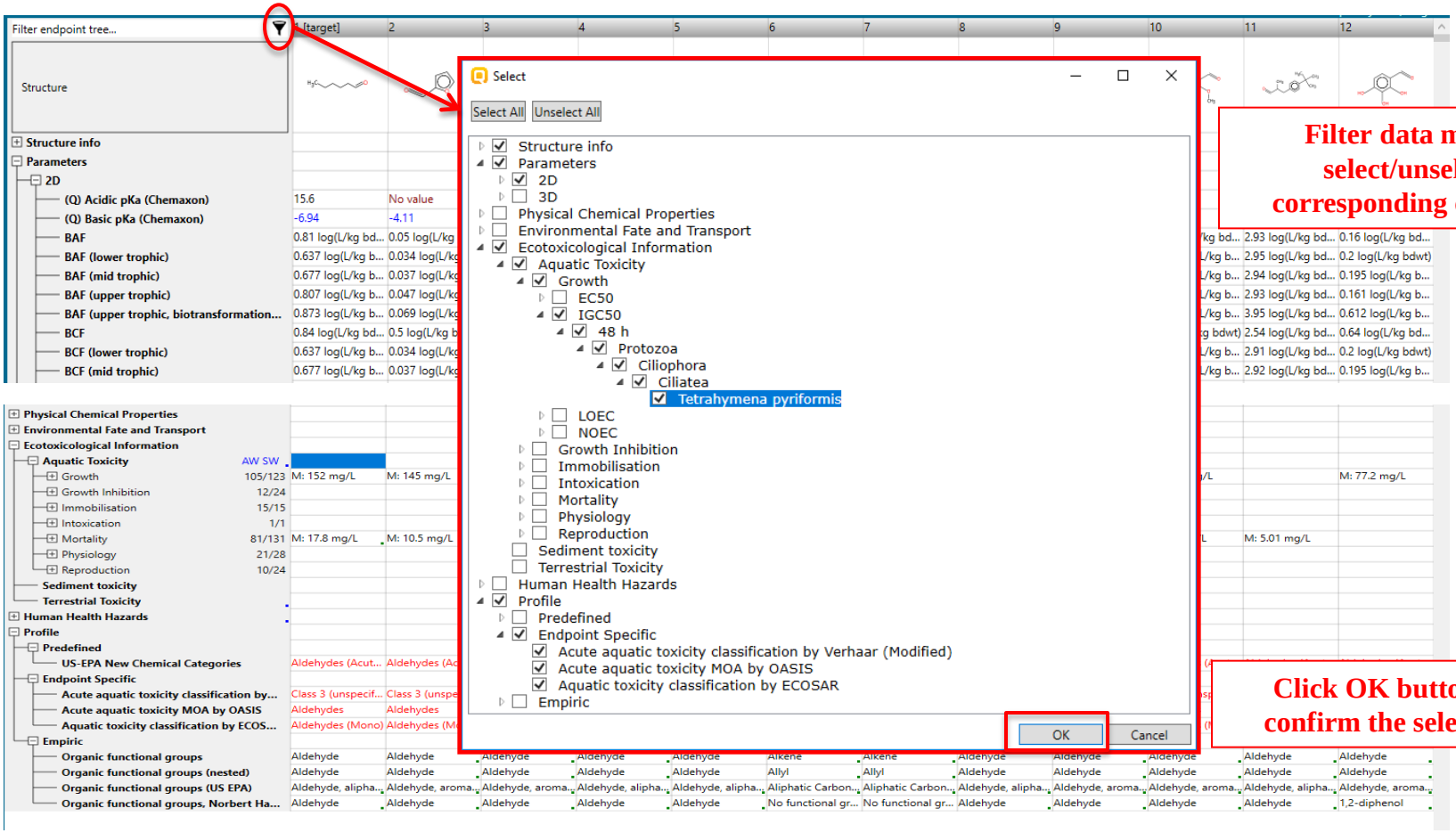
Filter endpoint tree...	1 [target]	2	3	4	5	6	7	8	9	10	11	12
Structure												
Structure info												
Parameters												
2D												
(Q) Acidic pKa (Chemaxon)	15.6	No value	No value	16.9	15.6	16.6	17.7	15.6	11.5	No value	15.8	7.22
(Q) Basic pKa (Chemaxon)	-6.94	-4.11	-7.1	-7.02	-6.94	-4.09	-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 bd...	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...

Data matrix before filtering

Physical Chemical Properties													
Environmental Fate and Transport													
Ecotoxicological Information													
Aquatic Toxicity	AW SW												
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					M: 1.66 mg/L	M: 14.6 mg/L	M: 5.64 mg/L					
Reproduction	10/24												
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...	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...	Class 3 (unspecif...
Acute aquatic toxicity MOA by OASIS		Aldehydes	Aldehydes	Aldehydes	Aldehydes	Aldehydes	Reactive unspeci...	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)	Aldehydes (Mono)
Empiric													
Organic functional groups		Aldehyde	Aldehyde	Aldehyde	Aldehyde	Aldehyde	Alkene	Alkene	Aldehyde	Aldehyde	Aldehyde	Aldehyde	Aldehyde
Organic functional groups (nested)		Aldehyde	Aldehyde	Aldehyde	Aldehyde	Aldehyde	Allyl	Allyl	Aldehyde	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...	Aldehyde, aroma...
Organic functional groups, Norbert Ha...		Aldehyde	Aldehyde	Aldehyde	Aldehyde	Aldehyde	No functional gr...	No functional gr...	Aldehyde	Aldehyde	Aldehyde	Aldehyde	1,2-diphenol

Manipulation of data matrix
Implementation in Toolbox

Filtering the data matrix

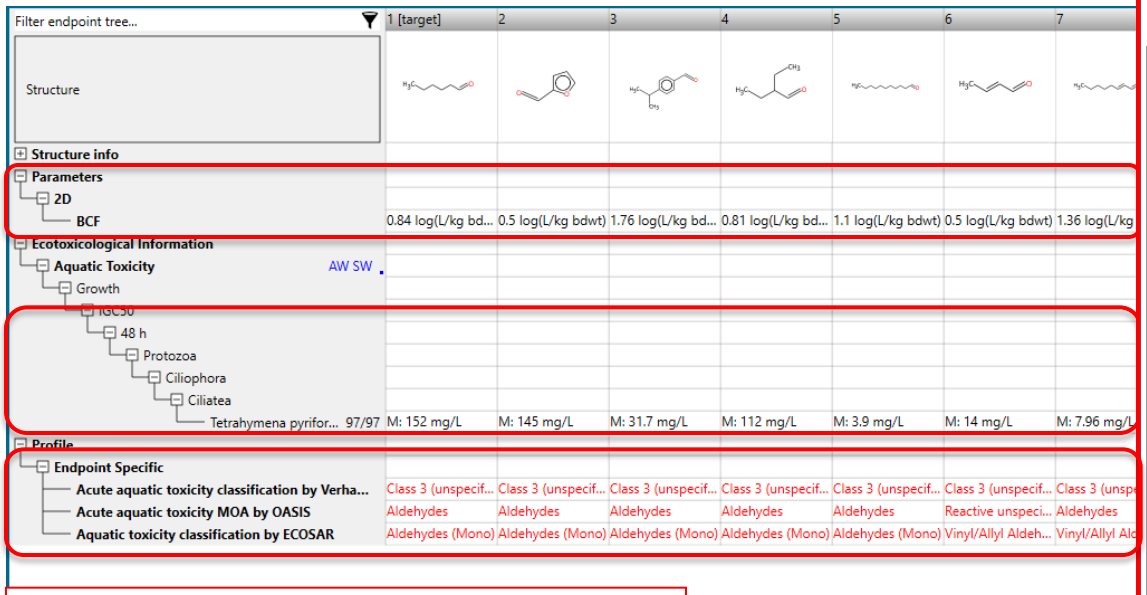


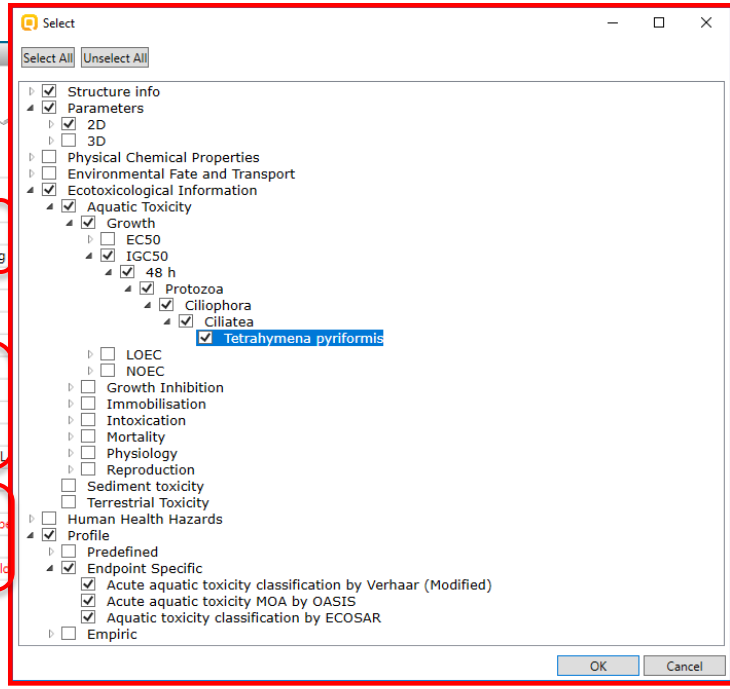
Filter data matrix by select/unselect the corresponding checkboxes

Click OK button to confirm the selection

Manipulation of data matrix
Implementation in Toolbox

Filtering the data matrix

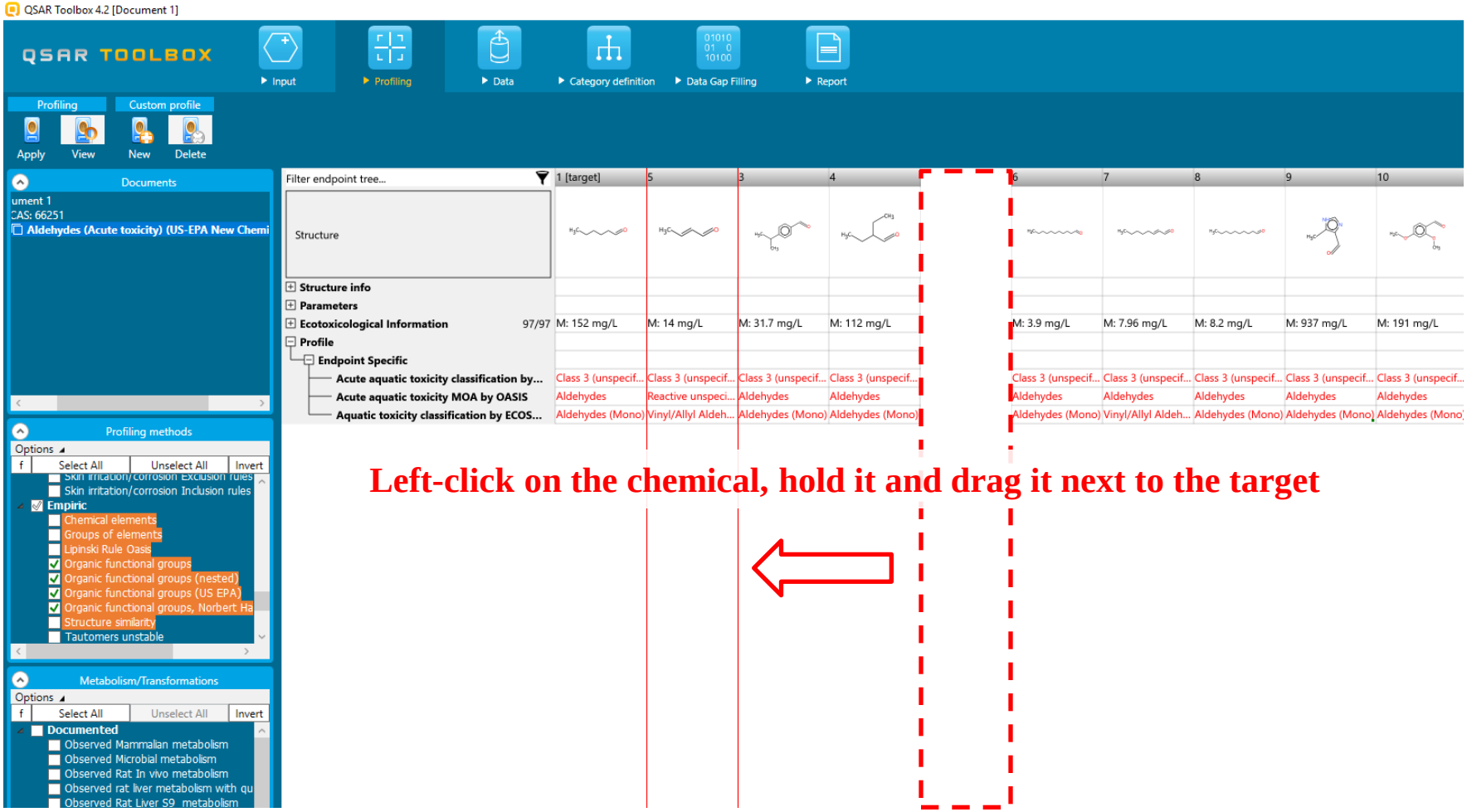




Data matrix after filtering includes the selected items (data/parameters/profilers) only

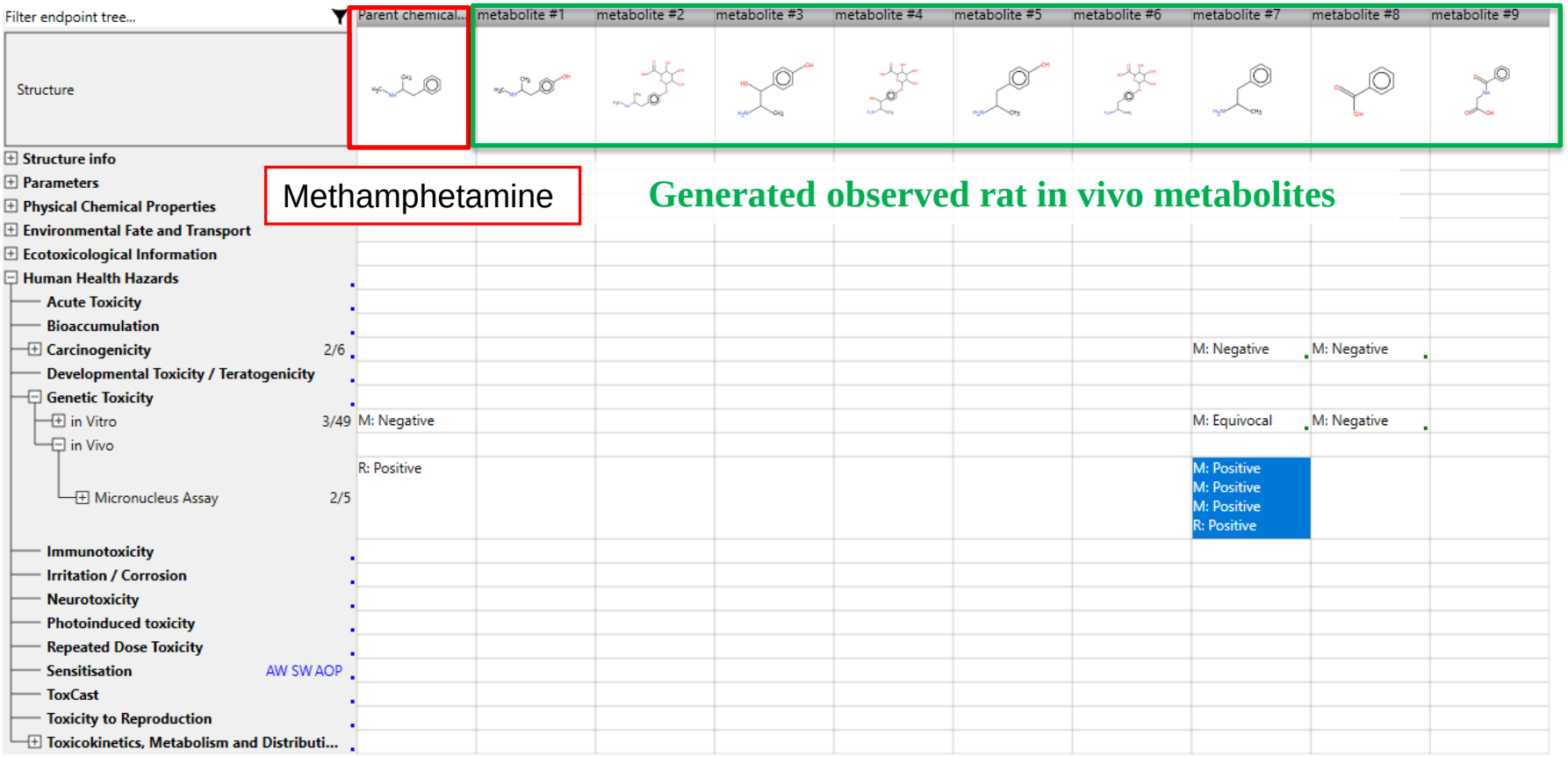
Manipulation of data matrix
Implementation in Toolbox

Reordering of the analogues in the category



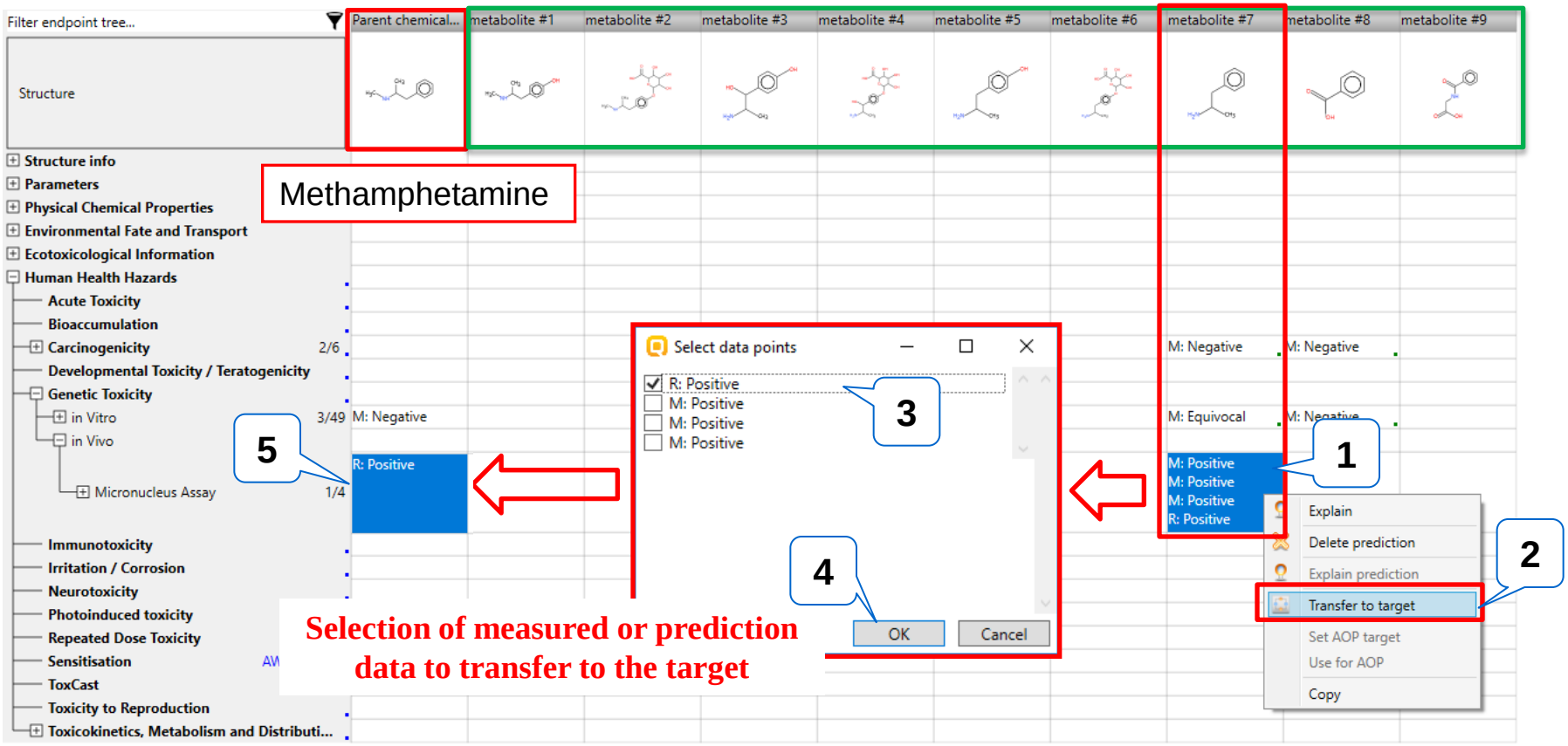
Manipulation of data matrix
Implementation in Toolbox

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



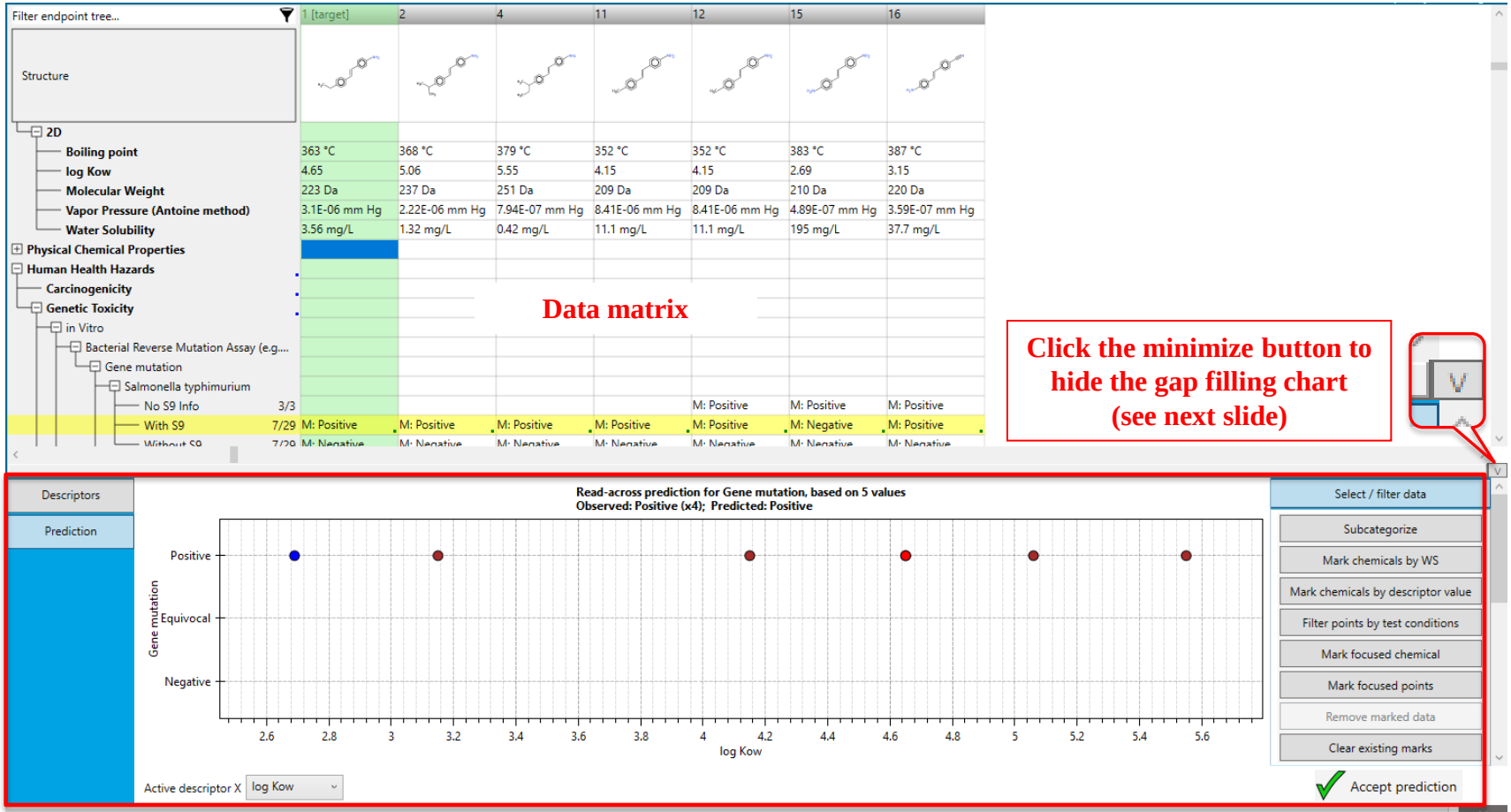
Manipulation of data matrix
Implementation in Toolbox

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



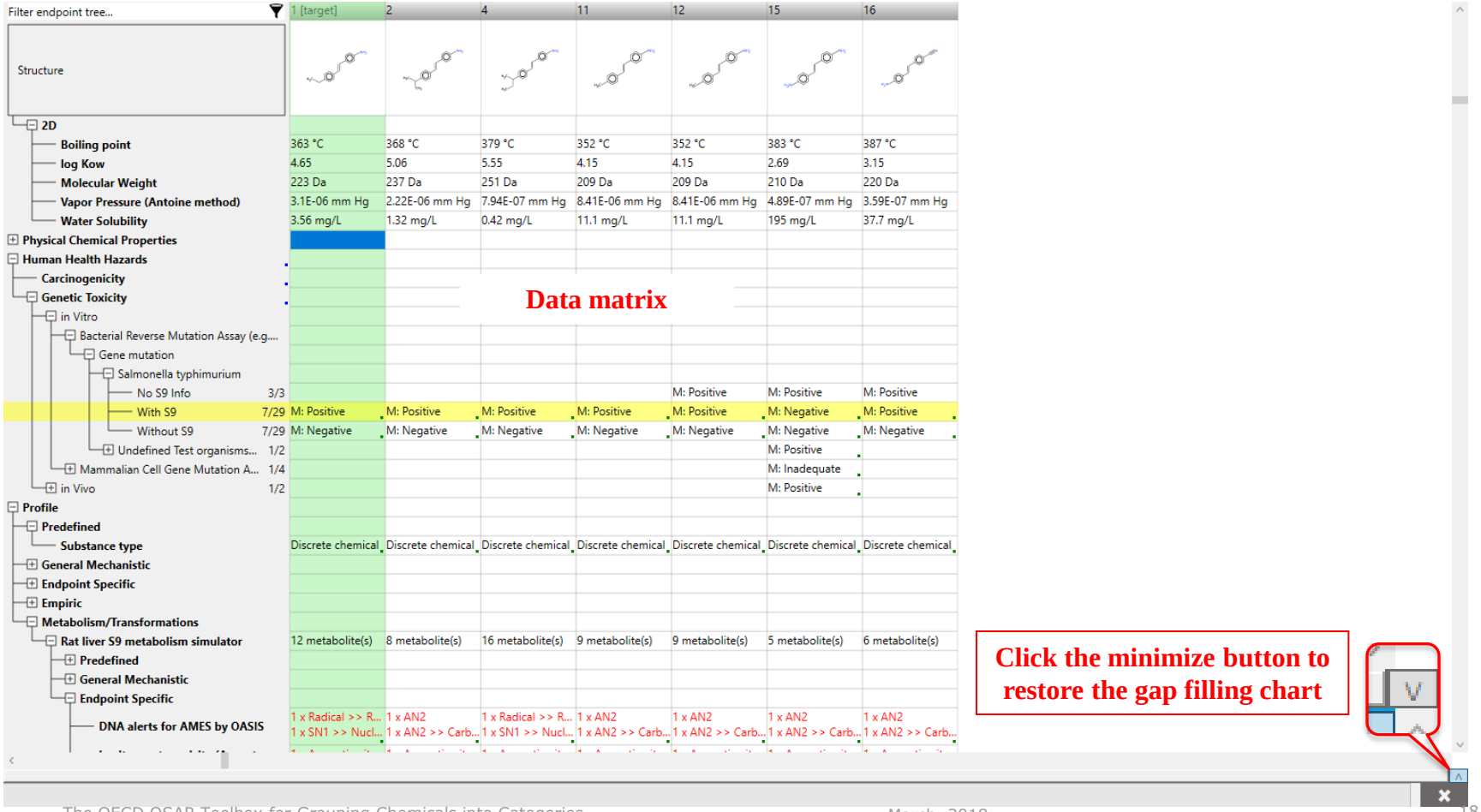
Manipulation of data matrix
Implementation in Toolbox

Hide/Show the chart while in data gap filling module



Manipulation of data matrix
Implementation in Toolbox

Hide/Show the chart while in data gap filling module



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

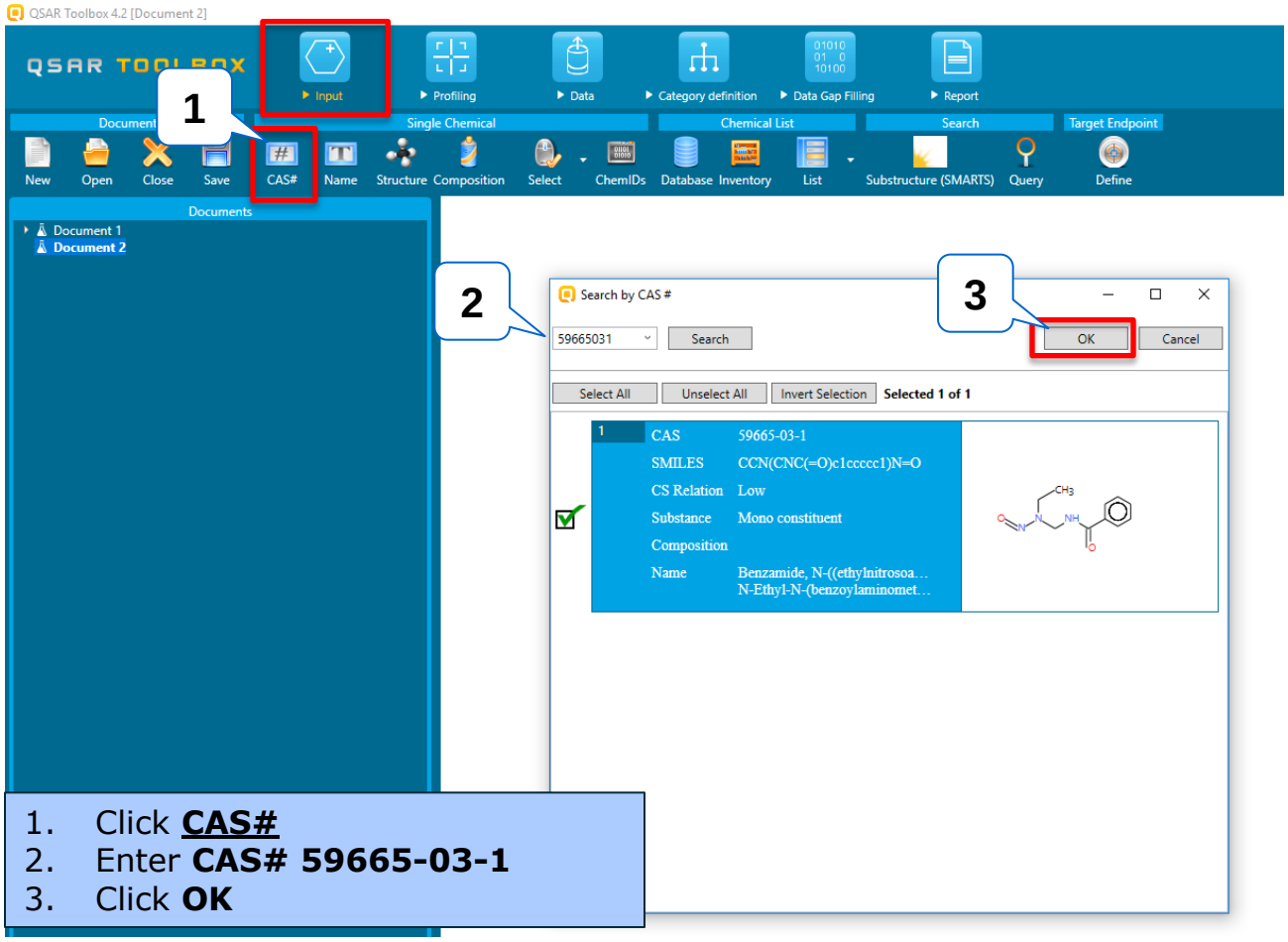
- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- Example
  - **Input**

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

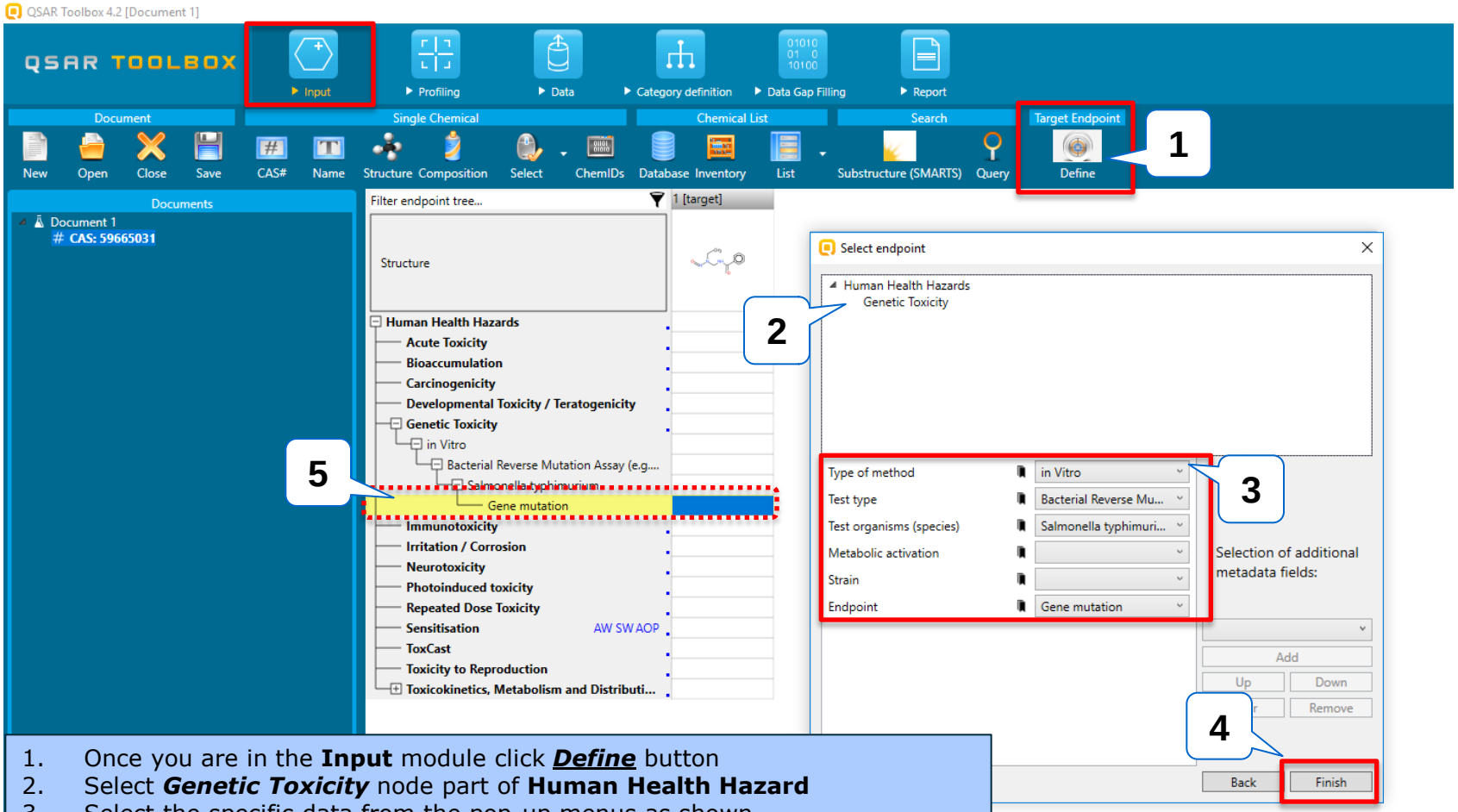
Entering a target chemical by CAS#



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**

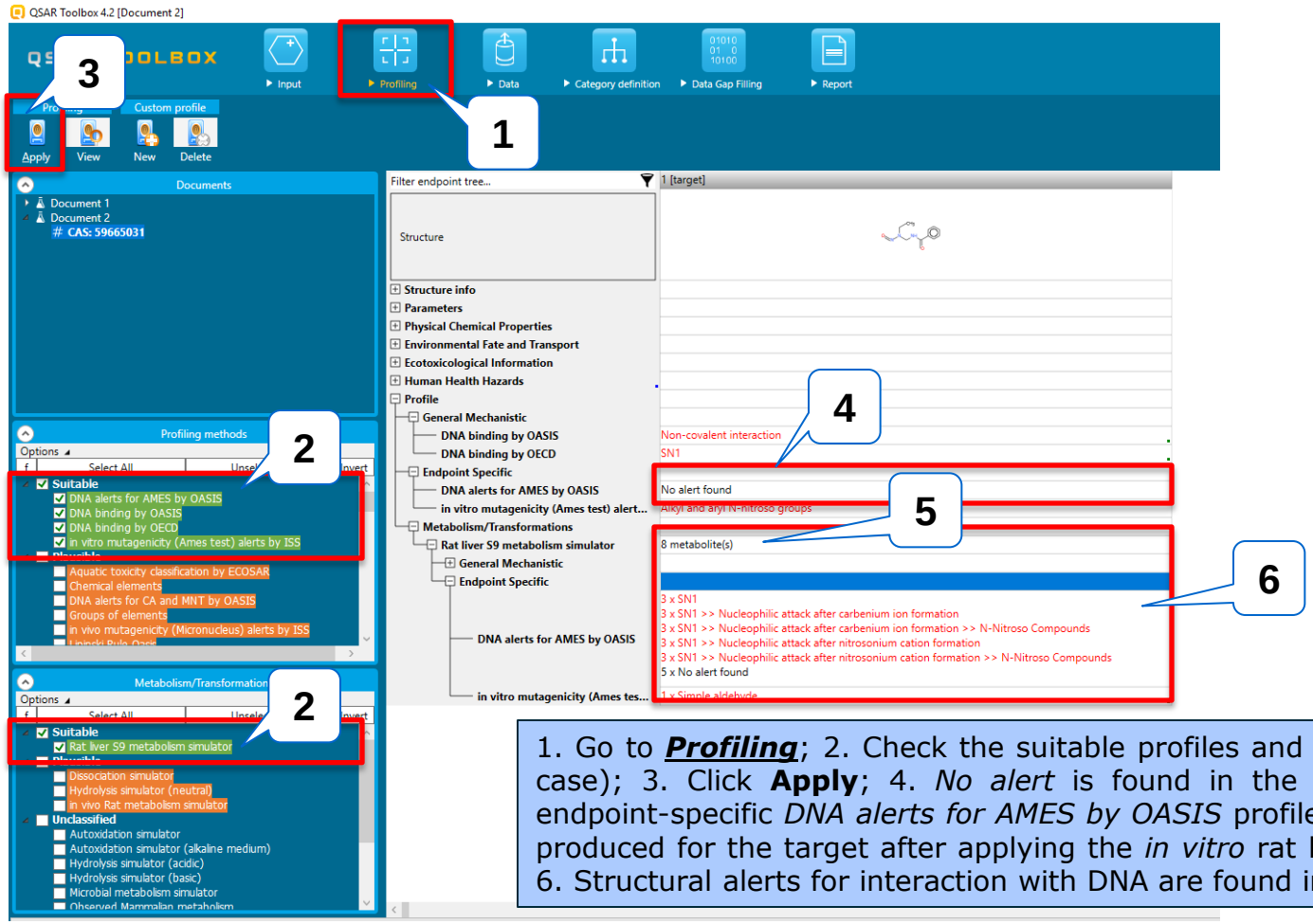
Profiling
Overview

- "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 once 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 relavancy of profilers see ppt: *Example for predicting skin sensitization taking into account alert performance*

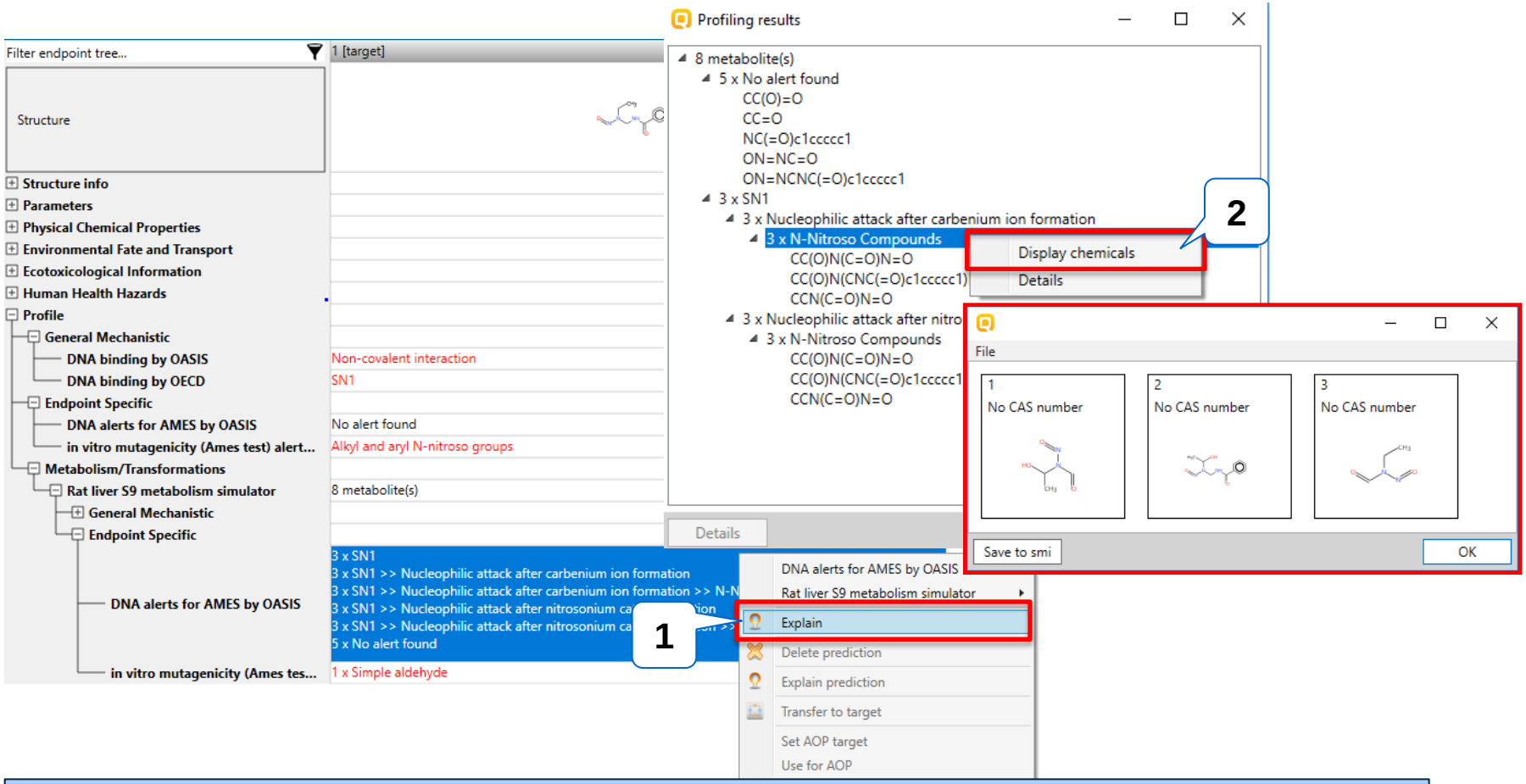
Profiling

Profiling the target chemical



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



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 genarted metabolites where active alert has been found.

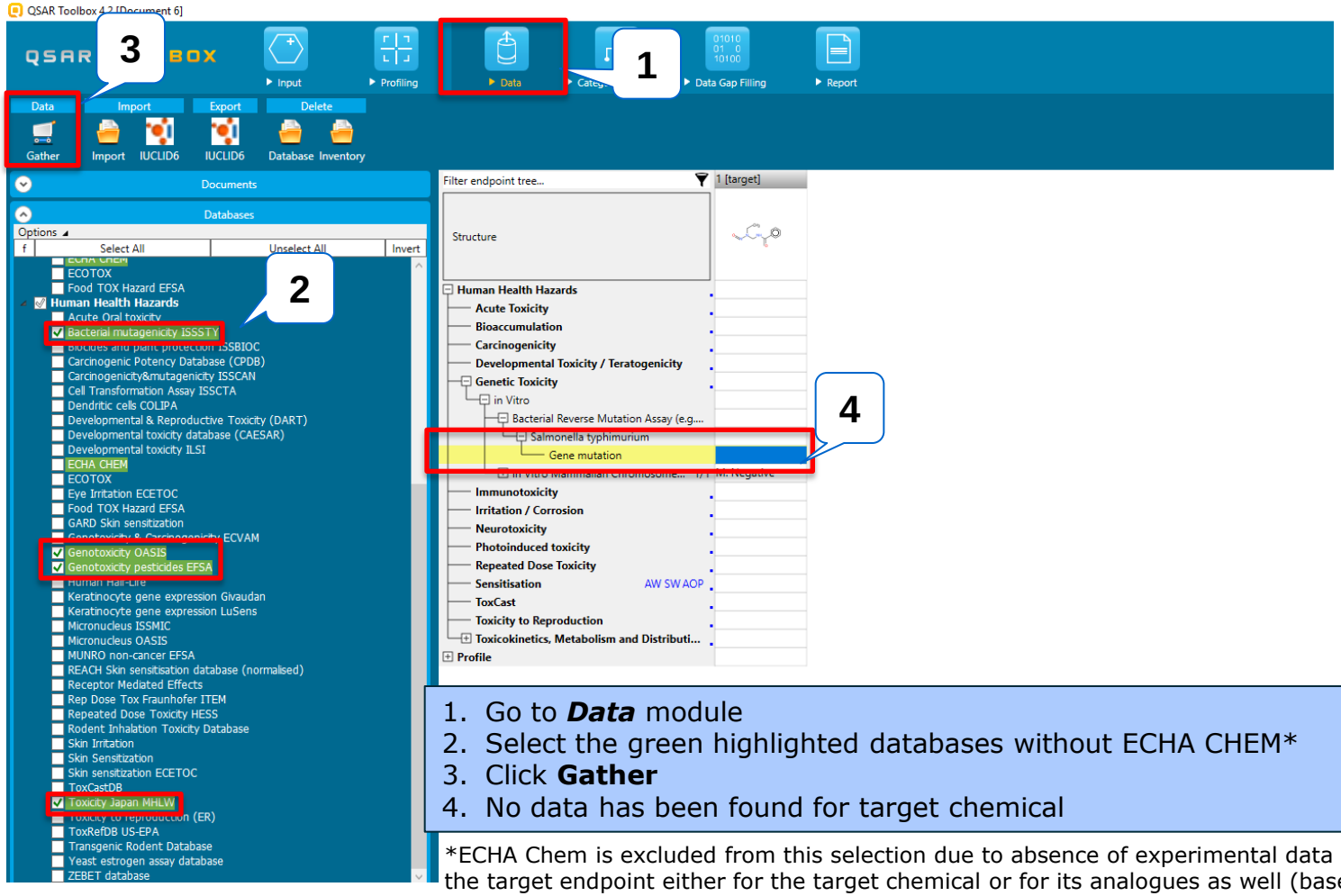
Outlook

- 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.

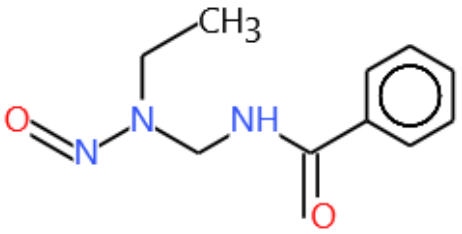
- In the third module (Data), you have found that the target chemical has no data associated with the target endpoint

- The study continues with investigating profiling results and data of generated rat liver metabolites of target chemical (see next slides).

Recap

Parent

Benzamide, N-((ethylnitrosoamino)methyl)
CAS# 59665-03-1



DNA binding alert for AMES

No alert found

Profiling Result

No alert found (no AMES mutagenicity effect is expected)

Rat liver metabolism – 3 of the metabolites have N-nitroso DNA binding SA

Investigating DNA prof + data for the metabolites

Outlook

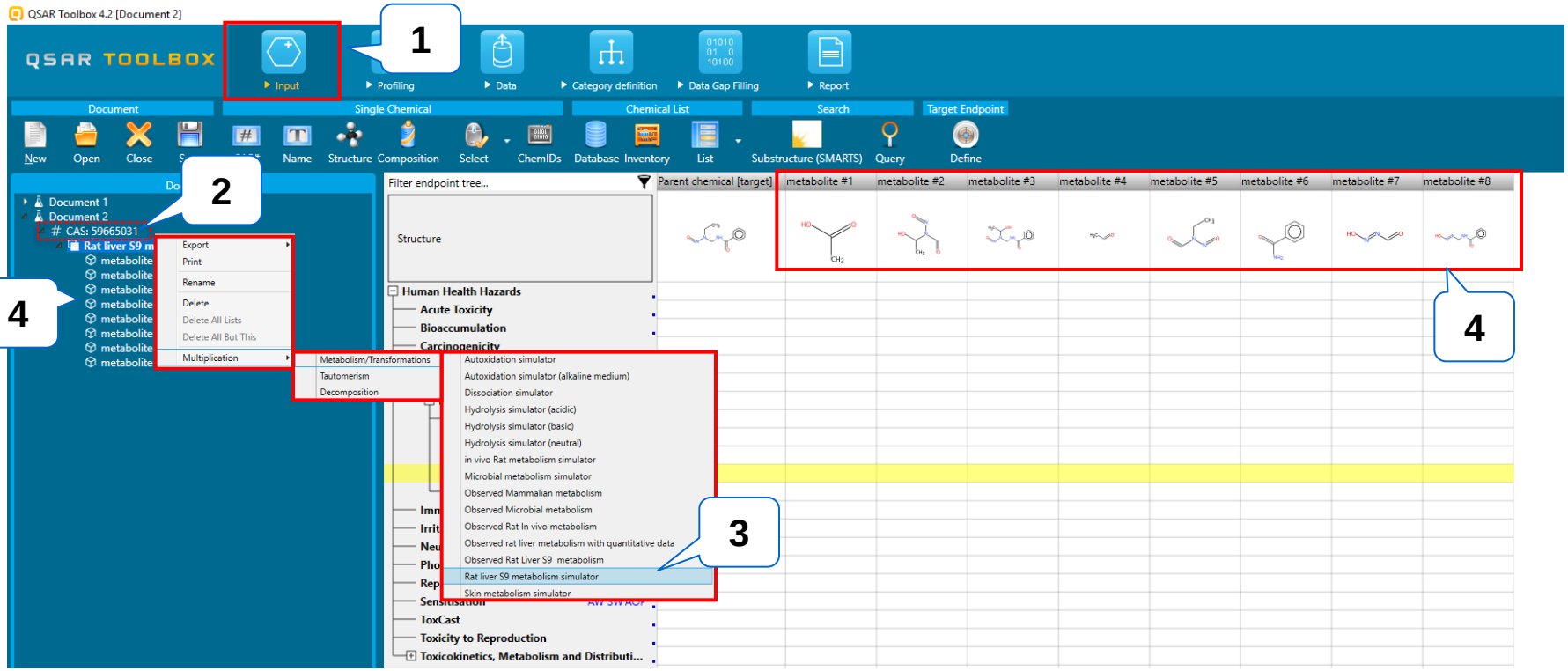
- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- Example
  - Input
  - Define target endpoint
  - Profiling
  - Data
  - **Simulation of rat liver S9 metabolism**

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



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

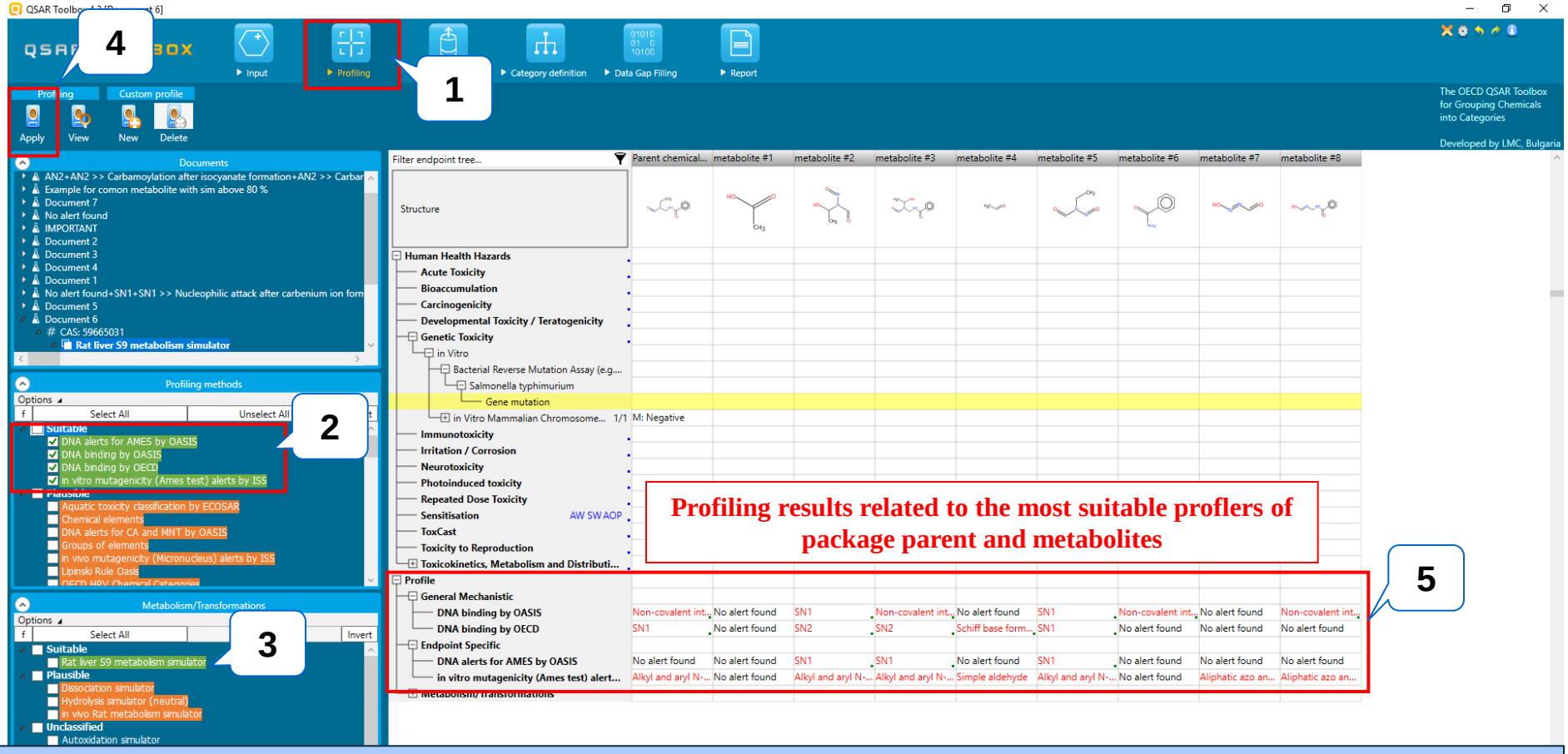
Next actions are focused on investigating the profilers of the generated metabolites and collecting data for them

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

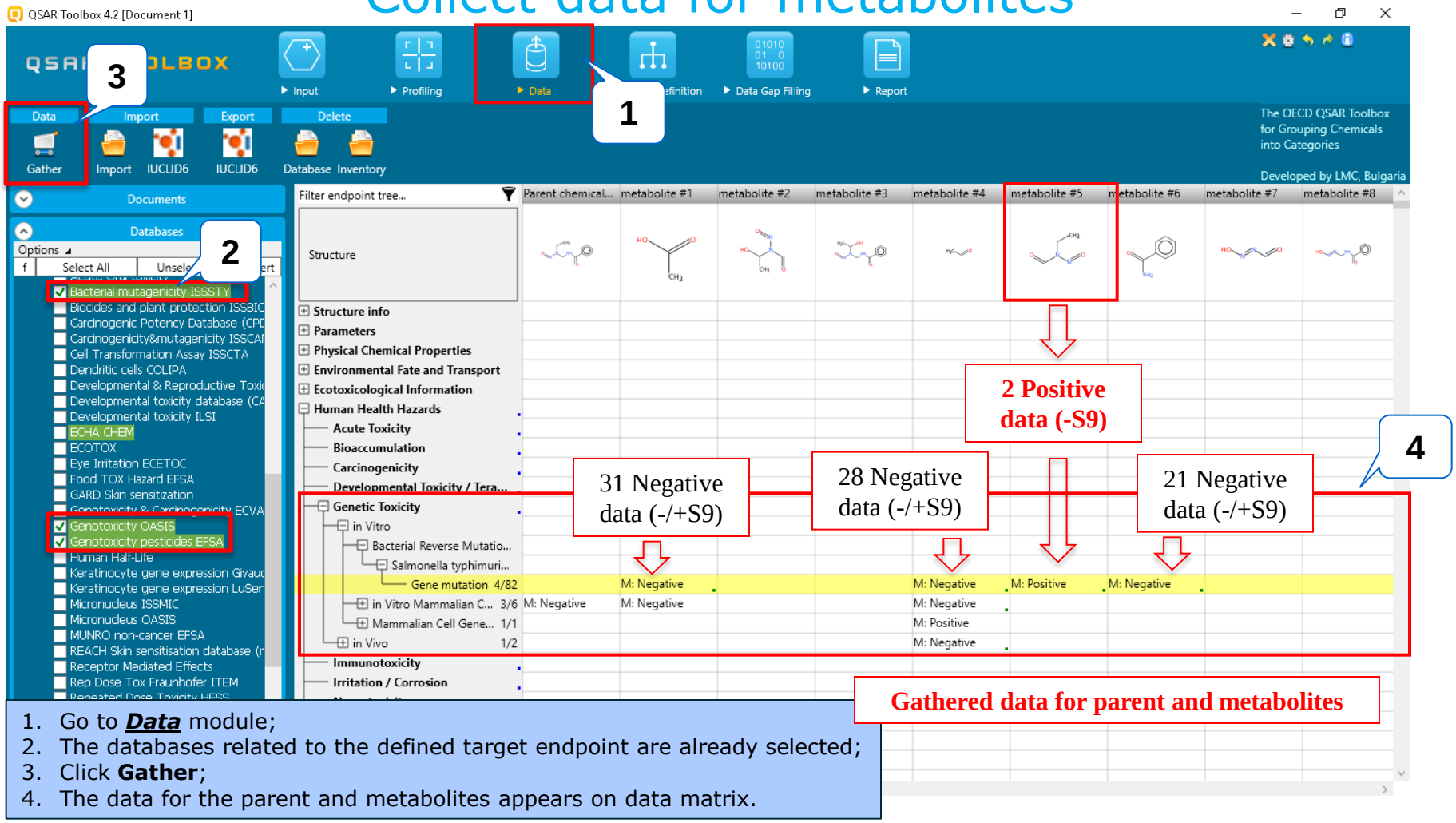
Profiling the package of metabolites



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

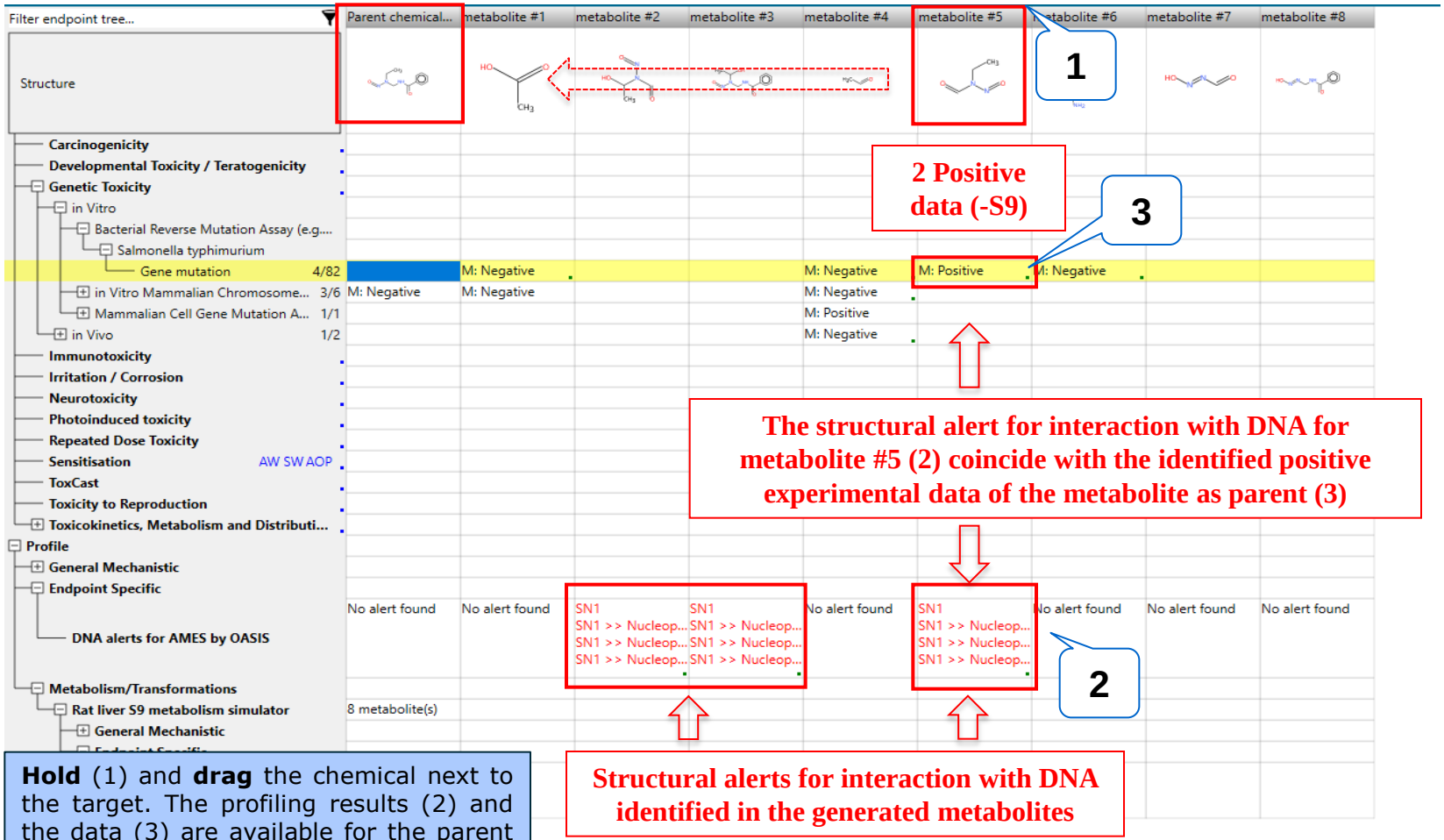


1. Go to ***Data*** module;
2. The databases related to the defined target endpoint are already selected;
3. Click **Gather**;
4. The data for the parent and metabolites appears on data matrix.

Outlook

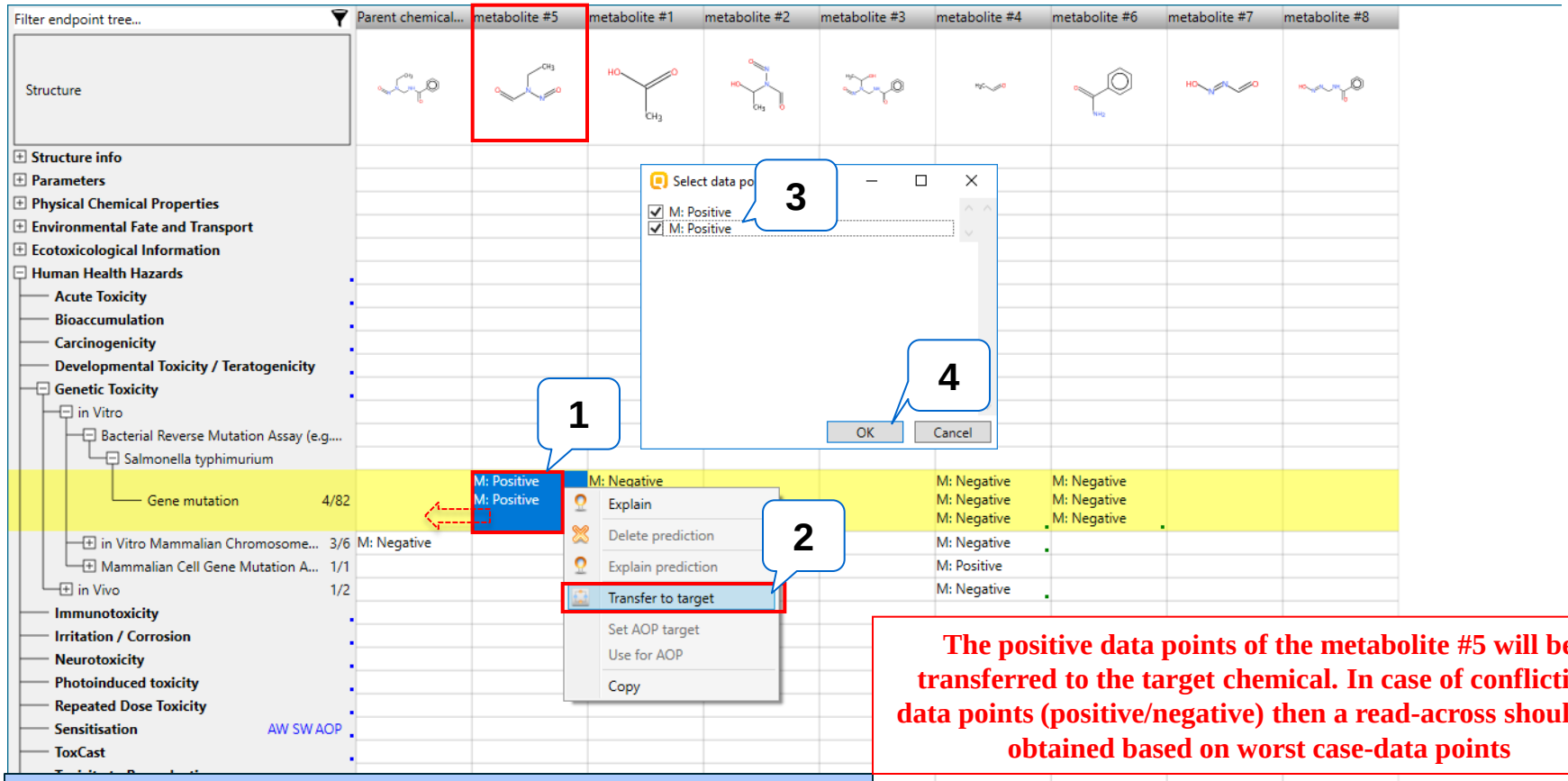
- Background
- Objectives
- Specific aim
- Manipulation of data matrix
- Example
  - Input
  - Define target endpoint
  - Profiling
  - Data
  - Generation of rat liver S9 metabolism
  - Profiling and collecting data for metabolites
  - **Transferring data to the target outside data gap filling module**

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).

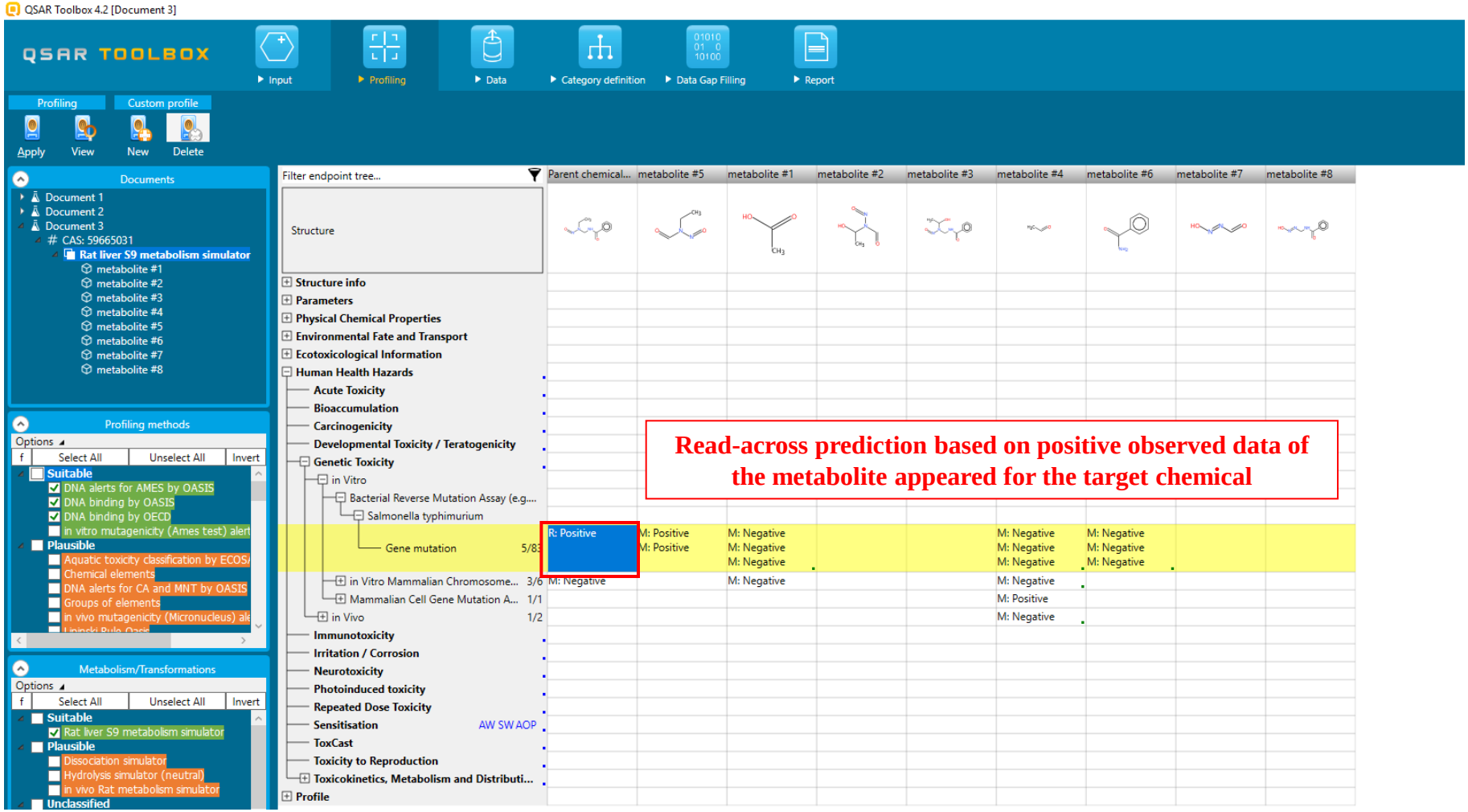
Transferring experimental data of metabolite to the target outside data gap filling



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

1. Right-click the cell with observed data of the metabolite #5
2. Select **Transfer to target**
3. Select which data points to be transfered to the target
4. Click **OK** button

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



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.