QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

OECD QSAR Toolbox v.3.4

Step-by-step example of how to predict Ames mutagenicity for a chemical by a qualitative read-across approach

Outlook

- **Background**
- Objectives
- Specific Aims
- Read-across
- The exercise
- Workflow of the exercise
- Save the prediction

Background

- This is a step-by-step presentation designed to take you through the workflow of the Toolbox in a data-gap filling exercise using read-across based on molecular similarity with data pruning.
- If you are a novice user of the Toolbox you may wish to review the "Getting Started" document available at [www.oecd.org/env/existingchemicals/qsar] as well as go through tutorials 1 and 2.

Outlook

- Background
- **Objectives**
- Specific Aims
- Read-across
- The exercise
- Workflow of the exercise
- Save the prediction

Objectives

- **This presentation demonstrates a number of functionalities of the Toolbox:**
  - Entering a target chemical by SMILES notation and Profiling
  - Identifying analogues for a target chemical by molecular similarity
  - Retrieving experimental results available for those analogues, and for multiple endpoints
  - Filling data gaps by read-across

Outlook

- Background
- Objectives
- **Specific Aims**
- Read-across
- The exercise
- Workflow of the exercise
- Save the prediction

Specific Aims

- To review the workflow of the Toolbox.
- To reacquaint the user with the six modules of the Toolbox.
- To reacquaint the user with the basic functionalities within each module.
- To introduce the user to new functionalities of selected modules.
- To explain to the rationale behind each step of the exercise.

Outlook

- Background
- Objectives
- Specific Aims
- **Read-across**
- The exercise
- Workflow of the exercise
- Save the prediction

Read-across & the Analogue Approach

- Remember, read-across is a method that can be used to estimate missing data from a single or limited number of chemicals using the analogue approach.
- In the analogue approach, experimental endpoint information for a single or small number of tested chemicals is used to predict the same endpoint for an untested chemical that is considered to be "similar" (i.e., within the same category).

Analogous Chemicals

- Previously you learned that analogous sets of chemicals are often selected based on the hypothesis that the toxicological effects of each member of the set will show a common behaviour.
- For this reason mechanistic profilers and grouping methods have been shown to be of great value in using the Toolbox.
- However, there are cases where the mechanistic profilers and grouping methods are inadequate and one is forced to rely on molecular similarity to form a category.
- The Toolbox allows one to develop a category by using either a mechanistic category like DNA binding or structural similarity.
- Since there is no preferred way of identifying structural similarity, the user is guided to use DNA binding as a first option.

Outlook

- Background
- Objectives
- Specific Aims
- Read-across
- **The exercise**
- Workflow of the exercise
- Save the prediction

Exercise

- In this exercise we will predict the Ames mutagenicity potential for an untested compound, (n-hexanal) [SMILES CCCCCC=O)], which is the "target" chemical.
- This prediction will be accomplished by collecting a small set of test data for chemicals considered to be in the same category as the target molecule.
- The category will be defined by empirical similarity, with respect to "Organic functional groups" profiler.
- The prediction itself will be made by "read-across" analysis.

Side-Bar On Mutagenesis

- Mutagens do not create mutations.
- Mutagens create DNA damage.
- Mutations are changes in nucleotide sequence.
- Mutagenesis is a cellular process requiring enzymes and/or DNA replication, thus cells create mutations.

Side-Bar On Mutagenesis

- Mutations within a gene are generally base-substitutions or small deletions/insertions (i.e., frame shifts).
- Such alteration are generally called point mutations.
- The Ames scheme based on strains of *Salmonella* provide the corresponding experimental data.

Side-Bar On Mutagenesis

- The Ames mutagenicity assay (see OECD guideline 471) is designed to assess the ability of a chemical to cause point mutations in the DNA of the bacterium *Salmonella typhimurium*.
- The Ames test includes a number of strains (TA1537, TA1535, TA100, TA98 and TA97) that have been engineered to detect differing classes of mutagenic chemicals.
- The basic test only detects direct acting mutagens (i.e., those chemicals able to interact with DNA without the need for metabolic activation).

Side-Bar on Metabolic Activation

- The inclusion of an S9 mix of rodent liver enzymes is designed to assess those chemicals requiring metabolic activation in order to be mutagenic.
- Typically, chemicals are assayed both without S9 and with S9 with results being reported in a binary fashion
- A positive result in any of the bacterial strains with or without S9 confirms mutagenic potential.

Outlook

- Background
- Objectives
- Specific Aims
- Read-across
- The exercise
- **Workflow of the exercise**
- Save the prediction

Workflow

- **The Toolbox has six modules which are used in a sequential workflow:**
  - Chemical Input
  - Profiling
  - Endpoints
  - Category Definition
  - Filling Data Gaps
  - Report

Outlook

- Background
- Objectives
- Specific Aims
- Read-across
- The exercise
- **Workflow of the exercise**
  - **Chemical input**

Chemical Input
Overview

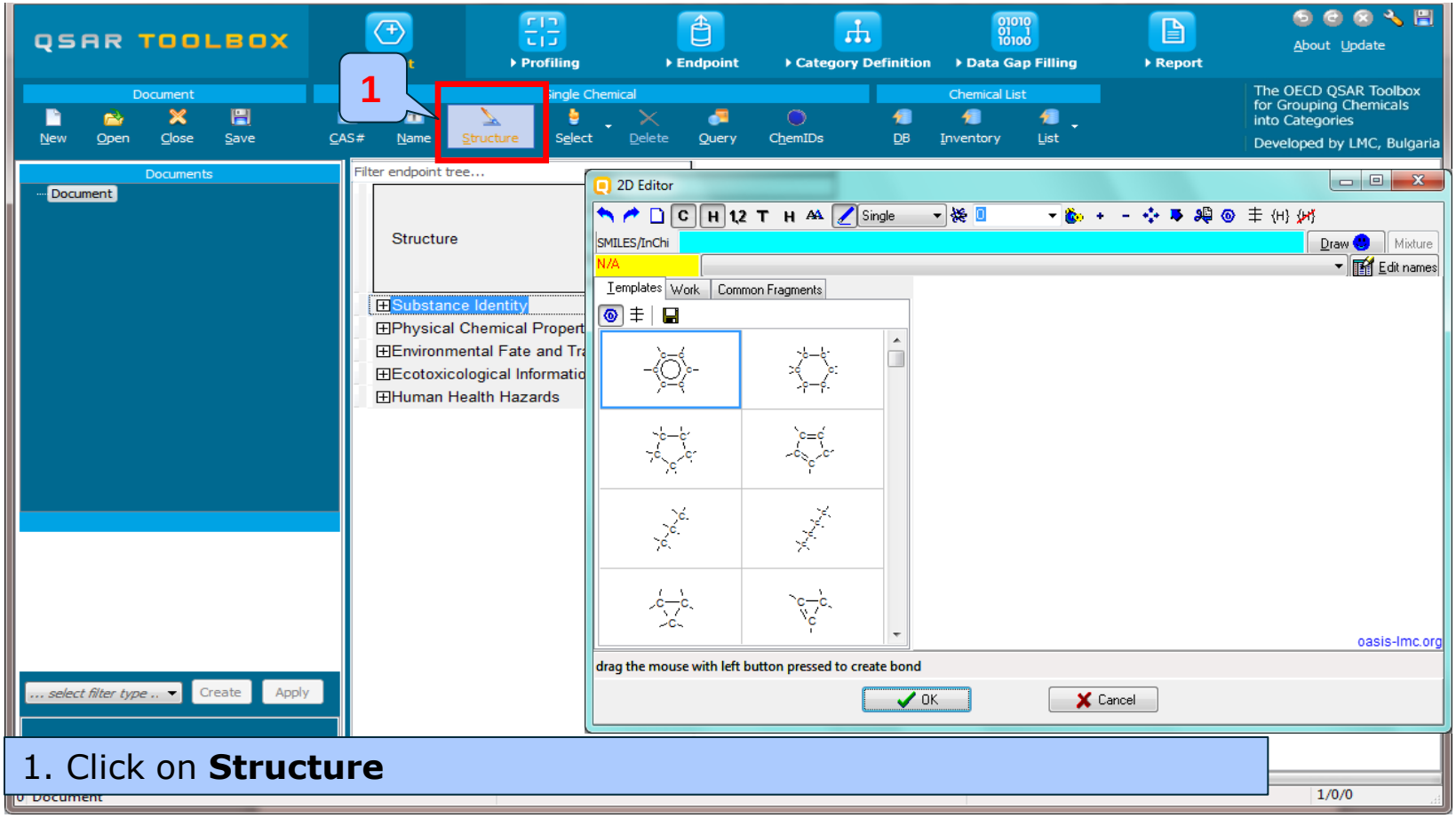
- As you leader in the previous tutorials, this module provides the user with several means of entering the chemical of interest or the target chemical.
- Since all subsequent functions are based on chemical structure, the goal here is to make sure the molecular structure assigned to the target chemicial is the correct one.

Chemical Input
Ways of Entering a Chemical

- **Remember there are several ways to enter a target chemical and the most often used are:**
  - CAS#,
  - SMILES (simplified molecular information line entry system) notation, and
  - Drawing the structure
- Click on **Structure**.
- This inserts the window entitled "2D editor" (see next screen shot).

Chemical Input

Input target chemical by drawing



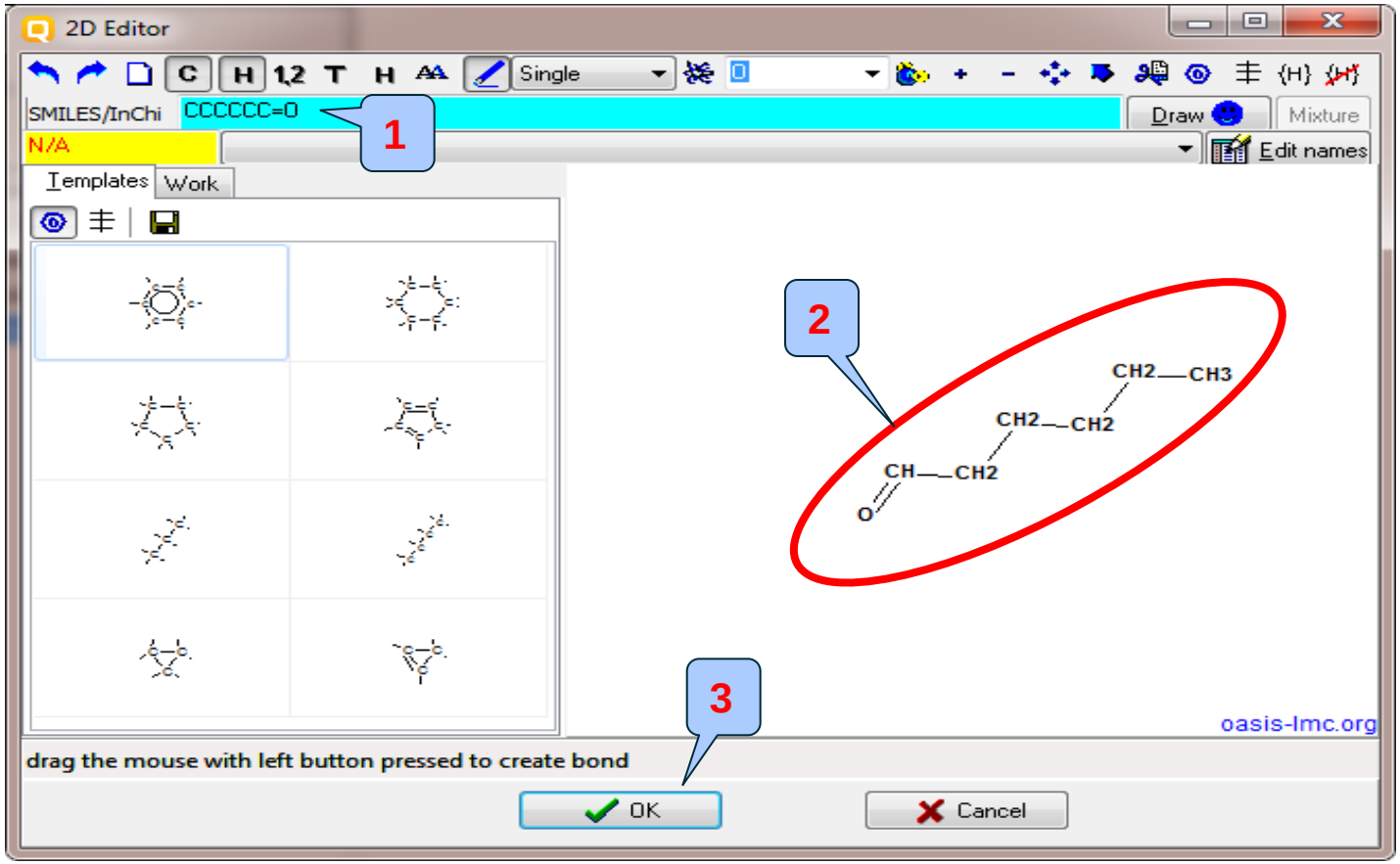
1. Click on **Structure**

Chemical Input
Input target chemical by SMILES

- In the Aqua-coloured area next to "SMILES/InChi" type CCCCCC=O.
- Note as you type the SMILES code the structure is being drawn in the centre of the structure field (see next screen shot).
- Click "OK" to accept the target chemical.

Chemical Input

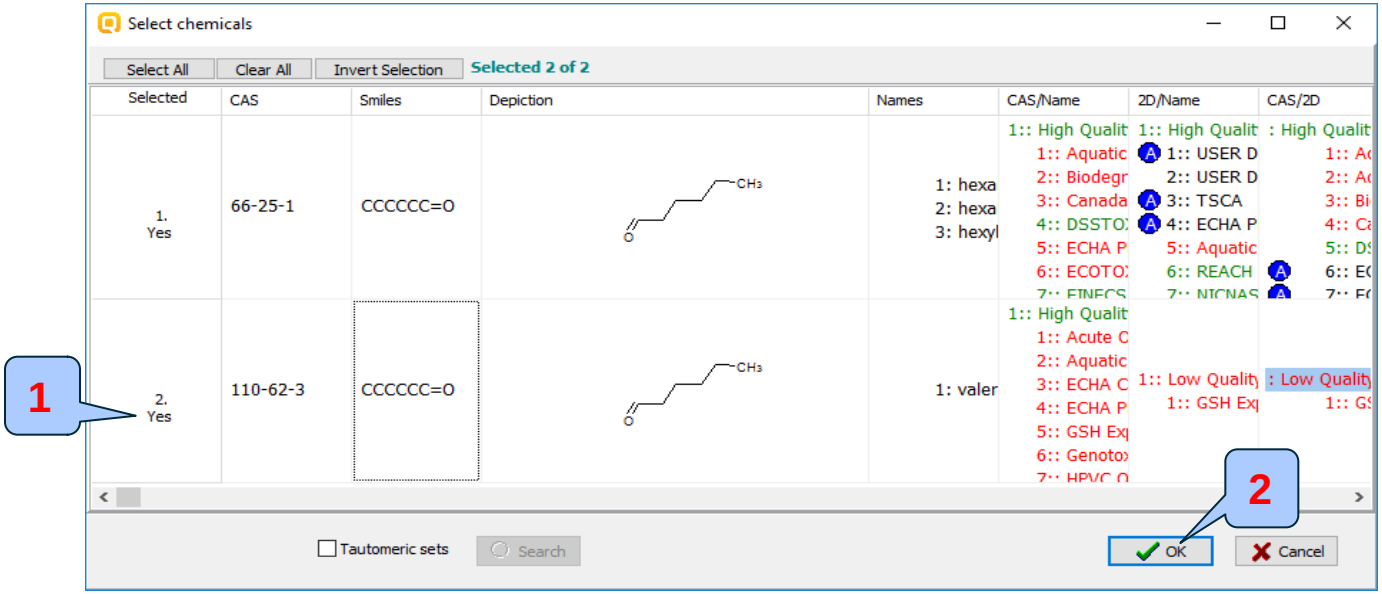
Input target chemical by SMILES



1. **Type** CCCCCC=O in SMILES/InChi window; 2. 2D structure; 3. **Click** OK

Chemical Input
Input target chemical by SMILES

The Toolbox now searches the Toolbox databases and inventories for the presence of the chemical with structure related to the current SMILES notation. It is depicted as a 2D image.
Two chemicals are found. All found chemicals are selected by default.

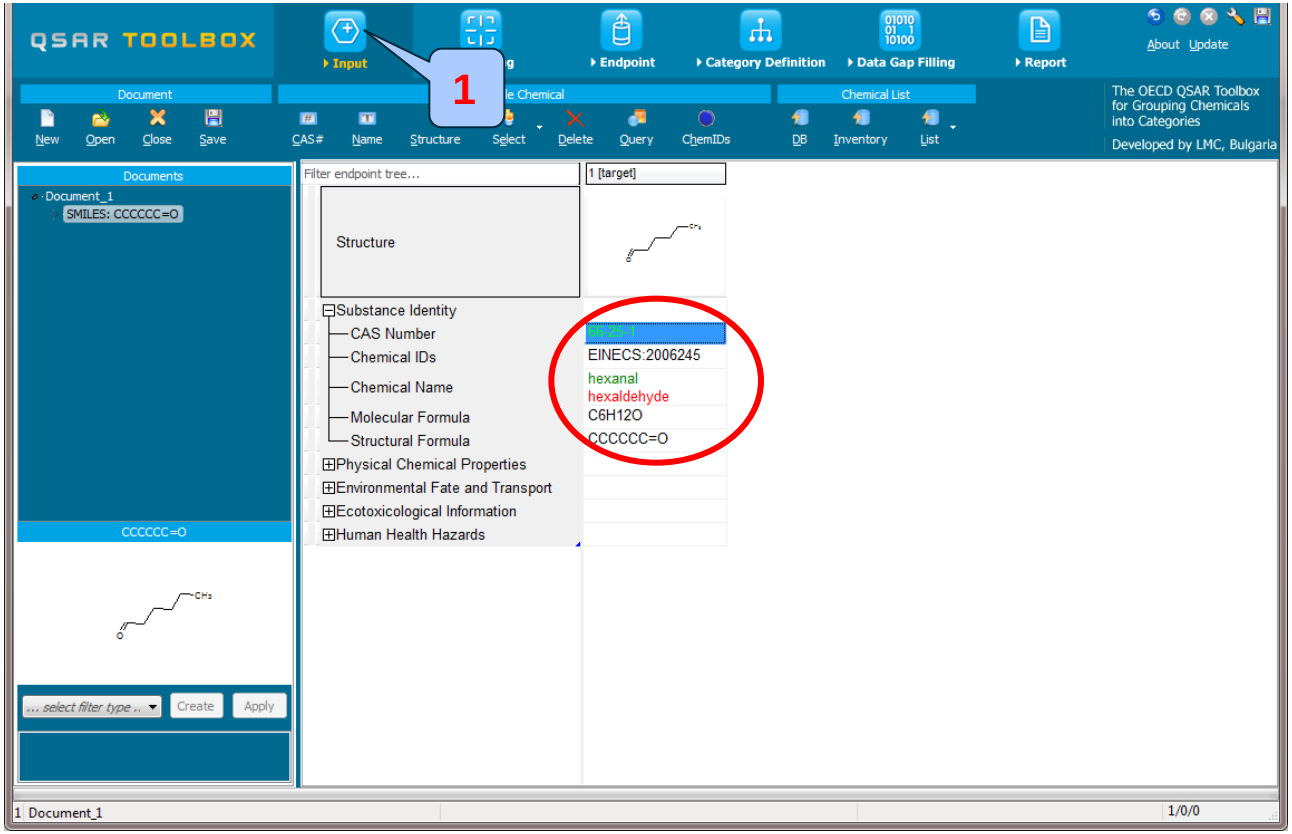


1. **Unselect** the second chemical by clicking on the "Yes"; 2. **Click** OK.

Chemical Input
Target chemical identity

- You have now selected your target chemical.
- Click on the box next to "Substance Identity"; this displays the chemical identification information (see next screen shot).
- It is important to remember that the workflow is based on the structure coded in SMILES.

Chemical Input
Target chemical identity



The workflow on the first module is now complete; click on "Profiling" [**1**] to move to the next module.

Outlook

- Background
- Objectives
- Specific Aims
- Read-across
- The exercise
- **Workflow of the exercise**
  - Chemical input
  - **Profiling**

Profiling
Overview

- As you may remember, “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.
- Available profilers includes likely mechanism(s) of action which have been show to be useful in forming categories that include the target chemical.

Profiling
Profiling the target chemical

- The outcome of the profiling determines the most appropriate way to search for analogues (detailed information in Manual for getting started (Chapter 4) http://www.oecd.org/dataoecd/58/56/46210452.pdf
- Table 4-1 in chapter 4 (Manual for getting started) lists a selection of profilers and their relevance for different endpoints of regulatory relevance.
- For this example, the following general mechanistic profiling methods are relevant to genetic toxicity:

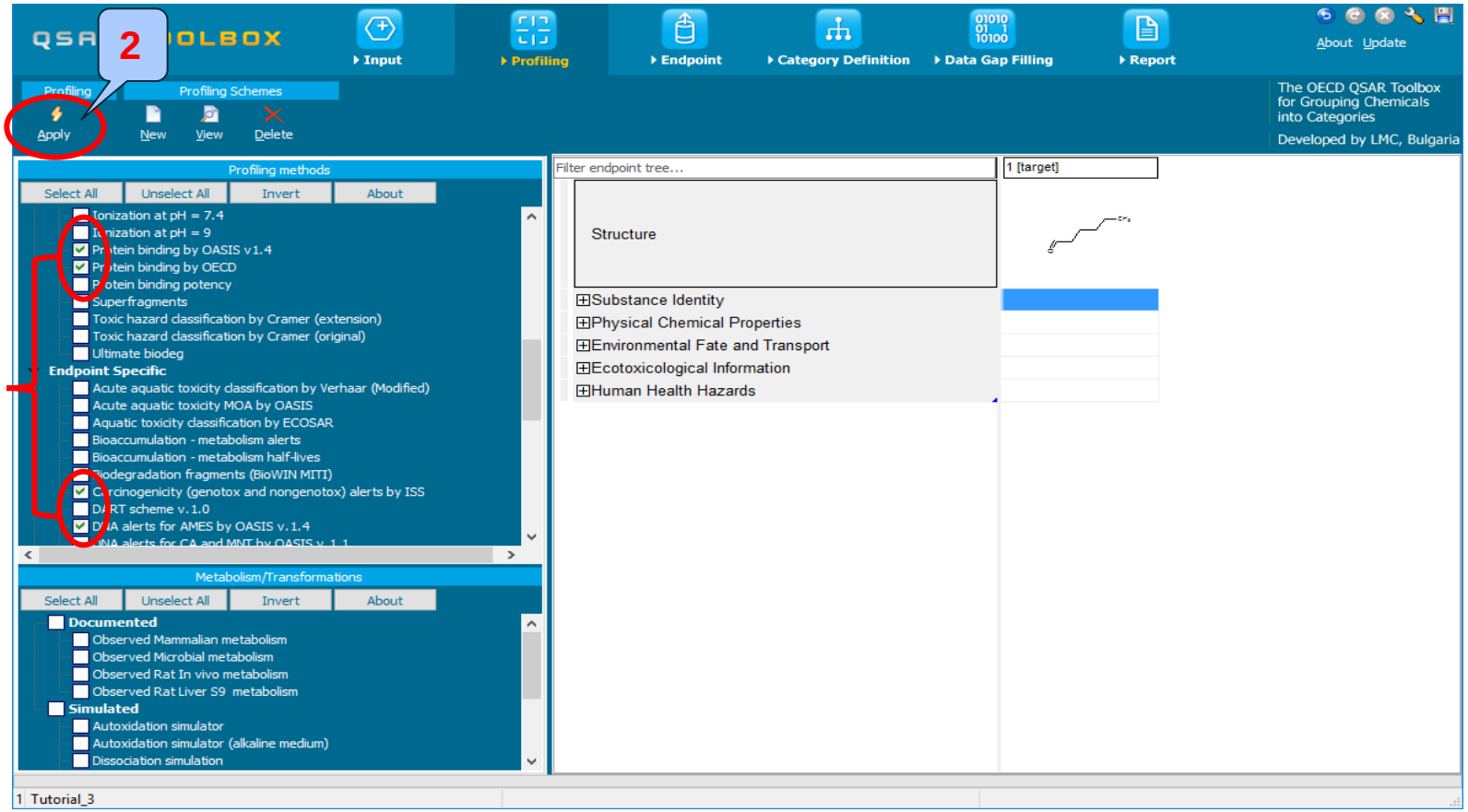
  - DNA binding by OASIS v1.4– mechanistic grouping
  - DNA binding by OECD – mechanistic grouping
  - Protein binding by OASIS v1.4 – mechanistic grouping
  - Protein binding by OECD – mechanistic grouping
  - Carcinogenicity (genotox and nongenotox) alerts by ISS - endpoint specific
  - DNA alerts for AMES by OASIS v.1.4 - endpoint specific
  - in vitro mutagenicity (Ames test) alerts by ISS - endpoint specific
  - in vivo mutagenicity (Micronucleus) alerts by ISS - endpoint specific
  - Organic function groups - empiric

Profiling
Profiling the target chemical

- Select the "Profiling methods" related to the target endpoint.
- This selects (a green check mark appears) or deselects(green check disappears) profilers.
- For this example, select the profilers relevant to genetic toxicity (see next screen shot).

Profiling

Profiling the target chemical

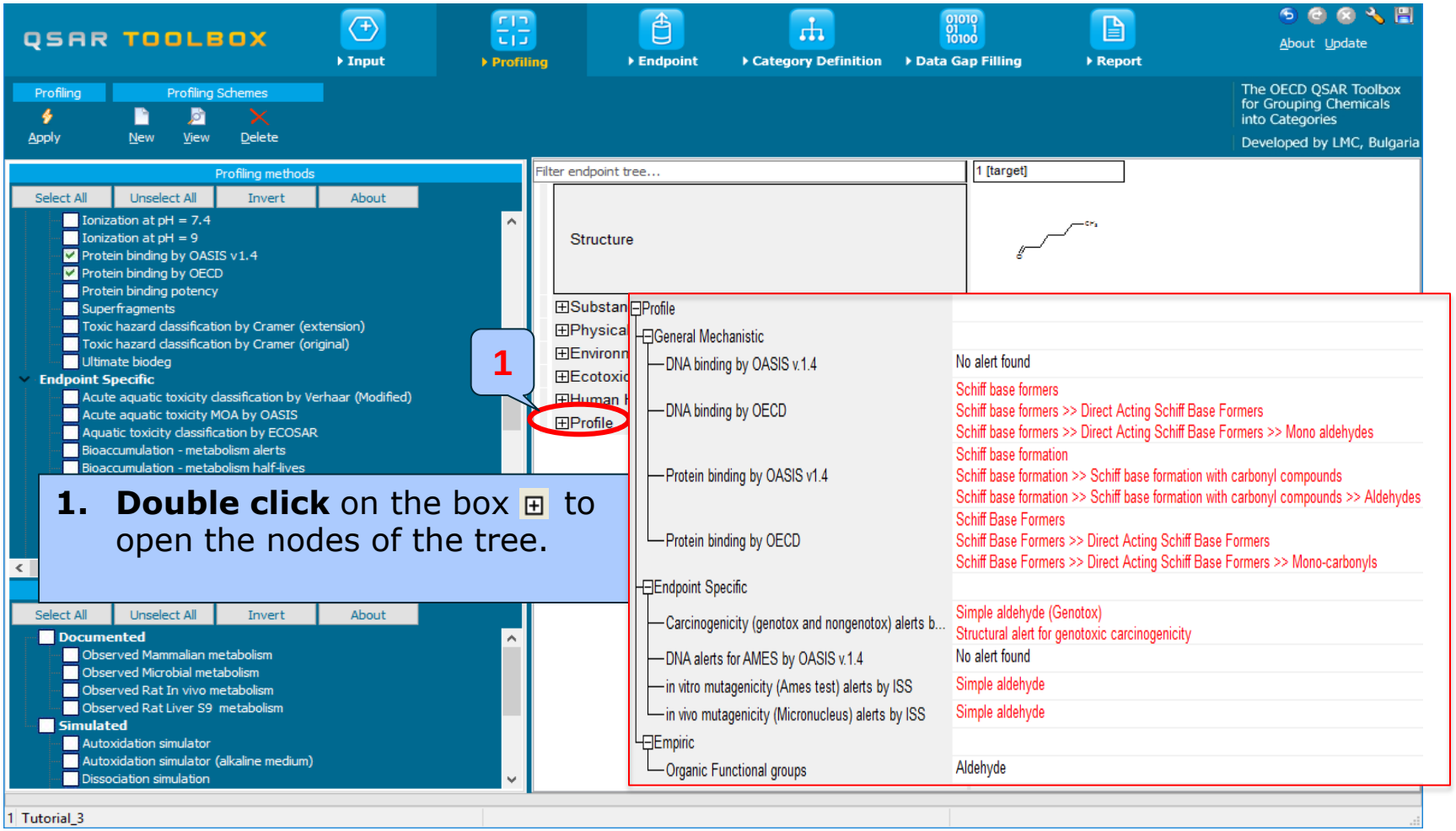


1. Check the profilers related to the target endpoint (see slide 30); **2. Click** Apply

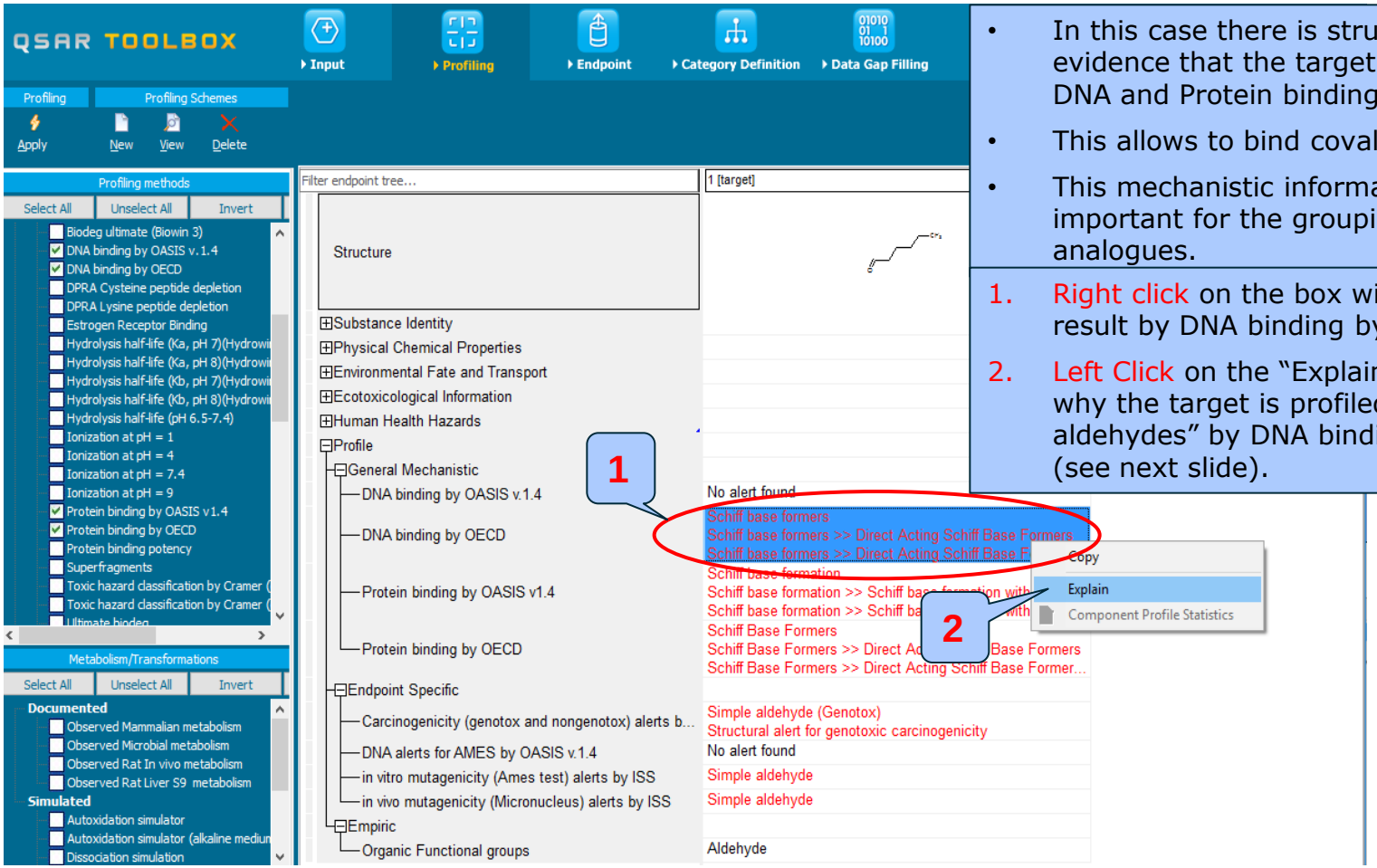
Profiling
Profiling the target chemical

- The actual profiling will take several seconds depending on the number and type of selected profilers.
- The results of profiling automatically appear as a dropdown box under the target chemical (see next slide).
- Please note the specific profiling results by DNA, Protein binding, and Organic functional groups.
- These results will be used to search for suitable analogues in the next steps of the exercise.

Profiling
Profiles of n-hexanal



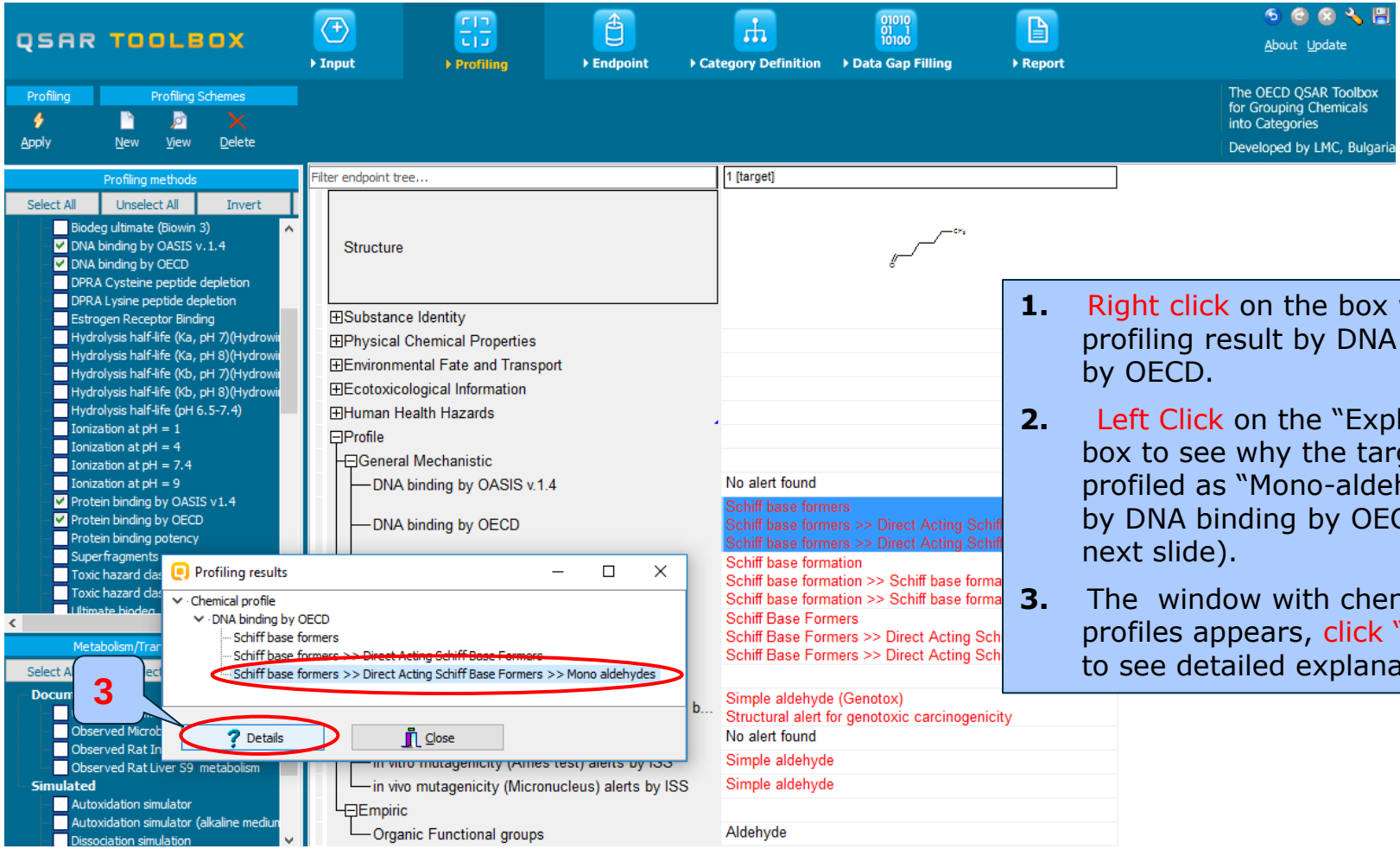
Profiling
Profiles of n-hexanal



- In this case there is structural evidence that the target has positive DNA and Protein binding alert
- This allows to bind covalently to DNA
- This mechanistic information is important for the grouping of analogues.

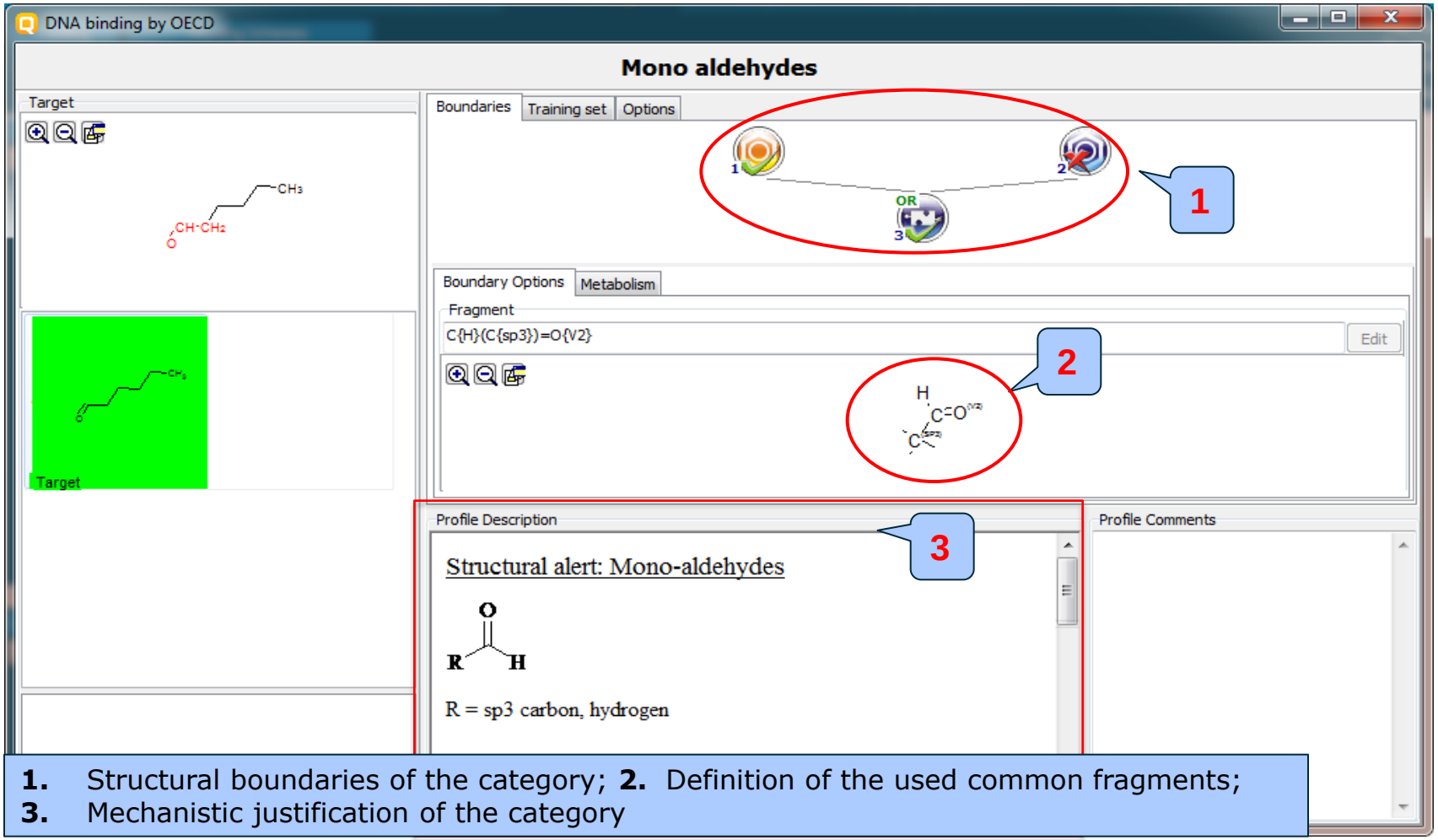
1. Right click on the box with profiling result by DNA binding by OECD.
2. Left Click on the "Explain" box to see why the target is profiled as "Mono-aldehydes" by DNA binding by OECD (see next slide).

Profiling
Profiles of n-hexanal



1. Right click on the box with profiling result by DNA binding by OECD.
2. Left Click on the "Explain" box to see why the target is profiled as "Mono-aldehydes" by DNA binding by OECD (see next slide).
3. The window with chemical profiles appears, click "Details" to see detailed explanation

Profiling
DNA binding by OECD of n-hexanal



1. Structural boundaries of the category; **2.** Definition of the used common fragments;
3. Mechanistic justification of the category

Outlook

- Background
- Objectives
- Specific Aims
- Read-across
- The exercise
- **Workflow of the exercise**
  - Chemical input
  - Profiling
  - **Endpoint**

Endpoint
Overview

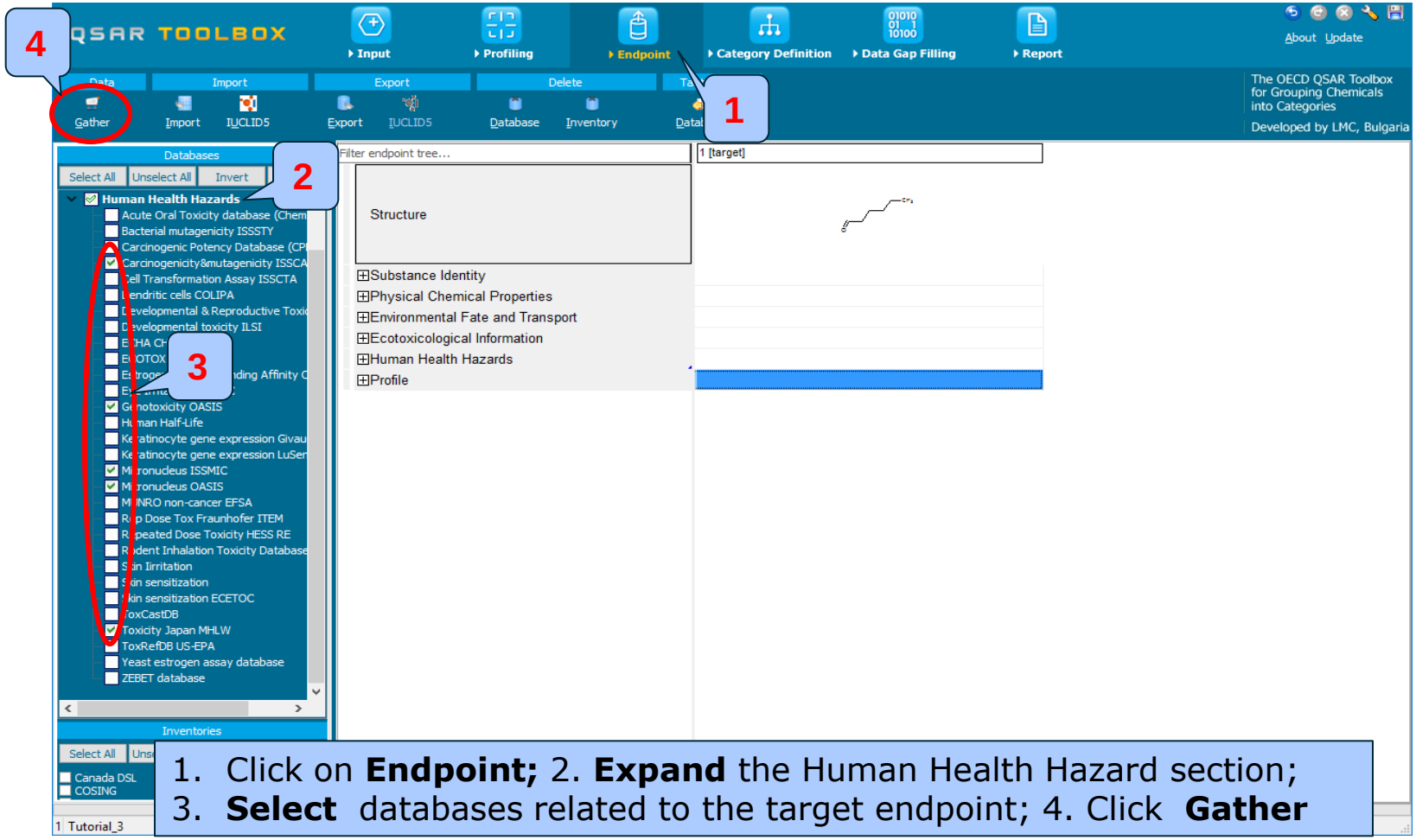
- As you should remember, “Endpoints” refer to the electronic process of retrieving the fate and toxicity data that are stored in the Toolbox database.

- Note, data can be gathered in a global fashion (i.e., collecting all data of all endpoints) or on more narrowly defined settings (e.g., collecting data for a single or limited number of endpoints).

Endpoint
Case study

- In this example, we limit our data gathering to the common genotoxicity endpoints from databases containing genotoxicity data (Carcinogenicity & Mutagenicity ISSCAN, Genotoxicity OASIS, Micronucleus ISSMIC, Micronucleus OASIS and Toxicity Japan MHLW).

Endpoint
Gather data

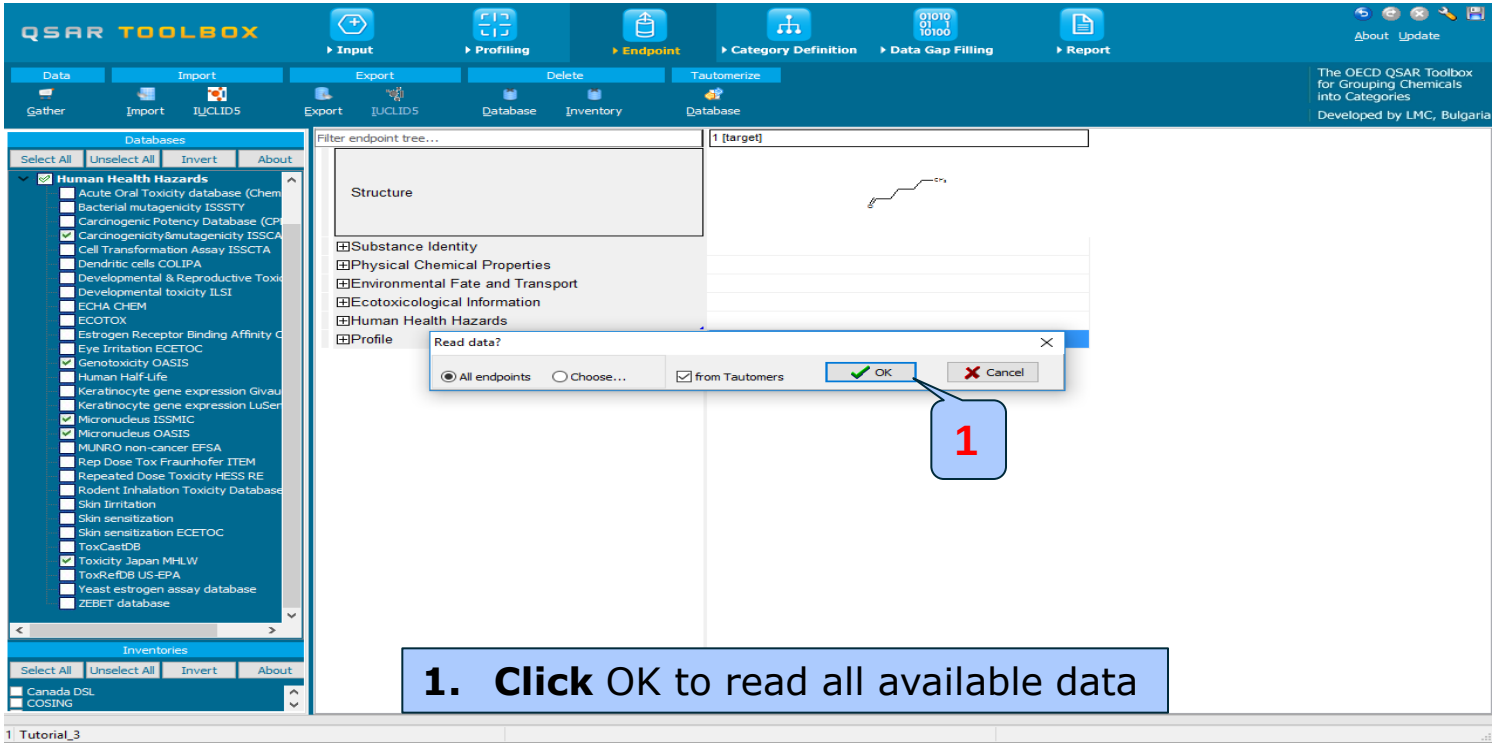


1. Click on **Endpoint;** 2. **Expand** the Human Health Hazard section;
3. **Select** databases related to the target endpoint; 4. Click **Gather**

Endpoint
Process of collecting data

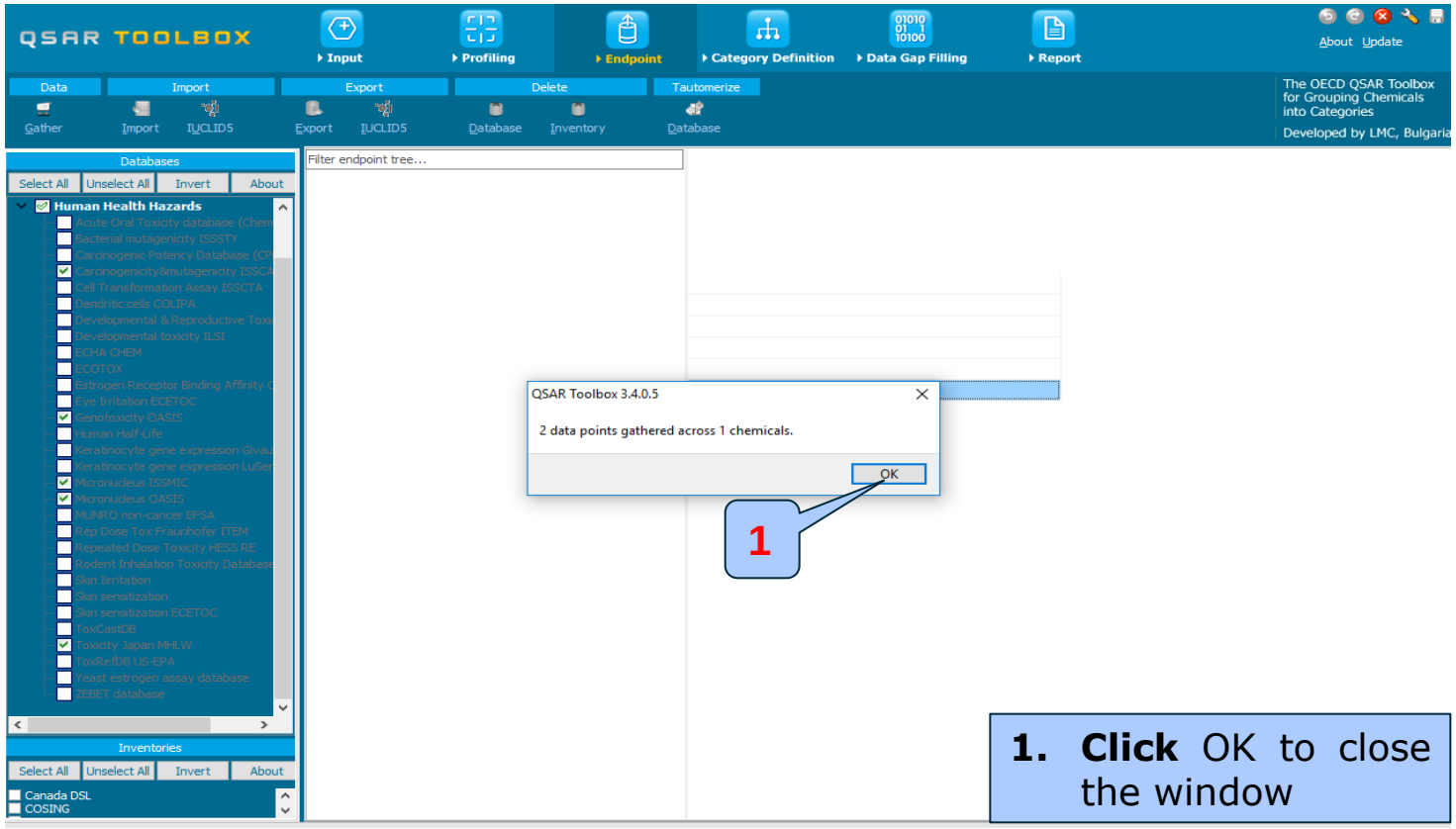
Toxicity information on the target chemical is electronically collected from the selected datasets.

A window with “Read data?” appears. Now the user could choose to collect “all” or “endpoint specific” data.



Endpoint
Process of collecting data

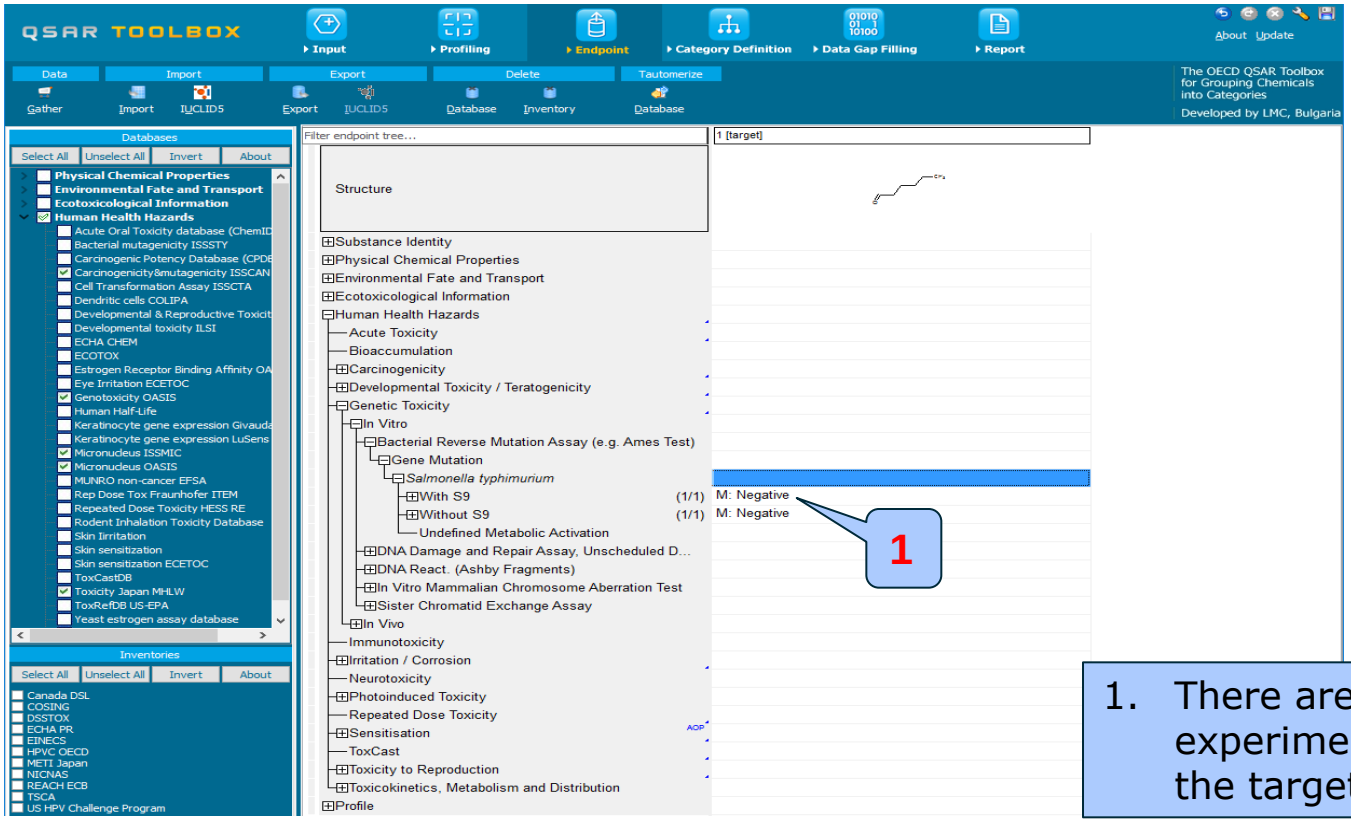
In this example, an insert window appears stating that there was 2 data points available for the target chemical appears.



1. **Click** OK to close the window

Endpoint
Process of collecting data

In this example, an insert window appears stating that there was 2 data points available for the target chemical appears.



1. There are two negative experimental data for the target chemical

Endpoints
Recap

- You have entered the target chemical by SMILES and found it to be n-hexanal with the CAS# [66-25-1].
- You have profiled the target chemical and found 2 experimental data is available for n-hexanal.
- In other words, we will try to reproduce the experimental data by using read-across approach.
- Click on "Category definition" to move to the next module.

Outlook

- Background
- Objectives
- Specific Aims
- Read-across
- The exercise
- **Workflow of the exercise**
  - Chemical input
  - Profiling
  - Endpoint
  - **Category definition**

Category Definition
Overview

- As stated in the previous tutorials, this module provides the user with several means of grouping chemicals into a toxicologically meaningful category that includes the target molecule.
- Remember, this is the critical step in the workflow of the Toolbox.
- Several options are available in the Toolbox to assist the user in defining the category definition.

Category Definition
Side-Bar on Mutagens

- It is important to remember that mutagens are really cell-damaging agents, which can create a wide array of adverse effects beyond damage to DNA.

- Lets take a moment to review our mechanistic profile of the target chemical (see next screen shots).

Category Definition
Grouping methods

- The different grouping methods allow the user to group chemicals into chemical categories according to different measures of "similarity" so that within a category data gaps can be filled by read-across.

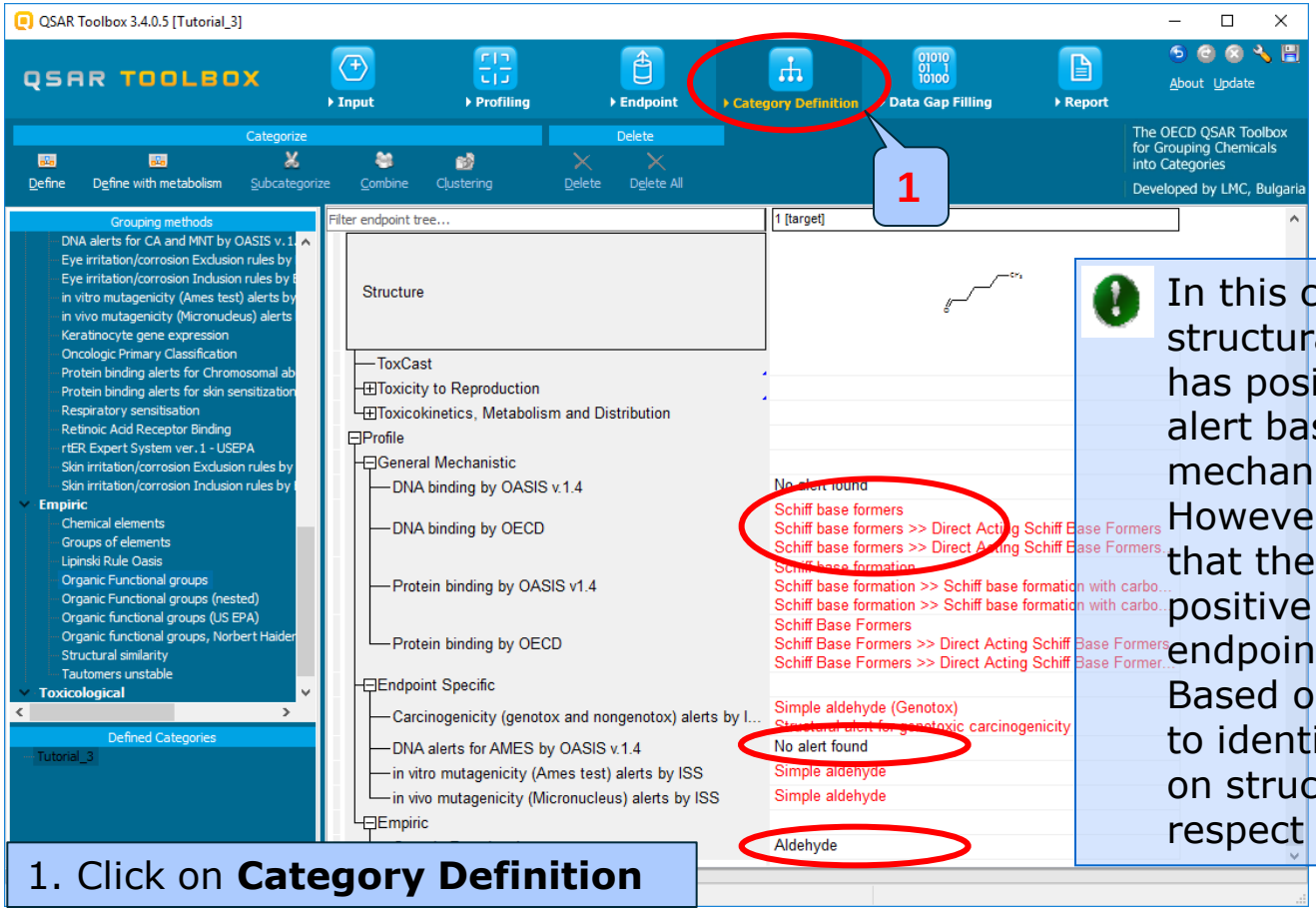
- Detailed information about grouping chemical (Chapter 4) could be downloaded from:
  http://www.oecd.org/dataoecd/58/56/46210452.pdf

- For this example, we will start from a broad group based on Organic functional group and after that

- Will refine the category by a specific DNA binding mechanism identified for the target chemical and find analogues which can bind by the same mechanism and for which experimental results are available.

Category Definition

Which of the category to be defined?

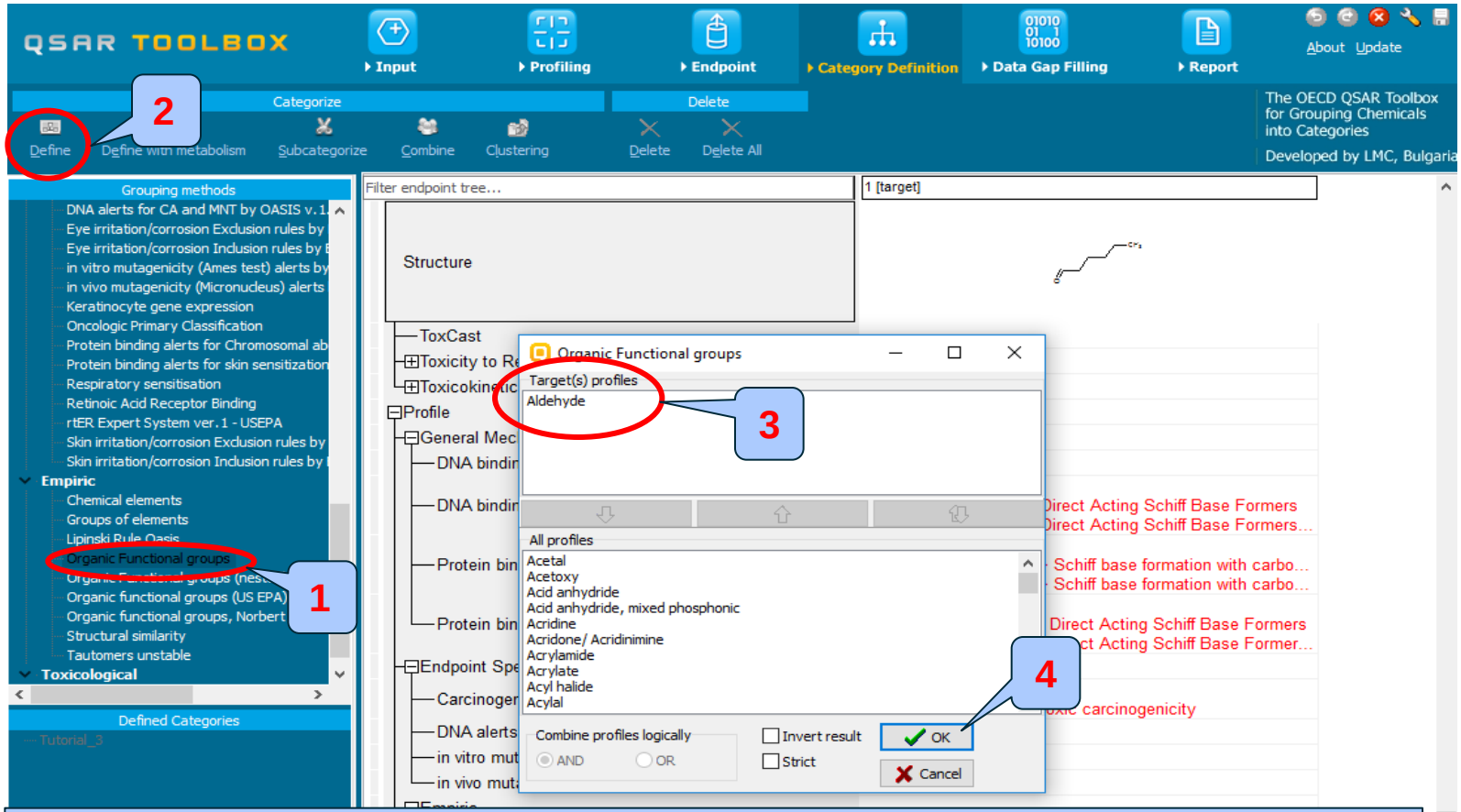


1. Click on **Category Definition**

In this case n-hexanal has structural evidence that it is has positive DNA binding alert based on general mechanistic DNA profilers. However there is no evidence that the target will elicit positive DNA effect based on endpoint specific DNA profiler. Based on this it is appropriate to identify analogues based on structural similarity with respect to OFG profiler.

Category Definition

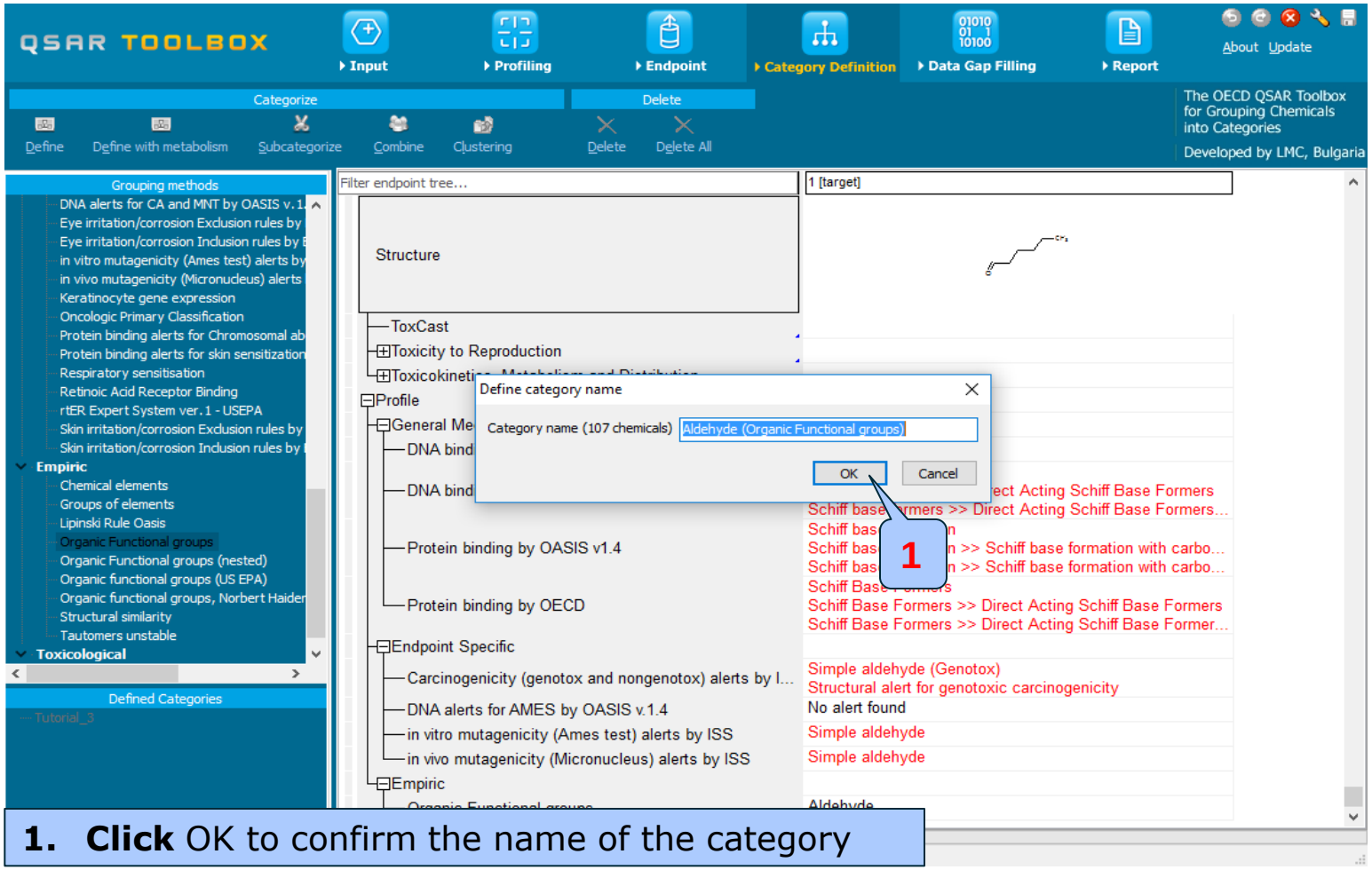
Defining Organic functional group



1. **Highlight** "OFG"; 2. **Click** Define; 3. The target category is Aldehydes Confirm the category; 4. **Click** OK

Category Definition

Defining Organic functional group category

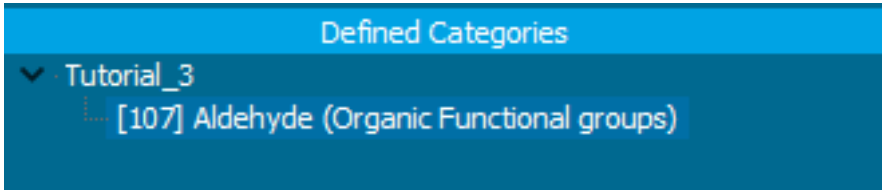


1. **Click** OK to confirm the name of the category

Category Definition
Analogues

- The Toolbox now identifies all chemicals corresponding to category "Aldehydes" by Organic functional groups listed in the databases selected under "Endpoints".

- The name of the category appears in the "Defined Categories" window, the number in brackets is the number of substances belonging to the category (107 analogues including the target chemical are identified)



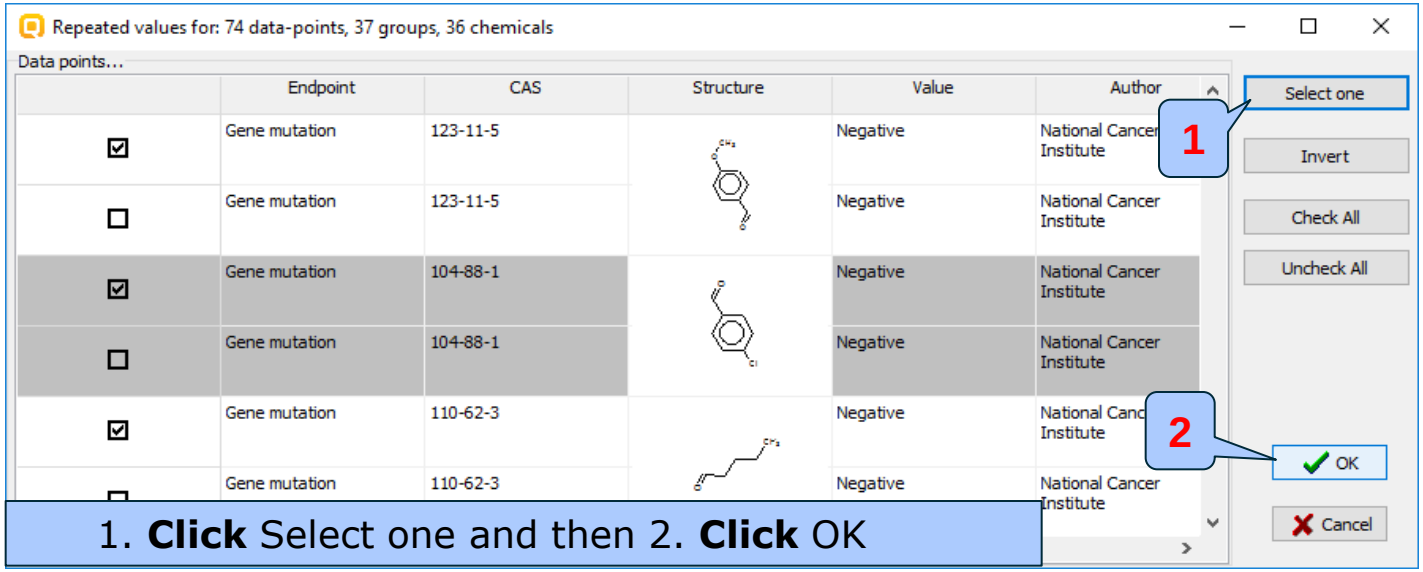
Category Definition
Read data for Analogues

- The Toolbox automatically request the user to select the endpoint that should be retrieved
- The user can either select the specific endpoint or by default choose to retrieve data on all endpoints (see below)

- **In this example, as only databases are selected that contain information for genetic toxicity endpoint, so both options give the same results.**

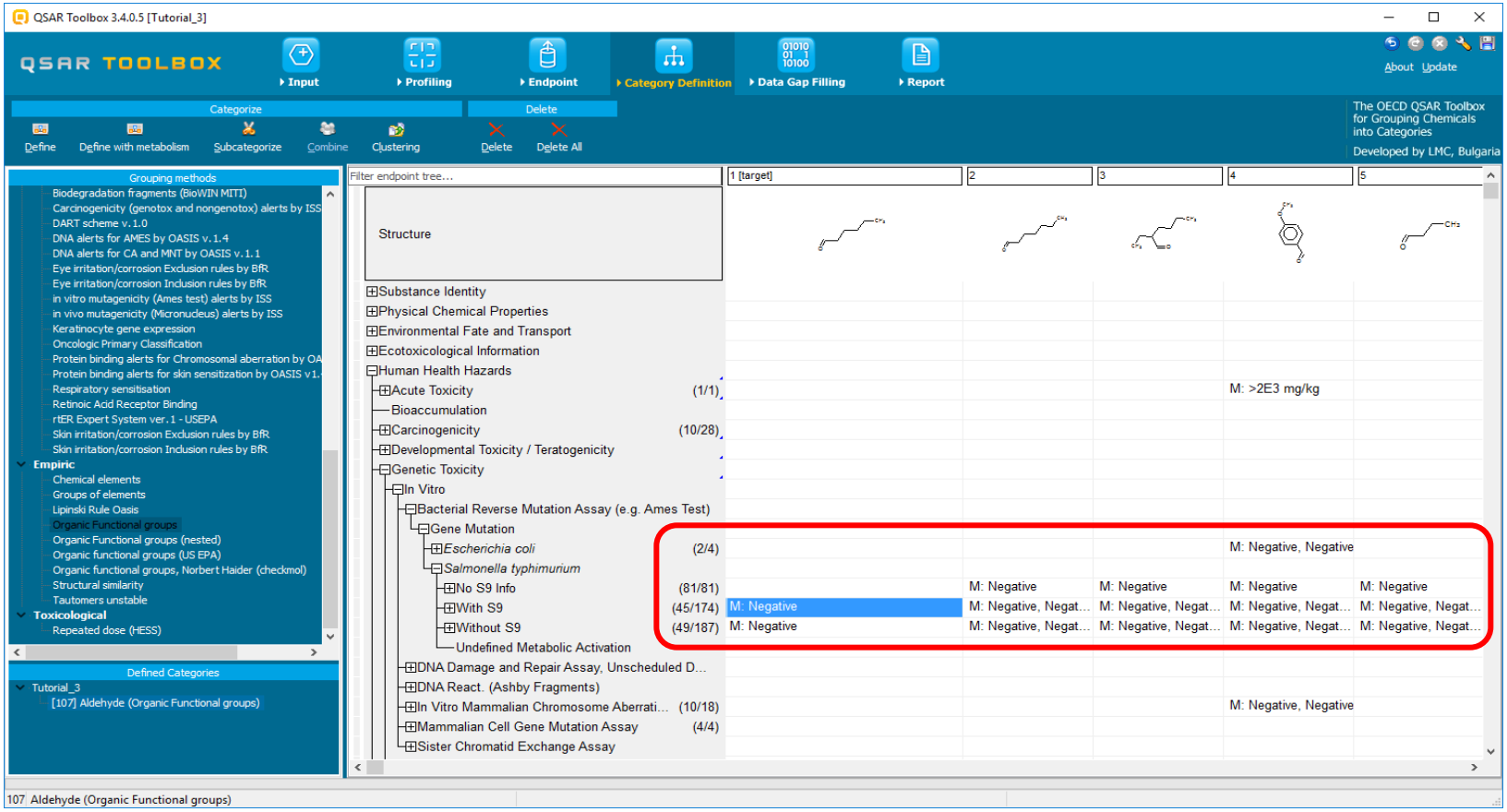
Category Definition
Read data for Analogues

Due to the overlap between the Toolbox databases same data for intersecting chemicals could be found simultaneously in more than one databases. The data redundancy is identified and the user has the opportunity to select either a single data value or all data values.



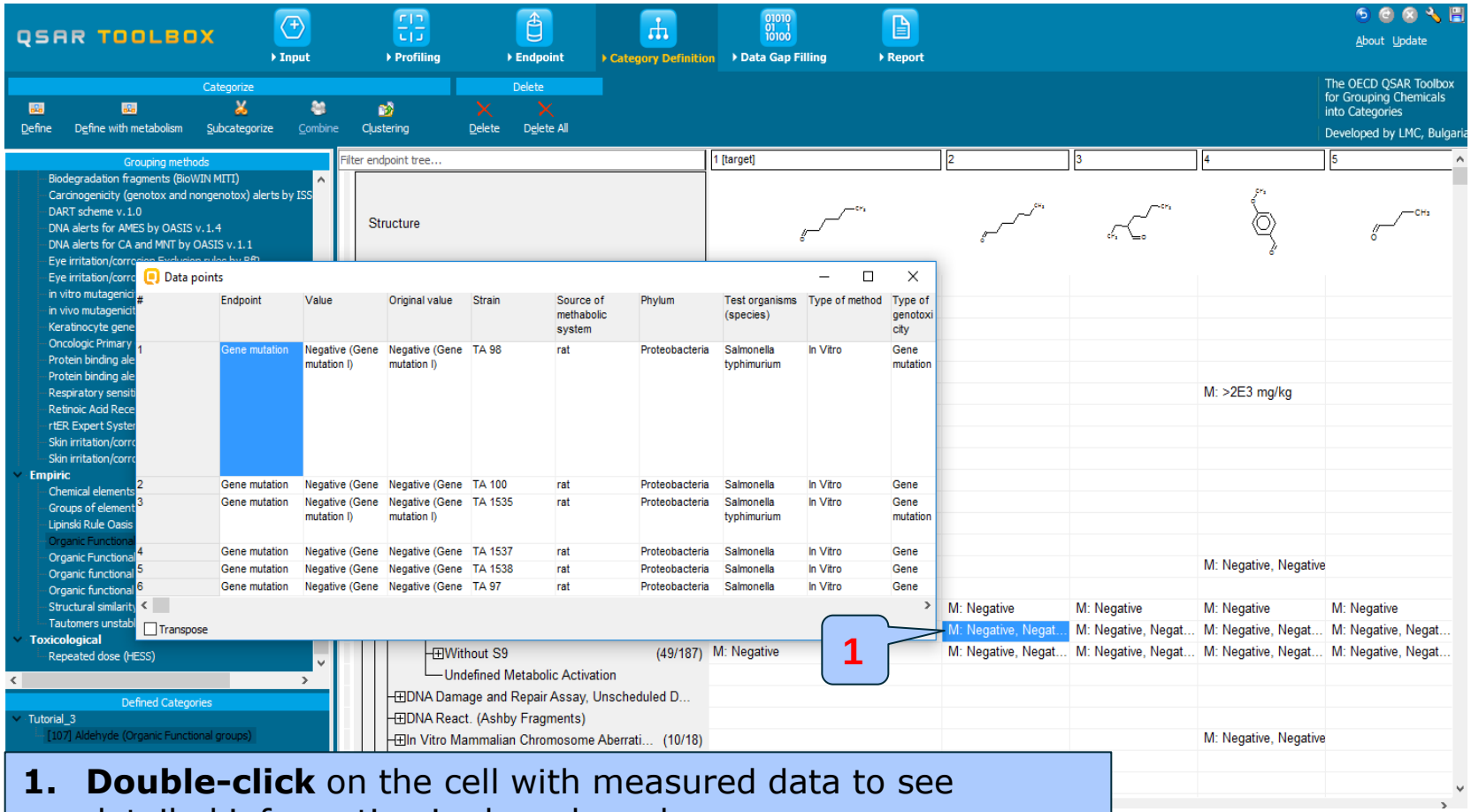
Category Definition
Summary information for Analogues

The experimental results for the analogues are inserted into the matrix.



Category Definition

Side-Bar of experimental data



#	Endpoint	Value	Original value	Strain	Source of methabolic system	Phylum	Test organisms (species)	Type of method	Type of genotoxicity
1	Gene mutation	Negative (Gene mutation I)	Negative (Gene mutation I)	TA 98	rat	Proteobacteria	Salmonella typhimurium	In Vitro	Gene mutation
2	Gene mutation	Negative (Gene	Negative (Gene	TA 100	rat	Proteobacteria	Salmonella	In Vitro	Gene
3	Gene mutation	Negative (Gene mutation I)	Negative (Gene mutation I)	TA 1535	rat	Proteobacteria	Salmonella typhimurium	In Vitro	Gene mutation
4	Gene mutation	Negative (Gene	Negative (Gene	TA 1537	rat	Proteobacteria	Salmonella	In Vitro	Gene
5	Gene mutation	Negative (Gene	Negative (Gene	TA 1538	rat	Proteobacteria	Salmonella	In Vitro	Gene
6	Gene mutation	Negative (Gene	Negative (Gene	TA 97	rat	Proteobacteria	Salmonella	In Vitro	Gene

1. **Double-click** on the cell with measured data to see detailed information in drop down box.

Category Definition
Recap

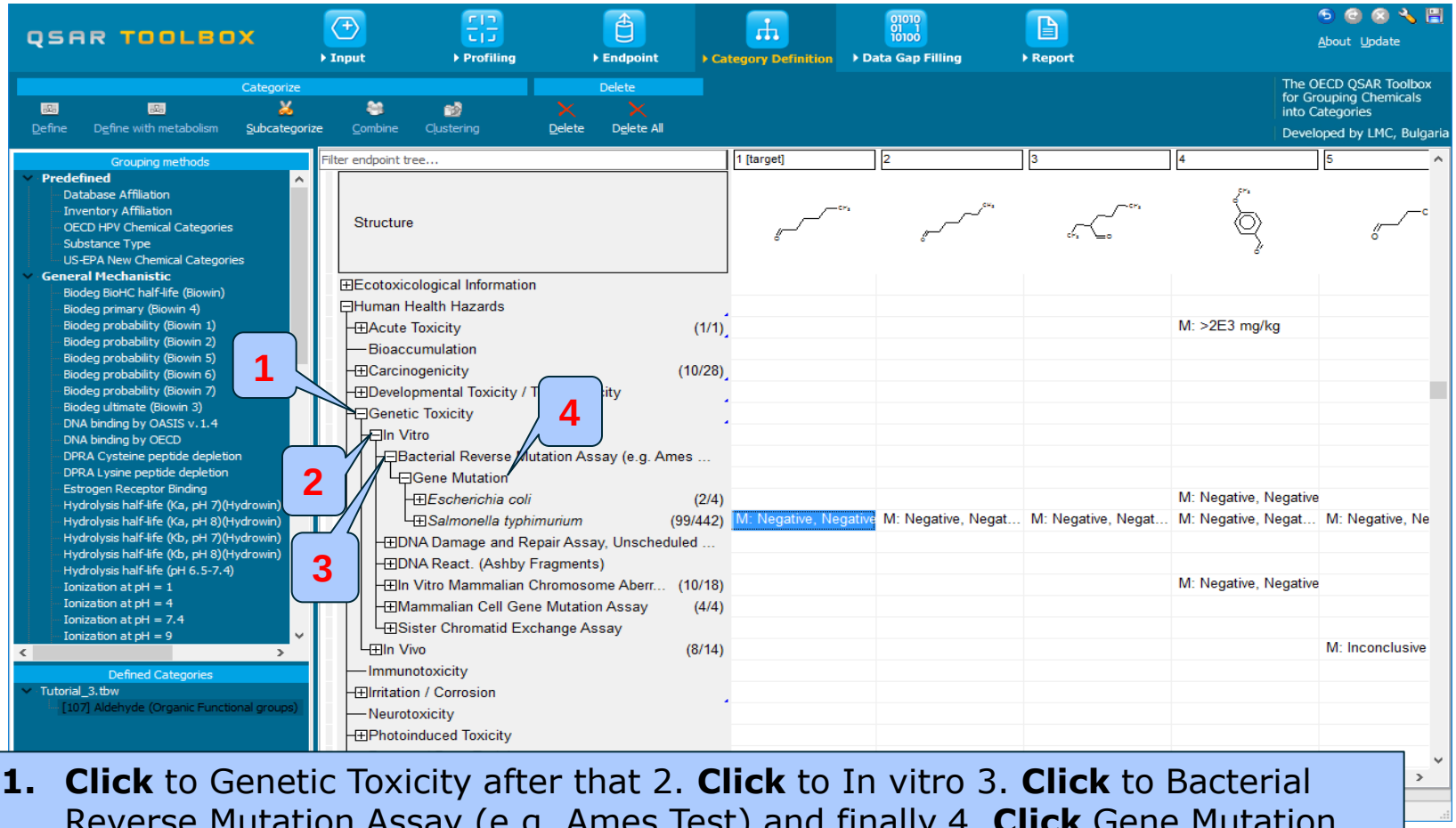
- You have identified a category consisting of 105 analogous (“Aldehydes” by OFG classification) with the target chemical (n-hexanal).
- The available experimental data for these 105 similar chemicals are collected from the previously selected databases under Endpoint section.
- The user can proceed with “Filling data gap” module, but before that he/she should navigate throw the endpoint tree and find the gap that will be filled in.

Category Definition
Navigation throw the endpoint tree

- The user can navigate through the data tree by closing or opening the nodes of the tree.
- In this example, results from genotox testing are available (see next screen shot).
- In this example to see does the target is mutagenic or not, it is recommended to check subsequently the two mutagenic endpoints:
  - Ames without S9
  - Ames with S9
- By double clicking on the nodes of endpoint tree open the tree to the target: **Bacterial reverse mutation (Ames) assay without S9** *(i.e., double click on Human Health Hazards then double click on Genetic Toxicity followed by In Vitro and Bacterial Reverse Mutation Assay (e.g. Ames Test),Gene Mutation Salmonella typhimurium, Without S9)* (see next screen shot).

Category Definition

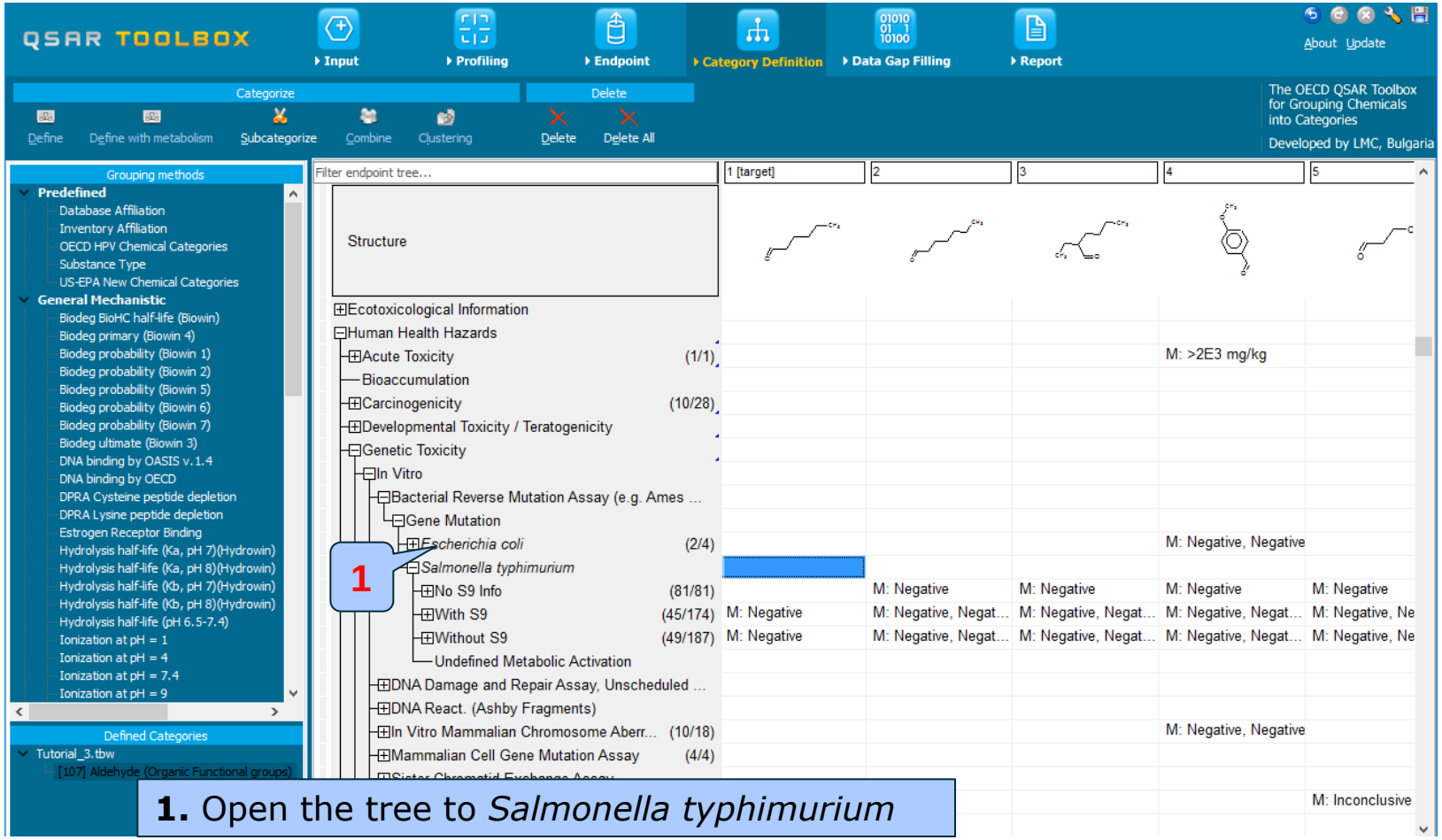
Navigation throw the endpoint tree



1. **Click** to Genetic Toxicity after that 2. **Click** to In vitro 3. **Click** to Bacterial Reverse Mutation Assay (e.g. Ames Test) and finally 4. **Click** Gene Mutation

Category Definition

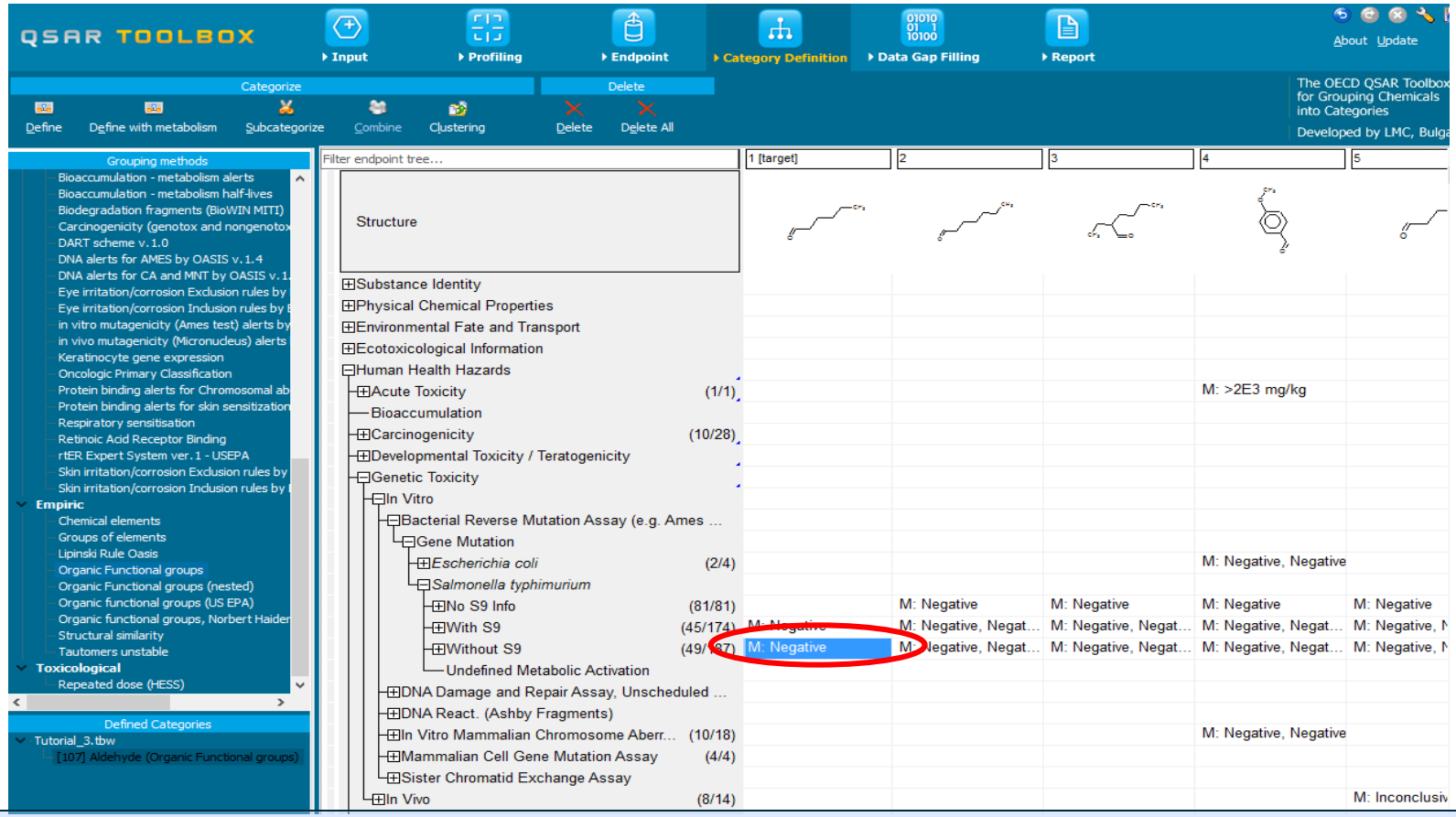
Navigation throw the endpoint tree



1. Open the tree to *Salmonella typhimurium*

Category Definition

Navigation throw the endpoint tree



In order to examine the target endpoint "Ames without S9", select the cell as shown.

Category Definition
Recap

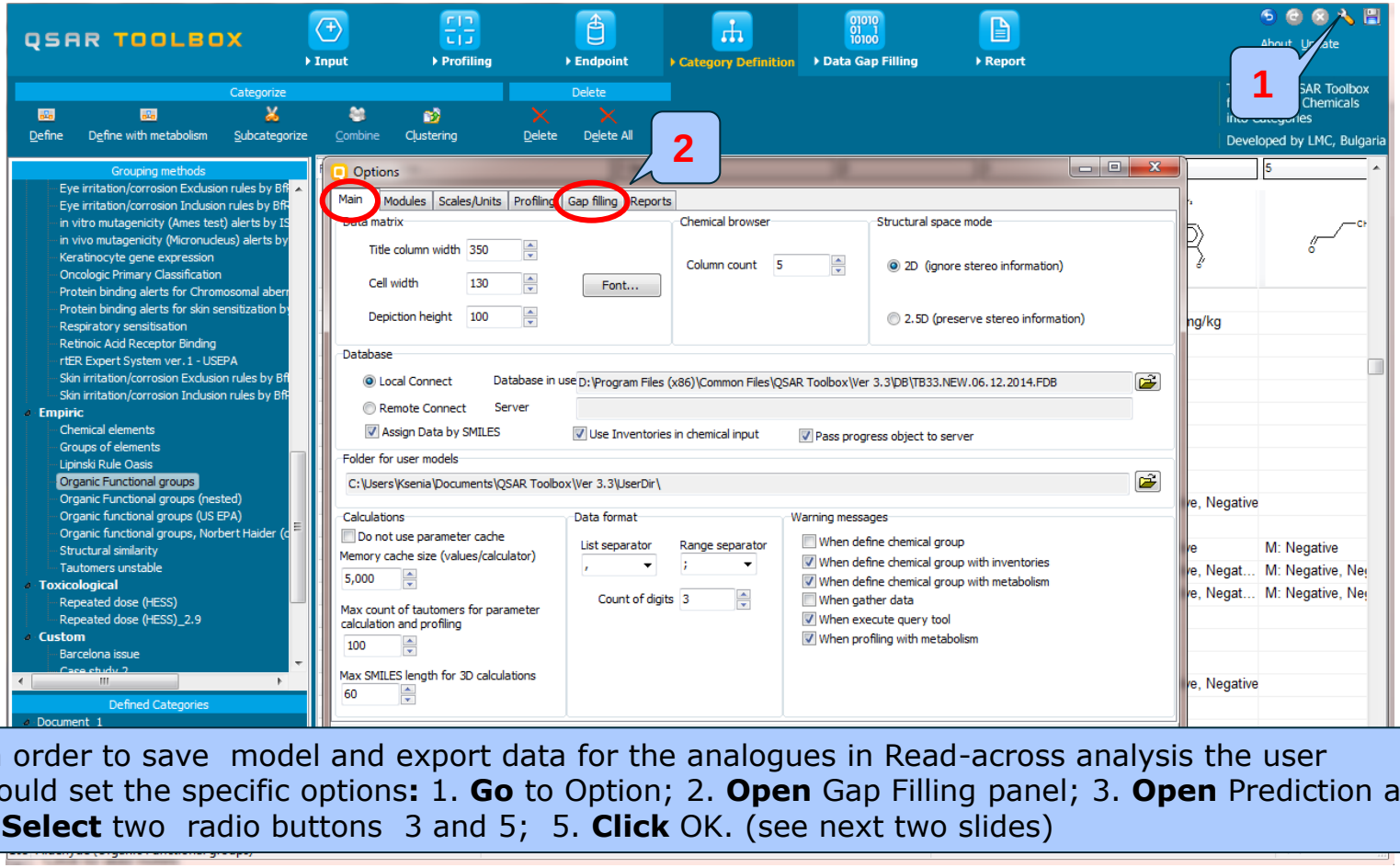
- You have now retrieved the available experimental data on genetic toxicity for 107 chemicals classified as "Aldehydes" by OFG, found in the databases containing mutagenicity data.
- Only 49 out of 107 analogues have experimental mutagenicity data related to the target.
- You are now ready to fill in the data gap and trying to reproduce the experimental data of the target.
- In this example with qualitative mutagenicity data we can only use read-across.

Outlook

- Background
- Objectives
- Specific Aims
- Read-across
- The exercise
- **Workflow of the exercise**
  - Chemical input
  - Profiling
  - Endpoint
  - Category definition
  - **Data Gap Filling**
    - **Ames without S9**

Data Gap Filling (Ames without S9)

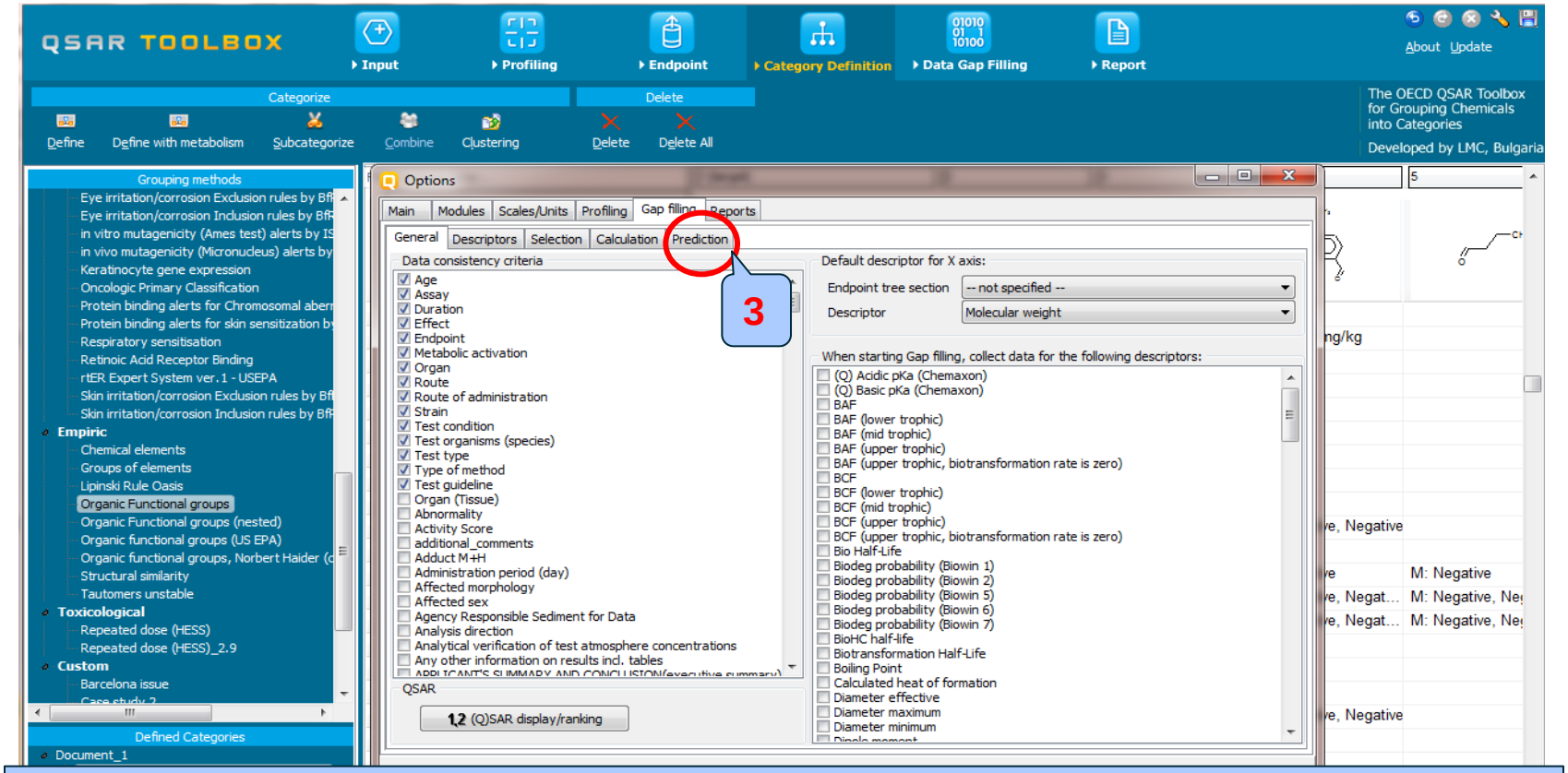
Apply read-across



In order to save model and export data for the analogues in Read-across analysis the user should set the specific options**:** 1. **Go** to Option; 2. **Open** Gap Filling panel; 3. **Open** Prediction and; 4. **Select** two radio buttons 3 and 5; 5. **Click** OK. (see next two slides)

Data Gap Filling (Ames without S9)

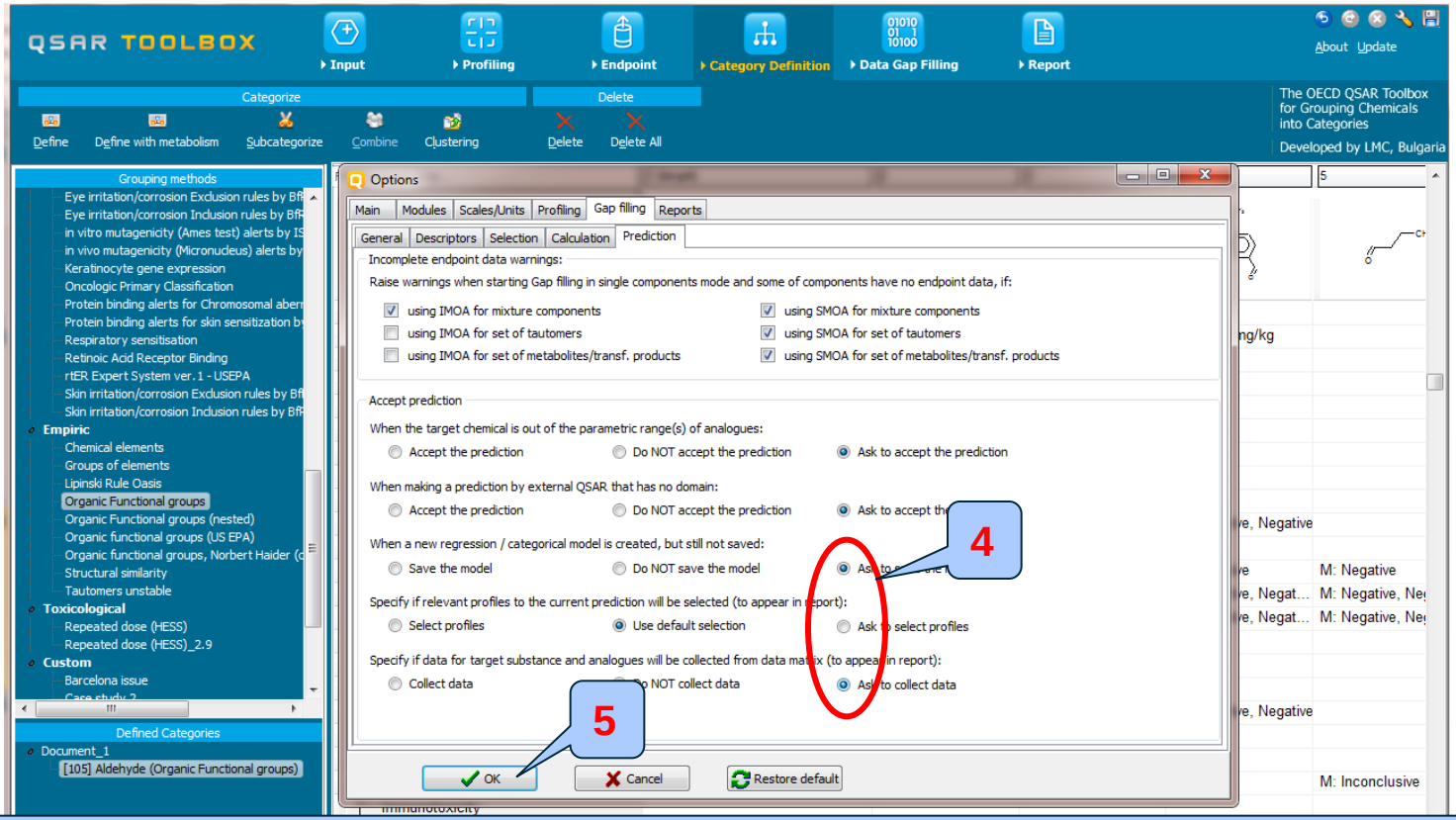
Apply read-across



In order to save model and export data for the analogues in Read-across analysis the user should set the specific options**:** 1. **Go** to Option; 2. **Open** Gap Filling panel; 3. **Open** Prediction and; 4. **Select** two radio buttons 3 and 5; 5. **Click** OK. (see next two slides)

Data Gap Filling (Ames without S9)

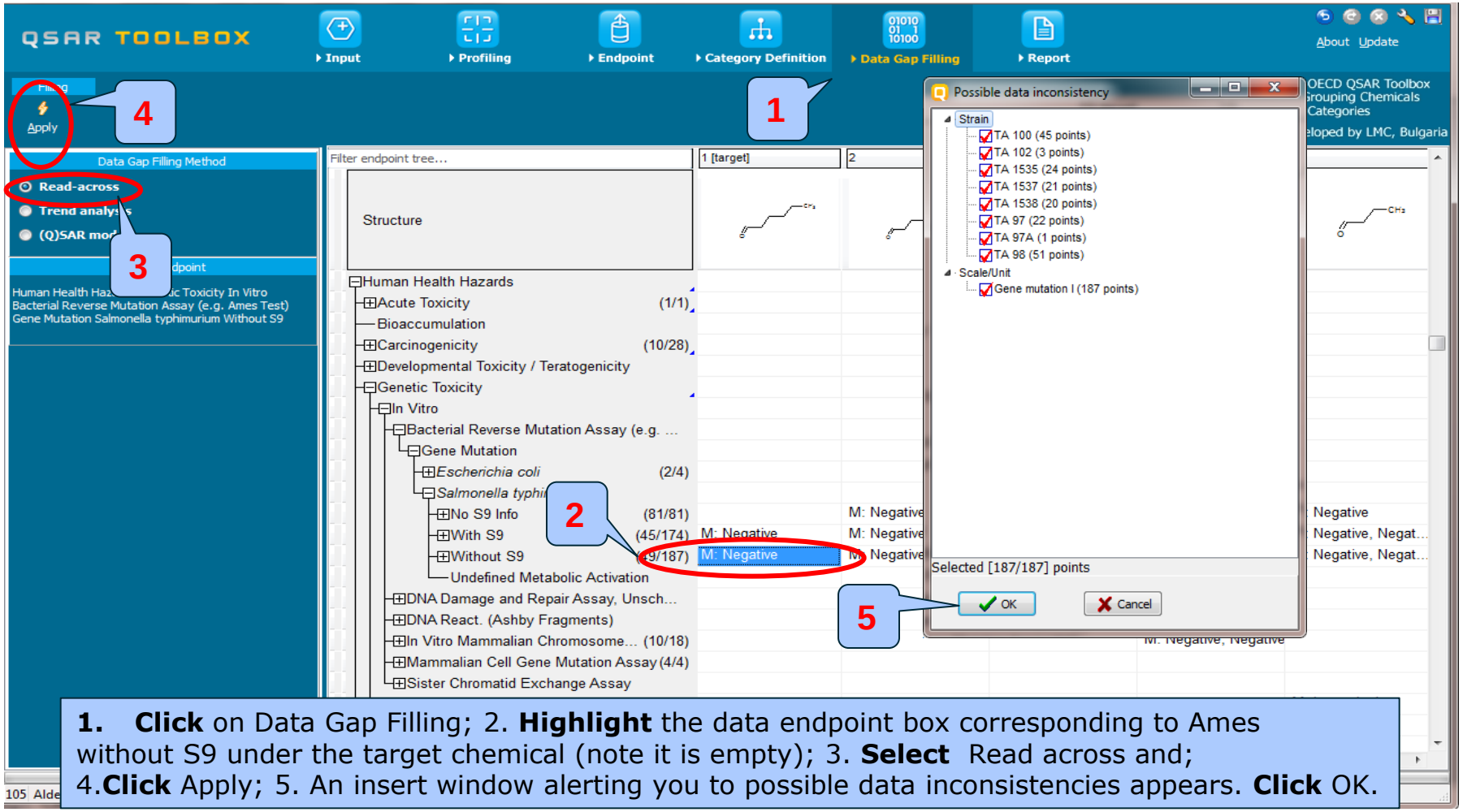
Apply read-across



In order to save model and export data for the analogues in Read-across analysis the user should set the specific options**:** 1. **Go** to Option; 2. **Open** Gap Filling panel; 3. **Open** Prediction and; 4. **Select** two radio buttons 3 and 5; 5. **Click** OK

Data Gap Filling (Ames without S9)

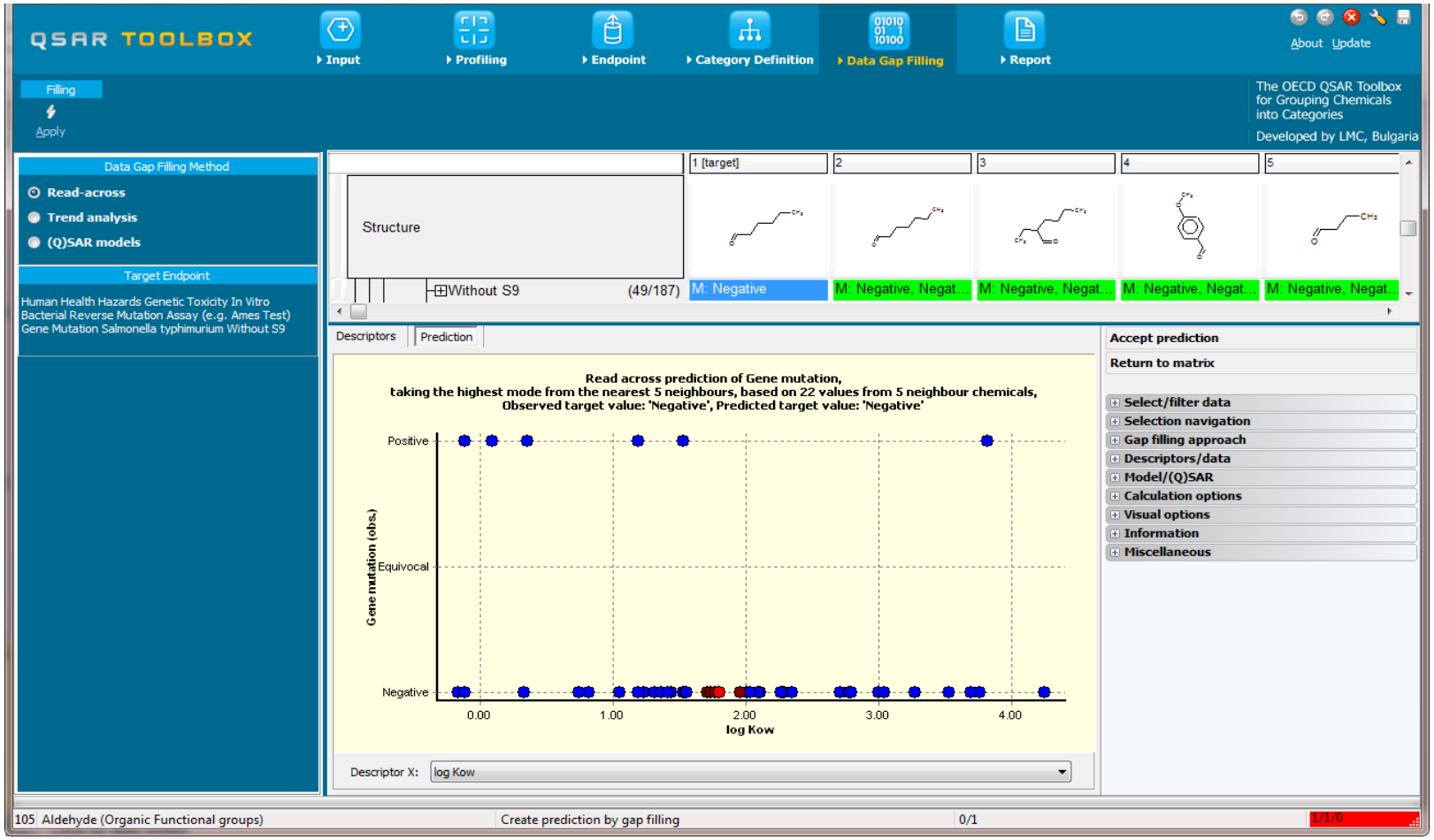
Apply read-across



1. **Click** on Data Gap Filling; 2. **Highlight** the data endpoint box corresponding to Ames without S9 under the target chemical (note it is empty); 3. **Select** Read across and; 4.**Click** Apply; 5. An insert window alerting you to possible data inconsistencies appears. **Click** OK.

Data Gap Filling (Ames without S9)

Results of Read across



Data Gap Filling (Ames without S9)
Interpreting Read-across

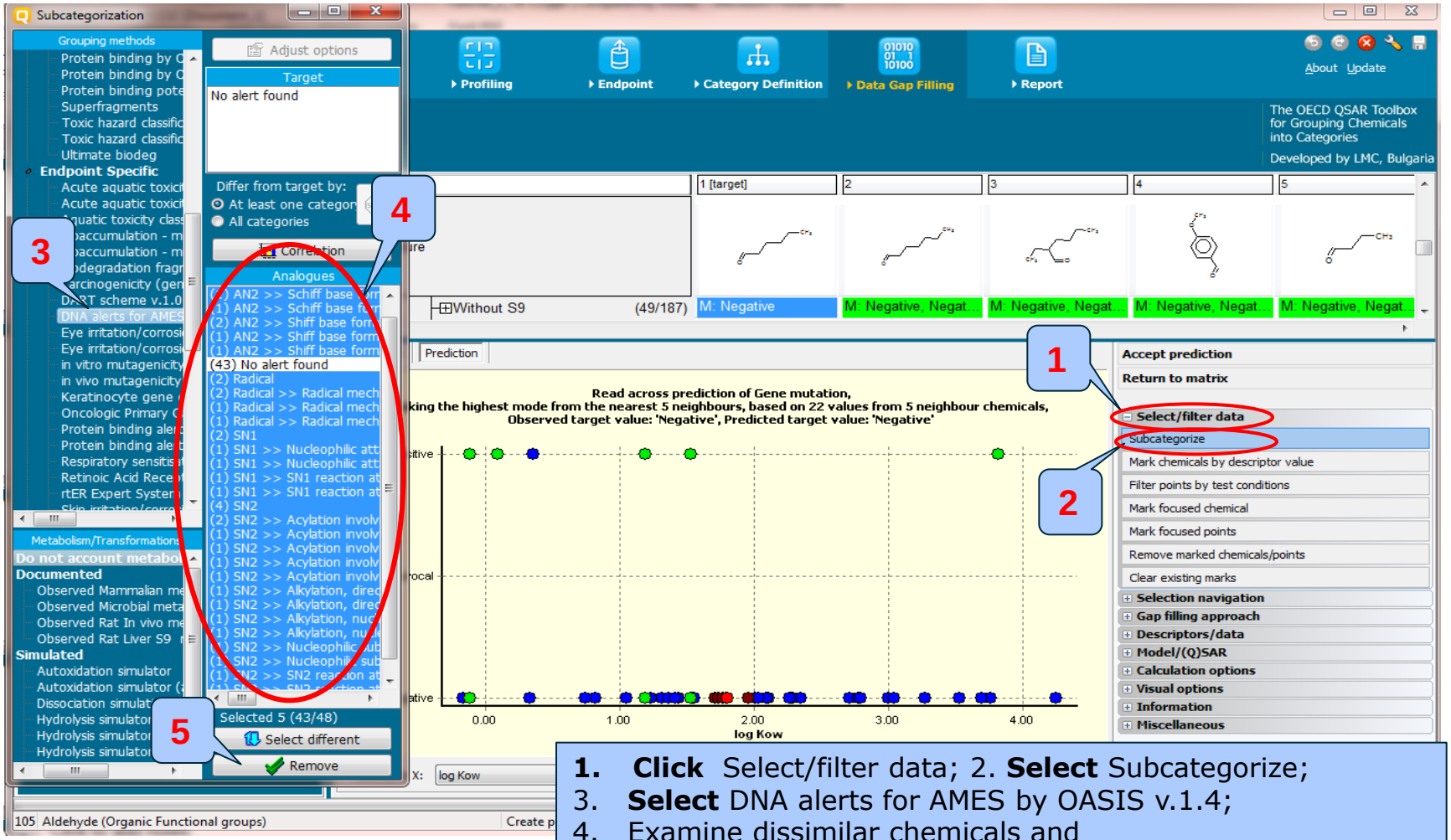
- The resulting plot outlines the experimental Ames results of all analogues (Y axis) according to a descriptor (X axis). Note, Log Kow is on the X-axis; while this descriptor is not significant to Ames data, it is the default descriptor for data gap filing (see next screen shot).
- The **RED** dot represents the predicted value for target chemical (see next screen shot).
- The **PURPLE** dots represent the observed value for the target neighbours(analogues) used for read-across (see next screen shot).
- The **BLUE** dots represent the experimental results available for the analogues but not used for read-across. (see next screen shot).
- Please note **GREEN** dots (which you will see shortly) represent analogues belonging to different subcategories.

Data Gap Filling (Ames without S9)
Interpretation of the Read across

- Six of the analogues are mutagenic in the Ames assays without S9, the rest analogues are non-mutagenic
- Non-mutagenic potential (Negative) is, therefore, predicted with confidence for the target chemical.
- However, before data gap filling it is recommended to check the similarity of the analogues used in the prediction (see next screen shot). This is performed in order to assure the category consists of analogues that are both mechanistically and structurally similar.

Data Gap Filling (Ames without S9)

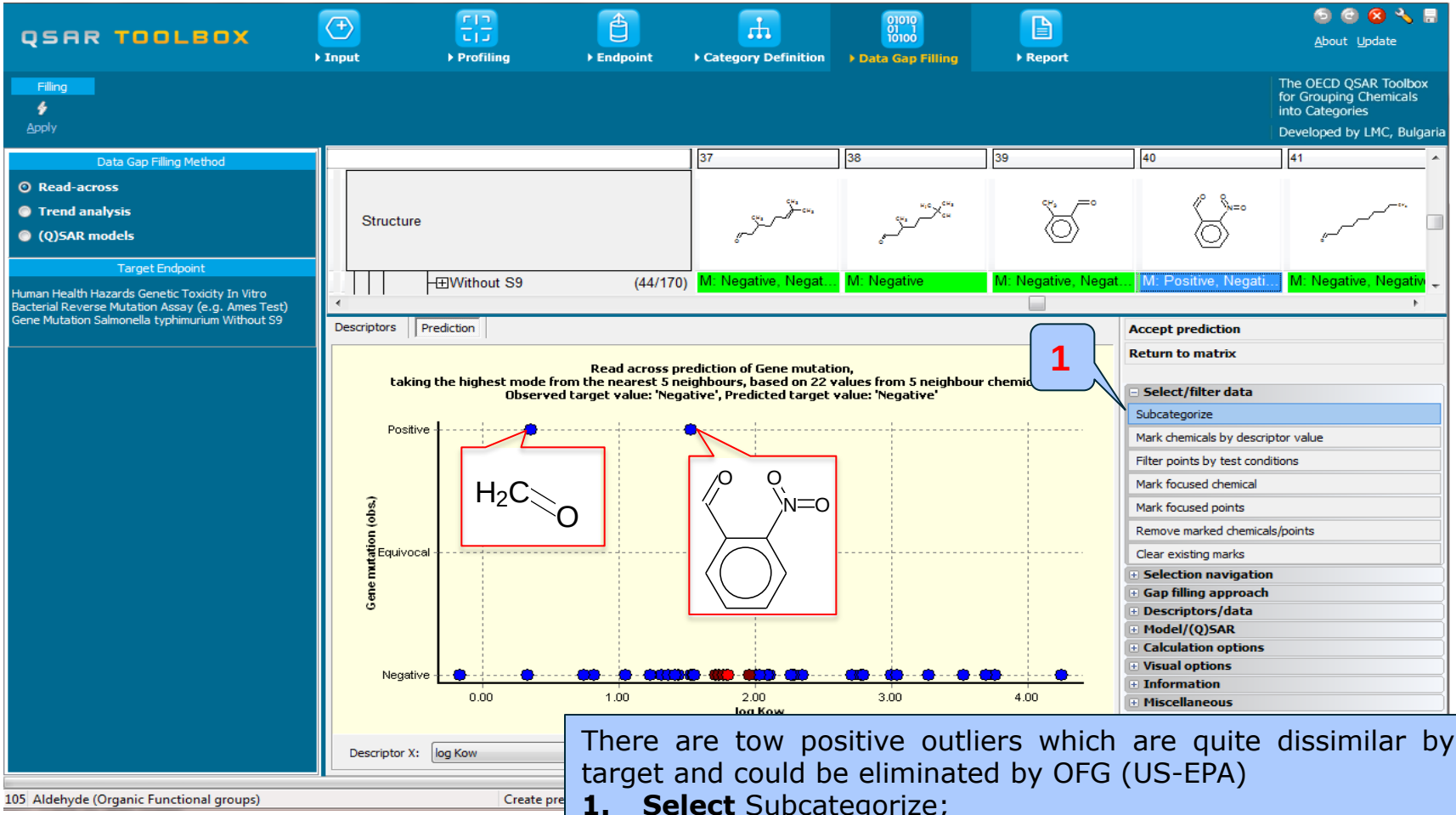
Subcategorization by DNA binding by OASIS (endpoint specific)



1. **Click** Select/filter data; 2. **Select** Subcategorize;
3. **Select** DNA alerts for AMES by OASIS v.1.4;
4. Examine dissimilar chemicals and
5. **Click** Remove to eliminate dissimilar chemical.

Data Gap Filling (Ames without S9)

Subcategorization by OFG (US-EPA)

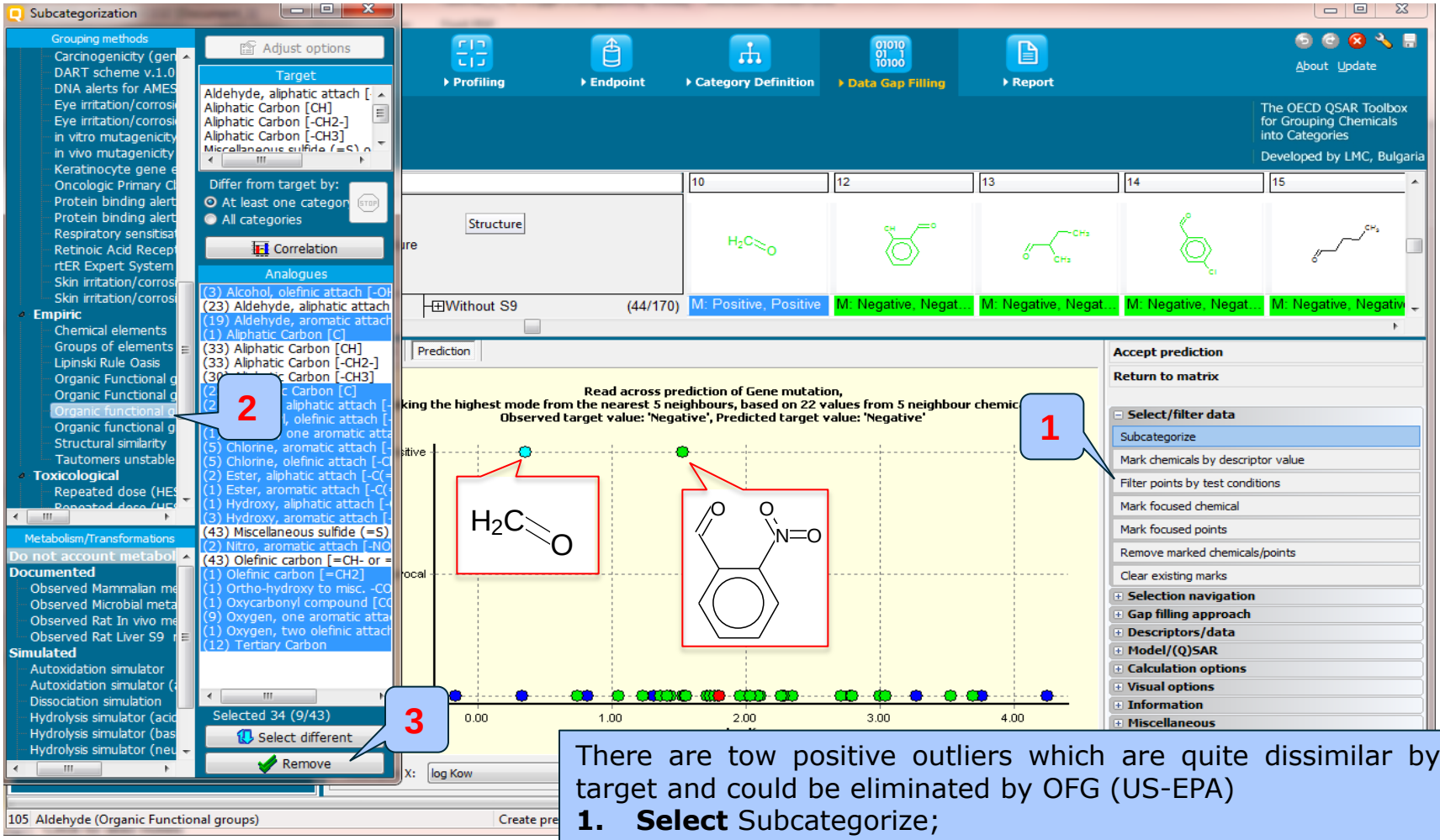


There are tow positive outliers which are quite dissimilar by the target and could be eliminated by OFG (US-EPA)

1. **Select** Subcategorize;
2. **Select** OFG (US-EPA);
3. **Click** Remove to eliminate dissimilar chemical. (see next slides)

Data Gap Filling (Ames without S9)

Subcategorization by OFG (US-EPA)

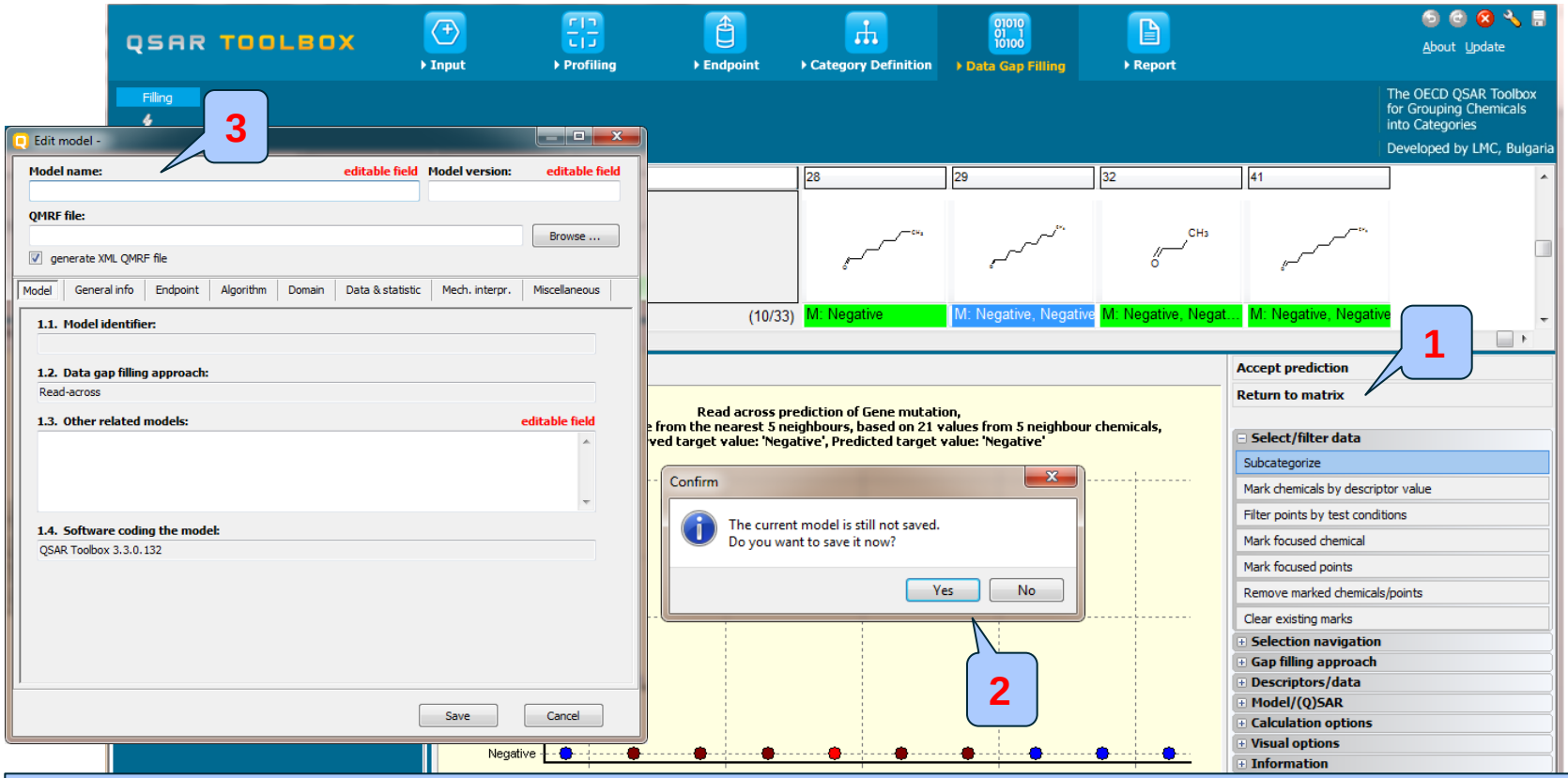


There are tow positive outliers which are quite dissimilar by the target and could be eliminated by OFG (US-EPA)

1. **Select** Subcategorize;
2. **Select** OFG (US-EPA);
3. **Click** Remove to eliminate dissimilar chemical. (see next slides)

Data Gap Filling (Ames without S9)

Interpretation of the Read across

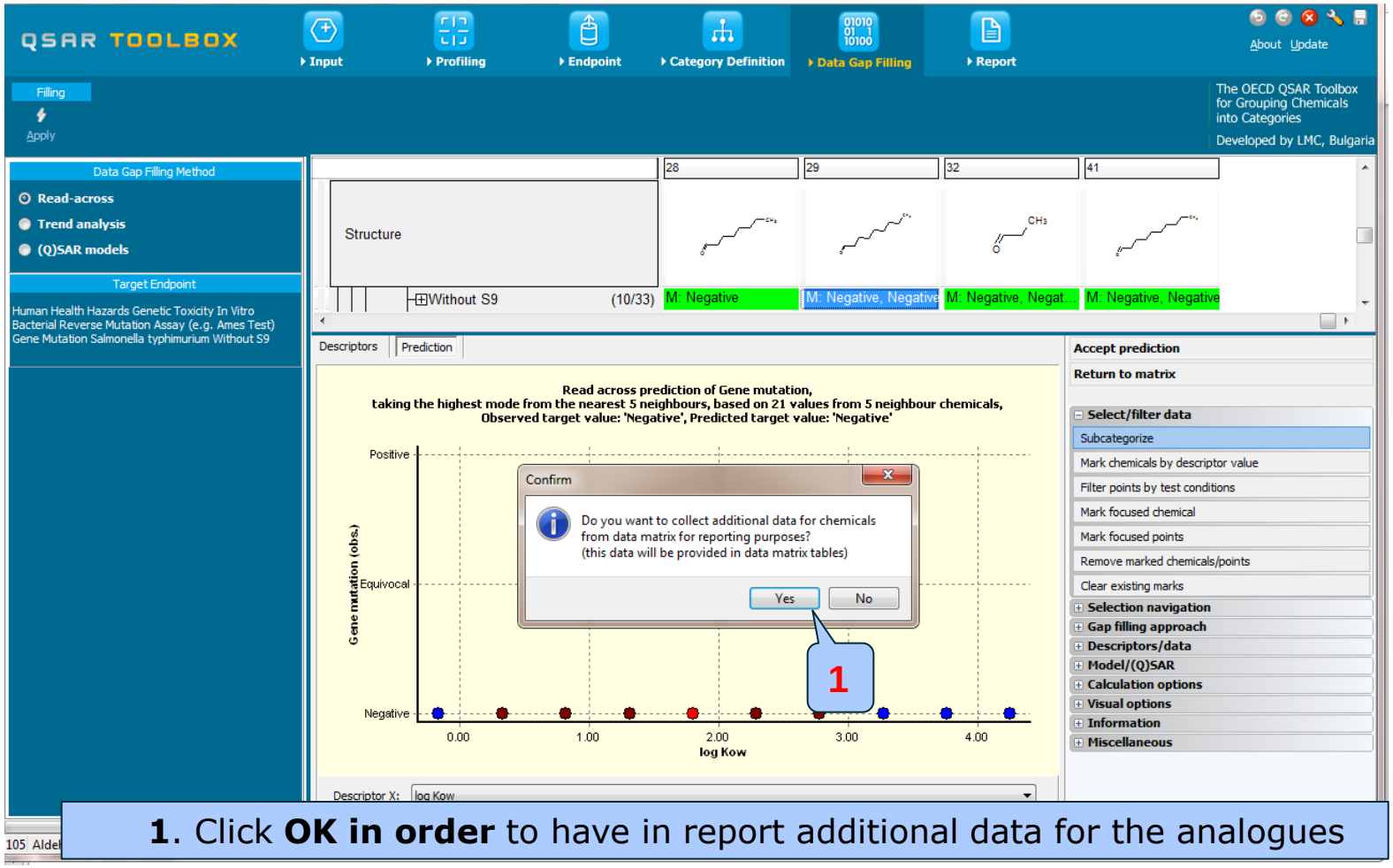


Now all analogues are structurally similar (Aldehydes) and negative by the experimental data. The prediction could be accepted by **1. Click** on Accept prediction and If you want to save the model, and use it for further predictions, then

2. Click Yes and then **3**. **Edit** the information about the model.

Data Gap Filling (Ames without S9)

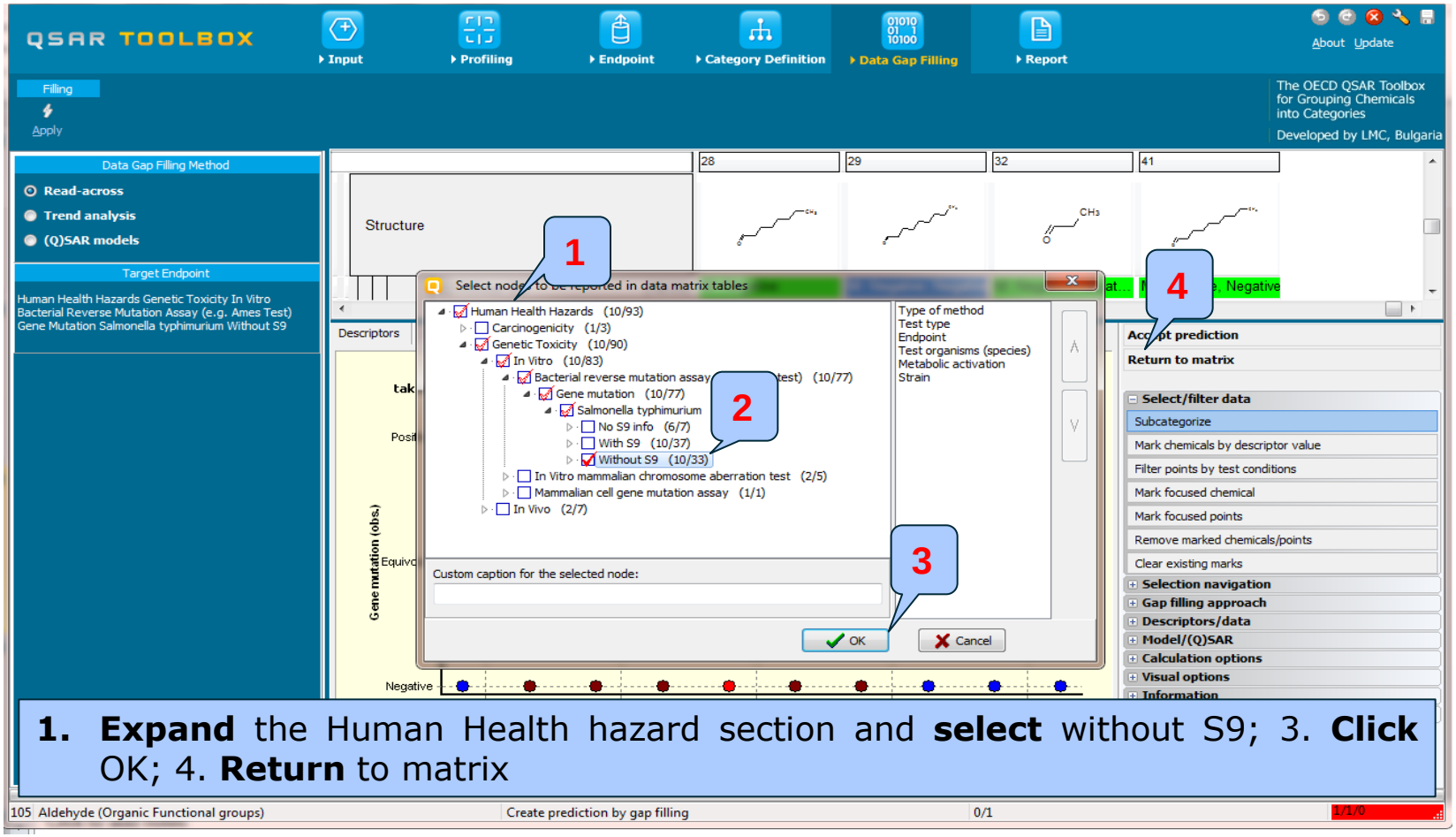
Interpretation of the Read across



1. Click **OK in order** to have in report additional data for the analogues

Data Gap Filling (Ames without S9)

Interpretation of the Read across

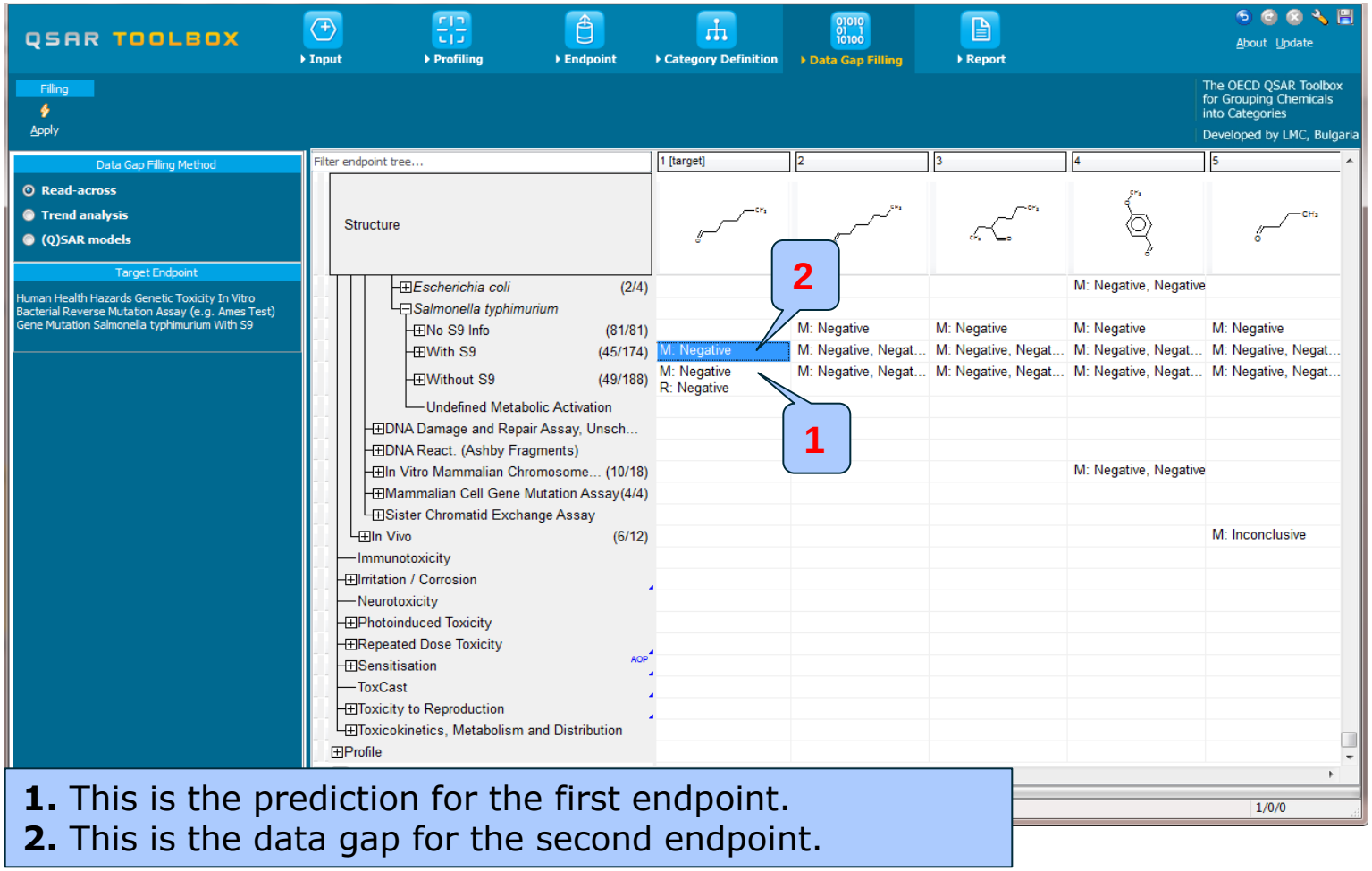


Data Gap Filling (Ames without S9)
Results

- By accepting the prediction the data gap is filled.

- By clicking on “Return to Matrix”, the user can close the read-across for the current endpoint and proceed with the workflow for the second endpoint, which in this case will be “Ames with S9” (see next screen shot).

Data Gap Filling (Ames without S9)
Results



1. This is the prediction for the first endpoint.
2. This is the data gap for the second endpoint.

Outlook

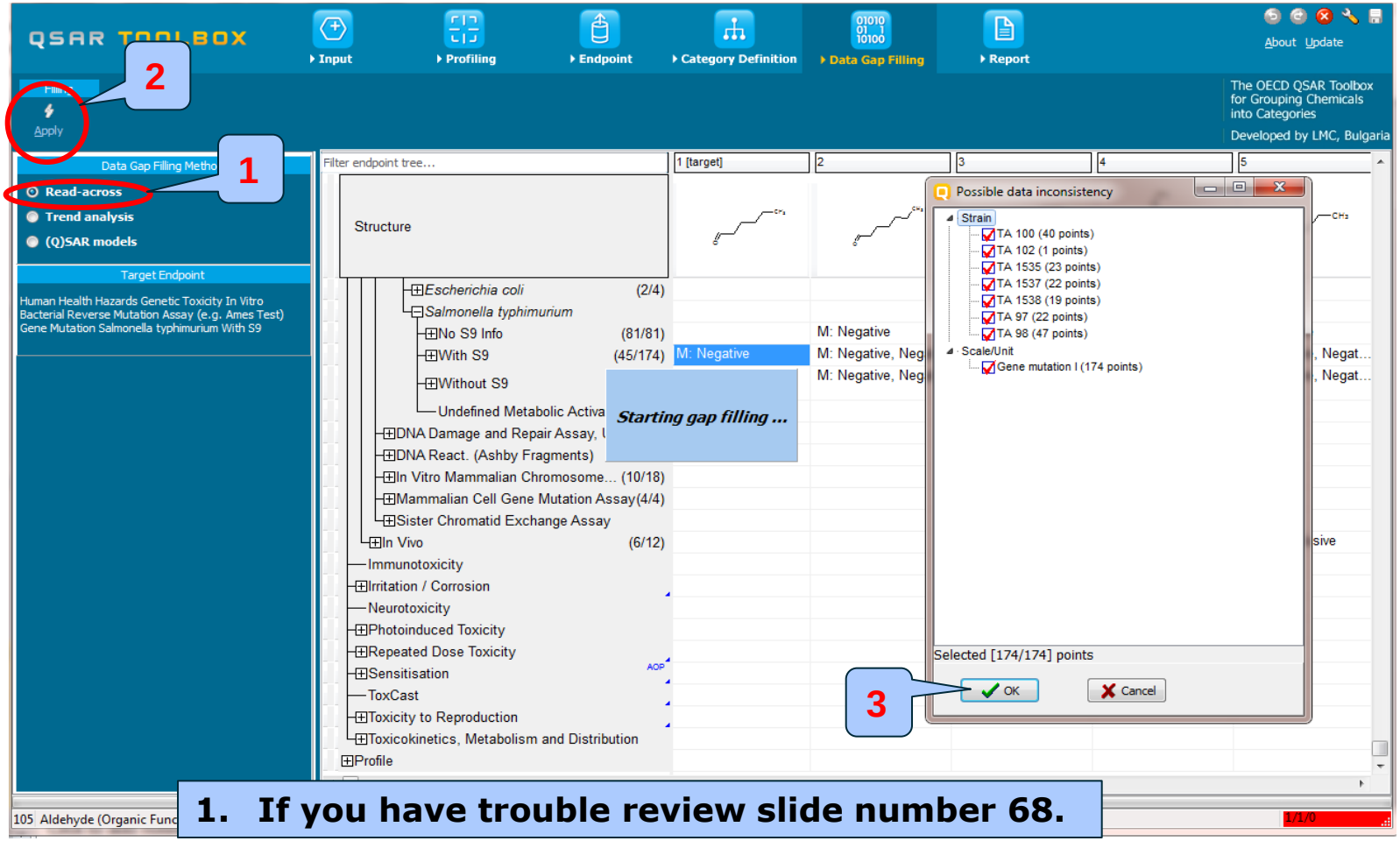
- Background
- Objectives
- Specific Aims
- Read-across
- The exercise
- **Workflow of the exercise**
  - Chemical input
  - Profiling
  - Endpoint
  - Category definition
  - **Data Gap Filling**
    - Ames without S9
    - **Ames with S9**

Data Gap Filling (Ames with S9)

- We do this the same way as with Ames without S9.
- Make sure **Data Gap Filling** is highlighted.
- Highlight the **data endpoint box**; this time corresponding to **Ames with S9**. Again the box under the structure is empty.
- Select **Read across** and Click **Apply.**
- As before an insert window alerting you to **possible data inconsistencies** appears. Click **OK** (see next screen shot).

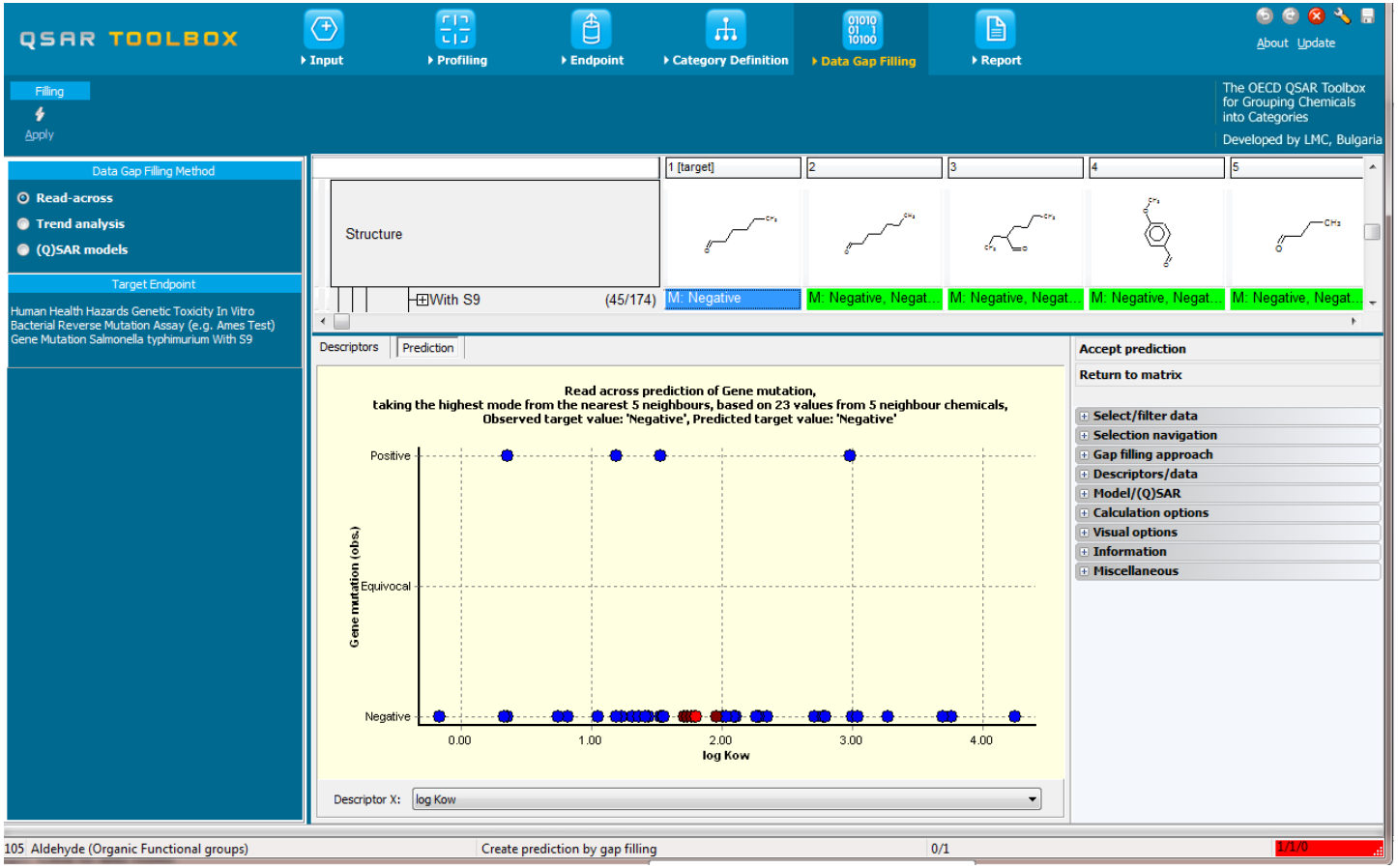
Data Gap Filling (Ames with S9)

Apply read-across



1. If you have trouble review slide number 68.

Data Gap Filling (Ames with S9)
Results of Read across



Data Gap Filling (Ames with S9)
Results of Read across

- As with Ames without S9, before accepting the estimated result for the target chemical, by read-across the user should refined the category by subcategorisation.
- Subcategorisation refers to the process of applying additional profilers to the previously defined category, identifying chemicals which have differing profiling results and eventually eliminating these chemicals from the category.
- In this example, we are going to use several different profilers to repeatedly subcategorise the data set.

Data Gap Filling (Ames with S9)
Side Bar of Subcategorization

The analogues which are dissimilar to the target chemical with respect to:

- DNA binding alerts (endpoint specific) taking into account liver metabolism – The categorization based on this profiler identifies analogues having same DNA binding alerts as the target after metabolic activation

- Organic functional groups (US-EPA) – The categorization based on this profiler identifies analogues having the same organic functional groups.

can be removed from the initial list of analogues previously defined by OFG.

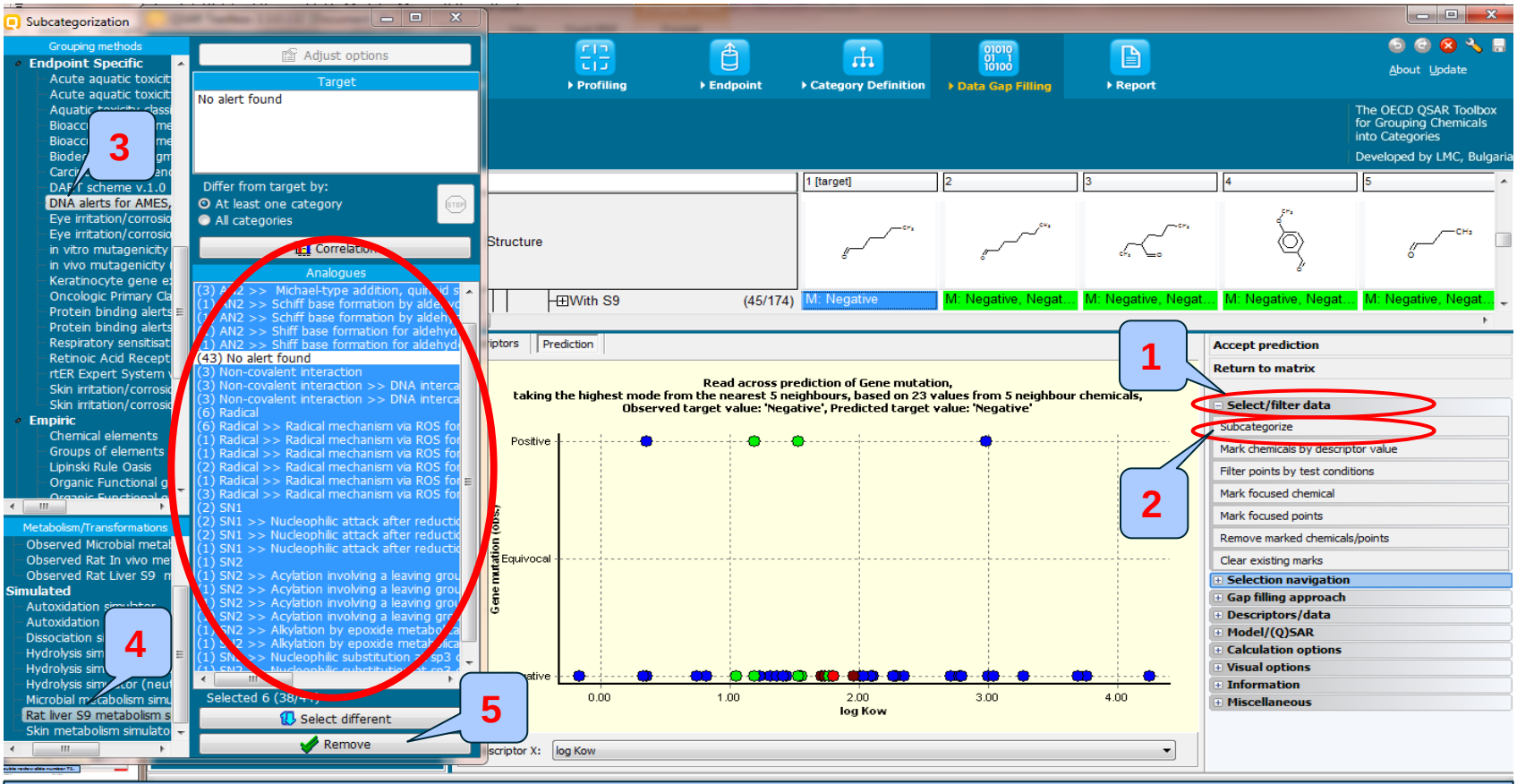
Data Gap Filling (Ames with S9)

Subcategorization by DNA alerts taking into account liver metabolism

- As with Ames without S9, we want to refined the category by subcategorisation with DNA binding by OASIS, taking into account liver metabolism
- Select **Select/filter data**
- Select **Subcategorize**
- Select **DNA binding alert**
- Select **Rat Liver S9 metabolism simulator**
- Look for dissimilar chemicals
- Click **Remove** to eliminate dissimilar chemical.

Data Gap Filling (Ames with S9)

Subcategorization by DNA binding alerts taking into account Rat liver metabolism

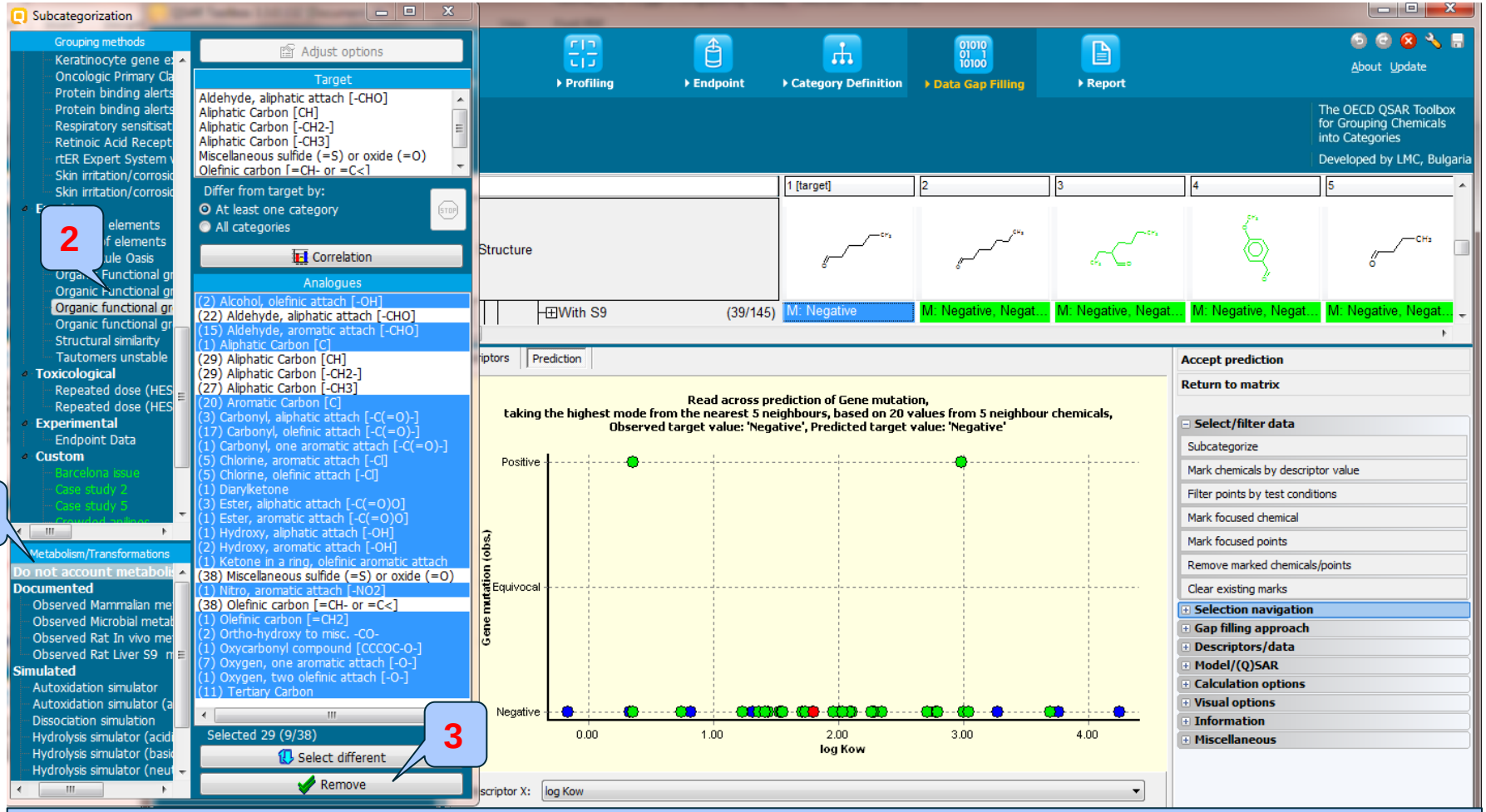


1. **Select** Select/Filter data **2. Click** Subcategorize **3. Select** DNA alerts for AMES by OASIS v.1.4 (endpoint specific) **4. Select** Rat liver metabolism simulator. **5. Click** Remove

Data Gap Filling (Ames with S9)
Subcategorization by OFG (US-EPA)

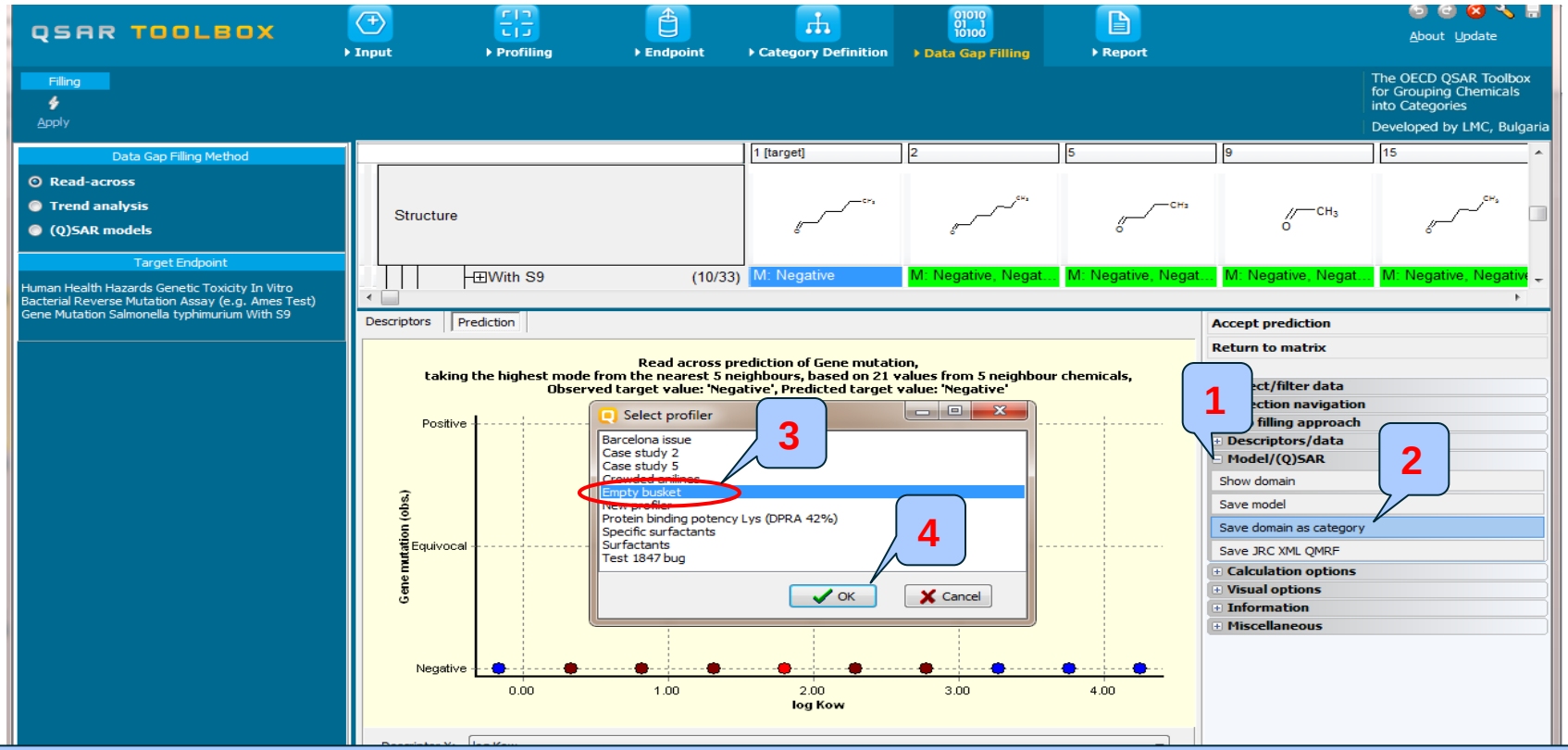
- As with Ames without S9, we want to refined the category by subcategorisation with OFG (US-EPA)
- Select **Select/filter data**
- Select **Subcategorize**
- Select **Organic functional groups (US-EPA)**
- Look for dissimilar chemicals
- Click **Remove** to eliminate dissimilar chemical.

Data Gap Filling (Ames with S9)
Subcategorization by OFG (US-EPA)



1. Click on Do not account metabolism; 2. Select OFG (US-EPA); **3. Click** Remove

Data Gap Filling (Ames with S9)
Result of read-across

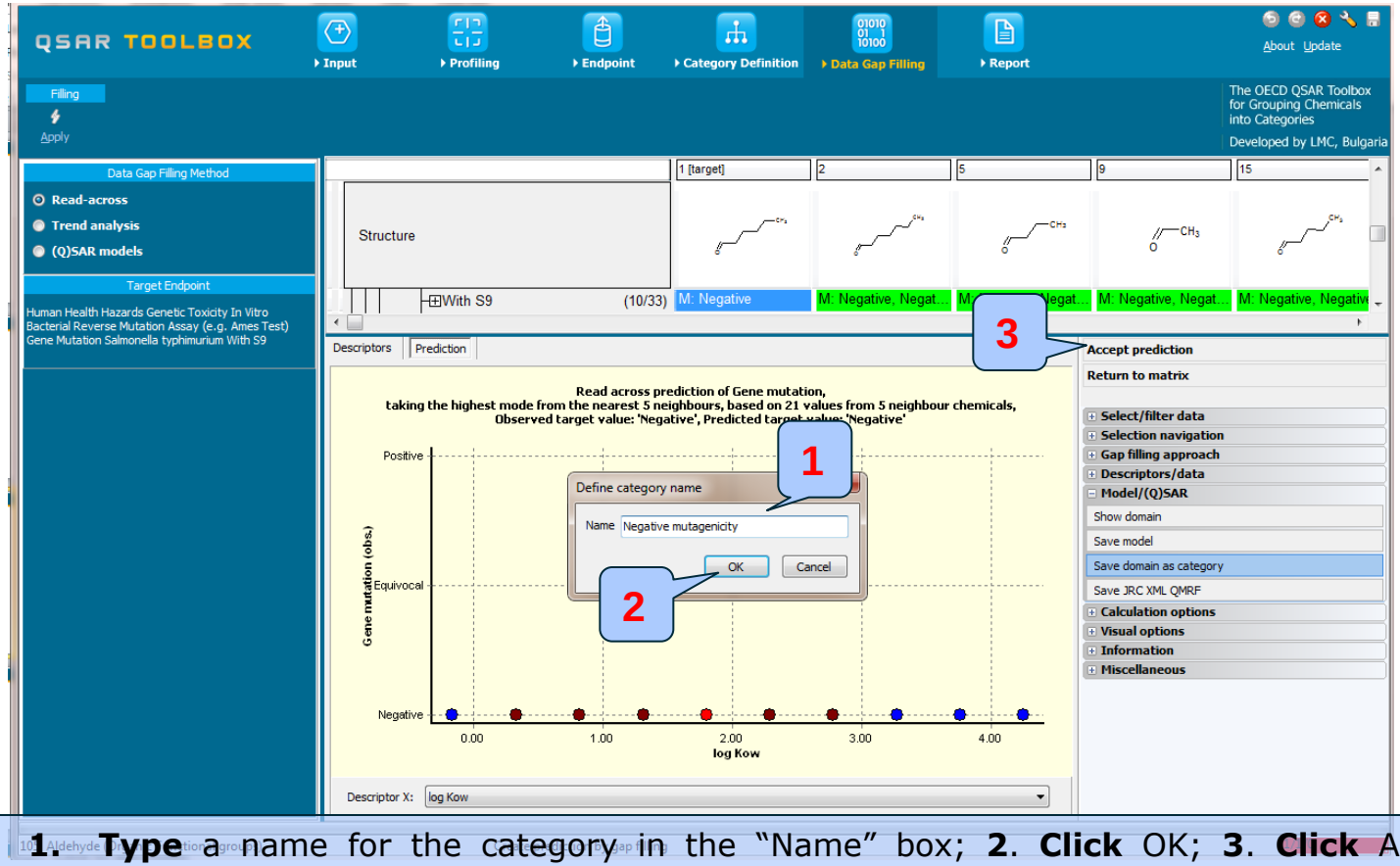


Now all 5 analogues are structurally and mechanistically similar, then the prediction could be accepted or saved as a category (domain) in the custom profiler, which could be used further for screening purposes. This could be done by

1. **Click** on Model/(Q)SAR and then; **2**. **Click** on Save domain as category

3. Since a custom profiler has previously been defined, **highlight** custom profiler and **4. Click** OK.

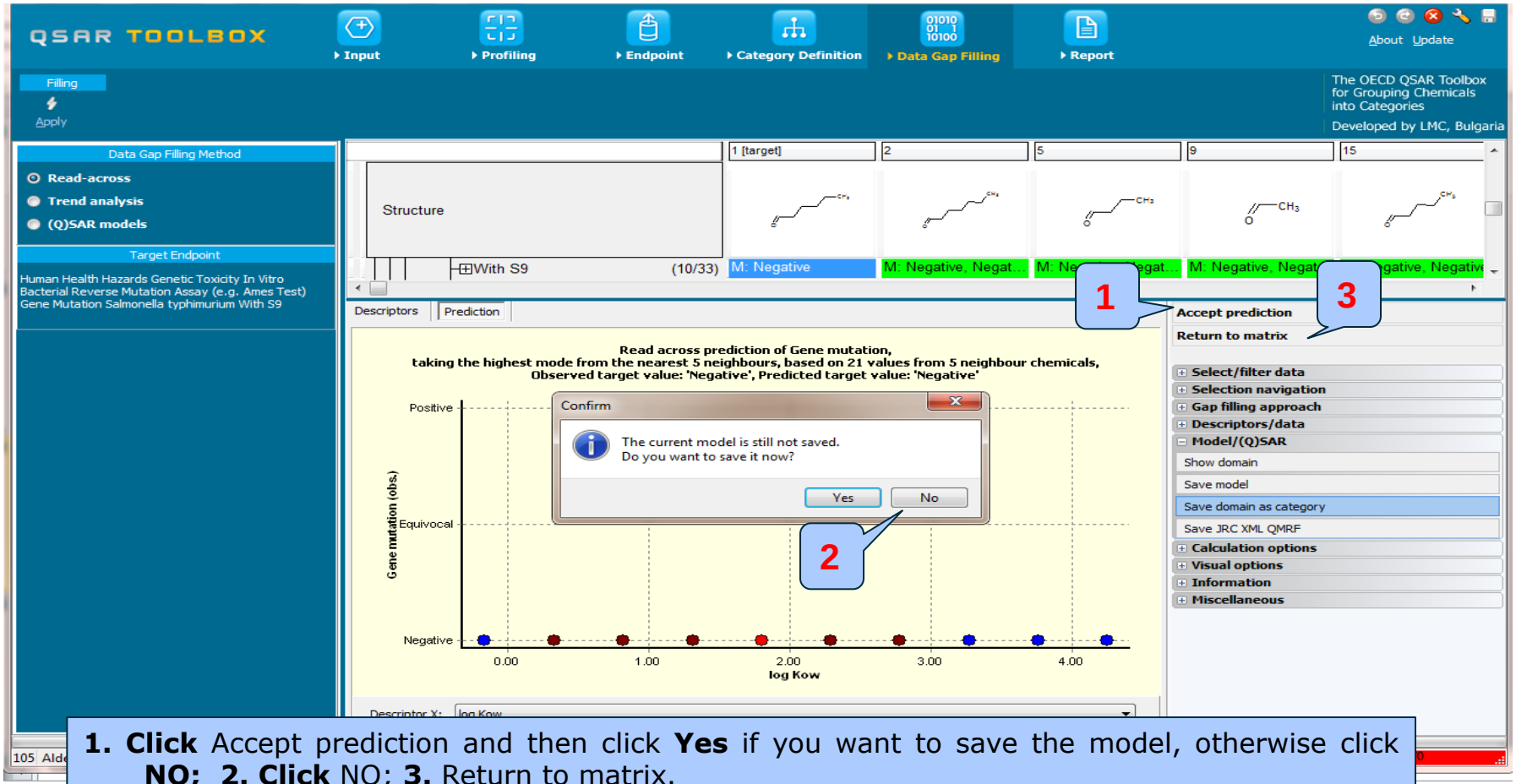
Data Gap Filling (Ames with S9)
Result of read-across



1. Type a name for the category in the "Name" box; **2. Click** OK; **3. Click** Accept prediction and Return to matrix.

Data Gap Filling (Ames with S9)

Result of read-across



1. Click Accept prediction and then click **Yes** if you want to save the model, otherwise click **NO; 2. Click** NO; **3.** Return to matrix.

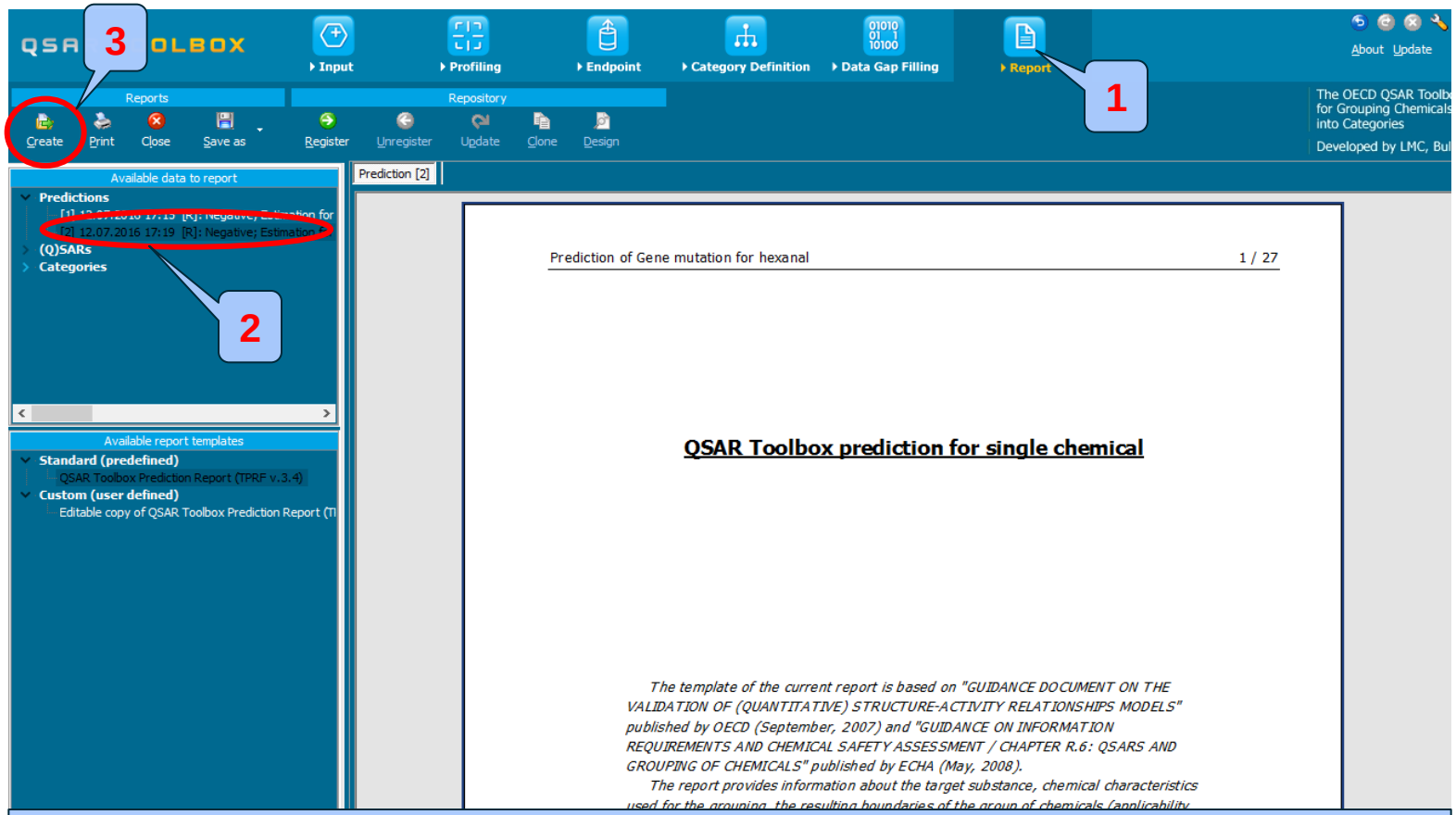
Outlook

- Background
- Objectives
- Specific Aims
- Read-across
- The exercise
- **Workflow of the exercise**
  - Chemical input
  - Profiling
  - Endpoint
  - Category definition
  - Data Gap Filling
  - **Report**

Report
Overview

- Report module could generate report on any of predictions performed with the Toolbox.
- Report module contains predefined report templates as well as a template editor with which users can define their own user defined templates.
- The report can then be printed or saved in different formats. (see next screen shot).

Report
Generate Report



1. **Select** "Report"; **2**. **Select** the current prediction from "Available data to report" window, and then **3**. **Click** Create .

Outlook

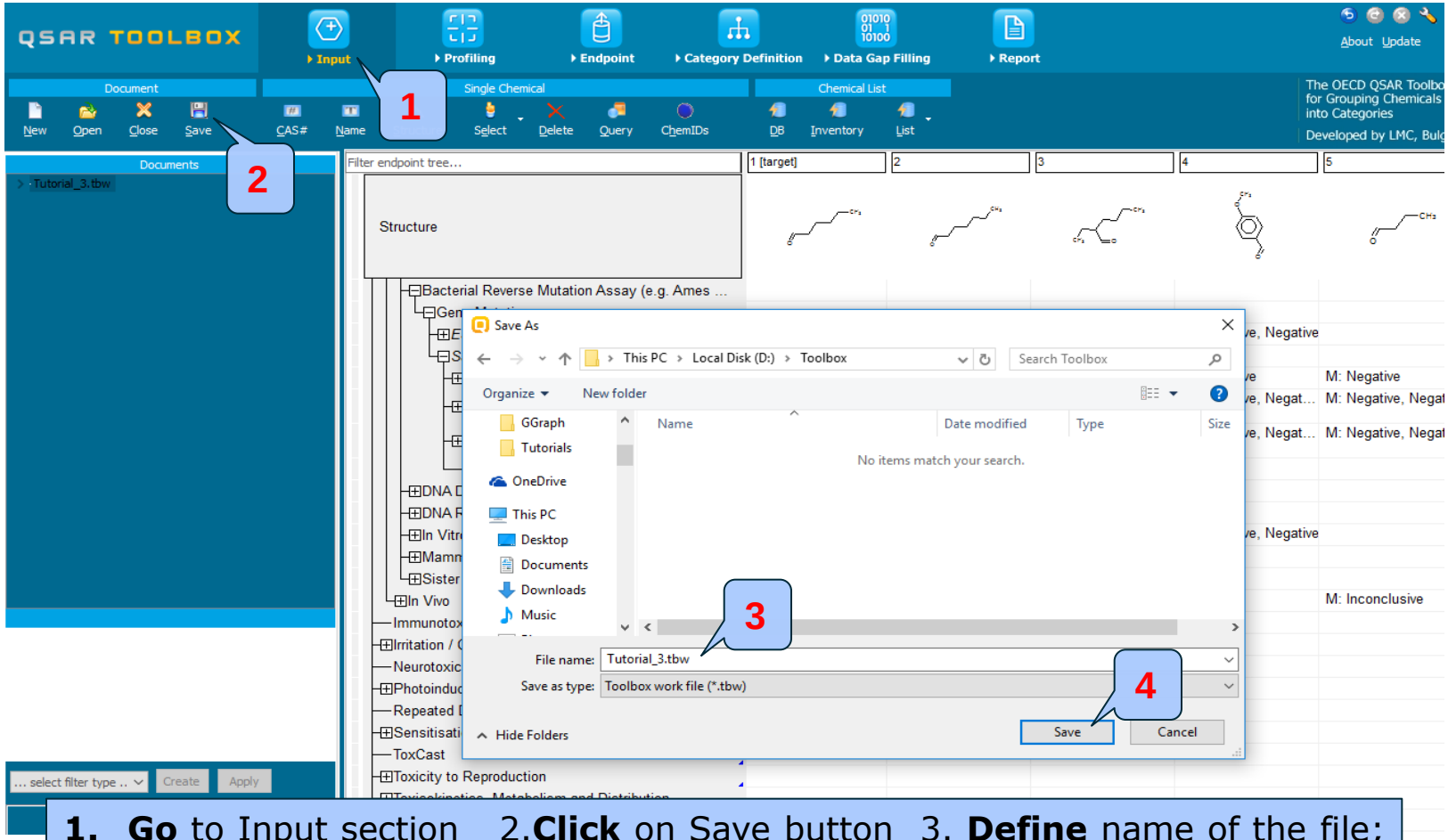
- Background
- Objectives
- Specific Aims
- Read-across
- The exercise
- Workflow of the exercise
- **Save the prediction**

Saving the prediction result

- This functionality allow storing/restoring the current state of Toolbox documents including loaded chemicals, experimental data, profiles, predictions etc, on the same computer. The functionality is implemented based on saving the sequence of actions that led to the current state of the Toolbox document and later executing these actions in the same sequence in order to get the same result(s).

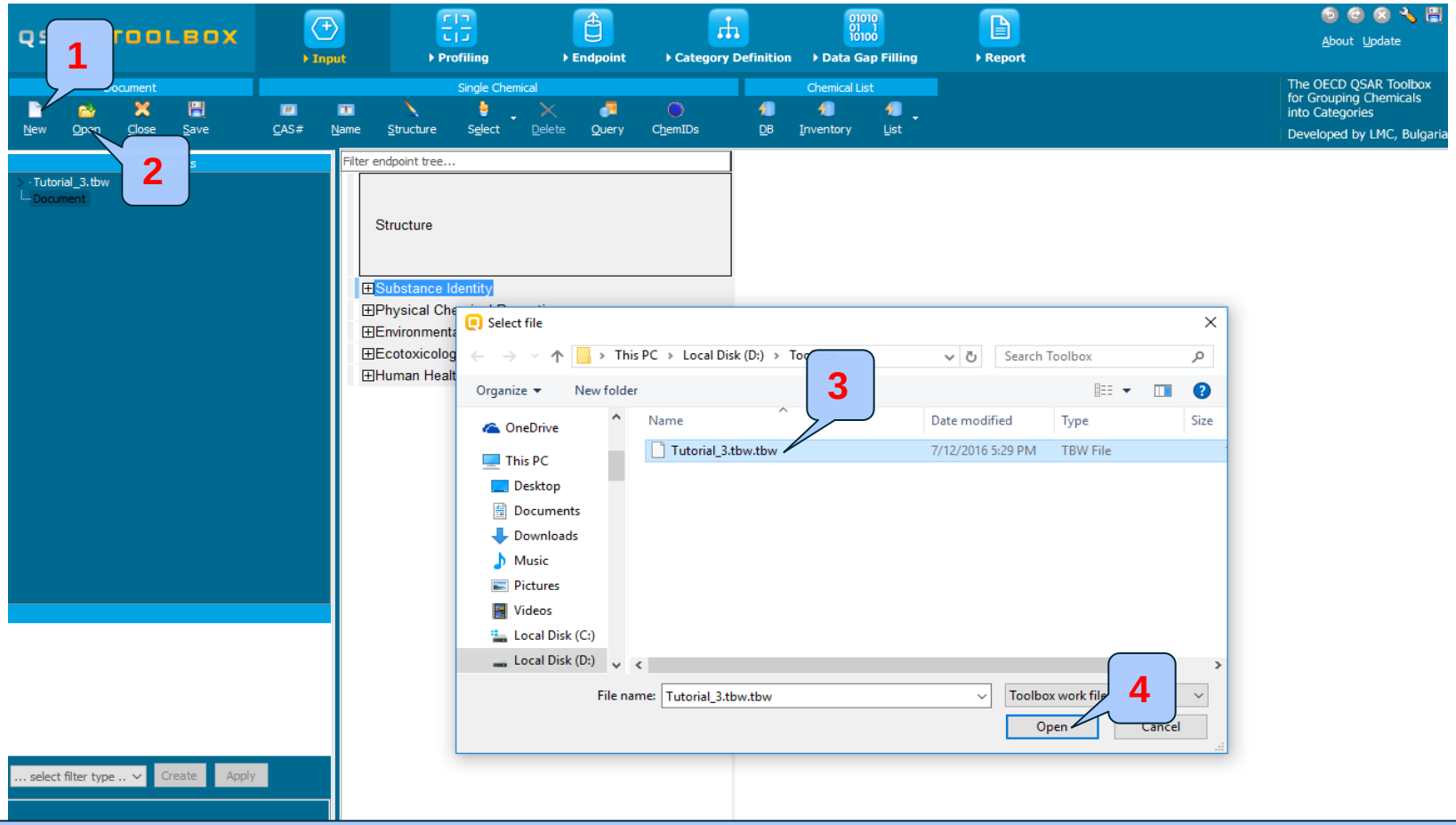
- Saving/Loading the file with TB prediction is shown on next screenshots

Saving the prediction result



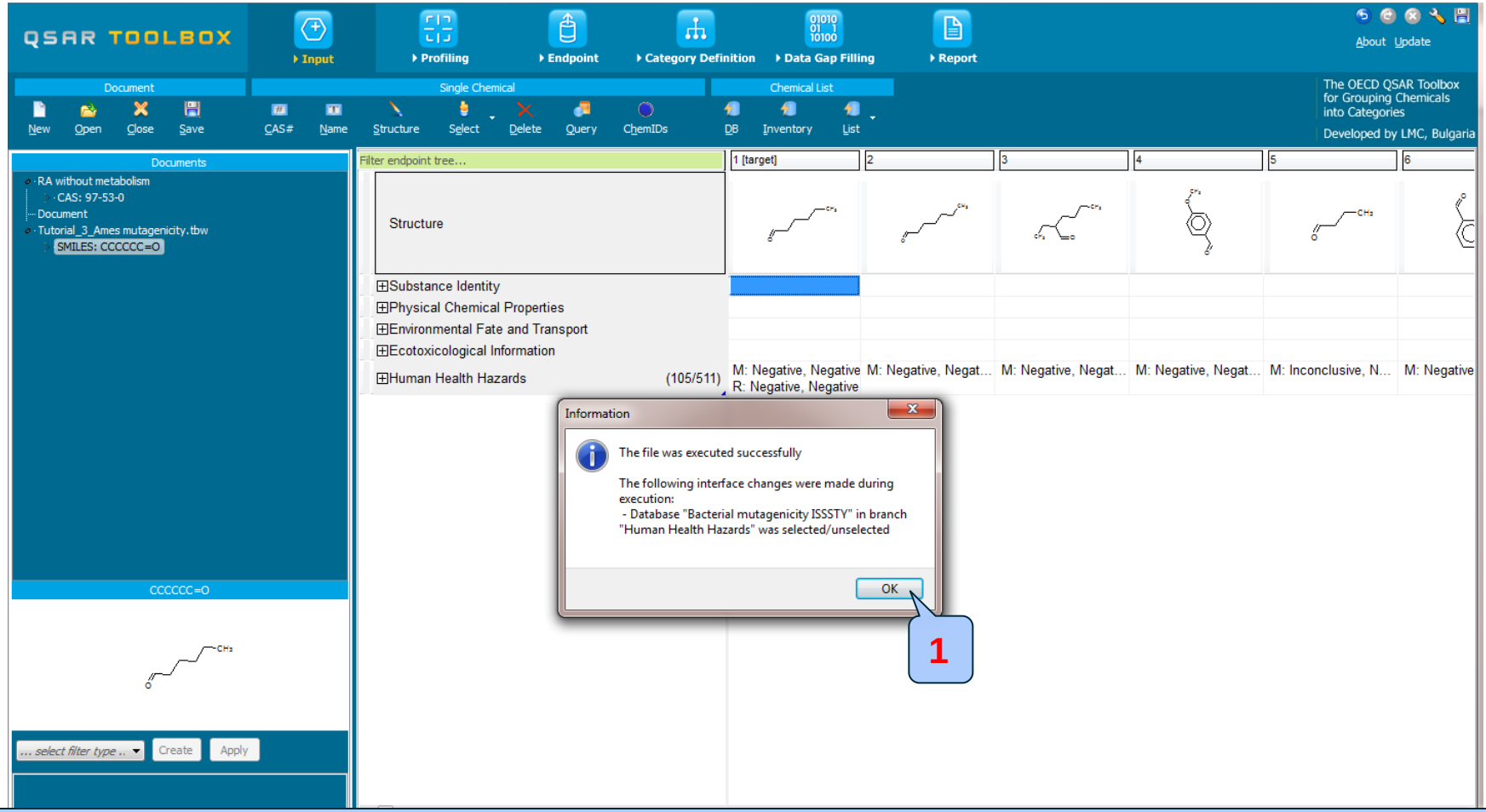
1. **Go** to Input section 2.**Click** on Save button 3. **Define** name of the file; 4. **Click** Save button

Open saved file



1. **Create** new document 2. **Click** Open; 3. **Find** and **select file**; 4. **Click** Open

Open saved file



1. The file is opened successfully

1. **Click** OK

Congratulation

- By now you should feel comfortable with the six basic modules of the toolbox and how they form the work flow of the Toolbox.
- In this tutorial you have now been introduced to several additional function in the Toolbox, especially using different profilers in subcategorizing the category of the target chemical.
- Remember proficiency in using the Toolbox will only come with practice.