QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

OECD QSAR Toolbox v.4.1

Step-by-step example of how to evaluate an ad-hoc category of aliphatic amines and to predict an ecotoxicological endpoint

Background

- This is a step-by-step presentation designed to take you through the workflow of the Toolbox for evaluating of an ad-hoc category.
- You will learn several new functionalities which we be repeated to assure a consistent category is defined.
- It is assumed that you now have some experience in using the Toolbox so there will be multiple key strokes between screen shots.

Outlook

- Background
- **Category evaluation - overview**

Category Evaluation
Two phases of evaluation process

Phase I. **Evaluation of Category & Data Gap Filling:**

- Investigating the structural consistency of an ad hoc category (e.g., a category submitted by an industry consortium to a regulatory assessment program).
- Implementation of an ad hoc category building and data gap filling.

Phase II. **Extension of Category & Data Gap Filling:**

- Search for other analogues which are consistent with the submitted category.
- Data gap filling using new data matrix.

Outlook

- Background
- Category evaluation – overview
  - **<u>Phase I.</u> Evaluation of Category & Data Gap Filling**
    - *<u>Step 1:</u> Investigating the structural consistency of an ad hoc category.*
      - ***Case study***

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Case Study

- The submission consists of a category with 19 aliphatic amines.
- The predicted ecotoxicological endpoint EC50, 48h, *D.magna* of 2-Butanamine(CAS 13952-84-6) will be reviewed.

Outlook

- Background
- Category evaluation – overview
  - **Phase I. Evaluation of Category & Data Gap Filling**
    - *Step 1: Investigating the structural consistency of an ad hoc category.*
      - *Case study*
      - ***Workflow***

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow

The following input workflow is used:

- Input the file for submission as an user list.(In this case input file Aliphatic amines.smi*)
- Evaluate the category applying the following profiling schemes:
  - US-EPA New Chemical Categories
  - Aquatic toxicity classification by ECOSAR
  - Aquatic toxicity MOA of action
  - Organic functional groups (nested)

Aliphatic amines.smi* - file is available with TB installation, located at C:\Program Files (x86)\Common Files\QSAR Toolbox 4.1\Config\Examples

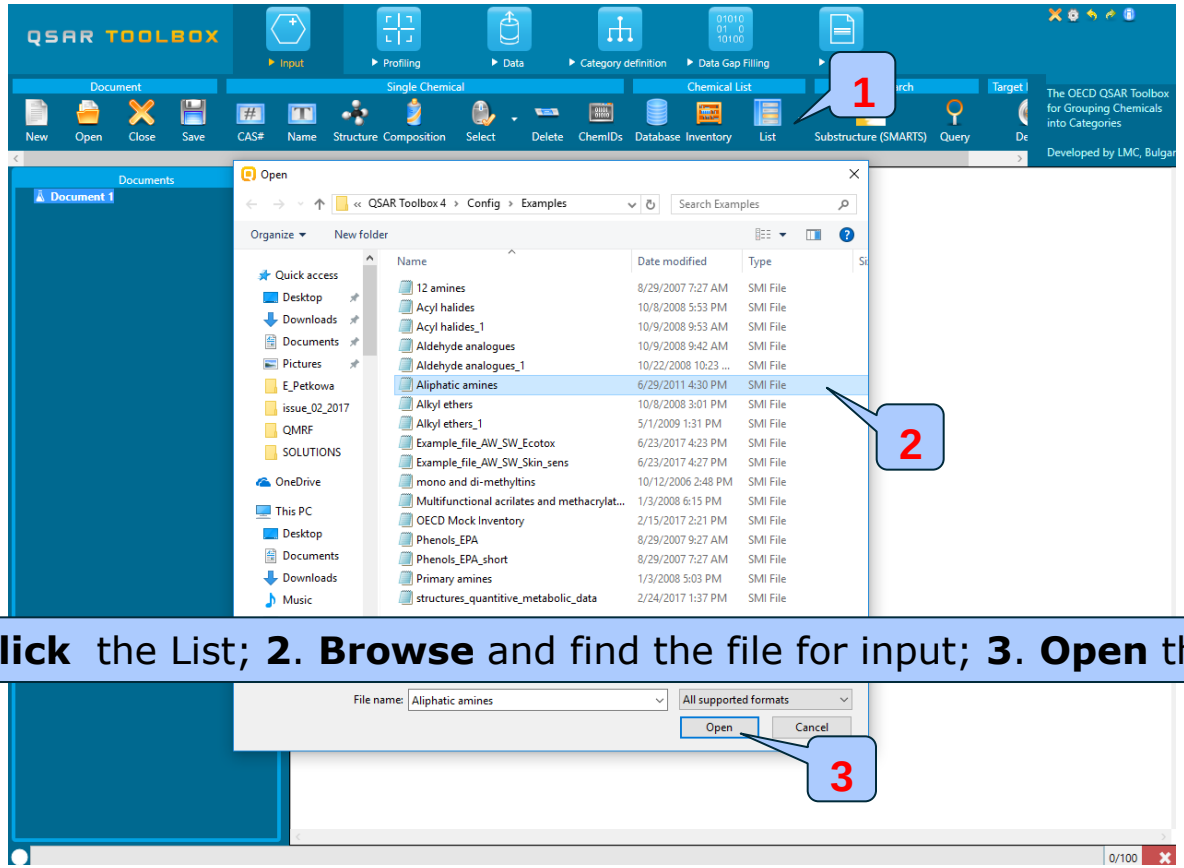
Outlook

- Background
- Category evaluation – overview
  - **Phase I. Evaluation of Category & Data Gap Filling**
    - *Step 1: Investigating the structural consistency of an ad hoc category.*
      - *Case study*
      - ***Workflow***
        - ***Input***

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Input



1. **Click** the List; **2**. **Browse** and find the file for input; **3**. **Open** the file.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

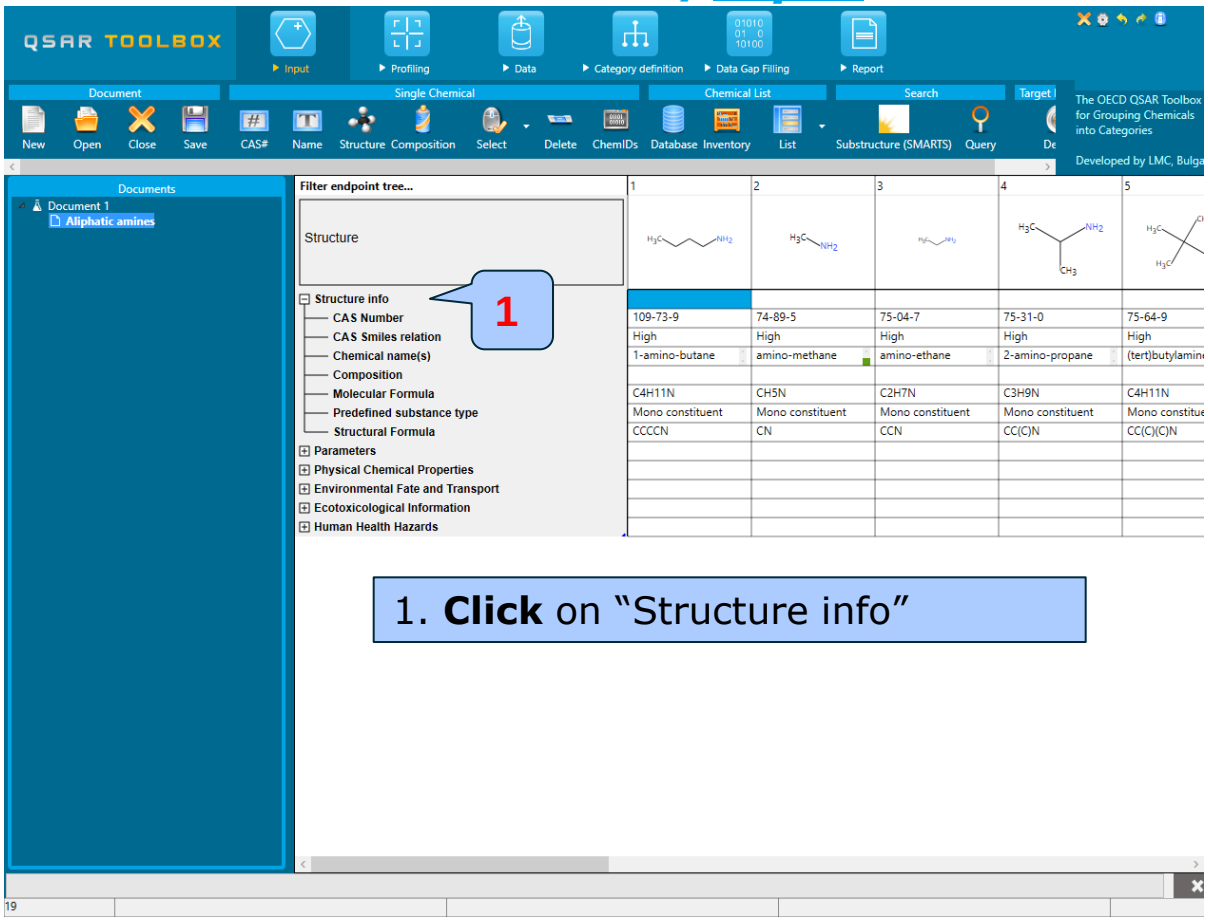
Workflow/Input

- You have now inserted your chemical list into the system.
- Click on the box next to "Structure info"; this displays the chemical identification information (see next screen shot).

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Input



Outlook

- Background
- Category evaluation – overview
  - **Phase I. Evaluation of Category & Data Gap Filling**
    - *Step 1: Investigating the structural consistency of an ad hoc category.*
      - *Case study*
      - ***Workflow***
        - *Input*
        - ***Profiling***

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

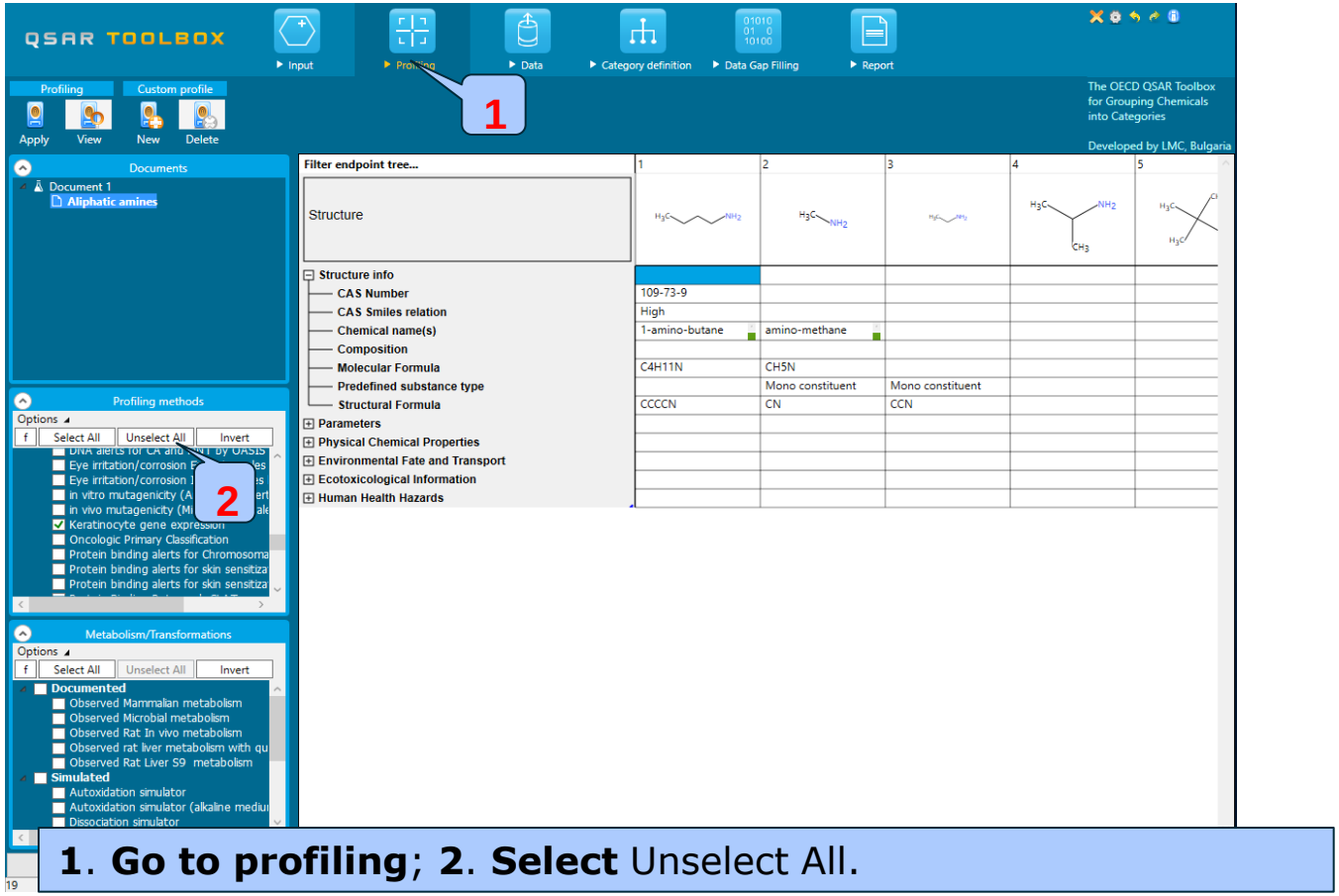
- The first step of the category evaluation according to Phase I is the consistency check of the category which starts with evaluation of the robustness with respect to structural functionalities. The following schemes could be used for this purpose:
  - US-EPA New Chemical Categories
  - Aquatic toxicity classification by ECOSAR
  - Aquatic toxicity MOA of action
  - Organic functional groups (nested)

- Select "profiling methods" by clicking on the boxes before the names of the profilers and Click "Apply". Before selecting the profiling methods unselect all (see next screen shot).

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

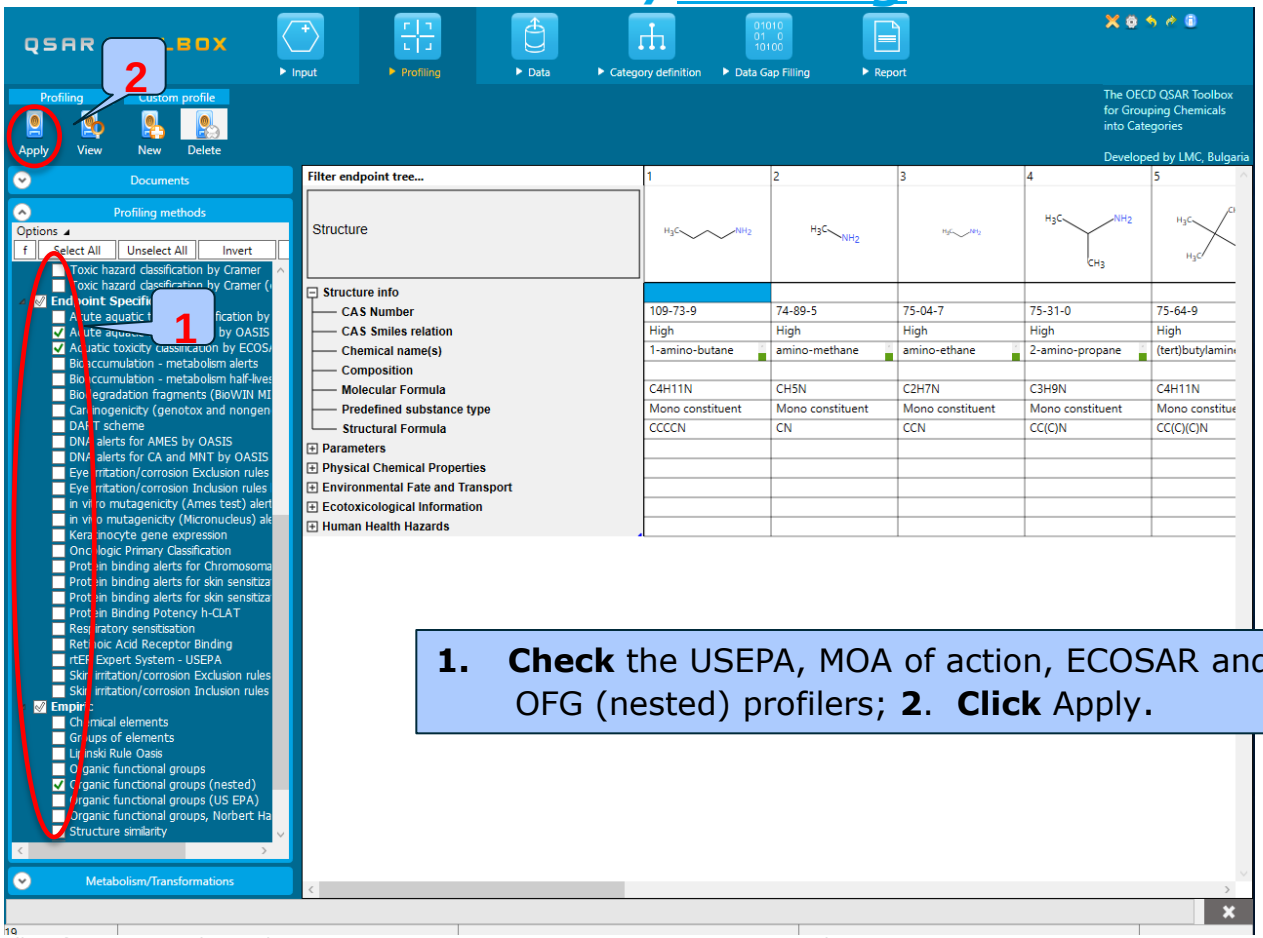


1. **Go to profiling**; **2**. **Select** Unselect All.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling



1. **Check** the USEPA, MOA of action, ECOSAR and OFG (nested) profilers; **2.** **Click** Apply.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

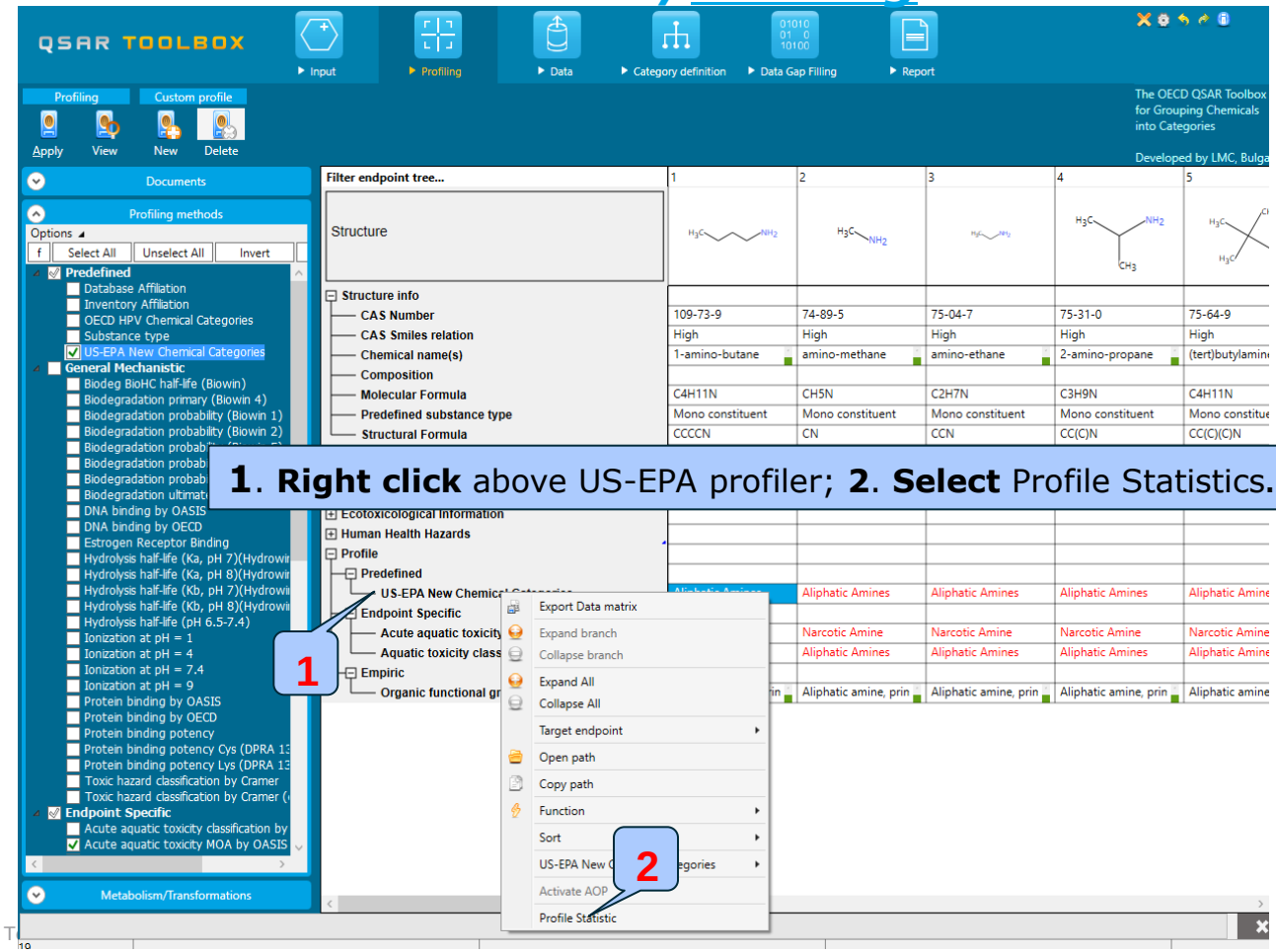
Workflow/Profiling

- The actual profiling will take several seconds depending on the number and type of selected profilers.
- The results of profiling automatically appeared as a dropdown box under the target chemical.
- The result from profile statistics, can be seen by right clicking in the space above the profiler in the endpoint tree and select Profile statistics from the dropdown menu (see next screen shot).

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

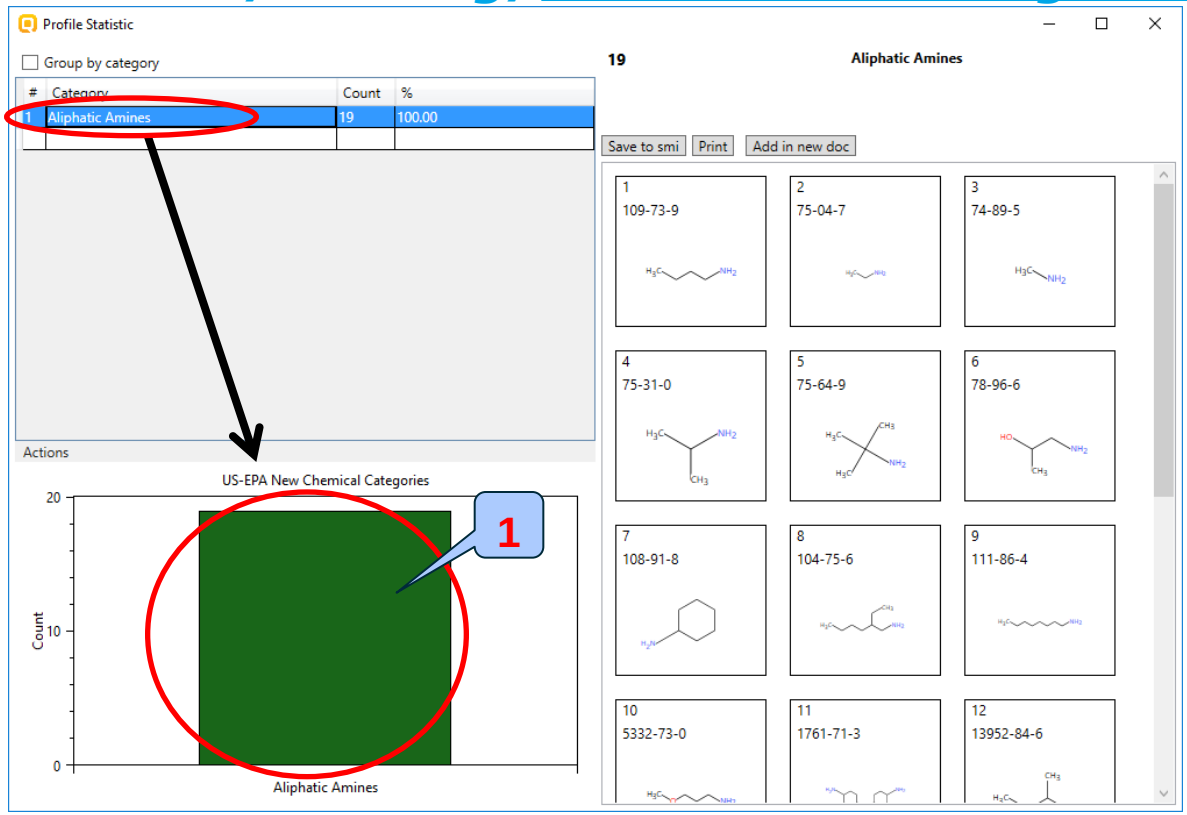
Workflow/Profiling



Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to US-EPA

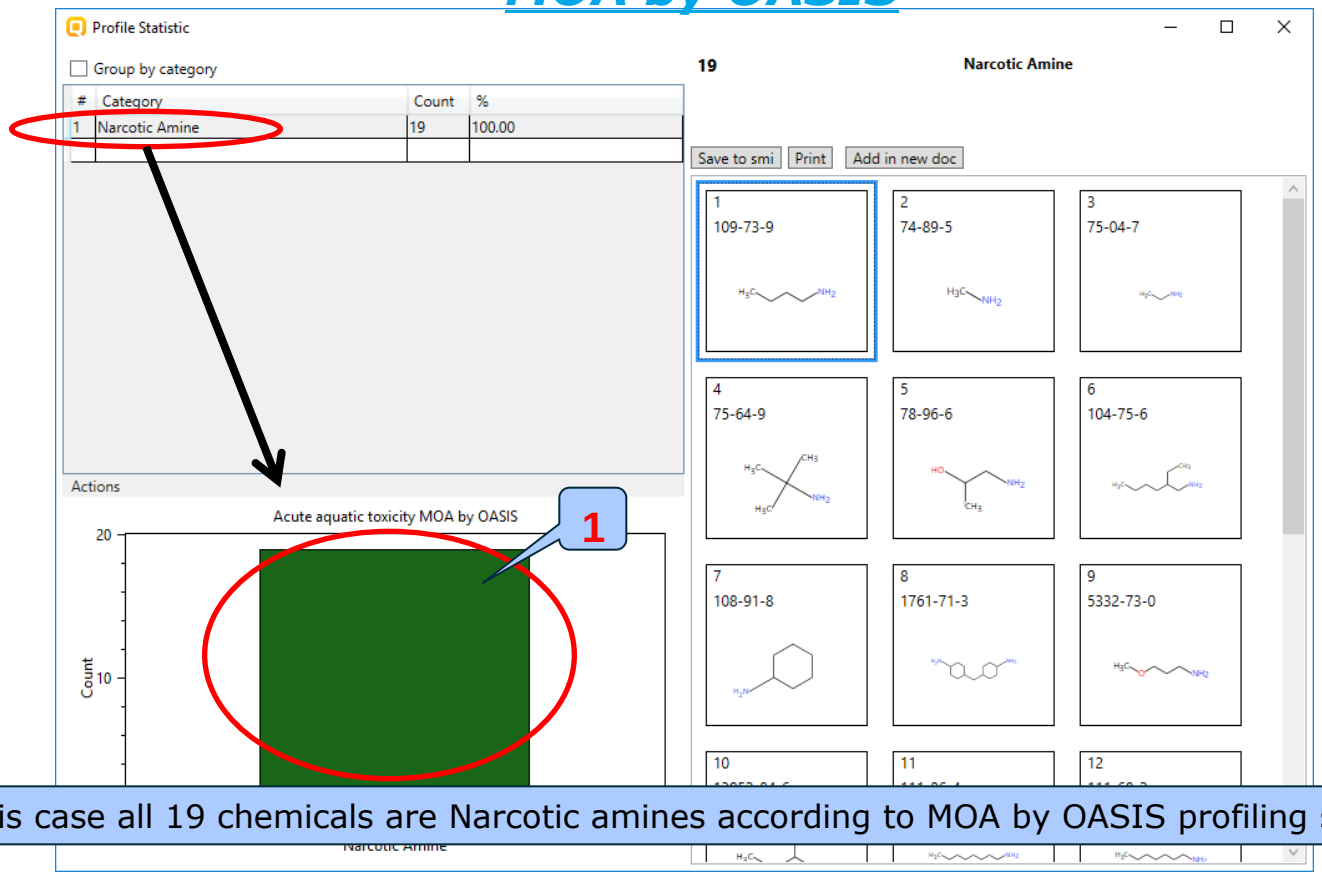


1. In this case all 19 chemicals are Aliphatic amines according to US-EPA profiler.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to Aquatic toxicity MOA by OASIS

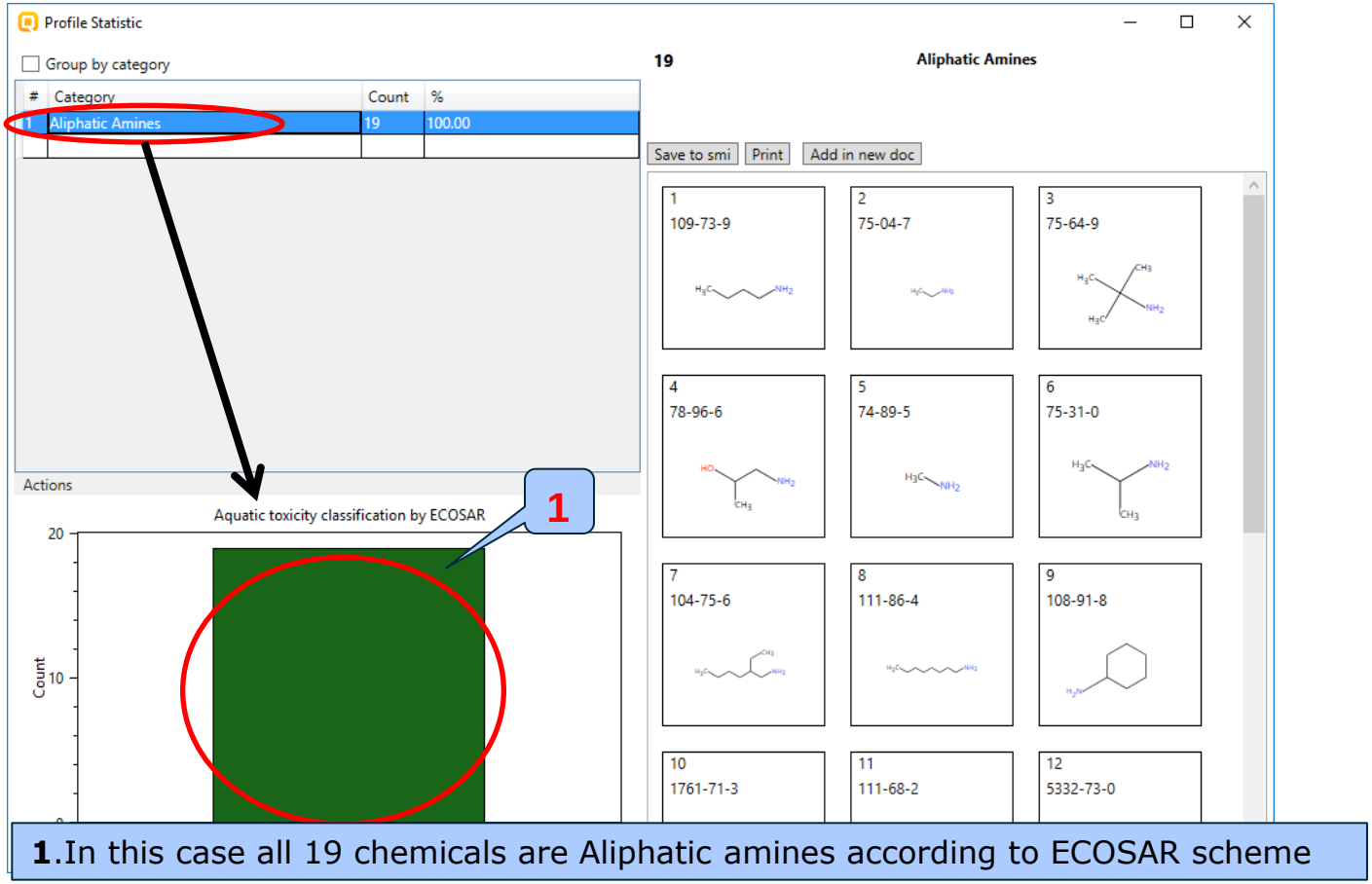


1. In this case all 19 chemicals are Narcotic amines according to MOA by OASIS profiling scheme

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to ECOSAR

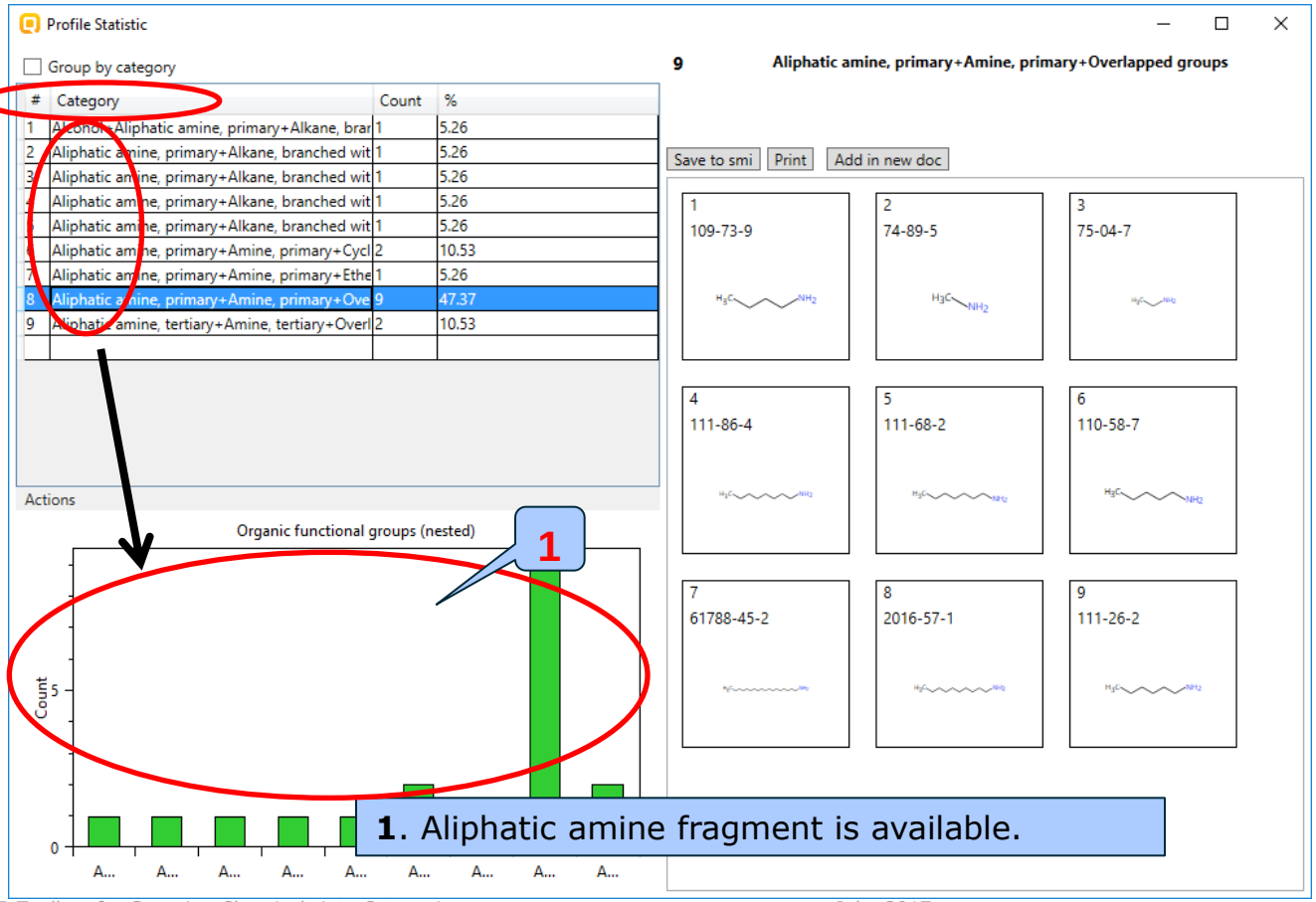


1.In this case all 19 chemicals are Aliphatic amines according to ECOSAR scheme

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to OFG(nested)



Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Recap

- **Chemicals are defined as:**
  - Aliphatic amines (broader category than primary amines) – US-EPA categories, ECOSAR classification
  - The statistics of organic functional groups provide detailed alert description of all 19 structures. However, all 19 chemicals have aliphatic amines fragment.
- **It could be concluded that the category is consistent with respect to structural functionalities (chemicals are empirically similar).**

Outlook

- Background
- Category evaluation – overview
  - **<u>Phase I.</u> Evaluation of Category & Data Gap Filling**
    - *<u>Step 1:</u> Investigating the structural consistency of an ad hoc category.*
    - *<u>Step 2:</u> Investigating the a*pplicability domain of an ad hoc category.

Phase I: Evaluation of Category & Data Gap Filling

Step 2. Investigating the applicability domain of an ad hoc category

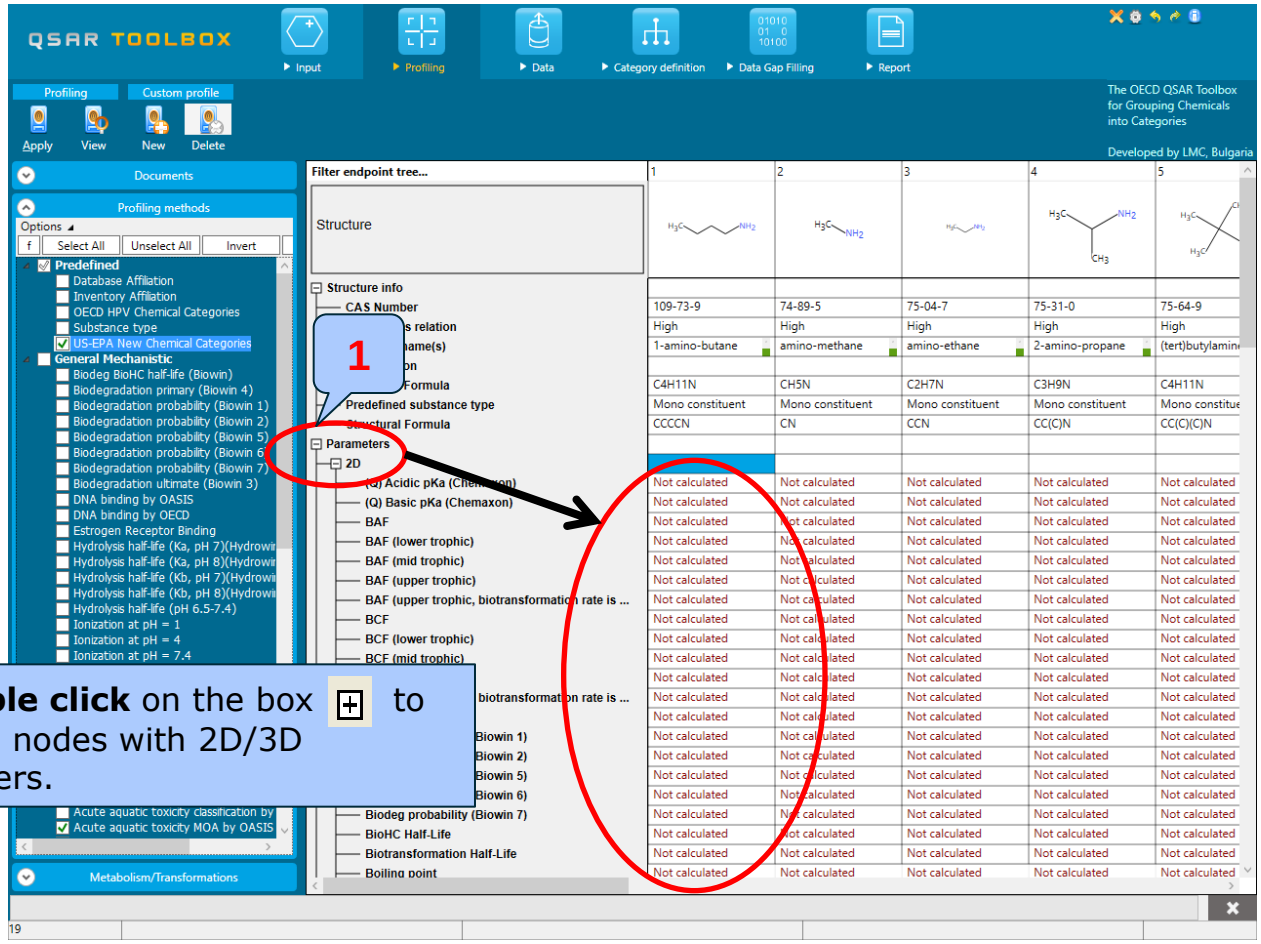
- Boundaries of structural functionalities
  - Aliphatic amines
- Parametric boundaries
  - log Kow (from 0.64 to 7.71)
  - Molecular weight (from 31 to 269 Da)
  - Water solubility (from $0.48x10^{-1}$ to $1x10^{6}$ mg/l)

You are now ready to extract the 2D and/or 3D parameters (see next screen shot).

Phase I: Evaluation of Category & Data Gap Filling

Step 2. Investigating the applicability domain of an ad hoc category

Extracting 2D and 3D parameters

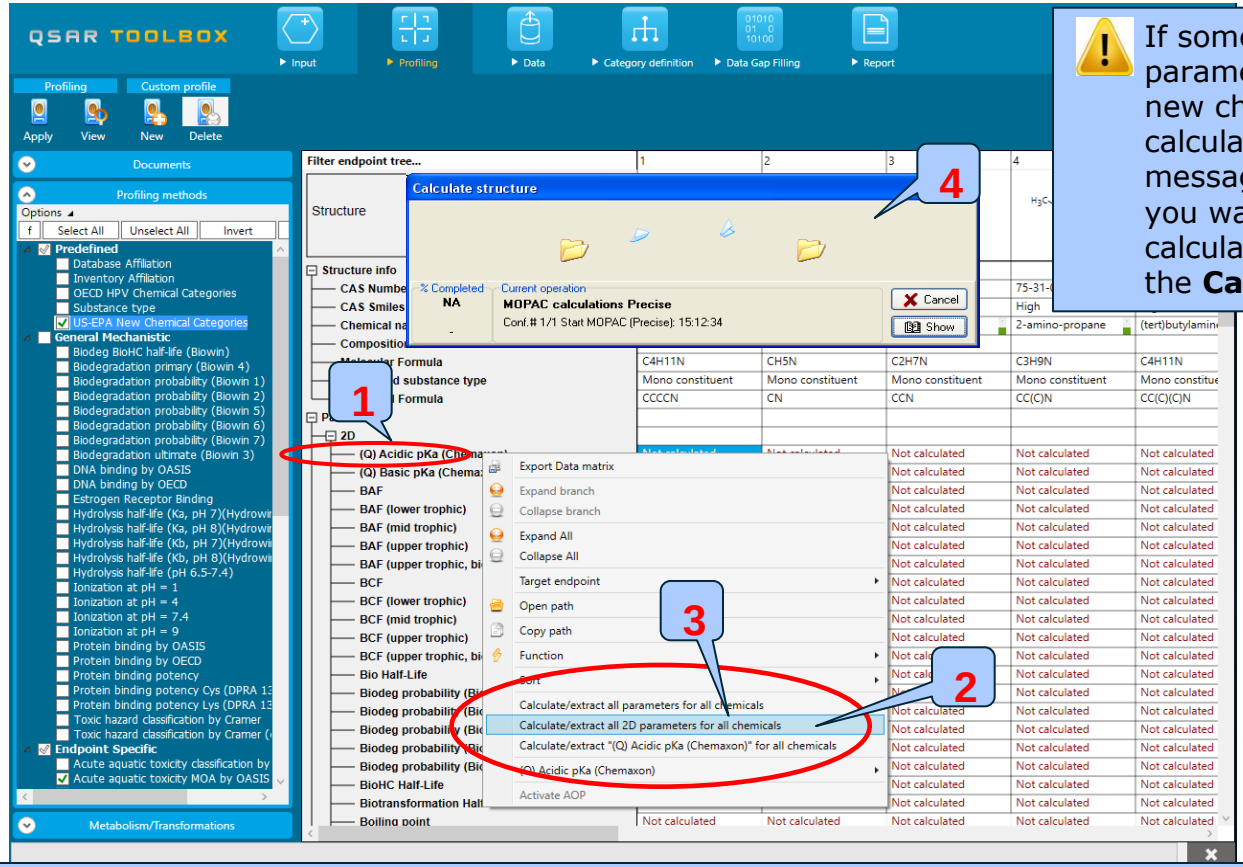


1. **Double click** on the box ⊞ to open the nodes with 2D/3D parameters.

Phase I: Evaluation of Category & Data Gap Filling

Step 2. Investigating the applicability domain of an ad hoc category

Extracting 2D and 3D parameters



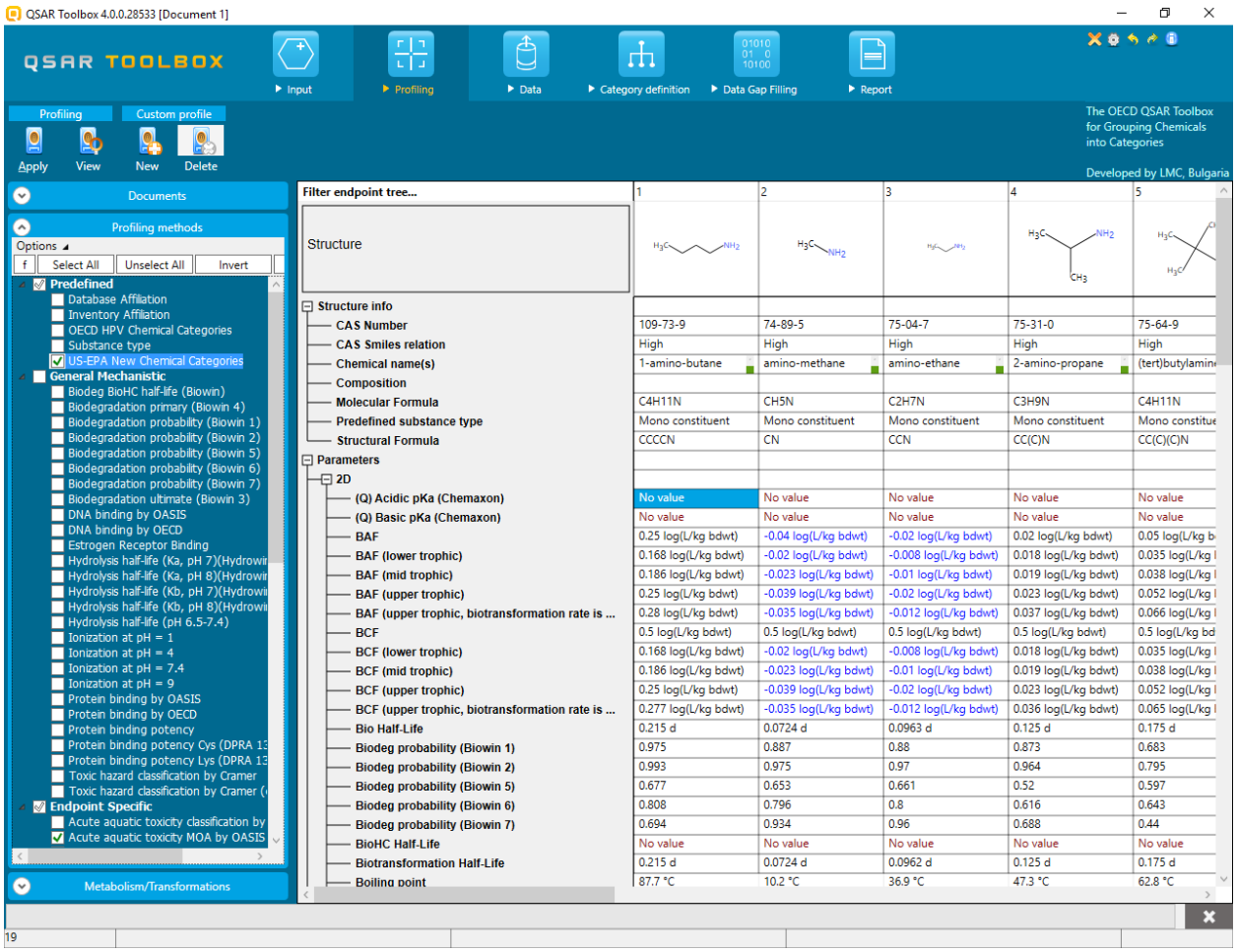
If some of the parameters (in case of new chemicals) are not calculated then the message (**4**) appears If you want to stop the calculation process, click the **Cancel** button.

1. **Right click** above the parameter in the endpoint tree; **2**. The user can calculate all parameters or extract current parameter for all chemicals in the current row. **3**. In this exercise calculate all (2D) parameters.

Phase I: Evaluation of Category & Data Gap Filling

Step 2. Investigating the applicability domain of an ad hoc category

Extracting 2D and 3D parameters - results



Outlook

- Background
- Category evaluation – overview
  - **Phase I. Evaluation of Category & Data Gap Filling**
    - *Step 1: Investigating the structural consistency of an ad hoc category.*
    - *Step 2: Investigating the a*pplicability domain of an ad hoc category.
    - *Step 3: Reading data for the analogues*

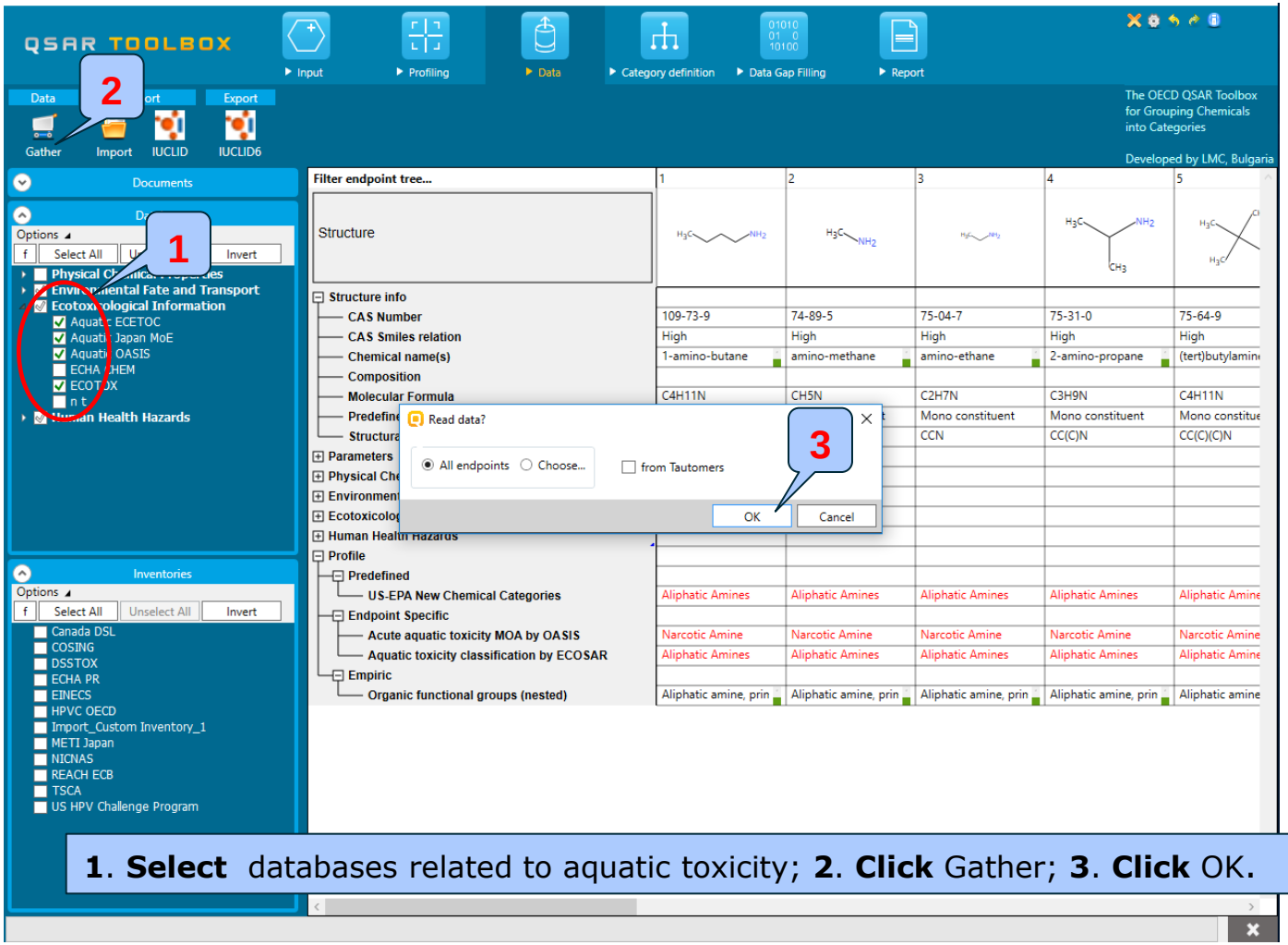
Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

- Next, the Data matrix is constructed by extracting available experimental results for all 19 members of the category.
- Based on preceding category evaluation, no outliers have been identified violating the structural and mechanistic consistency of the category.

Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

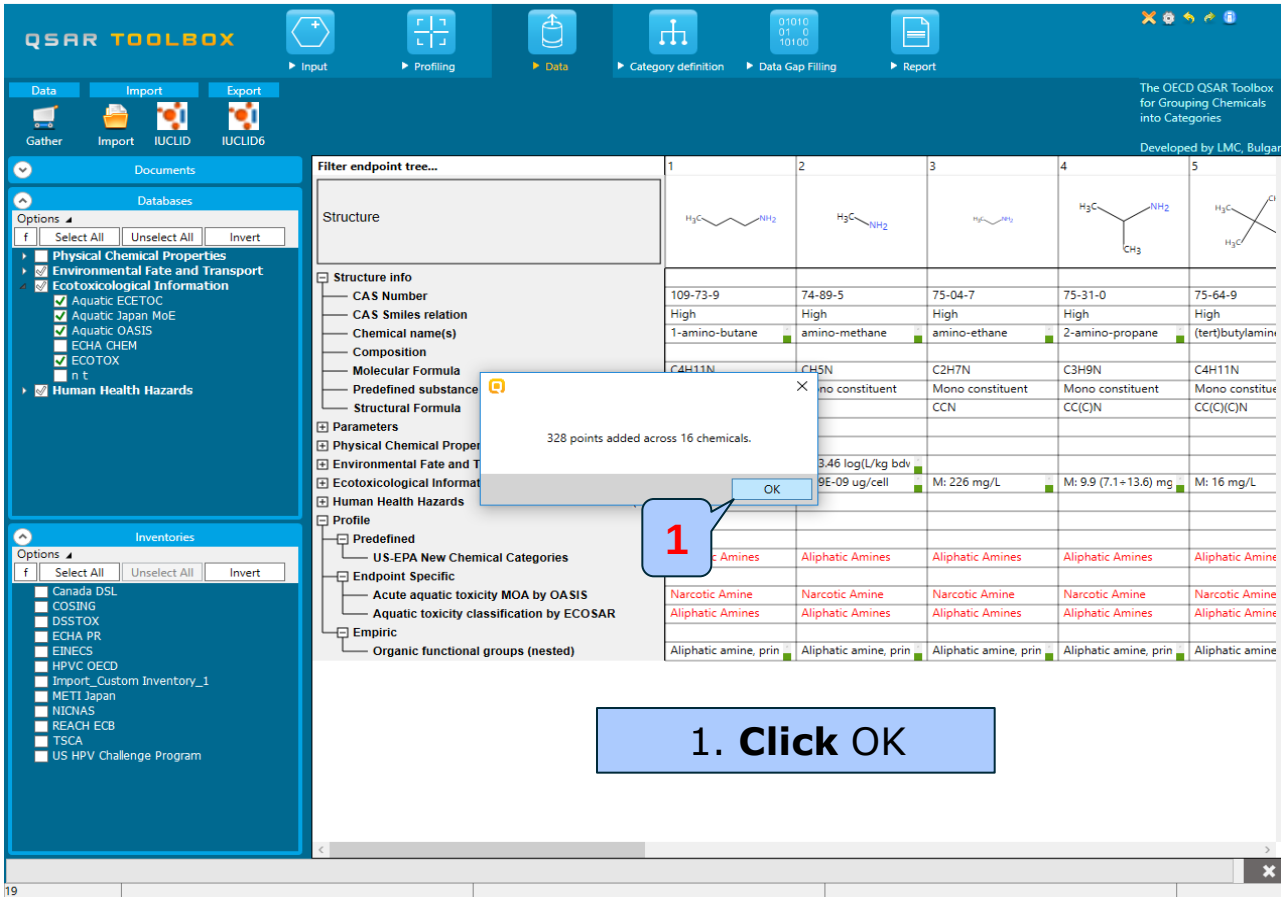


1. **Select** databases related to aquatic toxicity; **2**. **Click** Gather; **3**. **Click** OK.

Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

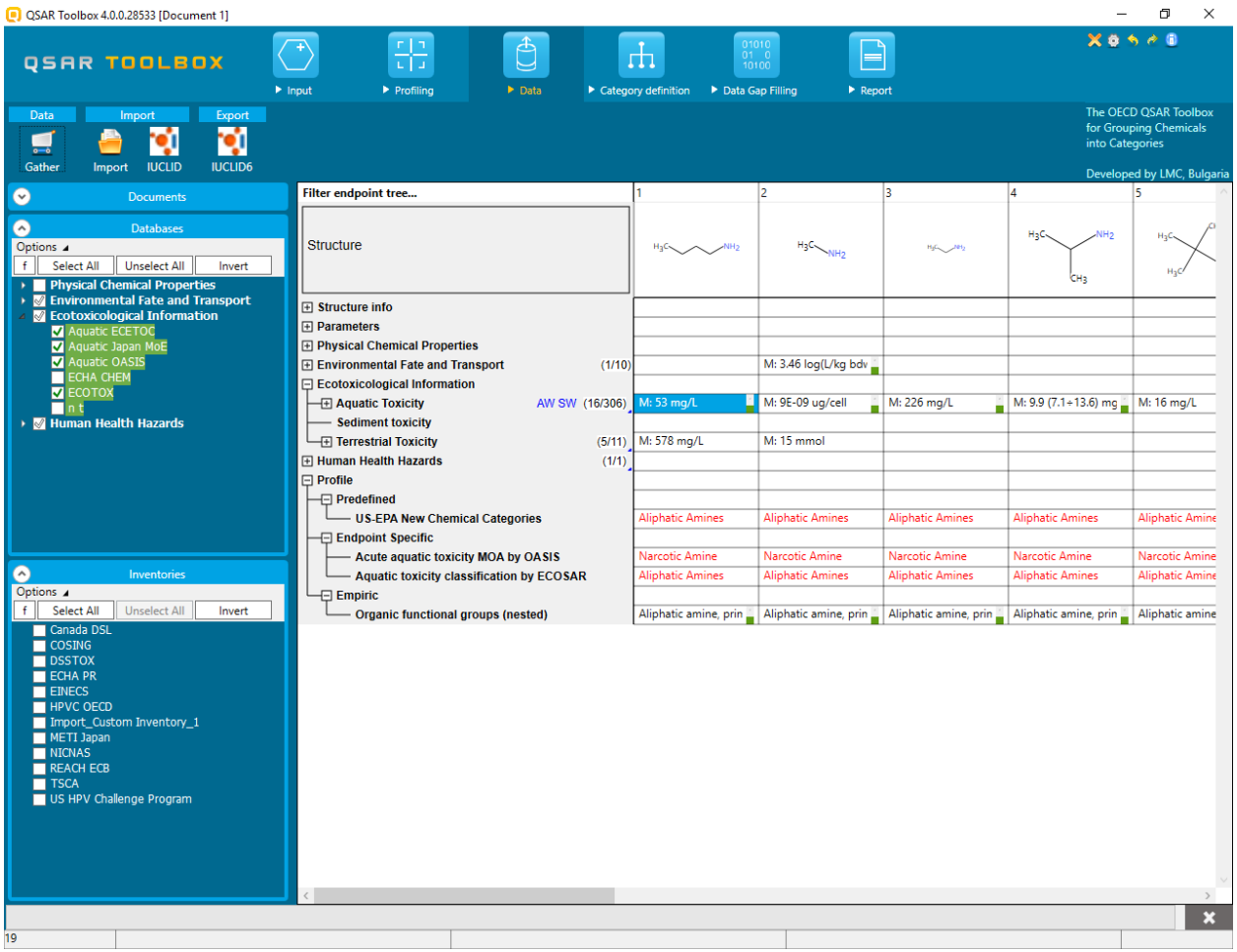
The system automatically gives indication for the number of gather experimental data points



Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data.

Inserting data into the data matrix



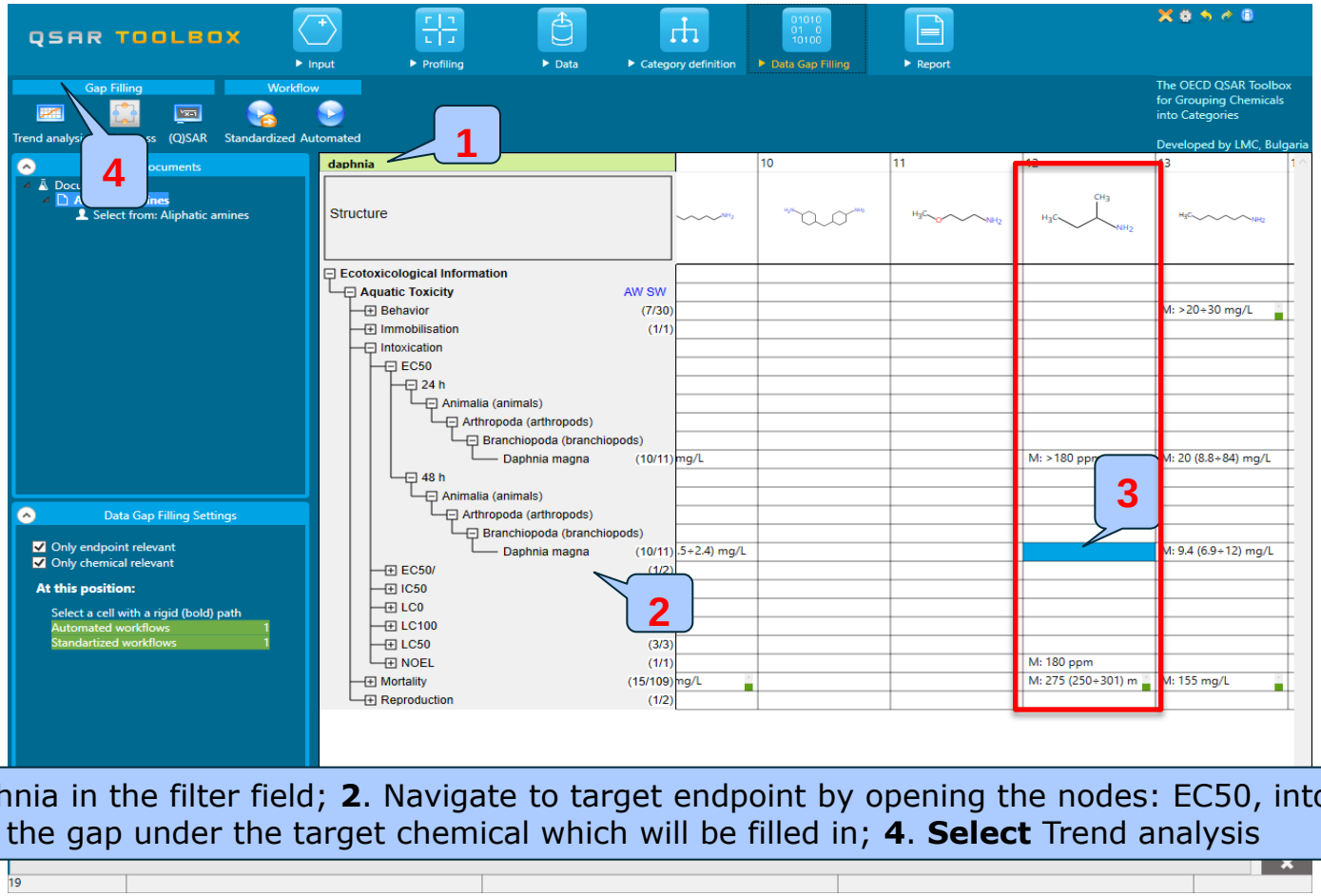
Outlook

- Background
- Category evaluation – overview
  - **Phase I. Evaluation of Category & Data Gap Filling**
    - *Step 1: Investigating the structural consistency of an ad hoc category.*
    - *Step 2: Investigating the a*pplicability domain of an ad hoc category.
    - *Step 3: Reading data for the analogues*
    - *Step 4: Data gap filling for 2-Butanamine*

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine

Navigate to the target endpoint: Intoxication, D.magna,EC50, 48h

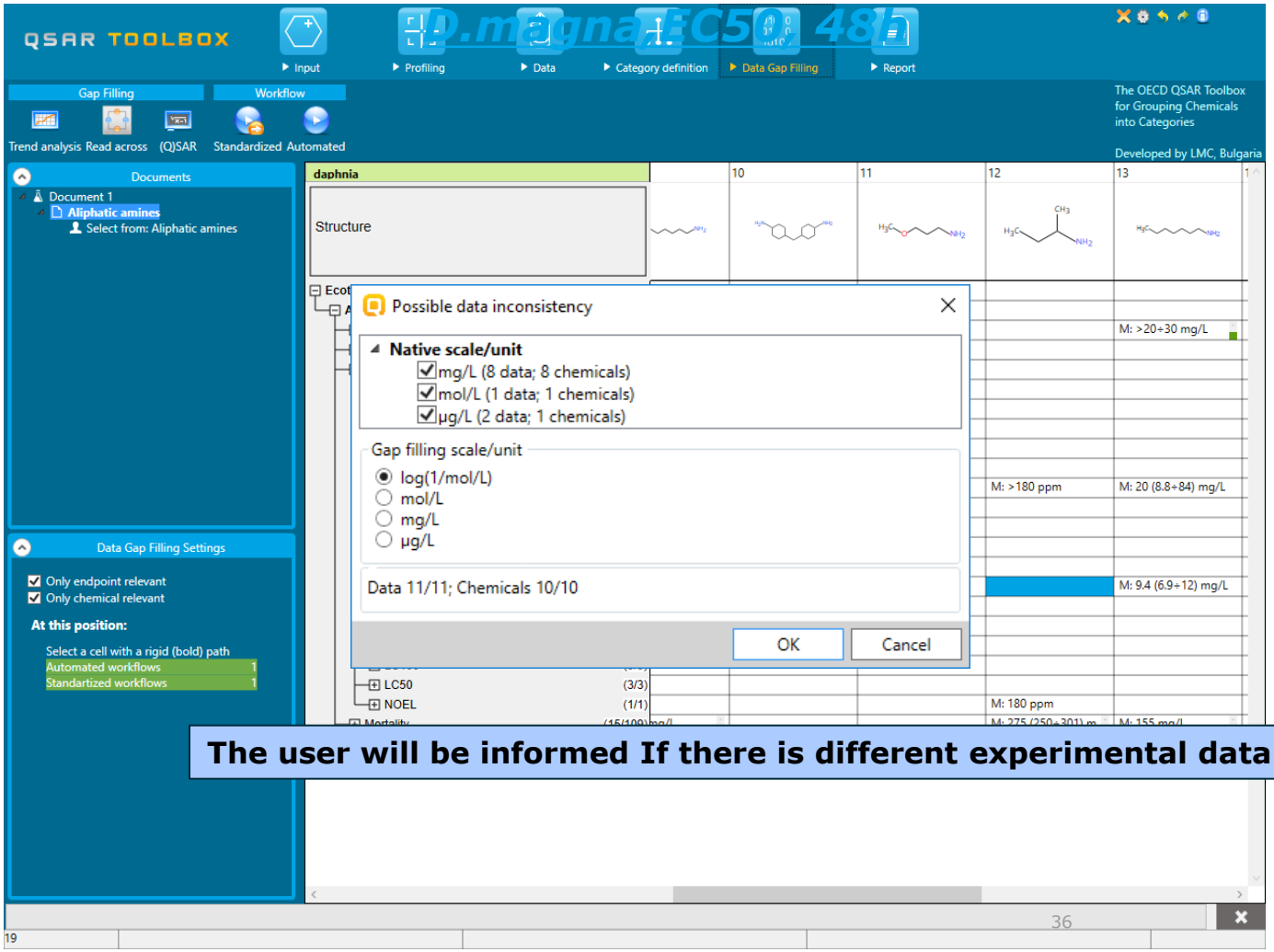


1. Type Daphnia in the filter field; **2**. Navigate to target endpoint by opening the nodes: EC50, intoxication, 48h
3. **Highlight** the gap under the target chemical which will be filled in; **4**. **Select** Trend analysis

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine

Navigate to the target endpoint: Intoxication,

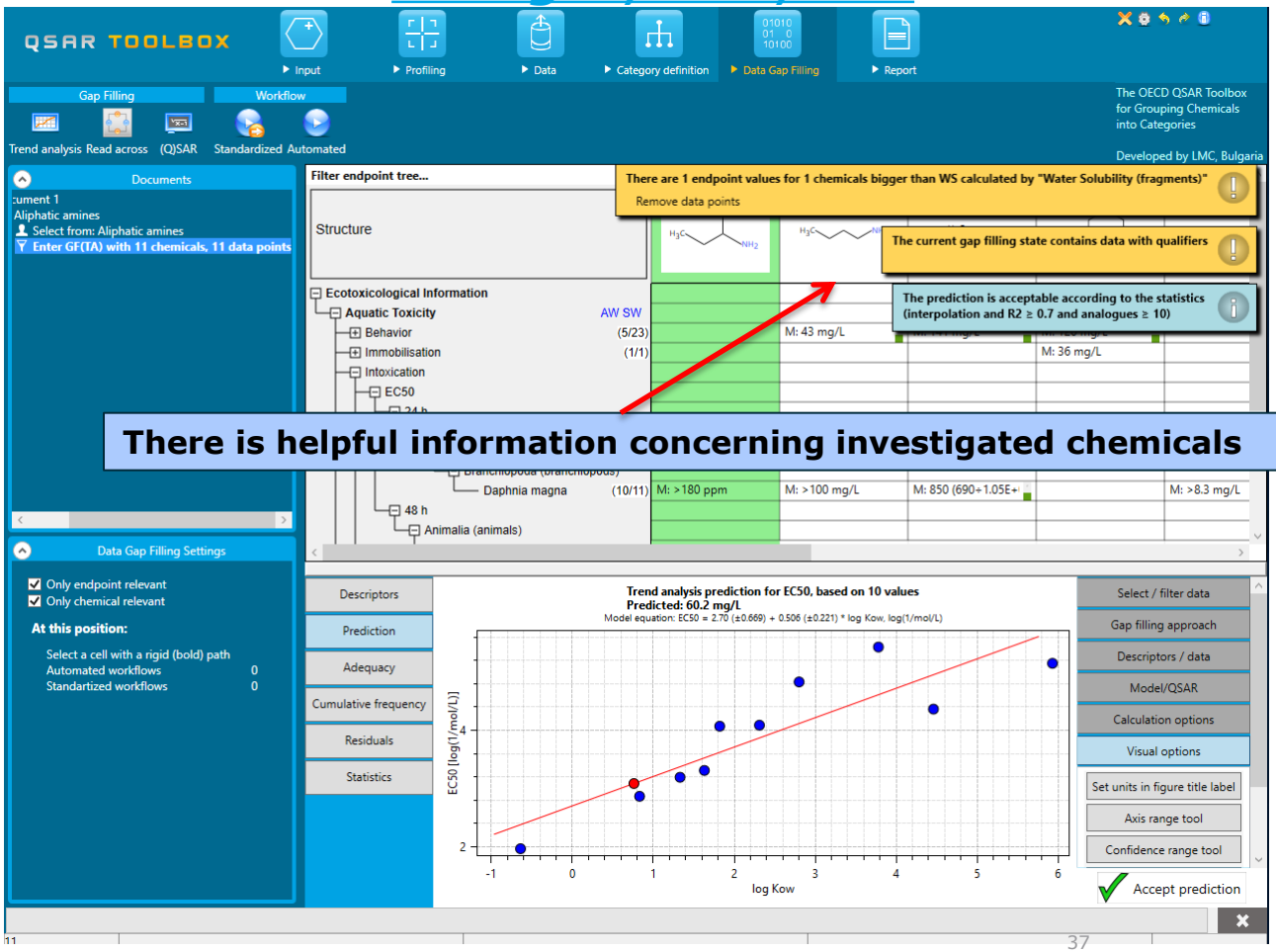


The user will be informed If there is different experimental data

Phase I: Evaluation of Category & Data Gap Filling

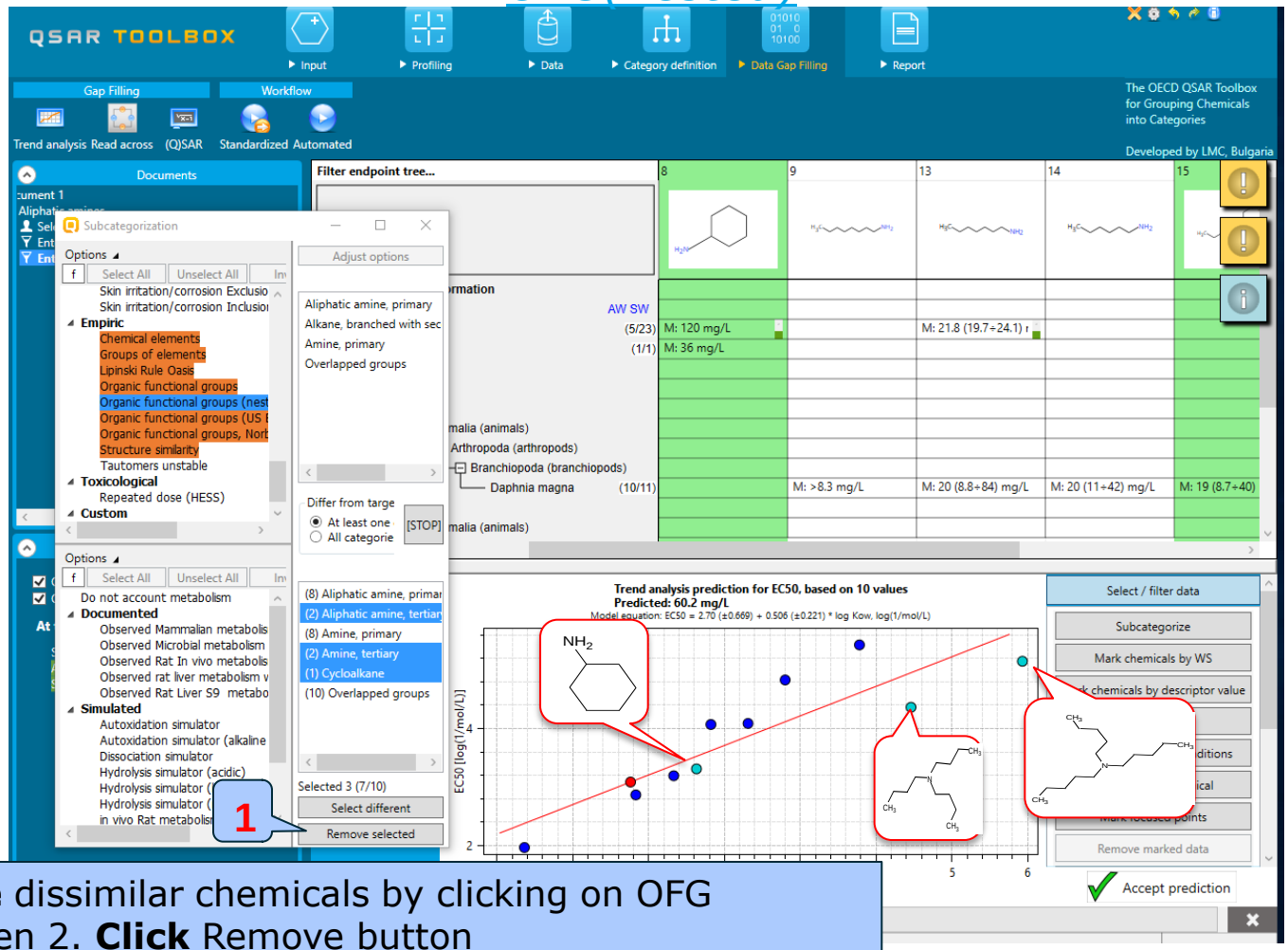
Step 4. Data Gap Filling for 2-Butanamine

Navigate to the target endpoint: Intoxication, D.magna,EC50, 48h



Phase I: Evaluation of Category & Data Gap Filling

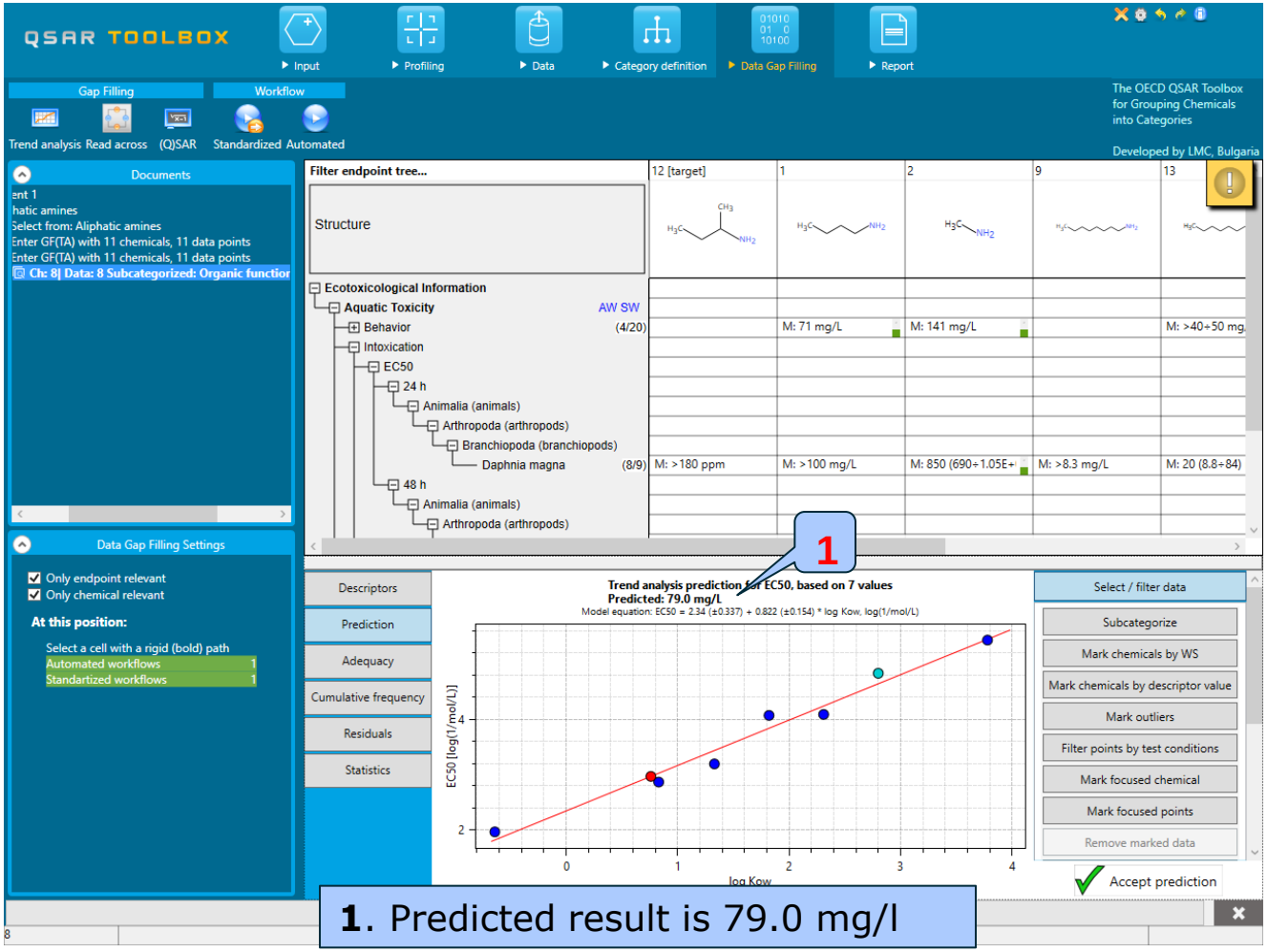
Step 4. Data Gap Filling for 2-Butanamine/*Subcategorize by OFG(nested)*



1. **Remove** dissimilar chemicals by clicking on OFG (nested) then 2. **Click** Remove button

Phase I: Evaluation of Category & Data Gap Filling

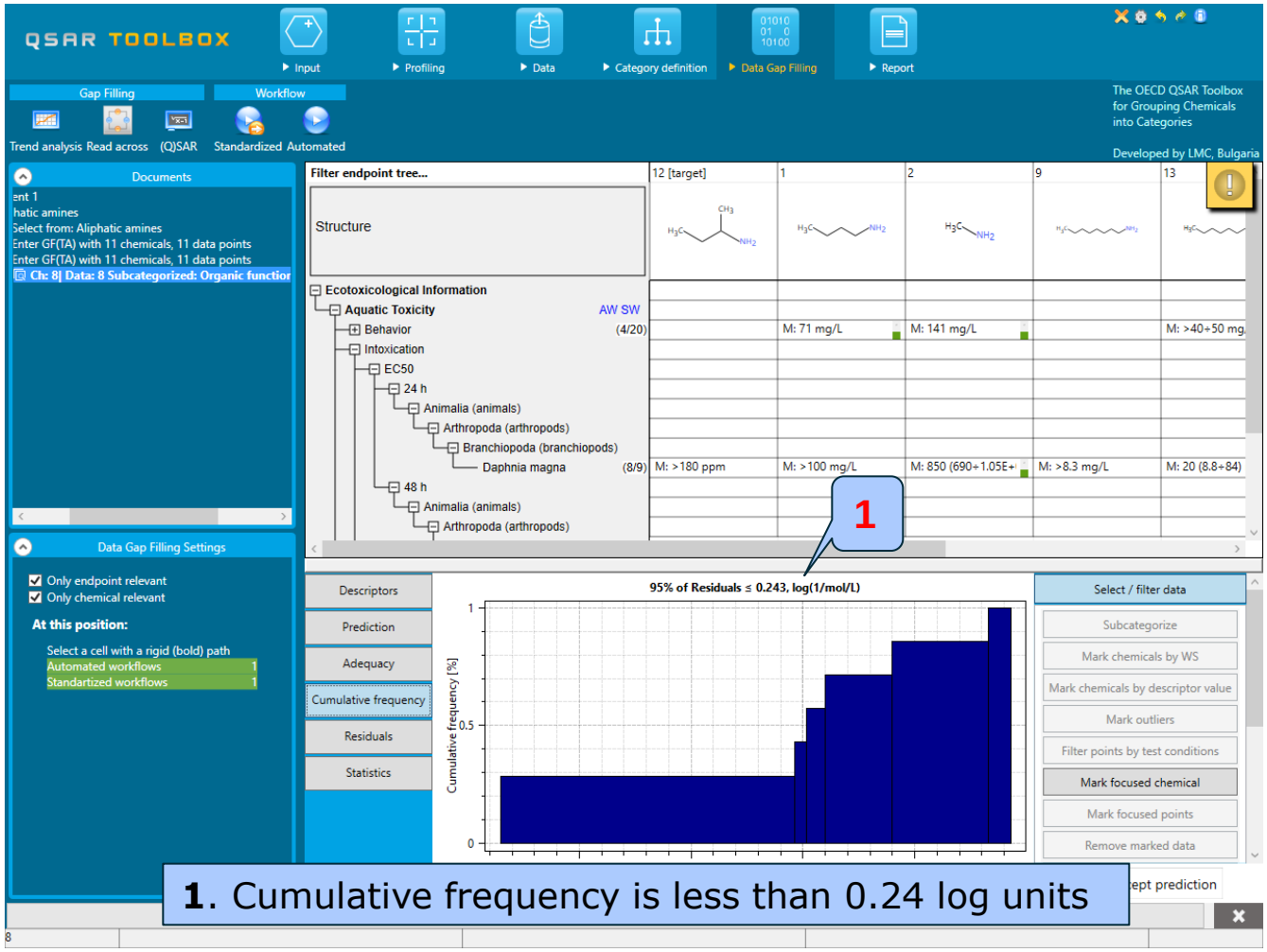
Step 4. Data Gap Filling for 2-Butanamine/*Predicted result*



1. Predicted result is 79.0 mg/l

Phase I: Evaluation of Category & Data Gap Filling

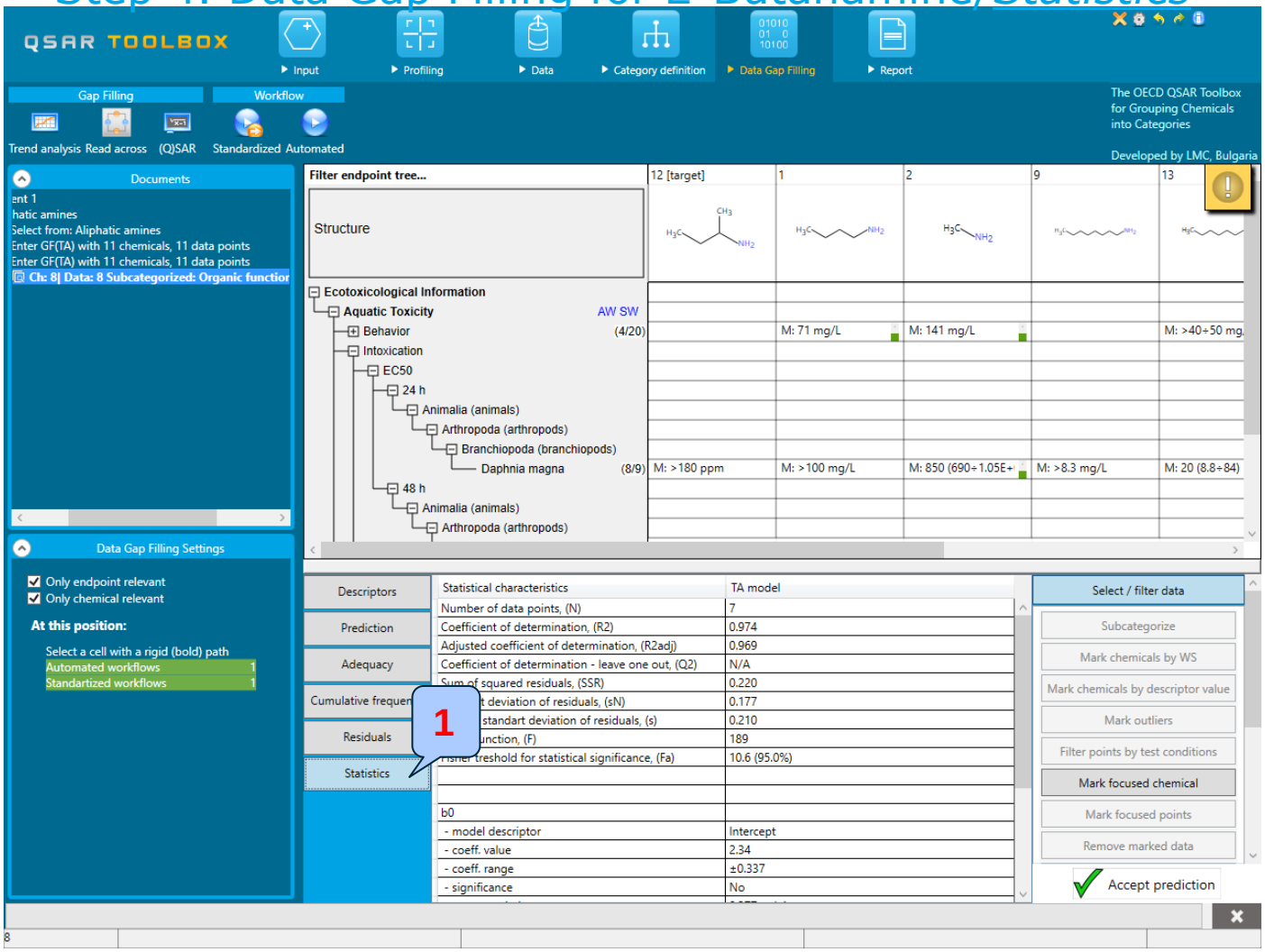
Step 4. Data Gap Filling for 2-Butanamine/*Cumulative frequency*



1. Cumulative frequency is less than 0.24 log units

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine/*Statistics*



Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine/*Interpretation of the result*

- The structurally similar analogs across category of aliphatic amines is used for data gap filling
- Subcategorization by Organic functional groups(nested) is applied
- The prediction based on the defined category is acceptable.
- The predicted value based on predefined category of aliphatic amines is 79.0 mg/l

Outlook

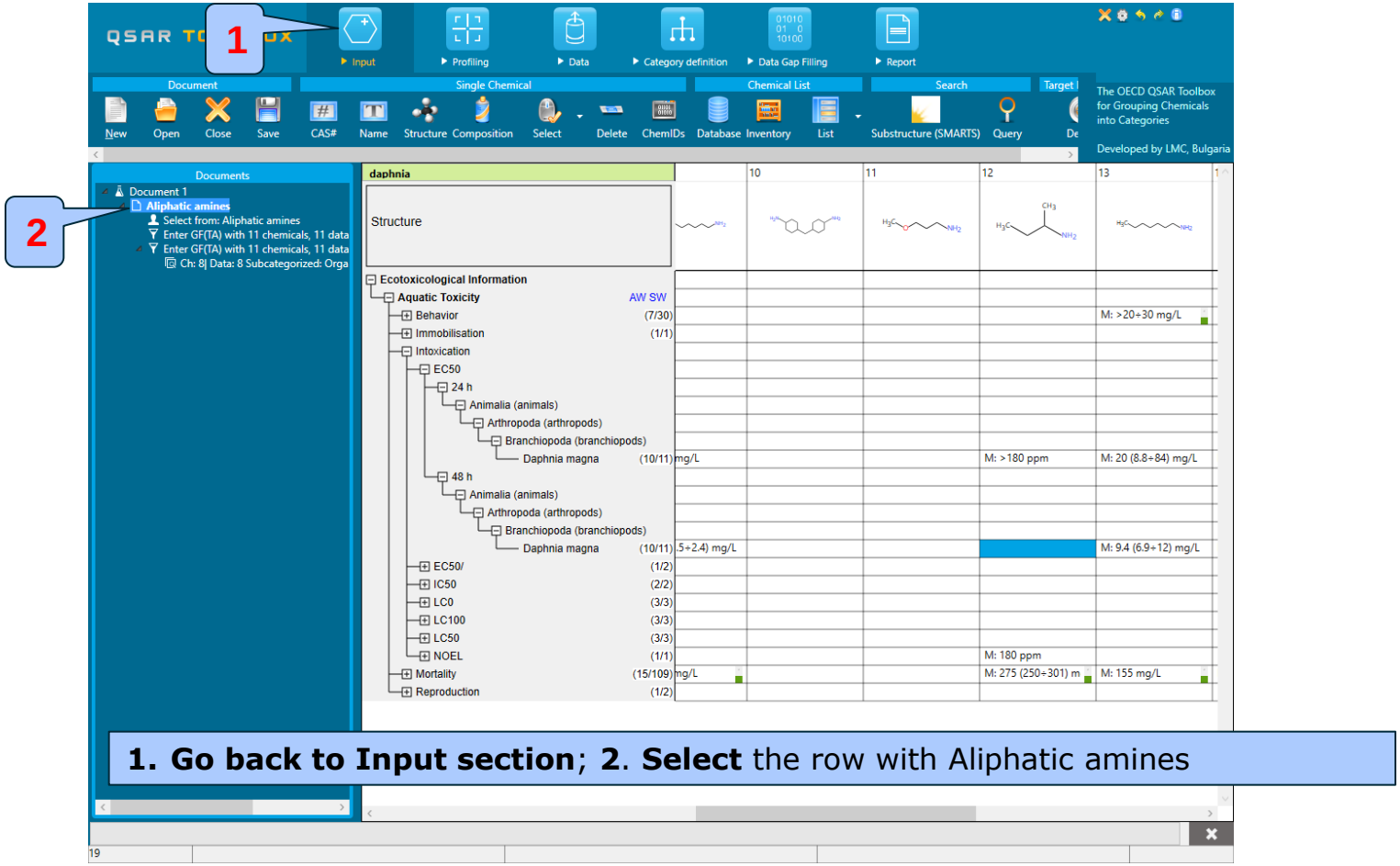
- Background
- Category evaluation – overview
  - Phase I. Evaluation of Category & Data Gap Filling
  - **Phase II. Extension of the Category& Data gap Filling**

Phase II: Extension Category & Data gap filling

- The extension of the category is performed by using Phase II of the category evaluation process (*Extension of Category & Data Gap Filling*). Other analogues are searched in the Toolbox, which are structurally and mechanistically consistent with the predefined category.
- The structural analogues could be defined using ECOSAR grouping

Phase II: Extension Category & Data gap filling

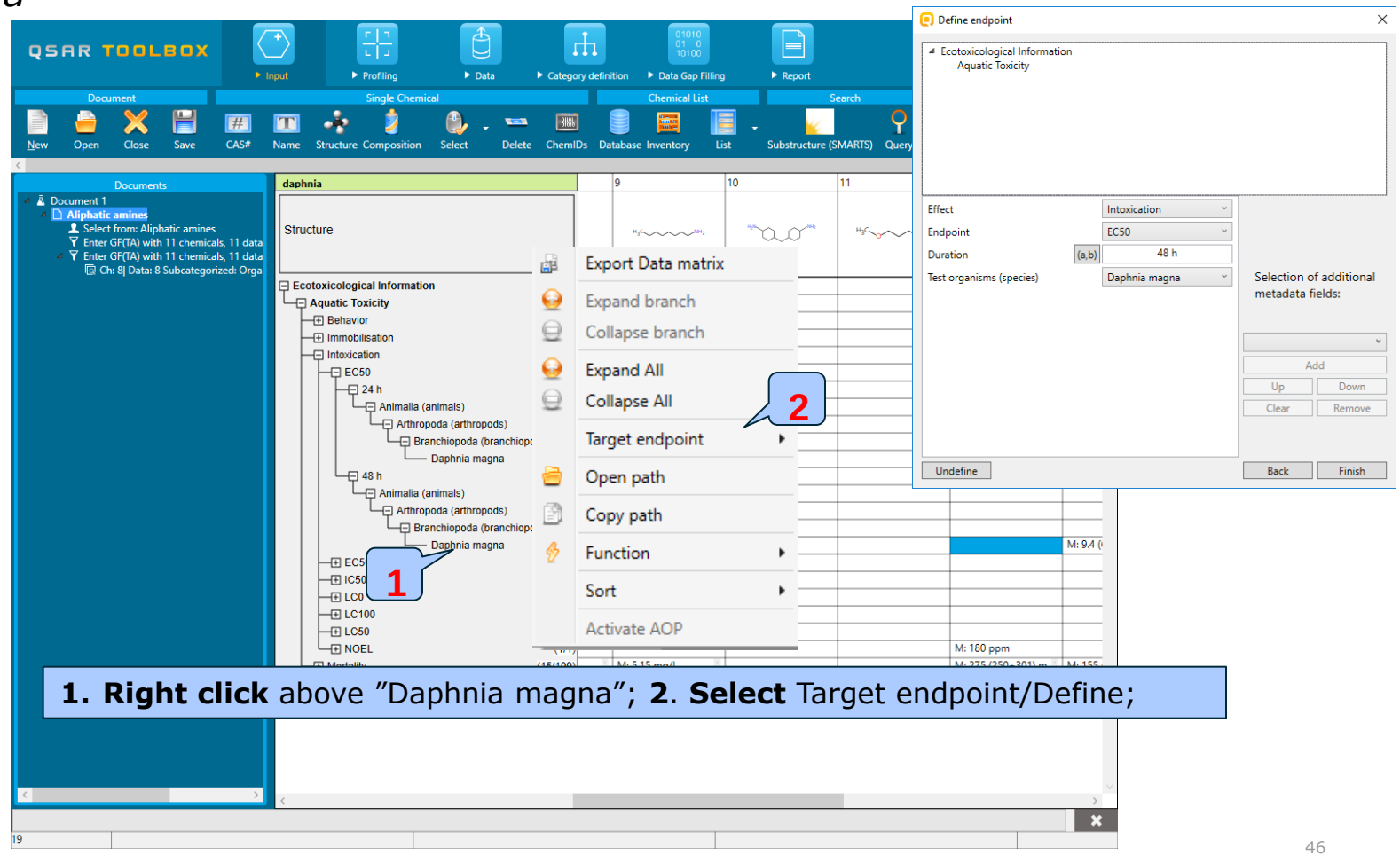
Step 1: Category definition



Phase II: Extension Category & Data gap filling

Step 1: Category definition

- Before defining the category, one should specify the endpoint - EC 50 48h *D.magna*

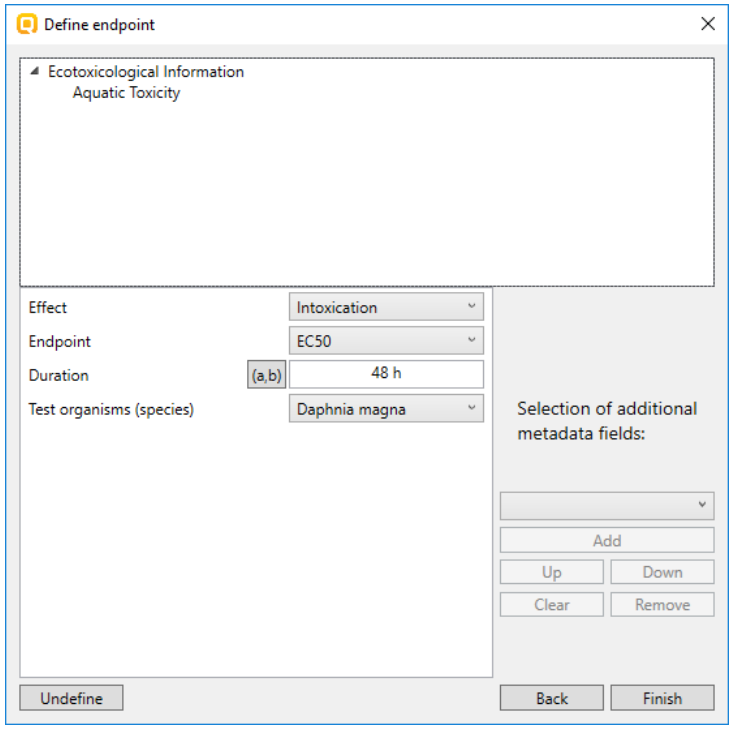


1. Right click above ”Daphnia magna”; **2. Select** Target endpoint/Define;

Phase II: Extension Category & Data gap filling

Step 1: Category definition

- Before defining the category, one should specify the endpoint - EC 50 48h *D.magna*

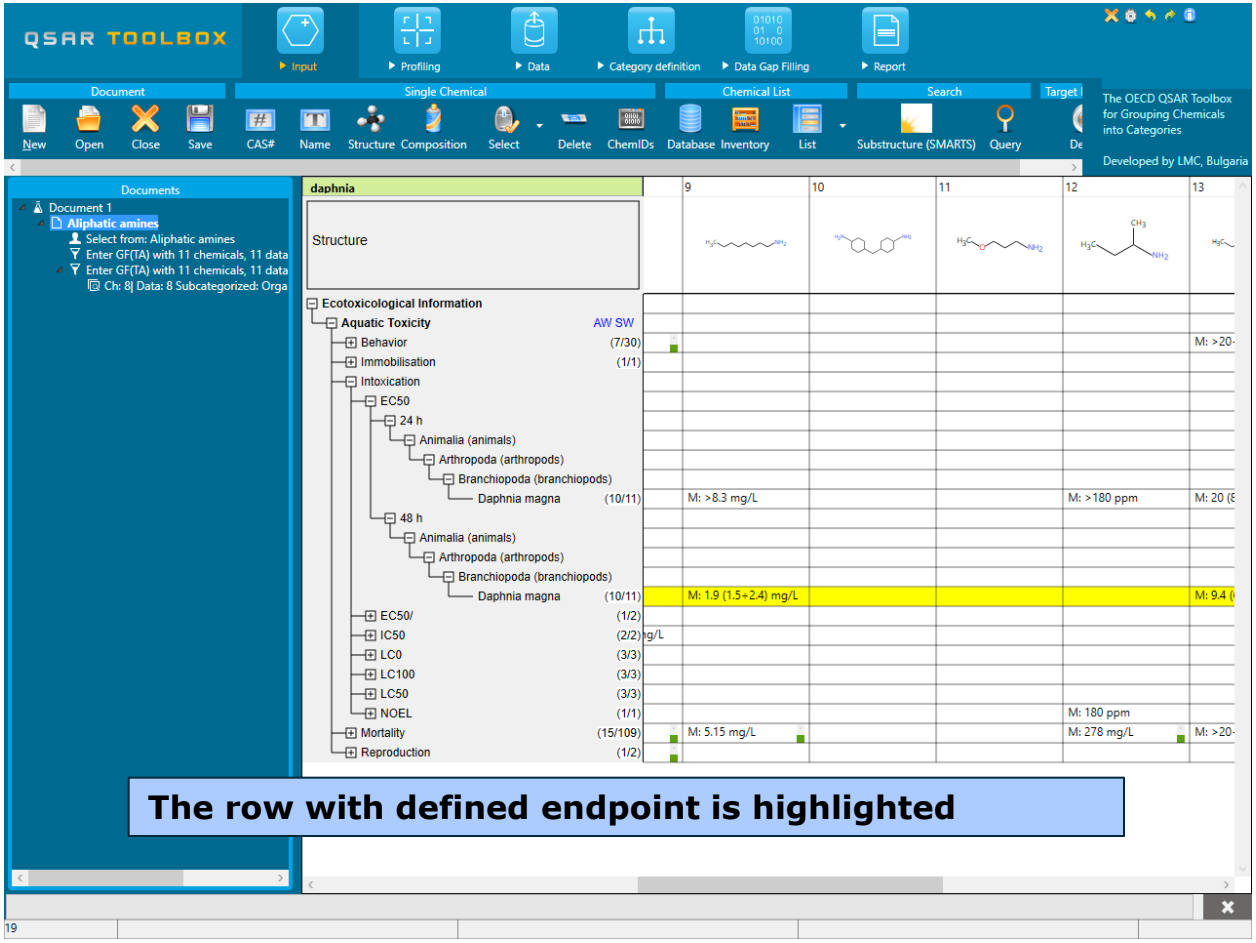


1. Confirm the endpoint by click "Finish"

Phase II: Extension Category & Data gap filling

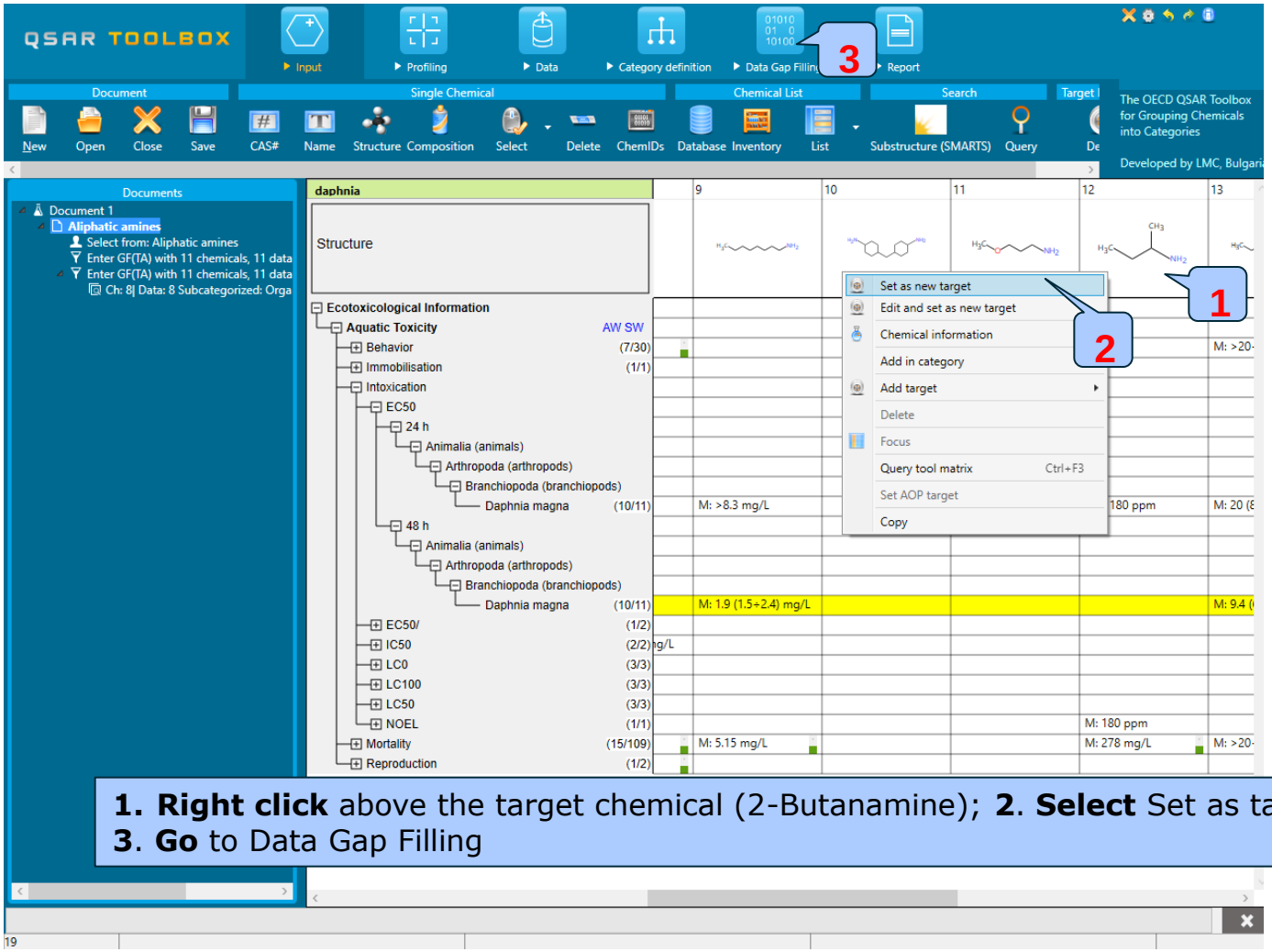
Step 1: Category definition

- Before defining the category, one should specify the endpoint - EC 50 48h *D.magna*



Phase II: Extension Category & Data gap filling

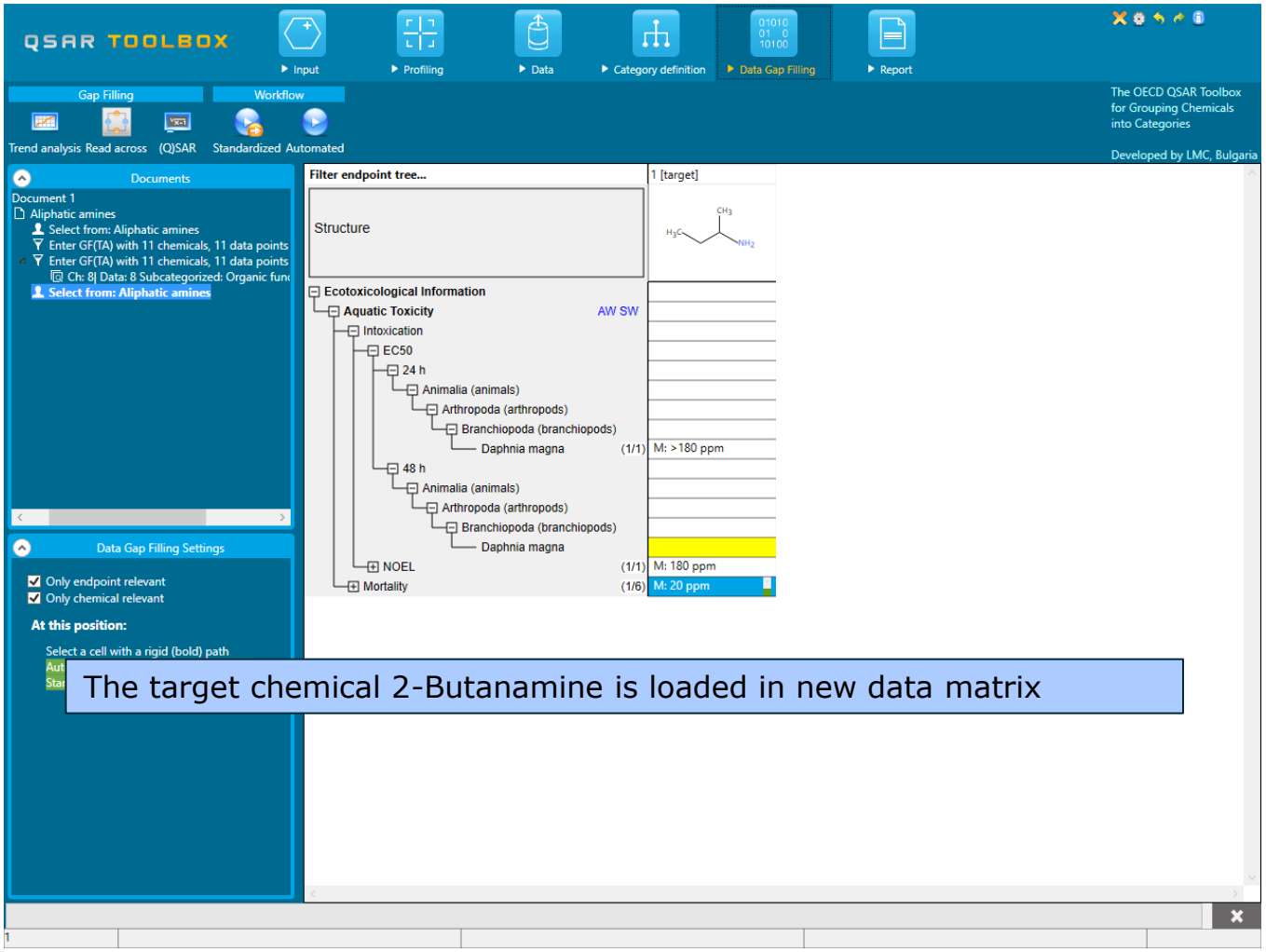
Step 1: Category definition



1. Right click above the target chemical (2-Butanamine); **2. Select** Set as target;
3. Go to Data Gap Filling

Phase II: Extension Category & Data gap filling

Step 1: Category definition



Phase II: Extension Category & Data gap filling

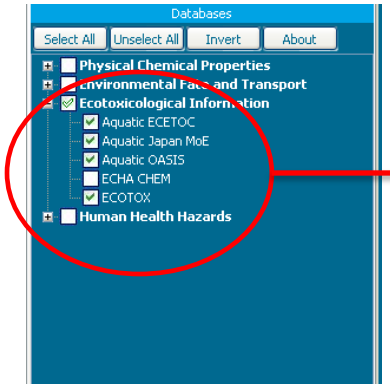
Step 1: Category definition

- The ECOSAR strict category is used to define an broader category used in further analysis.
- The same endpoint: EC 50 48h *D.magna* will be predicted as with the predefined category

Phase II: Extension Category & Data gap filling

Step 1: Category definition

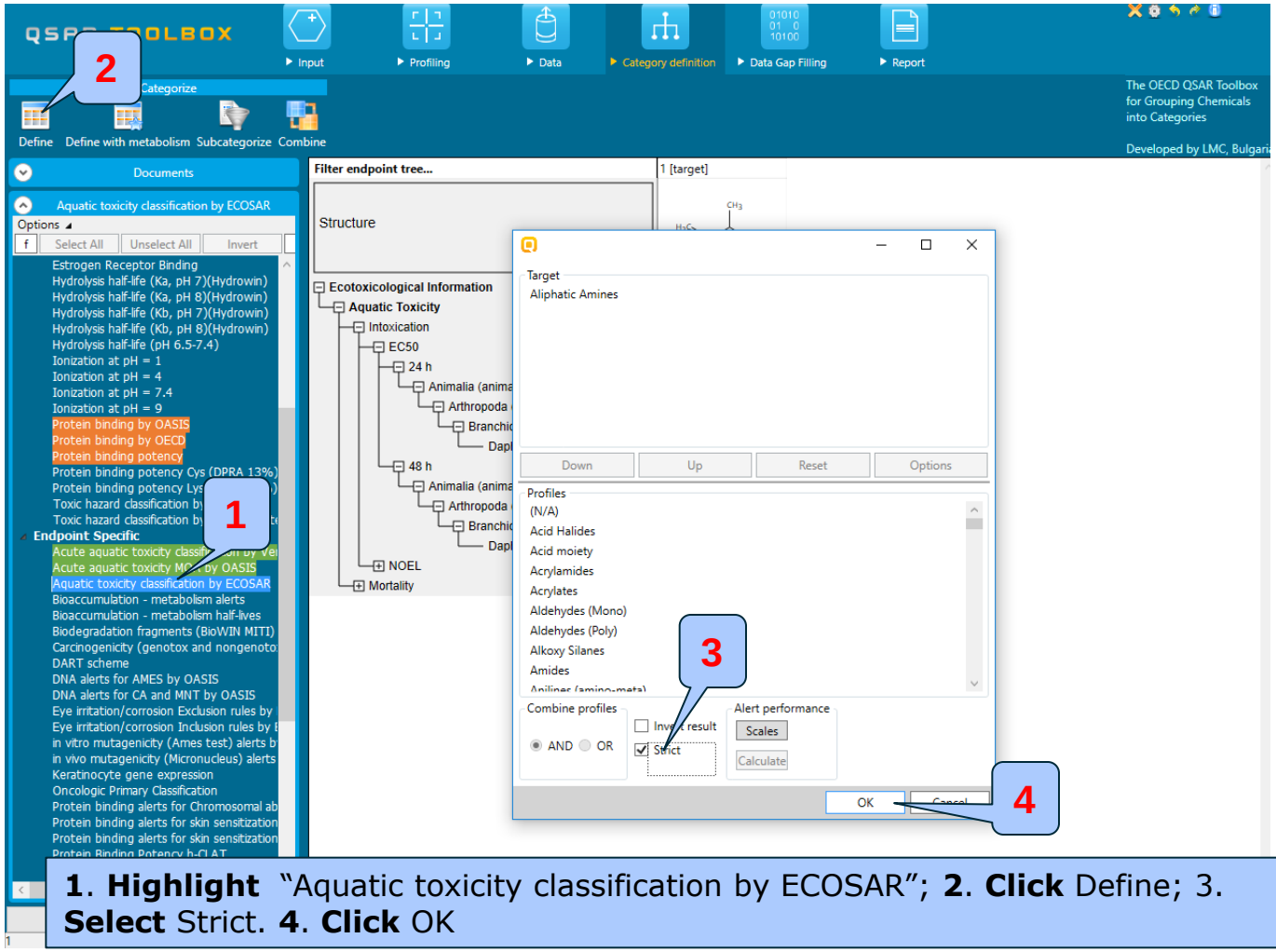
- Define ECOSAR category
- Before defining the category, the following databases related to the predicted endpoint are selected:



Aquatic ECETOC
Aquatic Japan MoE
Aquatic OASIS
ECOTOX

Phase II: Extension Category & Data gap filling

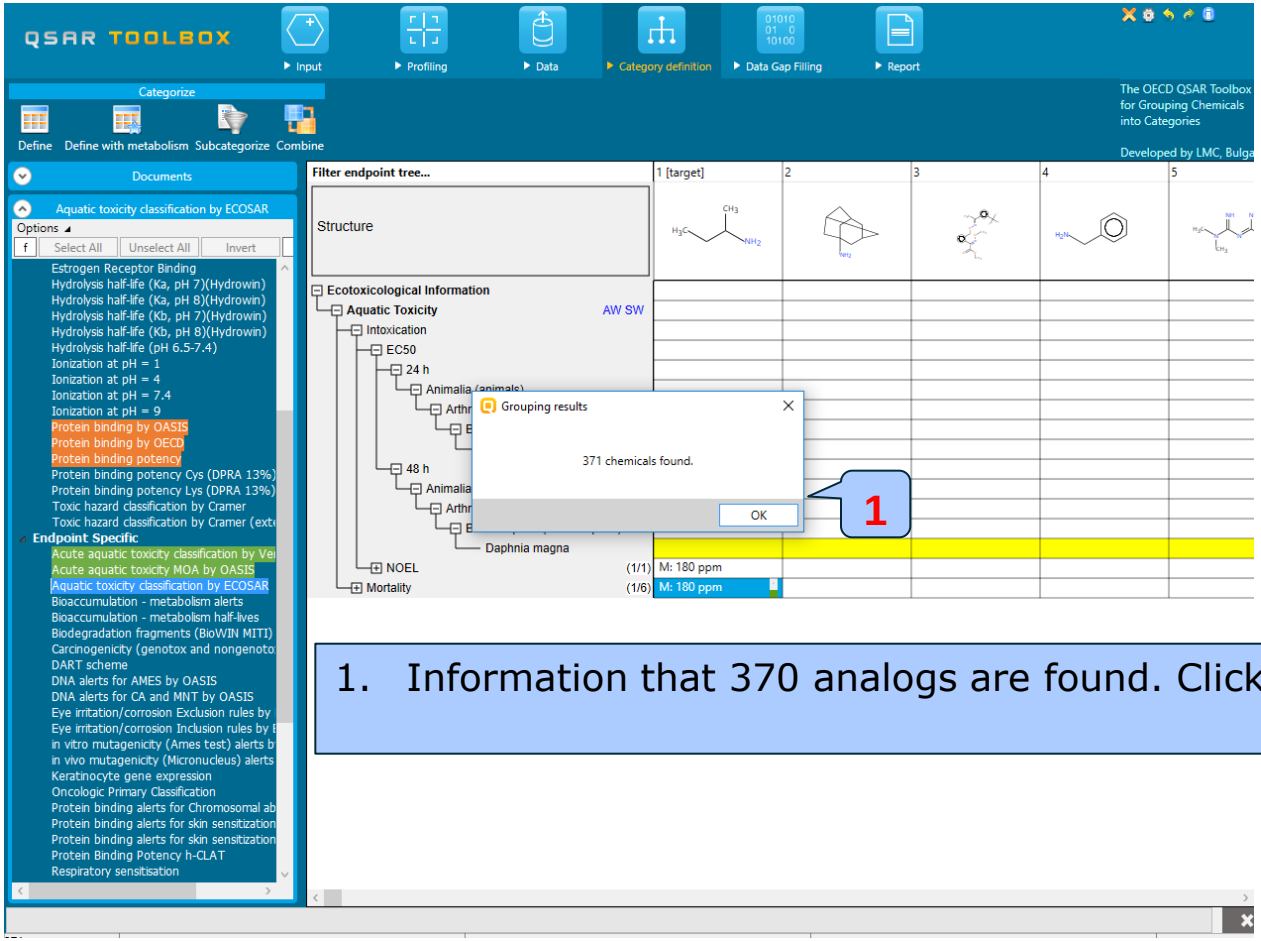
Step 1: Category definition/*Defining ECOSAR (strict)*



1. **Highlight** "Aquatic toxicity classification by ECOSAR"; **2**. **Click** Define; 3. **Select** Strict. **4**. **Click** OK

Phase II: Extension Category & Data gap filling

Step 1: Category definition/*Defining ECOSAR (strict)*

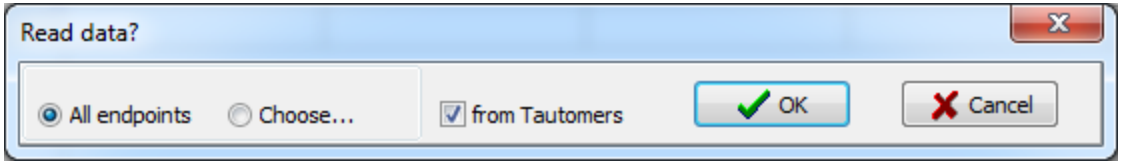


1. Information that 370 analogs are found. Click OK

Phase II: Extension Category & Data gap filling

Step 1: Category definition/*Reading data*

- The Toolbox will now retrieve those chemicals that have the same ECOSAR functionality as the target compound.
- 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).

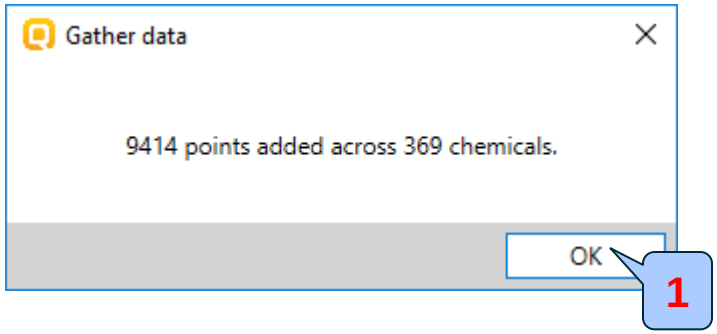


- In this example, as only databases are selected that contain information for aquatic toxicity endpoint, both options give the same results.

Phase II: Extension Category & Data gap filling

Step 1: Category definition/*Reading data*

The system automatically gives indication for the number of gather experimental data points



1. **Click** OK

Outlook

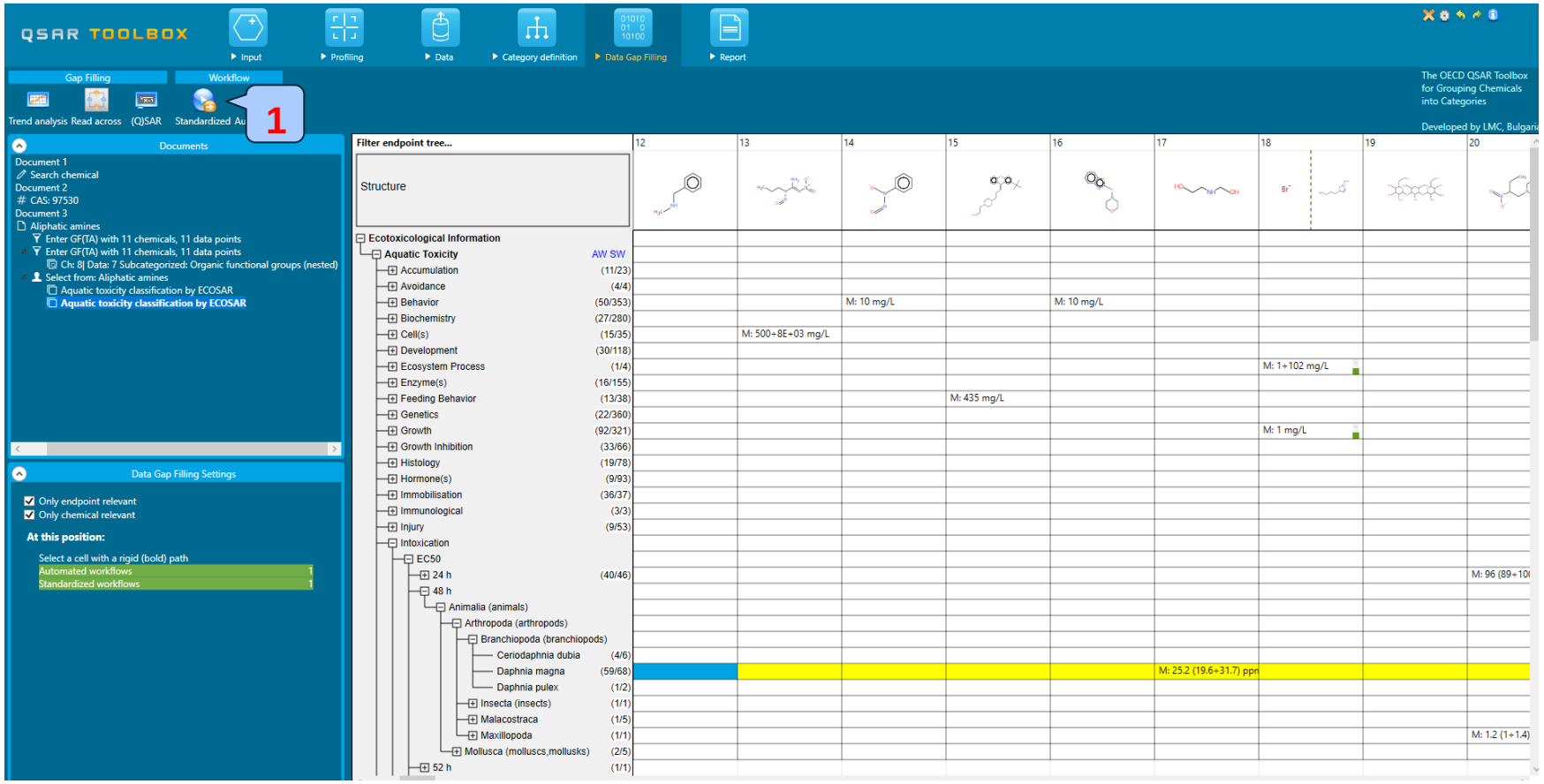
- Background
- Category evaluation – overview
  - Phase I. Evaluation of Category & Data Gap Filling
  - **Phase II. Extension of the Category& Data gap Filling**
    - Step 1: Category definition
    - Step 2: Navigate to the target endpoint – this is already done, so it is skipped

Outlook

- Background
- Category evaluation – overview
  - Phase I. Evaluation of Category & Data Gap Filling
  - **Phase II. Extension of the Category& Data gap Filling**
    - Step 1: Category definition
    - Step 2: Navigate to the target endpoint
    - Step 3: Data Gap Filling

Phase II: Extension Category & Data gap filling

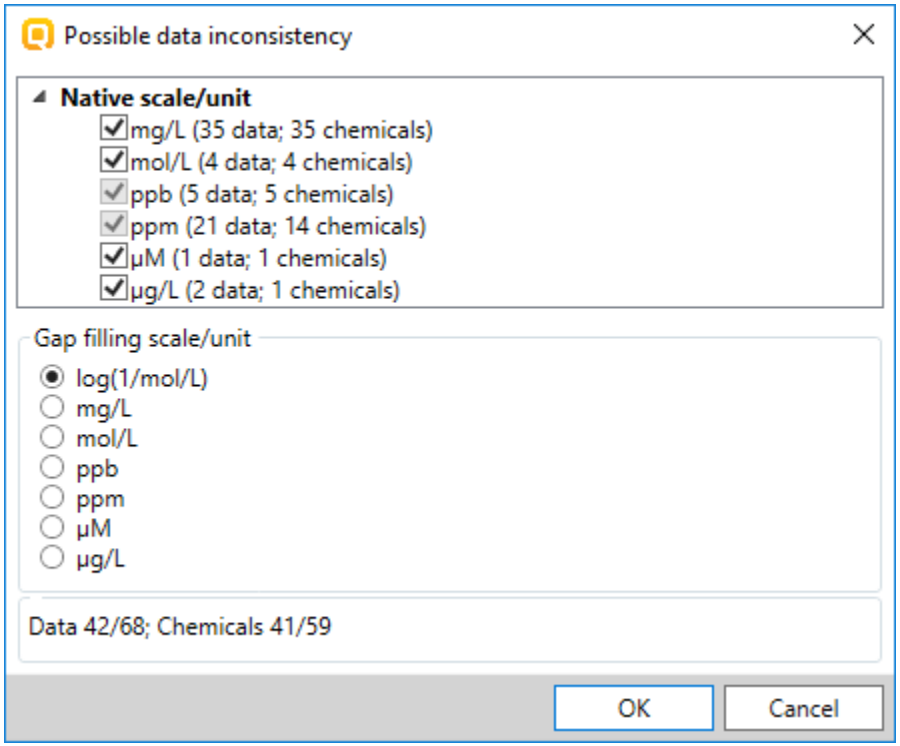
Step 3: Data Gap Filling



1. **Select** Trend analysis

Phase II: Extension Category & Data gap filling

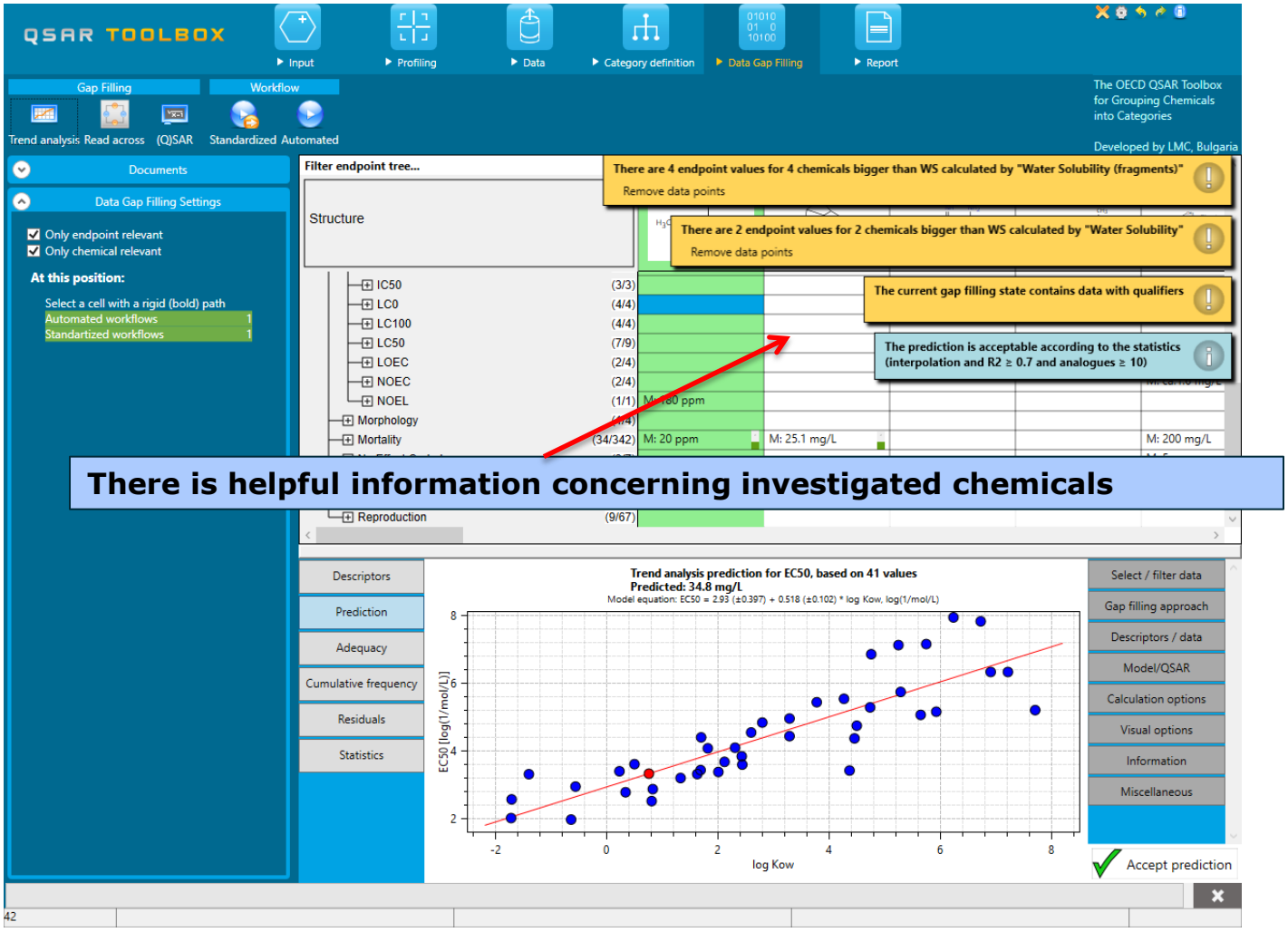
Step 3: Data Gap Filling



The user will be informed If there is different experimental data. Click OK

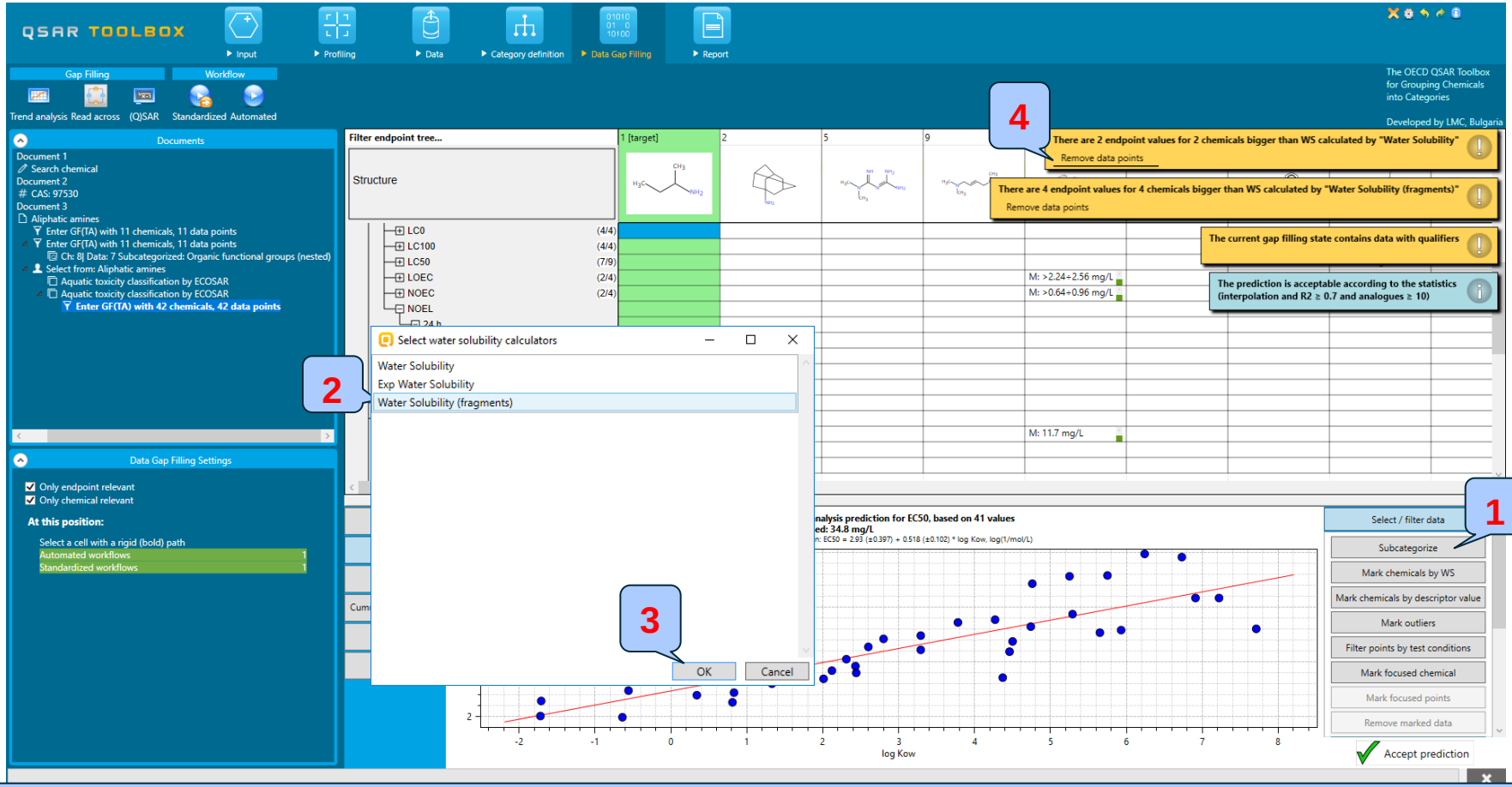
Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling



Phase II: Extension Category & Data gap filling

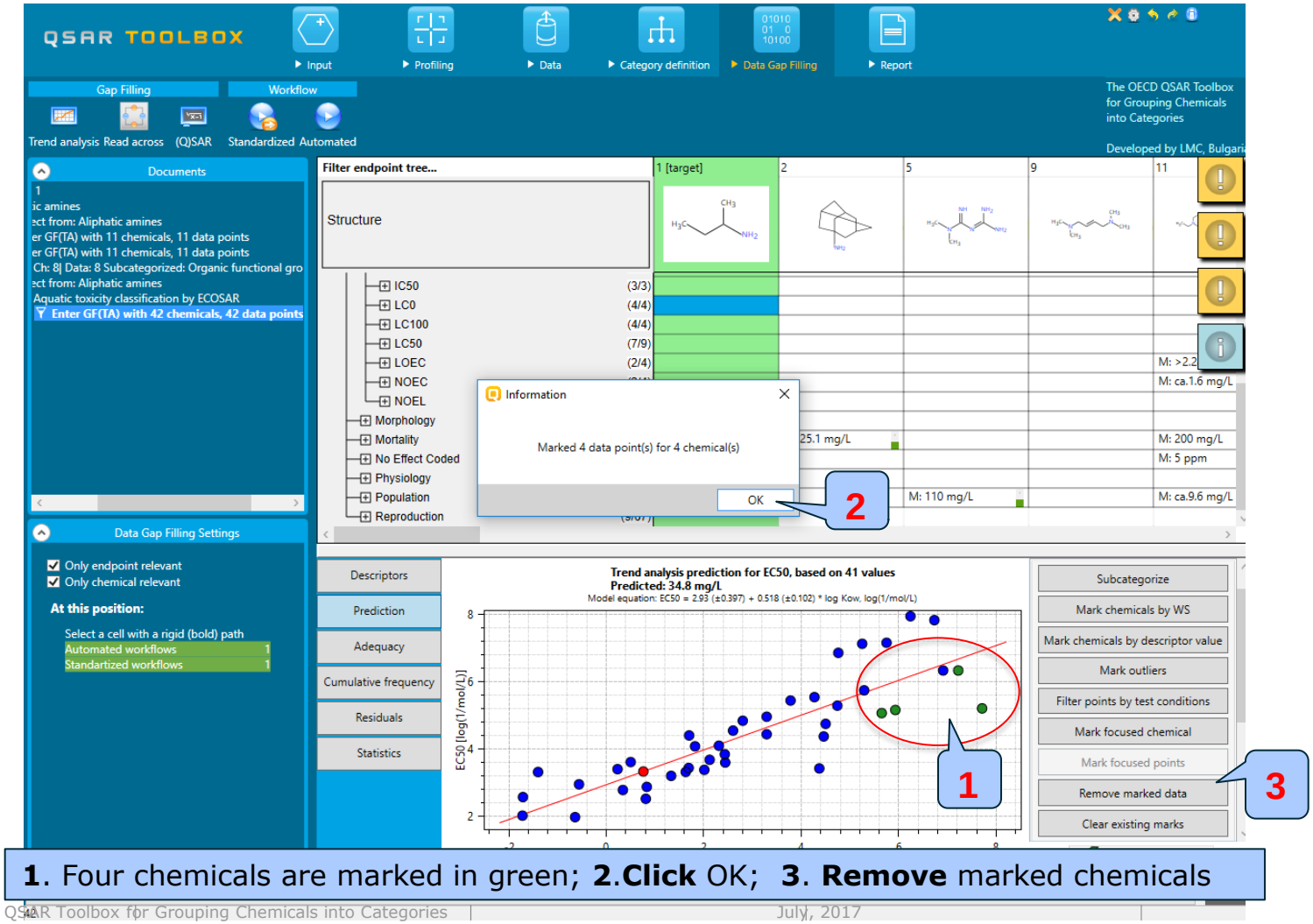
Step 3: Data Gap Filling/*Subcategorize by Water solubility*



1. Click Mark chemical by WS; **2. Select** Water solubility (fragments);
3. Click OK. 4. The same steps will be reproduced if you click on the link available in the notifying helper

Phase II: Extension Category & Data gap filling

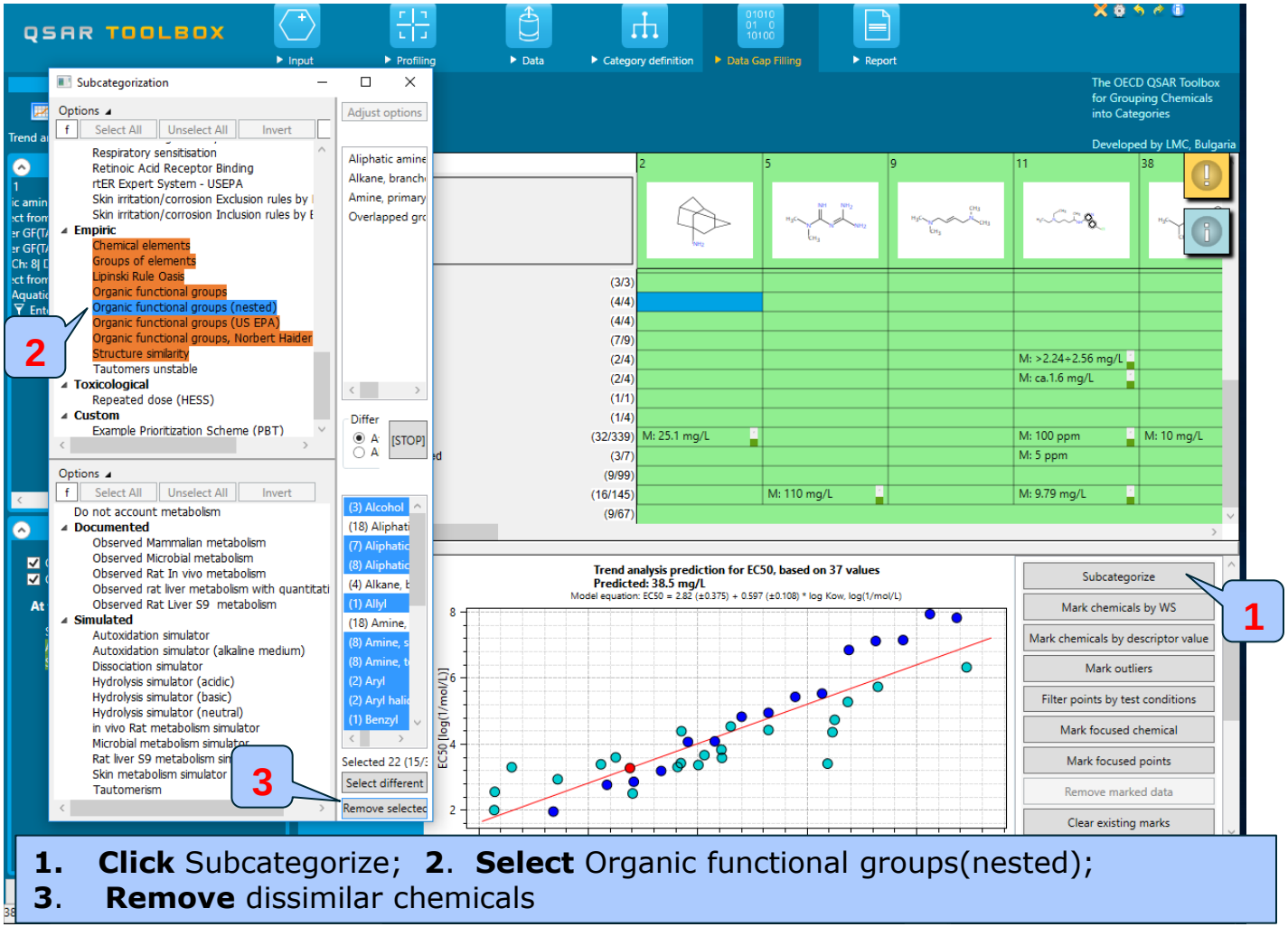
Step 3: Data Gap Filling/*Subcategorize by Water solubility*



1. Four chemicals are marked in green; **2.Click** OK; **3**. **Remove** marked chemicals

Phase II: Extension Category & Data gap filling

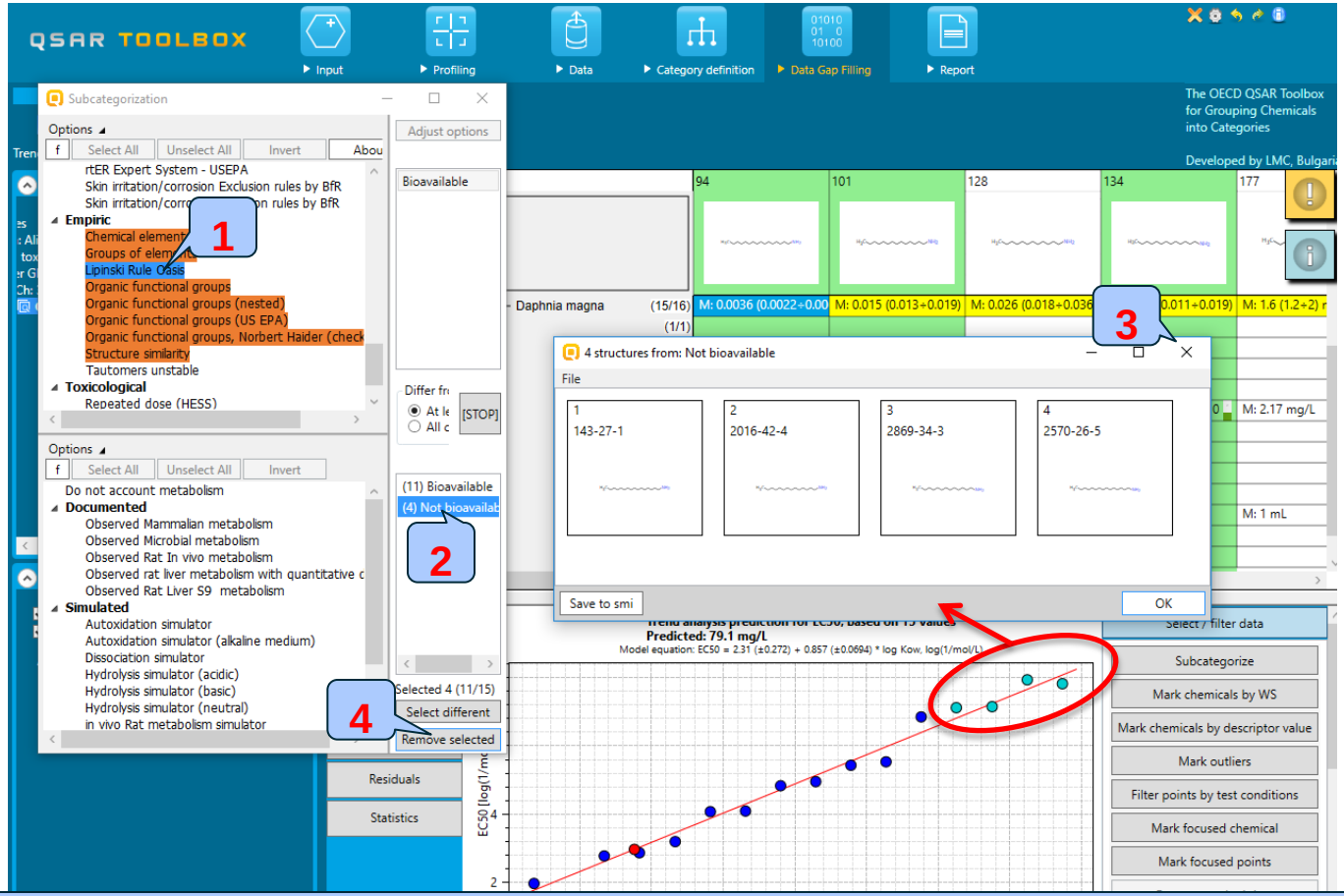
Step 3: Data Gap Filling/*Subcategorize by OFG(nested)*



1. **Click** Subcategorize; **2.** **Select** Organic functional groups(nested);
3. **Remove** dissimilar chemicals

Phase II: Extension Category & Data gap filling

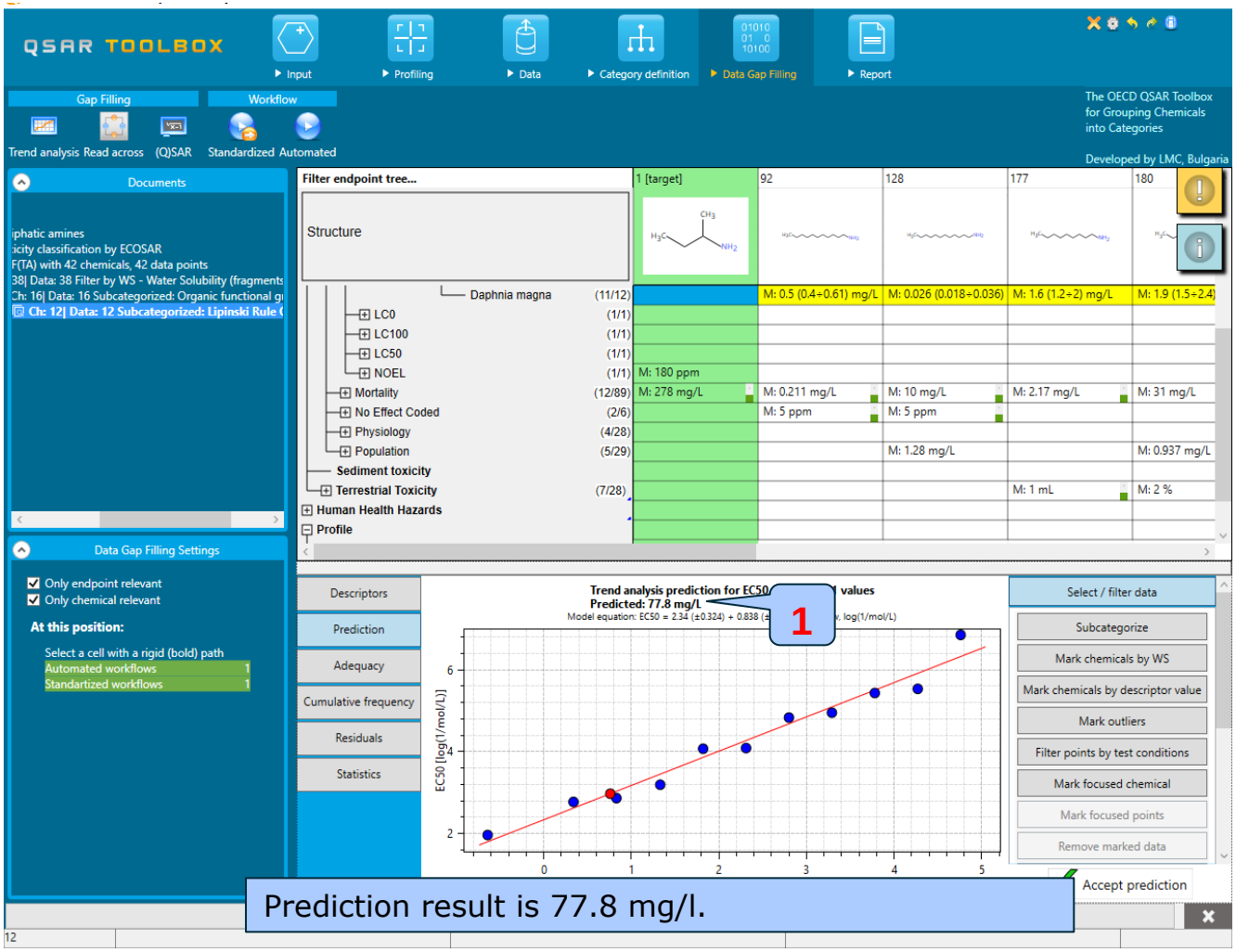
Step 3: Data Gap Filling/*Subcategorize by Lipinski rules*



Chemicals with very long chain could be removed from the category due to their non-bioavailability. **1**. **Select** Lipinski rules; **2**. **Double click** to see "Not- bioavailable" chemicals; **3**. **Close** the appeared window; **4**. **Remove** dissimilar chemicals

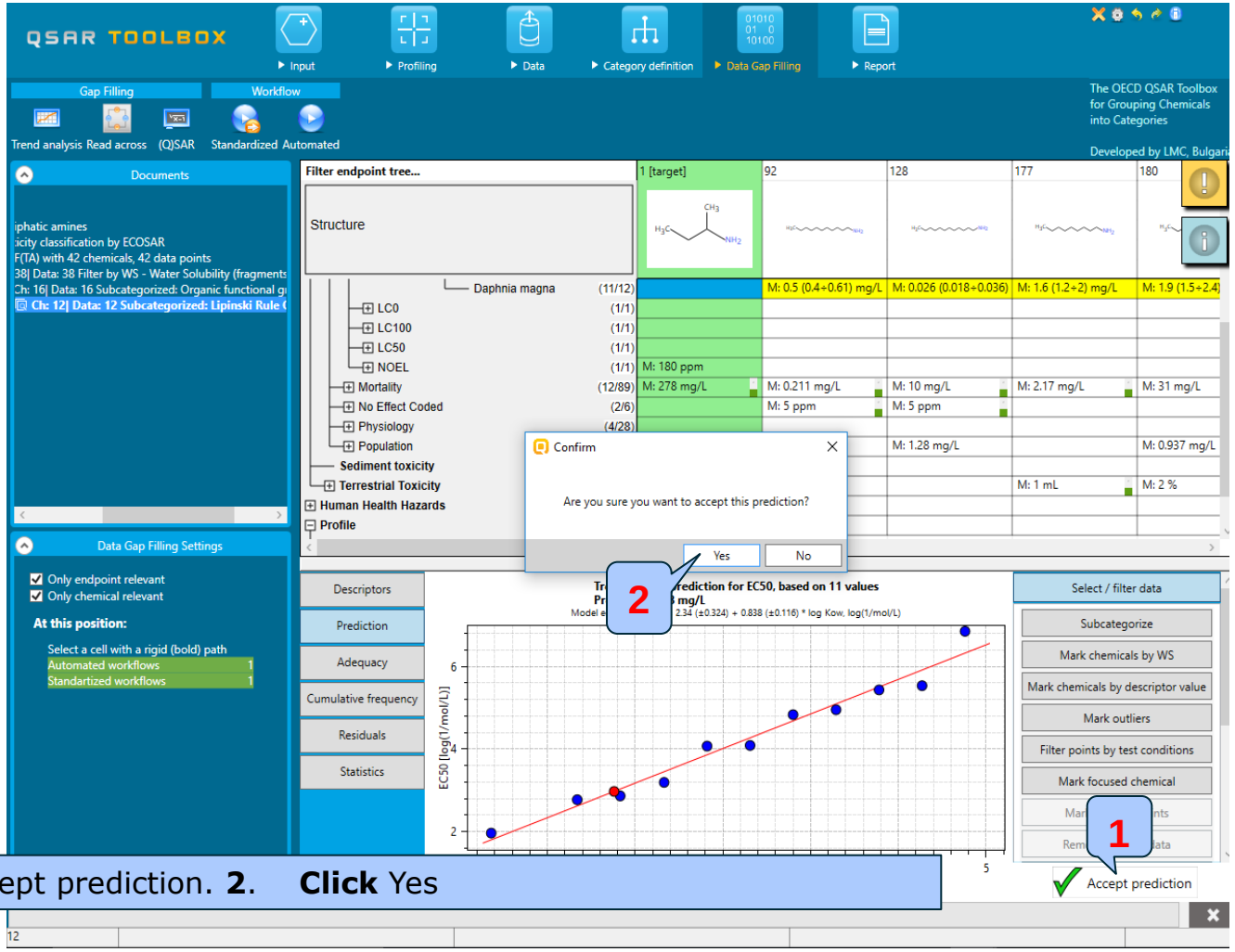
Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/*Prediction result*



Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/*Accept the Prediction result*



1. **Click** on Accept prediction. **2.** **Click** Yes

Outlook

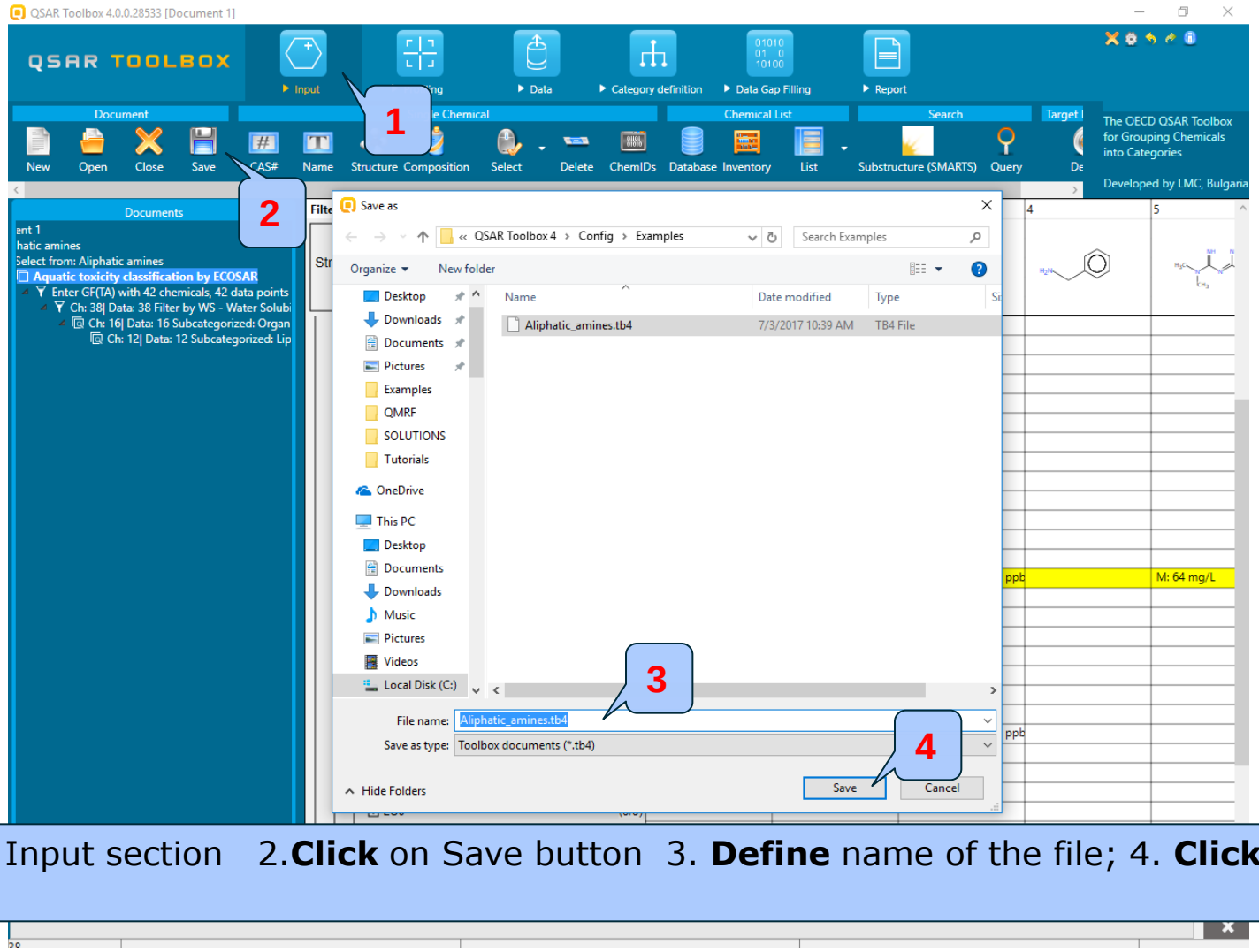
- Background
- Category evaluation – overview
  - Phase I. Evaluation of Category & Data Gap Filling
  - Phase II. Extension of the Category& Data gap Filling
- Save the prediction result

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).

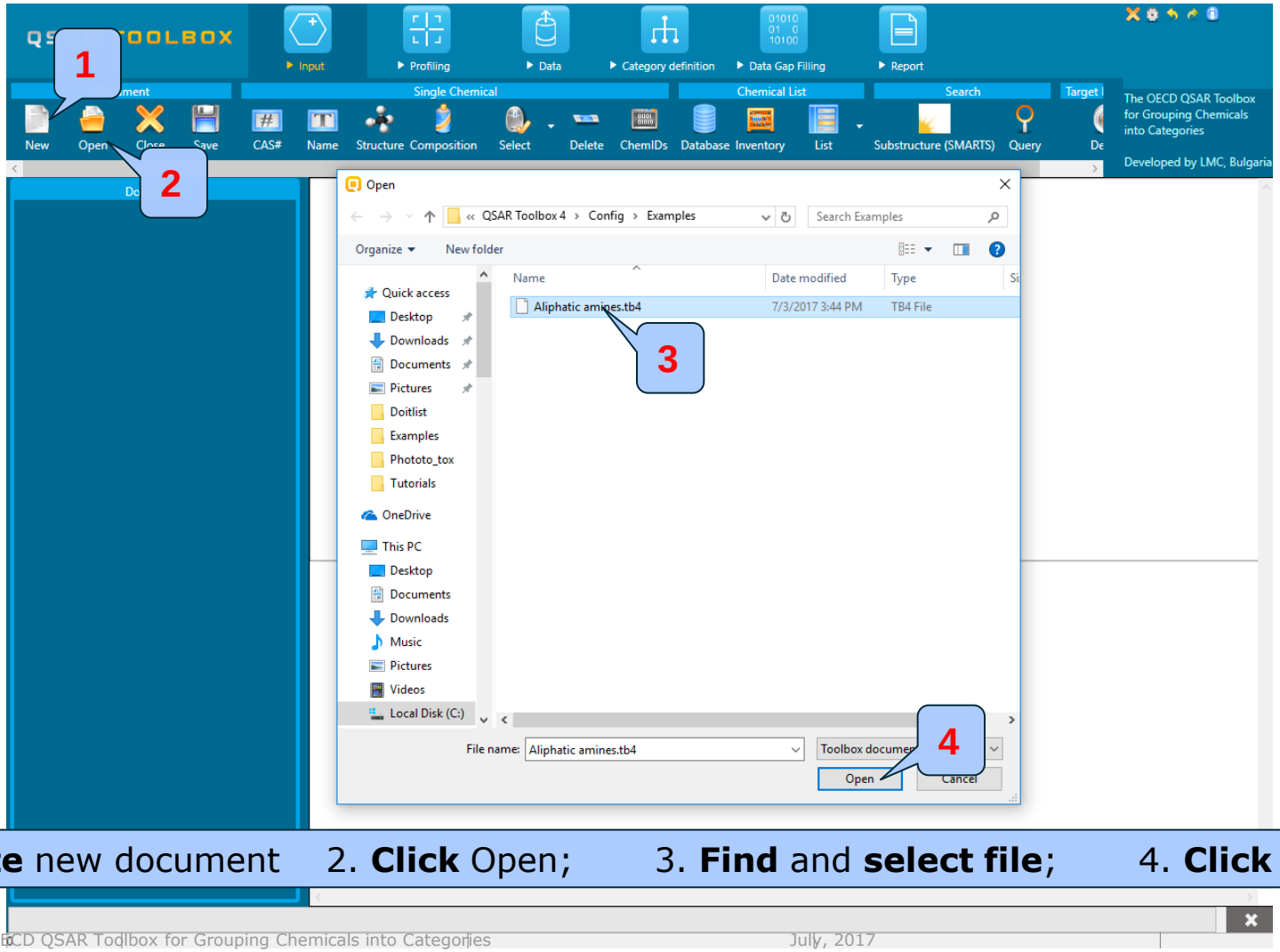
- Saving/Loading the file with TB prediction is shown on next screenshots

Saving the prediction result



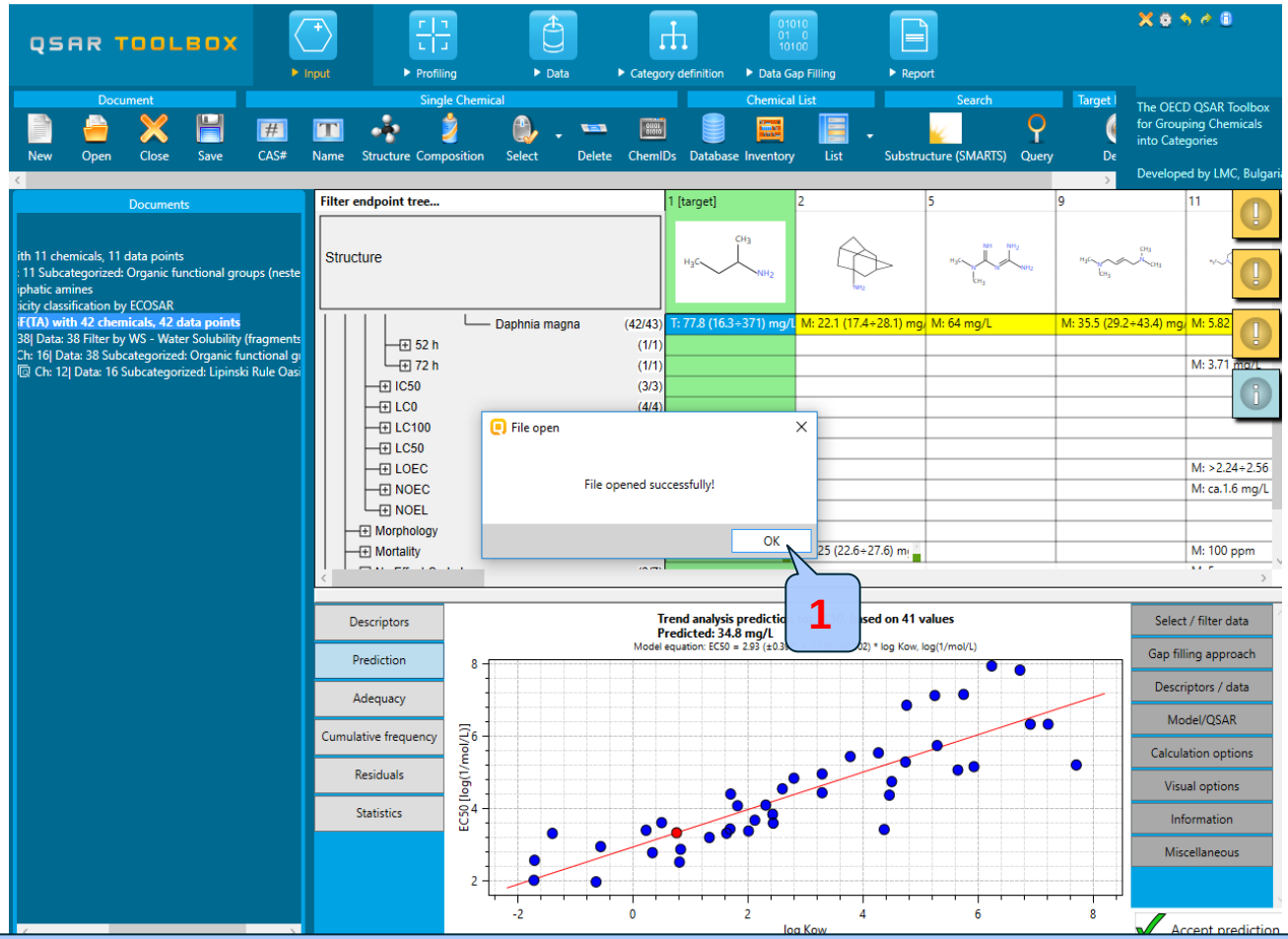
1. Go to Input section 2.**Click** on Save button 3. **Define** name of the file; 4. **Click** Save button

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