QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

OECD (Q)SAR Toolbox v.4.4.1

Tutorial on using the PBT prioritization scheme

Outlook

- **Aim**
- Keywords
- PBT scheme
- Workflow of the prediction
- Export of the results

Aim

This is a step-by-step presentation designed to take the user of Toolbox through the PBT prioritization scheme implemented in the software.

Outlook

- **Aim**
- Keywords
- PBT scheme
- Workflow of the prediction
- Export of the results

Keywords

TARGET CHEMICAL - chemical of interest

MODULE – a Toolbox module is a section dedicated to specific actions and options (e.g. Profiling)

WORKFLOW – the use, in combination, of the different modules (e.g. prediction workflow: from input to report)

PROFILER - algorithm (rule set) for the identification of specific features of the chemicals. Several types of profilers are available, such as structural (e.g. Organic functional groups), mechanistic (e.g. Protein binding by OECD) and endpoint-specific (e.g. in vitro in vitro mutagenicity (Ames test) alerts by ISS) profilers.

ALERT - the profilers consist of sets of rules or alerts. Each of the rules consists of a set of queries. The queries could be related to the chemical structure, physicochemical properties, experimental data, comparison with the target or list with substances and external queries from other predefined profilers (reference queries).

CATEGORY – "group" of substances sharing same characteristics (e.g. the same functional groups or mode of action). In a typical Toolbox workflow, it consists of the target chemical and its analogues gathered according to the selected profilers

ENDPOINT TREE – Endpoints are structured in a branched scheme, from a broader level (Phys-Chem properties, Environmental Fate and transport, Ecotoxicology, Human health hazard) to a more detailed one (e.g. EC3 in LLNA test under Human health hazard-Skin sensitization)

EXPORT – Toolbox allows to export data and predictions for chemicals in an Excel format

DATA MATRIX – Table reporting the chemical(s) and data (experimental results, profilers outcomes, predictions). Each chemical is in a different column and each data in a different row

PBT SCHEME – PBT prioritization scheme based on experimental data and (Q)SAR models for assessing persistency, bioaccumulation and toxicity of substances

Outlook

- Aim
- Keywords
- **PBT scheme**
  - **Background**
- Workflow of the prediction
- Export of the results

Background: PBT prioritization scheme

• The PBT prioritization scheme is based on experimental data and (Q)SAR models.

• A substance is classified based on thresholds for persistence, bioaccumulation and toxicity, and the following classifications could be assigned:

- very persistent (**vP**),
- persistent (**P**),
- very bioaccumulative (**vB**),
- bioaccumulative (**B**),
- very toxic (**vT**),
- toxic (**T**).

• The thresholds are illustrated in the table on the next slide.

Background: PBT assessment criteria

Classification	Criteria	Guidance	Data and models
vP	BOD ≤ 30%	REACH Annex XIII [1]	Experimental data only
P	30% < BOD ≤ 40% or biodegradation probability < 0.5	REACH Annex XIII [1]	Experimental data and BIOWIN 5 and 6 models
vB	BCF ≥ 3.699 log(L/kg wet)	REACH Annex XIII [1]	Experimental data and BCFWIN model
B	3.301 ≤ BCF < 3.699 log(L/kg wet) or log Kow > 4.5	REACH Annex XIII [1]	Experimental data and BCFWIN and KOWWIN models
vT	LC50 ≤ 1 mg/l	GHS classification [2]	Experimental data only
T	1 mg/l < LC50 ≤ 10 mg/l	GHS classification [2]	Experimental data only

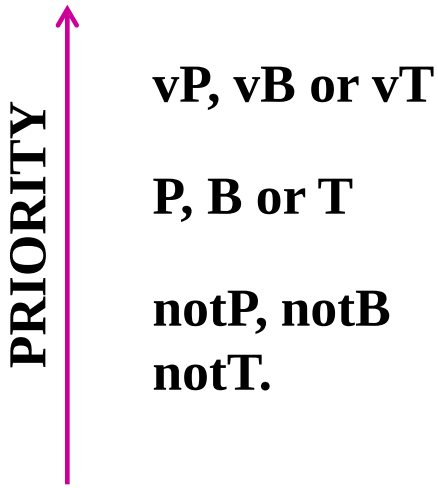
[1] https://echa.europa.eu/documents/10162/13632/information_requirements_r11_en.pdf
[2] https://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev04/English/ST-SG-AC10-30-Rev4e.pdf

Background: PBT assessment criteria

PBT priority grouping:

Based on the collected experimental data and results provided by (Q)SAR models, the system applies the worst case scenario to categorize chemicals using the following priority:

- With the **highest priority** are chemicals **with available experimental data** for P, B or T assessment.
- In case of unavailable experimental data for P, B or T, (Q)SAR models are applied for making predictions.

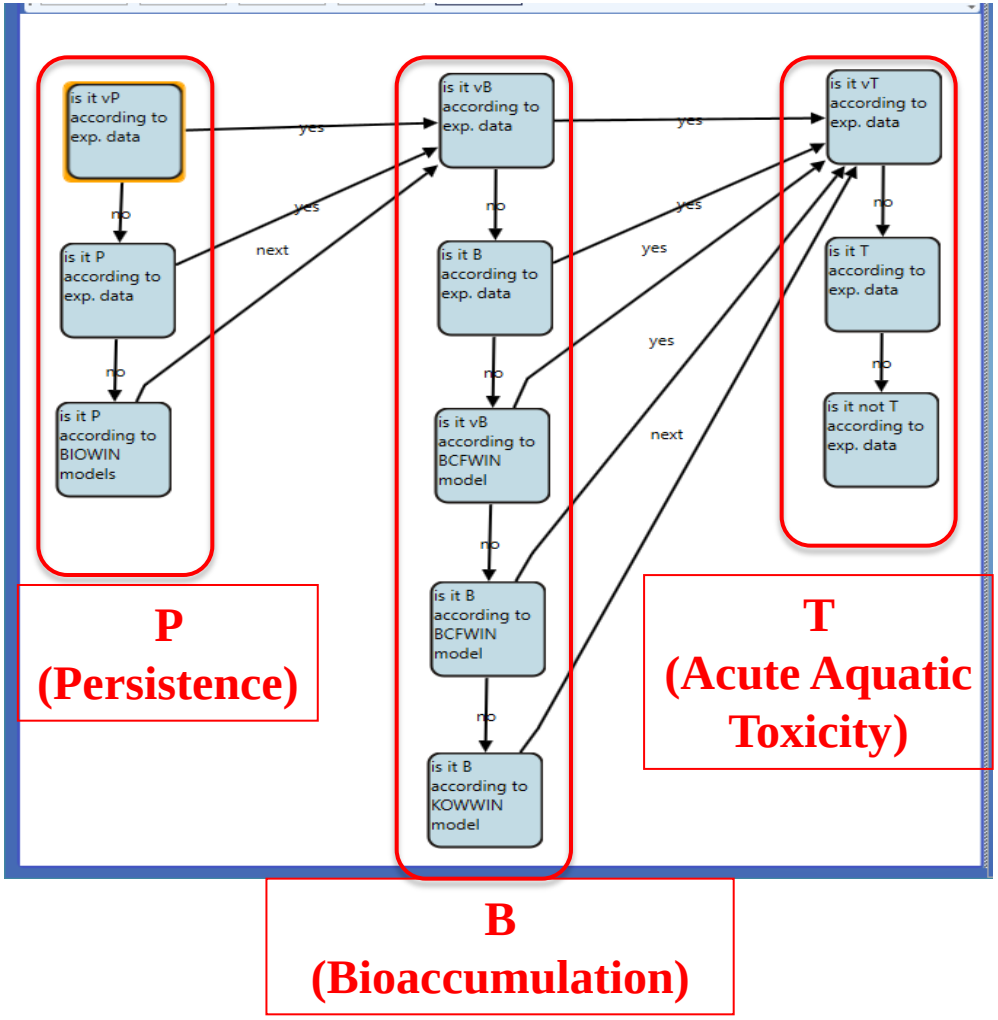


Outlook

- Aim
- Keywords
- **PBT scheme**
  - Background
  - **Implementation**
- Workflow of the prediction
- Export of the results

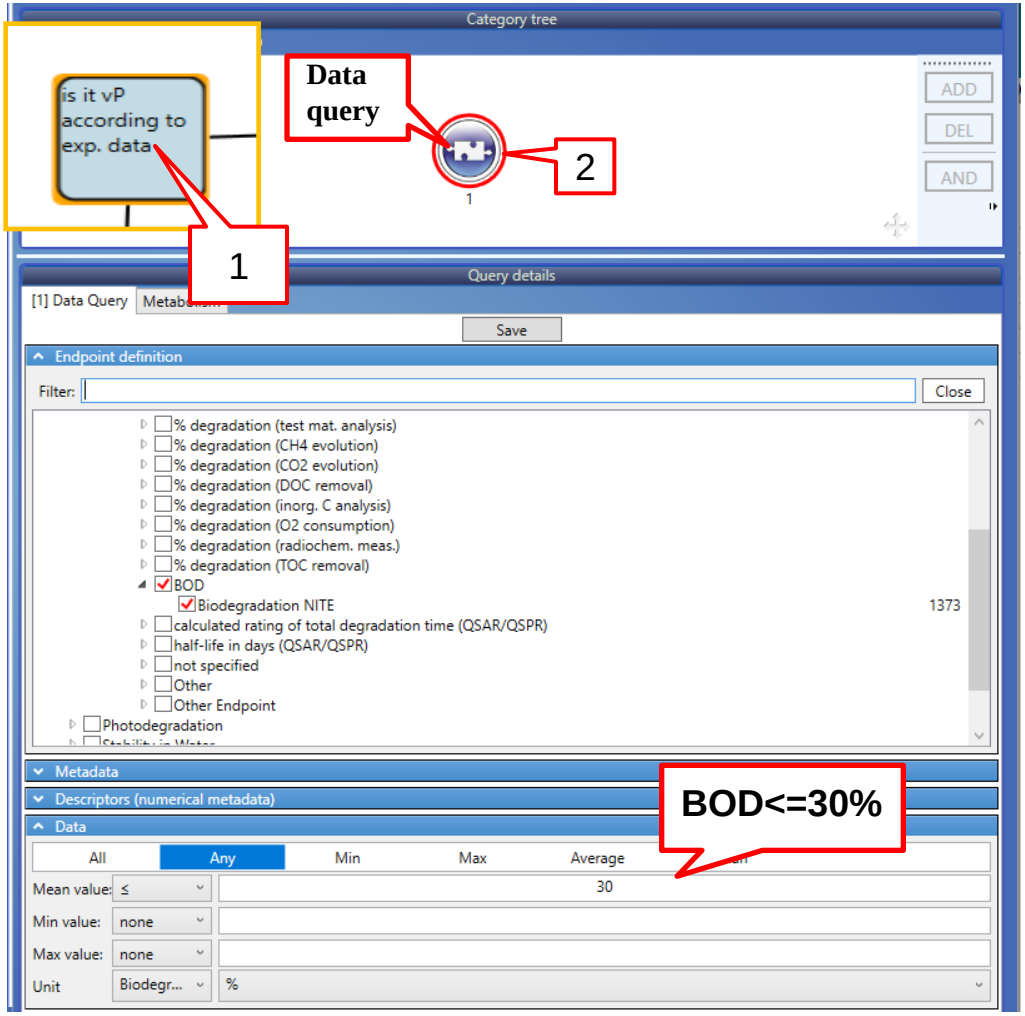
Implementation of PBT prioritization scheme

- The PBT prioritization scheme is a graph of logically connected nodes.
- The nodes are divided into three groups based on the P, B or T classification for hazard assessment.



Implementation of PBT prioritization scheme

- Each node (1) contains a data or parametric (model) boundary (2) where the criteria for PBT assessment are set.
- Data query is depicted here.



Outlook

- Aim
- Keywords
- PBT scheme
  - Background
  - Implementation
- **Workflow of the prediction**
- Export of the results

Workflow of the prediction: Steps

- Input of target chemical(s)
- Profiling: Example prioritization scheme (PBT)
- Report

Outlook

- Aim
- Keywords
- PBT scheme
  - Background
  - Implementation
- **Workflow of the prediction**
  - **Chemical input**
- Export of the results

Chemical Input

- 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.

Chemical Input
Ways of Entering a Chemical

I. Single target chemical:

- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- SMILES (simplified molecular information line entry system) notation
- Chemical with defined composition
- Drawing chemical structure
- Select from User List/Inventory/Databases

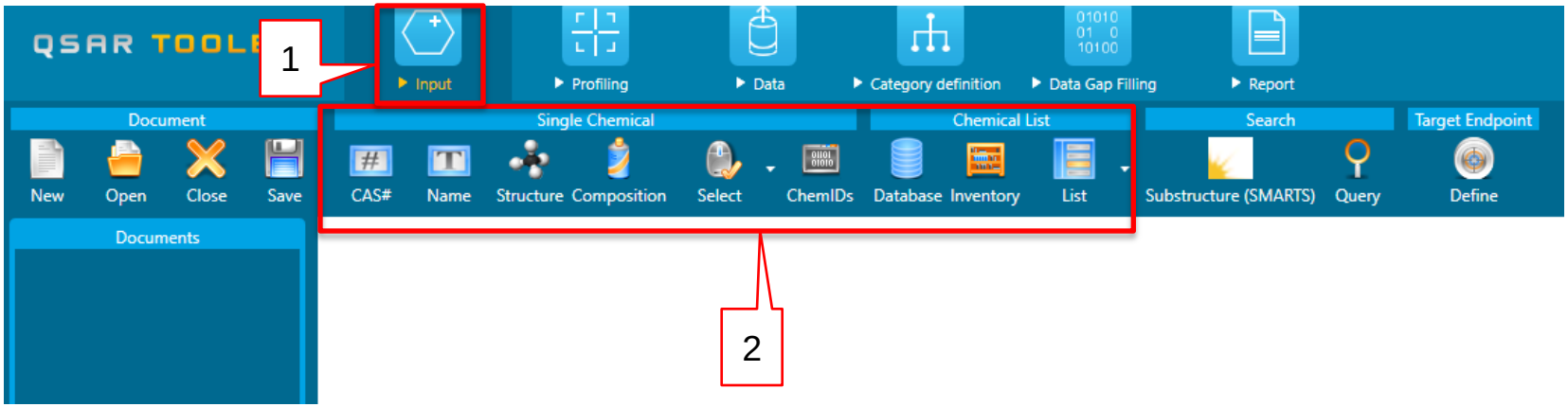
II. Group of chemicals:

- User's List
- Inventory/Database

Chemical Input: Single chemical

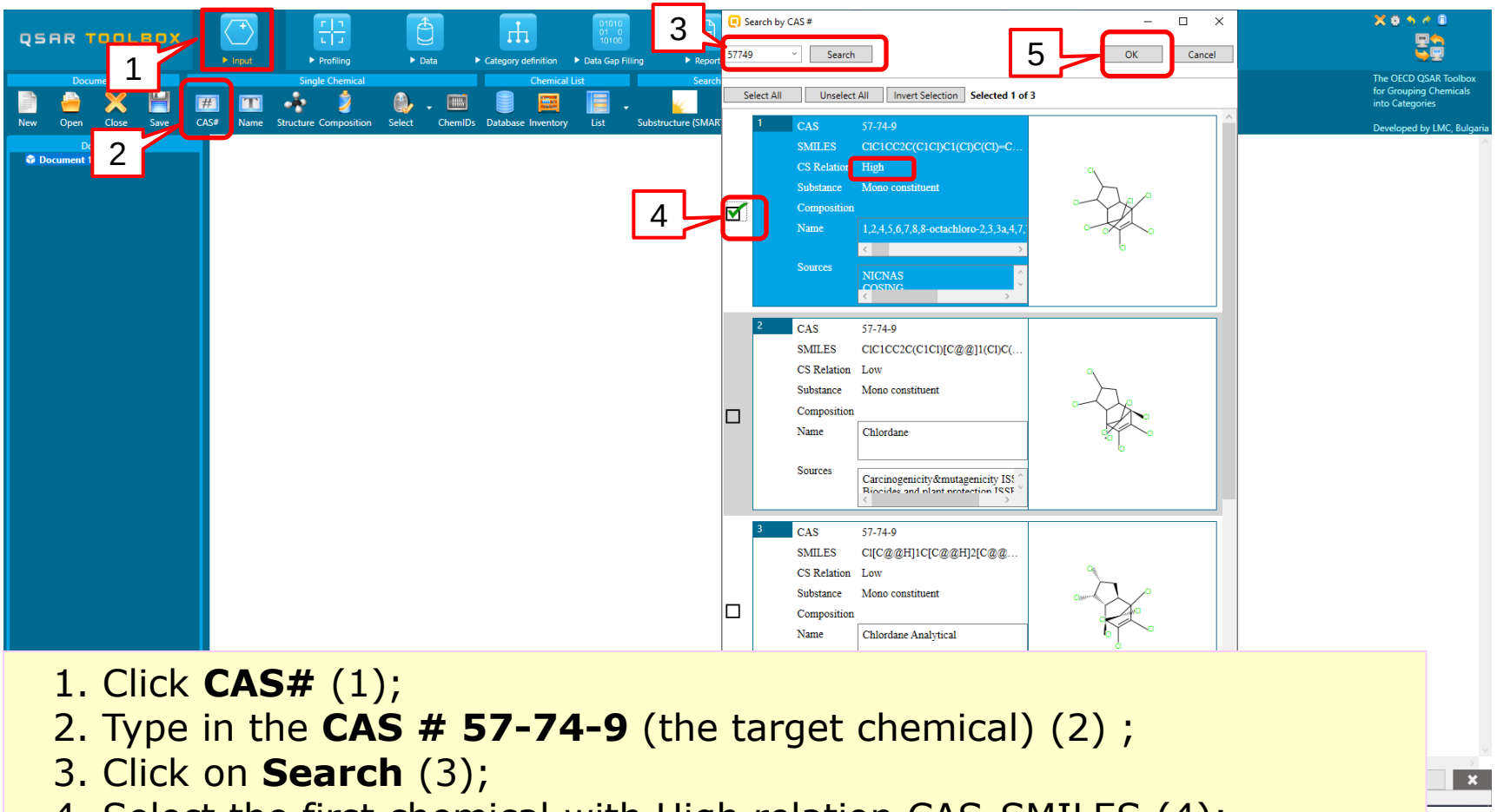
- Open the Toolbox.
- Click on “Input” (see next screen shot).

Chemical Input
Single chemical



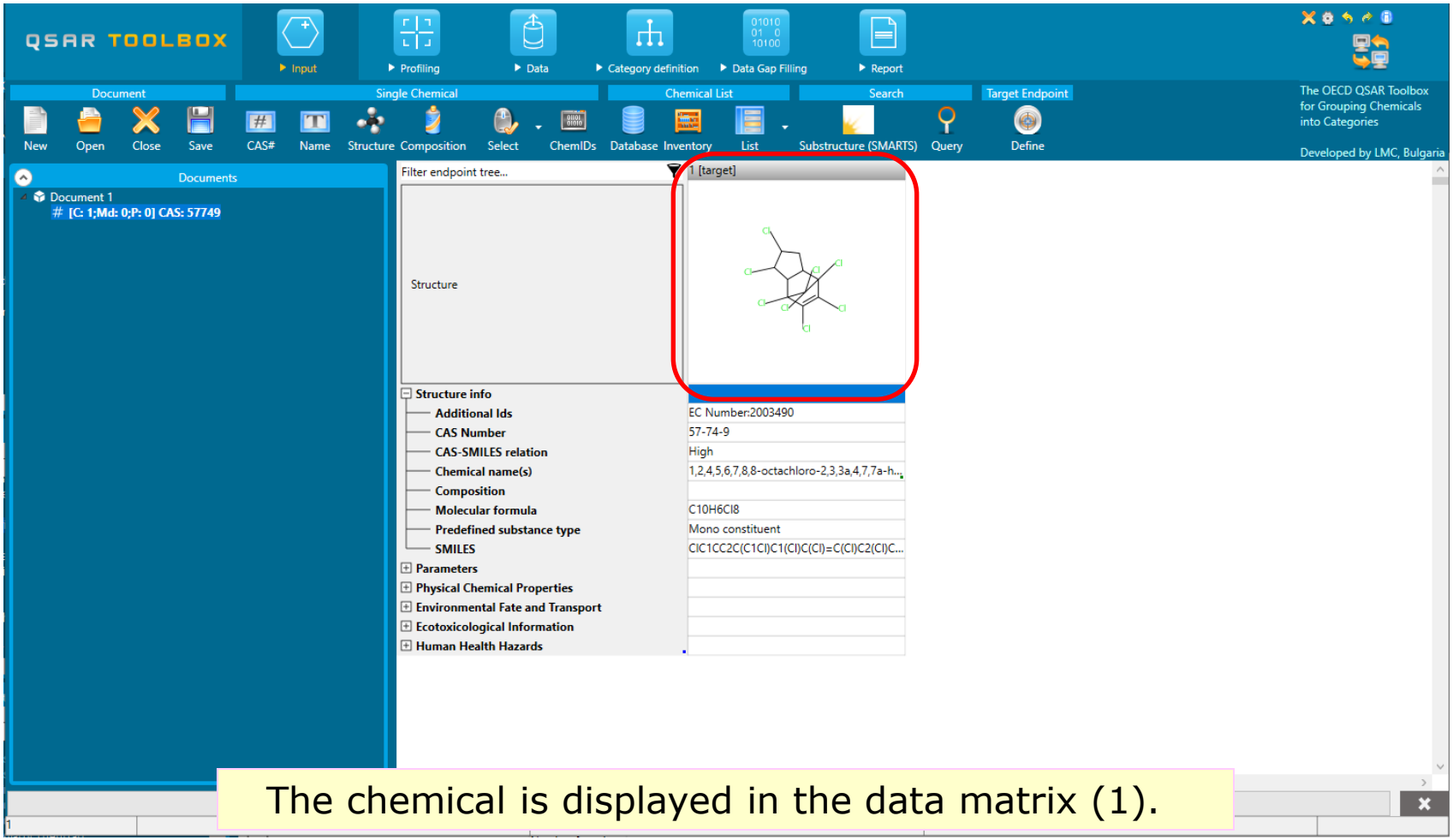
1. Click ***Input*** (1) to display the main Input section (2).

Chemical Input
Single chemical: CAS RN



1. Click **CAS#** (1);
2. Type in the **CAS # 57-74-9** (the target chemical) (2) ;
3. Click on **Search** (3);
4. Select the first chemical with High relation CAS-SMILES (4);
5. Click **OK** (5).

Chemical Input
Single chemical: CAS RN



The chemical is displayed in the data matrix (1).

Outlook

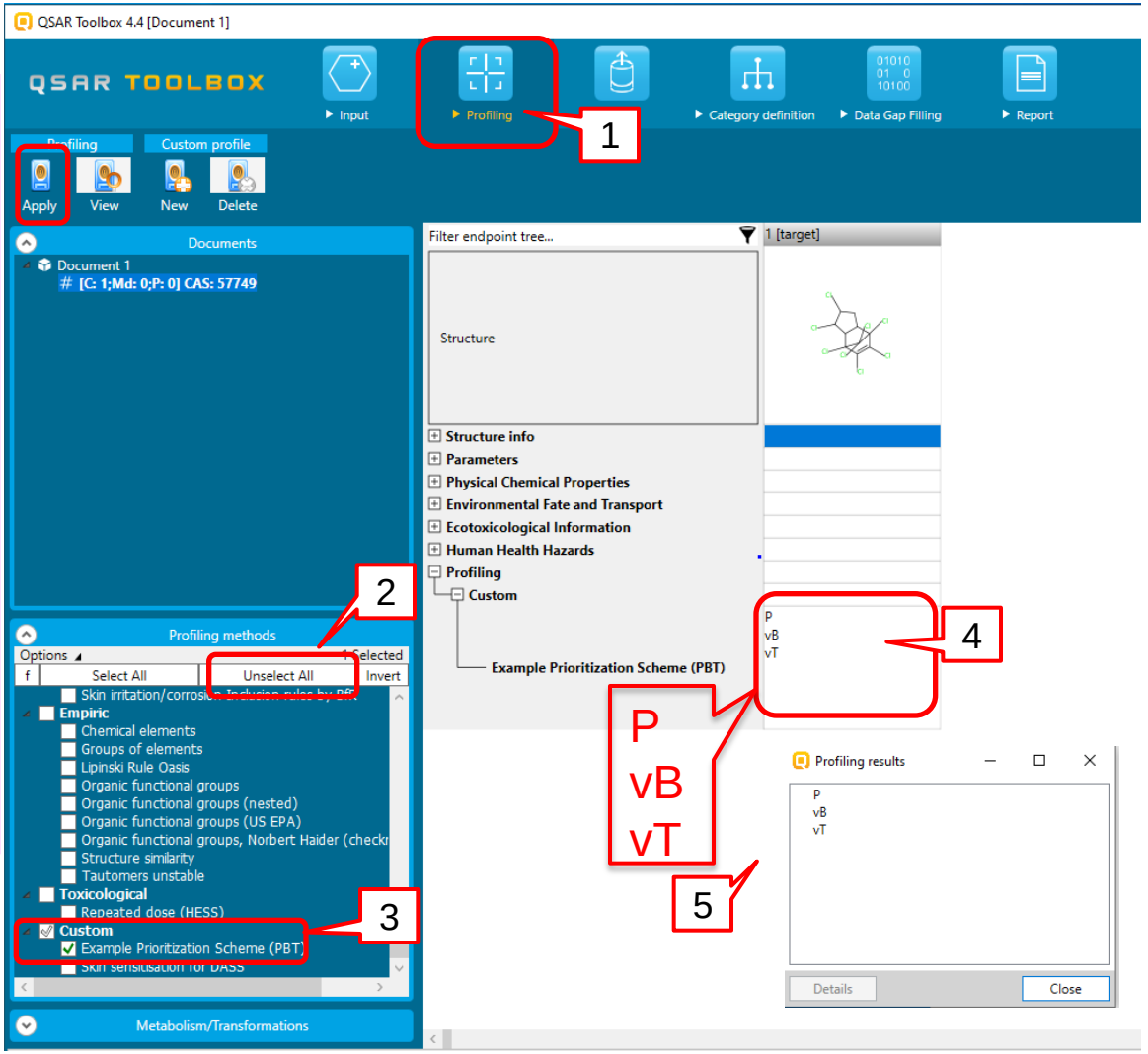
- Aim
- Keywords
- PBT scheme
  - Background
  - Implementation
- **Workflow of the prediction**
  - Chemical input
  - **Profiling**
- Export of the results

Profiling
Overview

- "Profiling" refers to the electronic process of retrieving relevant information on a compound which is stored in the Toolbox, other than its fate and (eco)toxicity data.
- Toolbox has many predefined profilers but it also allows the user to develop new profilers.

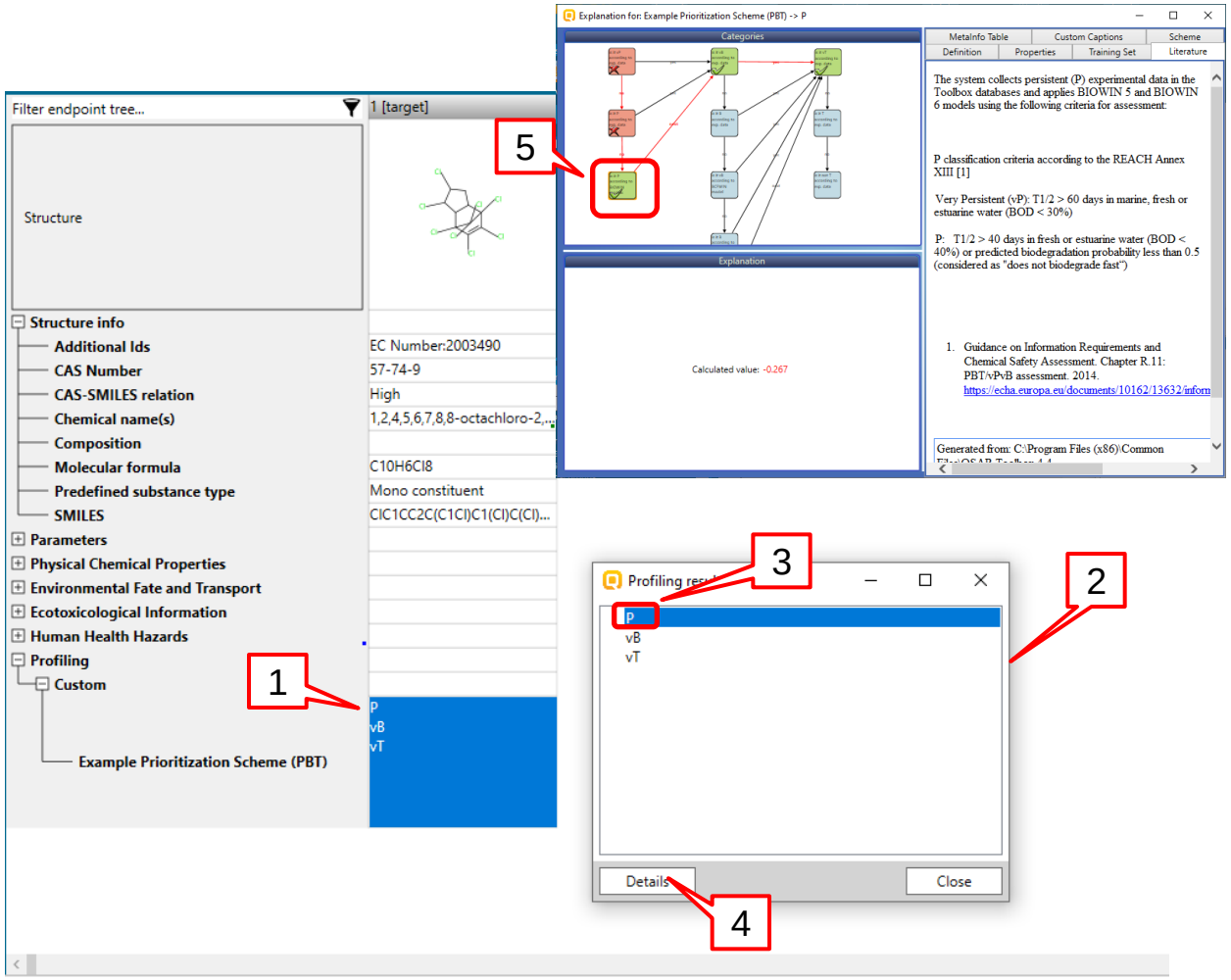
Profiling

1. Select ***Profiling tab*** (1);
2. Click **Unselect All** *(2);*
3. Tick ***Example prioritization Scheme (PBT)*** (3);
4. Click on **Apply** *(4).*
5. Expand the cell to see the results: P, vB, vT.
6. Double-click on the profiling results cell (5)
7. The results are explained in the next slides.

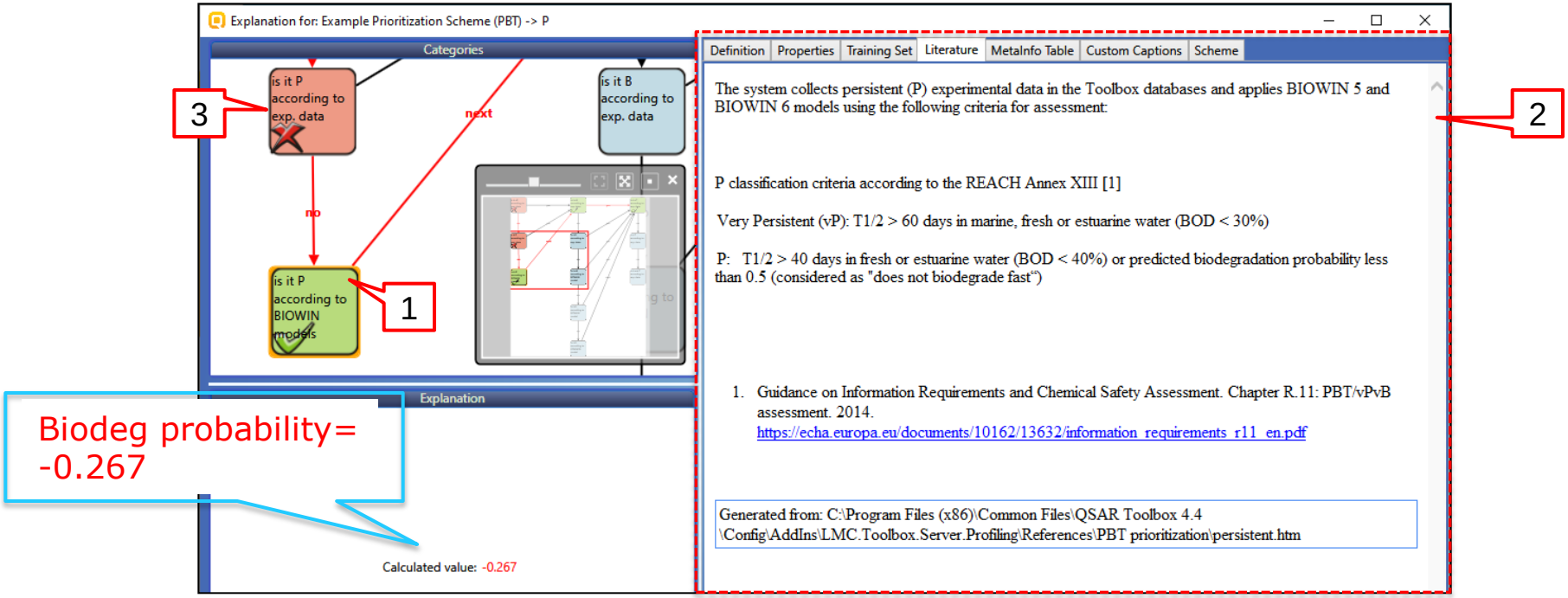


Profiling: explanation of "P" prediction

1. Double left-click on the cell with profiling results (1)
2. The **Profiling results** window is displayed (2)
3. Select *"P"* (4)
4. Click **Details** (5)
5. The PBT scheme is displayed and the query, which gives the prediction is marked (see next slide)



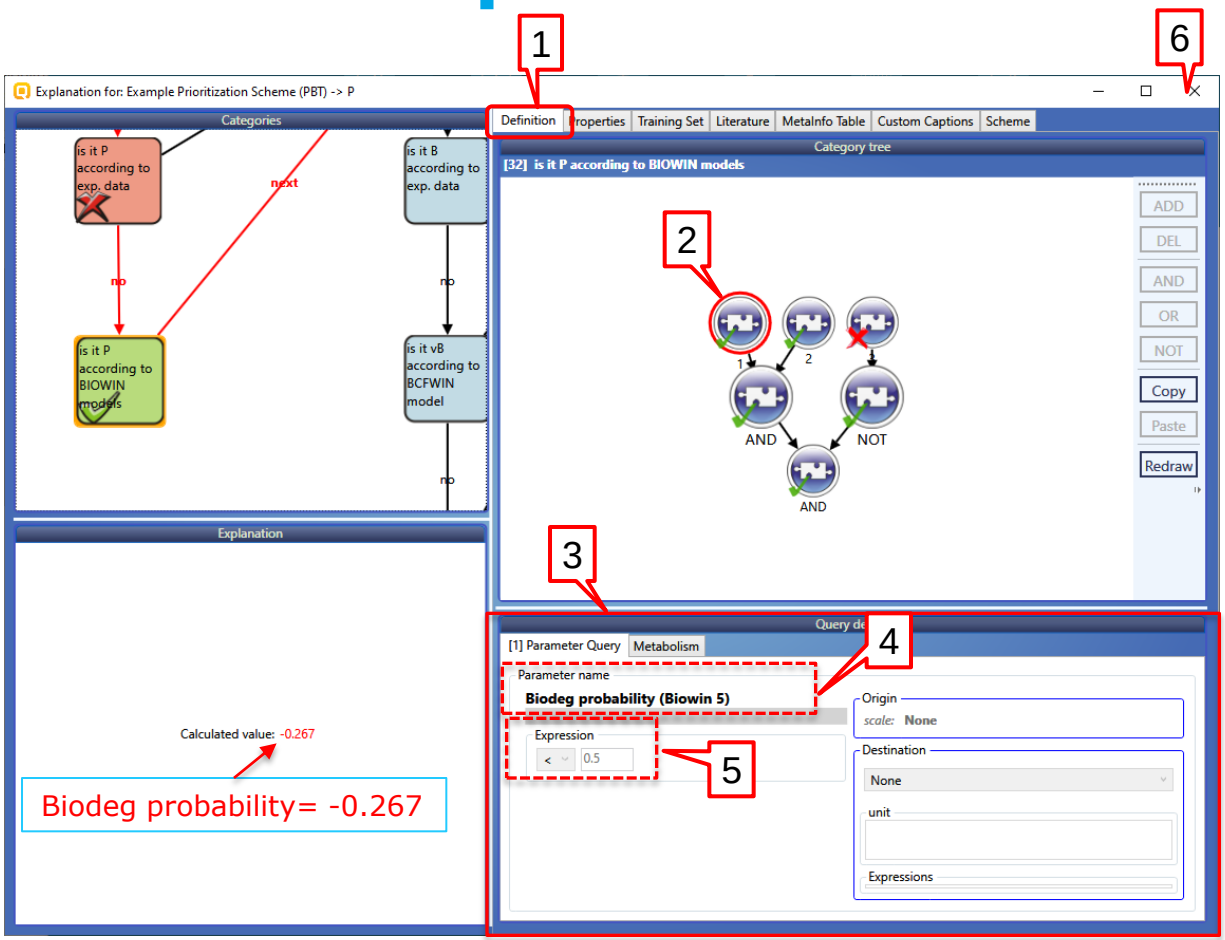
Profiling: explanation of "P" prediction



1. The "P" node (1), which criteria are fulfilled, is coloured in green and has a check mark. A yellow border around the node indicates that it is selected and its explanation is displayed in the Literature panel (2).
2. The nodes, which criteria are not fulfilled are colored in red and marked with "X" sign (3).

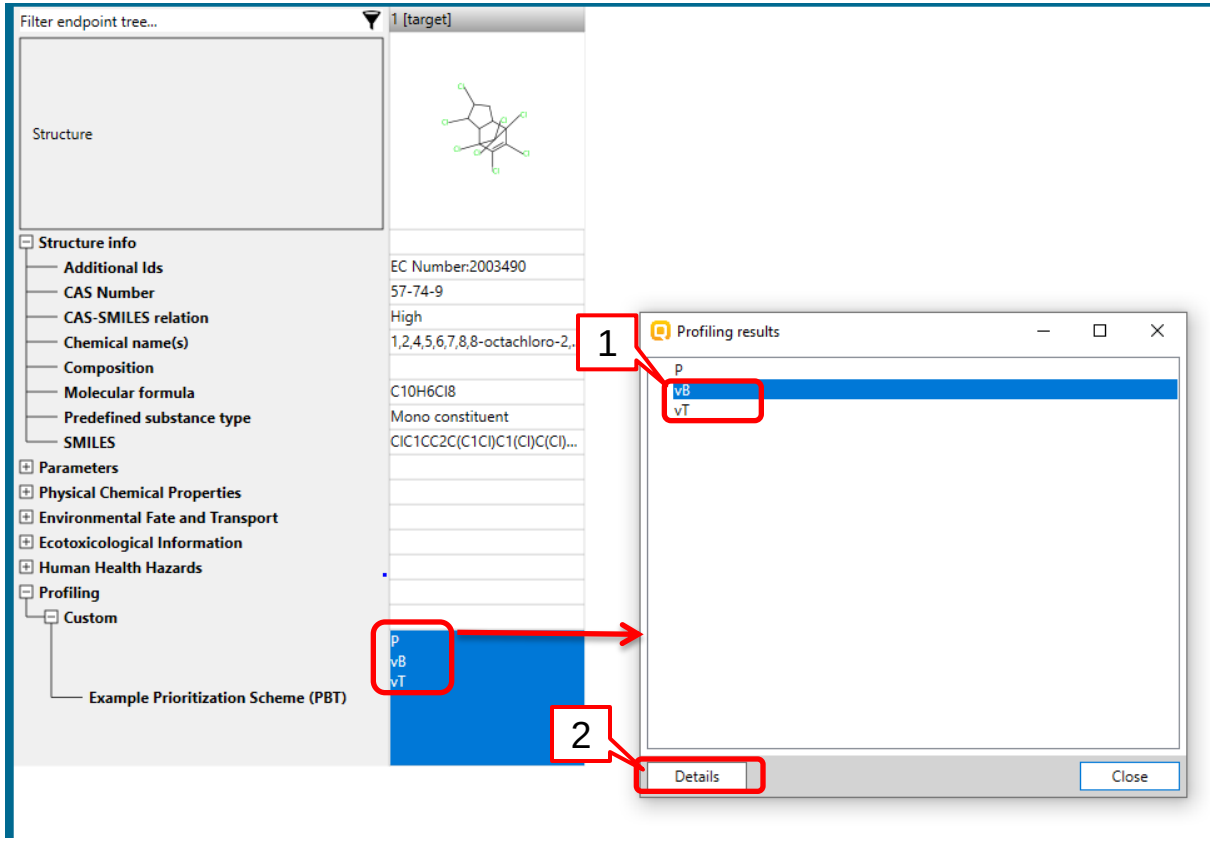
Profiling: explanation of "P" prediction

1. Click **Definition** tab (1).
2. A red circle around the query indicates that it is selected (2).
3. The selected query is a parametric query, which content is shown in the **Parameter query** tab (3).
4. Biodegradation probability (Biowin 5) model is incorporated in that query (4). The predicted outcome has to be below 0.5 (5) in order for the requirements to be fulfilled. Once finish with examination of the boundary the window could be closed (6)



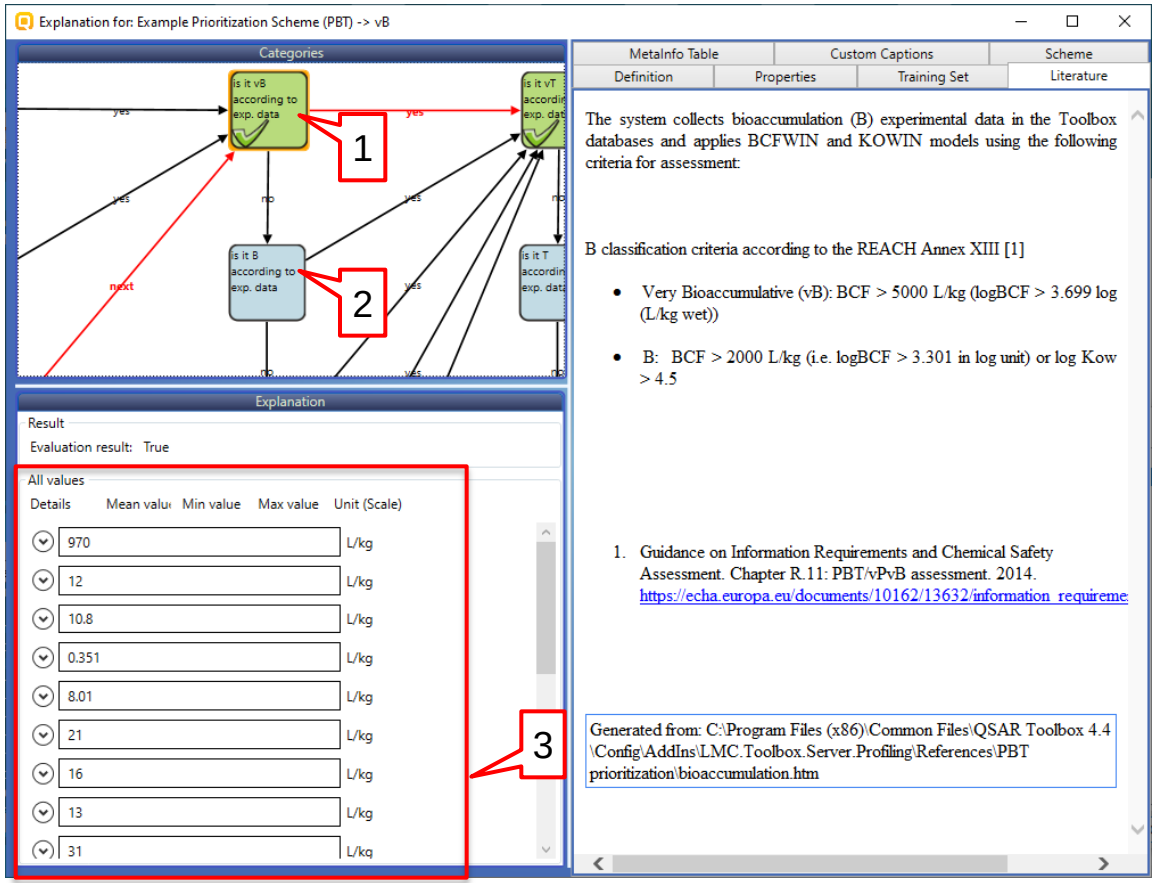
Profiling: explanation of "vB" prediction

1. Select *"vB" from the opened* **Profiling results** window
2. (4).
3. Click **Details** (5).
6. The PBT scheme is displayed and the query, which gives the prediction is shown (see next slide)

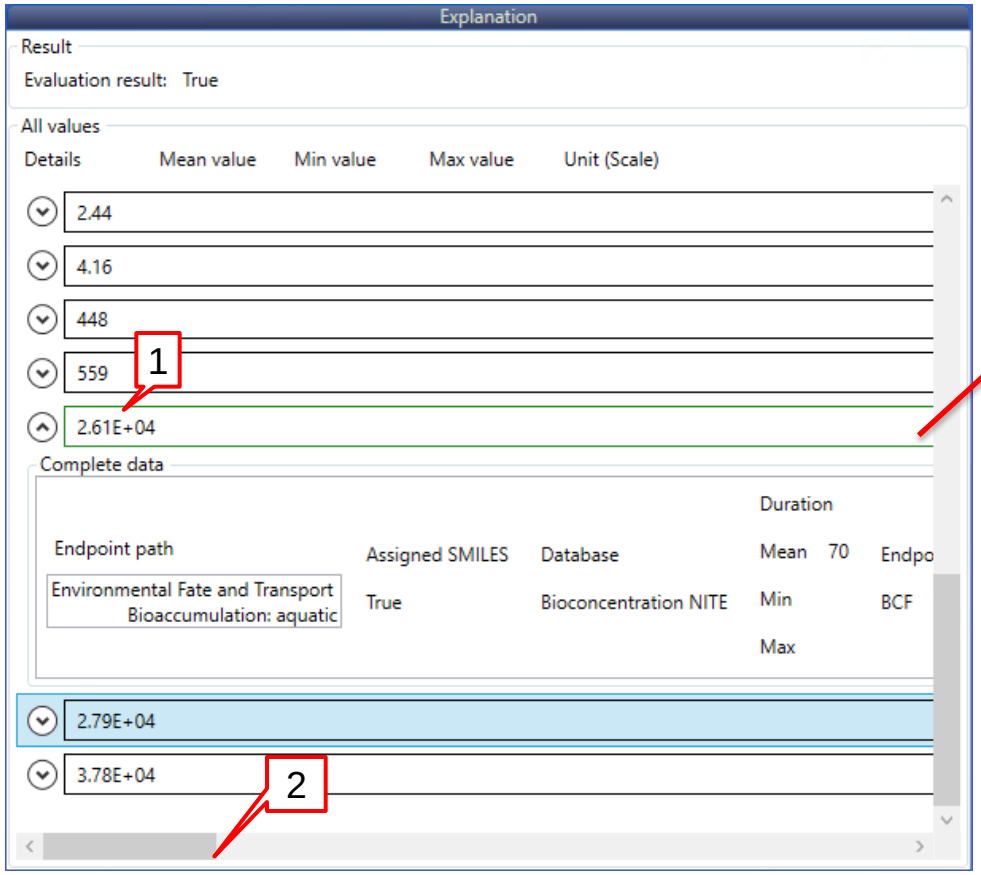


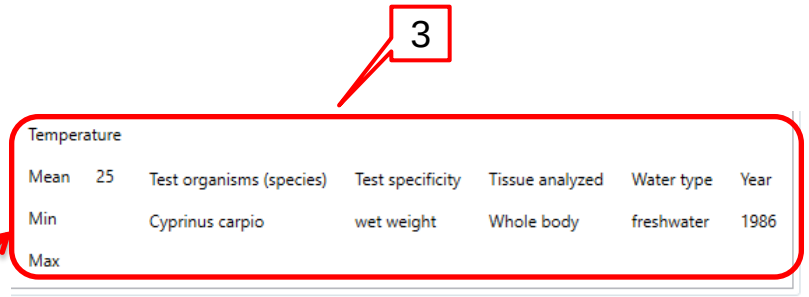
Profiling: explanation of "vB" prediction

1. The node, which criteria are fulfilled is green colored and marked with "√" sign (1).
2. Here it is a node containing experimental data. The logic implemented in the PBT scheme implies that if the target chemical has experimental data (in this case vB) then the consecutive nodes from the bioaccumulation assessment are not run. Hence they are colored in grey (2).
3. All experimental data available for the target chemical is listed in the **Explanation** panel(3). The unit is L/kg.



Profiling: explanation of "vB" prediction



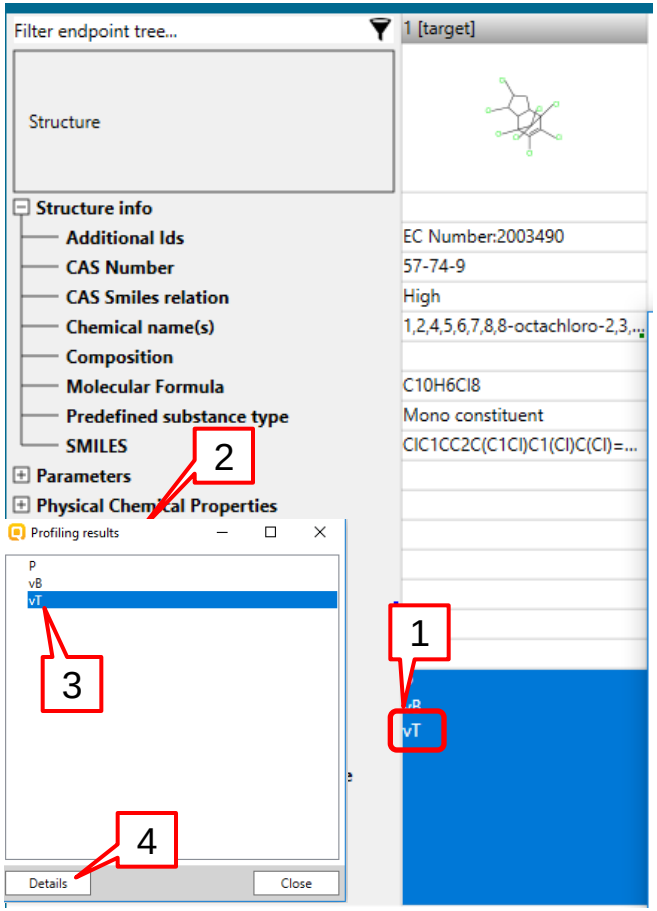


1. Expand the experimental result from the arrow next to the digits (1).
2. By moving the scroll bar (2) , you can see all the details of the measured data (3).

Profiling: explanation of "vT" prediction

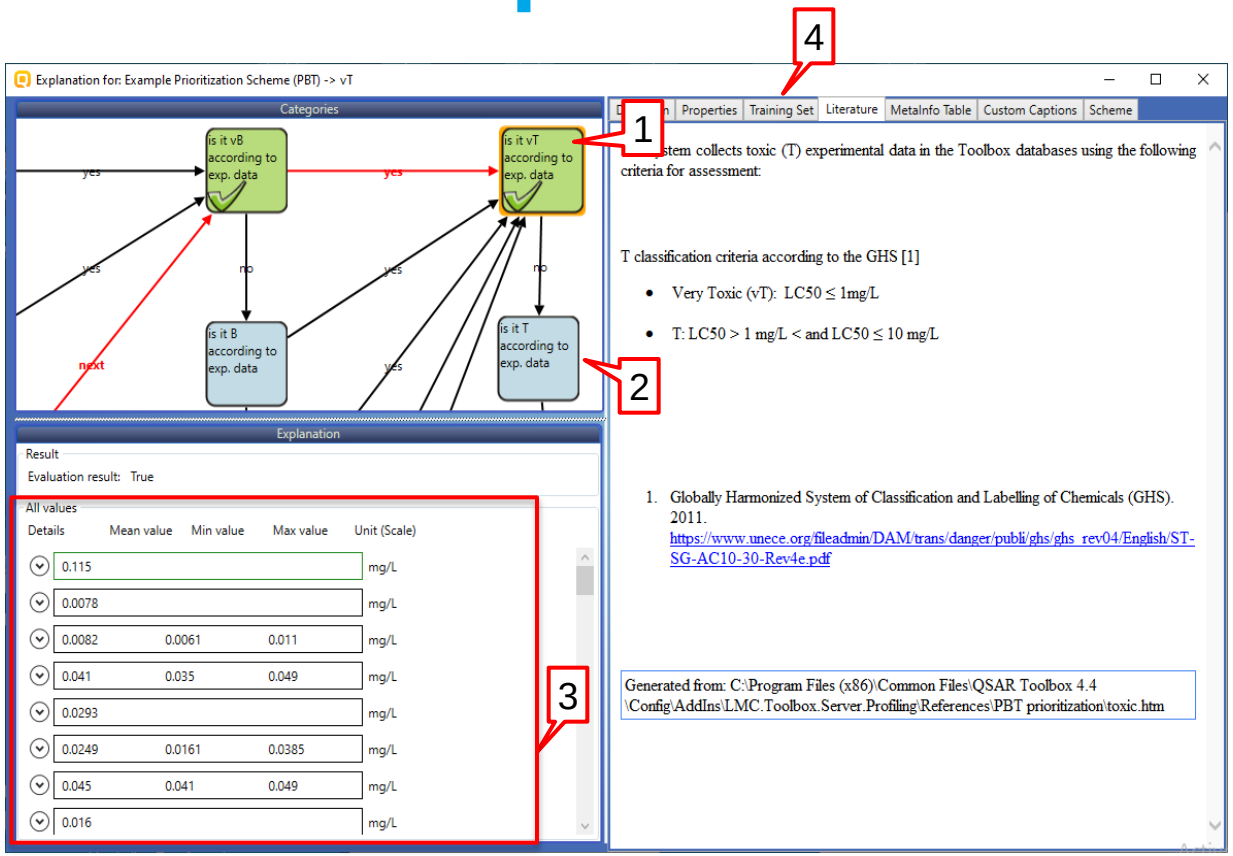
1. Double click on the profiling result cell (1).
2. The **Profiling results** window is displayed (2).
3. Select *"vT"* (3).
4. Click **Details** (4).

The PBT scheme is displayed and the query, which gives the prediction is shown (see next slide).



Profiling: explanation of "vT" prediction

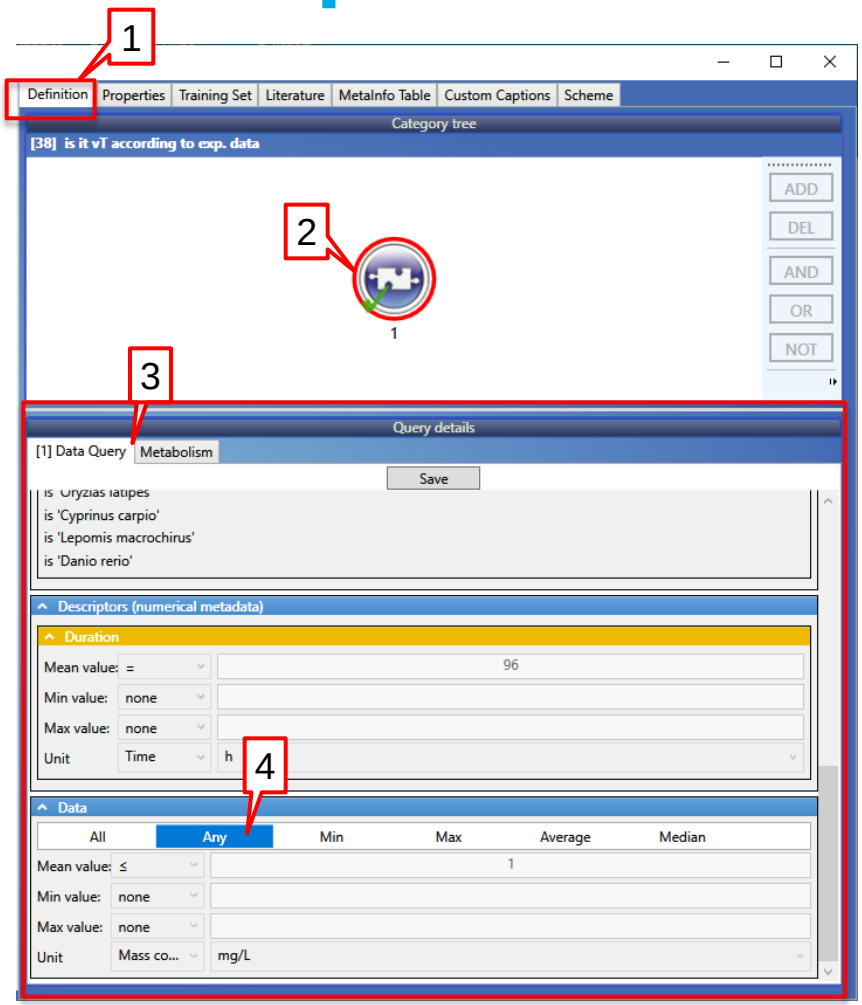
1. The node, which criteria are fulfilled is colored green and marked with "√" sign (1).
2. Here it is a node with experimental data. The logic implemented in the PBT scheme implies that if the target chemical has experimental data (in this case vT) then the consecutive nodes from the toxicity assessment are not run. Hence they are colored in grey (2).
3. All experimental data available for the target chemical is listed in the **Explanation** panel(3). The unit is mg/L or ppb. However only mg/L is considered (see slide 8).



4. The rules by which experimental data is collected were shown in the *Literature* tab (4).

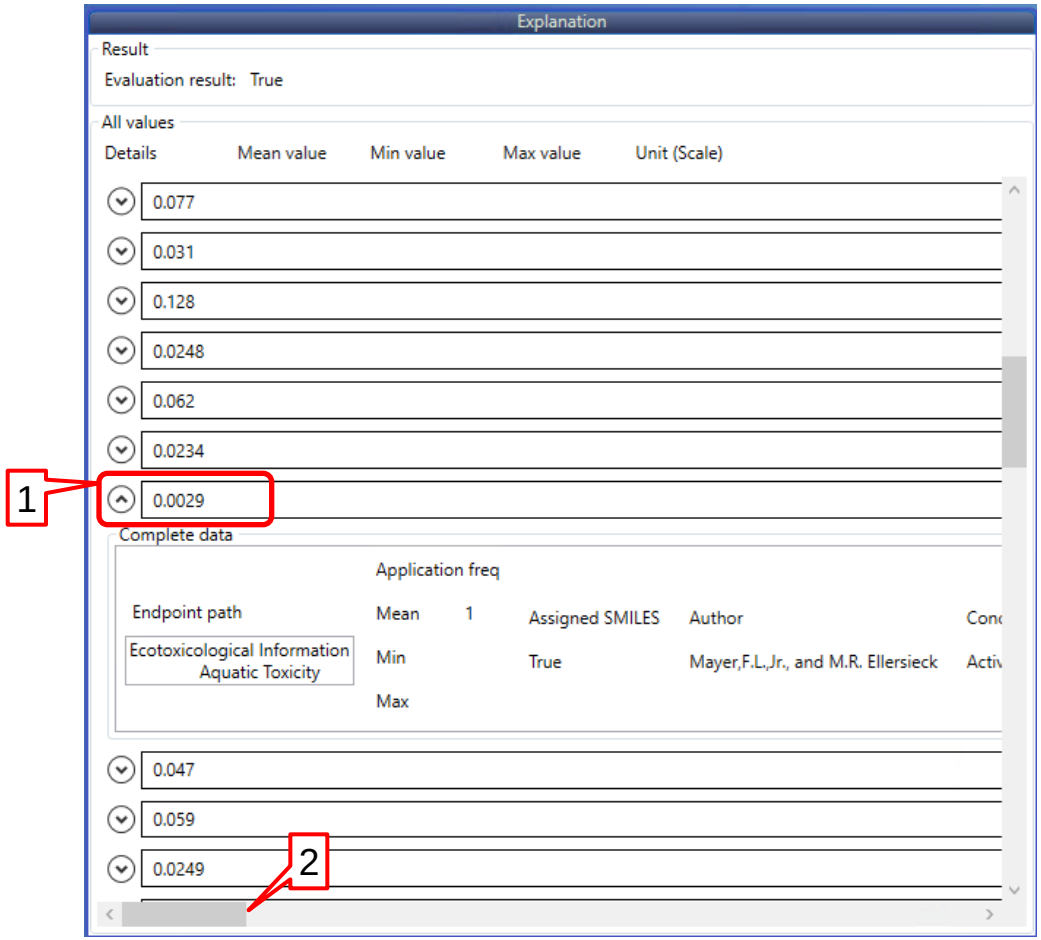
Profiling: explanation of "vT" prediction

1. Select **Definition** tab (1).
2. The details of the **Data query** (2), are shown in the Query details panel (3).
3. By moving the scrollbar, you can see the type of metadata included in the query.
4. In this example, the type of experimental data is **Any** (4) meaning that all available data is checked, but only data in mg/L and mean value ≤ 1 mg/L is considered.

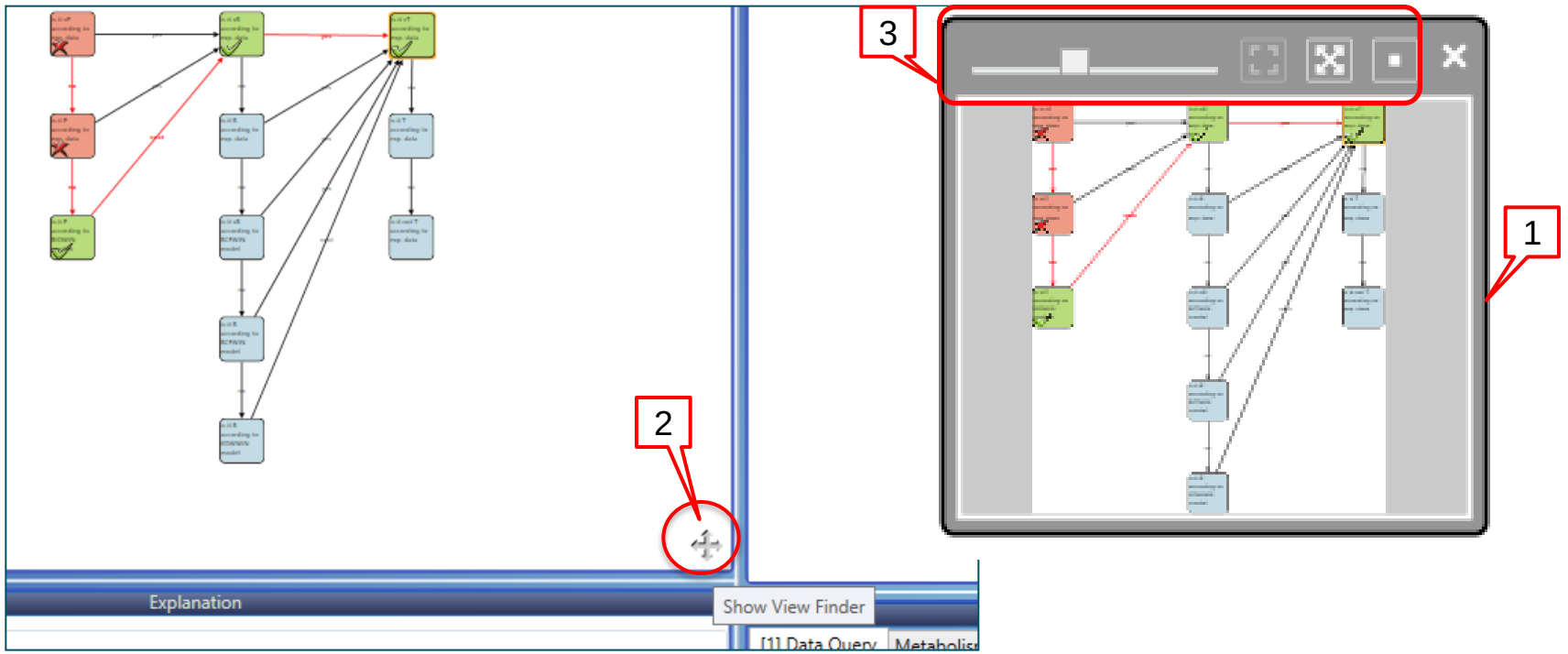


Profiling: explanation of "vT" prediction

1. Expand the experimental data from the arrow next to the digits (1)
2. By moving the scroll bar (2), you can see all details of the measured data

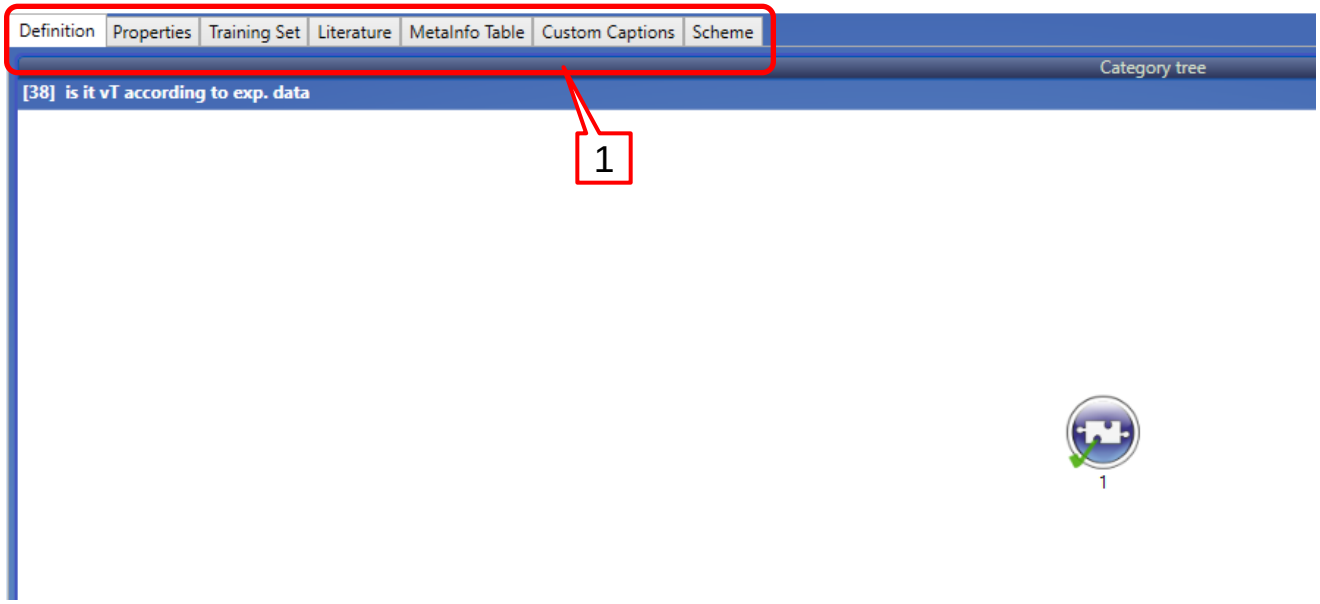


Profiling: common features of PBT interface



1. A **Viewfinder** (1) is displayed by clicking on the cross sign (2).
2. It contains navigation options for viewing the PBT scheme (3).

Profiling: common features of PBT interface



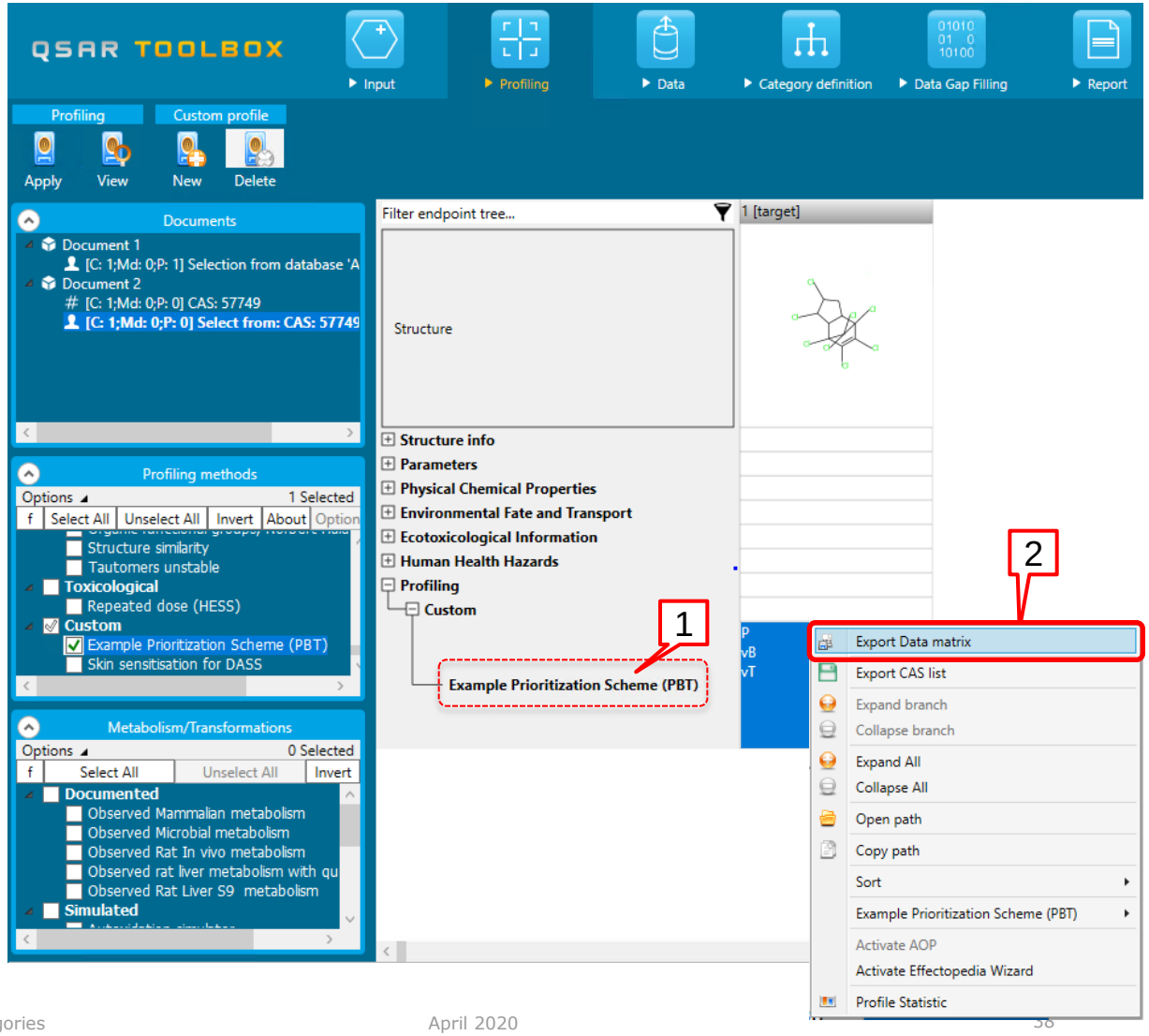
Properties, ***Literature*** and ***Scheme*** tabs gives more information about the query (properties), the node(literature) and the developer (scheme) (1).

Outlook

- Aim
- Keywords
- PBT scheme
  - Background
  - Implementation
- Workflow of the prediction
  - Chemical input
  - Profiling
- **Export of the results**

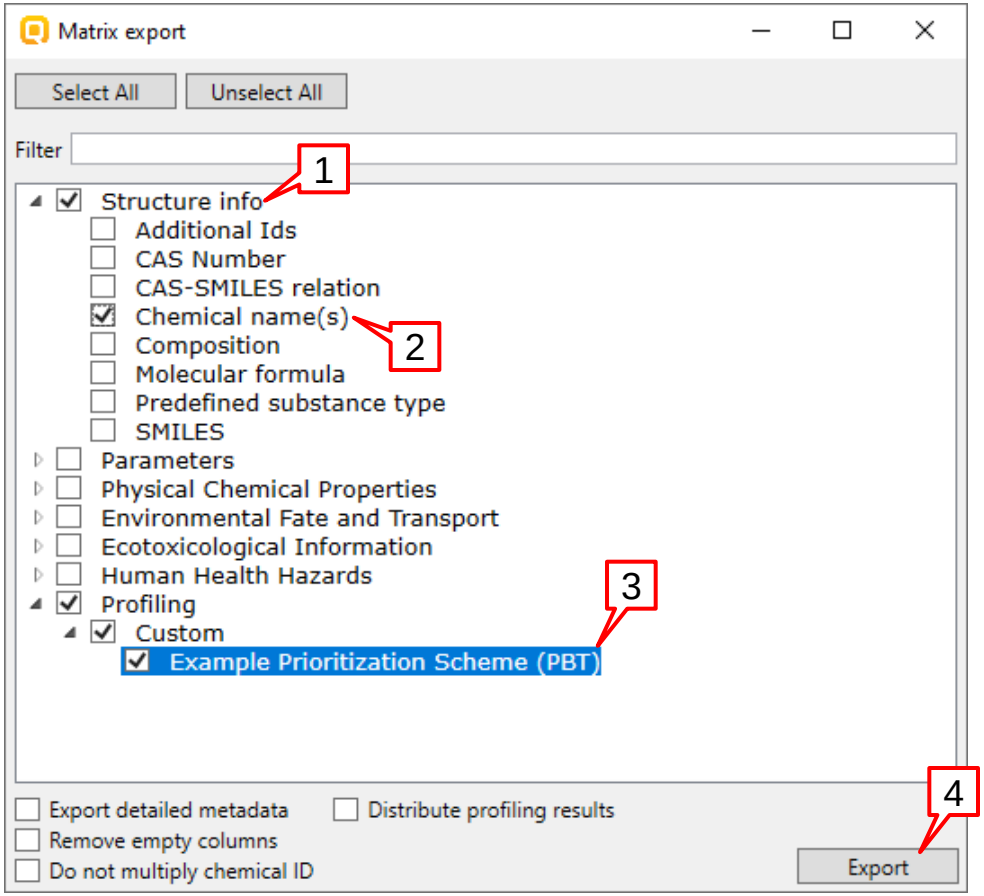
Export of the results

1. Right click next to the name of the scheme (1)
2. Select **Export Data matrix** (2)

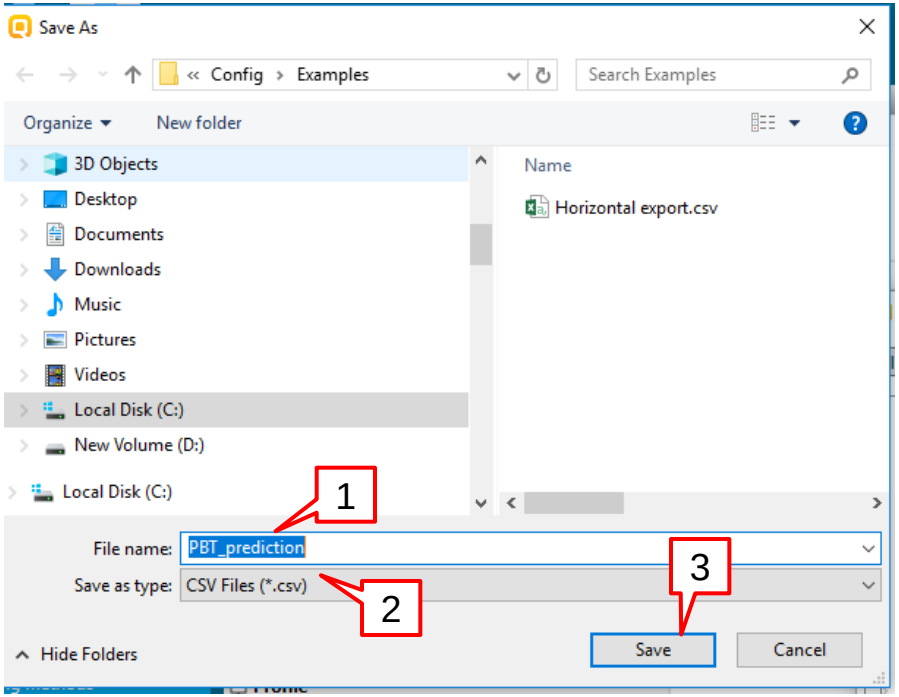


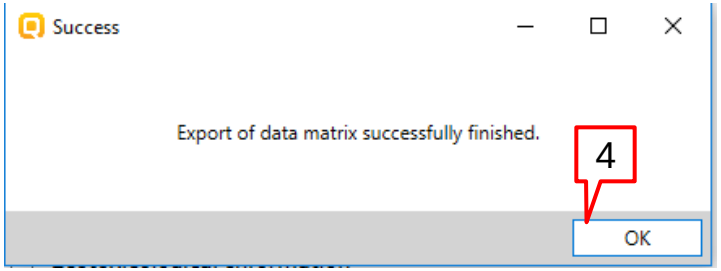
Export of the results

1. Expand structure info (1) and check **Chemical name(s)** (2).
2. **Example Prioritization Scheme (PBT)** is selected by default (3).
3. Click **Export** (4).



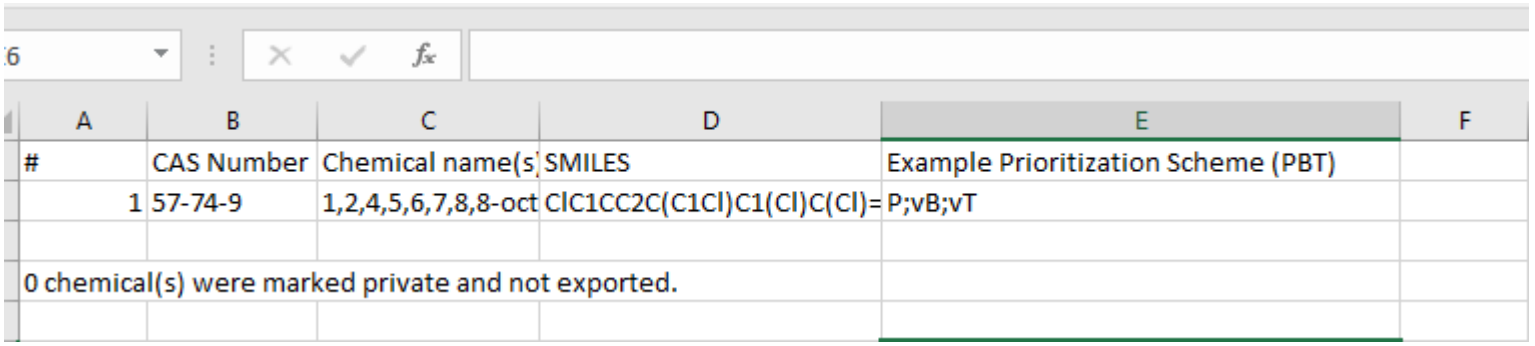
Export of the results





1. Type in the name of the file (1).
2. The files are only saved in *.csv* format(2).
3. Click **Save** (3).
4. Click **OK** in the confirmation message (4).

Export of the results



	A	B	C	D	E	F
	#	CAS Number	Chemical name(s	SMILES	Example Prioritization Scheme (PBT)	
	1	57-74-9	1,2,4,5,6,7,8,8-oct	ClC1CC2C(C1Cl)C1(Cl)C(Cl)=	P;vB;vT	
	0 chemical(s) were marked private and not exported.					

The file can be open as an Excel sheet.