

OECD QSAR Toolbox v.4.4.1

Step-by-step example on how to build a user-defined profiling scheme accounting for the (a)biotic activation of the chemicals

Outlook

- **Background**
- Objectives
- Profiling
- The exercise

Background

- This is a step-by-step presentation designed to provide guidance to the Toolbox user on how to create their own profiling scheme accounting for the (a)biotic activation of the chemicals.
- The created custom profiler can be used for searching precursors of a given target.
- Two examples will be illustrated:
 - Searching PBT precursors in the Toolbox databases;
 - Searching p-benzoquinone precursors in a specific file.

Note: Please note that building of custom items (such as profilers, (Q)SAR models as well as importing of custom databases) is only enabled in single user mode. So, if your Toolbox is installed in multiuser mode, you will be not able to follow this tutorial.

Outlook

- Background
- **Objectives**
- Profiling
- The exercise

Objectives

- **This presentation demonstrates how to build and use a new profiling scheme including:**
 - naming of the new scheme;
 - selection of specific (a)biotic activation;
 - searching of specific (a)biotic product;
 - saving the new profiler;
 - defining category by using the custom profiler;
 - profiling a chemical list.

Outlook

- Background
- Objectives
- **Profiling**
- The exercise

Profiling Overview

- As you know, “Profiling” refers to the electronic process of retrieving relevant information on a compound which is stored in the Toolbox, other than fate and toxicity data.
- The Toolbox has many predefined profilers but it also allows the user to develop new custom profilers.

Outlook

- Background
- Objectives
- Profiling
- **The exercise**

The Exercise

Example 1

- In the first example we will build a profiler that identifies precursors of 1,2,4-trichlorobenzene [CAS 120-82-1] (i.e. we will search for chemicals that are transformed to our target).
- The 1,2,4-trichlorobenzene (1,2,4-TCB) is considered as a substance fulfilling the PBT-criteria^[1,2]. Therefore, the potential precursors of 1,2,4-TCB could be prioritized for further PBT-related investigations.
- Precursors of 1,2,4-TCB will be searched in the Toolbox databases containing biodegradation (BOD) 301C data.

[1] [European Chemicals Agency \(ECHA\). PBT/vPvB assessments under the previous EU chemicals legislation.](#)

[2] [European Chemicals Agency \(ECHA\). ECB – Summary fact sheet. PBT working group –PBT list No.4. 2008. TC NES subgroup on identification of PBT and vPvB substances. Results of the evaluation of the PBT/vPvB properties of: 1,2,4-trichlorobenzene.](#)

The Exercise

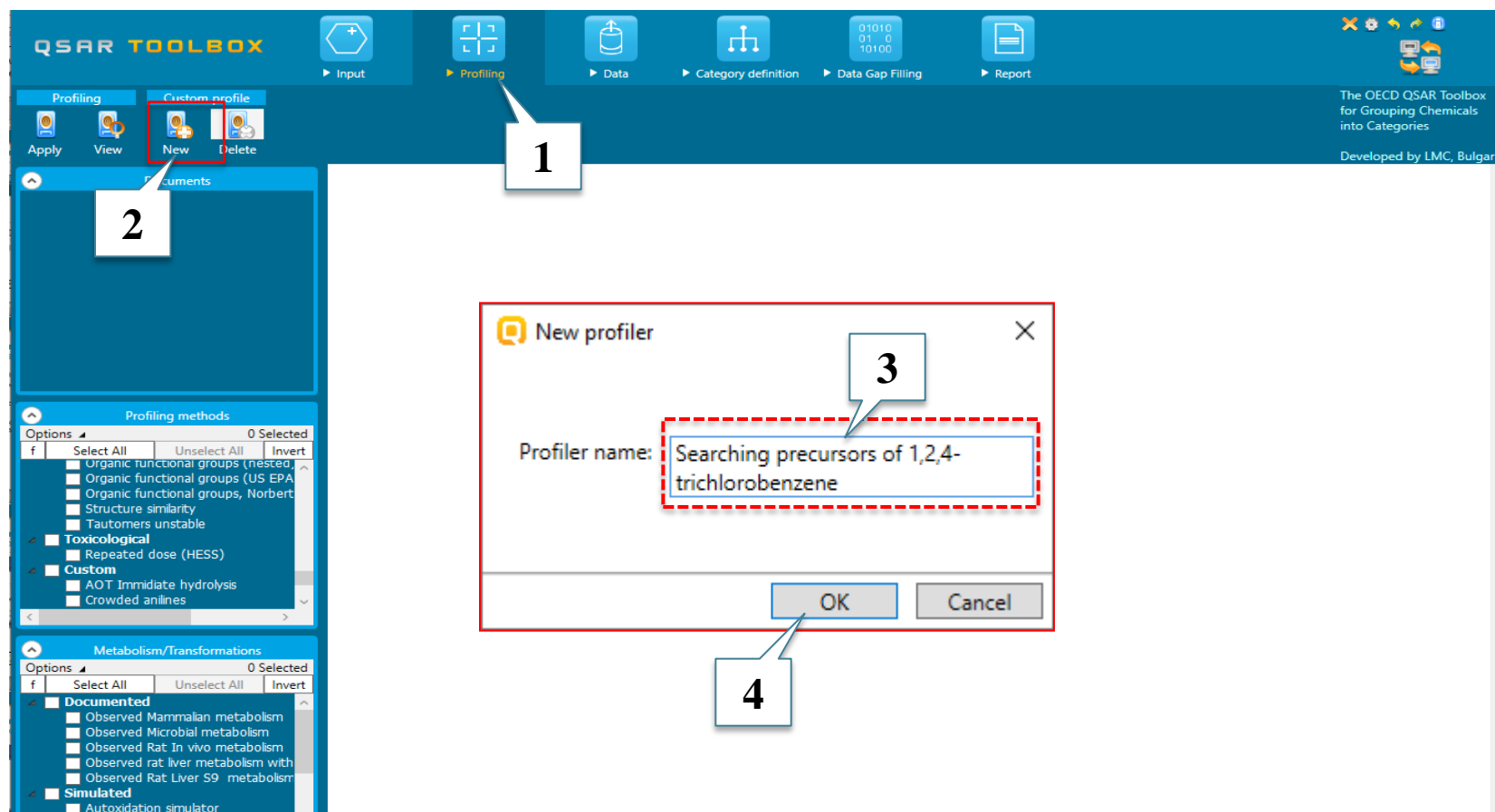
Start building a new profiler

We are going to create a new profiling scheme:

- Open the Toolbox.
- Move to the Profiling module
(see next screen shot).

Building of a new profiler

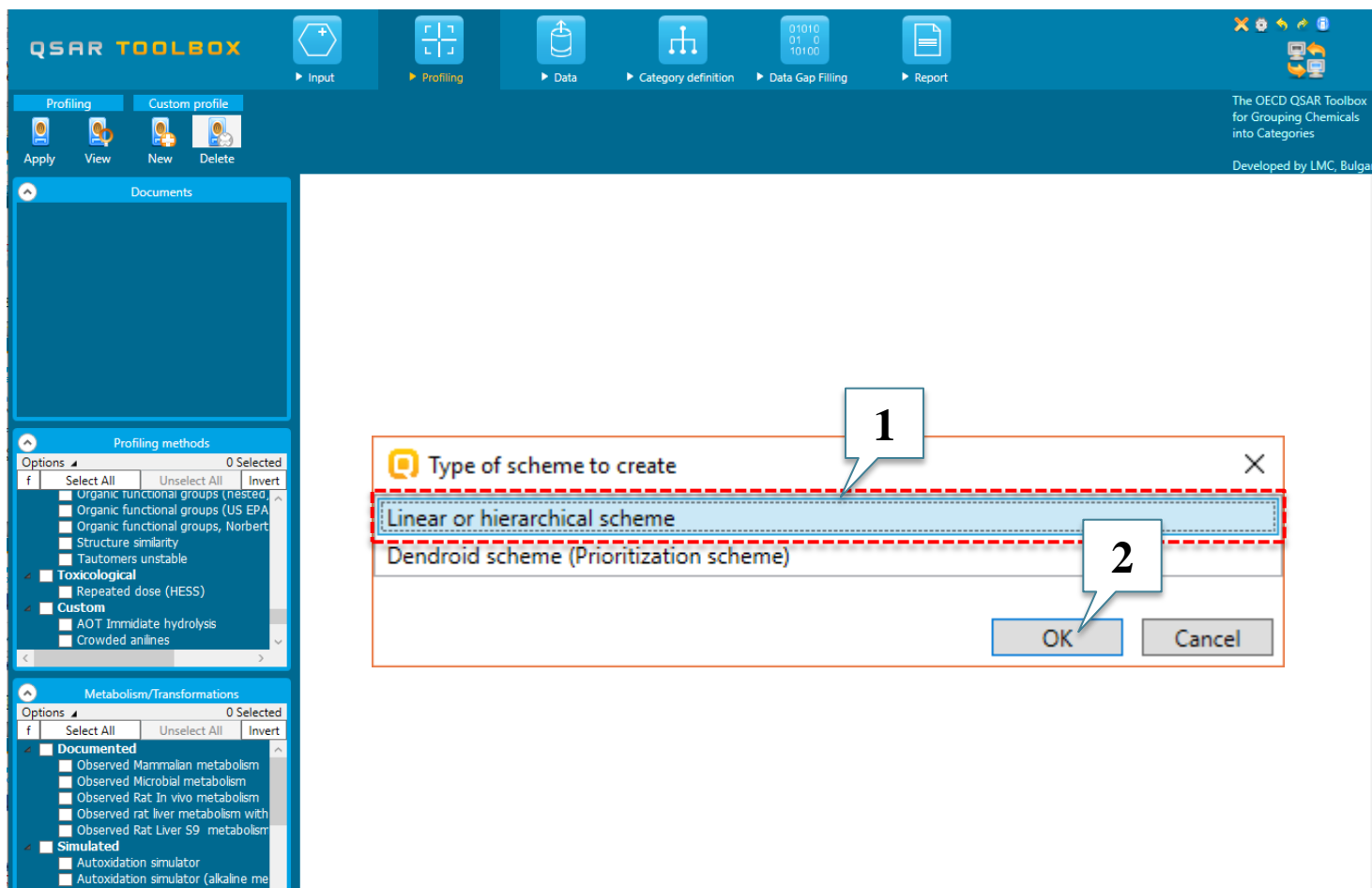
Define the name of the new profiler



1. Go to the **Profiling** module; 2. Click **New**; 3. Enter the name of the new scheme, for example **"Searching precursors of 1,2,4-trichlorobenzene"**; 4. Confirm by click **OK**.

Building of a new profiler

Define type of new profiler



1. Select the first option for scheme type: **Linear or hierarchical scheme**; 2. Click **OK**.

Building of a new profiler

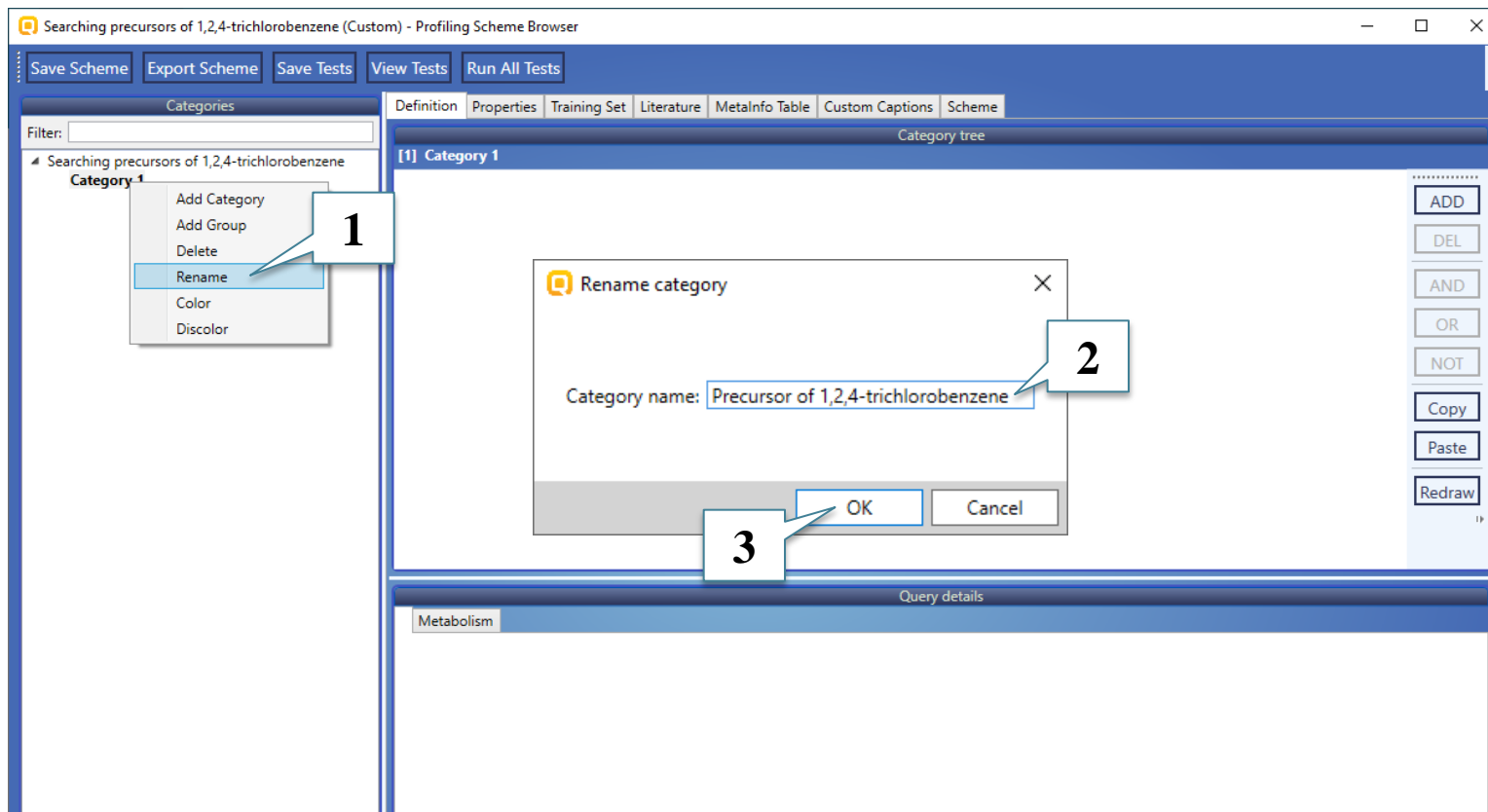
Creating a profiling category

Once the scheme is created, it will be defined:

- the name of the profiling category, which will appear on the data matrix;
- the structure which we will be searching for (i.e. *1,2,4-TCB*);
- the simulator, which will be used to simulate the transformations of the chemicals (i.e. *Microbial degradation simulator*).

Building of a new profiler

Name a profiling category



1. Right click on "Category 1" and select to **Rename** it; 2. Put a name that will appear on the data matrix when profile, e.g. *Precursor of 1,2,4-trichlorobenzene*; 3. Confirm by **OK**.

Building of a new profiler

Define the structural criteria

Searching precursors of 1,2,4-trichlorobenzene (Custom) - Profiling Scheme Browser

Save Scheme Export Scheme Save Tests View Tests Run All Tests

Categories

Filter:

Searching precursors of 1,2,4-trichlorobenzene
Precursor of 1,2,4-trichlorobenzene

Definition Properties Training Set Literature MetaInfo Table Custom Captions Scheme

Category tree

[1] Category 1

ADD
DEL
AND
OR
NOT
Copy
Paste
Redraw

Data Query
Parameter Query
QSAR Domain Query
QSAR Query
Reference Query
Similarity Query
Structure Query

2

Query details

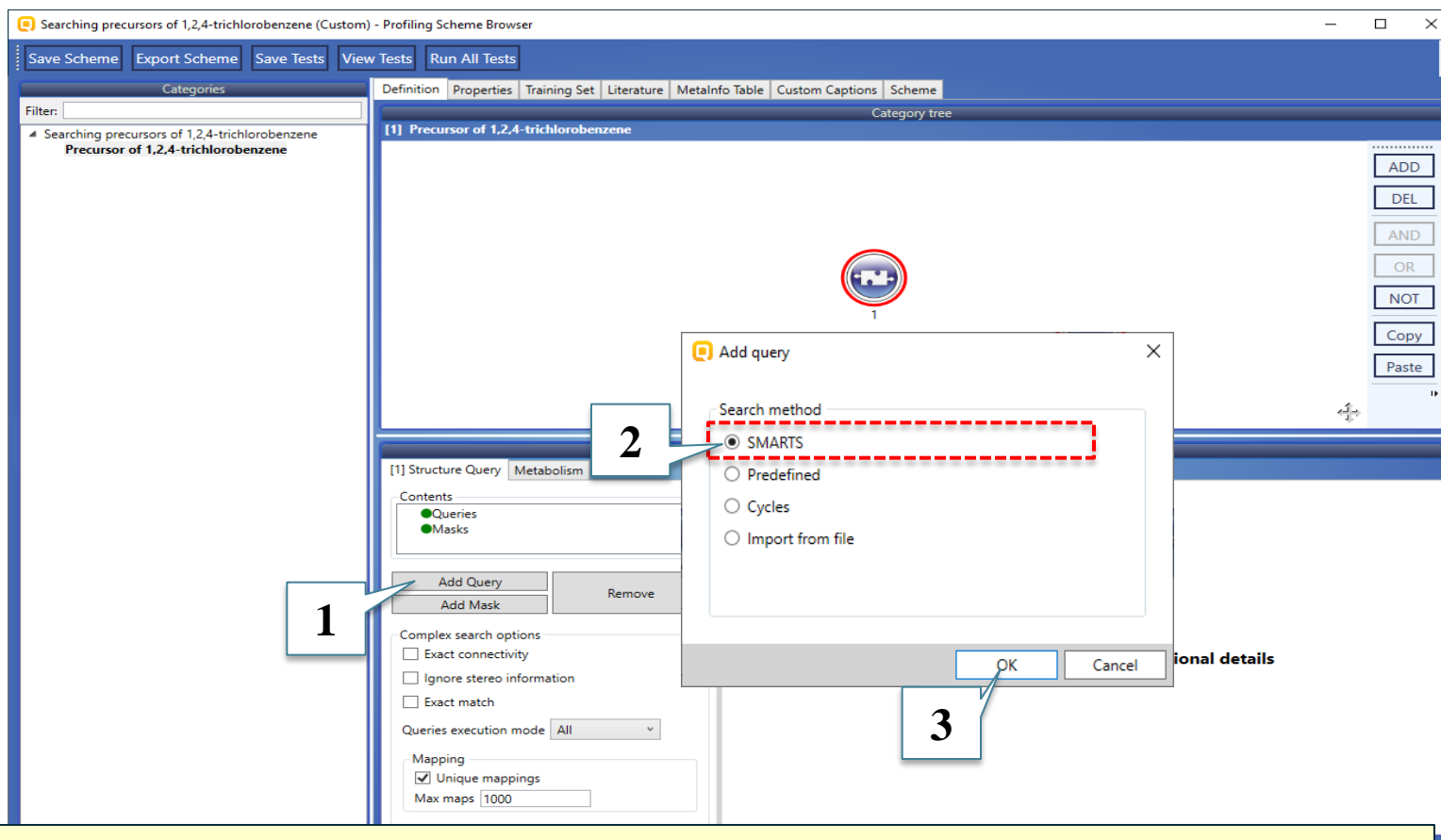
Metabolism

First we need to define the structural criteria, i.e. the structure of the metabolite which we will searching for. The structural criteria will be defined using the SMARTS language.

1. Click on the **ADD** button; 2. Select **Structure query**.

Building of a new profiler

Define the structural criteria



1. Click on the **Add Query** button; 2. Select to add SMARTS; 3. Confirm by **OK**.

Building of a new profiler

Define the structural criteria

The screenshot displays the 'Profiling Scheme Browser' window for 'Searching precursors of 1,2,4-trichlorobenzene (Custom)'. The interface includes a 'Categories' panel on the left, a 'Definition' tab at the top, and a 'Query details' section at the bottom. In the 'Query details' section, a red dashed box highlights the 'SMARTS' field containing 'c1ccccc1' and an 'Edit' button. A callout box labeled '1' points to the 'Queries' section in the 'Contents' panel, and another callout box labeled '2' points to the 'Edit' button. The 'Query details' section also shows 'View mode' set to 'Facade' and 'Navigation mode' set to 'Cascade'. A chemical structure of a benzene ring is displayed in the 'View mode' area.

1. Click on the created SMARTS query in order to define it. The SMARTS editor will appear;
2. Click **Edit** to define the structure.

Building of a new profiler

Define the structural criteria

Searching precursors of 1,2,4-trichlorobenzene (Custom) - Profiling Scheme Browser

Save Scheme Export Scheme Save Tests View Tests Run All Tests

Categories Definition Properties Training Set Literature MetaInfo Table Custom Captions Scheme

Filter: Searching precursors of 1,2,4-trichlorobenzene (Custom) - Profiling Scheme Browser

SMARTS Editor

Smarts: c1cc(c(cc1Cl)Cl)Cl

1a

1b

2b

2a

3

Once the SMART editor is opened, we need to draw the structure which we will searching for (i.e. 1,2,4-TCB)

Click on the drawing tool (1a) and draw three bonds to the benzene scaffold (1b). Then click the chlorine symbol (i.e. **Cl**) (2a) and click on the bonds' ends (2b). Finally click **OK** (3)

Building of a new profiler

Define the structural criteria

The structural criteria is defined. Now we need to specify the metabolic simulator, which will be used to simulate the microbial degradation of the chemicals.

Select the options for **Exact connectivity** and **Exact match** (1) in order to search for exactly the same structure, without any other substituents. Then go to the **Metabolism** tab (2).

Building of a new profiler

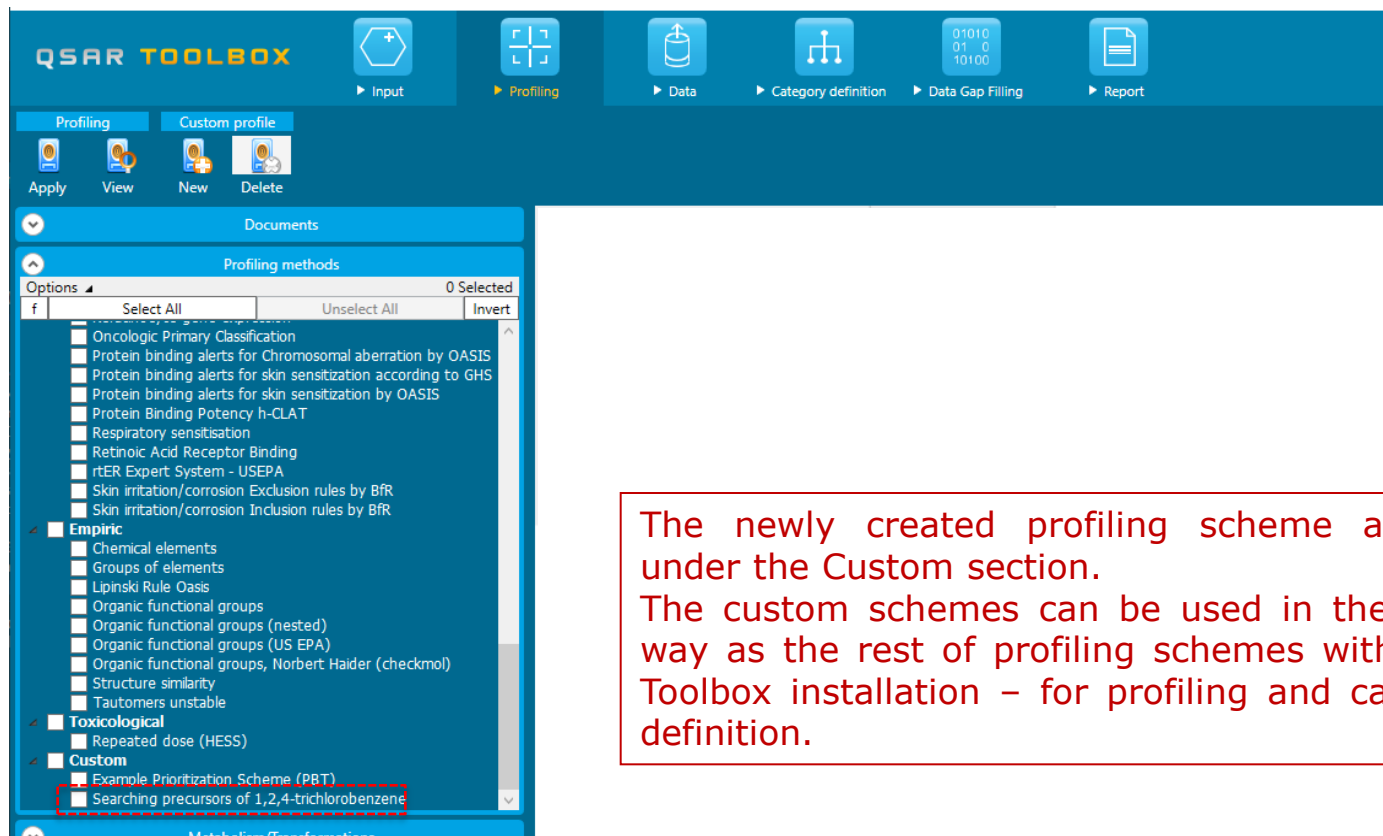
Define the metabolic criteria

In the current example we search for chemicals that are potential PBT substances based on their PBT degradation product. Therefore, the *Microbial metabolism simulator* will be used.

Select the **Microbial metabolism simulator** from the drop-down menu (1). This simulator will be applied on the chemicals as parents (i.e. *Input* >> *Parents* (2)) and we will search within the whole package – parent and produced metabolites (i.e. *Output* >> *All* (3)). The process mode should be *Any* (parent and produced metabolites) (4). Finally, click on **Save Scheme** (5).

Building of a new profiler

Custom profiler location



The newly created profiling scheme appears under the Custom section.

The custom schemes can be used in the same way as the rest of profiling schemes within the Toolbox installation – for profiling and category definition.

Building of a new profiler

Custom profiler usage

1

2

3

4

Precursors of 1,2,4-TCB will be searched in the databases containing BOD data. Therefore, we need first to define the target endpoint, which will highlight the relevant databases.

1) Enter the target by CAS (CAS 120-82-1); 2) Click on the **Define** button in order to define the target endpoint; 3) Select *Environmental Fate and Transport* >> *Biodegradation* >> *Biodegradation in Water: Screening Tests* >> *Ready biodegradability* and then specify the *BOD* endpoint for 28d according to *OECD 301C* guideline; 4) Click **Finish** to finalize the endpoint definition.

Building of a new profiler

Custom profiler usage

The screenshot displays the QSAR Toolbox software interface. The top toolbar includes icons for Input, Profiling, Data (highlighted with a red dashed box and labeled '1'), Category definition, Data Gap Filling, and Report. Below the toolbar, the 'Data' module is active, showing a 'Filter endpoint tree...' on the left and a 'Structure' panel on the right. The 'Structure' panel displays a chemical structure of 1,2,4,5-tetrachlorobenzene. The 'Filter endpoint tree...' lists various endpoints, with 'Biodegradation' and 'Ready Biodegradability' expanded. The 'Databases' panel on the left shows a list of databases, with 'Biodegradation NITE' and 'No data available' highlighted by a red dashed box and labeled '2'. The 'Inventories' panel at the bottom shows a list of inventories, with 'Canada DSL' and 'COSING' highlighted.

1) Go to the **Data** module; 2) Select the highlighted database.

Building of a new profiler

Custom profiler usage

1) Go to the **Category definition** module; 2) Select the custom profiler; 3) Click on the **Define** button; 4) Click **OK** button to initialize the search of 1,2,4-TCB precursors.

Building of a new profiler

Custom profiler usage

The screenshot displays the QSAR Toolbox interface. The 'Filter endpoint tree...' dialog is open, showing a hierarchical tree of endpoints. The 'Biodegradation' endpoint is selected, and its sub-endpoints are expanded. The 'Grouping results' dialog box is open, showing '6 chemical(s) found.' and an 'OK' button. The 'Read data?' dialog box is also open, with 'All endpoints' selected and 'OK' and 'Cancel' buttons. Callout numbers 1 and 2 point to the 'OK' buttons in the respective dialog boxes.

Six chemicals are found (the target plus five analogues. Confirm by **OK** (1). The system offers you to check for available experimental data for the found analogues in the selected databases. Click **OK** (2).

Note: The tutorial is consistent with Toolbox v.4.4 (and TB 4.4.1), where the *Biodegradation NITE* database is preliminary cached with the microbial metabolism. If older version of Toolbox is used, the search will take up to several minutes if you make this example for first time.

Building of a new profiler

Custom profiler usage

The screenshot shows the QSAR Toolbox interface. On the left, a list of documents is visible, including 'Searching precursors of 1,2,4-trichlorobenzene'. The main window displays a 'Filter endpoint tree...' dialog box. The tree structure shows 'Biodegradation' expanded, with 'BOD' selected. A callout box labeled '1' points to the 'OK' button. Below the dialog, a table shows data for BOD 301C across six chemicals, with values like '5/5' and 'M: 0 %'.

Chemical	1 [target]	2	3	4	5	6
Structure						
Structure info						
Parameters						
Physical Chemical Properties						
Environmental Fate and Transport						
Bioaccumulation: aq						
Bioaccumulation: ter						
Biodegradation						
Biodegradation in						
Biodegradation in						
Biodegradation in						
Biodegradation in						
Half-life						
Ready Biodegradability						
OECD 301C						
28 d						
BOD	5/5	M: 0 %	M: 0 %	M: 0 %	M: 1 %	M: 0 %
Photodegradation						

Experimental data for BOD 301C is available for all analogues identified in the selected databases. The analogues are not biodegradable having data <60%^[1].

[1] [OECD guideline for testing of chemicals 301. 1992. Ready biodegradability.](#)

The Exercise

Example 2: Background

- Micronucleus positive results for p-benzoquinone [CAS 106-51-4] are reported in the literature^[1-3] and are also available in the Toolbox.
- Therefore, it could be assumed that the chemicals producing p-benzoquinone (p-BQ) as a metabolite would also elicit positive effects in the micronucleus test.
- The hydroquinone/quinone molecular transformations are reproduced by the *in vivo* rat simulator available in the Toolbox.

[1] J.L. Bolton, M.A. Trush, T.M. Penning, G. Dryhurst, T.J. Monks. Role of quinones in toxicology. Chem. Res. Toxicol., 13 (2000), pp. 135-160

[2] J.L. Bolton, J.A. Thompson. Formation, biological targets, and resulting toxicity of quinone methides. S.E. Rokita (Ed.), Reactive intermediates in chemistry and biology: Quinone methides, Wiley (2009), pp. 329-356

[3] R. Ciranni, R. Barale, A. Marrazzini, N. Loprieno. Benzene and the genotoxicity of its metabolites. I. Transplacental activity in mouse fetuses and in their dams. Mutat. Res., 208 (1988), pp. 61-67

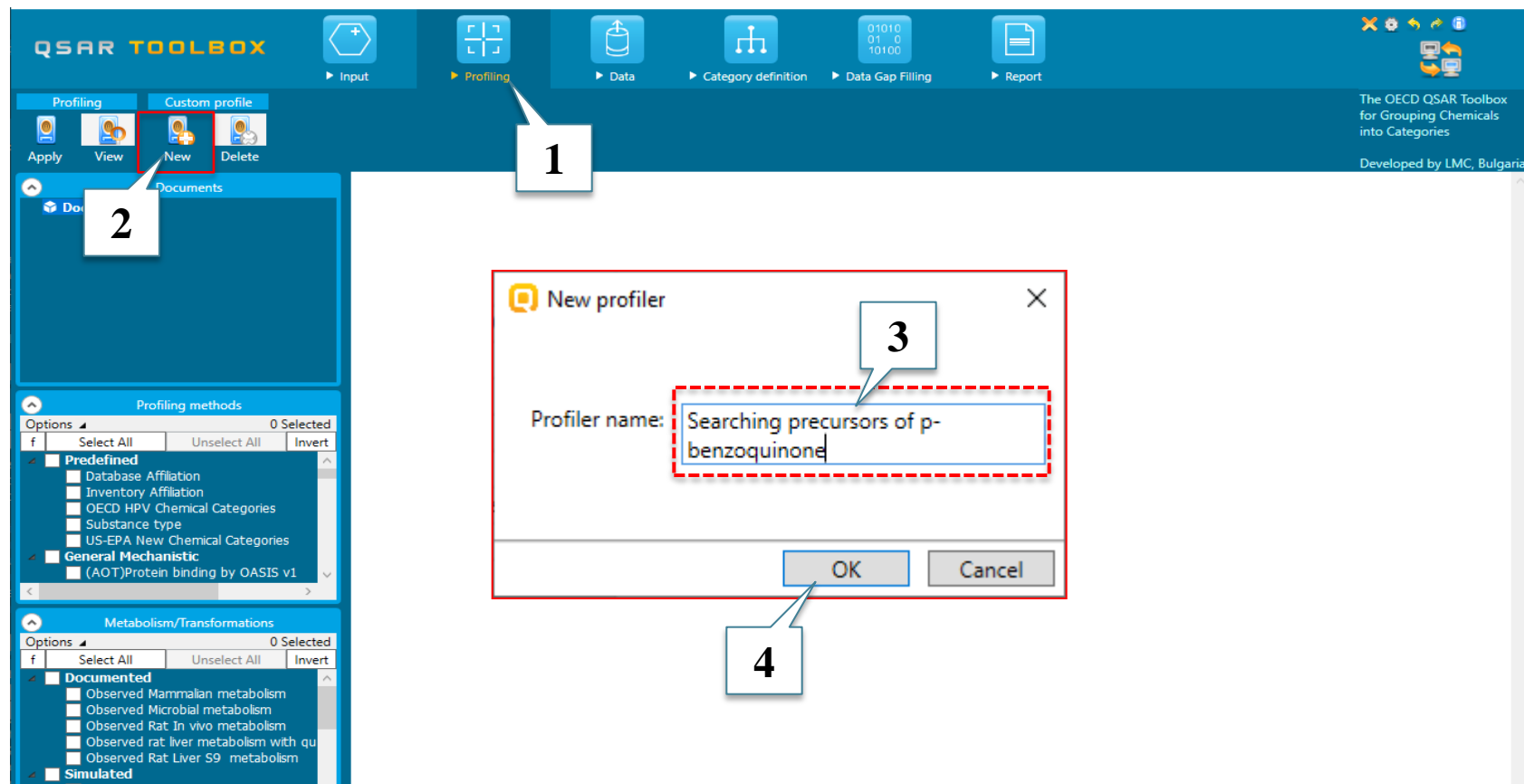
The Exercise

Example 2

- In the current example we will build a profiler that identifies chemicals producing p-BQ (i.e. precursors of p-BQ) as a result of *in vivo* rat metabolism.
- It will be illustrated how to search precursors in a specific database. The same could be done using a custom file.

Building of a new profiler

Define the name of the new profiler



1. Go to the **Profiling** module; 2. Click **New**; 3. Enter the name of the new scheme, for example **"Searching precursors of p-benzoquinone"**; 4. Confirm by click **OK**.

Building of a new profiler

Define type of new profiler

The screenshot shows the QSAR Toolbox software interface. On the left, there is a sidebar with 'Documents' (containing 'Document 1') and 'Profiling methods' (with 'Predefined' and 'General Mechanistic' categories). The main window displays a dialog box titled 'Type of scheme to create'. The dialog contains two radio button options: 'Linear or hierarchical scheme' (which is selected and highlighted with a red dashed box and a callout '1') and 'Dendroid scheme (Prioritization scheme)'. At the bottom right of the dialog are 'OK' and 'Cancel' buttons, with a callout '2' pointing to the 'OK' button.

1. Select the first option for scheme type: **Linear or hierarchical scheme**; 2. Click **OK**.

Building of a new profiler

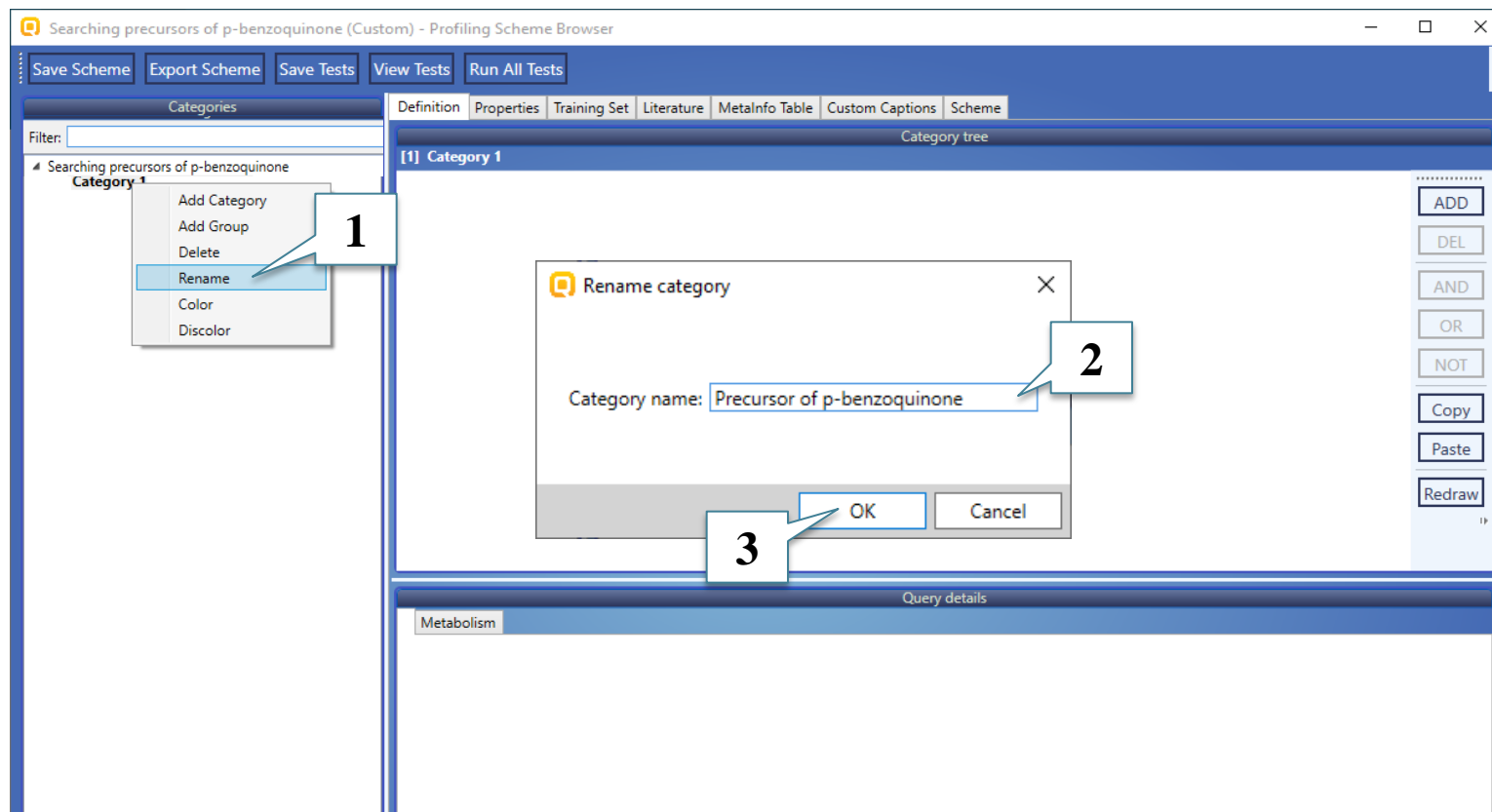
Creating a profiling category

Once the scheme is created, it will be defined:

- the name of the profiling category, which will appear on the data matrix;
- the structure which we will searching for (i.e. p-BQ);
- the simulator, which will be used to simulate the transformations of the chemicals (i.e. *in vivo Rat metabolism simulator*).

Building of a new profiler

Name a profiling category



1. Right click on "Category 1" and select to **Rename** it; 2. Put a name that will appear on the data matrix when profile, e.g. *Precursor of p-benzoquinone*; 3. Confirm by **OK**.

Building of a new profiler

Define the structural criteria

Searching precursors of p-benzoquinone (Custom) - Profiling Scheme Browser

Save Scheme Export Scheme Save Tests View Tests Run All Tests

Categories

Filter:

- Searching precursors of p-benzoquinone
- Precursor of p-benzoquinone

Definition Properties Training Set Literature MetaInfo Table Custom Captions Scheme

Category tree

[1] Category 1

Data Query ADD

Parameter Query DEL

QSAR Domain Query

QSAR Query AND

Reference Query OR

Similarity Query NOT

Structure Query

Copy

Paste

Redraw

Query details

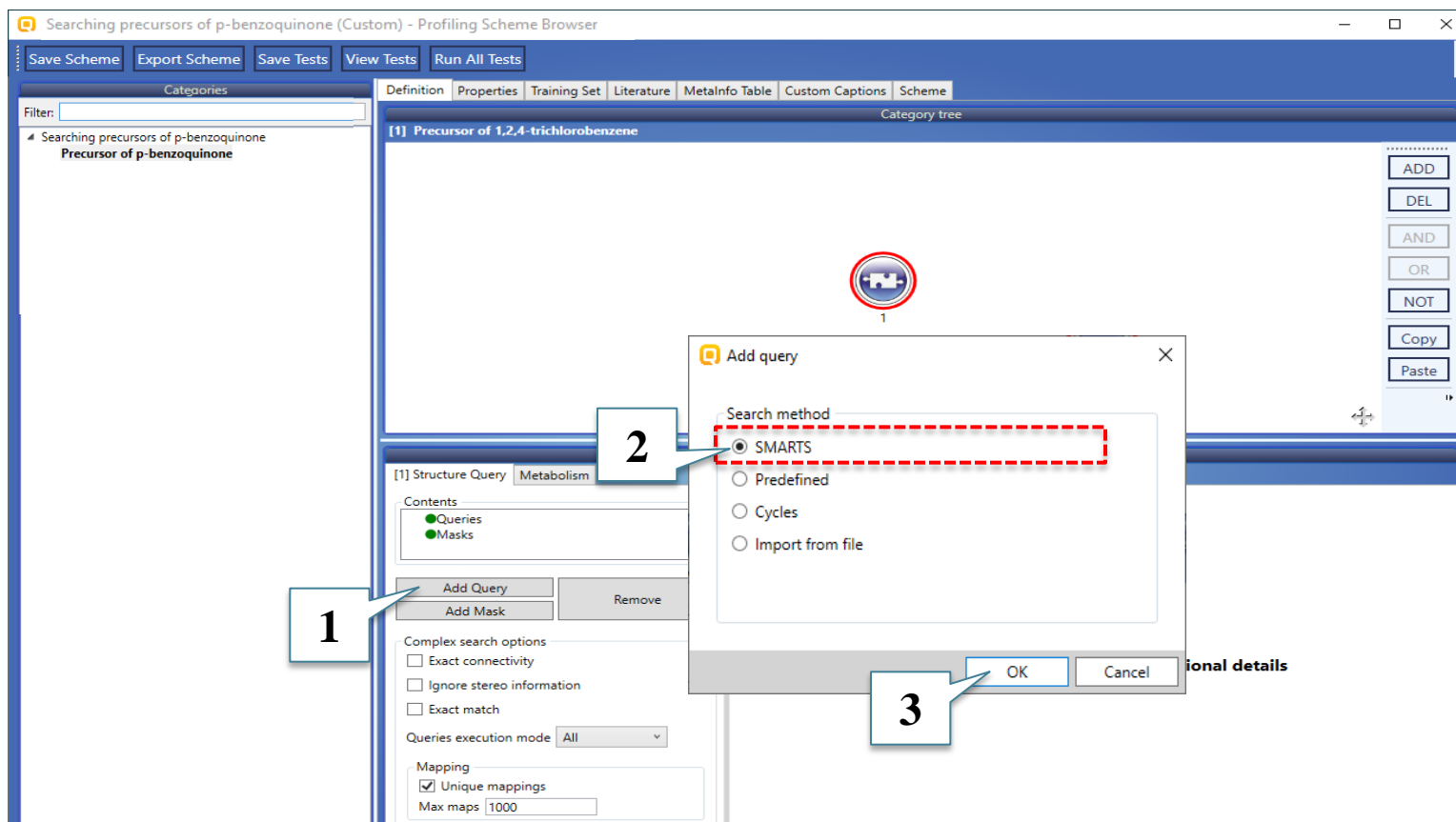
Metabolism

First we need to define the structural criteria, i.e. the structure of the metabolite which we will searching for. The structural criteria will be defined using the SMARTS language.

1. Click on the **ADD** button; 2. Select **Structure query**.

Building of a new profiler

Define the structural criteria



1. Click on the **Add Query** button; 2. Select to add SMARTS; 3. Confirm by **OK**.

Building of a new profiler

Define the structural criteria

1. Click on the created SMARTS query in order to define it. The SMARTS editor will appear;
2. Click **Edit** to define the structure.

Building of a new profiler

Define the structural criteria

Once the SMART editor is opened, we need to draw the structure which we will searching for (i.e. p-BQ).

Click on the cleaning tool (1) in order to clean everything. Select the cyclohexane template (2). Use the drawing tool (3a) to draw two bonds to the scaffold (3b) and the double bonds (second click over a bond converts it to double bond). Then click the oxygen symbol (i.e. **O**) (4a) and click on the bonds' ends (4b). Finally click **OK** (5).

Building of a new profiler

Define the structural criteria

The structural criteria is defined. Now we need to specify the metabolic simulator, which will be used for searching precursors of p-BQ.

1

2

1

1

The structures of p-benzoquinone is defined(1) Then go to the **Metabolism** tab (2).

Building of a new profiler

Define the metabolic criteria

5

Searching precursors of 1,2,4-trichlorobenzene (Custom) - Profiling Scheme Browser

Save Scheme Export Scheme Save Tests View Tests Run All Tests

Categories

Filter:

Searching precursors of p-benzoquinone
Precursor of p-benzoquinone

Definition Properties Training Set Literature MetaInfo Table Custom Captions Scheme

[1] Precursor of 1,2,4-trichlorobenzene

Category tree

ADD
DEL
AND
OR
NOT
Copy
Paste
Redraw

In the current example we search for chemicals that would be positive in the micronucleus test based on their positive metabolite. As the micronucleus is *in vivo* test, we will use the *in vivo* rat metabolic simulator for the purpose.

1

2

3

4

[1] Structure Query - Metabolism

Simulator: in vivo Rat metabolism simulator

Input

☒ Parent

☐ Metabolites

☐ All

Process mode

☒ Any

☐ All

☐ Accumulative

Output

☐ Parent

☐ Metabolites

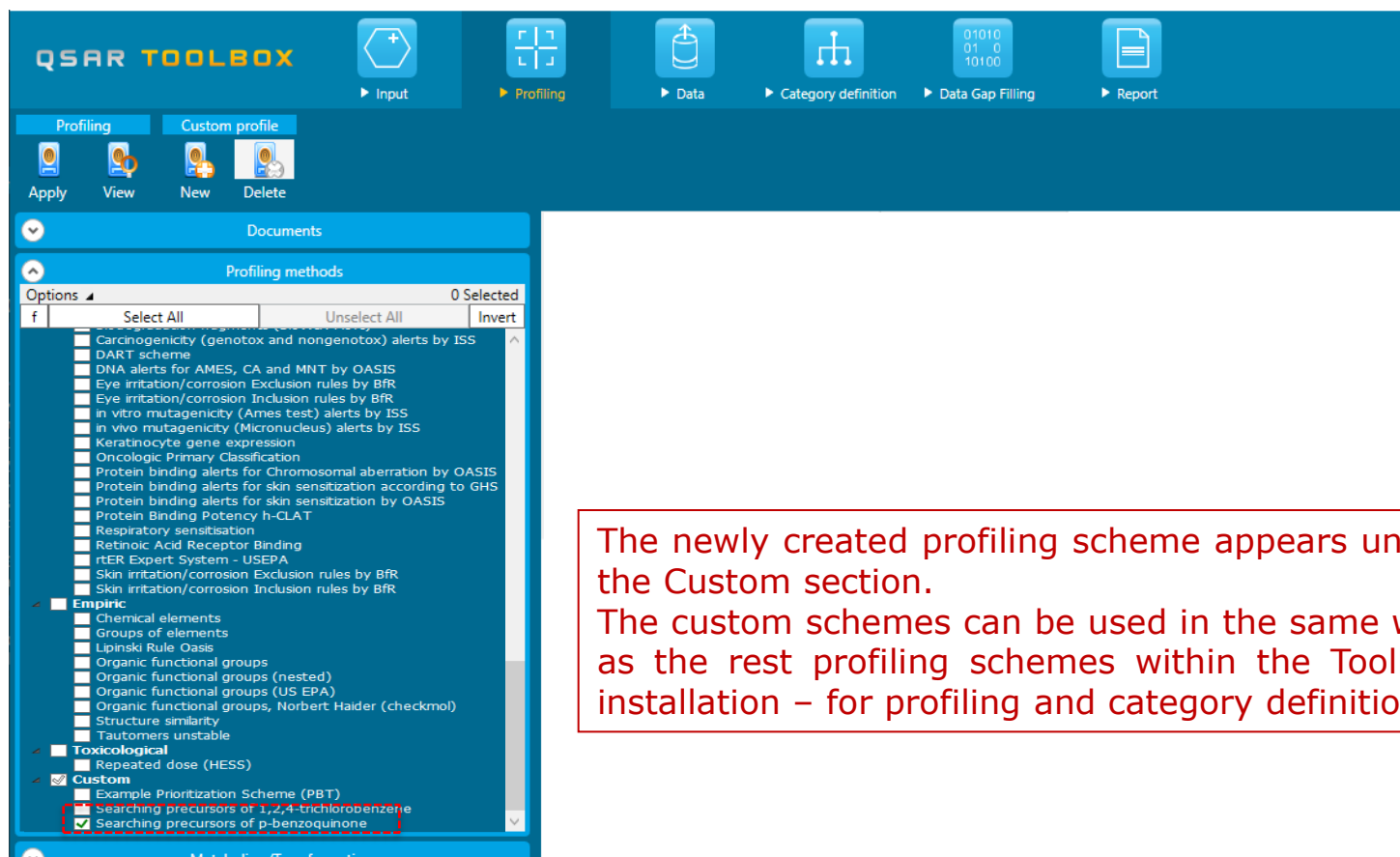
☒ All

When no metabolites: UseParent

Select the ***in vivo Rat metabolism simulator*** from the drop-down menu (1). This simulator will be applied on the chemicals as parents (i.e. *Input* >> *Parents* (2)) and we will search within the whole package – parent and produced metabolites (i.e. *Output* >> *All* (3)); select Any process mode (4). Finally, click on **Save Scheme** (5).

Building of a new profiler

Custom profiler location



The newly created profiling scheme appears under the Custom section.

The custom schemes can be used in the same way as the rest profiling schemes within the Toolbox installation – for profiling and category definition.

Building of a new profiler

Custom profiler usage

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes options like Document, Single Chemical, Chemical List, Search, and Target Endpoint. Below this is a toolbar with icons for New, Open, Close, Save, CAS#, Name, Structure, Composition, Select, ChemIDs, Database, Inventory, List, Substructure (SMARTS), Query, and Define. The 'Database' button is highlighted with a red box and a callout '1'. A 'Select database' dialog box is open, showing a list of databases under the 'Human Health Hazards' section. The 'Micronucleus ISSMIC' database is highlighted with a red box and a callout '2'. The 'OK' button is highlighted with a callout '3'.

Options: 1 Selected

Select All Unselect All Invert About Options

Human Health Hazards

- Acute Oral toxicity DB
- ADME Database
- Bacterial mutagenicity ISSSTY
- Biocides and plant protection ISSBIOC
- Carcinogenic Potency Database (CPDB)
- Carcinogenicity&mutagenicity ISSCAN
- Cell Transformation Assay ISSCTA
- Dendritic cells COLIPA
- Developmental & Reproductive Toxicity (DART)
- Developmental toxicity database (CAESAR)
- Developmental toxicity ILSI
- ECHA REACH
- ECOTOX
- Eye Irritation ECETOC
- Food TOX Hazard EFSA
- GARD Skin sensitization
- Genotoxicity & Carcinogenicity ECVAM
- Genotoxicity OASIS
- Genotoxicity pesticides EFSA
- Human Half-Life
- Keratinocyte gene expression Givaudan
- Keratinocyte gene expression LuSens
- Micronucleus ISSMIC**
- Micronucleus OASIS
- MUNRO non-cancer EFSA
- REACH Skin sensitisation database (normalised)
- Receptor Mediated Effects
- Rep Dose Tox Fraunhofer ITEM
- Repeated Dose Toxicity HESS
- Rodent Inhalation Toxicity Database
- Skin Irritation
- Skin Sensitization
- Skin sensitization ECETOC

OK Cancel

Precursors of p-BQ will be searched in a specific file. *Micronucleus ISSMIC* database will be used in the current example.

1) Click on the **Database** button*. List with all available databases will appear;
2) Select the *Micronucleus ISSMIC* database located under the *Human Health Hazards* section; 3) Confirm by **OK**.

Note: If you want to check for precursors in a custom file, you can use the *List* button to load a preliminary prepared .txt or .smi file

Building of a new profiler

Custom profiler usage

1) Go to the **Profiling** module; 2) Select the newly created profiler; 3) Click on the **Apply** button.

Building of a new profiler

Custom profiler usage

Once the chemicals are profiled by the created custom profiler, the results appear on the data matrix:

- "N/A" result for the chemicals that do not give p-BQ after in vivo metabolic activation
- "*Precursor of p-benzoquinone*" results for the chemicals that metabolize to p-BQ.

Building of a new profiler

Custom profiler usage

The screenshot illustrates the workflow for building a new profiler and using a custom profiler. The main window shows the 'Profile Statistic' table with the following data:

#	Category	Count	%
1	(N/A)	551	97.87
2	Precursor of p-benzoquinone	12	2.13

The secondary window, titled 'Precursor of p-benzoquinone', shows 12 chemical structures arranged in a grid. The 'Add in' button is highlighted, indicating the next step in the process.

1) Apply right click over the name of the custom profiler that appears in the grey field and select **Profile statistic**. It could be seen that the profiler identifies 12 chemicals among the chemicals in the data matrix; 2) Click on the *Precursor of p-benzoquinone* row; 3) Add the chemicals in a new document.

Building of a new profiler

Custom profiler usage

The screenshot displays the QSAR Toolbox interface. The top menu bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The left sidebar contains panels for 'Documents', 'Profiling methods', and 'Metabolism/Transformations'. The 'Documents' panel shows a list of documents, with 'Precursor of p-benzoquinone' selected. The 'Profiling methods' panel shows a list of methods, with 'Searching precursors of p-benzoquinone' selected. The 'Metabolism/Transformations' panel shows a list of transformations, with 'Observed Mammalian metabolism' selected. The main window displays a table of profiling results. The table has columns for chemical structures and their corresponding profiling results. The row for 'Precursor of p-benzoquinone' is highlighted with a red box and labeled '2'. The 'Documents' panel also has an 'Apply' button highlighted with a red box and labeled '1'.

1) Click on **Apply** to profile again the chemicals; 2) Double click on a profiling result to explain it.

Building of a new profiler

Explaining profiling results

In the new opened window it could be seen the target and all metabolites generated by the selected metabolic simulator (i.e. in vivo rat metabolic simulator in this case).

The status for each molecule is also shown:

- *Satisfied* – means that the corresponding chemical meet the criteria defined in the profiler;
- *Not Satisfied* – means that the corresponding chemical does NOT meet the criteria defined in the profiler.

Building of a new profiler

Explaining profiling results

The screenshot shows the 'Explaining' window in the QSAR Toolbox, titled 'Explanation for: Searching precursors of p-benzoquinone -> Precursor of p-benzoquinone'. The window has tabs for Definition, Properties, Training Set, Literature, MetaInfo Table, Custom Captions, and Scheme. The 'MetaInfo Table' tab is active, showing a table with columns for Chemical and Status. The table lists several chemical structures and their corresponding status.

Chemical	Status
<chem>Oc1ccccc1</chem> (Phenol)	NotSatisfied
<chem>O=C1C=CC(=O)C=C1</chem> (p-Benzoquinone)	Satisfied
<chem>Oc1ccc(O)cc1</chem> (Catechol)	NotSatisfied
<chem>Oc1ccccc1O</chem> (Resorcinol)	NotSatisfied

Annotations on the image:

- Parent chemical:** Points to the Phenol structure (NotSatisfied).
- the metabolite satisfying the defined criteria (i.e. the structure of p-BQ):** Points to the p-Benzoquinone structure (Satisfied).
- in vivo metabolites:** Points to the Catechol and Resorcinol structures (NotSatisfied).

The right side of the image shows a grid of chemical structures, likely representing the results of a search or a comparison of different chemical groups.

Congratulations

- Now you know how to:
 - create a new profiler taking into account the (a)biotic activation of the chemicals;
 - use the custom profiler for searching precursors of a target in selected databases;
 - use the custom profiler for searching precursors of a target in a specific file (database);
 - explain the results of the custom profiler.
- Continual use of the Toolbox will increase your skills.