QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

OECD QSAR Toolbox v.4.4.1

Category elements for assessing category consistency

Outlook

- **Background**
- Keywords
- Objectives
- Aim
- Category consistency
- Examples

Background

- This is a step-by-step presentation designed to introduce to the user the category elements applicable for assessing category consistency

- Further to illustrate some simple examples of using category consistency elements targeting Fate and Human toxicity endpoints

Outlook

- Background
- **Keywords**
- Objectives
- Aim
- Category consistency
- Examples

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) endpoint-specific (e.g. in vitro in vitro mutagenicity (Ames test) alerts by ISS) profilers and empiric (e.g. Organic functional groups).

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 during Category definition process.

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)

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

Outlook

- Background
- Keywords
- **Objectives**
- Aim
- Category consistency
- Examples

Objectives

This presentation demonstrates a number of functionalities of the Toolbox:

- Including category elements in the process of category consistency assessment.

- Generation of the category report including category consistency elements.

Outlook

- Background
- Keywords
- Objectives
- **Aim**
- Category consistency
- Examples

Aim

- To introduce and make the user familiar with:
  - the similarity elements (physicochemical parameters, profilers) used in category consistency assessment;
  - the default settings of category consistency elements for reporting.

Outlook

- Background
- Keywords
- Objectives
- Aim
- Category consistency
  - **Overview**
  - Adopted concept
  - Elements of category consistency
  - Implementation in Toolbox
- Examples

Overview
What is a category ?

- Based on REACH regulation [1]:

  "Substances whose physicochemical, toxicological and ecotoxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity may be considered as a group, or 'category' of substances."

[1] REACH regulation: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02006R1907-20190702&from=EN

Outlook

- Background
- Keywords
- Objectives
- Aim
- Category consistency
  - Overview
  - **Adopted concept**
  - Elements of category consistency
  - Implementation in Toolbox
- Examples

Category consistency
Adopted concept

- The read-across methodology is based on the premise that similar chemical structures (forming the category) are expected to have similar biological/toxicological action with respect to the endpoint
- Category consistency is endpoint specific;
- Category consistency check (using category elements) can be applied to:
  - a list with chemicals with a defined target endpoint - there is no need to enter into the data gap filling (i.e. it could be done without accepted prediction)
  - analogues developed during the read-across workflow (in the data gap filling)

Outlook

- Background
- Keywords
- Objectives
- Aim
- Category consistency
  - Overview
  - Adopted concept
  - **Elements of category consistency**
  - Implementation in Toolbox
- Examples

Category consistency elements

Three layers of information are considered important in analyzing consistency of the chemical categories:

- **Physicochemical similarity** (e.g. *MW, VP, logKow etc*.) and experimental data
- **Structural similarity** (e.g. *Organic functional groups (OFG)*)
- **Mechanistic similarity** (e.g. *Protein binding alerts for skin sensitization* accounting for (a)biotic activation in case the target endpoint is skin sensitization)

The layers mentioned above are implemented as elements of the category consistency (see next few slides)

Outlook

- Background
- Keywords
- Objectives
- Aim
- Category consistency
  - Overview
  - Adopted concept
  - Elements of category consistency
  - **Implementation in Toolbox**
- Examples

Category consistency

Implementation in Toolbox

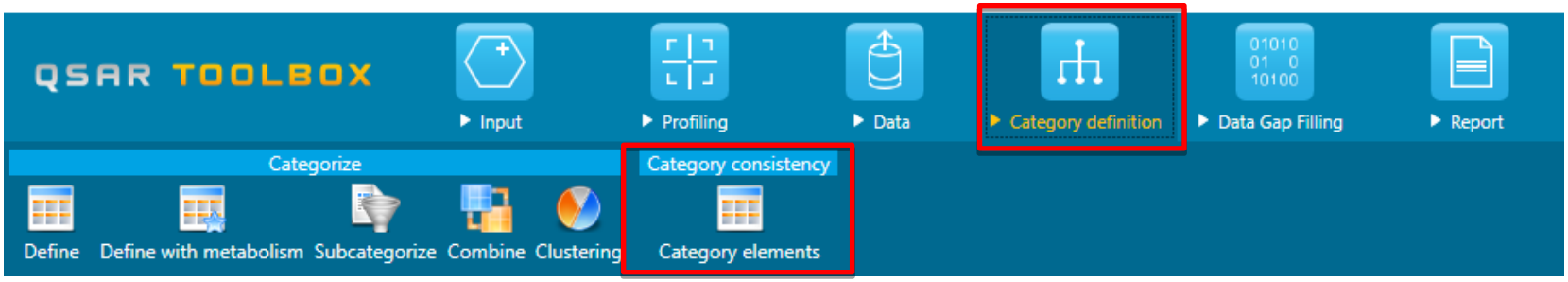
- New functionality related to Category consistency assessment has been implemented in version 4.2
- It is located in the "Category definition" module and provides default selection of category consistency elements as follows:
  - **Physicochemical similarity** – does not depend on the endpoint
    - ***Parameters* -** Boiling point, logKow, Molecular weight, Vapor pressure, Water solubility
    - ***Physico-chemical exp. data* -** Boiling point, logKow, Vapor pressure, Water solubility
  - **Structural similarity** – does not depend on the endpoint
    - Organic functional groups
    - Structure similarity
  - **Mechanistic similarity** – depends on the pre-defined target endpoint

  The profilers and metabolism simulators highlighted as "suitable" for the target endpoint are selected by default (e.g. for target endpoint Skin sensitization/LLNA/EC3 all *Protein binding alerts for skin sensitization* profilers along with the simulators for *Autoxidation* and *Skin metabolism* are selected by default)
  - **(Eco)toxicological experimental data** – depends on the pre-defined target endpoint
- Different selection of the category consistency elements than the default one is allowed

Category consistency

Implementation in Toolbox

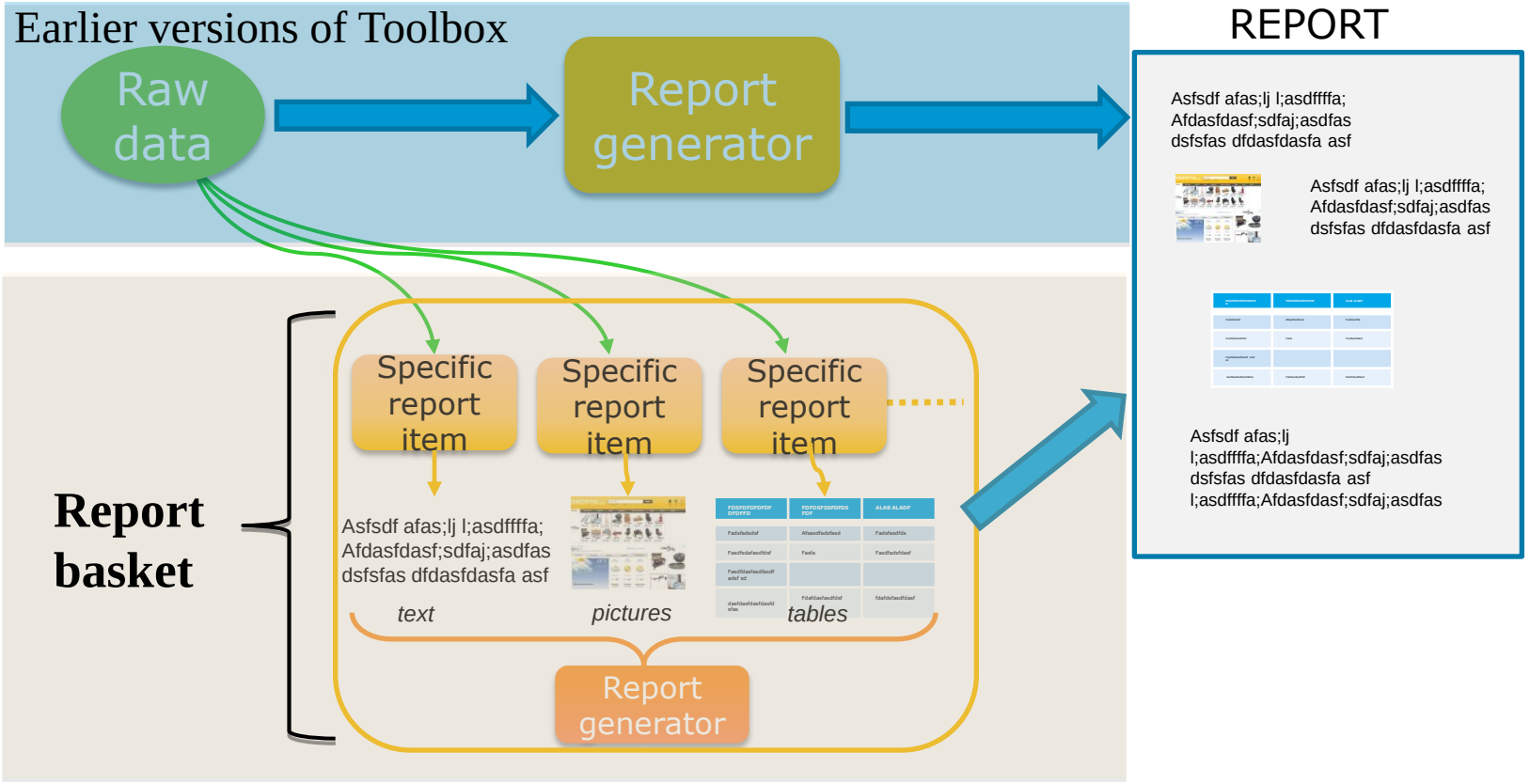
- "Category consistency" is located in *Category definition* module
- It allows to filter and collect physico-chemical properties, structural similarity, profiling results accounting for metabolic activation and experimental data
- All selected category elements are automatically stored in the so-called *"Report basket"*, which could be added later on in the report (overview of the Report basket is provided on next slide).
- Evaluation of category consistency is possible if the target endpoint is defined



Category consistency

Report basket - background

- The specific report items are collected during the workflow or from external modeling sources.
- All items are stored in the "*Report basket*" and can be used in the report to support or justify the consistency of a category.



Category consistency

Report basket – examples of the items stored here

Example of the specific report items and their description

#	Specific report item	Related to:	Options	Actions which trigger saving results in the basket	Result
1	Target information	Target chemical		Target is defined	Depiction of chemical; CAS; Names; SMILES
2	Purity/ Impurity	Target chemical; Category		Target or category members have compositions	Table with description of composition with depictions, quantities of each structure
3	Category members	Category		Category is defined or a list with structures is loaded/ created	Table with depiction; CAS; Name; SMILES of all chemicals
4	Physical chemical similarity	Category	Select 2D parameters and/or physical chemical properties	Applying Category elements	Table with calculated 2D parameters; Table with experimental physical chemical properties
5	Structural similarity	Category	Adjust structure similarity options	Applying Category elements	Table with calculated similarity and variation of similarity with other chemicals
6	Mechanistic similarity	Category	Select metabolism and profiler; Add table; add summary chart; add individual charts	Applying Category elements	Tables with generated metabolites and their profiling result; Table summarizing number of metabolites and found alerts; Graphic with distribution of alerts across metabolites
7	Common product	Metabolism		Category is defined with accounting for metabolism with map similarity option "exact" metabolite	Depiction of common compound (defined in grouping options) and used simulator

….continues on next slide

Category consistency

Report basket – examples of the items stored here

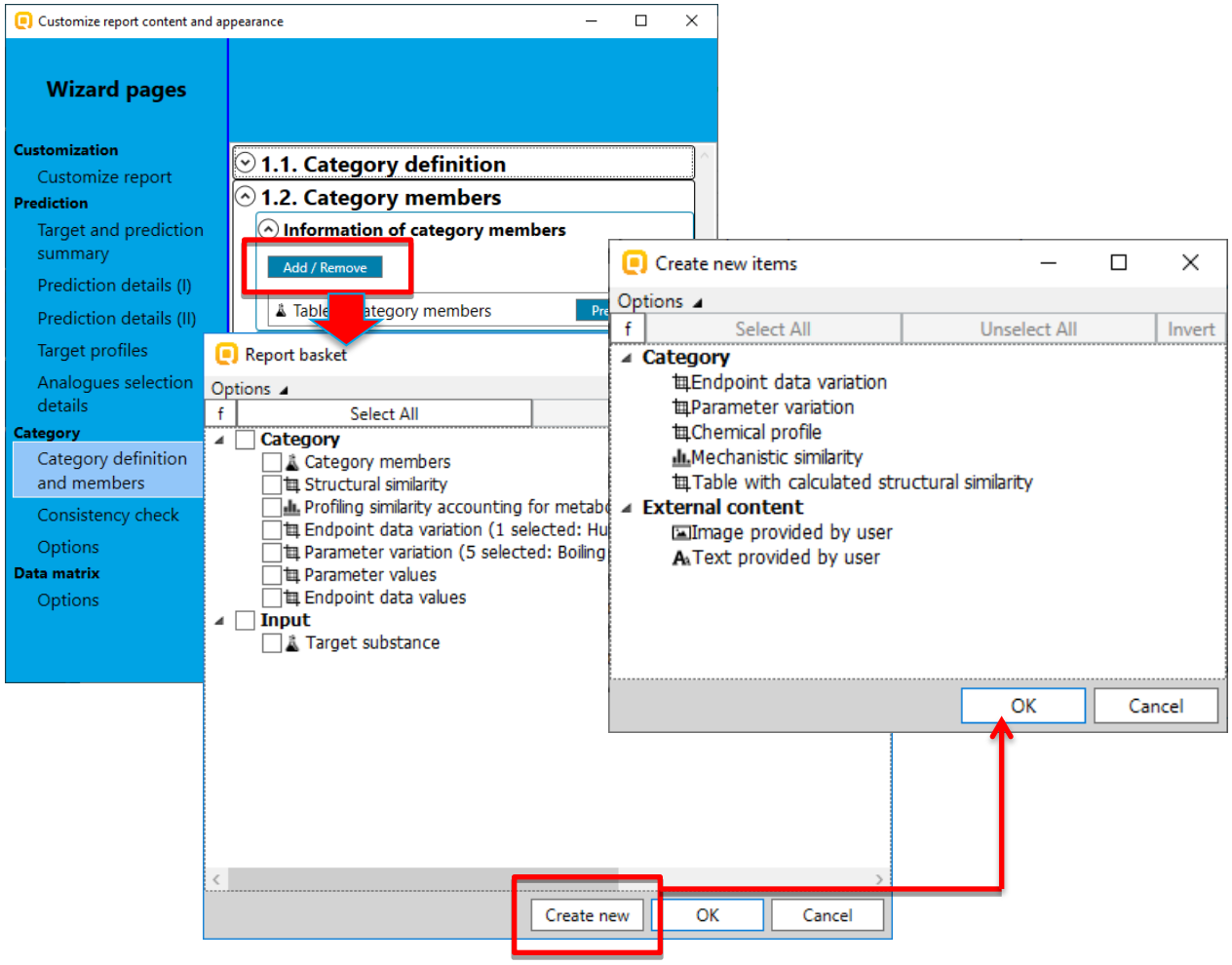
Example of the specific report items and their description

#	Specific report item	Related to:	Options	Actions which trigger saving results in the basket	Result
8	Alert performance	Category		Alert performance is calculated	Alert performance table
9	Selected endpoint data values	Data		Applying Category elements	Table with analogues used for retrieving the prediction with their observed data
10	Metabolic map	External QSAR model	Select metabolism generation options	Prediction from external model	Graphic with generated metabolic map
11	Metabolic distribution	External QSAR model	Select metabolism generation options	Prediction from external model	Graphic with quantitative distribution of metabolites
….	……..	……	…….		

etc….

Category consistency

Report basket – visualization



Report basket:

- The selected category elements are automatically stored in the so-called *"Report basket"* (the latter appears once click Add/Remove button included in the Reporting wizard)
- Additional report items (such as parameter variation for the closest analogues, text, pictures etc.) could be added via clicking on "Create new"

Examples of the report items generated in the basket are given on the forthcoming slides.

Outlook

- Background
- Keywords
- Objectives
- Aim
- Category consistency
- **Exercise**

Category consistency elements

Exercise

- Three examples will be illustrated addressing category consistency elements used in the read-across assessment of Fate and Human health endpoints:
  - *Fate endpoint* – BOD
  - *Ecotoxicity endpoint* – LC50
  - *Human health hazard* – skin sensitization (EC3)

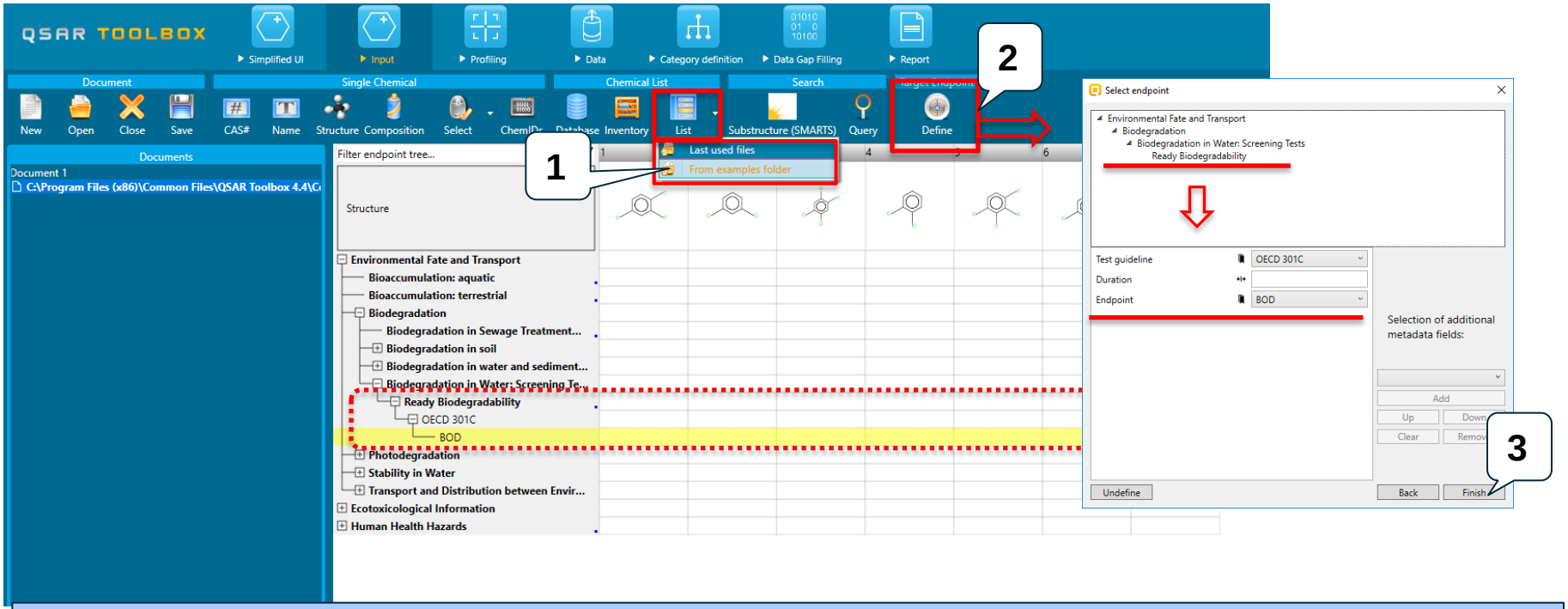
- Illustrating of category consistency elements will be applied on a predefined lists with chemicals without entering into the data gap filling

* Predefined lists with chemicals are available in the Toolbox installation folder located at: C:\Program Files (x86)\Common Files\QSAR Toolbox 4.4\Config\Examples

Category consistency elements

Example 1 – *Fate endpoint*

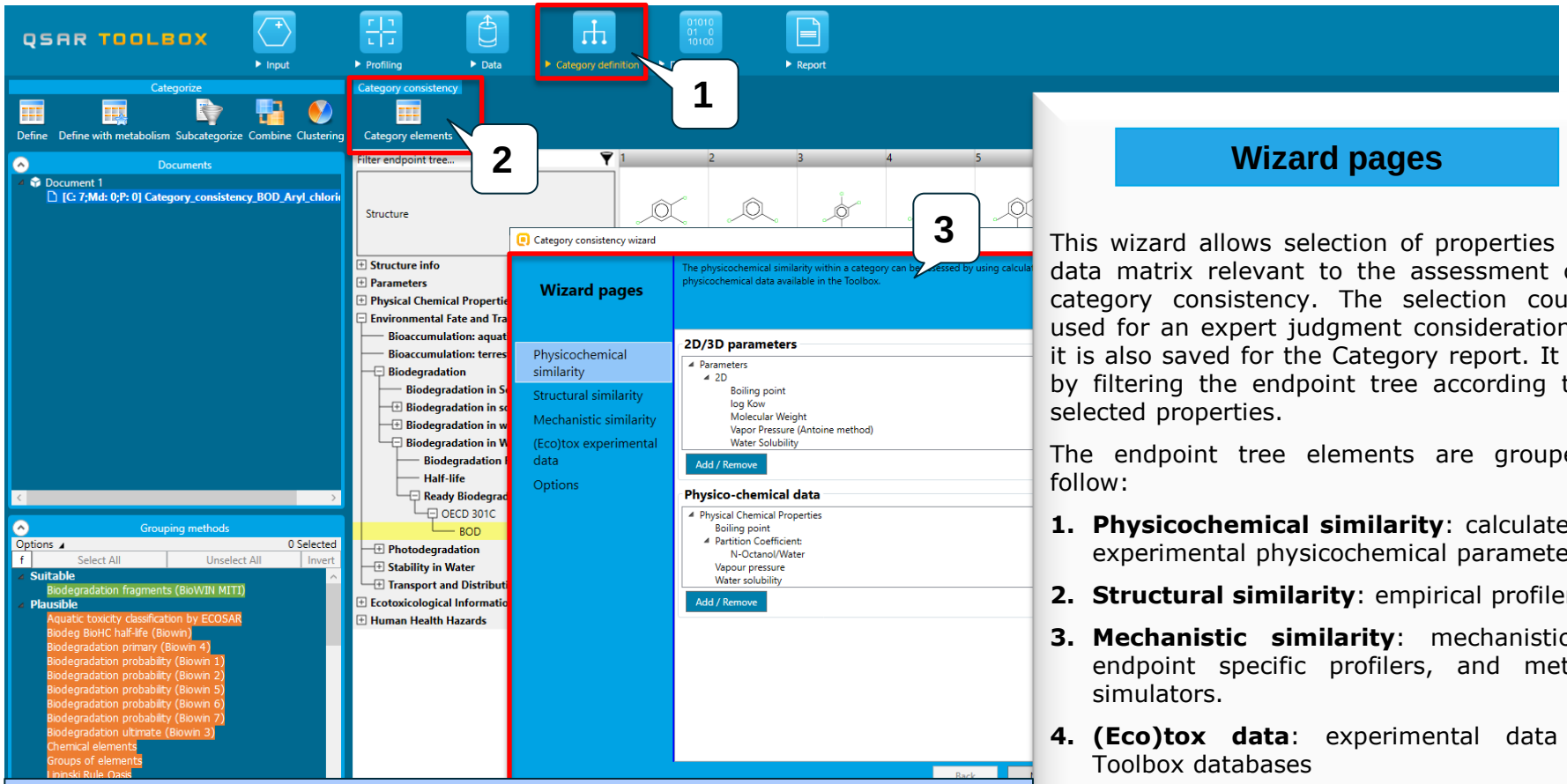
- A list of 7 aromatic halogens is submitted as a category
- The target endpoint is: Ready Biodegradability/BOD/OECD 301C



1. Open the list with chemicals **"Category_consistency_BOD_Aryl_chlorides.smi"** from the example folder
2. Click **Define,** open **Environmental Fate and Transport, Biodegradation** level and select **Ready Biodegradability**, after that select Endpoint: **BOD**; Test guideline: **OECD 301C**
3. Click **Finish**

Category consistency elements

Example 1 – *Fate endpoint*



1. Go to **Category Definition** module
2. Click ***Category elements*** button
3. The **"Category consistency wizard"** appears

Wizard pages

This wizard allows selection of properties in the data matrix relevant to the assessment of the category consistency. The selection could be used for an expert judgment considerations and it is also saved for the Category report. It works by filtering the endpoint tree according to the selected properties.

The endpoint tree elements are grouped as follow:

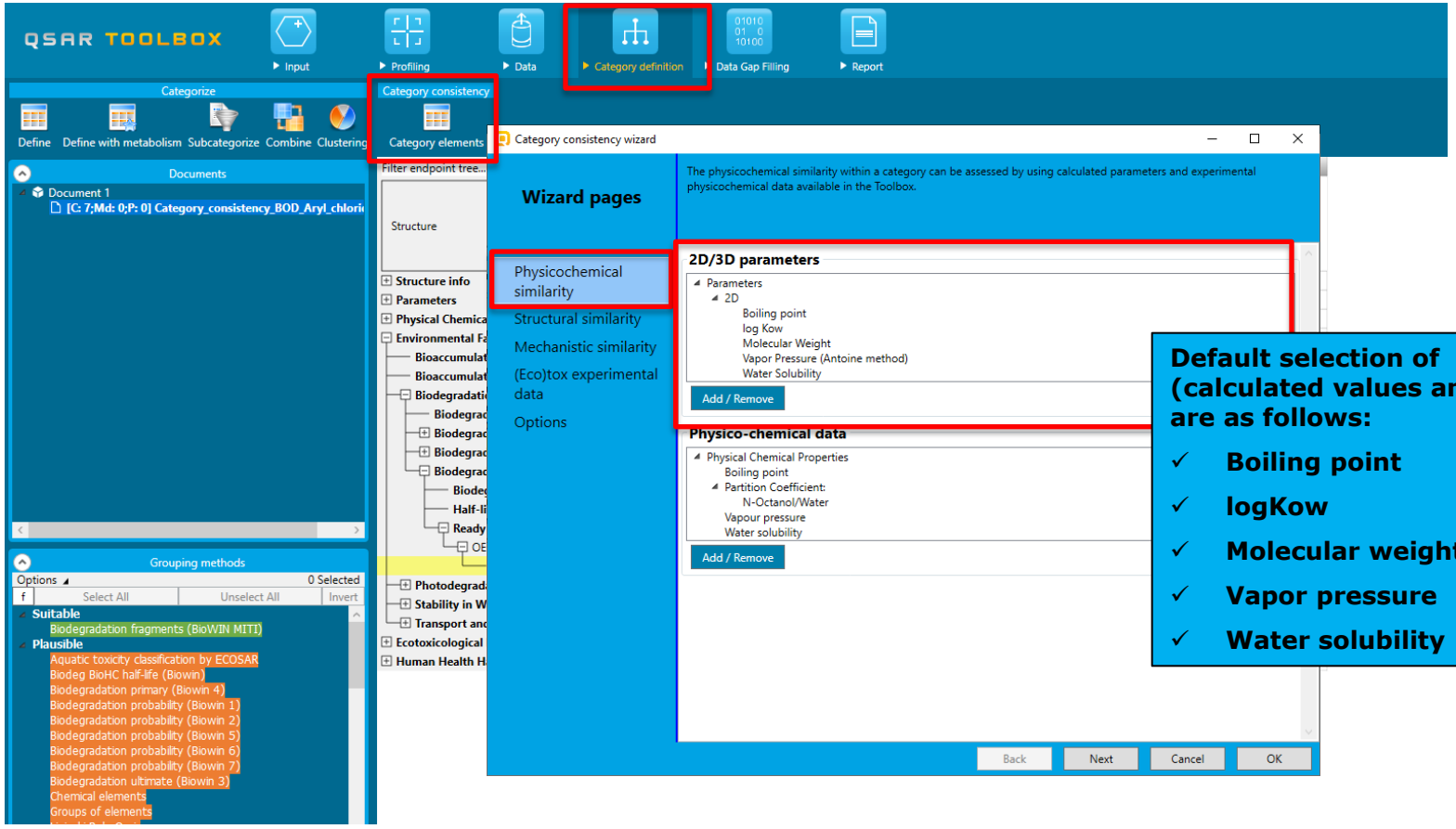
1. **Physicochemical similarity**: calculated and experimental physicochemical parameters.
2. **Structural similarity**: empirical profilers.
3. **Mechanistic similarity**: mechanistic and endpoint specific profilers, and metabolic simulators.
4. **(Eco)tox data**: experimental data from Toolbox databases

The Toolbox automatically highlights profilers and metabolic simulators relevant to the target endpoint.

Category consistency elements

Example 1 – *Fate endpoint*

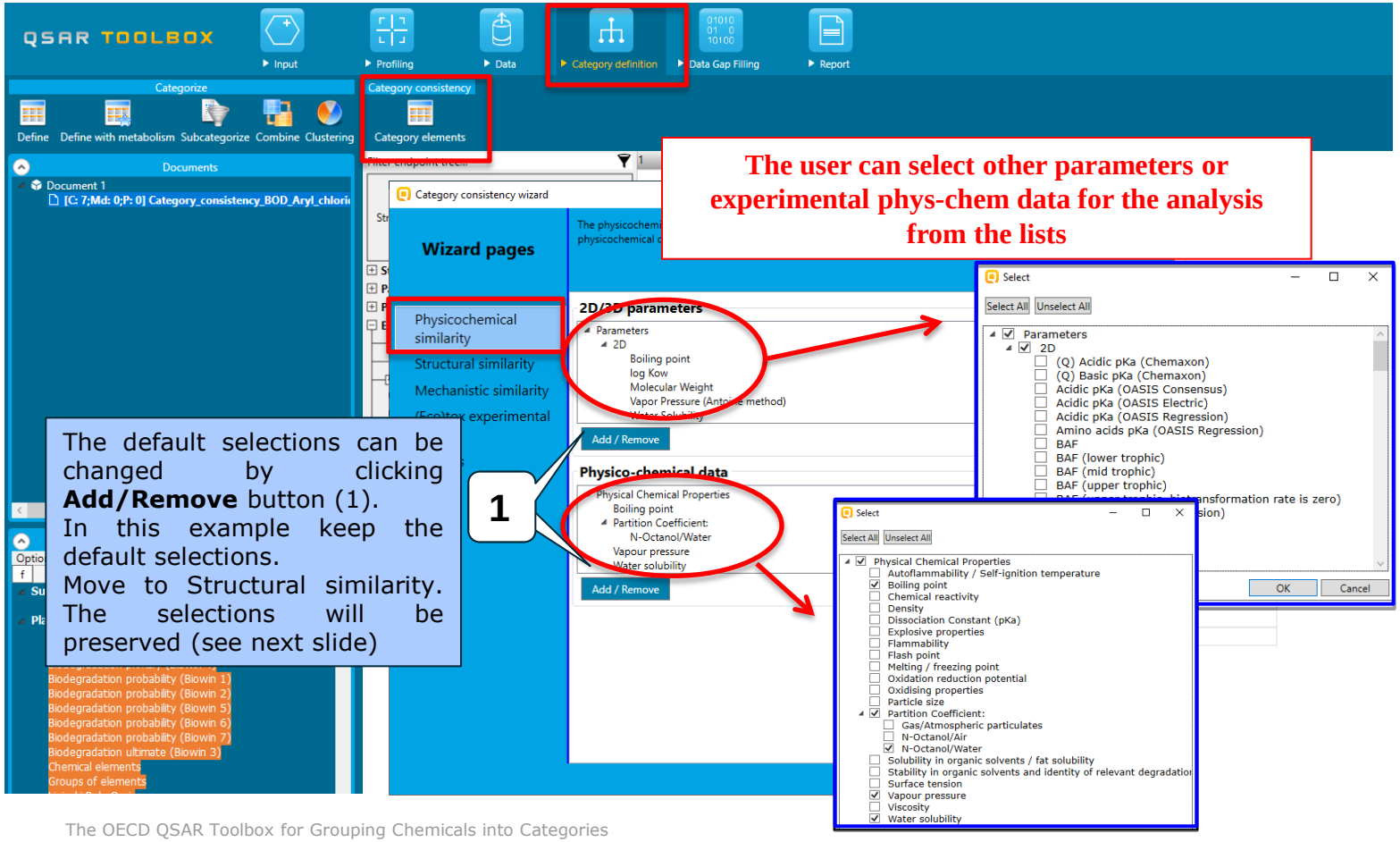
Category element: Physicochemical similarity



Category consistency elements

Example 1 – *Fate endpoint*

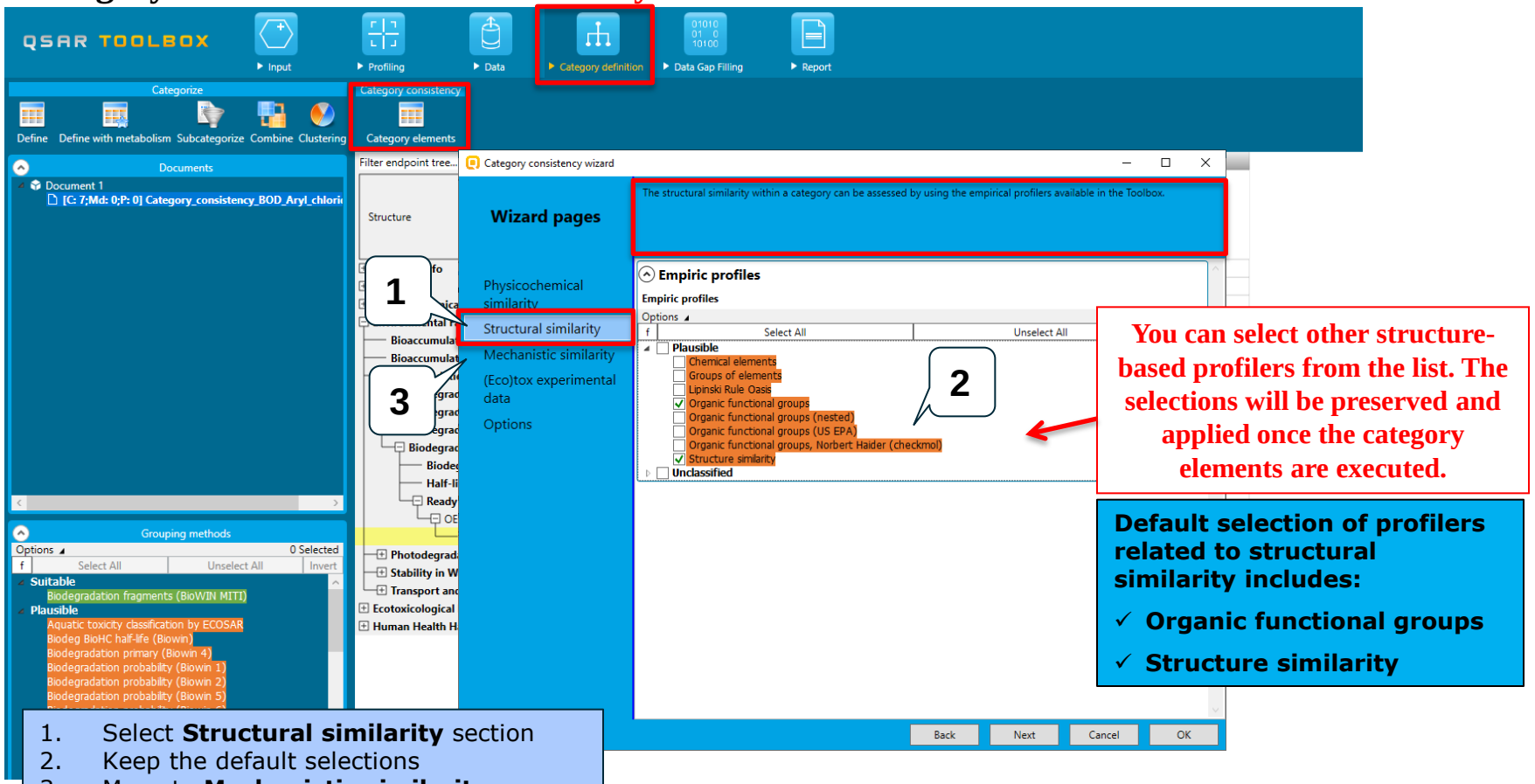
Category element: Physicochemical similarity



Category consistency elements

Example 1 – *Fate endpoint*

Category element: Structural similarity



You can select other structure-based profilers from the list. The selections will be preserved and applied once the category elements are executed.

Default selection of profilers related to structural similarity includes:

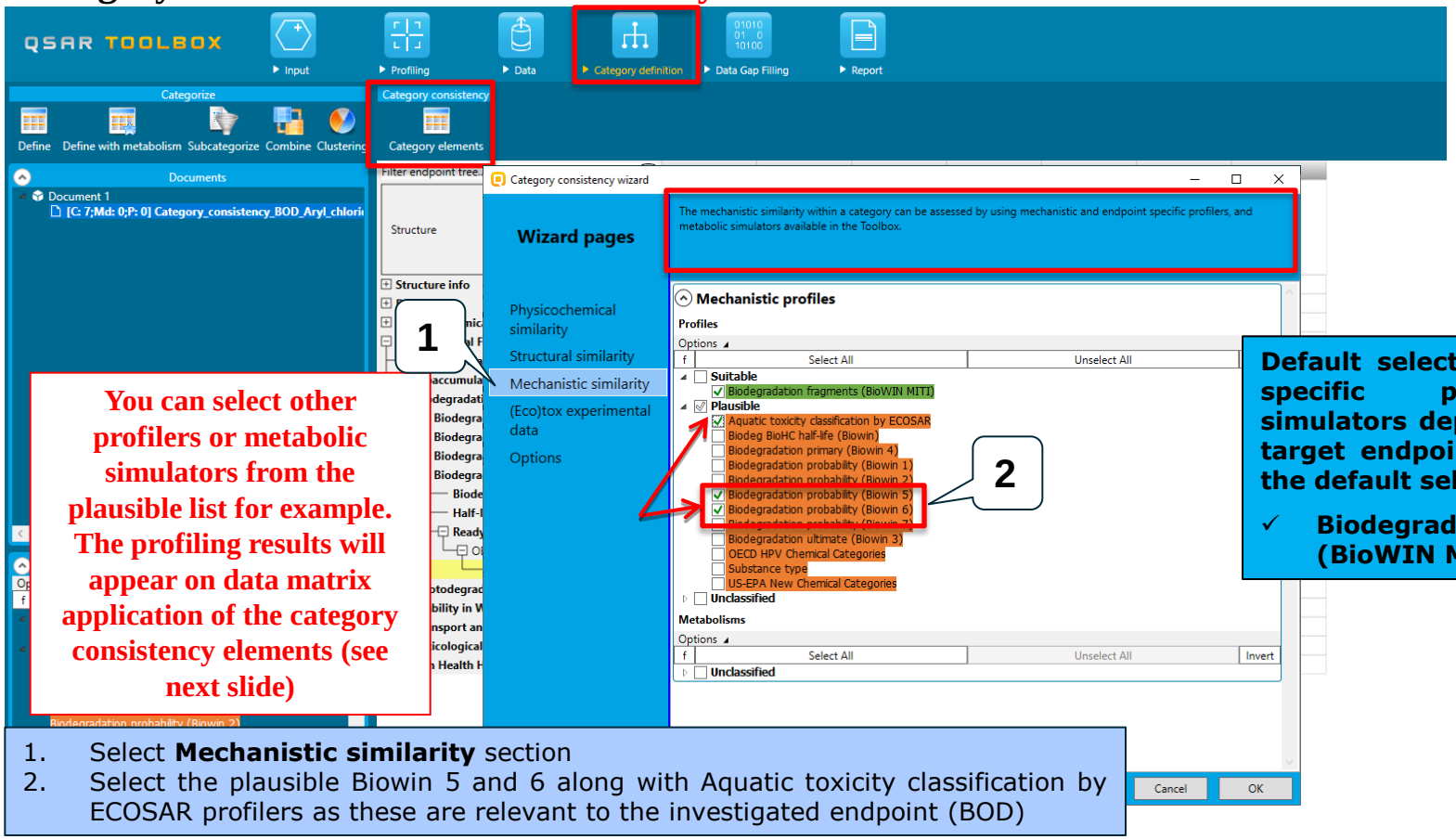
- ✓ **Organic functional groups**
- ✓ **Structure similarity**

1. Select **Structural similarity** section
2. Keep the default selections
3. Move to **Mechanistic similarity**

Category consistency elements

Example 1 – *Fate endpoint*

Category element: Mechanistic similarity



You can select other profilers or metabolic simulators from the plausible list for example. The profiling results will appear on data matrix application of the category consistency elements (see next slide)

Default selection of endpoint specific profilers and simulators depend on defined target endpoint. In this case the default selection includes:

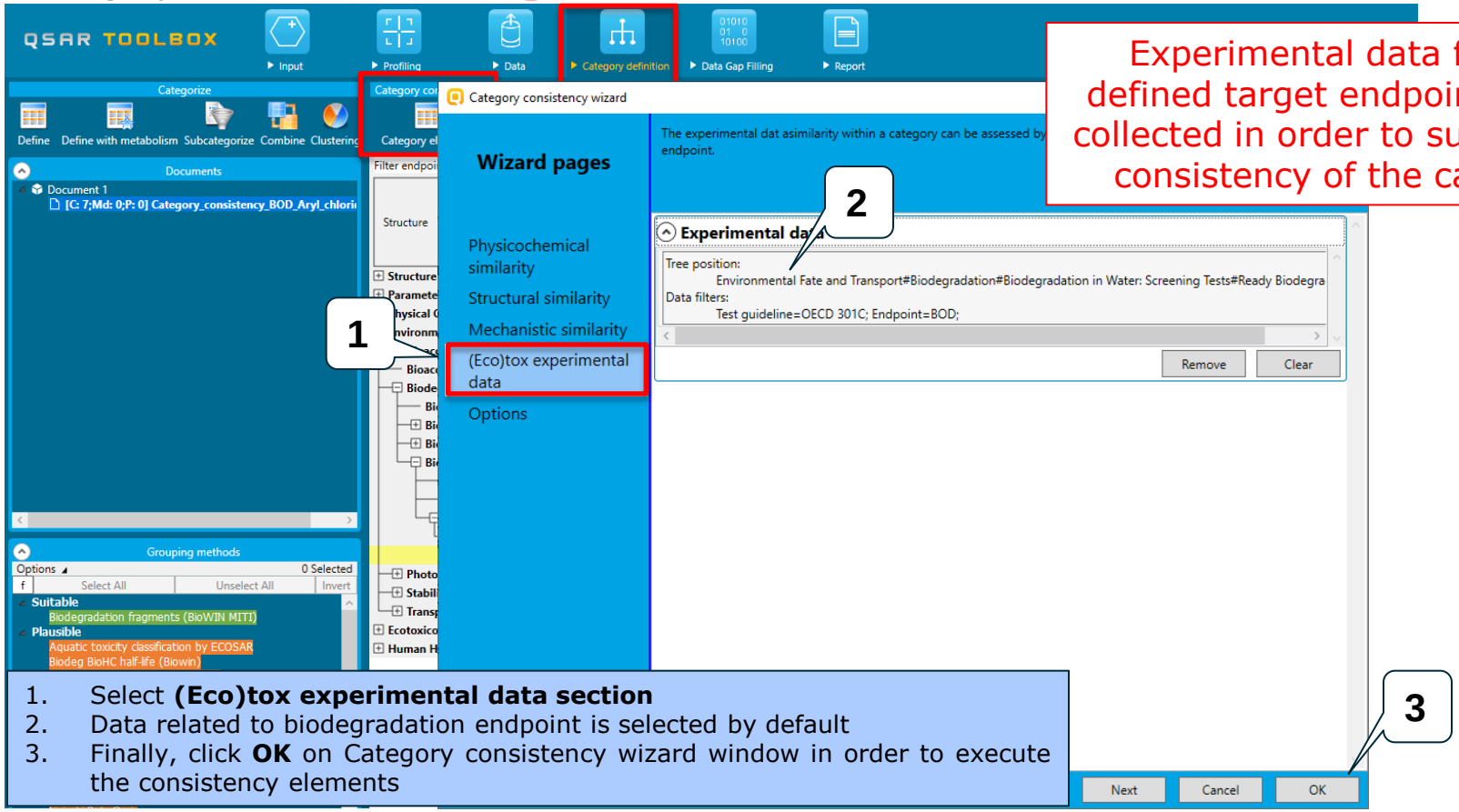
- ✓ **Biodegradation fragments (BioWIN MITI)**

1. Select **Mechanistic similarity** section
2. Select the plausible Biowin 5 and 6 along with Aquatic toxicity classification by ECOSAR profilers as these are relevant to the investigated endpoint (BOD)

Category consistency elements

Example 1 – *Fate endpoint*

Category element: (Eco)tox experimental data



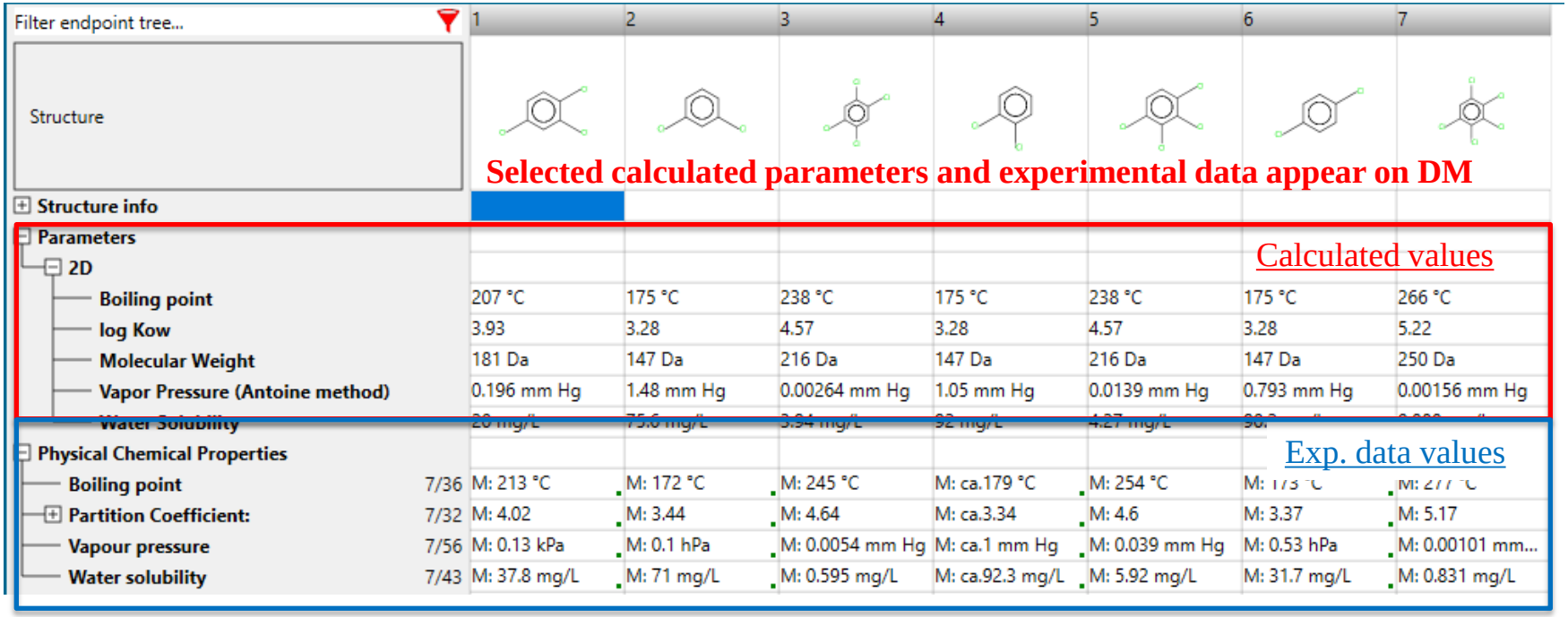
Experimental data for the defined target endpoint will be collected in order to support the consistency of the category

1. Select **(Eco)tox experimental data section**
2. Data related to biodegradation endpoint is selected by default
3. Finally, click **OK** on Category consistency wizard window in order to execute the consistency elements

Category consistency elements

Example 1 – *Fate endpoint*

Selected parameters and experimental data related to "Physicochemical similarity" (shown on slides #26-27) are calculated/extracted from databases and filtered* on the DM



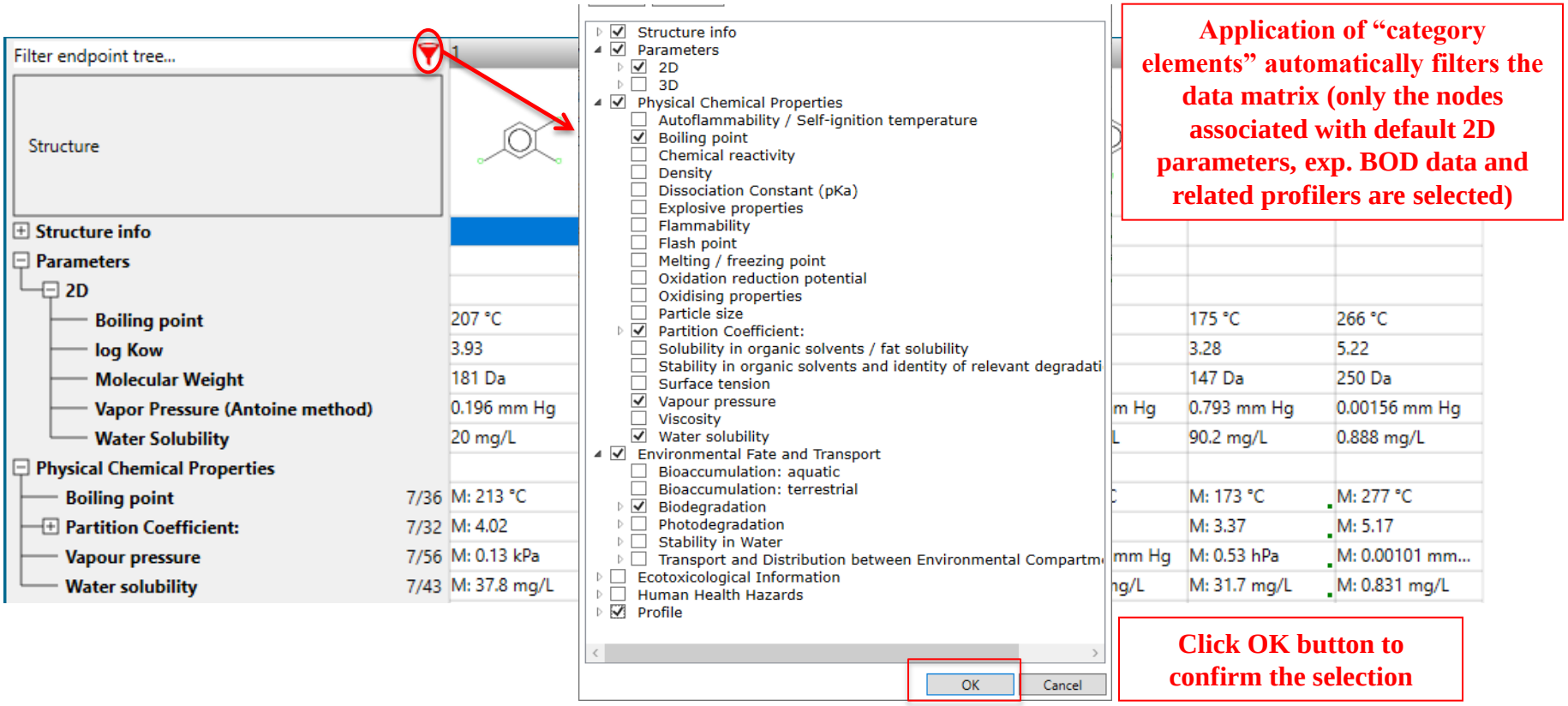
* Filtering of DM is illustrated on the next slide. More details could be found in tutorial: *Manipulation of data matrix and manual transferring of data to the target outside data gap filling module.*

Category consistency elements

Example 1 – *Fate endpoint*

Filtering of DM

Button for filtering the endpoint tree



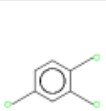
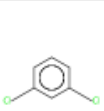
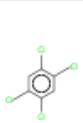
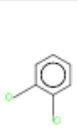
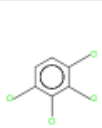
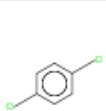
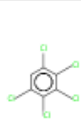
Application of "category elements" automatically filters the data matrix (only the nodes associated with default 2D parameters, exp. BOD data and related profilers are selected)

Click OK button to confirm the selection

Category consistency elements

Example 1 – *Fate endpoint*

Profiling results from the selected structure–based profilers appeared on DM (see slide #28)

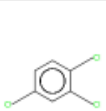
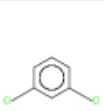
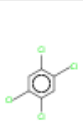
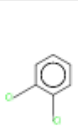
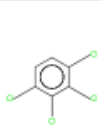
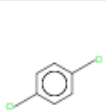
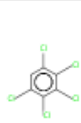
Filter endpoint tree...		1	2	3	4	5	6	7
Structure								
Structure info								
Parameters								
Physical Chemical Properties	7/167	M: 37.8 mg/L	M: 71 mg/L	M: 0.595 mg/L	M: ca.92.3 mg/L	M: 5.92 mg/L	M: 31.7 mg/L	M: 0.831 mg/L
Environmental Fate and Transport	6/6		M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %
Profiling								
Endpoint Specific								
Empiric								
Organic functional groups		Aryl Aryl halide	Aryl Aryl halide	Aromatic perhal... Aryl Aryl halide	Aryl Aryl halide	Aromatic perhal... Aryl Aryl halide	Aryl Aryl halide	Aromatic perhal... Aryl Aryl halide
Structure similarity [category]		58.8÷73.7 %	31.6÷75 %	22.2÷66.7 %	42.1÷70.6 %	44.4÷76.2 %	21.1÷75 %	21.1÷76.2 %

Profiling results from the selected OFG profiler appeared on DM

Category consistency elements

Example 1 – *Fate endpoint*

Profiling results from the selected structure–based profilers appeared on DM (see slide #28)

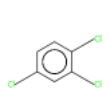
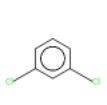
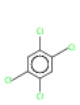
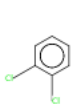
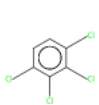
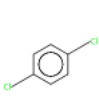
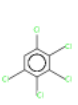
Filter endpoint tree...		1	2	3	4	5	6	7
Structure								
Structure info								
Parameters								
Physical Chemical Properties	7/167	M: 37.8 mg/L	M: 71 mg/L	M: 0.595 mg/L	M: ca.92.3 mg/L	M: 5.92 mg/L	M: 31.7 mg/L	M: 0.831 mg/L
Environmental Fate and Transport	6/6		M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %
Profiling								
Endpoint Specific								
Empiric								
Organic functional groups		Aryl Aryl halide	Aryl Aryl halide	Aromatic perhal... Aryl Aryl halide	Aryl Aryl halide	Aromatic perhal... Aryl Aryl halide	Aryl Aryl halide	Aromatic perhal... Aryl Aryl halide
Structure similarity [category]		58.8÷73.7 %	31.6÷75 %	22.2÷66.7 %	42.1÷70.6 %	44.4÷76.2 %	21.1÷75 %	21.1÷76.2 %

Profiling results from the selected "Structure similarity" profiler appeared on DM.
The similarity is estimated by analyzing variations of similarity values between category members (analogues)

Category consistency elements

Example 1 – *Fate endpoint*

Profiling results from the selected mechanism-based profilers appeared on DM (slide #29)

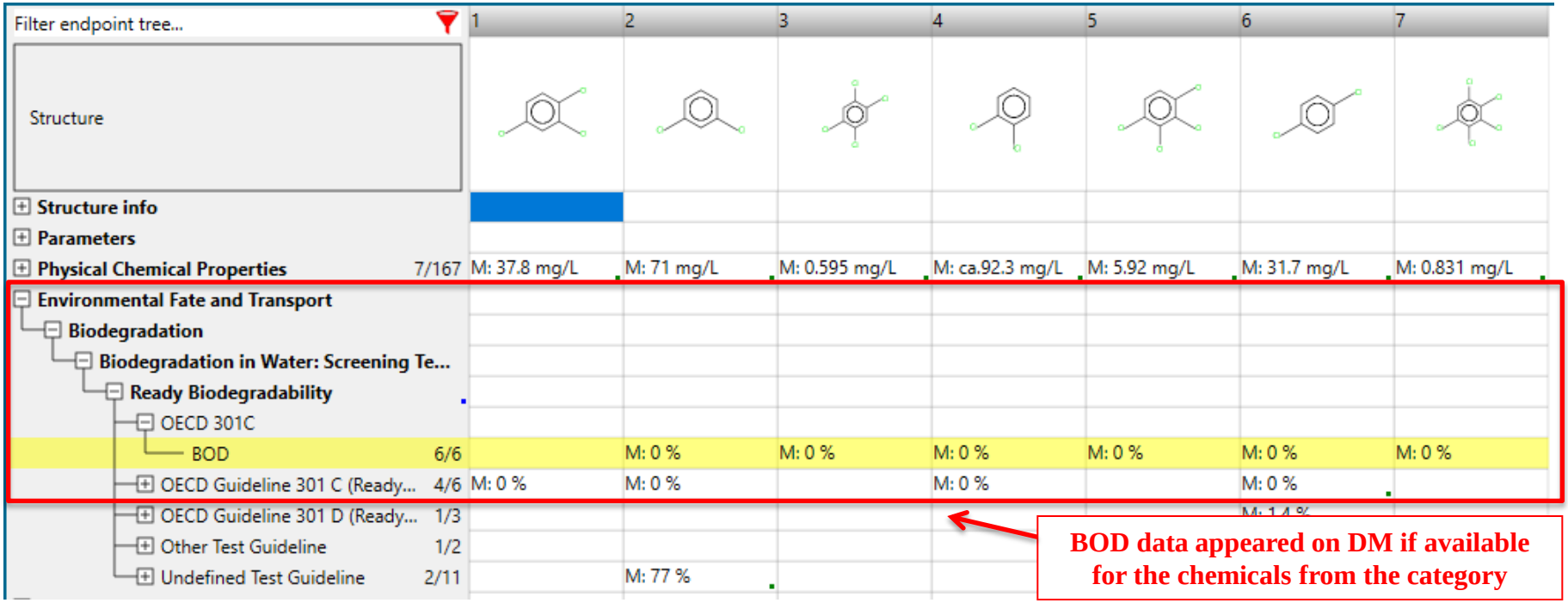
Filter endpoint tree...	1	2	3	4	5	6	7
Structure							
Structure info							
Parameters							
Physical Chemical Properties 7/167	M: 37.8 mg/L	M: 71 mg/L	M: 0.595 mg/L	M: ca.92.3 mg/L	M: 5.92 mg/L	M: 31.7 mg/L	M: 0.831 mg/L
Environmental Fate and Transport 6/6		M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %
Profiling							
General Mechanistic							
Biodegradation probability (Biowin 5)	Does NOT Biodegrade Fast	Does NOT Biodegrade Fast	Does NOT Biodegrade Fast	Does NOT Biodegrade Fast	Does NOT Biodegrade Fast	Does NOT Biodegrade F...	Does NOT Biodegrade Fast
Biodegradation probability (Biowin 6)	Does NOT Biodegrade Fast	Does NOT Biodegrade Fast	Does NOT Biodegrade Fast	Does NOT Biodegrade Fast	Does NOT Biodegrade Fast	Does NOT Biodegrade F...	Does NOT Biodegrade Fast
Endpoint Specific							
Aquatic toxicity classification by ECOSAR	Neutral Organics	Neutral Organics	Neutral Organics	Neutral Organics	Neutral Organics	Neutral Organics	Neutral Organics
Biodegradation fragments (BioWIN MITI)	Aromatic chloride [-CL] Aromatic-H	Aromatic chloride [-CL] Aromatic-H	Aromatic chloride [-CL] Aromatic-H	Aromatic chloride [-CL] Aromatic-H	Aromatic chloride [-CL] Aromatic-H	Aromatic chloride [-CL] Aromatic-H	Aromatic chloride [-CL] Aromatic-H

All the chemicals within the category are consistent with respect to the selected mechanism-based profilers

Category consistency elements

Example 1 – *Fate endpoint*

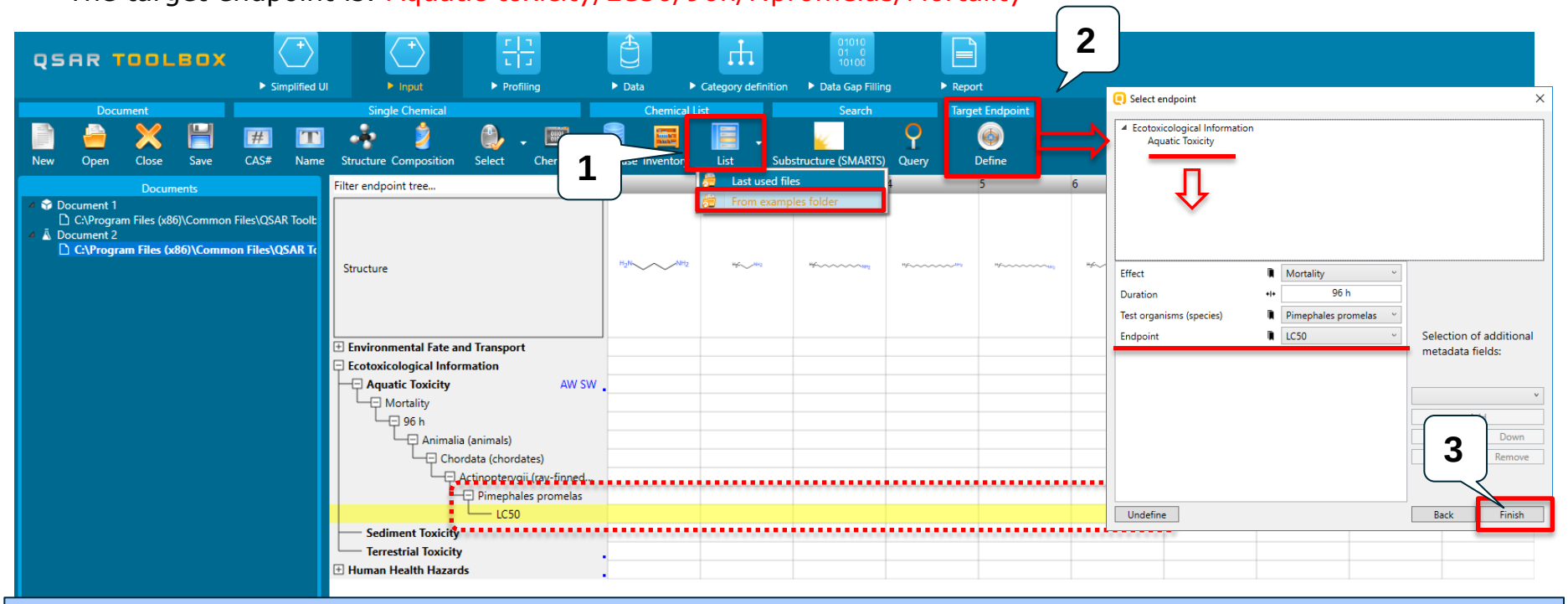
Experimental data related to the target endpoint appeared on DM



Category consistency elements

Example 2 – *Ecotox endpoint*

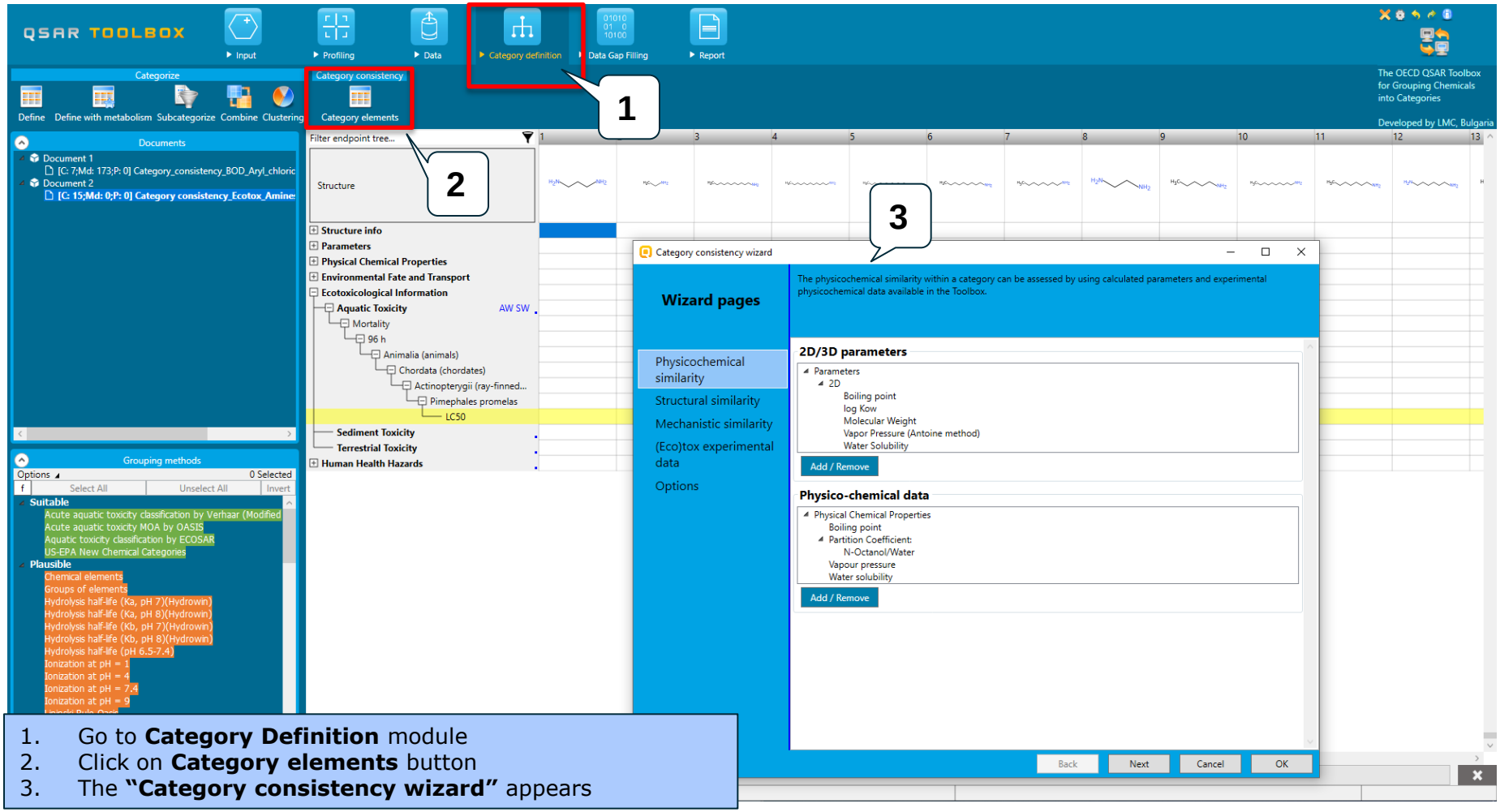
- A list of 15 Aliphatic amines is submitted as a category
- The target endpoint is: Aquatic toxicity/LC50/96h/P.promelas/Mortality



1. Open the list with chemicals **"Category consistency_Ecotox_Amines_LC50_96h_P.promelas.smi"** from the example folder
2. Click **Define** and fill in the fields of the target endpoint as shown on the snapshot on the right
3. Click **Finish**

Category consistency elements

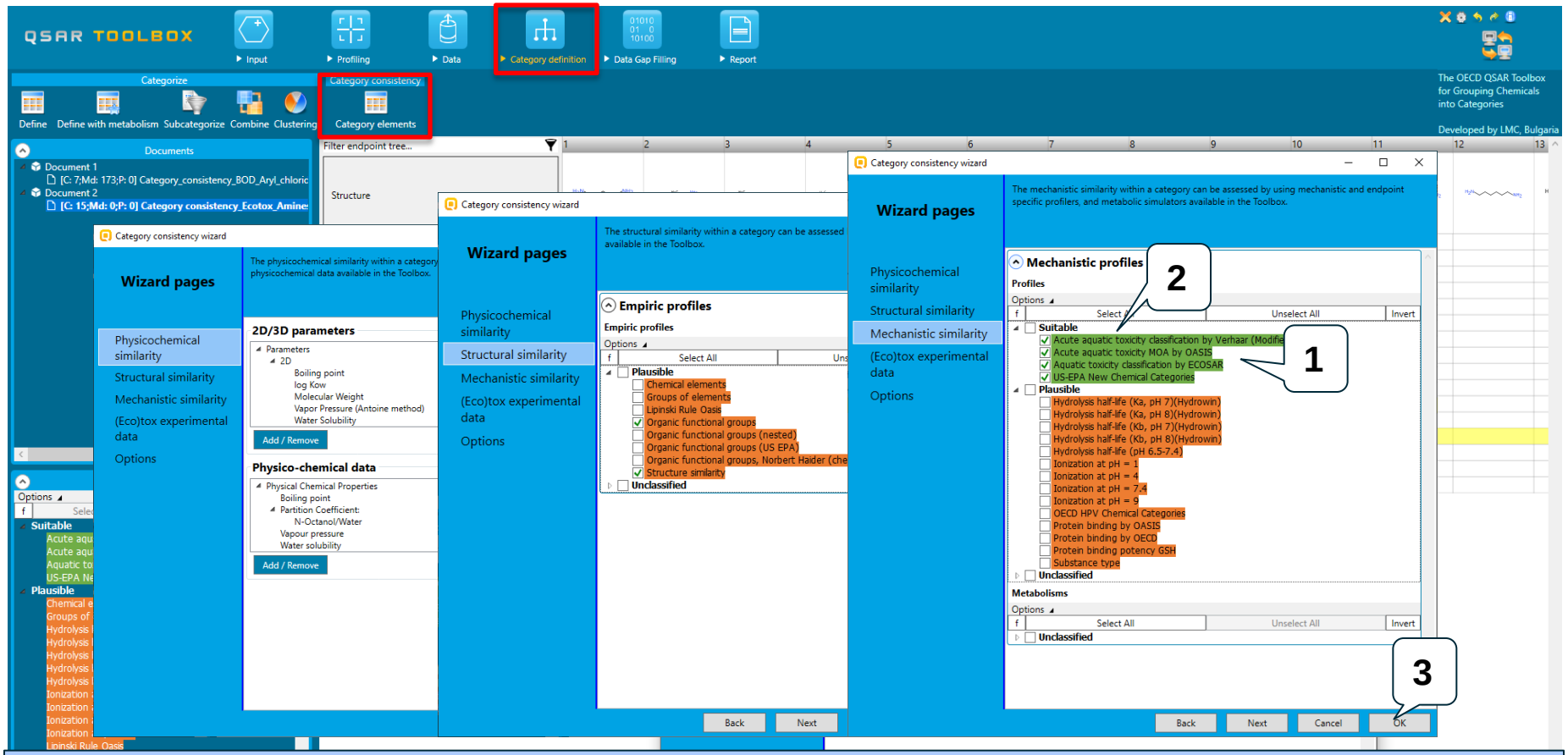
Example 2 – *Ecotox endpoint*



1. Go to **Category Definition** module
2. Click on **Category elements** button
3. The **"Category consistency wizard"** appears

Category consistency elements

Example 2 – *Ecotox endpoint*

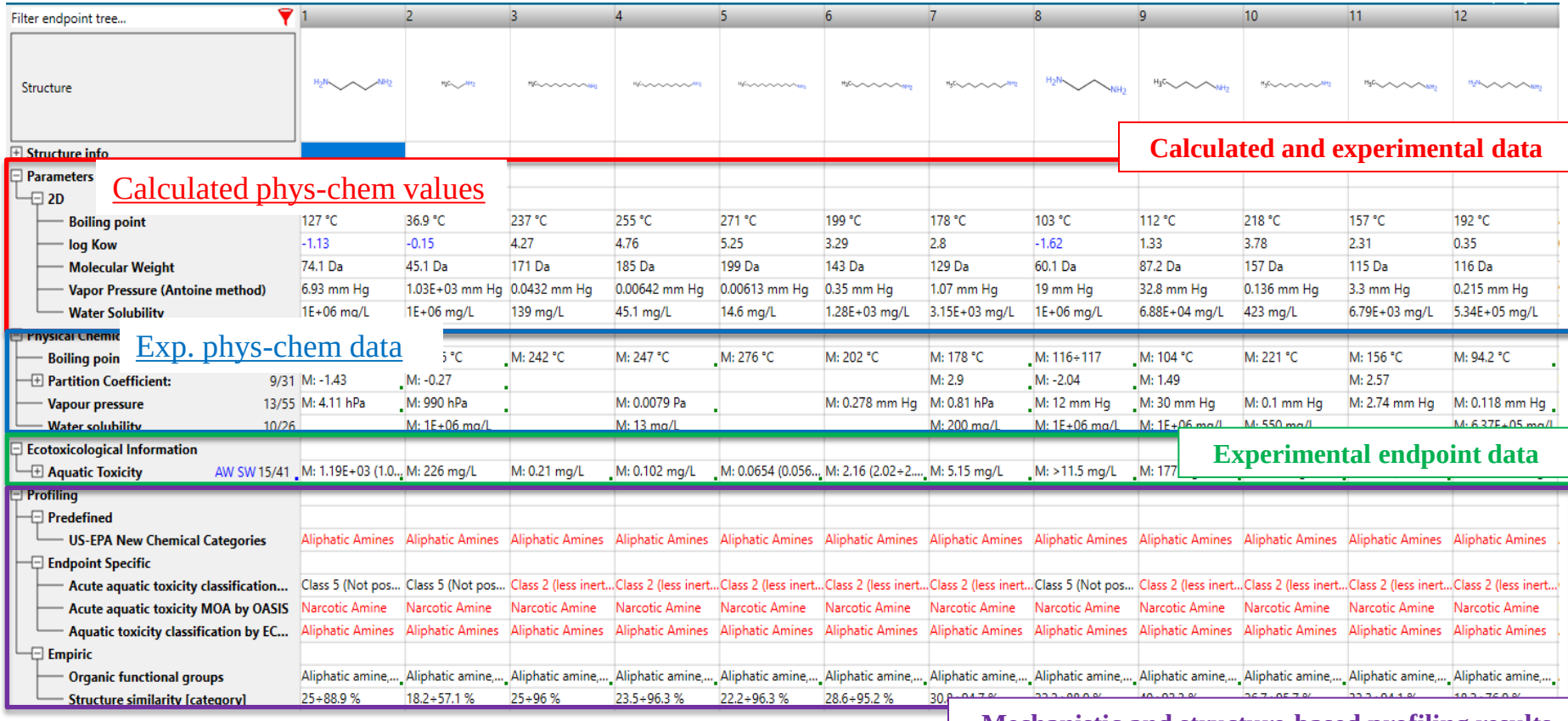


1. The only difference with the previous example is the default selection of **mechanism-based profilers**. The latter **depends on the defined target endpoint**
2. Keep the default selections of profilers and phys-chem parameters
3. Click **OK** button

Category consistency elements

Example 2 – *Ecotox endpoint*

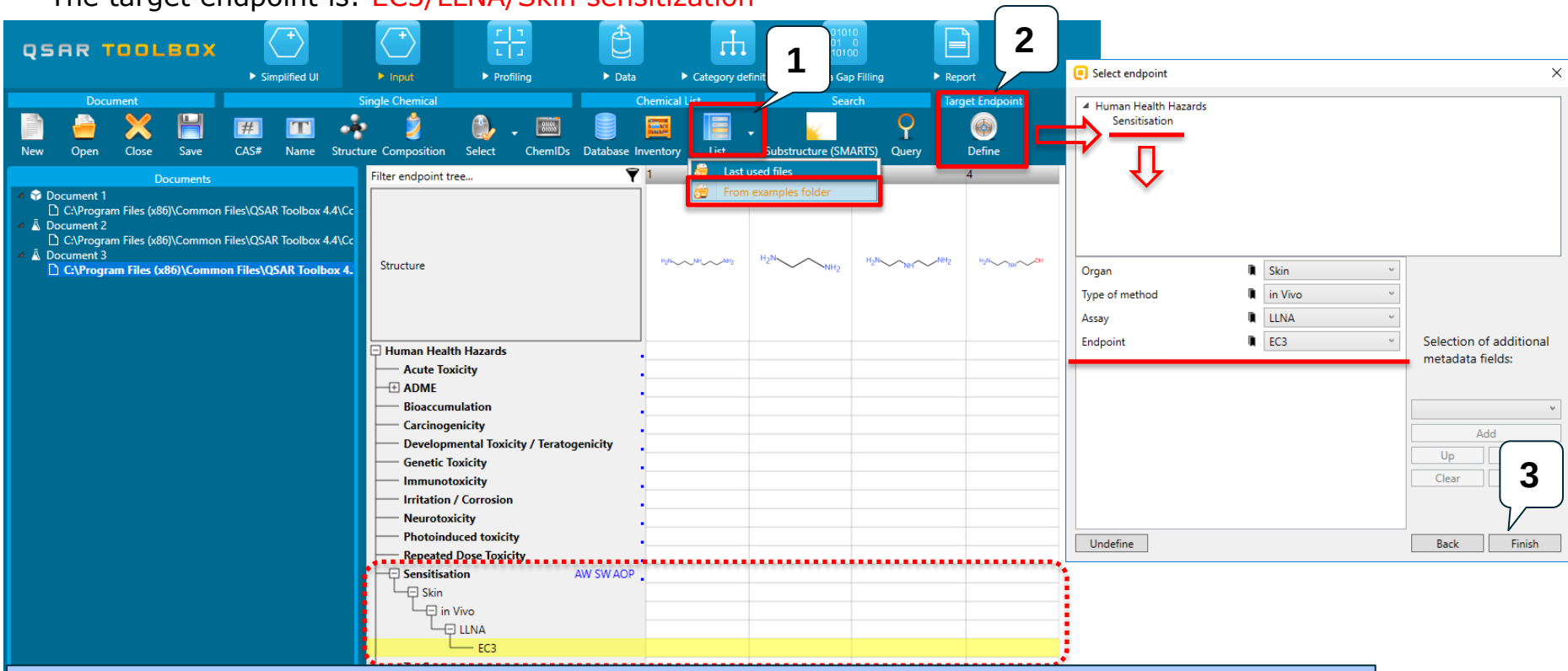
Selected data and profilers related to the specific Category elements appeared on DM



Category consistency elements

Example 3 – *Human health hazard endpoint*

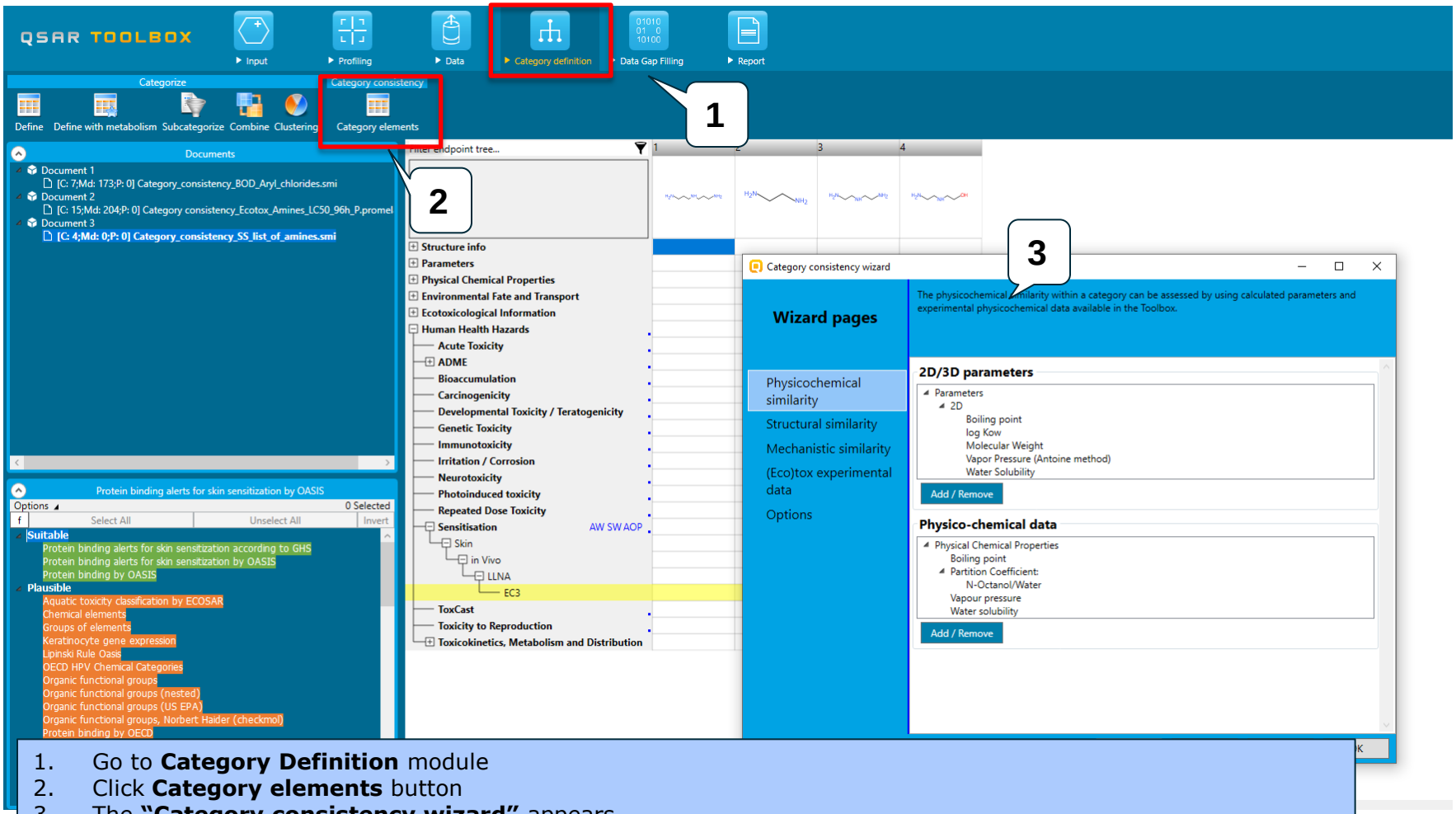
- A list of four amines is submitted as a category
- The target endpoint is: EC3/LLNA/Skin sensitization



1. Open the list with 4 amines **"Category_consistency_SS_list_of_amines.smi"** from the example folder
2. Click on **Define button** and define the target endpoint as shown on the snapshot on the right
3. Click **Finish**

Category consistency elements

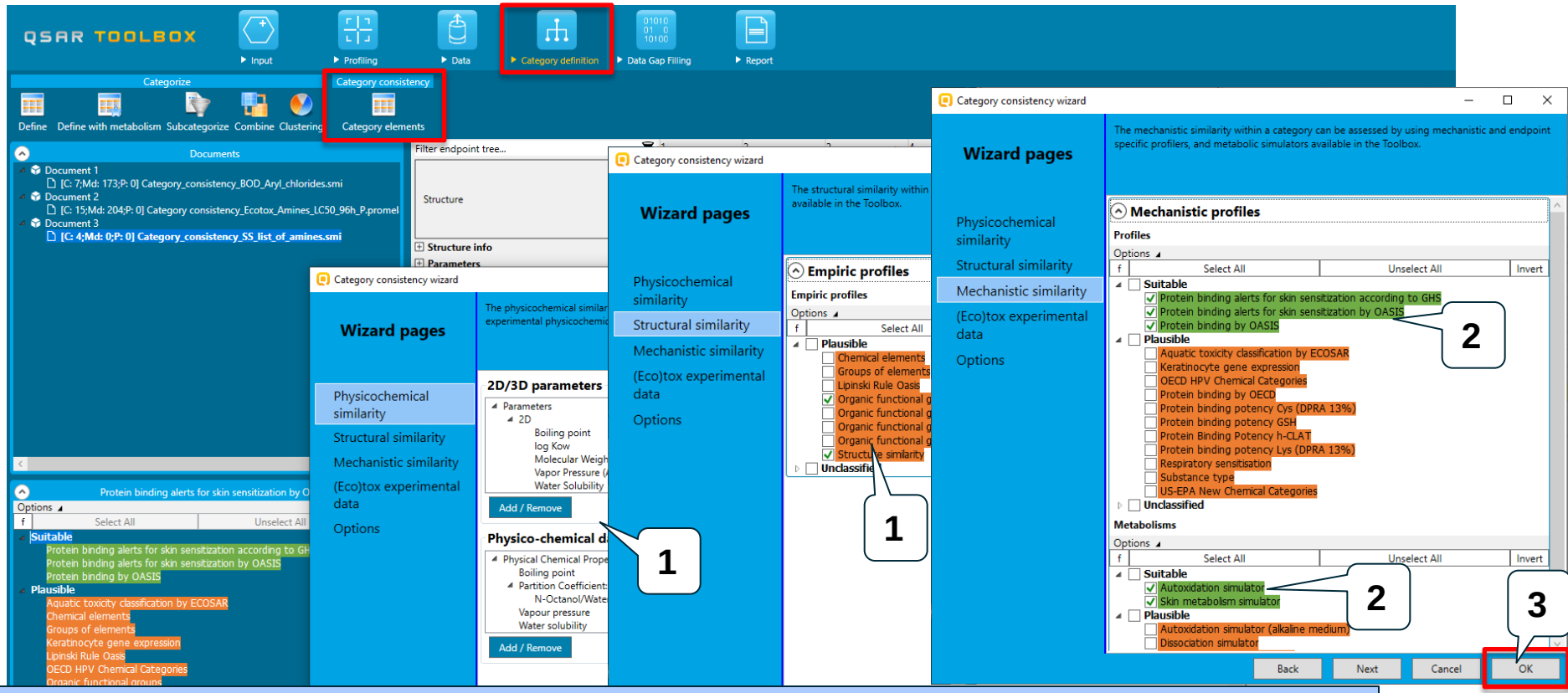
Example 3 – *Human health hazard endpoint*



1. Go to **Category Definition** module
2. Click **Category elements** button
3. The **"Category consistency wizard"** appears

Category consistency elements

Example 3 – *Human health hazard endpoint*

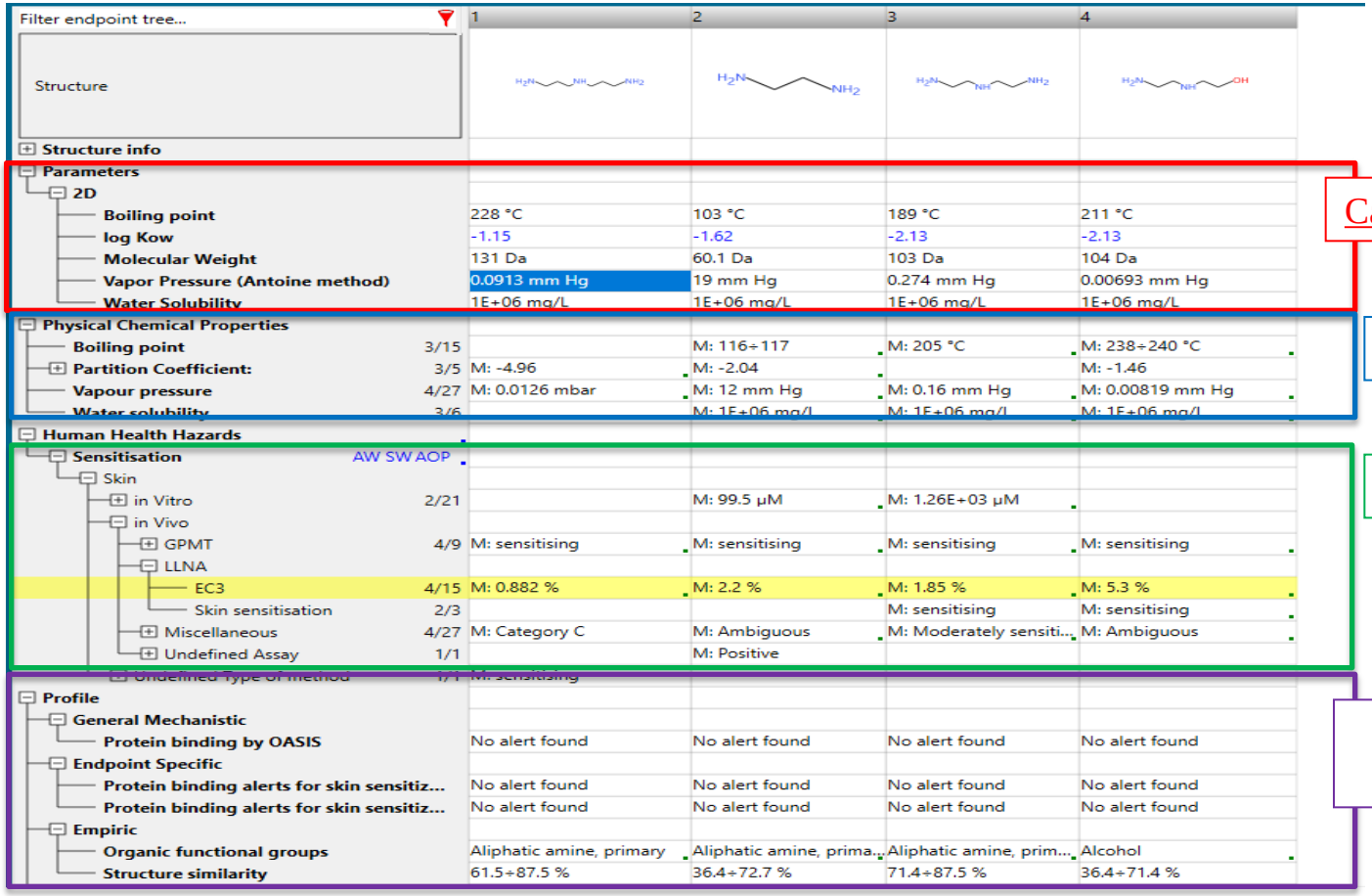


1. As in the previous two examples the default selection of 2D/3D parameters and structure-based profilers are the same and do not depend on the defined target endpoint, while the default selection of mechanism–based profilers are different in all the three examples. Moreover it depends on the defined target endpoint. As can be seen the mechanistic profilers could be combined with appropriate metabolism.
2. Keep the default selection of profilers, metabolic simulators and phys-chem properties
3. Click **OK** button.

Category consistency elements

Example 3 – *Human health hazard endpoint*

Selected data and profilers related to the specific Category elements appeared on DM



Filter endpoint tree...		1	2	3	4
Structure					
Structure info					
Parameters					
2D					
Boiling point		228 °C	103 °C	189 °C	211 °C
log Kow		-1.15	-1.62	-2.13	-2.13
Molecular Weight		131 Da	60.1 Da	103 Da	104 Da
Vapor Pressure (Antoine method)		0.0913 mm Hg	19 mm Hg	0.274 mm Hg	0.00693 mm Hg
Water Solubility		1E+06 mg/L	1E+06 mg/L	1E+06 mg/L	1E+06 mg/L
Physical Chemical Properties					
Boiling point	3/15		M: 116÷117	M: 205 °C	M: 238÷240 °C
Partition Coefficient:	3/5	M: -4.96	M: -2.04		M: -1.46
Vapour pressure	4/27	M: 0.0126 mbar	M: 12 mm Hg	M: 0.16 mm Hg	M: 0.00819 mm Hg
Water solubility	3/6		M: 1E+06 mg/l	M: 1E+06 mg/l	M: 1E+06 mg/l
Human Health Hazards					
Sensitisation	AW SW AOP				
Skin					
in Vitro	2/21		M: 99.5 µM	M: 1.26E+03 µM	
in Vivo					
GPMT	4/9	M: sensitising	M: sensitising	M: sensitising	M: sensitising
LLNA					
EC3	4/15	M: 0.882 %	M: 2.2 %	M: 1.85 %	M: 5.3 %
Skin sensitisation	2/3			M: sensitising	M: sensitising
Miscellaneous	4/27	M: Category C	M: Ambiguous	M: Moderately sensiti...	M: Ambiguous
Undefined Assay	1/1		M: Positive		
Profile					
General Mechanistic					
Protein binding by OASIS		No alert found	No alert found	No alert found	No alert found
Endpoint Specific					
Protein binding alerts for skin sensitiz...		No alert found	No alert found	No alert found	No alert found
Protein binding alerts for skin sensitiz...		No alert found	No alert found	No alert found	No alert found
Empiric					
Organic functional groups		Aliphatic amine, primary	Aliphatic amine, prima...	Aliphatic amine, prim...	Alcohol
Structure similarity		61.5÷87.5 %	36.4÷72.7 %	71.4÷87.5 %	36.4÷71.4 %

Calculated phys-chem values

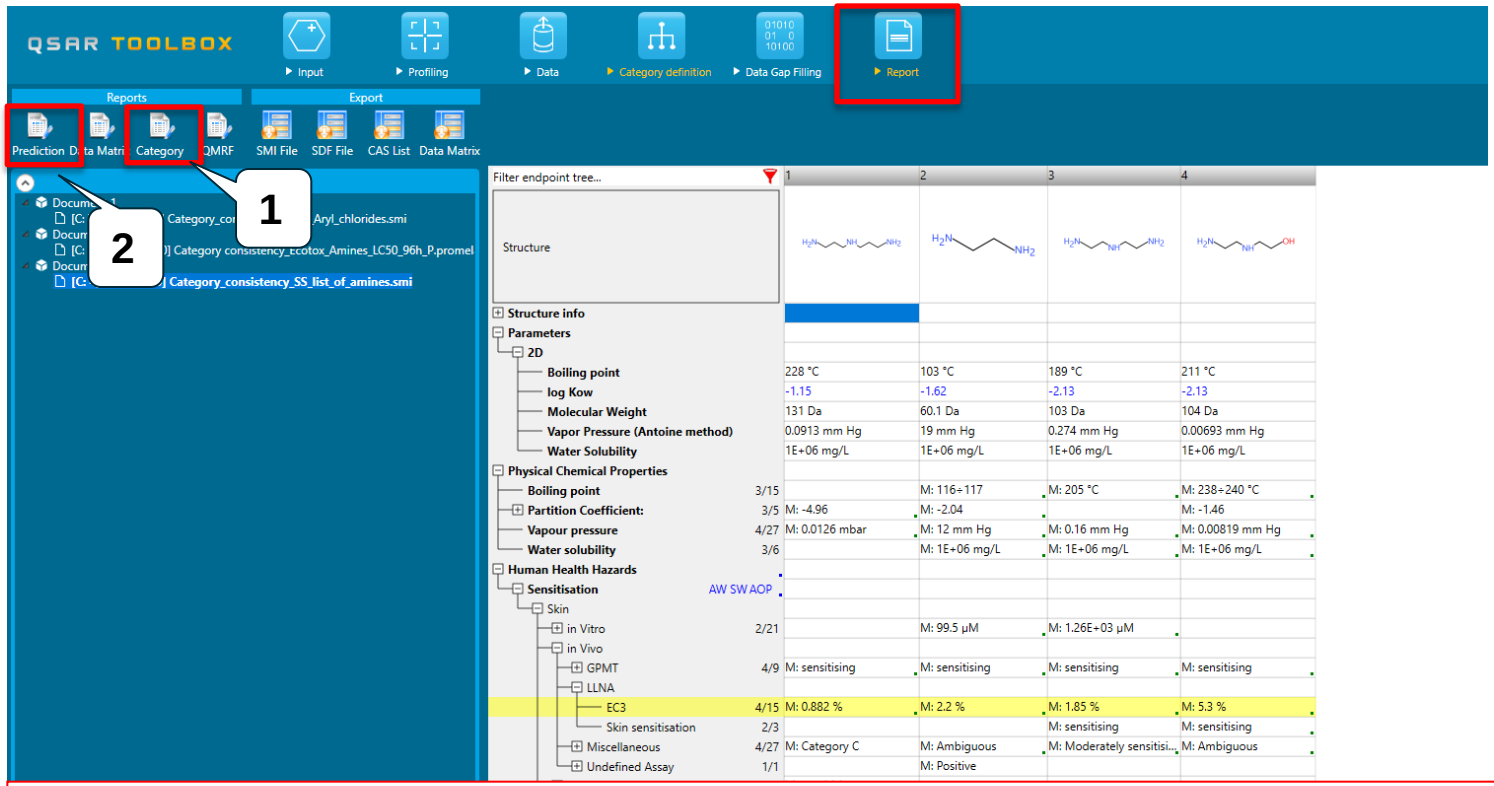
Exp. phys-chem data

Experimental data

Mechanistic and structure-based profiling results

Category consistency elements

Report



A report can be obtained for:

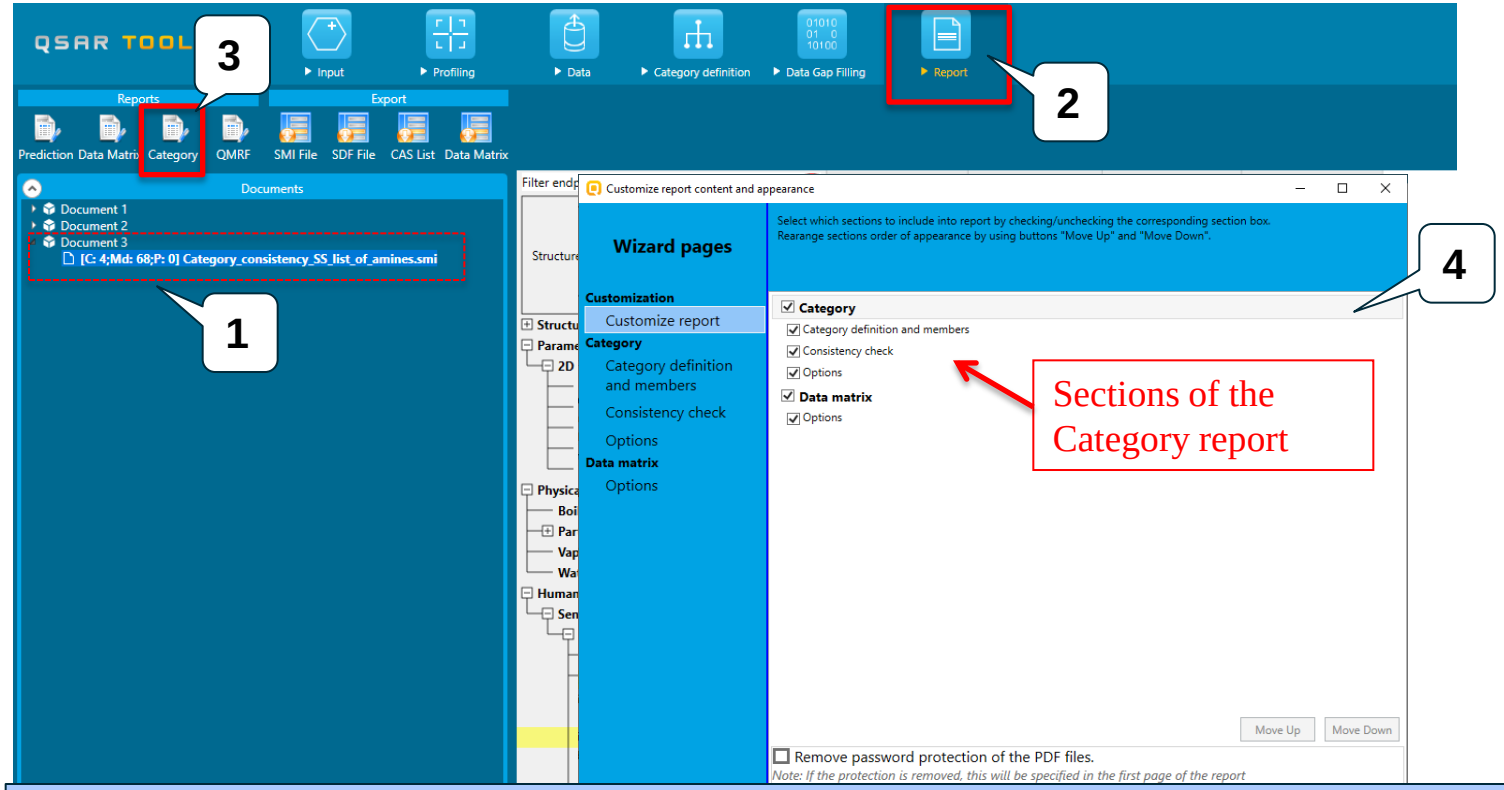
1) A category (list of chemicals) with defined target endpoint without entering into data gap filling (1) OR
2) The prediction based on the analogues grouped during the read-across workflow (in the data gap filling) (2)

A report for a list of chemicals without entering into data gap filling will be illustrated in the next few slides (point 1)

Category consistency elements

Report

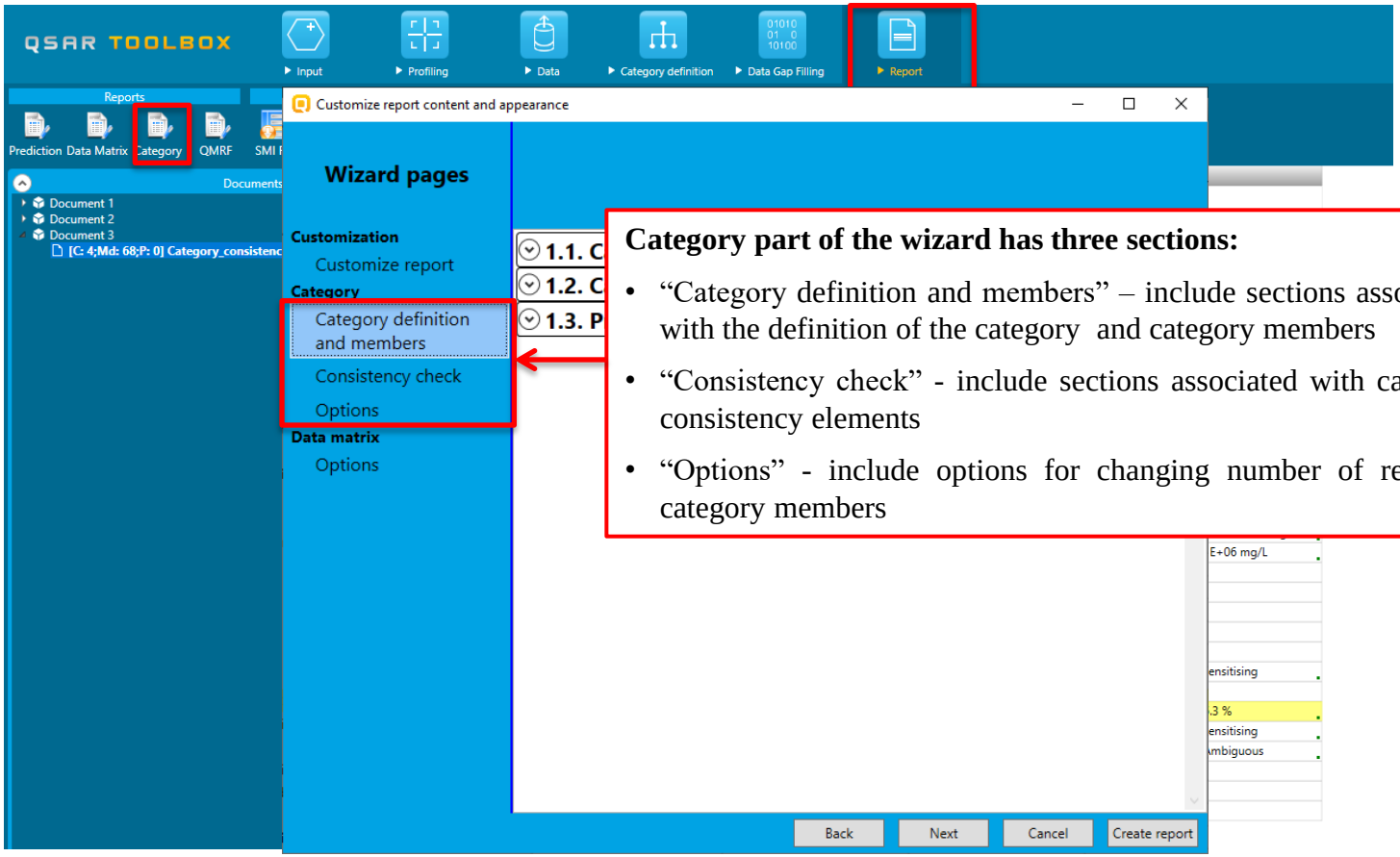
A report for the list of aliphatic amines associated with skin sensitization endpoint (Example 3) will be illustrated



1. Click on the document for assessment category consistency of the list with aliphatic amines
2. Go to **Report module**
3. Click on **Category report**
4. The wizard window with sections of the category report appeared

Category consistency elements

Report

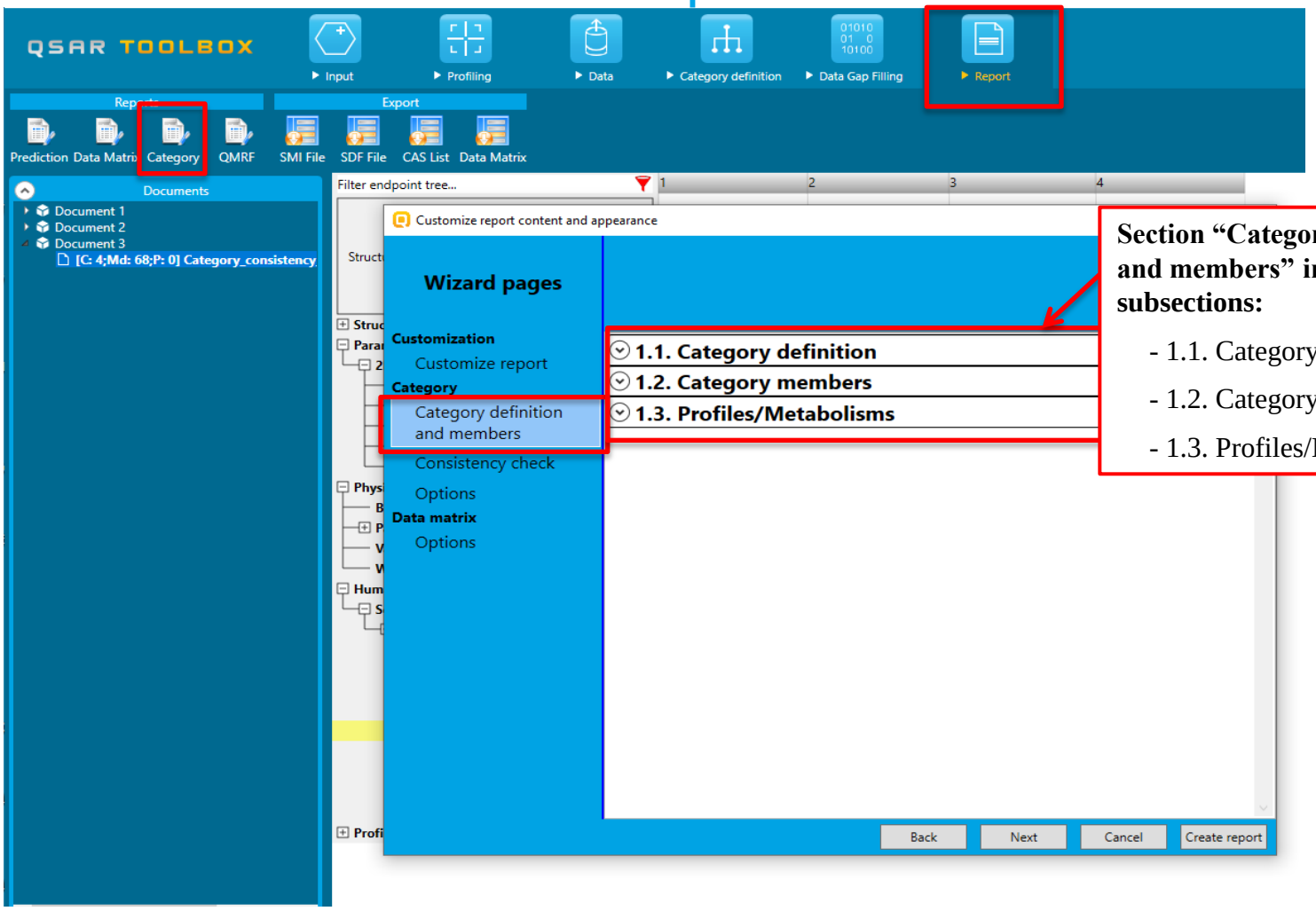


Category part of the wizard has three sections:

- "Category definition and members" – include sections associated with the definition of the category and category members
- "Consistency check" - include sections associated with category consistency elements
- "Options" - include options for changing number of reported category members

Category consistency elements

Report

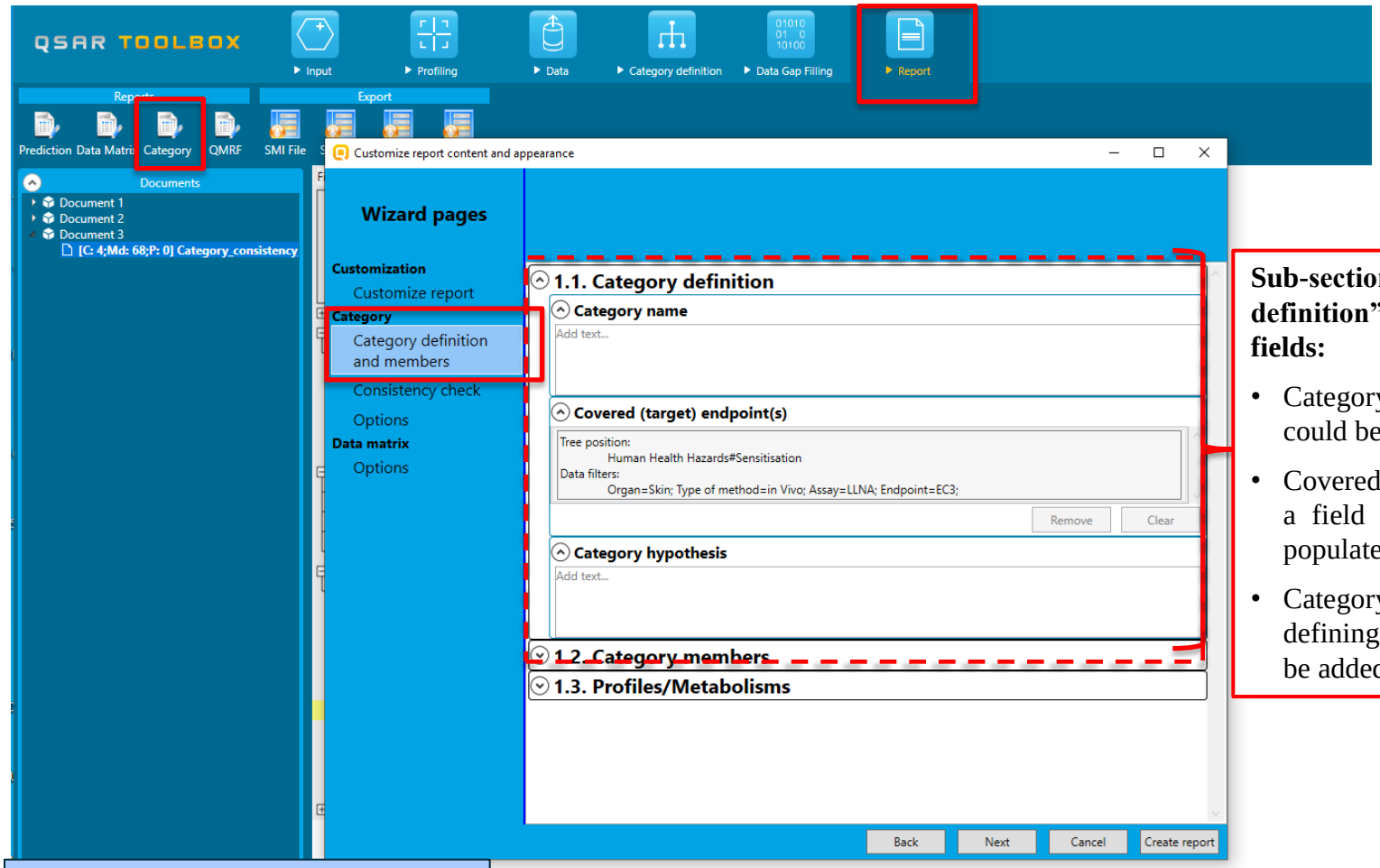


Section "Category definition and members" includes three subsections:

- 1.1. Category definition
- 1.2. Category members
- 1.3. Profiles/Metabolisms

Category consistency elements

Report



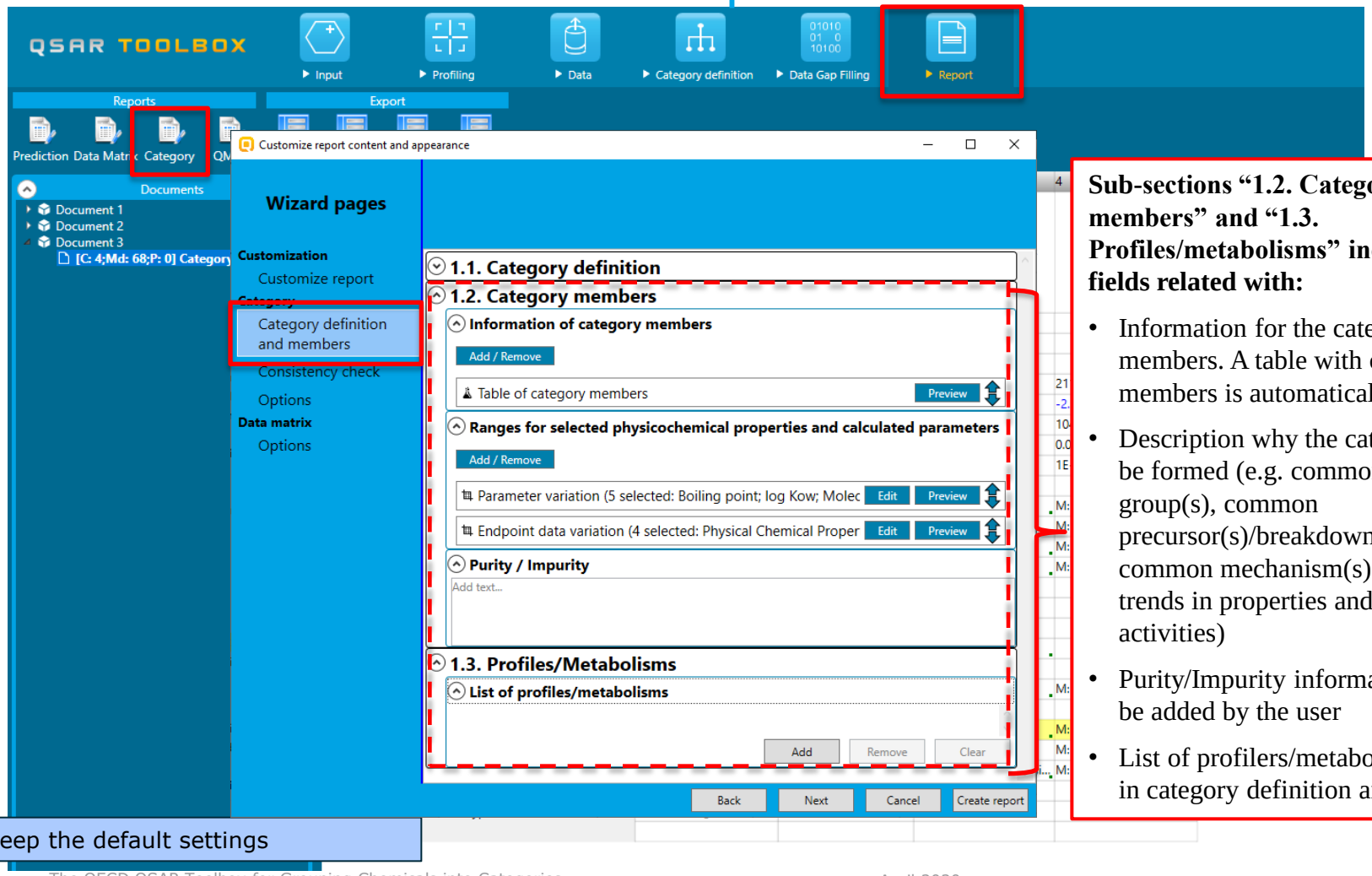
Sub-section "1.1. Category definition" includes the following fields:

- Category name – a field which could be fulfilled by the user
- Covered (Target) endpoint(s) – a field which is automatically populated by the system
- Category Hypothesis for defining the category – need to be added manually

Keep the default settings

Category consistency elements

Report



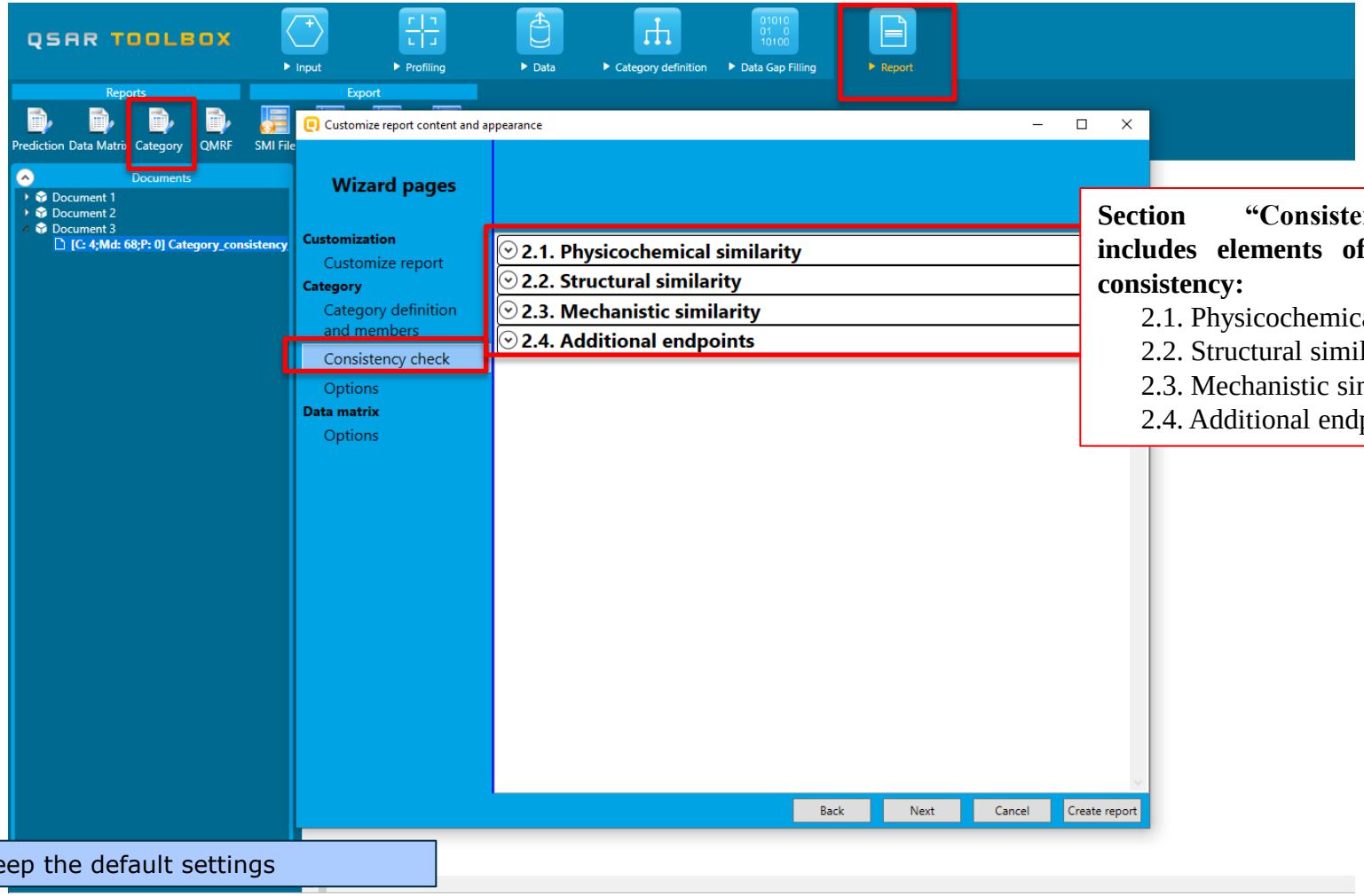
Sub-sections "1.2. Category members" and "1.3. Profiles/metabolisms" includes fields related with:

- Information for the category members. A table with category members is automatically provided
- Description why the category can be formed (e.g. common functional group(s), common precursor(s)/breakdown product(s), common mechanism(s) of action, trends in properties and/or activities)
- Purity/Impurity information should be added by the user
- List of profilers/metabolisms used in category definition are provided.

Keep the default settings

Category consistency elements

Report



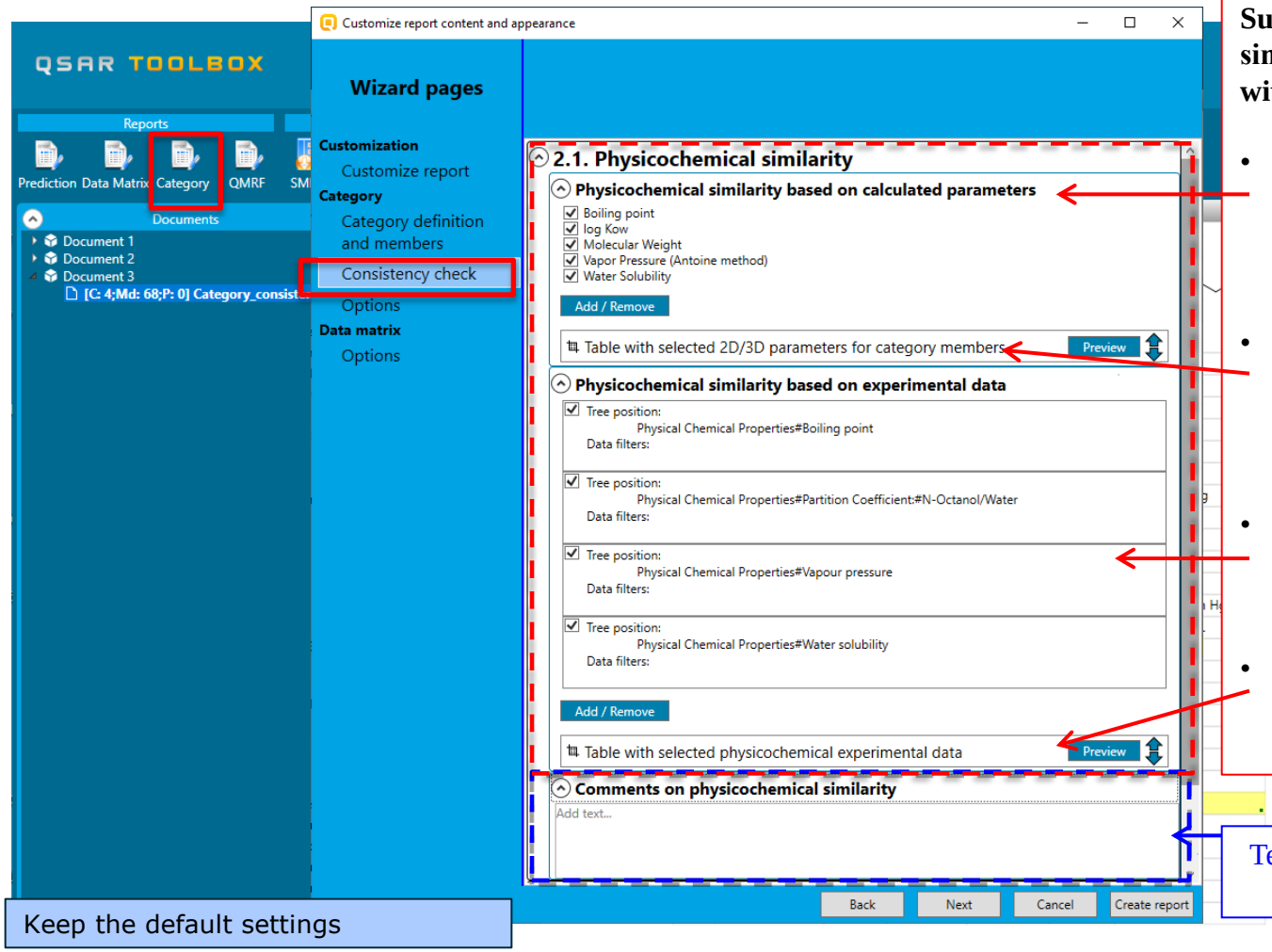
Section "Consistency check" includes elements of the category consistency:

- 2.1. Physicochemical similarity
- 2.2. Structural similarity
- 2.3. Mechanistic similarity
- 2.4. Additional endpoints

Keep the default settings

Category consistency elements

Report



Sub-section "2.1. Physicochemical similarity" includes fields related with:

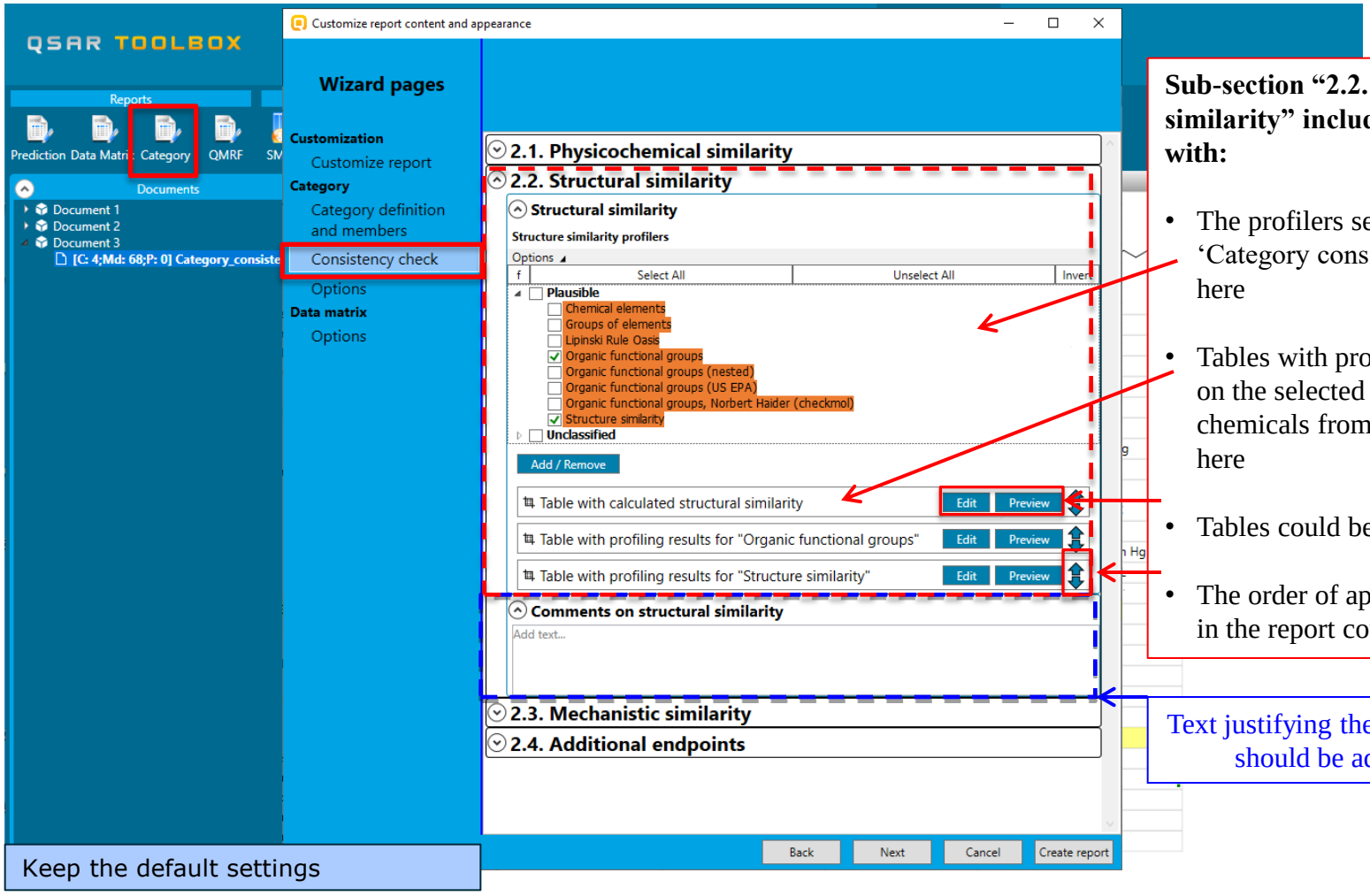
- Collection of calculated physicochemical parameters already selected in the 'Category consistency' tool appear here
- A table with calculated physicochemical parameters selected in the 'Category consistency' tool appear here
- The experimental physicochemical parameters selected in the 'Category consistency' tool appear here.
- A table with phys-chem parameters for category members is automatically provided.

Text justifying the selected parameters should be added by the user

Keep the default settings

Category consistency elements

Report



Sub-section "2.2. Structural similarity" includes fields related with:

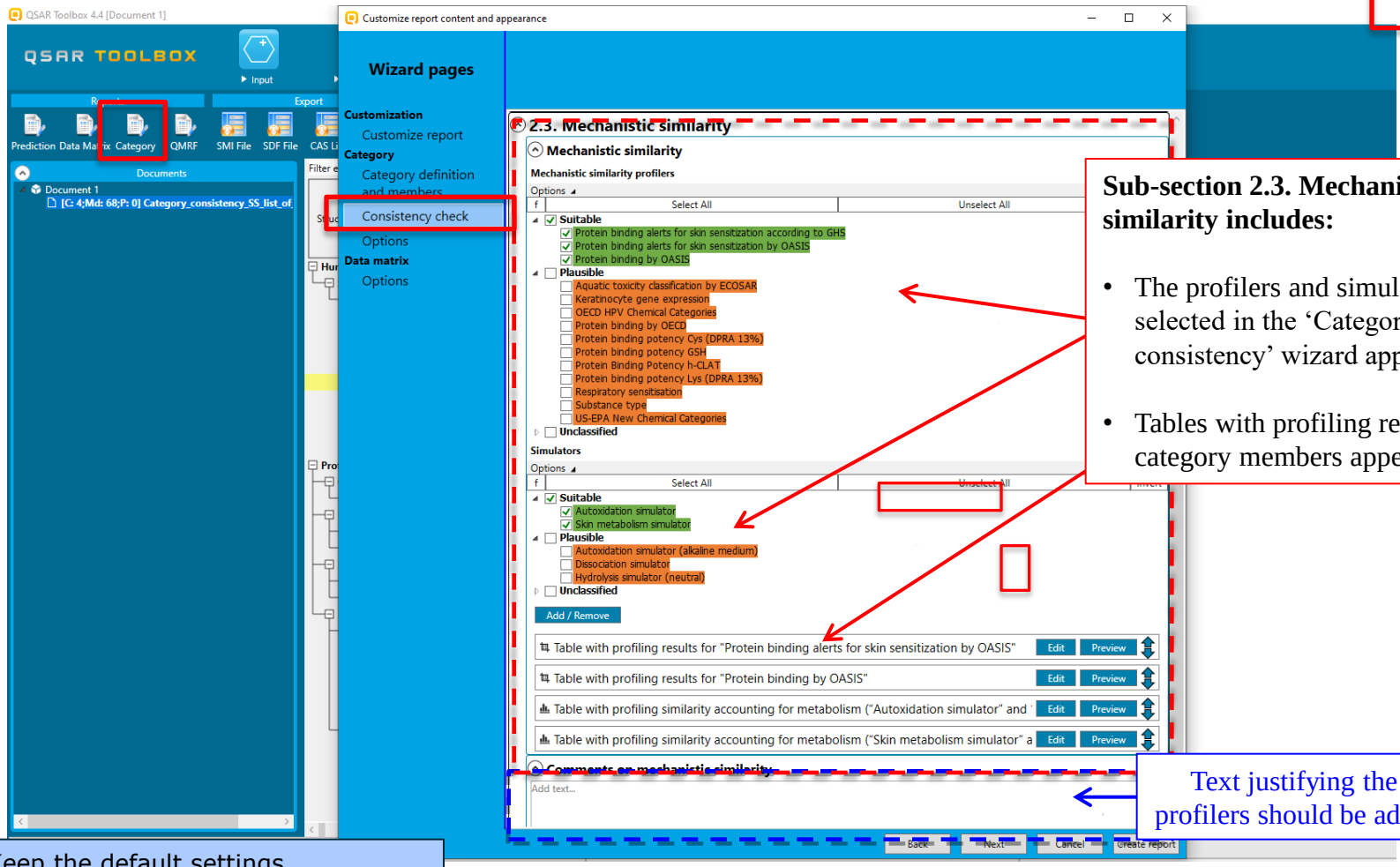
- The profilers selected in the 'Category consistency' tool appear here
- Tables with profiling results based on the selected profilers for the chemicals from the category appear here
- Tables could be edited or previewed
- The order of appearance of the tables in the report could be changed

Text justifying the selected profilers should be added by user

Keep the default settings

Category consistency elements

Report



Sub-section 2.3. Mechanistic similarity includes:

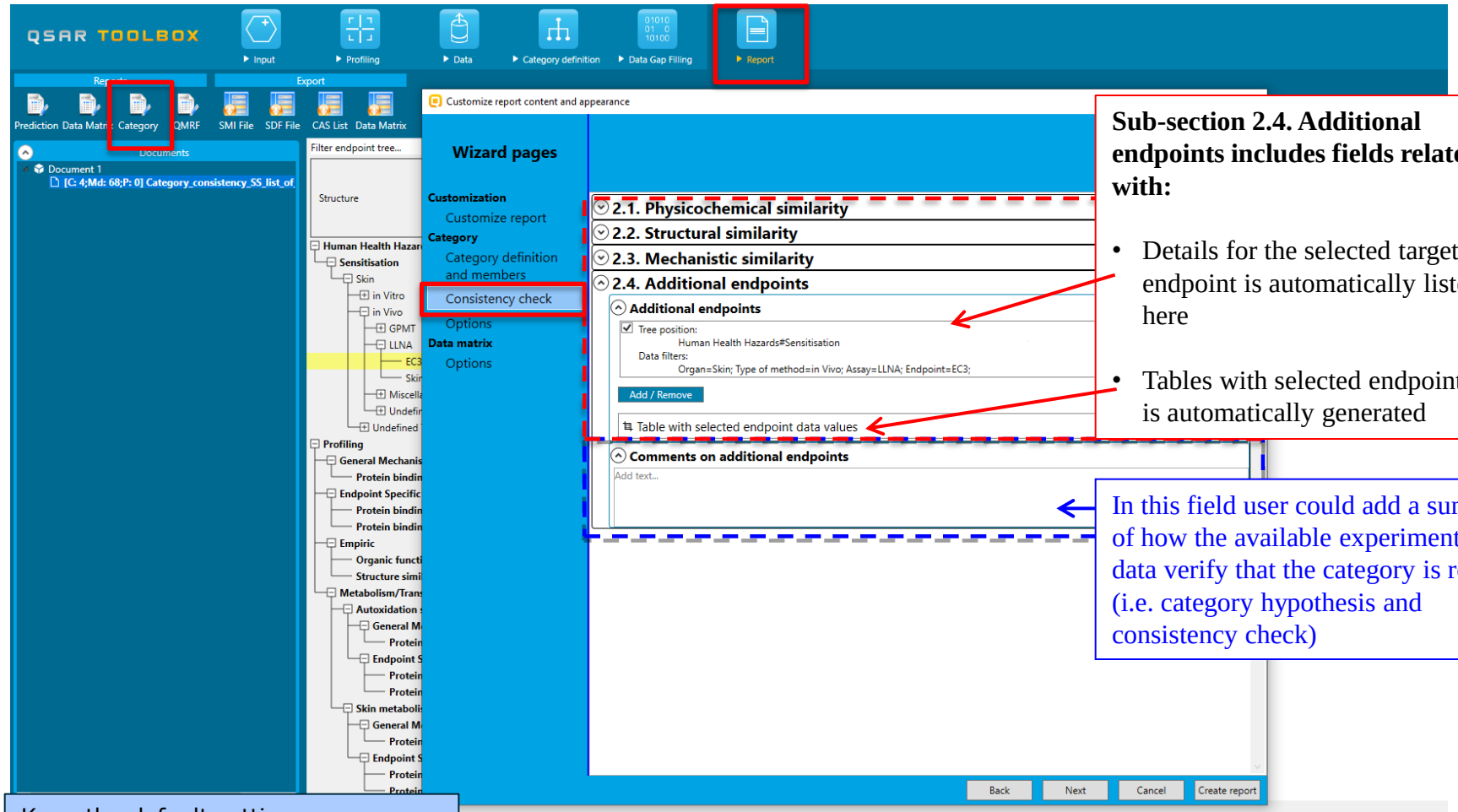
- The profilers and simulators selected in the 'Category consistency' wizard appear here
- Tables with profiling results for the category members appear here

Text justifying the selected profilers should be added by user

Keep the default settings

Category consistency elements

Report



Sub-section 2.4. Additional endpoints includes fields related with:

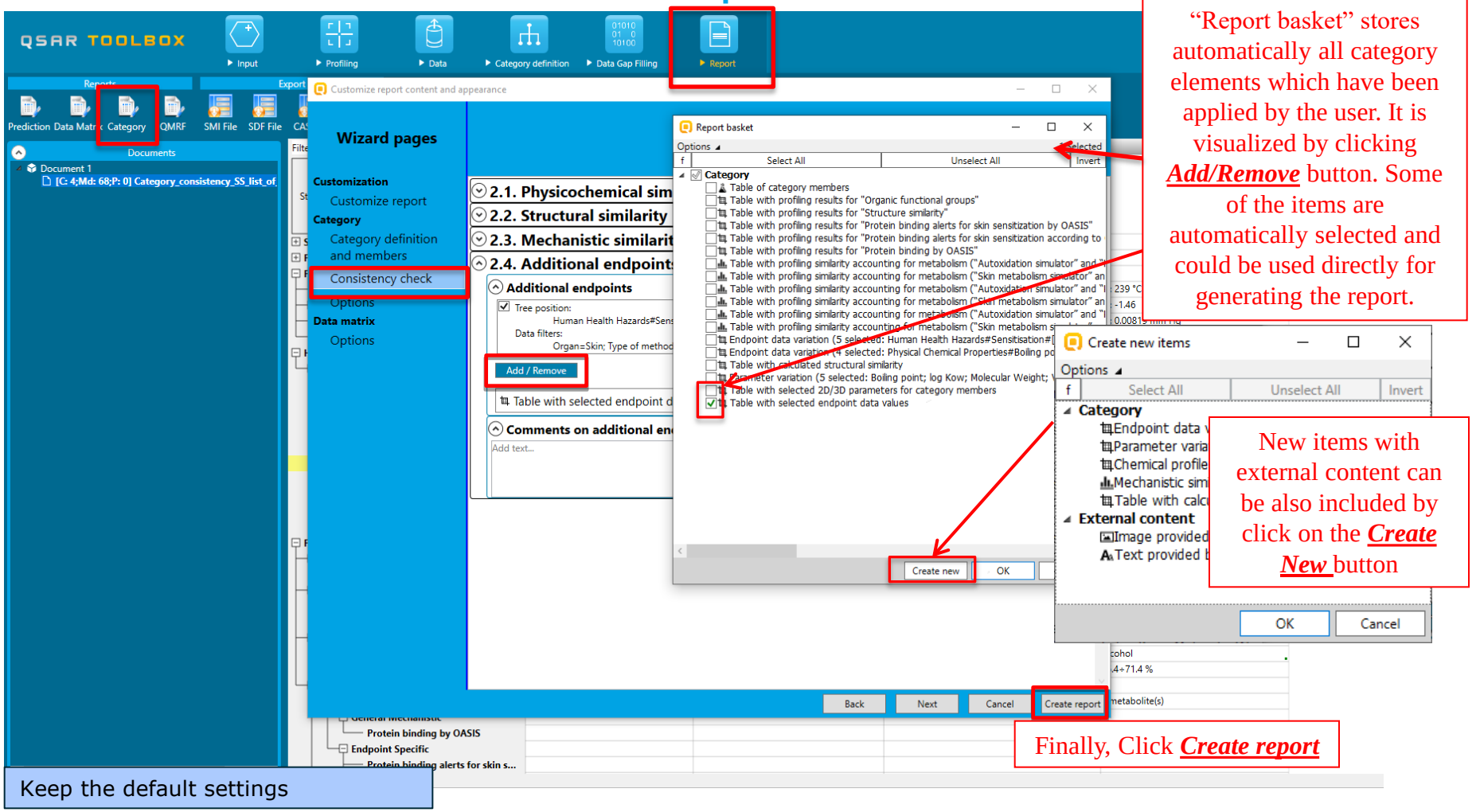
- Details for the selected target endpoint is automatically listed here
- Tables with selected endpoint data is automatically generated

In this field user could add a summary of how the available experimental data verify that the category is robust (i.e. category hypothesis and consistency check)

Keep the default settings

Category consistency elements

Report



"Report basket" stores automatically all category elements which have been applied by the user. It is visualized by clicking ***Add/Remove*** button. Some of the items are automatically selected and could be used directly for generating the report.

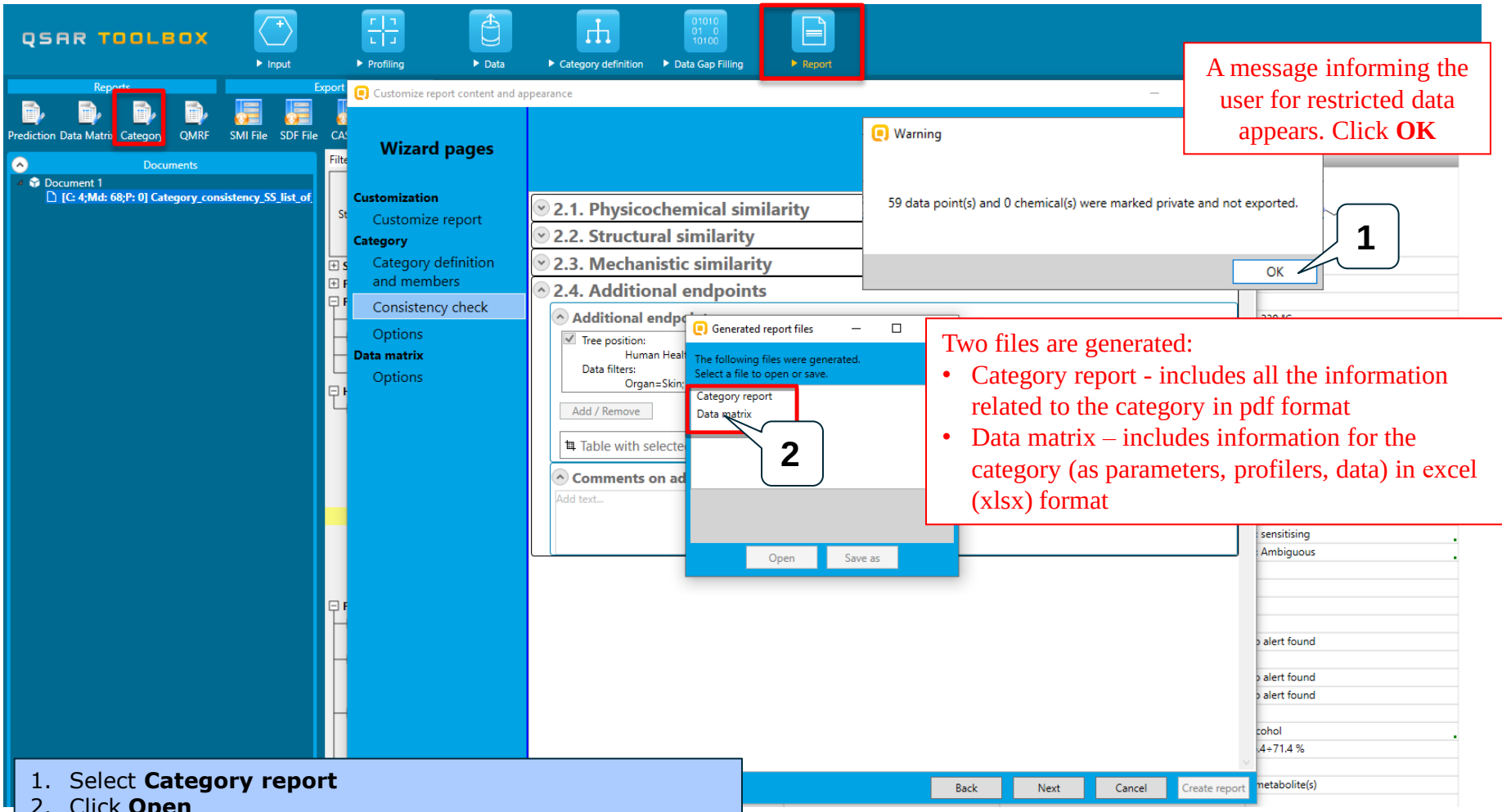
New items with external content can be also included by click on the ***Create New*** button

Finally, Click ***Create report***

Keep the default settings

Category consistency elements

Report



A message informing the user for restricted data appears. Click **OK**

Two files are generated:

- Category report - includes all the information related to the category in pdf format
- Data matrix – includes information for the category (as parameters, profilers, data) in excel (xlsx) format

1. Select **Category report**
2. Click **Open**

Category consistency elements

Report

1. Category definition section:

QSAR Toolbox report for category

1. Category definition

1.1. Category definition

Category name *manually editable field*

Not provided by the user

Covered (target) endpoint(s)

- Human Health Hazards/Sensitisation: Skin, in Vivo, LLNA, EC3

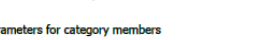
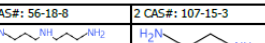
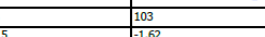
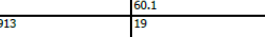
Category hypothesis *manually editable field*

Not provided by the user

1.2. Category members

Information of category members

Table of category members

#	CAS	Name	SMILES	Structure
1	56-18-8	Iminobis-3-propylamine	NCCCNCCCN	
2	107-15-3	Ethylenediamine	NCCN	
3	111-40-0	DETA	NCCNCCN	
4	111-41-1	Aminoethyl ethanolamine	NCCNCCO	

Ranges for selected physicochemical properties and calculated parameters

Table with 2D parameters data variation

Parameter name	Variation	unit (family)
Boiling point	103 ÷ 228	°C(Temperature)
log Kow	-2.13 ÷ -1.15	
Molecular Weight	60.1 ÷ 131	Da(Mass)
Vapor Pressure (Antoine method)	0.00693 ÷ 19	mm Hg(Pressure)
Water Solubility	1E+06	mg/L(Mass concentration)

Table with Physical Chemical Properties data variation

Position	Variation	unit (family)	Number of chemicals
Boiling point	117 ÷ 244	°C(Temperature)	3
Boiling point	117		1

Partition Coefficient:#N-Octanol/Water	-4.96 ÷ -1.25		3
Vapour pressure	0.0126 ÷ 99.8	mbar(Pressure)	1
Vapour pressure	0.00819 ÷ 12	mm Hg(Pressure)	3
Vapour pressure	0.01 ÷ 46.9	hPa(Pressure)	3
Vapour pressure	29.3 ÷ 1.25E+03	Pa(Pressure)	2
Water solubility	1E+06	mg/L(Mass concentration)	3
Water solubility	100	vol%(Unknown)	1

Purity / Impurity *manually editable field*

Not provided by the user

1.3. Profiles/Metabolisms

List of profiles/metabolisms

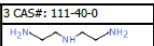
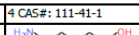
2. Consistency check

2.1. Physicochemical similarity

Physicochemical similarity based on calculated parameters

Table with selected 2D/3D parameters for category members

	1 CAS#: 56-18-8	2 CAS#: 107-15-3
Boiling point, °C	228	103
log Kow,	-1.15	-1.62
Molecular Weight, Da	131	60.1
Vapor Pressure (Antoine method), mm Hg	0.0913	19
Water Solubility, mg/L	1E+06	1E+06

	3 CAS#: 111-40-0	4 CAS#: 111-41-1
		
Boiling point, °C	189	211
log Kow,	-2.13	-2.13
Molecular Weight, Da	103	104
Vapor Pressure (Antoine method), mm Hg	0.274	0.00693
Water Solubility, mg/L	1E+06	1E+06

Physicochemical similarity based on experimental data

1. Category definition section:

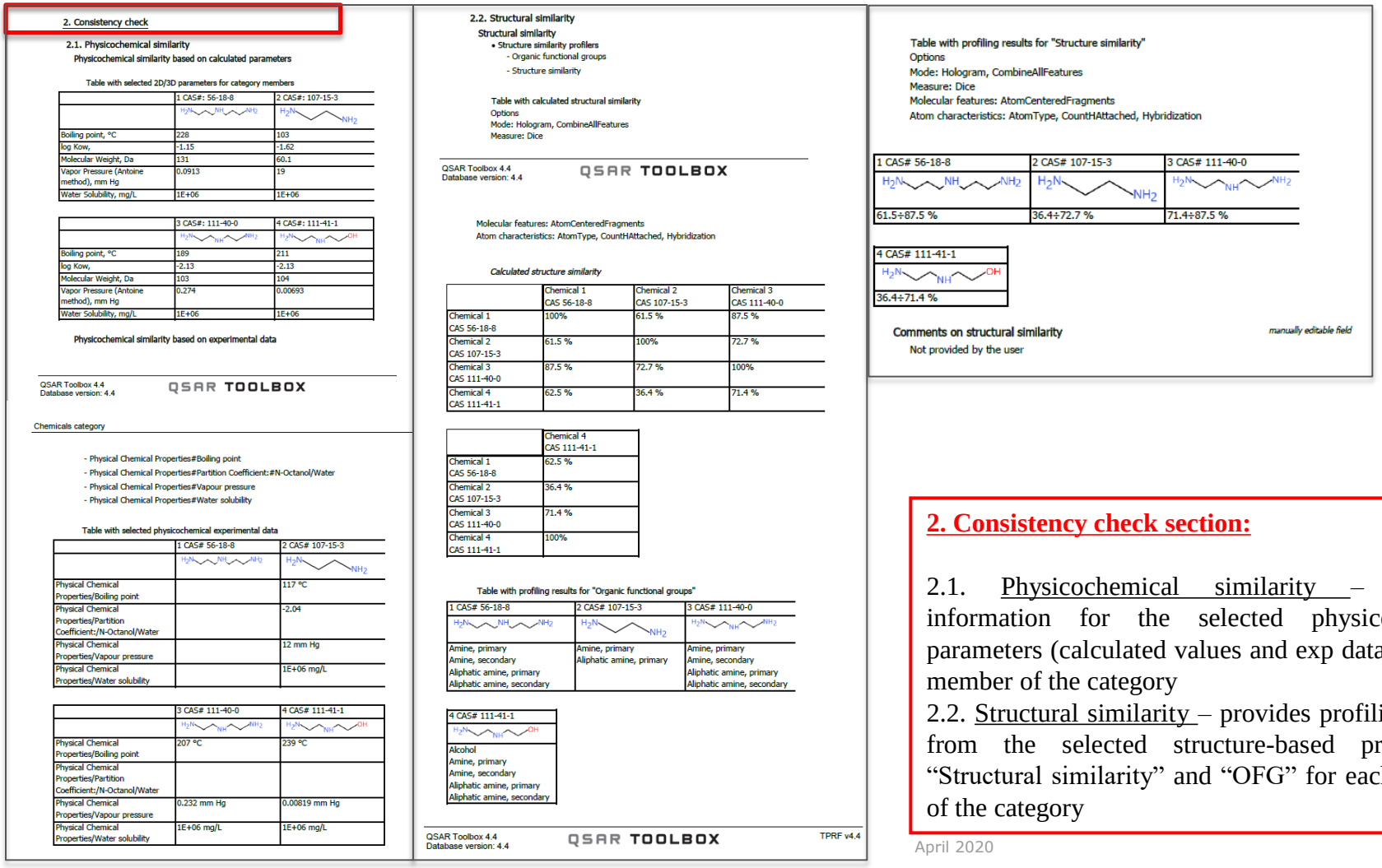
1.1. Category definition – includes:
- name of the category if provided by the user);
- Covered (target) endpoint(s) – automatically fulfilled
- Category hypothesis – information provided by the user

1.2. Category members
- a table with category members is provided;
- ranges of selected experimental and calculated physicochemical properties in table format;
- information for Purity/Impurity manually added by the user

1.3. Profiling/Metabolisms –
- additional profilers/metabolisms could be provided by the user

Category consistency elements

2. Consistency check section: Report



2. Consistency check

2.1. Physicochemical similarity

Physicochemical similarity based on calculated parameters

Table with selected 2D/3D parameters for category members

	1 CAS#: 56-18-8	2 CAS#: 107-15-3
Boiling point, °C	228	103
log Kow,	-1.15	-1.62
Molecular Weight, Da	131	60.1
Vapor Pressure (Antoine method), mm Hg	0.0913	19
Water Solubility, mg/L	1E+06	1E+06

	3 CAS#: 111-40-0	4 CAS#: 111-41-1
Boiling point, °C	189	211
log Kow,	-2.13	-2.13
Molecular Weight, Da	103	104
Vapor Pressure (Antoine method), mm Hg	0.274	0.00693
Water Solubility, mg/L	1E+06	1E+06

Physicochemical similarity based on experimental data

QSAR Toolbox 4.4 Database version: 4.4 QSAR TOOLBOX

Chemicals category

- Physical Chemical Properties#Boiling point
- Physical Chemical Properties#Partition Coefficient:#N-Octanol/Water
- Physical Chemical Properties#Vapour pressure
- Physical Chemical Properties#Water solubility

Table with selected physicochemical experimental data

	1 CAS# 56-18-8	2 CAS# 107-15-3
Physical Chemical Properties/Boiling point		117 °C
Physical Chemical Properties/Partition Coefficient:/N-Octanol/Water		-2.04
Physical Chemical Properties/Vapour pressure		12 mm Hg
Physical Chemical Properties/Water solubility		1E+06 mg/L

	3 CAS# 111-40-0	4 CAS# 111-41-1
Physical Chemical Properties/Boiling point	207 °C	239 °C
Physical Chemical Properties/Partition Coefficient:/N-Octanol/Water		
Physical Chemical Properties/Vapour pressure	0.232 mm Hg	0.00819 mm Hg
Physical Chemical Properties/Water solubility	1E+06 mg/L	1E+06 mg/L

2.2. Structural similarity

Structural similarity

- Structure similarity profilers
  - Organic functional groups
  - Structure similarity

Table with calculated structural similarity

Options

Mode: Hologram, CombineAllFeatures

Measure: Dice

QSAR Toolbox 4.4 Database version: 4.4 QSAR TOOLBOX

Molecular features: AtomCenteredFragments

Atom characteristics: AtomType, CountHAttached, Hybridization

Calculated structure similarity

	Chemical 1 CAS 56-18-8	Chemical 2 CAS 107-15-3	Chemical 3 CAS 111-40-0
Chemical 1 CAS 56-18-8	100%	61.5 %	87.5 %
Chemical 2 CAS 107-15-3	61.5 %	100%	72.7 %
Chemical 3 CAS 111-40-0	87.5 %	72.7 %	100%
Chemical 4 CAS 111-41-1	62.5 %	36.4 %	71.4 %

	Chemical 4 CAS 111-41-1
Chemical 1 CAS 56-18-8	62.5 %
Chemical 2 CAS 107-15-3	36.4 %
Chemical 3 CAS 111-40-0	71.4 %
Chemical 4 CAS 111-41-1	100%

Table with profiling results for "Organic functional groups"

1 CAS# 56-18-8	2 CAS# 107-15-3	3 CAS# 111-40-0
Amine, primary Amine, secondary Aliphatic amine, primary Aliphatic amine, secondary	Amine, primary Aliphatic amine, primary	Amine, primary Amine, secondary Aliphatic amine, primary Aliphatic amine, secondary

4 CAS# 111-41-1
Alcohol Amine, primary Amine, secondary Aliphatic amine, primary Aliphatic amine, secondary

QSAR Toolbox 4.4 Database version: 4.4 QSAR TOOLBOX TPRF v4.4

Table with profiling results for "Structure similarity"

Options

Mode: Hologram, CombineAllFeatures

Measure: Dice

Molecular features: AtomCenteredFragments

Atom characteristics: AtomType, CountHAttached, Hybridization

1 CAS# 56-18-8	2 CAS# 107-15-3	3 CAS# 111-40-0
61.5÷87.5 %	36.4÷72.7 %	71.4÷87.5 %

4 CAS# 111-41-1
36.4÷71.4 %

Comments on structural similarity *manually editable field*

Not provided by the user

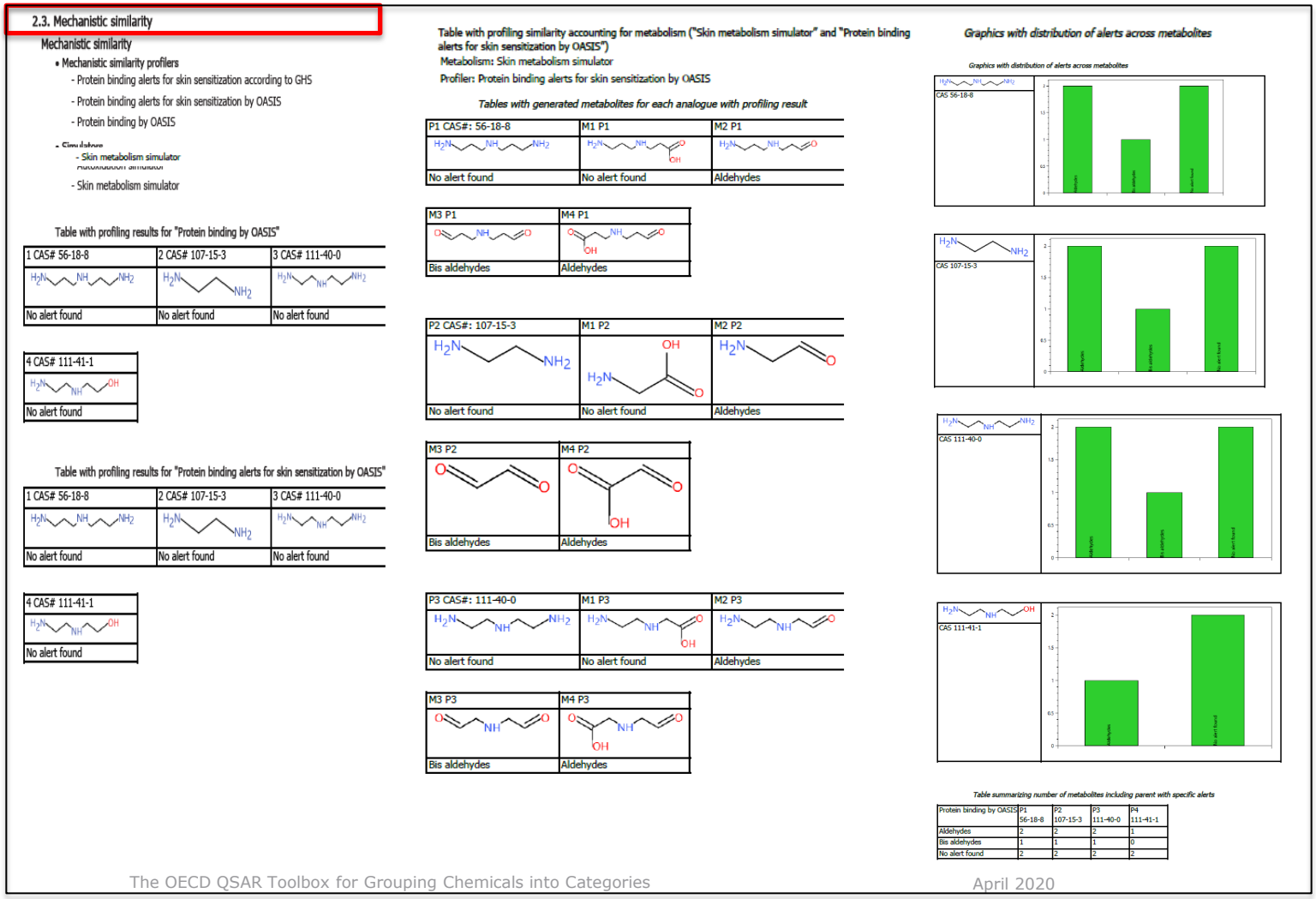
2. Consistency check section:

2.1. Physicochemical similarity – provides information for the selected physicochemical parameters (calculated values and exp data) for each member of the category

2.2. Structural similarity – provides profiling results from the selected structure-based profilers – "Structural similarity" and "OFG" for each member of the category

Category consistency elements

2. Consistency check section (cond.): Report



2.3. Mechanistic similarity

Mechanistic similarity

- Mechanistic similarity profilers
  - Protein binding alerts for skin sensitization according to GHS
  - Protein binding alerts for skin sensitization by OASIS
  - Protein binding by OASIS
- Simulators
  - Skin metabolism simulator
  - Skin metabolism simulator

Table with profiling results for "Protein binding by OASIS"

1 CAS# 56-18-8	2 CAS# 107-15-3	3 CAS# 111-40-0
No alert found	No alert found	No alert found

4 CAS# 111-41-1
No alert found

Table with profiling results for "Protein binding alerts for skin sensitization by OASIS"

1 CAS# 56-18-8	2 CAS# 107-15-3	3 CAS# 111-40-0
No alert found	No alert found	No alert found

4 CAS# 111-41-1
No alert found

Table with profiling similarity accounting for metabolism ("Skin metabolism simulator" and "Protein binding alerts for skin sensitization by OASIS")

Metabolism: Skin metabolism simulator

Profiler: Protein binding alerts for skin sensitization by OASIS

Tables with generated metabolites for each analogue with profiling result

P1 CAS#: 56-18-8	M1 P1	M2 P1
No alert found	No alert found	Aldehydes

M3 P1	M4 P1
Bis aldehydes	Aldehydes

P2 CAS#: 107-15-3	M1 P2	M2 P2
No alert found	No alert found	Aldehydes

M3 P2	M4 P2
Bis aldehydes	Aldehydes

P3 CAS#: 111-40-0	M1 P3	M2 P3
No alert found	No alert found	Aldehydes

M3 P3	M4 P3
Bis aldehydes	Aldehydes

Graphics with distribution of alerts across metabolites

Graphics with distribution of alerts across metabolites

CAS 56-18-8

CAS 107-15-3

CAS 111-40-0

CAS 111-41-1

Table summarizing number of metabolites including parent with specific alerts

Protein binding by OASIS	P1 56-18-8	P2 107-15-3	P3 111-40-0	P4 111-41-1
Aldehydes	2	2	2	1
Bis aldehydes	1	1	1	0
No alert found	2	2	2	2

2. Consistency check section:

2.3. Mechanistic similarity – provides information for the selected profiles (assigned as suitable with respect to the target endpoint and/or manually added by the user to justify the consistency). This section includes:

- Profiling results for all category members
- Tables for parents and generated metabolites (if any)
- Graphics with distribution of alerts across metabolites
- Tables summarizing the number of metabolites (including parent) with specific alerts

Category consistency elements

2. Consistency check section (cond.):

Report

2.4. Additional endpoints

Additional endpoints

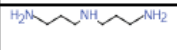
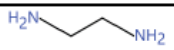
- Human Health Hazards#Sensitisation#[s]Organ:Skin#[s]Type of method:in Vivo#[s]Assay:LLNA#[s]Endpoint:EC3

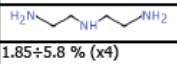
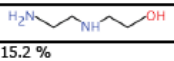
NOTE:

[s] - string metadata

[v] - numerical metadata

Table with selected endpoint data values

	1 CAS# 56-18-8	2 CAS# 107-15-3
		
Human Health Hazards/Sensitisation: Skin, in Vivo, LLNA, EC3	0.882÷0.9 % (x2)	2.2÷2.7 % (x4)

	3 CAS# 111-40-0	4 CAS# 111-41-1
		
Human Health Hazards/Sensitisation: Skin, in Vivo, LLNA, EC3	1.85÷5.8 % (x4)	15.2 %

Comments on additional endpoints *manually editable field*

Not provided by the user

2. Consistency check section:

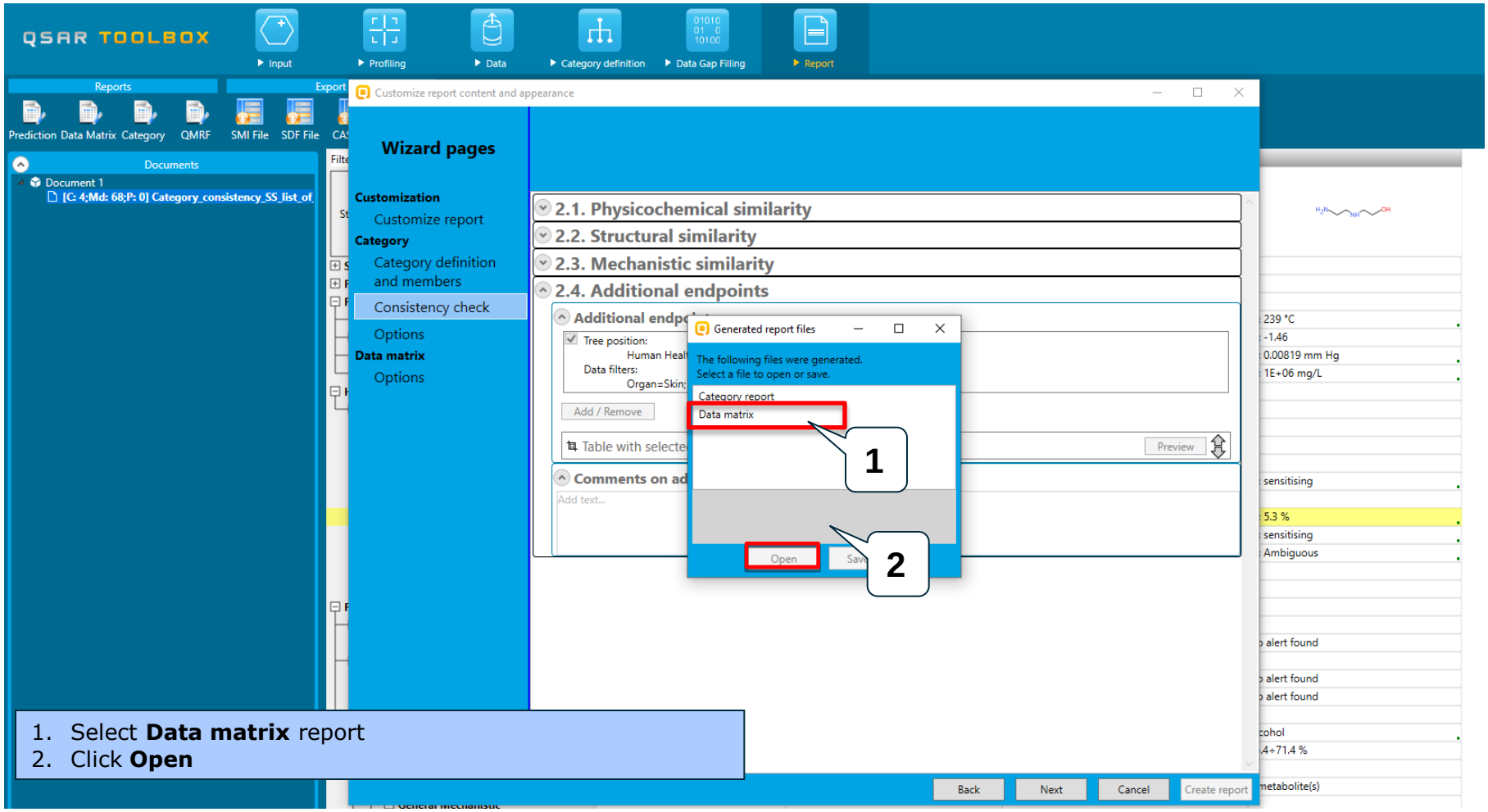
2.4. Additional endpoints – provides experimental data for all category members related to the additionally added endpoints

The number in brackets shows how many times the same experimental result is recorded

Category consistency elements

Report

Data matrix



1. Select **Data matrix** report
2. Click **Open**

Category consistency elements

Report

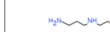
Data matrix report

		Chemical #1	Chemical #2	Chemical #3	Chemical #4
Substance identity					
Structure					
CAS number		56-18-8	107-15-3	111-40-0	111-41-1
Chemical name		Iminobis-3-propylamine	Ethylenediamine		
Other identifier					
SMILES		NCCCNCCCN	NCCN		
Parameters	**unit**				
Boiling point	°C	228	103		
log Kow		-1.15	-1.62		
Molecular Weight	Da	131	60.1		
Vapor Pressure (Antoine method)	mm Hg	0.0913	19		
Water Solubility	mg/L	1E+06	1E+06	1E+06	1E+06
Profilers					
General Mechanistic					
Protein binding by OASIS, with Autoxidation		No alert found	No alert found	No alert found	No alert found
Protein binding by OASIS, with Skin metabolism simulator		No alert found; Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base	No alert found; Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base	No alert found; Schiff base formation >> Schiff base formation with carbonyl compounds	No alert found; Schiff base formation >> Schiff
Endpoint Specific					
Protein binding alerts for skin sensitization by		No alert found	No alert found		
Protein binding alerts for skin sensitization by OASIS, with Skin metabolism simulator		No alert found; Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base formation with carbonyl compounds	No alert found; Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base formation with carbonyl compounds >> Bis	No alert found; Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base formation with carbonyl compounds	No alert found; Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes
Protein binding alerts for skin sensitization		No alert found	No alert found	No alert found	No alert found
Protein binding alerts for skin sensitization according to GHS, with Skin metabolism simulator		Skin sensitization Category 1B >> Aldehydes; Skin sensitization Category 1A >> Bis	Skin sensitization Category 1B >> Aldehydes; Skin sensitization Category 1A >> Bis	Skin sensitization Category 1B >> Aldehydes; Skin sensitization Category 1A >> Bis	No alert found; Skin sensitization Category 1B >> Aldehydes
Empiric					

Parameters values and profilers used in the workflow

Profilers accounting for metabolism if applied

Measured data for the chemicals in the category

		Chemical #1	Chemical #2	Chemical #3	Chemical #4
Substance identity					
Structure					
CAS number		56-18-8	107-15		
Chemical name		Iminobis-3-propylamine	Ethylenediamine	NCCNCCN	N-(Aminoethyl)ethanolamine
Other identifier					
SMILES		NCCCNCCCN	NCCN	NCCNCCN	NCCNCCO
Measured and predicted data					
Physical Chemical					
sublevel	*endpoint*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*
Boiling point	Boiling point	117 °C Boiling point SRC	207 °C Boiling point SRC	239 °C Boiling point SRC	
Physical Chemical Properties#Partition Coefficient:#N-					
sublevel	*endpoint*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*
N-Octanol/Water	Log P	-2.04 Log P SRC			
Physical Chemical					
sublevel	*endpoint*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*
Vapour pressure	Vapor Pressure	12 mm Hg Vapor Pressure SRC	0.232 mm Hg Vapor Pressure SRC	0.00819 mm Hg Vapor Pressure SRC	
Physical Chemical Properties#Water					
sublevel	*endpoint*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*	*value* *unit* *species, duration, test type, type of method, assay, strain, test guideline, year, reference*
Water solubility	Water solubility	1E+06 mg/L Water solubility SRC	1E+06 mg/L Water solubility SRC	1E+06 mg/L Water solubility SRC	
Human Health					

Congratulations!

- You have now been introduced to the defining of a target endpoint;
- You have now been introduced to the elements of the category consistency;
- You have now been introduced to the Report basket;
- Note, proficiency comes with practice!