

## 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

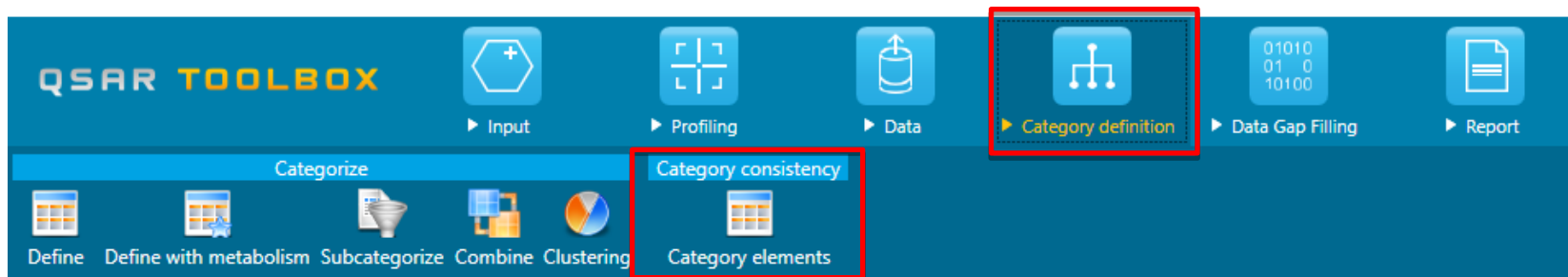
The profilers and metabolism simulators highlighted as “suitable” for the target endpoint are selected by default (e.g. for target endpoint Skin sensitization/LLNA/EC3 all *Protein binding alerts for skin sensitization* profilers along with the simulators for *Autoxidation* and *Skin metabolism* are 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

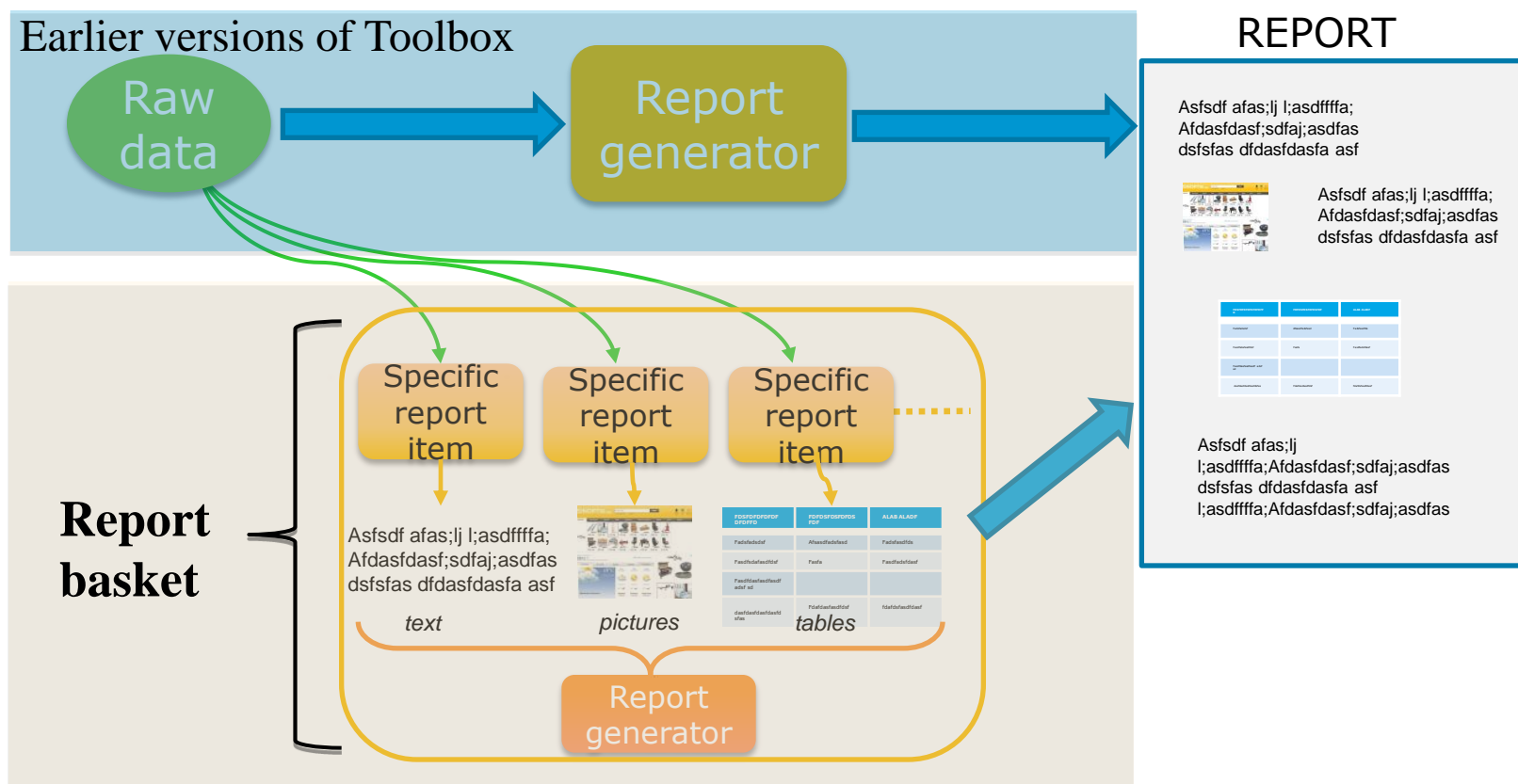
- “Category consistency” is located in *Category definition* module
- It allows to filter and collect physico-chemical properties, structural similarity, profiling results accounting for metabolic activation and experimental data
- All selected category elements are automatically stored in the so-called “*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*

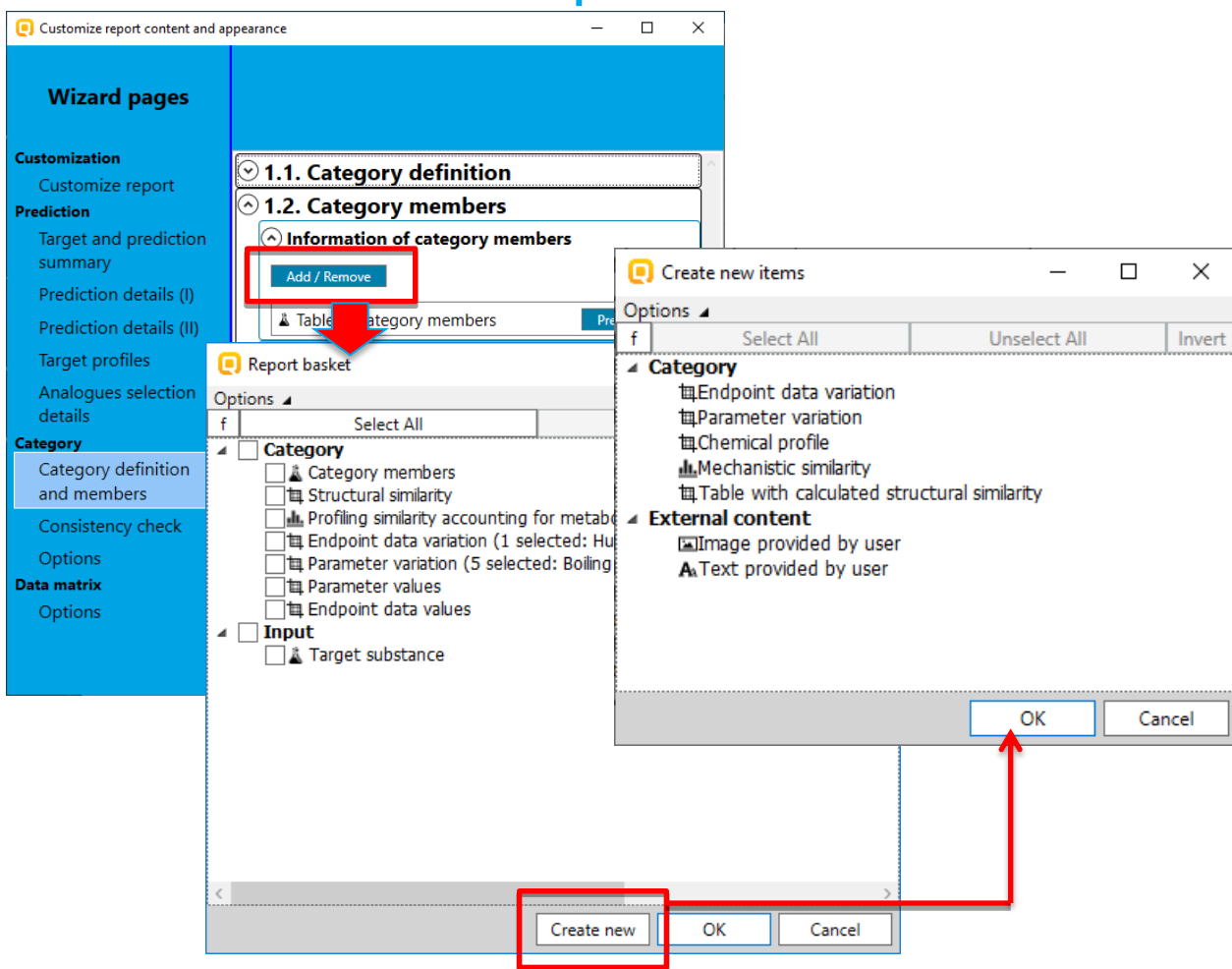
## Example of the specific report items and their description

#	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**

The screenshot shows the QSAR Toolbox software interface. The 'List' menu is open, and the 'From examples folder' option is selected. The 'Define' button in the 'Category definition' menu is highlighted. The 'Select endpoint' dialog box is open, showing the 'Environmental Fate and Transport' category. The 'Ready Biodegradability' endpoint is selected, and the 'Test guideline' is set to 'OECD 301C' and the 'Endpoint' is set to 'BOD'. The 'Selection of additional metadata fields' section is also visible.

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*

The screenshot displays the QSAR Toolbox interface. The top toolbar includes buttons for Input, Profiling, Data, Category definition (highlighted with a red box and callout 1), and Report. Below the toolbar, the 'Category consistency' section is active, showing 'Category elements' (highlighted with a red box and callout 2). The 'Filter endpoint tree...' panel on the left shows a hierarchical tree of endpoints, with 'Biodegradation' and 'BOD' highlighted. The 'Grouping methods' panel on the bottom left shows a list of methods, including 'Suitable' and 'Plausible'. The 'Wizard pages' panel on the right is open, showing 'Physicochemical similarity' and '2D/3D parameters' (highlighted with a red box and callout 3). The '2D/3D parameters' section lists various properties like Boiling point, log Kow, Molecular Weight, Vapor Pressure (Antoine method), and Water Solubility.

### 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.

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

# Category consistency elements

## Example 1 – Fate endpoint

Category element: **Physicochemical similarity**

The screenshot displays the QSAR Toolbox interface with the 'Category consistency wizard' dialog box open. The 'Physicochemical similarity' option is selected in the 'Wizard pages' list. The '2D/3D parameters' section shows a list of parameters: Boiling point, log Kow, Molecular Weight, Vapor Pressure (Antoine method), and Water Solubility. The 'Physico-chemical data' section shows a list of properties: Boiling point, Partition Coefficient: N-Octanol/Water, Vapour pressure, and Water solubility. A callout box on the right lists the default selection of 2D parameters: Boiling point, logKow, Molecular weight, Vapor pressure, and Water solubility.

**Default selection of 2D parameters (calculated values and exp. data) are as follows:**

- ✓ Boiling point
- ✓ logKow
- ✓ Molecular weight
- ✓ Vapor pressure
- ✓ Water solubility

# Category consistency elements

## Example 1 – Fate endpoint

### Category element: **Physicochemical similarity**

The user can select other parameters or experimental phys-chem data for the analysis from the lists

The default selections can be changed by clicking **Add/Remove** button (1). In this example keep the default selections. Move to Structural similarity. The selections will be preserved (see next slide)

1

2D/3D parameters

- Parameters
  - 2D
    - Boiling point
    - log Kow
    - Molecular Weight
    - Vapor Pressure (Antoine method)
    - Water Solubility

Physico-chemical data

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

Select

Select All Unselect All

Parameters

- 2D
  - (Q) Acidic pKa (Chemaxon)
  - (Q) Basic pKa (Chemaxon)
  - Acidic pKa (OASIS Consensus)
  - Acidic pKa (OASIS Electric)
  - Acidic pKa (OASIS Regression)
  - Amino acids pKa (OASIS Regression)
  - BAF
  - BAF (lower trophic)
  - BAF (mid trophic)
  - BAF (upper trophic)
  - BAF (upper trophic, biotransformation rate is zero)

Select

Select All Unselect All

Physical Chemical Properties

- Autoflammability / Self-ignition temperature
- Boiling point
- Chemical reactivity
- Density
- Dissociation Constant (pKa)
- Explosive properties
- Flammability
- Flash point
- Melting / freezing point
- Oxidation reduction potential
- Oxidizing properties
- Particle size
- Partition Coefficient:
  - Gas/Atmospheric particulates
  - N-Octanol/Air
  - N-Octanol/Water
  - Solubility in organic solvents / fat solubility
  - Stability in organic solvents and identity of relevant degradation
  - Surface tension
  - Vapour pressure
  - Viscosity
  - Water solubility

# Category consistency elements

## Example 1 – Fate endpoint

### Category element: **Structural similarity**

The screenshot shows the QSAR Toolbox interface with the 'Category consistency' wizard open. The 'Category definition' button in the top toolbar is highlighted with a red box. The wizard's 'Wizard pages' list on the left has 'Structural similarity' selected, also highlighted with a red box and labeled with a '1'. The 'Empiric profiles' list on the right shows 'Structure similarity' selected, labeled with a '2'. A red arrow points from the '2' label to a text box explaining that selections are preserved. Another text box at the bottom right lists the default selection of profilers for structural similarity: 'Organic functional groups' and 'Structure similarity'. A third text box at the bottom left provides a three-step instruction: 1. Select 'Structural similarity' section, 2. Keep the default selections, 3. Move to 'Mechanistic similarity'.

**Wizard pages**

- Physicochemical similarity
- Structural similarity**
- Mechanistic similarity
- (Eco)tox experimental data
- Options

**Empiric profiles**

Options ▾ Select All Unselect All

- ☐ Plausible
  - ☐ Chemical elements
  - ☐ Groups of elements
  - ☐ Lipinski Rule of Five
  - ☒ Organic functional groups
  - ☐ Organic functional groups (nested)
  - ☐ Organic functional groups (US EPA)
  - ☐ Organic functional groups, Norbert Haider (checkmol)
- ☒ Structure similarity
- ☐ Unclassified

**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**

**Category consistency wizard**

The mechanistic similarity within a category can be assessed by using mechanistic and endpoint specific profilers, and metabolic simulators available in the Toolbox.

**Wizard pages**

- Physicochemical similarity
- Structural similarity
- Mechanistic similarity**
- (Eco)tox experimental data
- Options

**Mechanistic profiles**

Profiles

Options

Select All Unselect All

- ☐ Suitable
- ☒ Biodegradation fragments (BioWIN MITI)
- ☒ Plausible
- ☒ Aquatic toxicity classification by ECOSAR
- ☐ Biodeg BioHC half-life (Biowin)
- ☐ Biodegradation primary (Biowin 4)
- ☐ Biodegradation probability (Biowin 1)
- ☐ Biodegradation probability (Biowin 2)
- ☒ Biodegradation probability (Biowin 5)
- ☒ Biodegradation probability (Biowin 6)
- ☐ Biodegradation ultimate (Biowin 3)
- ☐ OECD HPV Chemical Categories
- ☐ Substance type
- ☐ US-EPA New Chemical Categories
- ☐ Unclassified

**Metabolisms**

Options

Select All Unselect All Invert

- ☐ Unclassified

**1** 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)

**2** 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**

The screenshot shows the QSAR Toolbox interface with the 'Category consistency wizard' open. The 'Data' tab is selected in the top menu. The 'Wizard pages' list on the left includes 'Physicochemical similarity', 'Structural similarity', 'Mechanistic similarity', '(Eco)tox experimental data' (highlighted with a red box and labeled '1'), and 'Options'. The 'Experimental data' page (labeled '2') shows a tree position: 'Environmental Fate and Transport#Biodegradation#Biodegradation in Water: Screening Tests#Ready Biodegradation'. The 'Data filters' section contains 'Test guideline=OECD 301C; Endpoint=BOD;'. At the bottom right, the 'OK' button is highlighted with a red box and labeled '3'. A red callout box on the right states: '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

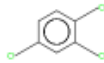
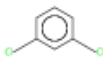


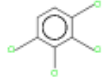
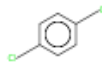

3. Finally, click **OK** on Category consistency wizard window in order to execute the consistency elements

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

# 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

Filter endpoint tree...	1	2	3	4	5	6	7
Structure							
Selected calculated parameters and experimental data appear on DM							
Structure info							
Parameters							
2D							
Boiling point	207 °C	175 °C	238 °C	175 °C	238 °C	175 °C	266 °C
log Kow	3.93	3.28	4.57	3.28	4.57	3.28	5.22
Molecular Weight	181 Da	147 Da	216 Da	147 Da	216 Da	147 Da	250 Da
Vapor Pressure (Antoine method)	0.196 mm Hg	1.48 mm Hg	0.00264 mm Hg	1.05 mm Hg	0.0139 mm Hg	0.793 mm Hg	0.00156 mm Hg
Water Solubility	20 mg/L	75.6 mg/L	3.94 mg/L	92 mg/L	4.27 mg/L	90.7 mg/L	0.000 mg/L
Physical Chemical Properties							
Boiling point	7/36 M: 213 °C	M: 172 °C	M: 245 °C	M: ca.179 °C	M: 254 °C	M: 173 °C	M: 277 °C
Partition Coefficient:	7/32 M: 4.02	M: 3.44	M: 4.64	M: ca.3.34	M: 4.6	M: 3.37	M: 5.17
Vapour pressure	7/56 M: 0.13 kPa	M: 0.1 hPa	M: 0.0054 mm Hg	M: ca.1 mm Hg	M: 0.039 mm Hg	M: 0.53 hPa	M: 0.00101 mm...
Water solubility	7/43 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

\* 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.*

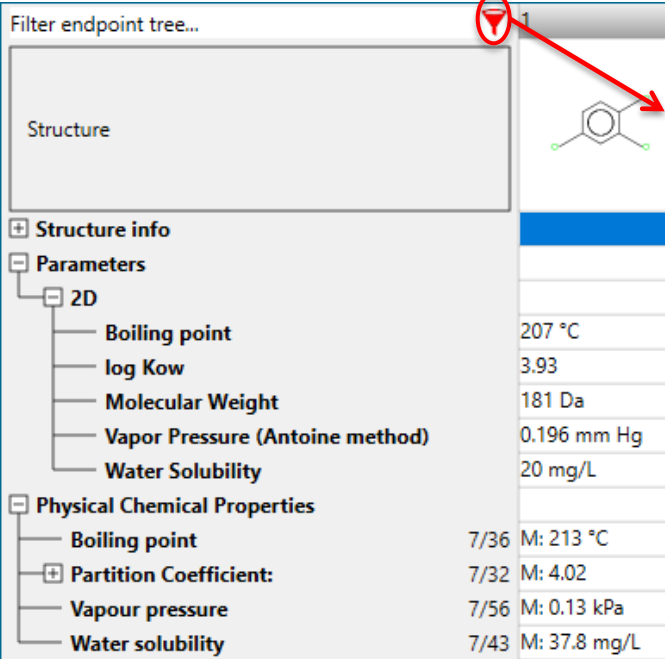


# Category consistency elements

## Example 1 – *Fate endpoint*

### Filtering of DM

### Button for filtering the endpoint tree



Filter endpoint tree...

Structure

**Structure info**

**Parameters**

☒ 2D

Boiling point 207 °C

log Kow 3.93

Molecular Weight 181 Da

Vapor Pressure (Antoine method) 0.196 mm Hg

Water Solubility 20 mg/L

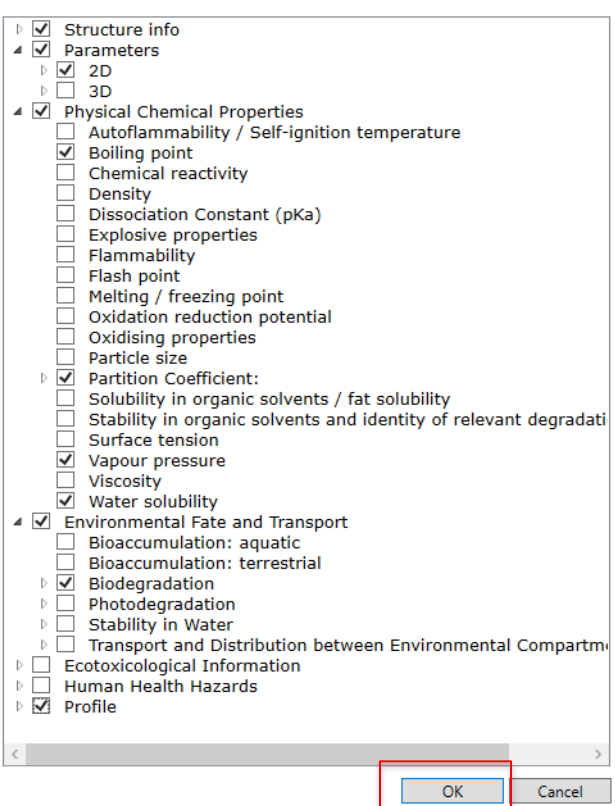
**Physical Chemical Properties**

Boiling point 7/36 M: 213 °C

**Partition Coefficient:** 7/32 M: 4.02

Vapour pressure 7/56 M: 0.13 kPa

Water solubility 7/43 M: 37.8 mg/L



☒ Structure info

☒ Parameters

☒ 2D

☒ 3D

**Physical Chemical Properties**

☐ Autoflammability / Self-ignition temperature

☒ Boiling point

☐ Chemical reactivity

☐ Density

☐ Dissociation Constant (pKa)

☐ Explosive properties

☐ Flammability

☐ Flash point

☐ Melting / freezing point

☐ Oxidation reduction potential

☐ Oxidising properties

☐ Particle size

☒ Partition Coefficient:

☐ Solubility in organic solvents / fat solubility

☐ Stability in organic solvents and identity of relevant degradati

☐ Surface tension

☒ Vapour pressure

☐ Viscosity

☒ Water solubility

**Environmental Fate and Transport**

☐ Bioaccumulation: aquatic

☐ Bioaccumulation: terrestrial

☒ Biodegradation

☐ Photodegradation

☐ Stability in Water

☐ Transport and Distribution between Environmental Compartm

☐ Ecotoxicological Information

☐ Human Health Hazards

☒ Profile

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)








	175 °C	266 °C
	3.28	5.22
	147 Da	250 Da
m Hg	0.793 mm Hg	0.00156 mm Hg
L	90.2 mg/L	0.888 mg/L
C	M: 173 °C	M: 277 °C
	M: 3.37	M: 5.17
mm Hg	M: 0.53 hPa	M: 0.00101 mm...
ng/L	M: 31.7 mg/L	M: 0.831 mg/L

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)

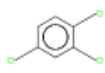
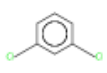
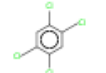
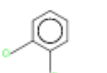
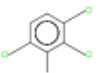
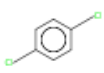
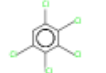
Filter endpoint tree...	1	2	3	4	5	6	7	
Structure								
<b>Structure info</b>								
<b>Parameters</b>								
<b>Physical Chemical Properties</b>	7/167	M: 37.8 mg/L	M: 71 mg/L	M: 0.595 mg/L	M: ca.92.3 mg/L	M: 5.92 mg/L	M: 31.7 mg/L	M: 0.831 mg/L
<b>Environmental Fate and Transport</b>	6/6		M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %
<b>Profiling</b>								
<b>Endpoint Specific</b>								
<b>Empiric</b>								
<b>Organic functional groups</b>	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	
<b>Structure similarity [category]</b>	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)

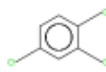
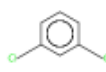
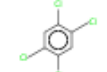
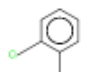
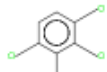
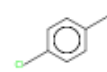
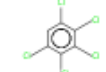
Filter endpoint tree...	1	2	3	4	5	6	7	
Structure								
+ Structure info								
+ Parameters								
+ Physical Chemical Properties	7/167	M: 37.8 mg/L	M: 71 mg/L	M: 0.595 mg/L	M: ca.92.3 mg/L	M: 5.92 mg/L	M: 31.7 mg/L	M: 0.831 mg/L
+ Environmental Fate and Transport	6/6	M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %	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	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	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	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	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

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								
Biodegradation								
Biodegradation in Water: Screening Te...								
Ready Biodegradability								
OECD 301C								
BOD	6/6	M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %	M: 0 %
OECD Guideline 301 C (Ready...	4/6	M: 0 %	M: 0 %	M: 0 %		M: 0 %		
OECD Guideline 301 D (Ready...	1/3					M: 1.4 %		
Other Test Guideline	1/2							
Undefined Test Guideline	2/11		M: 77 %					

BOD data appeared on DM if available for the chemicals from the category

BOD data appeared on DM if available for the chemicals from the category

# 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**

The screenshot shows the QSAR Toolbox interface with three numbered steps:

- 1**: The 'List' menu is highlighted in the top toolbar.
- 2**: The 'Define' button in the 'Target Endpoint' section is highlighted.
- 3**: The 'Select endpoint' dialog box is open, showing the 'Ecotoxicological Information' tree. The 'Aquatic Toxicity' endpoint is selected, and the 'Mortality' effect is chosen. The 'Duration' is set to '96 h', the 'Test organisms (species)' is 'Pimephales promelas', and the 'Endpoint' is 'LC50'. The 'Finish' button is highlighted.

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*

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The 'Category definition' module is active, showing a 'Filter endpoint tree...' on the left and a grid of chemical structures on the right. The 'Category consistency wizard' dialog box is open, displaying 'Wizard pages' and '2D/3D parameters'.

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*

The screenshot shows the QSAR Toolbox interface with the 'Category consistency wizard' dialog box open. The wizard is in the 'Mechanistic profiles' step. The 'Suitable' section is selected, and the 'Plausible' section is also visible. The 'Options' section is at the bottom. Callout boxes 1, 2, and 3 highlight specific elements: 1 points to the 'US-EPA New Chemical Categories' checkbox, 2 points to the 'Suitable' section header, and 3 points to the 'OK' button.

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

Filter endpoint tree...	1	2	3	4	5	6	7	8	9	10	11	12
Structure												
Structure info	Calculated and experimental data											
Parameters	Calculated phys-chem values											
2D												
Boiling point	127 °C	36.9 °C	237 °C	255 °C	271 °C	199 °C	178 °C	103 °C	112 °C	218 °C	157 °C	192 °C
log Kow	-1.13	-0.15	4.27	4.76	5.25	3.29	2.8	-1.62	1.33	3.78	2.31	0.35
Molecular Weight	74.1 Da	45.1 Da	171 Da	185 Da	199 Da	143 Da	129 Da	60.1 Da	87.2 Da	157 Da	115 Da	116 Da
Vapor Pressure (Antoine method)	6.93 mm Hg	1.03E+03 mm Hg	0.0432 mm Hg	0.00642 mm Hg	0.00613 mm Hg	0.35 mm Hg	1.07 mm Hg	19 mm Hg	32.8 mm Hg	0.136 mm Hg	3.3 mm Hg	0.215 mm Hg
Water Solubility	1E+06 mg/L	1E+06 mg/L	139 mg/L	45.1 mg/L	14.6 mg/L	1.28E+03 mg/L	3.15E+03 mg/L	1E+06 mg/L	6.88E+04 mg/L	423 mg/L	6.79E+03 mg/L	5.34E+05 mg/L
Physical Chem	Exp. phys-chem data											
Boiling point		5 °C	M: 242 °C	M: 247 °C	M: 276 °C	M: 202 °C	M: 178 °C	M: 116+117	M: 104 °C	M: 221 °C	M: 156 °C	M: 94.2 °C
Partition Coefficient:	9/31 M: -1.43	M: -0.27					M: 2.9	M: -2.04	M: 1.49		M: 2.57	
Vapour pressure	13/55 M: 4.11 hPa	M: 990 hPa		M: 0.0079 Pa		M: 0.278 mm Hg	M: 0.81 hPa	M: 12 mm Hg	M: 30 mm Hg	M: 0.1 mm Hg	M: 2.74 mm Hg	M: 0.118 mm Hg
Water solubility	10/26	M: 1E+06 mg/L		M: 13 mg/L			M: 200 mg/L	M: 1E+06 mg/L	M: 1E+06 mg/L	M: 550 mg/L		M: 6.37E+05 mg/L
Ecotoxicological Information	Experimental endpoint data											
Aquatic Toxicity	AW SW 15/41	M: 1.19E+03 (1.0...	M: 226 mg/L	M: 0.21 mg/L	M: 0.102 mg/L	M: 0.0654 (0.056...	M: 2.16 (2.02+2...	M: 5.15 mg/L	M: >11.5 mg/L	M: 177		
Profiling	Mechanistic and structure-based profiling results											
Predefined												
US-EPA New Chemical Categories	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines
Endpoint Specific												
Acute aquatic toxicity classification...	Class 5 (Not pos...	Class 5 (Not pos...	Class 2 (less inert...	Class 2 (less inert...	Class 2 (less inert...	Class 2 (less inert...	Class 2 (less inert...	Class 5 (Not pos...	Class 2 (less inert...	Class 2 (less inert...	Class 2 (less inert...	Class 2 (less inert...
Acute aquatic toxicity MOA by OASIS	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine
Aquatic toxicity classification by EC...	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines
Empiric												
Organic functional groups	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...	Aliphatic amine,...
Structure similarity [category]	25+88.9 %	18.2+57.1 %	25+96 %	23.5+96.3 %	22.2+96.3 %	28.6+95.2 %	30.8+94.7 %	22.2+88.9 %	10+92.3 %	26.7+95.7 %	22.2+94.1 %	10.2+75.9 %

# 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**

The screenshot shows the QSAR Toolbox software interface. The main window has a 'Documents' panel on the left and a 'Filter endpoint tree...' panel in the center. The 'Filter endpoint tree...' panel shows a tree structure with 'Human Health Hazards' expanded, and 'Sensitisation' selected. The 'Sensitisation' node is further expanded, showing 'Skin', 'in Vivo', 'LLNA', and 'EC3'. The 'EC3' node is highlighted. A 'Select endpoint' dialog box is open on the right, showing the selection of 'Human Health Hazards Sensitisation' and 'EC3'. The dialog box has a 'Define' button and a 'Finish' button. Numbered callouts 1, 2, and 3 highlight key steps: 1. Clicking 'List' in the 'Category definition' tab; 2. Clicking 'Define' in the 'Target Endpoint' tab; 3. Clicking 'Finish' in the 'Select endpoint' dialog.

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*

The screenshot displays the QSAR Toolbox interface. The top toolbar includes icons for Input, Profiling, Data, Category definition (highlighted with a red box and callout 1), Data Gap Filling, and Report. Below the toolbar, the 'Categorize' menu is open, showing options like Define, Define with metabolism, Subcategorize, Combine, Clustering, and Category elements (highlighted with a red box and callout 2). The 'Documents' panel on the left lists several documents, including 'Category\_consistency\_BOD\_Aryl\_chlorides.smi'. The 'Endpoint tree' on the right shows a hierarchy of categories, with 'Human Health Hazards' expanded to show 'Acute Toxicity', 'ADME', 'Bioaccumulation', 'Carcinogenicity', 'Developmental Toxicity / Teratogenicity', 'Genetic Toxicity', 'Immunotoxicity', 'Irritation / Corrosion', 'Neurotoxicity', 'Photoinduced toxicity', 'Repeated Dose Toxicity', 'Sensitisation', 'ToxCast', 'Toxicity to Reproduction', and 'Toxicokinetics, Metabolism and Distribution'. The 'Category consistency wizard' dialog box (callout 3) is open, showing the 'Physicochemical similarity' wizard page. It includes sections for '2D/3D parameters' (Boiling point, log Kow, Molecular Weight, Vapor Pressure, Water Solubility) and 'Physico-chemical data' (Boiling point, Partition Coefficient, Vapour pressure, Water solubility).

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*



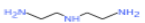

The screenshot displays the QSAR Toolbox interface. The 'Category consistency' tab is selected in the top menu. The 'Category consistency wizard' dialog is open, showing the 'Mechanistic profiles' and 'Metabolisms' sections. Callout 1 points to the '2D/3D parameters' and 'Physico-chemical data' sections. Callout 2 points to the 'Mechanistic profiles' and 'Metabolisms' sections. Callout 3 points to the 'OK' button.

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 sensi...	M: Ambiguous
Undefined Assay	1/1	M: Positive		
Undefined Type of method	1/1 M: sensitising			
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

The screenshot shows the QSAR Toolbox interface. The top toolbar has a 'Report' icon highlighted with a red box and labeled '1'. The left sidebar has a 'Category' icon highlighted with a red box and labeled '2'. Below the 'Category' icon, a file named 'Category\_consistency\_SS\_list\_of\_amines.smi' is selected, also labeled '1'. The main window displays a table of chemical properties for four chemical structures.

Filter endpoint tree...	1	2	3	4
Structure	<chem>NCCNCCN</chem>	<chem>NCCNCCN</chem>	<chem>NCCNCCN</chem>	<chem>NCCNCCN</chem>
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: 0.00819 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				
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
LLNA				
EC3	4/15	M: 0.882 %	M: 2.2 %	M: 1.85 %
Skin sensitisation	2/3		M: sensitising	M: sensitising
Miscellaneous	4/27	M: Category C	M: Ambiguous	M: Moderately sensiti...
Undefined Assay	1/1		M: Positive	M: Ambiguous

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

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The 'Report' icon is highlighted with a red box and a callout '2'. Below the menu bar, the 'Reports' section contains icons for 'Prediction Data Matrix', 'Category', 'QMRF', 'SMI File', 'SDF File', 'CAS List', and 'Data Matrix'. The 'Category' icon is highlighted with a red box and a callout '3'. In the left sidebar, the 'Documents' list shows 'Document 1', 'Document 2', and 'Document 3'. A document titled '[C: 4;Md: 68;P: 0] Category\_consistency\_SS\_list\_of\_amines.smi' is highlighted with a red box and a callout '1'. The main window displays the 'Customize report content and appearance' wizard. The 'Wizard pages' list includes 'Customize report', 'Category', 'Data matrix', and 'Options'. The 'Category' page is selected, showing sections for 'Category definition and members', 'Consistency check', 'Options', 'Data matrix', and 'Options'. The 'Category' and 'Data matrix' sections are checked. A red box highlights these sections with the text 'Sections of the Category report' and a callout '4'. At the bottom of the wizard, there is a checkbox for 'Remove password protection of the PDF files.' and a note: 'Note: If the protection is removed, this will be specified in the first page of the report'.

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

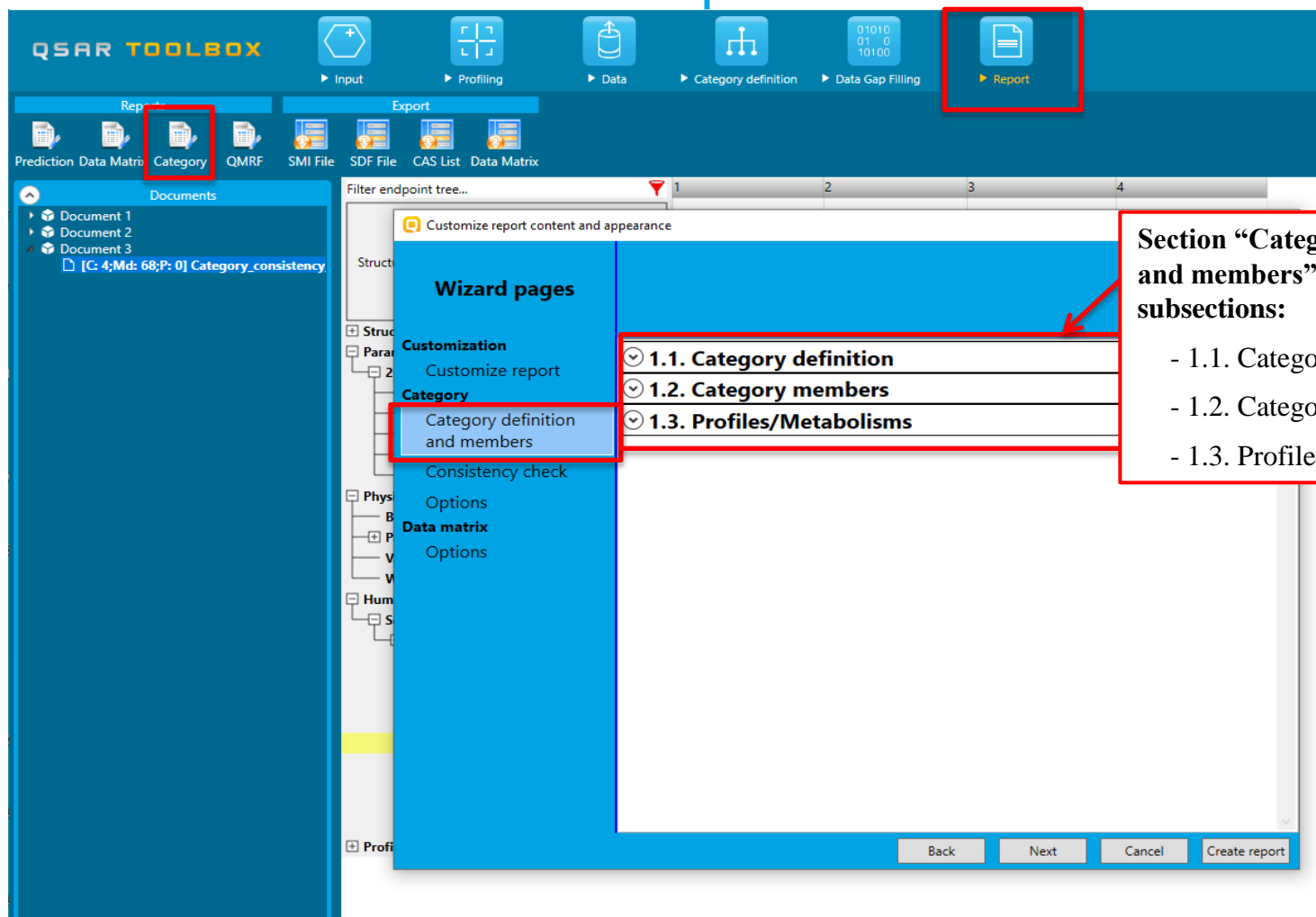
**Category part of the wizard has three sections:**

- “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

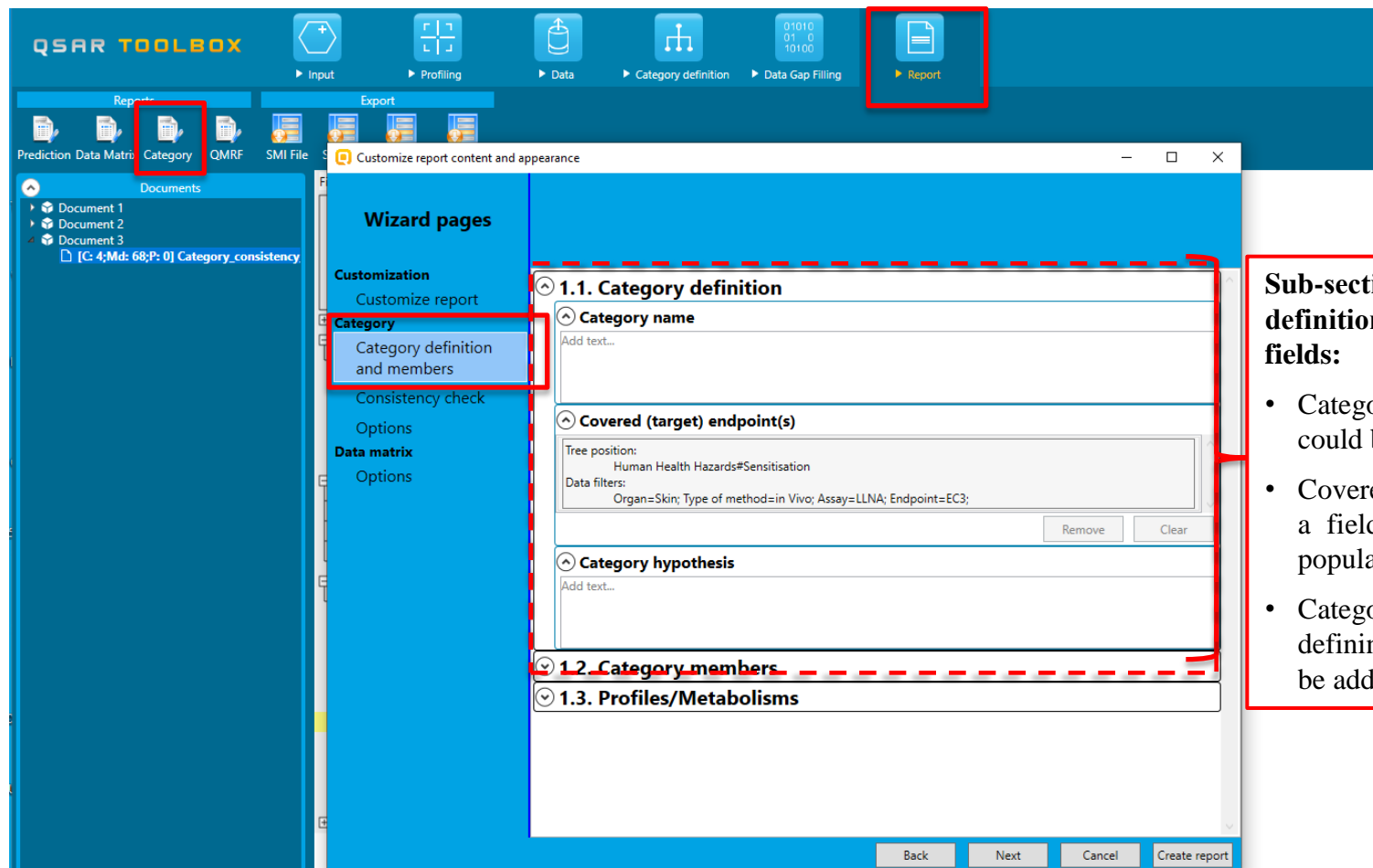


Section “Category definition and members” includes three subsections:

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

# Category consistency elements

## Report



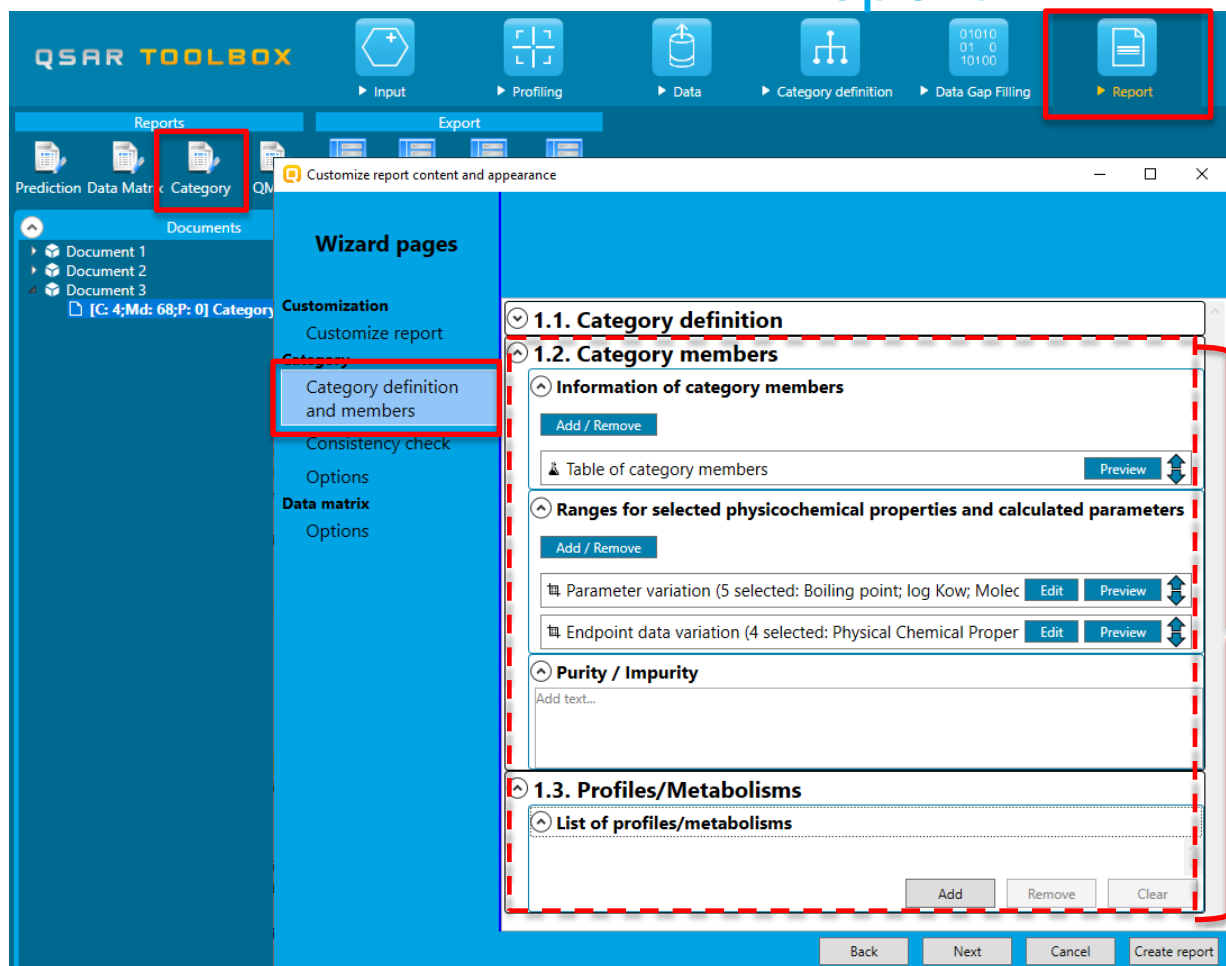
**Sub-section “1.1. Category definition” includes the following fields:**

- 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



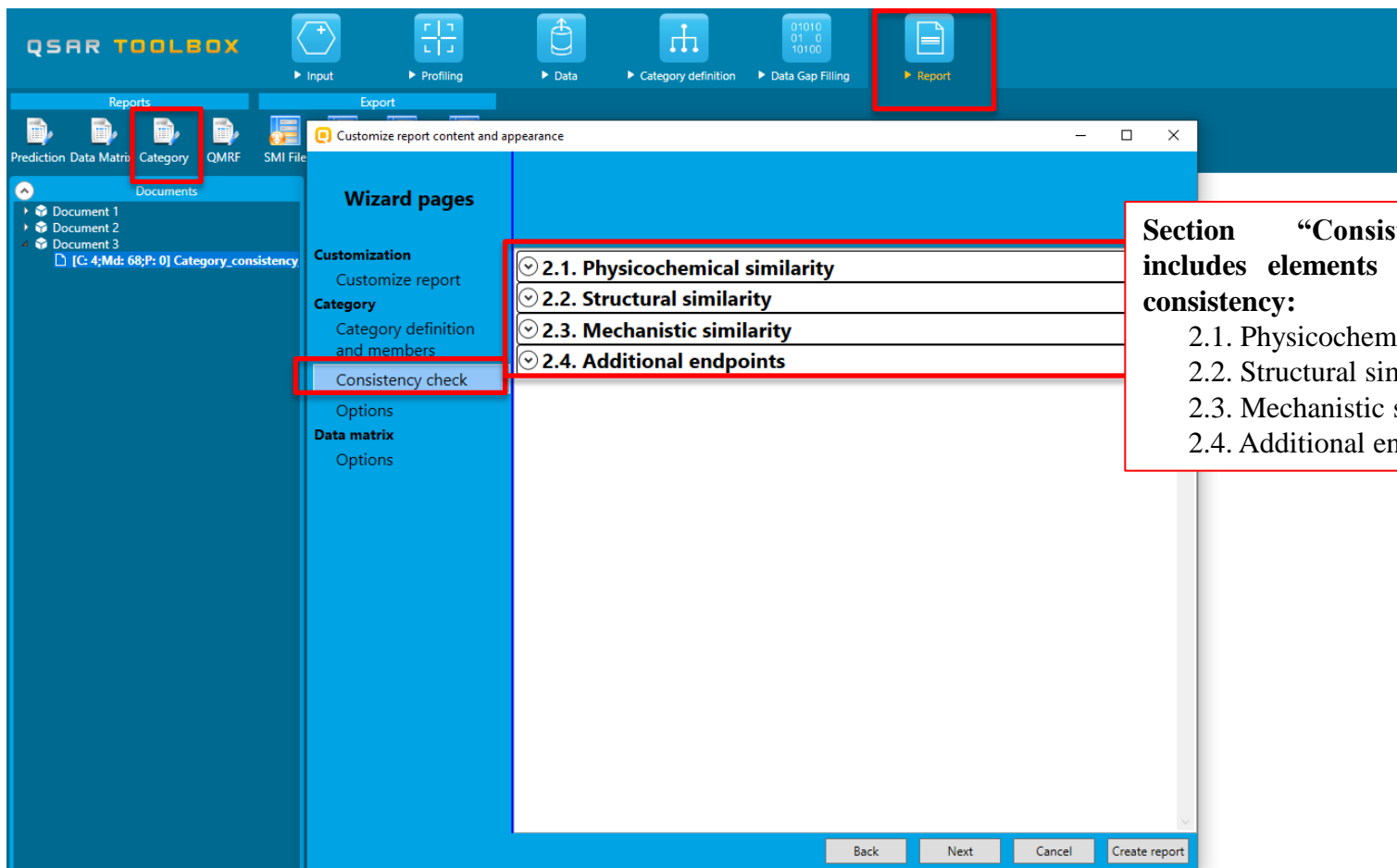
**Sub-sections “1.2. Category members” and “1.3. Profiles/metabolisms” includes fields related with:**

- 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



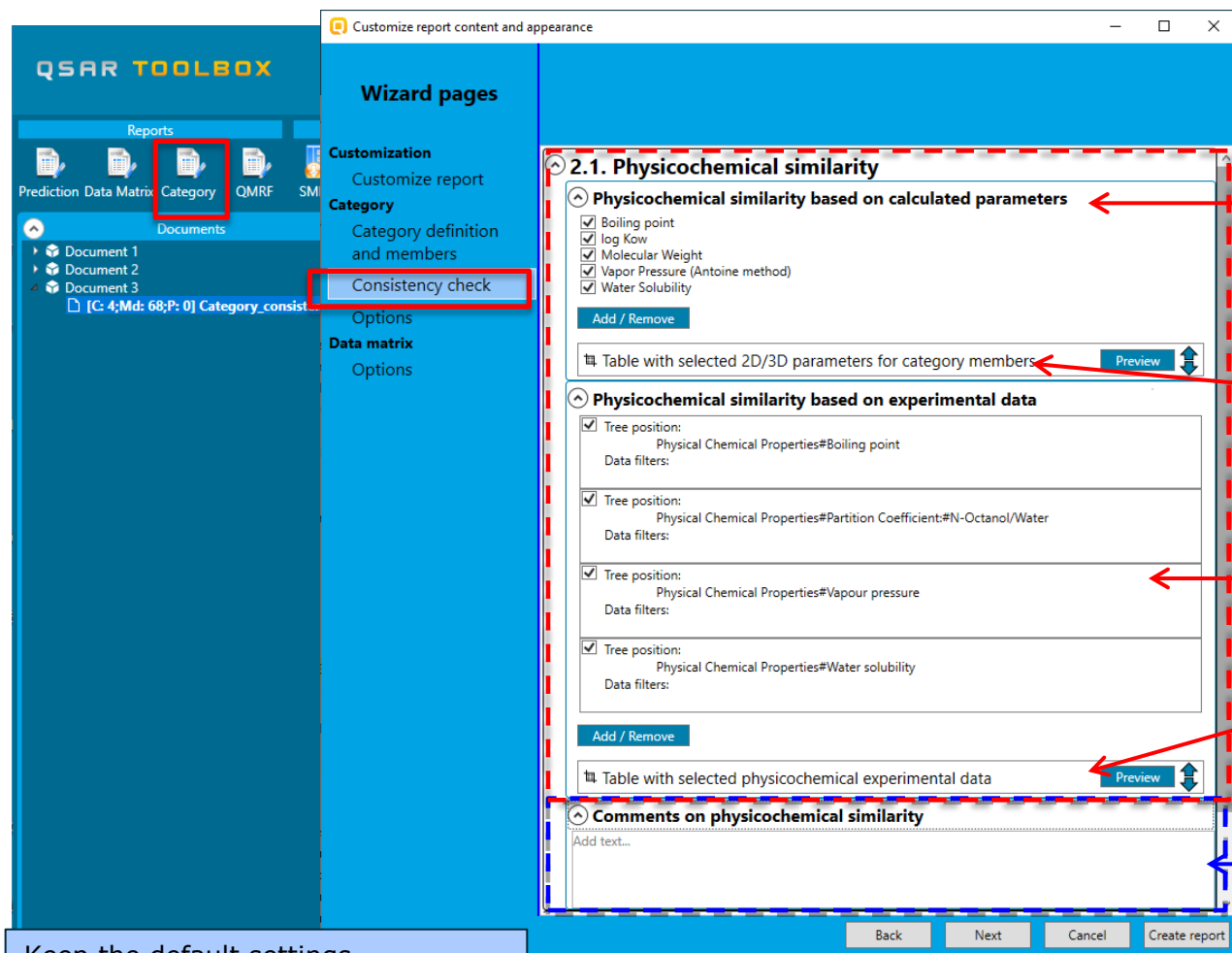
**Section “Consistency check” includes elements of the category consistency:**

- 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

Keep the default settings

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

# Category consistency elements

## Report

**Wizard pages**

- Customization
  - Customize report
  - Category**
    - Category definition and members
    - Consistency check**
    - Options
  - Data matrix
    - Options

**2.3. Mechanistic Similarity**

**Mechanistic similarity**

Mechanistic similarity profilers

Options: Select All Unselect All

- ☒ **Suitable**
  - ☒ Protein binding alerts for skin sensitization according to GHS
  - ☒ Protein binding alerts for skin sensitization by OASIS
  - ☒ Protein binding by OASIS
- ☐ **Plausible**
  - ☐ Aquatic toxicity classification by ECOSAR
  - ☐ Keratinocyte gene expression
  - ☐ OECD HPV Chemical Categories
  - ☐ Protein binding by OECD
  - ☐ Protein binding potency Cys (DPRA 13%)
  - ☐ Protein binding potency GSH
  - ☐ Protein Binding Potency h-CLAT
  - ☐ Protein binding potency Lys (DPRA 13%)
  - ☐ Respiratory sensitization
  - ☐ Substance type
  - ☐ US-EPA New Chemical Categories
- ☐ **Unclassified**

**Simulators**

Options: Select All Unselect All

- ☒ **Suitable**
  - ☒ Autoxidation simulator
  - ☒ Skin metabolism simulator
- ☐ **Plausible**
  - ☐ Autoxidation simulator (alkaline medium)
  - ☐ Desorption simulator
  - ☐ Hydrolysis simulator (neutral)
- ☐ **Unclassified**

Add / Remove

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

Table with profiling results for "Protein binding by OASIS" Edit Preview

Table with profiling similarity accounting for metabolism ("Autoxidation simulator" and "Skin metabolism simulator") Edit Preview

Table with profiling similarity accounting for metabolism ("Skin metabolism simulator" and "Autoxidation simulator") Edit Preview

**Comments on mechanistic similarity**

Add text...

Back Next Cancel Create 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

**1. Select **Category report****

**2. Click **Open****

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

# Category consistency elements

## Report

### 1. Category definition section:

Chemicals category 1 / 20

#### QSAR Toolbox report for category

##### 1. Category definition

###### 1.1. Category definition

###### Category name

Not provided by the user

###### Covered (target) endpoint(s)

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

###### Category hypothesis

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	NCCNCCCN	
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

QSAR Toolbox 4.4 QSAR TOOLBOX TPBE v4.4

Chemicals category 2 / 20

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

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

QSAR Toolbox 4.4 QSAR TOOLBOX TPBE v4.4

### 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
	<chem>NC(=O)CCNC(=O)CC</chem>	<chem>NC(=O)CCNC(=O)CC</chem>
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
	<chem>NC(=O)CCNC(=O)CC</chem>	<chem>NC(=O)CCNC(=O)CC</chem>
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
	<chem>NC(=O)CCNC(=O)CC</chem>	<chem>NC(=O)CCNC(=O)CC</chem>
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
	<chem>NC(=O)CCNC(=O)CC</chem>	<chem>NC(=O)CCNC(=O)CC</chem>
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	0.00619 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
<chem>NC(=O)CCNC(=O)CC</chem>	<chem>NC(=O)CCNC(=O)CC</chem>	<chem>NC(=O)CCNC(=O)CC</chem>
Amine, primary	Amine, primary	Amine, primary
Amine, secondary	Aliphatic amine, primary	Amine, secondary
Aliphatic amine, primary		Aliphatic amine, primary
Aliphatic amine, secondary		Aliphatic amine, secondary

4 CAS# 111-41-1
<chem>NC(=O)CCNC(=O)CC</chem>
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
<chem>NC(=O)CCNC(=O)CC</chem>	<chem>NC(=O)CCNC(=O)CC</chem>	<chem>NC(=O)CCNC(=O)CC</chem>
61.5+87.5 %	36.4+72.7 %	71.4+87.5 %

4 CAS# 111-41-1
<chem>NC(=O)CCNC(=O)CC</chem>
36.4+71.4 %

Comments on structural similarity

Not provided by the user

manually editable field

## 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

- Consistency
  - 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
<chem>NCNCCN</chem>	<chem>NCNCCN</chem>	<chem>NCNCCN</chem>
No alert found	No alert found	No alert found

4 CAS# 111-41-1
<chem>NCNCCN</chem>
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
<chem>NCNCCN</chem>	<chem>NCNCCN</chem>	<chem>NCNCCN</chem>
No alert found	No alert found	No alert found

4 CAS# 111-41-1
<chem>NCNCCN</chem>
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
<chem>NCNCCN</chem>	<chem>NCNCCN</chem>	<chem>NCNCCN</chem>
No alert found	No alert found	Aldehydes

M3 P1	M4 P1
<chem>NCNCCN</chem>	<chem>NCNCCN</chem>
Bis aldehydes	Aldehydes

P2 CAS#: 107-15-3	M1 P2	M2 P2
<chem>NCNCCN</chem>	<chem>NCNCCN</chem>	<chem>NCNCCN</chem>
No alert found	No alert found	Aldehydes

M3 P2	M4 P2
<chem>NCNCCN</chem>	<chem>NCNCCN</chem>
Bis aldehydes	Aldehydes

P3 CAS#: 111-40-0	M1 P3	M2 P3
<chem>NCNCCN</chem>	<chem>NCNCCN</chem>	<chem>NCNCCN</chem>
No alert found	No alert found	Aldehydes

M3 P3	M4 P3
<chem>NCNCCN</chem>	<chem>NCNCCN</chem>
Bis aldehydes	Aldehydes

Graphics with distribution of alerts across metabolites

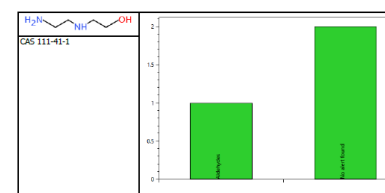
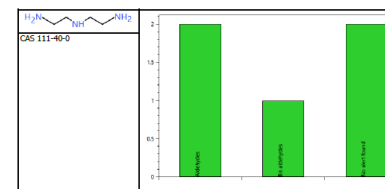
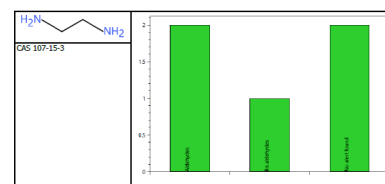
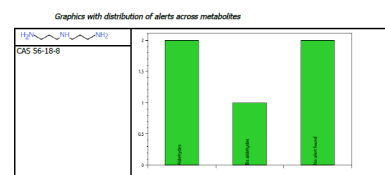


Table summarizing number of metabolites including parent with specific alerts

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

Not provided by the user

*manually editable field*

### 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

The screenshot shows the QSAR Toolbox software interface. The main window is titled 'Customize report content and appearance'. On the left, there is a 'Wizard pages' sidebar with the following sections: Customization, Category, Category definition and members, Consistency check (selected), Options, and Data matrix. The main area displays a list of report sections: 2.1. Physicochemical similarity, 2.2. Structural similarity, 2.3. Mechanistic similarity, and 2.4. Additional endpoints. A smaller dialog box titled 'Generated report files' is open over the 'Additional endpoints' section. It contains a list of generated files: 'Category report' and 'Data matrix'. The 'Data matrix' file is highlighted with a red box and a callout '1'. The 'Open' button is also highlighted with a red box and a callout '2'.





1. Select **Data matrix** report

2. Click **Open**

# Category consistency elements



## Report

### Data matrix report

A	B	C	D	E	F	G	H	I	J	K	L	M	N
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		Da		-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			
Profiles													
General Mechanistic													
Protein binding by OASIS, with Autoxidation		No alert found		No alert found		No alert found		No alert found		No alert found			
Protein binding by OASIS, with Skin metabolism simulator		Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes		Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes		Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes		Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes		Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes			
Endpoint Specific													
Protein binding alerts for skin sensitization by OASIS, with Skin metabolism simulator		No alert found		No alert found		No alert found		No alert found		No alert found			
Protein binding alerts for skin sensitization according to GH5, with Skin metabolism simulator		Skin sensitization Category 1B >> Aldehydes; Skin sensitization Category 1A >> Bis Aldehydes		Skin sensitization Category 1B >> Aldehydes; Skin sensitization Category 1A >> Bis Aldehydes		Skin sensitization Category 1B >> Aldehydes; Skin sensitization Category 1A >> Bis Aldehydes		Skin sensitization Category 1B >> Aldehydes; Skin sensitization Category 1A >> Bis Aldehydes		Skin sensitization Category 1B >> Aldehydes; Skin sensitization Category 1A >> Bis Aldehydes			
Empiric													

Parameters values and profilers used in the workflow

Profilers accounting for metabolism if applied

Chemical #1		Chemical #2		Chemical #3		Chemical #4	
Substance identity							
Structure							
CAS number		56-18-8		107-15-3		NCCNCCCN	
Chemical name		Iminobis-3-propylamine		Ethylenediamine		N,N'-bis(3-aminopropyl)ethane-1,2-diamine	
Other identifier							
SMILES		NCCCNCCCN		NCCN		NCCNCCCN	
Measured and predicted data							
Physical Chemical							
sublevel	endpoint	value	unit	value	unit	value	unit
Boiling point	Boiling point	117	°C	207	°C	239	°C
Physical Chemical Properties#Partition Coefficient: #N							
sublevel	endpoint	value	unit	value	unit	value	unit
N-Octanol/Water	Log P	-2.04					
Physical Chemical							
sublevel	endpoint	value	unit	value	unit	value	unit
Vapour pressure	Vapor Pressure	12	mm Hg	0.232	mm Hg	0.008	mm Hg
Physical Chemical Properties#Water							
sublevel	endpoint	value	unit	value	unit	value	unit
Water solubility	Water solubility	1E+06	mg/L	1E+06	mg/L	1E+06	mg/L
Human Health							

Measured data for the chemicals in the category



# 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!