

OECD QSAR Toolbox v.3.0

Step-by-step example of how to build a category for more than one target chemicals and predict acute toxicity to fish

Outlook

- **Background**
- Objectives
- Specific Aims
- The exercise
- Workflow of the exercise 1
- Workflow of the exercise 2

Background

- This is a step-by-step presentation designed to take you through the workflow of the Toolbox for evaluating an ad-hoc analogue approach.

Background

- This is a step-by-step presentation designed to take you through the workflow of the Toolbox for evaluating an ad-hoc analogue approach.
- By now you are experienced in using the Toolbox so there will be multiple key strokes between screen shots.
- By now you are experienced in using the Toolbox so there will be multiple key strokes between screen shots.

Outlook

- Background
- **Objectives**
- Specific Aims
- The exercise
- Workflow of the exercise 1
- Workflow of the exercise 2

Objectives

- To demonstrate how to use the Toolbox to evaluate whether a data gap filling with read-across from potential analogues of target chemicals is robust.

Outlook

- Background
- Objectives
- **Specific Aims**
- The exercise
- Workflow of the exercise 1
- Workflow of the exercise 2

Specific Aims

- To examine the workflow of evaluating an analogue approach.
- To introduce the user to new functionalities within selected modules.
- To explain the rationale behind each step of the exercises.
- To demonstrate with two practical examples how to use the Toolbox to evaluate whether a read-across from a potential analogue to a target chemical is robust.

Outlook

- Background
- Objectives
- Specific Aims
- **The exercise**
- Workflow of the exercise 1
- Workflow of the exercise 2

Exercise

- In this exercise we will perform two examples of ad-hoc read-across for acute toxicity to fish.
- We will do this by first entering the source and target chemicals and analysing the available data for the source chemicals.
- We will then profile the source and target chemicals and evaluate whether the read-across is robust.

Exercise

Side-Bar on the Robustness of a Potential Analogue

- **According to the OECD Guidance on Grouping of Chemicals, the following issues should be taken into account when evaluating the robustness of an analogue approach:**
 - Quality of the experimental result of the source chemical
 - Differences in functionalities in the molecules of the source and target compound (*)
 - Purity and impurity profiles
 - Differences in physical chemical properties
 - Differences in experimental results for other (eco)toxicological endpoints
 - Differences in mode of action (*)
 - Differences in toxicokinetics
- **Some of the issues above (those marked with an *) will be addressed in the current examples with the help of the Toolbox.**

Outlook

- Background
- Objectives
- Specific Aims
- The exercise
- **Workflow of the exercise 1**
- Workflow of the exercise 2

Workflow of the Exercise 1

- **As you know the Toolbox has 6 modules which are typically used in sequence:**
 - Chemical Input
 - Profiling
 - Endpoint
 - Category Definition
 - Data Gap Filling
 - Report
- **In this example we will use the modules in a different order, tailored to the aims of the example.**

Outlook

- Background
- Objectives
- Specific Aims
- The exercise
- **Workflow of the exercise 1**
 - **Chemical Input**

Chemical Input Overview

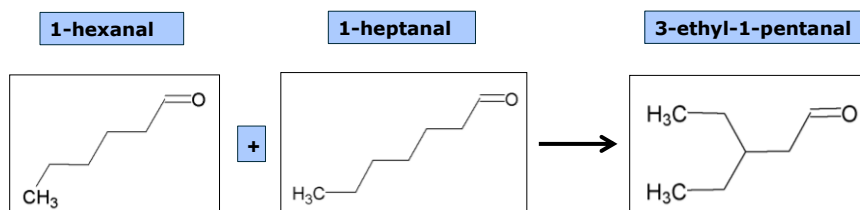
- As you know this module provides the user with several means of entering the chemical of interest or the target chemical.
- It is essential to remember that since all subsequent functions are based on chemical structure, the goal here is to make sure the molecular structure assigned to the target chemical is the correct one.

Chemical Input Ways of Entering a Chemical

- **Remember there are several ways to enter a target chemical and the most often used are:**
 - CAS#,
 - SMILES (simplified molecular information line entry system) notation, and
 - Drawing the structure.

Chemical Input Exercise 1

Read-across of acute toxicity to fish from 1-hexanal and 1-heptanal to 3-ethyl-1-pentanal.



Chemical Input Exercise 1

- In this example, we are entering the structure using the SMILES notation.
- **Click** on Structure, then
- **Enter** CCCCCC=O for n-hexanal on "SMILES/InChi" window.
- The structure is drawn simultaneously while entering the SMILES (see next screen shot).

Chemical Input

Input target chemical#1 by SMILES

1. Click on Structure; 2. Type CCCCC=O in SMILES/InChi window; 3. Click OK.

Chemical Input

Target chemical identity

The Toolbox now searches the Toolbox databases and inventories for the presence of the chemical with structure related to the current SMILES notation. It is displayed as a 2D image.

Selected	CAS/ZD	Names	IALName	ICFName	CAS/ZD
<input checked="" type="checkbox"/>	C ₆ -O=C	1: hexaldehyde	1: Modern QM1	1: Low Quality	1: High Quality
<input checked="" type="checkbox"/>	CAS: 64751	2: hexanal	1: Aquatic US	1: Aquatic US	1: AICS

In this case Toolbox found two chemicals answering the required SMILES. This panel displays QA information for presented chemicals. The user can decide which substance is to be retained for the subsequent workflow.

Chemical Input

Target chemical identity

- Click OK to add your target to data matrix
- **Click** on the box next to "Substance Identity"; this displays the chemical identification information. (see next screen shot).

Chemical Input

Target chemical identity

The screenshot shows the QSAR Toolbox software interface. The 'Target' tab is active, displaying chemical identification information. A red circle highlights the 'Substance Identity' section, which includes the following details:

Substance Identity	16-21-1
CAS Number	Enecis Number 2006245
Chemical ID	hexaldehyde
Chemical Name	hexanal
Structural Formula	C ₆ H ₁₂ O

The interface also shows a chemical structure diagram of hexanal and a list of categories on the left side of the window.

Chemical Input

Input target chemical#2 by SMILES

- To add additional chemicals by hand into the matrix, **right-click** above the structure and **select** "Add target" and then "Structure".
- Enter** the SMILES for 1-heptanal: CCCCCCC=O and **click** "OK" (see next screen shot).

Chemical Input

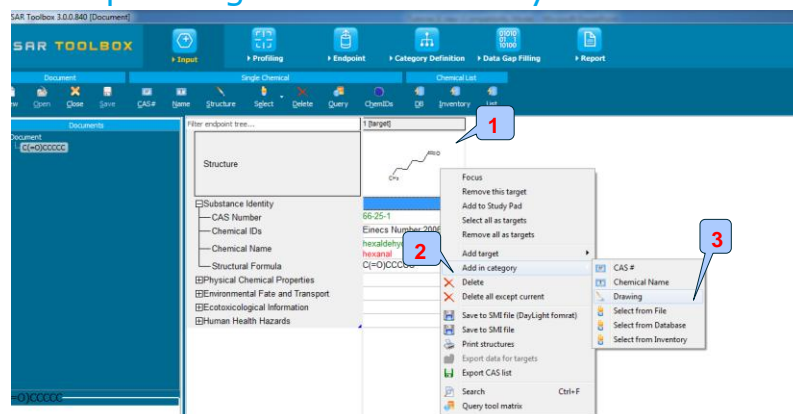
Input target chemical#2 by SMILES

The screenshot shows the QSAR Toolbox software interface. The main window displays a chemical structure of 1-heptanal. A context menu is open over the structure, with three numbered callouts: 1 points to the right-click action, 2 points to the 'Add target' option, and 3 points to the 'Structure' sub-option. The 'Structure' sub-option is highlighted in blue.

1. Right-click in the space above the structure; **2. Select Add target**; **3. Select Structure.**

Chemical Input

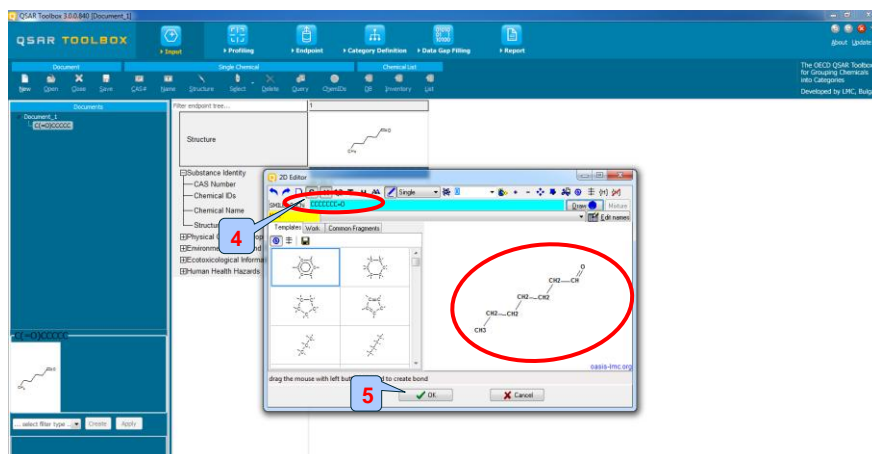
Input target chemical#2 by SMILES



1. Right-click in the space above the structure; 2. Select Add in category; 3. Select Drawing.

Chemical Input

Input target chemical#2 by SMILES



4. Type CCCCC=O in SMILES/InChi window; 5. Click OK.

Chemical Input

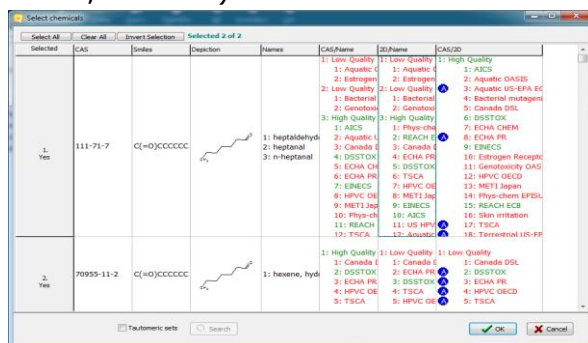
Input target chemical#2 by SMILES

- The Toolbox now consults its chemical ID database and finds all chemicals with the structure CCCCCCC=O.
- The Toolbox finds three chemicals with the same structure for 1-heptanal but with different CAS numbers and chemical names. Therefore, the Toolbox find two chemicals with different QA relations (CAS-Name; 2D -Name; CAS-2D (see next screen shot).

Chemical Input

Input target chemical#2 by SMILES

- The Toolbox finds two chemicals with the same structure and with different QA relations (CAS-Name; 2D -Name; CAS-2D).



Chemical Input QA

Note: The last 3 columns represent the chemical identification relations: CAS/Name, 2D/Name, and CAS/2D.

Chemical Input QA

- The columns represent chemical relations
- The colors represent the quality of relation

Text Color	Evaluated Q
Black	N/A Quality
Red	Low Q
Orange	Moderate Q
Green	High Q
Blue	Conflict

Chemical Input QA

CAS	Smiles	Depiction	Names	CAS/Name	2D/Name	CAS/2D	Quality
111-71-7	<chem>C(=O)CCCCC</chem>		1: heptaldehyde 2: heptanal 3: n-heptanal	1: Low Quality 2: Estrogen Re 3: Canada DSL 4: DSSTOX 5: ECHA CHEM 6: ECHA PR 7: ENRCS 8: HPVC OECD 9: METI Japan 10: Phys-chem 11: REACH ECF	1: Aquatic OAS 2: Estrogen Re 1: Bacterial mu 2: Genotoxicity 3: High Quality 1: Phys-chem 2: REACH ECF 3: Canada DSL 4: DSSTOX 5: ECHA CHEM 6: TSCA 7: HPVC OECD 8: METI Japan 9: ENRCS 10: AICS 11: US HPV CH	1: High Quality 1: Canada DSL 2: DSSTOX 3: ECHA PR 4: HPVC OECD 5: TSCA	1: Low Quality 2: Estrogen Re 1: Bacterial mu 2: Genotoxicity 3: High Quality 1: Phys-chem 2: REACH ECF 3: Canada DSL 4: DSSTOX 5: ECHA CHEM 6: TSCA 7: HPVC OECD 8: METI Japan 9: ENRCS 10: AICS 11: US HPV CH
70955-11-2	<chem>C(=O)CCCCC</chem>		1: hexene, hydr...	1: High Quality 1: Canada DSL 2: DSSTOX 3: ECHA PR 4: HPVC OECD 5: TSCA	1: Low Quality 1: Canada DSL 2: ECHA PR 3: DSSTOX 4: TSCA 5: HPVC OECD	1: Low Quality 1: Canada DSL 2: DSSTOX 3: ECHA PR 4: HPVC OECD 5: TSCA	1: Low Quality 1: Canada DSL 2: DSSTOX 3: ECHA PR 4: HPVC OECD 5: TSCA

Text Color	Evaluated Q
Black	N/A Quality
Red	Low Q
Orange	Moderate Q
Green	High Q
Blue	Conflict

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 31

Chemical Input QA

Selected	CAS	Smiles	Depiction	Names	CAS/Name	2D/Name	CAS/2D	Quality
1 Yes	111-71-7	<chem>C(=O)CCCCC</chem>		1: heptaldehy 2: heptanal 3: n-heptanal	1: Low Quality 2: Estrogen Recept 3: genotoxicity OAS 4: DSSTOX 5: ECHA CHEM 6: ECHA PR 7: ENRCS 8: HPVC OECD 9: METI Japan 10: Phys-chem EPI 11: REACH ECF	1: Aquatic OAS 2: Estrogen Recept 1: Bacterial mutag 2: genotoxicity OAS 3: High Quality 1: Aquatic US-EPA 2: Canada DSL 4: DSSTOX 5: ECHA CHEM 6: ECHA PR 7: ENRCS 8: HPVC OECD 9: METI Japan 10: Phys-chem EPI 11: REACH ECF	1: AICS 2: REACH ECF 3: Canada DSL 4: DSSTOX 5: ECHA CHEM 6: ECHA PR 7: ENRCS 8: Estrogen Re 9: ENRCS 10: Estrogen Re 11: Genotoxicity 12: HPVC OECD 13: METI Japan 14: Phys-chem 15: REACH ECF 16: Skin irritatic 17: TSCA 18: Terrestrial I	
3 Yes	70955-11-2	<chem>C(=O)CCCCC</chem>		1: hexene, hyc	1: High Quality 1: Canada DSL 2: DSSTOX 3: ECHA PR 4: HPVC OECD 5: TSCA	1: Low Quality 1: Canada DSL 2: ECHA PR 3: DSSTOX 4: TSCA 5: HPVC OECD	1: Low Quality 1: Canada DSL 2: DSSTOX 3: ECHA PR 4: HPVC OECD 5: TSCA	1: Low Quality 1: Canada DSL 2: DSSTOX 3: ECHA PR 4: HPVC OECD 5: TSCA

Text Color	Evaluated Q
Black	N/A Quality
Red	Low Q
Orange	Moderate Q
Green	High Q
Blue	Conflict

1. Double click on the column to see sources (2) of the representing names for the current QA relation (CAS/Name in this case).

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 32

Chemical Input

Input target chemical#2 by SMILES

Back to our target chemical, the first one is the actual 1-heptanal while the second one is a mixture containing 1-heptanal. As we are not interested in the mixture this chemical can be removed from the exercise. (see next screenshot)

The screenshot shows the 'Select chemicals' dialog box with two entries. The first entry is '1: Yes' with CAS 111-71-7 and SMILES C(=O)CCCC. The second entry is '2: Yes' with CAS 70955-11-2 and SMILES C(=O)CCCC, which is circled in red. The dialog box also shows a table of chemical names and their associated CAS/2D names.

Selected	CAS	Smiles	Depiction	Names	CAS/Name	2D/Name	CAS/2D
1: Yes	111-71-7	<chem>C(=O)CCCC</chem>		1: heptaldeh 2: heptanal 3: n-heptanal	3: High Qual 1: AICS 2: Aquat 3: Canad 4: DSST 5: ECHA 6: ECHA	3: High Qual 1: Phys- 2: REAC 3: Canad 4: ECHA 5: DSST 6: TSCA	6: DS 7: EC 8: EC 9: EII 10: E 11: G 12: H
2: Yes	70955-11-2	<chem>C(=O)CCCC</chem>		1: hexene, h	1: High Qual 1: Canad 2: DSST 3: ECHA 4: HPMC 5: TSCA	1: Low Qual 1: Canad 2: ECHA 3: DSST 4: TSCA 5: HPMC	1: Low Qual 1: Ca 2: DS 3: EC 4: HF 5: TS

Chemical Input

Input target chemical#2 by SMILES

The screenshot shows the 'Select chemicals' dialog box with two entries. The first entry is '1: Yes' with CAS 111-71-7 and SMILES C(=O)CCCC. The second entry is '2: No' with CAS 70955-11-2 and SMILES C(=O)CCCC. A blue box labeled '1' points to the 'Yes' column, and another blue box labeled '2' points to the 'OK' button.

Selected	CAS	Smiles	Depiction	Names	CAS/Name	2D/Name	CAS/2D
1: Yes	111-71-7	<chem>C(=O)CCCC</chem>		1: heptaldeh 2: heptanal 3: n-heptanal	3: High Qual 1: AICS 2: Aquat 3: Canad 4: DSST 5: ECHA 6: ECHA	3: High Qual 1: Phys- 2: REAC 3: Canad 4: ECHA 5: DSST 6: TSCA	6: DS 7: EC 8: EC 9: EII 10: E 11: G 12: H
2: No	70955-11-2	<chem>C(=O)CCCC</chem>		1: hexene, h	1: High Qual 1: Canad 2: DSST 3: ECHA 4: HPMC 5: TSCA	1: Low Qual 1: Canad 2: ECHA 3: DSST 4: TSCA 5: HPMC	1: Low Qual 1: Ca 2: DS 3: EC 4: HF 5: TS

1. **Click** over the first column with label Yes, then the column become unmarked (labeled with No); 2. **Click** OK

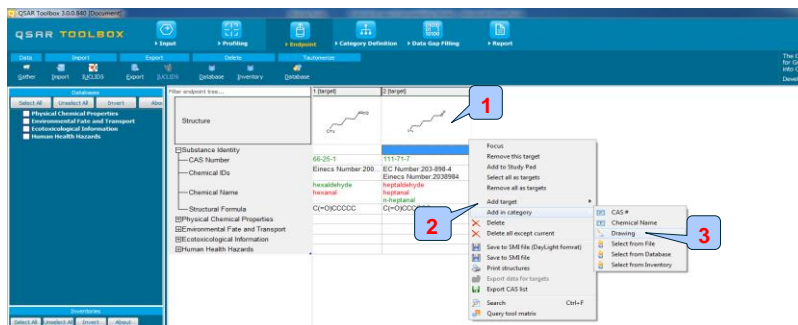
Chemical Input

Input target chemical#2 by SMILES

- To add the third chemical by hand into the matrix, **right-click** above the structure and **select** "Add target" and then "Structure".
- Enter** the SMILES for 3-ethyl-1-pentanal : CCC(CC)CC=O and **click** "OK".
- Your data matrix should now contain your three chemicals (see next screen shot).

Chemical Input

Input target chemical#3 by SMILES



- Right-click** in the space above the structure; **2. Select** Add in category; **3. Select** Drawing.

Chemical Input

Input target chemical#3 by SMILES

4. Type CCCCC=O in SMILES/InChi window; 5. Click OK.

Chemical Input

Input target chemical#3 by SMILES

Outlook

- Background
- Objectives
- Specific Aims
- The exercise
- **Workflow of the exercise 1**
 - Chemical Input
 - Profiling
 - **Endpoints**

Endpoints

- Move directly to the module "Endpoints".
- Remember, "Endpoints" refer to the electronic process of retrieving fate and toxicity data stored in the Toolbox and it can be gathering in a global fashion or on a more defined basis.
- In this example we only want to retrieve data on toxicity to fish so **select** the following databases containing information on aquatic toxicity:
 - Aquatic ECETOC
 - Aquatic Japan MoE
 - Aquatic OASIS
 - Aquatic US-EPA ECOTOX
- **Click** "Gather Data" (see next screen shot).

Endpoints Gather data

The screenshot shows the QSAR Toolbox 3.0.0.80 interface. The 'Databases' list on the left is expanded to show 'Human Health Hazards' (1). The 'Gather' button is circled in red (2). A 'Read data?' dialog box is open with 'All endpoints' selected and 'OK' clicked (3). The main window shows a table of chemical data for three targets.

Structure	1 Target	2 Target	3 Target
Structure	<chem>CCCCC</chem>	<chem>CCCCCC</chem>	<chem>CCCCC</chem>
Substance Identity			
-CAS Number	66-25-1	111-71-7	N/A
Chemical IDs	Einecs Number 200-111-71-7	EC Number 203-898-4	N/A
Chemical Name	hexanaldehyde	heptaldehyde	hexanal
Structural Formula	<chem>C=OCCCC</chem>	<chem>C=OCCCCC</chem>	<chem>C=OCCCC</chem>

1. Select databases related to the target endpoint; 2. Click Gather; 3. Click OK.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 41

Endpoints Available experimental data

- Results are available for two effects:
 - Growth** for *Tetrahymena pyriformis* for both n-hexanal and n-heptanal.
 - Mortality** for two species: *Pimephales promelas* and *Poecilia reticulata* for both n-hexanal and n-heptanal (see next screen shot).
- These can potentially be used for read-across to fill in the data gap for the third target: **3-ethyl-1-pentanal** (e.g. using the lowest available LC50 result).

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 42

Endpoints
Available experimental data

The screenshot shows the QSAR Toolbox interface with the following data in the 'Endpoints' table:

Endpoint	1 Target	2 Target	3 Target
Salatylena pyrifoma (22)	M: 152 mg/L		M: 114 mg/L
Acropora (Vertebrates)			M: 51.187 mg/kg

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 43

Outlook

- Background
- Objectives
- Specific Aims
- The exercise
- **Workflow of the exercise 1**
 - Chemical Input
 - **Profiling**
 - Endpoints

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 44

Profiling

- Click on "Profiling" to move back (**yes back**) to the previous module.
- Remember that "Profiling" refers to the process of retrieving information on the target compounds, other than and toxicity data.
- Available information includes likely mechanism(s) of action.
- In this exercise we will use the profiling results to evaluate the robustness of the analogue approach.

Profiling

Profiling the target chemical

- As you remember, the outcome of the profiling determines the most appropriate way to search for analogues.
- For this example the following mechanistic and endpoint specific profiling methods should be selected:
 - Protein binding by OASIS v1.1 – mechanistic grouping
 - Acute aquatic toxicity MOA by OASIS
 - Aquatic toxicity classification by ECOSAR
 - Acute aquatic toxicity classification by Verhaar
- Select those 4 "profiling methods" by **clicking** on the boxes before the names of the profilers before **clicking** "Apply" (see next screen shot).

Profiling

Profiling the target chemical

The screenshot shows the QSAR Toolbox interface during the Profiling step. The 'Profiling methods' list on the left includes various endpoints such as 'Estrogen Receptor Binding', 'Acute Toxicity (LC50, LD50, etc.)', and 'Aquatic Toxicity'. The 'Apply' button is highlighted with a red circle and the number '1'. The top toolbar also has an 'Apply' button highlighted with a red circle and the number '2'. The right panel displays a tree view of chemical categories and a table of results for 'Tetrahymena pyriformis'.

Filter endpoint tree...	1 Target	2 Target	3 Target
Structure	<chem>CC1=CC=CC=C1</chem>	<chem>CC1=CC=CC=C1</chem>	
Substance Identity			
Physical Chemical Properties			
Environmental Fate and Transport			
Ecotoxicological Information			
Aquatic Toxicity			
Growth			
HBC50			
H48 h			
Ligandotoxicity			
Cilicilia			
Ciliata			
Tetrahymena pyriformis	(2) M: 152 mg/L	M: 114 mg/L	
Immobility			
HBC50			
HCS0	(2) M: 17.8 mg/L, 9.79	M: 12 mg/L, 3.86 m	
Undefined Endpoint	(1/1)	M: 51.187 mg/kg	
Physiology			
Undefined Effect	(1/1)	M: 16.5 mg/L	
Sediment Toxicity			
Terrestrial Toxicity			
Human Health Hazards			

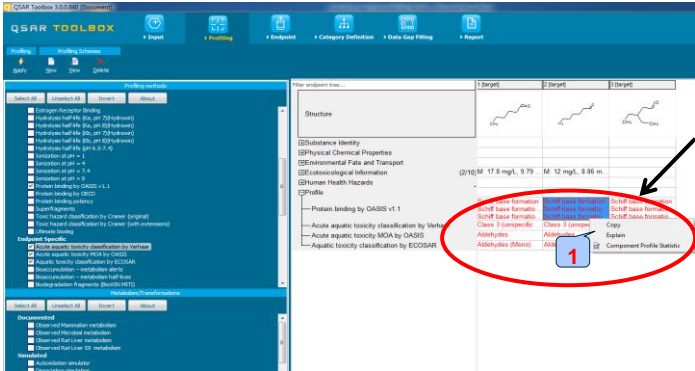
1. Select the profilers related to the target endpoint; 2. Click Apply.

Profiling

Profiling the target chemical

- The actual profiling will take several seconds depending on the number and type of selected profilers.
- The results of profiling automatically appeared as a dropdown box under the target chemical.
- The target and source chemicals have the same mechanisms or modes of action relevant for acute aquatic toxicity.
- The Toolbox does not provide any arguments against read-across (see next screen shot).

Profiling Profiles of the targets

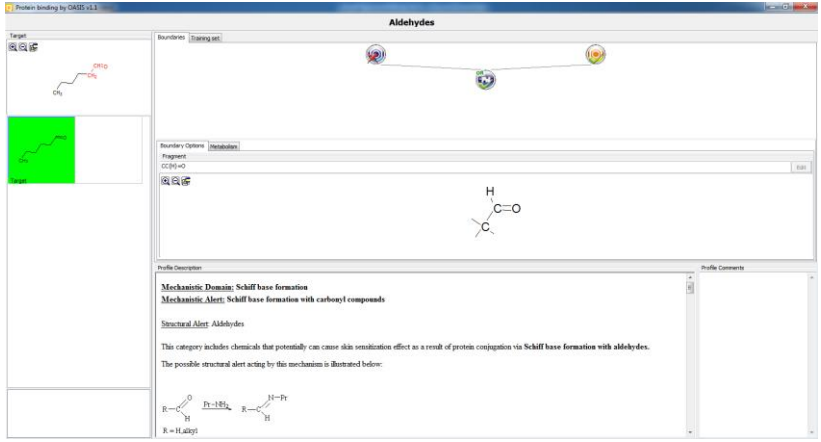


- In this case the target and source chemicals have the same mechanisms and modes of action.
- So the Toolbox does not provide any arguments against read-across.
- This step is critical for next grouping of analogues.

1. Right click to see why this target is Protein binding alert v1.1(see next screen shot).

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 49

Profiling Protein binding by OASIS v1.1 of "n-hexanal"



Profile Description

Mechanistic Domain: Schiff base formation

Mechanistic Alert: Schiff base formation with carbonyl compounds

Structural Alert: Aldehydes

This category includes chemicals that potentially can cause skin sensitization effect as a result of protein conjugation via Schiff base formation with aldehydes. The possible structural alert acting by this mechanism is illustrated below:

$$R-\overset{\text{O}}{\parallel}{C}-H + H_2N-Pr \rightarrow R-\overset{\text{O}}{\parallel}{C}-N(Pr)-H$$

R = H, alkyl

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 50

Profiling Recap

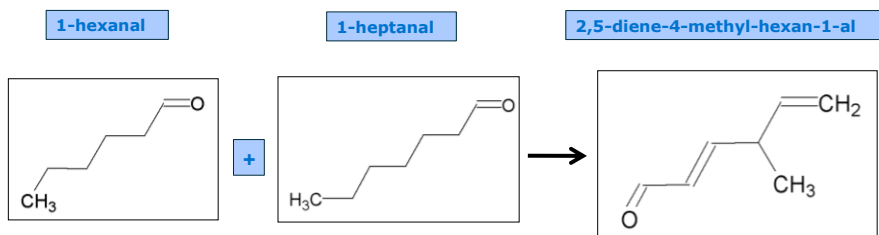
- You have entered the source and target chemicals being sure of the correct structures.
- You have checked the relevant databases for available experimental results.
- You have profiled the source and target chemicals.
- You have evaluated the robustness of the analogue approach and concluded that the read-across may be acceptable.

Outlook

- Background
- Objectives
- Specific Aims
- The exercise
- Workflow of the exercise 1
- **Workflow of the exercise 2**
 - **Chemical Input**
 - Profiling
 - Endpoints

Chemical Input Exercise 2

Read-across of acute toxicity to fish from 1-hexanal and 1-heptanal to 2,5-diene-4-methyl-hexan-1-al.



Chemical Input Exercise 2

- In the second example, we use the same source chemicals and a different target chemical.
- We can therefore simply delete the previous target chemical and enter the identity of the new target chemical.
- **Right-click** above the structure of chemical 3-ethyl-1-pentanal and **select** "Delete chemical" (see next screen shot).

Chemical Input Exercise 2

The screenshot shows the QSAR Toolbox interface with a table of chemical targets. The table has columns for 'Structure', 'Target', and 'Target'. The first target is highlighted. A right-click context menu is open over the first target, with the 'Delete chemical' option selected. Red callouts 1 and 2 indicate the right-click and the selection of 'Delete chemical' respectively.

1. Right click on the previous target; 2. Select Delete chemical.

Chemical Input Exercise 2

The screenshot shows the QSAR Toolbox interface with a table of chemical targets. The table has columns for 'Structure', 'Target', and 'Target'. The first target is highlighted. A right-click context menu is open over the first target, with the 'Delete chemical' option selected. Red callouts 1 and 2 indicate the right-click and the selection of 'Delete chemical' respectively.

1. Right click on the previous target; 2. Select Delete chemical.

Chemical Input Exercise 2

- Add the new target chemical as in the previous exercise.
- To add the third chemical by hand into the matrix, **right-click** above the structure and **select** "Add target" and then "Structure".
- **Enter** the SMILES for 2,5-diene-4-methyl-hexan-1-ol: O=CC=CC(C)C=C and **click** "OK".

Chemical Input Exercise 2

The screenshot shows the QSAR Toolbox software interface. The main window displays a list of chemical structures in a table with columns for 'Structure', 'CAS#', 'Mol. Weight', and 'Mol. Weight'. A context menu is open over the second structure, showing options like 'Focus', 'Remove this target', 'Add to Study List', 'Select all in target', 'Remove all in target', 'Add target', 'Add in category', 'Delete', 'Delete all except current', 'Save to Grid file (Ctrl+I) (right mouse)', 'Save to Grid file', 'Print structure', 'Export data for targets', 'Export CAS list', and 'Query tool matrix'. Three numbered callouts are present: 1. A red circle with the number '1' points to the right-click action above the structure. 2. A red circle with the number '2' points to the 'Add category' option in the menu. 3. A red circle with the number '3' points to the 'Drawing' option in the sub-menu.

1. **Right-click** above the structure; 2. **Select** Add category; 3. **Select** Drawing

Chemical Input Exercise 2

4. Type O=C=C=C=C in SMILES/InChi window; 5. Click OK.

Outlook

- Background
- Objectives
- Specific Aims
- The exercise
- Workflow of the exercise 1
- **Workflow of the exercise 2**
 - Chemical Input
 - **Profiling**
 - Endpoints

QSAR TOOLBOX

Profiling

- In the module profiling, profile the new target chemical with the 4 profilers relevant for aquatic toxicity, in the same way as for the previous example.
- See these results in the next screen shot.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 61

QSAR TOOLBOX

Profiling

1. Check the profilers related to the target endpoint; 2. Click Apply.

In this case 2,5-diene-4-methylhexan-1-al has different mechanism and mode of action than the other two targets. So the analogue approach is not justified.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 62

Profiling

Conclusions from profiling

- The profiling results indicate differences in protein binding and ECOSAR classification between target and source chemicals.
- The analogue approach may therefore not be justified.
- It is recommended to search for other source chemicals which have the same profile.

Recap

- You have replaced a target chemical with another target chemical in the data matrix.
- You have profiled the new target chemical.
- You have evaluated the robustness of the analogue approach and concluded that the read-across may not be acceptable.

Searching for More Suitable Analogues

- Before searching for more suitable analogues, delete n-hexanal and n-heptanal from the data matrix by **right-clicking** above each of them and **select** "Delete chemical" or **right-clicking** above the target (2,5-diene-4-methyl-hexan-1-al) and **select** "Delete all except current" (see next screen shot).
- The aim of the next part of the exercise will be to find analogues which have the same profiling results as the target chemical.

Searching for More Suitable Analogues

The screenshot shows the QSAR Toolbox interface with a table of chemical profiles. The table has columns for 'Structure', 'Target', and 'Target'. A context menu is open over the 'Target' column, with two callouts: '1' pointing to the right-click action and '2' pointing to the 'Delete All except current' option.

Structure	Target	Target
	<chem>CCCCC=CC=O</chem>	<chem>CCCCC=CC=O</chem>
	<chem>CCCCC=CC=O</chem>	<chem>CCCCC=CC=O</chem>
	<chem>CCCCC=CC=O</chem>	<chem>CCCCC=CC=O</chem>

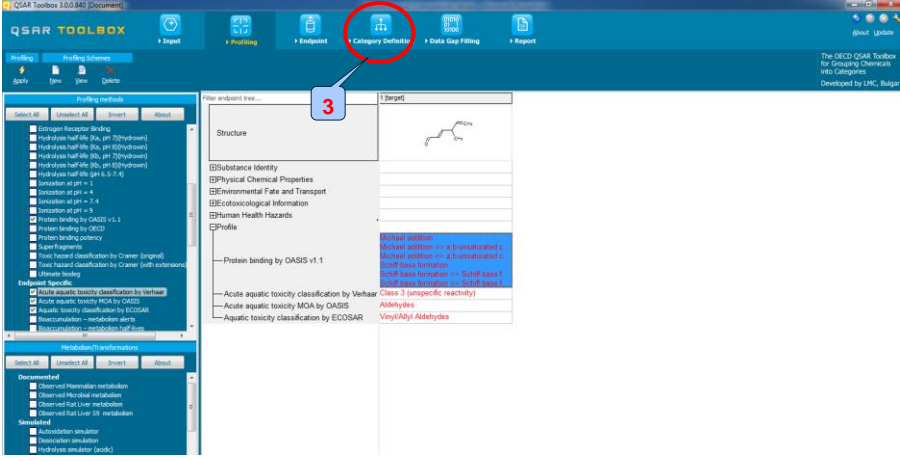
Context menu options:

- Focus
- Remove this target
- Add to Study Pad
- Select all as targets
- Remove all as targets
- Add target
- Add in category
- Delete
- Delete all except current
- Save to SKI file (De/Light format)
- Save to SKI file
- Print structures
- Export data for targets
- Export CAS list
- Search
- Query tool matrix

1. **Right click** in the space above the target chemical; 2. **Select** Delete All except current.

QSAR TOOLBOX

Searching for More Suitable Analogues



3. Move to the module Category Definition to launch a search for more suitable analogues.

The OECD QSAR Toolbox for Grouping Chemicals into Categories
5.10.2012 67

QSAR TOOLBOX

Outlook

- Background
- Objectives
- Specific Aims
- The exercise
- Workflow of the exercise 1
- **Workflow of the exercise 2**
 - Chemical Input
 - Profiling
 - Endpoints
 - **Category definition**

The OECD QSAR Toolbox for Grouping Chemicals into Categories
5.10.2012 68

Category definition

Searching for More Suitable Analogues

- Currently it is not possible to query directly by several profiling results in parallel. The user has first to query according to one profiler and then subcategorise the results step-by-step according to other profilers.
- For this example, the user could first select the protein binding mechanism of the target chemical and query for all the chemicals with the same mechanism in the selected inventory and databases (see next screen shot).

Category definition

Searching for More Suitable Analogues

The screenshot shows the QSAR Toolbox 2.3.0.13190 interface. On the left, the 'Consistent methods' list includes 'Protein binding by OASIS', which is highlighted with a red circle and the number 1. In the center, the 'Define' button is circled in red with the number 2. On the right, a dialog box titled 'Protein binding by OASIS' is open, showing a list of categories and the 'OK' button, which is circled in blue with the number 3.

1. **Highlight** "Protein binding by OASIS"; 2. **Click** Define and confirm the category from Protein binding by OASIS profiler; 3. **Click** OK.

Category definition
Searching for More Suitable Analogues

1. Highlight "Protein binding by OASIS v1.1"; 2. Click Define and confirm the category from Protein binding by OASIS v1.1 profiler; 3. Click OK.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 71

Category definition
Searching for More Suitable Analogues

1. Click OK to confirm the name of the category.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 72

Category definition

Searching for More Suitable Analogues

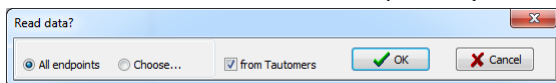
- The Toolbox now identifies all chemicals corresponding to mechanism "Michael addition<AND>Schiff base formation" by Protein binding by OASIS v1.1 listed in the databases selected under "Endpoints".
- 36 analogues are identified. Along with the target they form a mechanistic category used for gap filling.
- The name of the category appear in the "Defined Categories" window, indicating the number of substances belonging to the category.

[36] Michael addition<AND>Michael addition >> a,b-unsaturated carbonyl compounds<AND>Michael addition >> a,b-unsaturated carbonyl compounds >> a,b-unsaturated aldehydes<AND>Schiff base formation<AND>Schiff base formation >> Schiff base formation with carbonyl

Category definition

Reading data for Analogues

- The Toolbox will now retrieve those chemicals that have the same protein binding mechanism than the target compound.
- The Toolbox automatically request the user to select the endpoint that should be retrieved.
- The user can either select the specific endpoint or by default choose to retrieve data on all endpoints (see below).



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

Category definition

Reading data for Analogues

Due to the overlap between the Toolbox databases same data for intersecting chemicals is found simultaneously in more than one database. The data redundancy is identified and the user has the opportunity to select either a single data value or all data values.

Repeated values for: 38 data-points, 10 groups, 4 chemicals

	Endpoint	CAS	Structure	Value	Age
<input checked="" type="checkbox"/>	LC50	107-02-8	<chem>C=CC=O</chem>	27(24;30) micrograms per liter	1 day(s)
<input checked="" type="checkbox"/>	LC50	107-02-8	<chem>C=CC=O</chem>	27(24;30) micrograms per liter	30 day(s)
<input checked="" type="checkbox"/>	IGC50	N/A	<chem>C=CC=O</chem>	0.000832 mol/L	
<input checked="" type="checkbox"/>	IGC50	N/A	<chem>C=CC=O</chem>	0.000832 mol/L	
<input checked="" type="checkbox"/>		107-02-8	<chem>C=CC=O</chem>	1.5E4;2E4 micrograms per liter	
<input checked="" type="checkbox"/>		107-02-8	<chem>C=CC=O</chem>	1.5E4;2E4 micrograms per liter	
<input checked="" type="checkbox"/>		107-02-8	<chem>C=CC=O</chem>	1.4;42 micrograms per liter	

1. Click Select one and then 2. Click OK.

Category Definition

Read data for Analogues

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

QSAR Toolbox 3.0.0.840

335 data points gathered across 30 chemicals.

1. Click OK

QSAR TOOLBOX

Category definition

Defined category

The OECD QSAR Toolbox for Grouping Chemicals into Categories
Developed by LMC, Belgium

5.10.2012

77

QSAR TOOLBOX

Category definition

Subcategorisation

- After the available data has been retrieved, the user can then further subcategorize the results according to the following subcategorisations:
 - EcoSAR Classification
 - MOA of action
 - Verhaar classification
- These steps are summarized in the next screen shots.

The OECD QSAR Toolbox for Grouping Chemicals into Categories

5.10.2012

78

Category definition Subcategorisation by ECOSAR

1. Select current category; **2. Click** Subcategorise; **3. Select** ECOSAR profiler. Note all analogues are in the same category as the target chemical so no further action is required.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 79

Category definition Subcategorisation by Acute aquatic toxicity by OASIS

1. Select current category; **2. Click** Subcategorise; **3. Select** MOA by OASIS profiler; **4. Remove** dissimilar chemicals and **5. Confirm** new category by clicking OK.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 80

Category definition

Subcategorisation by Acute aquatic classification by Verhaar

1. Select current category; 2. Click Subcategorise; 3. Select Verhaar profiler; 4. Remove dissimilar chemicals and 5. Confirm new category by clicking OK.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 81

Category definition

Results after subcategorisation

Structure	Target	n	M
Chordata/Vertebrates		(1/1)	M: 5.94 mg/L, 5.95... M: 3.4 mg/L
Amphibia/Fish		(3/4)	
Mollusca/Vertebrates		(2/2)	M: 0.91 mg/L

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 82

Category definition

Interpretation of the results

- Following the above-described subcategorisation exercise, 28 chemicals are left in the category. All have same mechanisms of action.
- The result is a group of chemicals that can bind to proteins by via Michael addition and Schiff base mechanisms, belong to the Vinyl/Allyl class by ECOSAR category and have same mode of action according to the MOA profiler.
- For 6 chemicals, experimental results for acute toxicity to fish are available- 3 chemicals have 96h-LC50 results from 3.4 to 7.9 mg/l for *Pimephales promelas*; 2 chemicals have 96h-LC50 results from 7.62 to 9.81 mg/l for *Poecilia reticulata*; 1 chemical has 96h-LC50 0.91mg/l for *Oryzias latipes*.
- (see next two screen shots)

Category Definition

Interpretation of the results

96h-LC50 for *Pimephales promelas*

The screenshot shows the QSAR Toolbox interface with a hierarchical tree of chemical categories. The tree is sorted by priority, and the 'Sort' menu is open, showing 'Descending' selected. Three numbered callouts (1, 2, 3) indicate the steps: 1. Right click above the current fish; 2. Select Sort (targets priority), then 3. Descending.

1. Right click above the current fish; 2. Select Sort (targets priority), then 3. Descending.

Category Definition Interpretation of the results

96h-LC50 from 3.4 to 7.29 mg/l for *Pimephales promelas*.

The screenshot shows the QSAR Toolbox interface. On the left, there is a 'Pimephales' tree view with various endpoints selected. The main window displays a list of chemicals with their structures and associated data. The entry for *Pimephales promelas* is circled in red, and its 96h-LC50 value of 7.29 mg/l is highlighted. Other chemicals listed include *Jordanella floridae*, *Leuciscus idus*, *Morone chrysops*, *Oncorhynchus kisutch*, *Oncorhynchus mykiss*, *Oryzias latipes*, and *Poecilia reticulata*.

Category Definition Interpretation of the results

96h-LC50 results for *Poecilia reticulata*

The screenshot shows the QSAR Toolbox interface. A context menu is open over the entry for *Poecilia reticulata*. Three numbered callouts indicate the steps: 1. Right click above the current fish; 2. Select Sort (targets priority); 3. Descending. The menu options include 'Sort targets priority', 'Sort', 'Function...', 'Set tree hierarchy...', 'Export CAS list', 'Export', 'Wiki search species', and 'Copy path'.

1. Right click above the current fish; 2. Select Sort (targets priority), then 3. Descending.

Category Definition

Interpretation of the results

96h-LC50 results from 7.6 to 9.8 mg/l for *Poecilia reticulata*.

Structure	1 Target	2
Jordaniella floridae		
Leuciscus idus		
Mendilia beryllina		
Micropterus salmoides		
Oncorhynchus kisutch		
Oncorhynchus mykiss		
Oryzias latipes (1/1)		
Pimephales promelas (3/4)		
Poecilia reticulata (2/2)	M: 9.81 mg/L	M: 7.62 mg/L
Amphibia(Frog)		
Mollusca(Invertebrates)		
EB96 Days		
EB27 Days		
EB32 Days		
EBLOEC		
EBLT50		

The OECD QSAR Toolbox for Grouping Chemicals into Categories

5.10.2012

87

Category Definition

Interpretation of the results

96h-LC50 results for *Oryzias latipes*

Structure	16	17	18	19
Jordaniella floridae				
Leuciscus idus				
Mendilia beryllina				
Micropterus salmoides				
Oncorhynchus kisutch				
Oncorhynchus mykiss				
Oryzias latipes (1/1)				
Pimephales promelas (3/4)				
Poecilia reticulata (2/2)				
Amphibia(Frog)				
Mollusca(Invertebrates)				
EB96 Days				
EB27 Days				
EB32 Days				
EBLOEC				
EBLT50				
EBMARC				
EBNOEC				
EBNR-LETH				
EBNR-ZERO				
EBUndefined Endpoint				

1. Right click above the current fish; 2. Select Sort (targets priority), then 3. Descending.

The OECD QSAR Toolbox for Grouping Chemicals into Categories

5.10.2012

88

QSAR TOOLBOX

Category Definition

Interpretation of the results

96h-LC50 result for *Oryzias latipes*

The screenshot shows the QSAR Toolbox interface with the 'Category Definition' window open. The window displays a list of chemical categories and their corresponding 96h-LC50 values for *Oryzias latipes*. The value for *Pimephales promelas* (34) is highlighted with a red circle, showing a result of 7.3 mg/l. Other visible results include *Poecilia reticulata* (2/2) at 7.6 mg/l and M. 81163 mg/kg.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 89

QSAR TOOLBOX

Category Definition

Interpretation of the results

- Further visual analysis of the structures (see next two screen shots) could indicate that the following results are most suitable for read-across:
 - *Pimephales promelas* : (E)-3,7-Dimethyl-2,6-octadienal - 96hLC50 = 7.3 mg/l
 - *Poecilia reticulata* : 2-Ethyl-2-butenal - 96hLC50 = 7.6 mg/l
- Indeed those chemicals are structurally most similar based on branching and functional groups in the molecule.

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 90

Category Definition Interpretation of the results

The screenshot shows the QSAR Toolbox interface. On the left, the 'Filter endpoint tree' is expanded to show endpoints like 'Hydrolysis half-life (K_a, pH 7)(Hydrow)', 'Ionization at pH = 1', and 'Acute aquatic toxicity classification by Verhaar'. The 'Structure' column displays a list of chemical names, with 'Jordanella floridae' at the top. A red circle highlights a chemical structure in the 'Structure' column, labeled '1'. The 'Table' on the right shows the following data:

Structure	1 target	2	3
Jordanella floridae			
Legomis macrochirus			
Leuciscus idus			
Meridia beryllina			
Micropterus salmoides			
Oncorhynchus kisutch			
Oncorhynchus mykiss			
Oryzias latipes (1/1)			
Pimephales promelas (3/4)	M. 7.29 mg/L	M. 5.94 mg/L, 5.9/5.38.6.	M. 3.4 mg/L
Poecilia reticulata (2/2)			
Amphibia(Frog)			
Mollusca(Invertebrates)			
EB6 Days			
EB7 Days			
EB32 Days			
EB.OEC			
EB.T50			
EB.MATC			
EB.OEC			
EB.NR.LETH			
EB.NR.TEFO			

1. Chemical 2 is most structurally similar to the **target** analogue.

Category Definition Interpretation of the results

The screenshot shows the QSAR Toolbox interface. On the left, the 'Filter endpoint tree' is expanded to show endpoints like 'Hydrolysis half-life (K_a, pH 7)(Hydrow)', 'Ionization at pH = 1', and 'Acute aquatic toxicity classification by Verhaar'. The 'Structure' column displays a list of chemical names, with 'Jordanella floridae' at the top. A red circle highlights a chemical structure in the 'Structure' column, labeled '1'. The 'Table' on the right shows the following data:

Structure	1 target	2	3
Jordanella floridae			
Legomis macrochirus			
Leuciscus idus			
Meridia beryllina			
Micropterus salmoides			
Oncorhynchus kisutch			
Oncorhynchus mykiss			
Oryzias latipes (1/1)			
Pimephales promelas (3/4)			
Poecilia reticulata (2/2)	M. 9.81 mg/L		M. 7.62 mg/L
Amphibia(Frog)			
Mollusca(Invertebrates)			
EB6 Days			
EB7 Days			
EB32 Days			
EB.OEC			
EB.T50			

1. Chemical 3 is most structurally similar to the **target** analogue.

Category Definition Recap

- You have searched for suitable analogues having the same profile than the target compound by successive subcategorisation with 4 profilers.
- You have chosen the most suitable candidates to be used for read-across based on a visual analysis of their molecular structure.

Report

- Remember the report module (not reviewed in this exercise) allows you to generate a report on the predictions performed with the Toolbox. This module contains predefined report templates as well as a template editor with which users can define their own user defined templates. The report can then be printed or saved in different formats.

Congratulation

- You have used some more functions of the Toolbox and changed up the workflow to address new issues.
- By now you should feel comfortable moving the curser around the basic screens for each one the modules.
- Continue to using the Toolbox and you speed and confidence will increase sharply.