

QSAR TOOLBOX

The OECD QSAR Toolbox
for Grouping Chemicals
into Categories

OECD QSAR Toolbox v.3.0

Step-by-step example of how to evaluate an
ad-hoc category of aliphatic amines and to
predict an ecotoxicological endpoint

Background

- This is a step-by-step presentation designed to take you through the workflow of the Toolbox for evaluating of an ad-hoc category.
- You will learn several new functionalities which we be repeated to assure a consistent category is defined.
- It is assumed that you now have some experience in using the Toolbox so there will be multiple key strokes between screen shots.

Outlook

- Background
- **Category evaluation - overview**

Category Evaluation

Two phases of evaluation process

Phase I. Evaluation of Category & Data Gap Filling:

- Investigating the structural consistency of an ad hoc category (e.g., a category submitted by an industry consortium to a regulatory assessment program).
- Implementation of an ad hoc category building and data gap filling.

Phase II. Extension of Category & Data Gap Filling:

- Search for other analogues which are consistent with the submitted category.
- Data gap filling using new data matrix.

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - ***Case study***

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Case Study

- The submission consists of a category with **19 aliphatic amines**.
- The predicted ecotoxicological endpoint EC50, 48h, *D.magna* of **2-Butanamine(CAS 13952-84-6)** will be reviewed.

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Case study*
 - **Workflow**

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category
Workflow

The following input workflow is used:

- Input the file for submission as an user list.(In this case input file Aliphatic amines.smi*)
- Evaluate the category applying the following profiling schemes:
 - USEPA categories
 - ECOSAR categorization
 - Aquatic toxicity MOA of action
 - Organic Functional Groups (nested)

Aliphatic amines.smi* - file is packed in the zip file with all power point tutorials

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Case study*
 - **Workflow**
 - **Input**

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/ Input

The screenshot displays the QSAR Toolbox 3.0.0.840 interface. The 'List' menu item is highlighted with a red '1'. A tooltip for 'Load chemicals from file' is visible, listing supported formats (SDF, MOL, MOL2, RDF, XYZ) and file extensions (.smi). An 'Open' dialog box is shown with the file 'aliphatic_amines.smi' selected, indicated by a red '2'. The 'Open' button is highlighted with a red '3'.

1. **Click** the List; 2. **Browse** and find the file for input 3. **Open** the file.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Input

- You have now inserted your chemical list into the system.
- **Click** on the box next to "Substance Identity"; this displays the chemical identification information (see next screen shot).

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/ Input

QSAR Toolbox 3.0.0.840 [aliphatic_amines.sm]

QSAR TOOLBOX

Document: [aliphatic_amines.sm]

Single Chemical: [aliphatic_amines.sm]

Chemical List: [aliphatic_amines.sm]

Filter endpoint tree...

	1 [target]	2 [target]	3 [target]	4 [target]	5 [target]	6 [target]	7 [target]	8 [target]
Structure								
Substance Identity								
CAS Number	74-89-5	75-04-7	75-31-0	75-64-9	78-96-6	102-82-9	104-75-6	108-91-8
Chemical IDs	EC Number 200-820-0 Eines Number 2006200	EC Number 200-834-7 Eines Number 200...	EC Number 200-860-9 Eines Number 200...	EC Number 200-888-1 Eines Number 200...	EC Number 201-162-7 Eines Number 201...	EC Number 203-058-7 Eines Number 203...	EC Number 203-233-8 Eines Number 203...	EC Number 203-233-8 Eines Number 203...
Chemical Name	amino-methane methanamine ethylamine monomethylamine	amino-ethane ethanamine ethylamine monoethylamine	2-amino-propane 2-propanamine 2-propylamine isopropylamine mono-isopropylamine propan-2-amine	(tert)butylamine 2-methylpropan-2-amine 2-propanamine, 2-amino- 1-butylamine tert-butylamine	1-amino-2-propanol 2-propanol, 1-amino- monoisopropanolamine	n,n-dibutylbutan-1-amine tri-n-butylamine tributylamine	1-hexanamine, 2-ethyl- 2-ethyl-1-hexanamine 2-ethylhexan-1-amine 2-ethylhexylamine hexylamine, 2-ethyl-	cyclohexylamine cyclohexylamine
Structural Formula	CN	C(C)N	C(C)C(N)	C(C)(C)C(N)	C(C)O(C)N	C(CCC)N(CCCC)C...	C(CCCC)C(C)C(N)	C1(N)CCCCC1
Physical Chemical Properties								
Environmental Fate and Transport								
Ecotoxicological Information								
Human Health Hazards								
Acute Toxicity								
Carcinogenicity								
Developmental Toxicity / Teratogenicity								

1. Click on "Substance identity"

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Case study*
 - **Workflow**
 - *Input*
 - **Profiling**

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

- The first step of the category evaluation according to Phase I is the consistency check of the category which starts with evaluation of the **robustness with respect to structural functionalities**. The following schemes could be used for this purpose:
 - US EPA categories
 - MOA of action
 - ECOSAR categorization
 - Organic Functional Groups (nested)
- **Select "profiling methods"** by clicking on the boxes before the names of the profilers and **Click "Apply"**. Before selecting the profiling methods unselect all (see next screen shot).

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

The screenshot displays the QSAR Toolbox 3.0.0.840 interface. The main window is titled 'QSAR Toolbox 3.0.0.840 [aliphatic_amines.om]'. The 'Profiling' workflow is active, showing a table of chemical structures and their identifiers. The table has 8 columns labeled '1 Target' through '8 Target'. The first column shows the chemical structure of methylamine (CN). The second column shows the chemical structure of ethylamine (CCN). The third column shows the chemical structure of isopropylamine (CC(C)N). The fourth column shows the chemical structure of isobutylamine (CC(C)CN). The fifth column shows the chemical structure of propylamine (CCCN). The sixth column shows the chemical structure of n-butylamine (CCCCN). The seventh column shows the chemical structure of n-hexylamine (CCCCCCN). The eighth column shows the chemical structure of n-octylamine (CCCCCCN). The table also includes columns for 'Number', 'EC Number', 'EINECS Number', and 'Environ. Formula'. The 'Environ. Formula' column contains the following values: CN, CCN, CC(C)N, CC(C)CN, C(C)CN, C(C)CCN, C(C)CCCCN, C(C)CCCCCN. The 'Profiling methods' panel on the left has a context menu open with 'Unselect All' selected. A blue box highlights the 'Unselect All' option in the context menu.

1. Right click in space above the profilers; 2. Select Unselect All.

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

The screenshot shows the QSAR Toolbox 3.0.0.840 interface. The 'Profiling' menu is open, and the 'Apply' button is highlighted with a red circle and the number '1'. The main window displays a table with columns for different target categories (1 Target to 7 Target) and rows for various chemical and physical properties. The table contains chemical structures and their corresponding identifiers and names.

	1 Target	2 Target	3 Target	4 Target	5 Target	6 Target	7 Target
Structure	<chem>CCN</chem>	<chem>CCNCC</chem>	<chem>CCN(C)C</chem>	<chem>CCN(C)C</chem>	<chem>CCN(C)C</chem>	<chem>CCN(C)C</chem>	<chem>CCN(C)C</chem>
Substance Identity							
CAS Number	4463-9	75-94-7	75-31-0	75-64-9	78-96-6	102-82-9	104-75-6
Chemical IDs	EC Number 200-820-0 Enics Number 2006200 amino-methane methanamine methylamine monomethylamine	EC Number 200-834-7 Enics Number 200 amino-ethane ethanamine ethylamine methylamine monomethylamine	EC Number 200-869-9 Enics Number 200 2-amino-propane amino-propane 2-propanamine propylamine isopropylamine isopropylamine mono-isopropylamine propen-2-amine C(C)C(N)	EC Number 200-888-1 Enics Number 200 tert-butylamine 1-amino-2-propanol 2-methylpropan-2-amine 2-propanamine, 2-amino- tert-butylamine monoisopropanolamine tert-butylamine	EC Number 201-162-7 Enics Number 201 1-butanamine, n-hexyl- 1-butanamine, n-hexyl- 2-ethyl-1-hexanamine, cyclohexyl- tri-n-butylamine 2-ethylthiothylamine tributylamine	EC Number 203-058-7 Enics Number 203 1-hexanamine, 2-ethyl- 1-hexanamine, 2-ethyl- 2-ethylhexan-1-amine, cyclohexyl- tri-n-butylamine 2-ethylthiothylamine tributylamine	EC Number 203-233-8 Enics Number 203 1-hexanamine, 2-ethyl- 1-hexanamine, 2-ethyl- 2-ethylhexan-1-amine, cyclohexyl- tri-n-butylamine 2-ethylthiothylamine tributylamine
Chemical Name							
Structural Formula	CN	C(C)N	C(C)C(N)	C(C)(C)N	C(C)(C)N	C(CCC)N(C)CC(C)C	C(C)CC(C)C(C)C(N)C(C)C
Physical Chemical Properties							
Environmental Fate and Transport							
Ecotoxicological Information							
Human Health Hazards							

1. Check the USEPA, MOA of action, ECOSAR and OFG(nested) profilers; 2. Click Apply.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

- 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 result from profile statistics, can be seen by **right clicking** in the space above the profiler in the endpoint tree and **select Profile statistics** from the dropdown menu (see next screen shot).

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

The screenshot shows the QSAR Toolbox interface with the following components:

- Top Navigation:** Input, Profiling, Endpoint, Category Definition, Data Gap Filling, Report.
- Left Panel (Profiling methods):**
 - Selected: US-EPA New Chemical Categories
 - Other methods include: General Mechanistic, Biodeg primary, Biodeg probability, DNA binding, Estrogen Receptor Binding, Hydrolysis half-life, Ionization, and Mutagenicity.
- Main Table:**

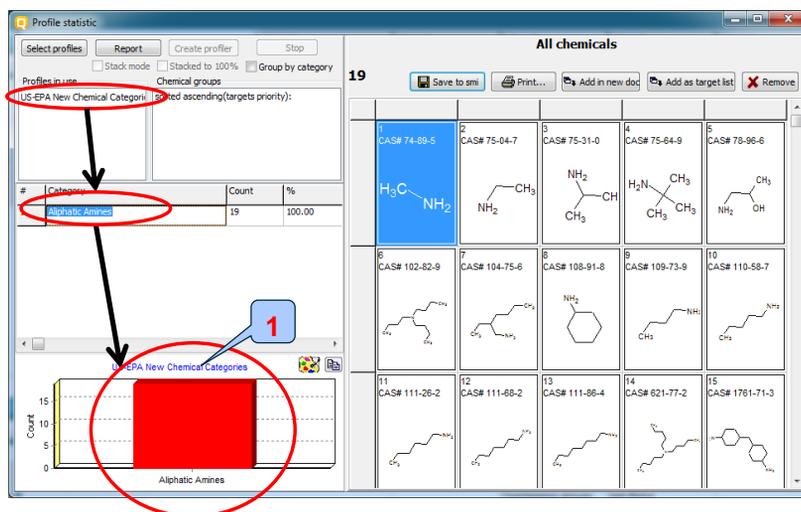
Filter endpoint tree...	1 (target)	2 (target)	3 (target)	4 (target)	5 (target)	6 (target)
Structure	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>
Substance Identity						
CAS Number	74-89-5	75-04-7	75-31-0	75-64-9	78-96-6	102-82-9
Chemical IDs	EC Number 200-820-0 Einecs Number 200/8200	EC Number 200-834-7 Einecs Number 200...	EC Number 200-869-9 Einecs Number 200...	EC Number 200-888-1 Einecs Number 200...	EC Number 201-162-7 Einecs Number 201...	EC Number 203-658-7 Einecs Number 203...
Chemical Name	amino-ethane ethanamine ethyl amine ethylamine monoethylamine	2-amino-propane 2-amino-propane ethyl amine 2-propanamine 2-propylamine isopropyl amine isopropylamine mono-isopropylamine propan-2-amine	(tert)butylamine 2-methylpropan-2-amine 2-propanamine, 2-... 1-butylamine tert-butylamine	1-amino-2-propanol 1-amino-propan-2-ol 2-propanol, 1-amino- monoisopropanola	1-butanamine, n,n...	1-dibutylaminobutane n,n-dibutylutan-1-... tri-n-butylamine tri-n-butylamine tributylamine tributylamine
Structural Formula	C(C)N	C(C)C(N)	C(C)(C)C(N)	C(C)(O)CN	C(CCC)N(CCCC)C...	
Profile	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines

1. Right click above US-EPA profiler; 2. Select Profile Statistics.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to US-EPA



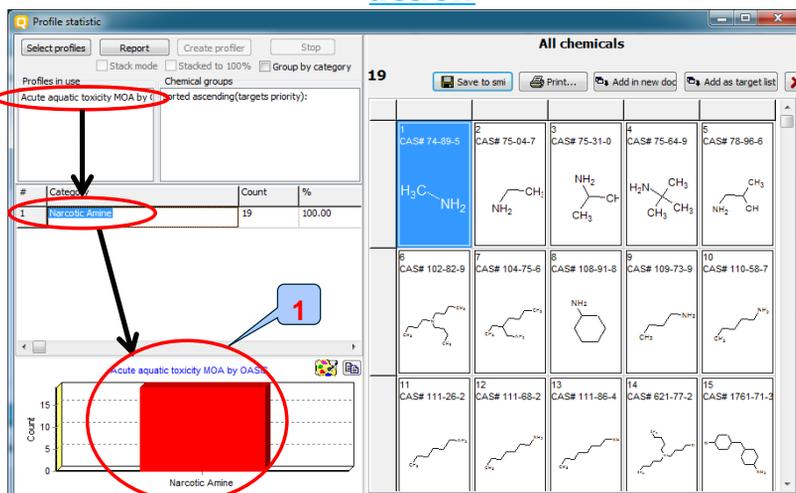
1. In this case all 19 chemicals are Aliphatic amines according to US-EPA profiler.

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to OASIS MOA of action



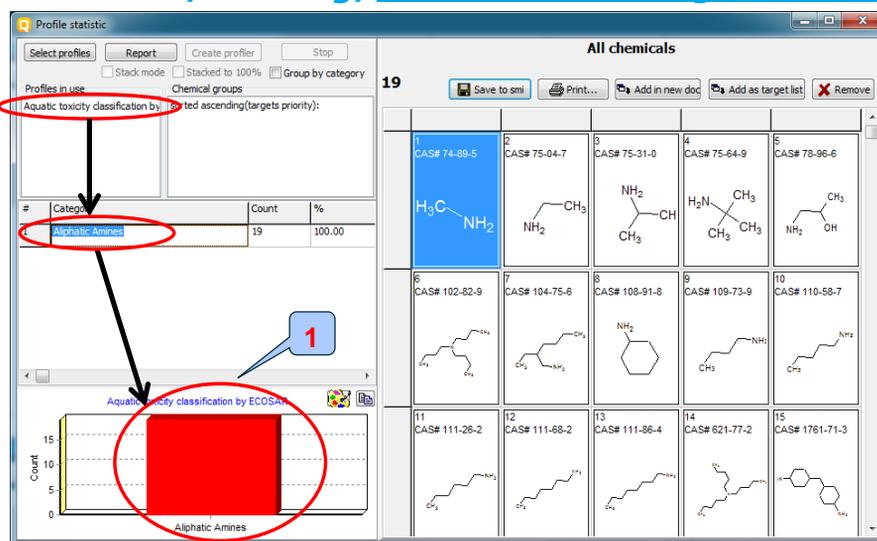
1. In this case all 19 chemicals are Narcotic amines according to MOA of action profiling scheme

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to ECOSAR

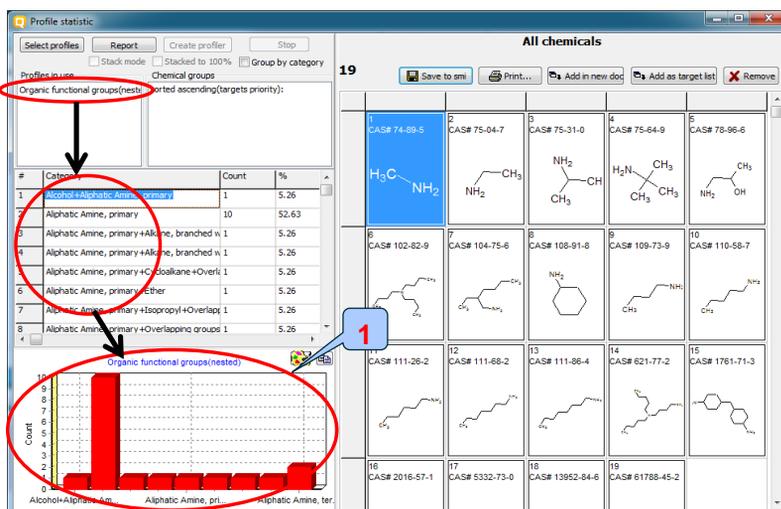


1. In this case all 19 chemicals are Aliphatic amines according to ECOSAR scheme

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to OFG(nested)



1. Aliphatic amine fragment is available.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Recap

- **Chemicals are defined as:**
 - Aliphatic amines (broader category than primary amines)
 - US EPA categories, ECOSAR classification
 - The statistics of organic functional groups provide detailed alert description of all 19 structures. However, all 19 chemicals have aliphatic amines fragment.
- **It could be concluded that the category is consistent with respect to structural functionalities (chemicals are empirically similar).**

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*

Phase I: Evaluation of Category & Data Gap Filling

Step 2. Investigating the applicability domain of an ad hoc category

- Boundaries of structural functionalities
 - Aliphatic amines
- Parametric boundaries
 - log Kow (from 0.64 to 7.71)
 - Molecular weight (from 31 to 269 Da)
 - Water solubility (from 0.48×10^{-1} to 1×10^6 mg/l)

You are now ready to extract the 2D and/or 3D parameters (see next screen shot).

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 2. Investigating the applicability domain of an ad hoc category *Extracting 2D and 3D parameters*

The screenshot displays the QSAR Toolbox interface. On the left, the 'Endpoint tree' is visible, with 'Log Kow' selected and circled in red (1). A context menu is open over the 'Log Kow' parameter, showing options like 'Calculate for all chemicals' (2), 'Calculate all parameters' (3), and 'Calculate structure' (4). The main window shows a table of chemical structures and their corresponding parameter values, with many cells containing 'Not calculated'. A message box (4) is overlaid on the table, indicating that 5 parameters were calculated.

1. **Right click** above the parameter in the endpoint tree; 2. The user can calculate all parameters or extract current parameter for all chemicals in the current row. 3. In this exercise calculate all (2D) parameters.



If some of the parameters (in case of new chemicals) are not calculated then the message (4) appears. If you want to stop the calculation process, click the **Cancel** button.

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 2. Investigating the applicability domain of an ad hoc category *Extracting 2D and 3D parametrs - results*

QSAR Toolbox 3.0.0.840 [aliphatic_amines.sm]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Profiling Profiling Schemes

Apply New View Delete

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

Profiling methods

Select All Unselect All Invert About

Prefdefined

- Database Affiliation
- Inventory Affiliation
- EC50/EC10 Chemical Categories
- Substance Type
- OEPA New Chemical Categories

General Mechanistic

- Biodeg BioC Half-Life (Blowin)
- Biodeg primary (Blowin 4)
- Biodeg probability (Blowin 1)
- Biodeg probability (Blowin 2)
- Biodeg probability (Blowin 5)
- Biodeg probability (Blowin 6)
- Biodeg probability (Blowin 7)
- Biodeg ultimate (Blowin 3)
- DNA binding by QSAR v.1.1
- DNA binding by OECD
- Estrogen Receptor Binding
- Hydrolysis half-life (Ka, pH 7)(Hydrowin)
- Hydrolysis half-life (Ka, pH 5)(Hydrowin)
- Hydrolysis half-life (Kb, pH 7)(Hydrowin)
- Hydrolysis half-life (Kb, pH 5)(Hydrowin)
- Hydrolysis half-life (Kt 6.5-7.0)

Metabolism/Transformations

Select All Unselect All Invert About

Documented

- Observed Mammalian metabolism
- Observed Microbial metabolism
- Observed Rat Liver metabolism
- Observed Rat Liver S9 metabolism

Simulated

- Autodissol simulation
- Dissociation simulation
- Hydrolysis simulator (acdk)

Filter endpoint tree...

Structure

Parameters

2D

	1 [target]	2 [target]	3 [target]	4 [target]	5 [target]	6 [target]	7 [target]
—(Q) Acidic pKa (Chemaxon)	Can not calculate	15.3	Can not calculate				
—(Q) Basic pKa (Chemaxon)	10.2	10.1	10.2	10.4	10.7	9.6	10.2
—BAF	1.78 L/kg wet	0.912 L/kg wet	0.955 L/kg wet	1.05 L/kg wet	1.12 L/kg wet	0.891 L/kg wet	56.2 L/kg wet
—BAF (lower trophic)	1.47 L/kg wet	0.955 L/kg wet	0.982 L/kg wet	1.04 L/kg wet	1.08 L/kg wet	0.944 L/kg wet	37 L/kg wet
—BAF (mid trophic)	1.53 L/kg wet	0.948 L/kg wet	0.977 L/kg wet	1.04 L/kg wet	1.09 L/kg wet	0.938 L/kg wet	41.2 L/kg wet
—BAF (upper trophic)	1.78 L/kg wet	0.914 L/kg wet	0.955 L/kg wet	1.05 L/kg wet	1.13 L/kg wet	0.899 L/kg wet	56.8 L/kg wet
—BAF (upper trophic, biotransformation rate is zero)	1.91 L/kg wet	0.923 L/kg wet	0.973 L/kg wet	1.09 L/kg wet	1.16 L/kg wet	0.966 L/kg wet	78.2 L/kg wet
—BCF	3.16 L/kg wet	33.9 L/kg wet					
—BCF (lower trophic)	1.47 L/kg wet	0.955 L/kg wet	0.982 L/kg wet	1.04 L/kg wet	1.08 L/kg wet	0.944 L/kg wet	36.9 L/kg wet
—BCF (mid trophic)	1.53 L/kg wet	0.948 L/kg wet	0.977 L/kg wet	1.04 L/kg wet	1.09 L/kg wet	0.938 L/kg wet	41.2 L/kg wet
—BCF (upper trophic)	1.78 L/kg wet	0.914 L/kg wet	0.955 L/kg wet	1.05 L/kg wet	1.13 L/kg wet	0.899 L/kg wet	56.8 L/kg wet
—BCF (upper trophic, biotransformation rate is zero)	1.89 L/kg wet	0.923 L/kg wet	0.973 L/kg wet	1.09 L/kg wet	1.16 L/kg wet	0.906 L/kg wet	71.4 L/kg wet
—Bio Half-Life	0.215 Days	0.0724 Days	0.0962 Days	0.125 Days	0.175 Days	0.025 Days	0.723 Days
—Biodeg probability (Blowin 1)	0.975	0.887	0.88	0.873	0.683	1.02	0.948
—Biodeg probability (Blowin 2)	0.993	0.975	0.97	0.964	0.795	0.985	0.984
—Biodeg probability (Blowin 5)	0.677	0.653	0.661	0.52	0.597	0.682	0.558
—Biodeg probability (Blowin 6)	0.808	0.796	0.8	0.616	0.643	0.806	0.645
—Biodeg probability (Blowin 7)	0.694	0.634	0.96	0.688	0.44	0.927	0.501
—BioC Half-life	Can not calculate						
—Biotransformation Half-Life	0.215 Days	0.0724 Days	0.0962 Days	0.125 Days	0.175 Days	0.025 Days	0.723 Days
—Boiling Point	87.7 °C	10.2 °C	36.9 °C	47.3 °C	62.8 °C	131 °C	166 °C
—Exp Boiling Point	78 °C	-6.3 °C	16.6 °C	31.8 °C	45 °C	160 °C	169 °C
—Exp Henrys Law Constant	1.76 Pa-m3/mole	1.12 Pa-m3/mole	1.25 Pa-m3/mole	4.57 Pa-m3/mole	3.63 Pa-m3/mole	Can not calculate	9.65 Pa-m3/mole

The OECD QSAR Toolbox for Grouping Chemicals into Categories

5.10.2012

29

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*
 - *Step 3: Reading data for the analogues*

Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

- Next, the Data matrix is constructed by extracting available experimental results for all 19 members of the category.
- Based on preceding category evaluation, no outliers have been identified violating the structural and mechanistic consistency of the category.

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 3. Reading data

The screenshot displays the QSAR Toolbox interface. On the left, the 'Databases' list includes 'Physical Chemical Properties', 'Environmental Fate', 'Ecotoxicology', 'Aquatic Toxicity', 'Human Health Hazards', and 'Human Health Hazards'. The 'Human Health Hazards' database is selected, indicated by a red circle and the number '1'. The 'Gather' button is highlighted with a red circle and the number '2'. A 'Read data?' dialog box is open, showing 'All endpoints' selected and 'From Tautomers' checked, with an 'OK' button highlighted and the number '3' next to it. The main window shows a table with columns for 'Structure' and multiple 'Target' columns, displaying chemical structures and their corresponding data.

1. **Select** databases related to aquatic toxicity; 2. **Click** Gather; 3. **Click** OK.

Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

Due to the overlap between the Toolbox databases same data for intersecting chemicals could be 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: 21 data-points, 10 groups, 8 chemicals.

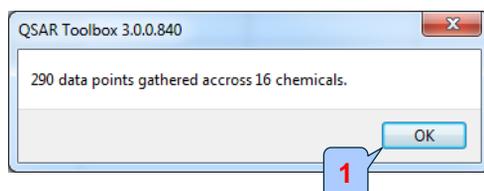
<input checked="" type="checkbox"/>	Endpoint	CAS	Structure	Value	Age
<input checked="" type="checkbox"/>		75-04-7	<chem>CCN</chem>	4.5E4 micrograms per liter	
<input checked="" type="checkbox"/>		75-04-7	<chem>CCN</chem>	4.5E4 micrograms per liter	
<input checked="" type="checkbox"/>		104-75-6	<chem>CCCC</chem>	360 micrograms per liter	
<input checked="" type="checkbox"/>		104-75-6	<chem>CCCC</chem>	360 micrograms per liter	
<input checked="" type="checkbox"/>		104-75-6	<chem>CCCC</chem>	1.2E4 micrograms per liter	
<input checked="" type="checkbox"/>		104-75-6	<chem>CCCC</chem>	1.2E4 micrograms per liter	
<input checked="" type="checkbox"/>		109-73-9	<chem>CCCC</chem>	9E3 micrograms per liter	

1. Click Select one; 2. Click OK
3. The indication for repeating values

Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

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



1. **Click OK**

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 3. Reading data. *Inserting data into the data matrix*

The screenshot shows the QSAR Toolbox interface with the 'Data Gap Filling' step selected. The main window displays a table with columns for '1 (target)', '2 (target)', '3 (target)', '4 (target)', '5 (target)', '6 (target)', '7 (target)', '8 (target)', and '9 (target)'. The table contains chemical structures and numerical data for various properties like 'Substance Identity', 'Physical Chemical Properties', 'Environmental Fate and Transport', 'Ecotoxicological Information', 'Human Health Hazards', and 'Profile'.

Filter endpoint tree...	1 (target)	2 (target)	3 (target)	4 (target)	5 (target)	6 (target)	7 (target)	8 (target)	9 (target)
Structure									
Substance Identity									
Physical Chemical Properties									
Environmental Fate and Transport									
Ecotoxicological Information (16/290)	M: 4.8(3.5, 6.6) mg/...	M: 20 mg/L, 40 mg/...	M: 178 mg/L, 16.6 ...	M: 1.04 mg/L, 1.04 ...		M: 155 mg/L, 22 m...	M: 278 mg/L, 275(2...		M: 56.9 mg/L, 56.6
Human Health Hazards									
Profile									

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*
 - *Step 3: Reading data for the analogues*
 - *Step 4: Data gap filling for 2-Butanamine*

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 4. Data Gap Filling for 2-Butanamine

The screenshot shows the QSAR Toolbox 3.0.0.840 interface. The 'Data Gap Filling' step is active. The 'Structure' panel displays the chemical structure of 2-Butanamine. The main workspace shows a table with columns labeled '1 Target', '2 Target', etc. A context menu is open over the first target column, with callouts 1, 2, and 3 indicating the steps: 1. Right click above the target chemicals, 2. Select 'Remove all as targets', and 3. Go to 'Data Gap Filling'.

1. **Right click** above the target chemicals. 2. **Select** Remove all as targets. 3. **Go** to Data Gap Filling

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine

The screenshot shows the QSAR Toolbox 3.0.0.840 interface. The main workspace displays a table with 10 columns representing different chemical structures. The first column is labeled 'Structure'. Below the structures, there is a list of properties and their corresponding values across the columns.

Filter endpoint tree...	1	2	3	4	5	6	7	8	9	10
Structure										
Substance Identity										
Physical Chemical Properties										
Environmental Fate and Transport										
Ecotoxicological Information										
Human Health Hazards										
Profile										
Ecotoxicological ... (16/29)	M: 4.8(3.5,6.6) mg/...	M: 20 mg/L, 40 mg/...	M: 178 mg/L, 16.6 ...	M: 1.04 mg/L, 1.04 ...		M: 155 mg/L, 22 m...	M: 278 mg/L, 275(2...			M: 56.9 mg/L, 56.5

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 4. Data Gap Filling for 2-Butanamine

The screenshot shows the QSAR Toolbox 3.0.0.840 interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The left sidebar has a tree view with categories like 'Physical Chemical Properties', 'Environmental Fate and Transport', and 'Ecotoxicological Information'. The central workspace displays the chemical structure of 2-Butanamine (labeled '1') and a table of data points. A context menu (labeled '2') is open over the structure, with 'Add as target' selected. A red '3' is in the top-left corner of the software window.

1. **Right click** above the target chemical (2-Butanamine); 2. **Select** Add as target. 3. **Go** to Data Gap Filling

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*
 - *Step 3: Reading data for the analogues*
 - *Step 4: Data gap filling for 2-Butanamine*
 - *Step 4.1*: How to reorder the endpoint tree*

Phase I: Evaluation of Category & Data Gap Filling

Step 4.1*. How to reorder the endpoint tree

* Detailed information describing how to reproduce the order of endpoint tree shown in the current presentation

Endpoint tree consist of two parts

Predefined

Ecotoxicological Information
Aquatic Toxicity

This part cannot be reordered or changed.

Dynamic part

Animalia }
Arthropoda(Invertebrates) }
Branchiopoda(branchiopods)}
Intoxication }
Daphnia magna }
24 h }
EC 50 }

Kingdom }
Phylum }
Class }
Effect }
Test organism }
Duration }
Endpoint }

Scientific information automatically collected from the Taxonomic bank.

Dynamic part, includes fields coming from databases. These fields have different labels like effect for intoxication, or endpoint for EC50. The position of these field in endpoint tree is managing using the **Set tree hierarchy** option (see next screen shot).

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

The screenshot displays the QSAR Toolbox software interface during the 'Data Gap Filling' phase. The main window shows a table with columns for chemical structures and various endpoints. A tree view on the left side of the table shows the hierarchy of endpoints. Three blue callout boxes with numbers 1, 2, and 3 are overlaid on the interface to indicate specific actions:

- 1:** A small blue triangle in the 'Structure' column of the table, indicating that the endpoint tree below is ordered.
- 2:** A right-click context menu over the 'Ecotoxicological Information' tree, indicating the user should right-click above the endpoint tree near this triangle.
- 3:** The 'Set tree hierarchy...' option in the context menu, indicating the user should select this option.

1. The little blue triangle is indication that the endpoint tree below is ordered. To see order of nodes of the tree, 2. **Right click** above endpoint tree near this triangle; 3. Select Set tree hierarchy.

Phase I: Evaluation of Category & Data Gap Filling

Step 4.1. How to reorder the endpoint tree

Intoxication Effect

Daphnia magna Test organism

24 h Duration

EC50 Endpoint

Left panel contains most usable labels of fields coming from databases.

The endpoint tree is organized following the order shown in right panel

The OECD QSAR Toolbox for Grouping Chemicals into Categories 5.10.2012 44

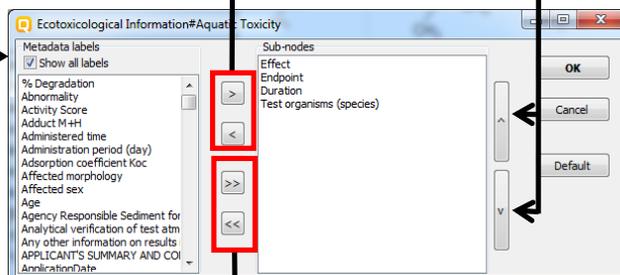
QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 4.1. How to reorder the endpoint tree

All available fields coming from databases could be shown when the box **Show all labels** is selected.

To remove or change some of the fields which are visualizing on the tree, highlight the field from left panel and move it to the right panel using the **Right** or **Left** button.

Reordering of the fields is possible using the **Up** and **Down** buttons.



To move all fields from one panel to another use **Right** and **Left** button.

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*
 - *Step 3: Reading data for the analogues*
 - *Step 4: Data gap filling for 2-Butanamine*
 - *Back to Trend analysis*

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 4. Data Gap Filling for 2-Butanamine

The screenshot shows the QSAR Toolbox software interface. The main workspace displays a list of chemical structures and their corresponding predicted values. The structures are arranged in a grid, with the first structure being 2-Butanamine. The predicted values are shown in a table below the structures:

Structure	M	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Structure															
Dap... (11/13)	M	1.9(1.5-2.4)	mg/L	M	46.6	mg/L	M	763(147-180)	mg/L	M	100	mg/L	M	91.6(77.1-104)	mg/L
	M	56(27-170)	mg/L	M	0.58(0.38-0.79)	mg/L	M	1.6(1.1-2.1)	mg/L	M	1.6(1.1-2.1)	mg/L	M	1.6(1.1-2.1)	mg/L

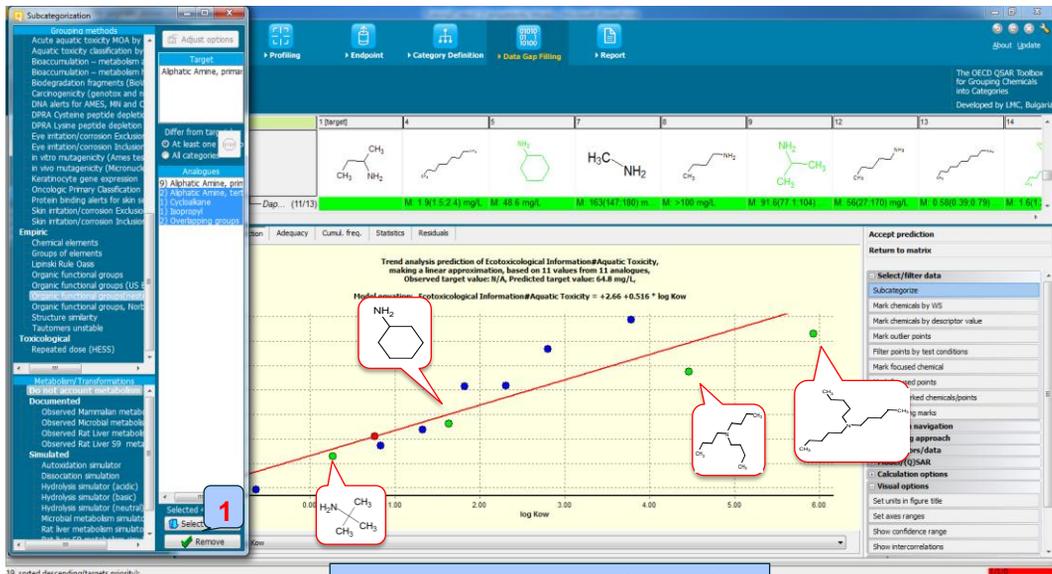
The scatter plot shows the relationship between log Kow (x-axis) and Ecotoxicological Information#Aquatic Toxicity (y-axis). The model equation is: $\text{Ecotoxicological Information\#Aquatic Toxicity} = +2.66 + 0.516 \cdot \log \text{Kow}$. The predicted target value is 64.8 mg/L. The interface also shows a 'Visual options' panel with a 'Set unit in figure title' option selected.

1. To see predicted value in log (1 mol/L) go to **Visual options** and select **Set unit in figure title**; 2. Select **mg/L**.

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

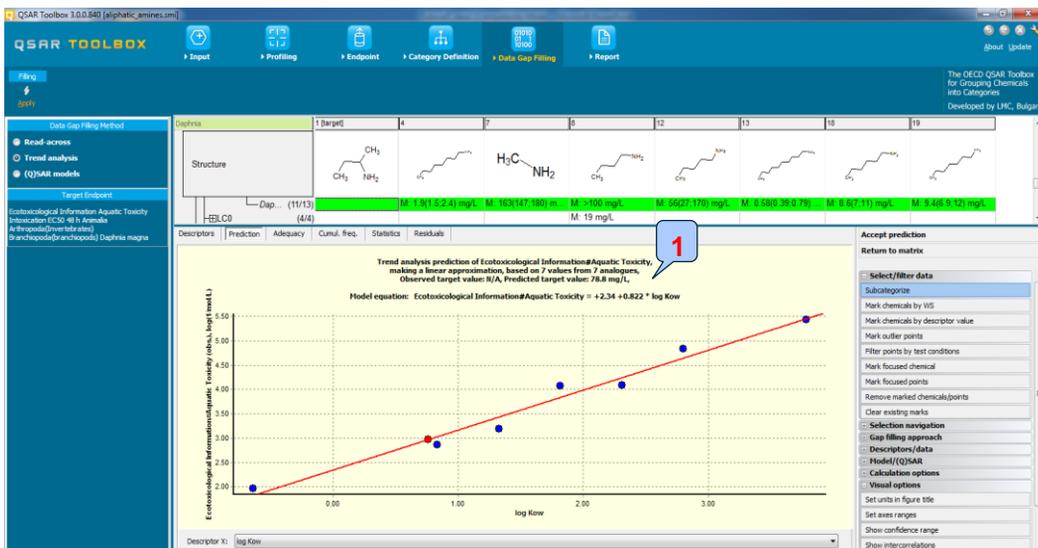
Step 4. Data Gap Filling for 2-Butanamine/Subcategorize by OFG(nested)



1. Remove dissimilar chemicals

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 4. Data Gap Filling for 2-Butanamine/*Predicted result*

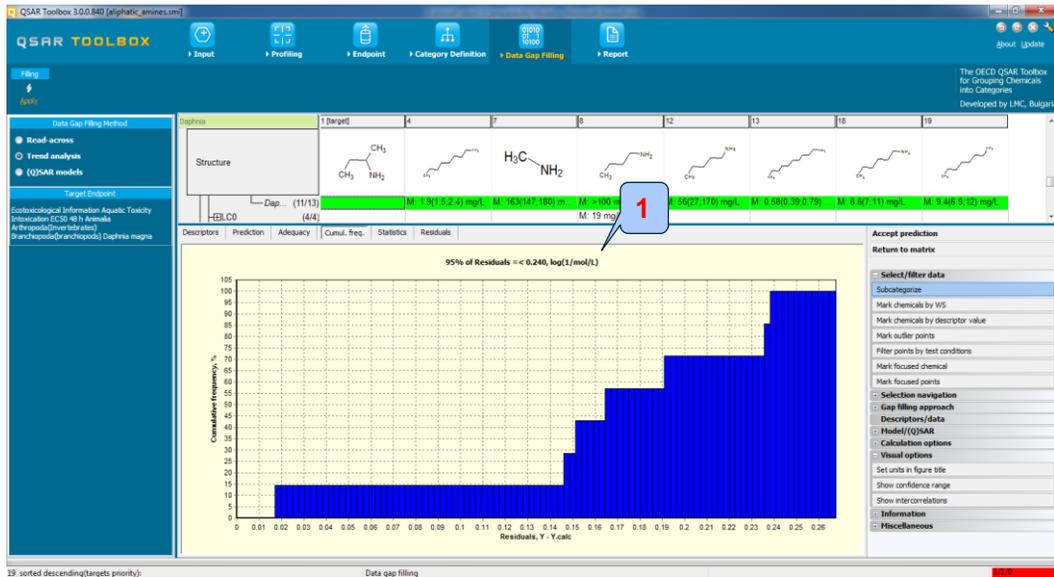


1. Predicted result is 78.8 mg/l

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine/*Cumulative frequency*



1. Cumulative frequency is less than 0.25 log units

QSAR TOOLBOX

Phase I: Evaluation of Category & Data Gap Filling Step 4. Data Gap Filling for 2-Butanamine/*Statistics*

QSAR Toolbox 3.0.0.940 [q:\phatic_amhms.sm]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition **Data Gap Filling** Report

Apply

The OECD QSAR Toolbox for Grouping Chemicals into Categories
Developed by UMC, Bulgaria

File Edit View Help

Target Endpoint: Ecotoxicological Information Aquatic Toxicity
Invertebrates LC50 (96 h) Acute
Alphapoda(Invertebrates)
Branchiopoda(Branchiopoda) Daphnia magna

Structure: CCCCN (2-Butanamine)

Target: Dap. (11/13) (44)

Target	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Structure	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>
Target	Dap. (11/13)	0.46 mg/L	M: 163(147-196) mg/L	M: 2100 mg/L	M: 2927(170) mg/L	M: 0.28(0.36-0.78)	M: 8.6(7.11) mg/L	M: 3.4(6.5-12) mg/L											
Descriptors	Prediction	Adequacy	Cumul. Freq.	Statistics	Residuals														

19 sorted descending(targets priority) Data gap filling

Accept prediction
Return to matrix

- Select/filter data
- Subcategorize
- Mark chemicals by WS
- Mark chemicals by descriptor value
- Mark outlier points
- Filter points by test conditions
- Mark focused chemical
- Mark focused points
- Selection navigation
- Gap filling approach
- Descriptors/data
- Model(QSAR)
- Calculation options
- Visual options
- Set units in figure title
- Show confidence range
- Show intercorrelations
- Information
- Miscellaneous

Statistical characteristics

Statistical characteristics	TA model
Number of data points, (N)	7
Coefficient of determination, (R ²)	0.974
Adjusted coefficient of determination, (R ² adj)	0.969
Coefficient of determination - leave one out, (Q ²)	0.949
Coefficient of correlation for external set, (Q ² ext)	-
Sum of squared residuals, (SSR)	0.220
Standard deviation of residuals, (sR)	0.210
Simple standard deviation of residuals, (s)	188
Fisher function, (F)	10.6
Fisher threshold for statistical significance, (F _α)	

b0

model descriptor	Intercept
- coeff. value	2.34
- coeff. range	± 0.34
- significance	Yes
- max. covariation	0.377 (vs b1)

b1

model descriptor	log K _{ow}
- coeff. value	0.822
- coeff. range	± 0.154
- significance	Yes
- max. covariation	0.377 (vs b0)

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine/*Interpretation of the result*

- The structurally similar analogs across category of aliphatic amines is used for data gap filling
- Subcategorization by OFG(nested) is applied
- The prediction based on the defined category is acceptable.
- The predicted value based on predefined category of aliphatic amines is 78.8 mg/l

Outlook

- Background
- Category evaluation – overview
 - Phase I. Evaluation of Category & Data Gap Filling
 - **Phase II. Extension of the Category& Data gap Filling**

Phase II: Extension Category & Data gap filling

- The extension of the category is performed by using Phase II of the category evaluation process (*Extension of Category & Data Gap Filling*). Other analogues are searched in the Toolbox, which are structurally and mechanistically consistent with the predefined category.
- The structural analogues could be defined using ECOSAR grouping

Outlook

- Background
- Category evaluation – overview
 - Phase I. Evaluation of Category & Data Gap Filling
 - **Phase II. Extension of the Category& Data gap Filling**
 - Step 1: Category definition

Phase II: Extension Category & Data gap filling

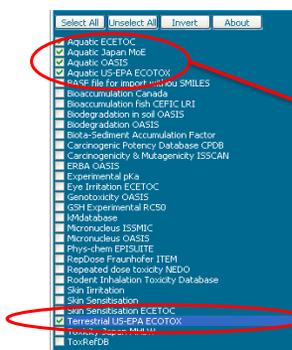
Step 1: Category definition

- The ECOSAR strict category is used to define an broader category used in further analysis.
- According to ECOSAR categorization **a list of 318 aliphatic amines have been selected** for which data are available in the Toolbox (identified in the Toolbox databases).
- The same endpoint: EC 50 48h *D.magna* will be predicted as with the predefined category

Phase II: Extension Category & Data gap filling

Step 1: Category definition

- Define ECOSAR category
- Before defining the category, the following databases related to the predicted endpoint are selected:



Aquatic ECETOC
 Aquatic Japan MoE
 Aquatic OASIS
 Aquatic US-EPA ECOTOX
 Terrestrial US-EPA ECOTOX

QSAR TOOLBOX

Phase II: Extension Category & Data gap filling Step 1: Category definition/Defining ECOSAR (strict)

The screenshot shows the QSAR Toolbox software interface. The main window is titled 'QSAR Toolbox 3.0.0.840 [aliphatic_amines.sm]'. The 'Category Definition' step is active. The 'Grouping methods' list on the left includes 'Aquatic toxicity classification by ECOSAR'. The 'Target' table shows a list of chemicals with their structures and molecular weights. The 'Aquatic toxicity classification by ECOSAR' dialog box is open, showing a list of target profiles and the 'Strict' option selected. The 'Define' button is highlighted in the top toolbar.

Target	1	2	3	4	5	6	7	8
Structure	<chem>CC1=CC=C(C=C1)N</chem>							
Substance Identity								
Ecotoxicological Information								
Aquatic Toxicity								
Avoidance	(1/1)							
Behavior	(4/12)							
Immobilisation	(1/1)							
Intoxication								
LC50	(11/12)							
- 24 h								
- 48 h								
- 96 h								
- 14 d								
- 21 d								
- 28 d								
- 35 d								
- 42 d								
- 49 d								
- 56 d								
- 63 d								
- 70 d								
- 77 d								
- 84 d								
- 91 d								
- 98 d								
- 105 d								
- 112 d								
- 119 d								
- 126 d								
- 133 d								
- 140 d								
- 147 d								
- 154 d								
- 161 d								
- 168 d								
- 175 d								
- 182 d								
- 189 d								
- 196 d								
- 203 d								
- 210 d								
- 217 d								
- 224 d								
- 231 d								
- 238 d								
- 245 d								
- 252 d								
- 259 d								
- 266 d								
- 273 d								
- 280 d								
- 287 d								
- 294 d								
- 301 d								
- 308 d								
- 315 d								
- 322 d								
- 329 d								
- 336 d								
- 343 d								
- 350 d								
- 357 d								
- 364 d								
- 371 d								
- 378 d								
- 385 d								
- 392 d								
- 399 d								
- 406 d								
- 413 d								
- 420 d								
- 427 d								
- 434 d								
- 441 d								
- 448 d								
- 455 d								
- 462 d								
- 469 d								
- 476 d								
- 483 d								
- 490 d								
- 497 d								
- 504 d								
- 511 d								
- 518 d								
- 525 d								
- 532 d								
- 539 d								
- 546 d								
- 553 d								
- 560 d								
- 567 d								
- 574 d								
- 581 d								
- 588 d								
- 595 d								
- 602 d								
- 609 d								
- 616 d								
- 623 d								
- 630 d								
- 637 d								
- 644 d								
- 651 d								
- 658 d								
- 665 d								
- 672 d								
- 679 d								
- 686 d								
- 693 d								
- 700 d								
- 707 d								
- 714 d								
- 721 d								
- 728 d								
- 735 d								
- 742 d								
- 749 d								
- 756 d								
- 763 d								
- 770 d								
- 777 d								
- 784 d								
- 791 d								
- 798 d								
- 805 d								
- 812 d								
- 819 d								
- 826 d								
- 833 d								
- 840 d								
- 847 d								
- 854 d								
- 861 d								
- 868 d								
- 875 d								
- 882 d								
- 889 d								
- 896 d								
- 903 d								
- 910 d								
- 917 d								
- 924 d								
- 931 d								
- 938 d								
- 945 d								
- 952 d								
- 959 d								
- 966 d								
- 973 d								
- 980 d								
- 987 d								
- 994 d								
- 1001 d								
- 1008 d								
- 1015 d								
- 1022 d								
- 1029 d								
- 1036 d								
- 1043 d								
- 1050 d								
- 1057 d								
- 1064 d								
- 1071 d								
- 1078 d								
- 1085 d								
- 1092 d								
- 1099 d								
- 1106 d								
- 1113 d								
- 1120 d								
- 1127 d								
- 1134 d								
- 1141 d								
- 1148 d								
- 1155 d								
- 1162 d								
- 1169 d								
- 1176 d								
- 1183 d								
- 1190 d								
- 1197 d								
- 1204 d								
- 1211 d								
- 1218 d								
- 1225 d								
- 1232 d								
- 1239 d								
- 1246 d								
- 1253 d								
- 1260 d								
- 1267 d								
- 1274 d			</					

QSAR TOOLBOX

Phase II: Extension Category & Data gap filling

Step 1: Category definition/Defining ECOSAR (strict)

The screenshot shows the QSAR Toolbox 3.0.0.840 interface. The 'Category Definition' step is active. A dialog box titled 'Define category name (021 chemicals)' is open, showing the 'Category name (021 chemicals)' field with the text 'Acute aquatic toxicity classification by ECOSAR'. The 'OK' button is highlighted with a red '1'.

Filter endpoint tree...	6	7 target	8	9	10	11	12	13	
Structure	<chem>CCCCCCCC</chem>	<chem>CC(C)N</chem>	<chem>C1=CC=C(C=C1)N</chem>	<chem>CCCCCCCC</chem>	<chem>CCCCCCCC</chem>	<chem>CCCCCCCC</chem>	<chem>C1=CC=CC=C1N</chem>	<chem>CCCCCCCC</chem>	
Substance Identity									
Physical Chemical Properties									
Environmental Fate and									
Ecotoxicological	(15/290) M: 155 mg/L, 22 m...			M: 56.9 mg/L, 56.9			M: 31 mg/L, 5.15 m...	M: 1.6 mg/L, 3.2 m...	M: 15.9 mg/L, 0.36...
Human Health Hazards									
Profile									

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

Phase II: Extension Category & Data gap filling

Step 1: Category definition/Reading data

- 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 endpoint, both options give the same results.

Phase II: Extension Category & Data gap filling

Step 1: Category definition/Reading data

Repeated values for: 1349 data-points, 439 groups, 129 chemicals

Data points...

	Endpoint	CAS	Structure	Value	Age
<input checked="" type="checkbox"/>		59-89-2		1E3 mM	
<input checked="" type="checkbox"/>		59-89-2		1E3 mM	
<input checked="" type="checkbox"/>		614-00-6		20 mM	
<input checked="" type="checkbox"/>		614-00-6		20 mM	
<input checked="" type="checkbox"/>		924-16-3		30 mM	
<input checked="" type="checkbox"/>		924-16-3		30 mM	
<input checked="" type="checkbox"/>		62-75-9		490 mM	
<input checked="" type="checkbox"/>		62-75-9		490 mM	

1

2

Select one

Invert

Check All

Uncheck All

OK

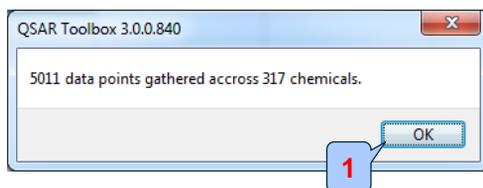
Cancel

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

Phase II: Extension Category & Data gap filling

Step 1: Category definition/Reading data

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



1. **Click OK**

Outlook

- Background
- Category evaluation – overview
 - Phase I. Evaluation of Category & Data Gap Filling
 - **Phase II. Extension of the Category& Data gap Filling**
 - Step 1: Category definition
 - Step 2: Navigate to the target endpoint

QSAR TOOLBOX

Phase II: Extension Category & Data gap filling Step 2: Navigate to the target endpoint/EC 50, 48h, *D.magna*

The screenshot shows the QSAR Toolbox interface with the following components:

- Top Bar:** Navigation tabs for Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report.
- Left Sidebar:** A tree view of endpoints. The 'Daphnia' node is expanded, and 'Daphnia magna' is selected. A red callout '1' points to this node.
- Main Table:** A table with columns for chemical structures and data points. A red callout '2' points to a gap in the data for the 'Daphnia magna' endpoint.
- Bottom Left:** A blue box with a red callout '3' pointing to the 'Apply' button.

1. Navigate to target endpoint by **opening the nodes**; 2. **Highlight the gap** which will be fill in. 3. **Move** to Data Gap Filling

Outlook

- Background
- Category evaluation – overview
 - Phase I. Evaluation of Category & Data Gap Filling
 - **Phase II. Extension of the Category& Data gap Filling**
 - Step 1: Category definition
 - Step 2: Navigate to the target endpoint
 - Step 3: Data Gap Filling

QSAR TOOLBOX

Phase II: Extension Category & Data gap filling Step 3: Data Gap Filling

The screenshot displays the QSAR Toolbox 2.3.0.1130 interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The left sidebar shows 'Read across', 'Trend analysis', and '(Q)SAR models'. The main area features a tree view of 'Daphnia' categories and a data table with columns for different endpoints. A callout box with '1' points to the 'Trend analysis' option, and another callout box with '2' points to the 'Apply' button in the top left.

Structure	1 (Target)	2	3	4	5	6	7	8
Structure	<chem>CC(N)C</chem>	<chem>Cc1ccccc1</chem>	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(O)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>
Substance Identity								
Parameters								
Ecotoxicological Information								
Aquatic Toxicity								
Avoidance	(5/5)						M: 100 mg/L	
Behavior	(12/40)						M: =1.2 mg/L, =19 ...	M: =268 mg/L, =77... M: =286 mg/L, =98 ...
Immobilisation	(25/25)		M: 15 mg/L				M: 58 mg/L	
Intoxication								
Animalia								
Anthropoda(Invertebrates)								
Branchiopoda(branchiopods)								
Ceriodaphnia dubia	(2/2)							
Daphnia magna	(48/76)						M: 14 mg/L, 16 mg/L ...	M: 164(161-168) mg/L, M: 48(44-52) mg/L
24 h								
48 h	(52/64)						M: 226 mg/L	
EC50	(3/3)						M: 46 mg/L	
LC50	(1/1)							M: 46(40.1-52.8) mg/L
96 h	(3/4)							
Daphnia pulex	(28/127)						M: 26.5(20.4-34.4) ...	M: 164 mg/L, 55.5 ... M: 48.3 mg/L, 49.4 ...
Mortality	(35/96)		M: 4.7 mg/L, 8.4 mg/L				M: 0.16 mg/L, 0.32 ...	M: 4.2 mg/L, 5.7 mg/L, M: 1-15 mg/L
Reproduction								
Profile								

1. Select Trend analysis; 2. Click Apply

QSAR TOOLBOX

Phase II: Extension Category & Data gap filling Step 3: Data Gap Filling

QSAR Toolbox 3.0.0.840 [aliphatic_amines.sm]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

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

Target Endpoint: Daphnia magna (52/64)

Ecotoxicological Information: Aquatic Toxicity (inhibition EC50, log10 Aquatic A@90pods@invertebrates), Branchpoints@branchpoints, Daphnia magna

Structure: CC(C)N CC(C)N C1CCN1 C1CCN1 C1CCN1 C1CCN1 C1CCN1 C1CCN1 C1CCN1 C1CCN1

Descriptors: Prediction Adequacy Cumul. Freq. Statistics Residuals

Trend analysis prediction of Ecotoxicological Information#Aquatic Toxicity, making a linear approximation, based on 52 values from 52 analogues, Observed target value: 8/A, Predicted target value: 10.0 mg/L.

Model equation: Ecotoxicological Information#Aquatic Toxicity = +3.63 + 0.309 * log Kow

Descriptor X: log Kow

Accept prediction

Return to matrix

- Select/filter data
- Selection navigation
- Gap filling approach
- Descriptors/data
- Model/QSAR
- Calculation options
- Visual options
- Information
- Miscellaneous

QSAR TOOLBOX

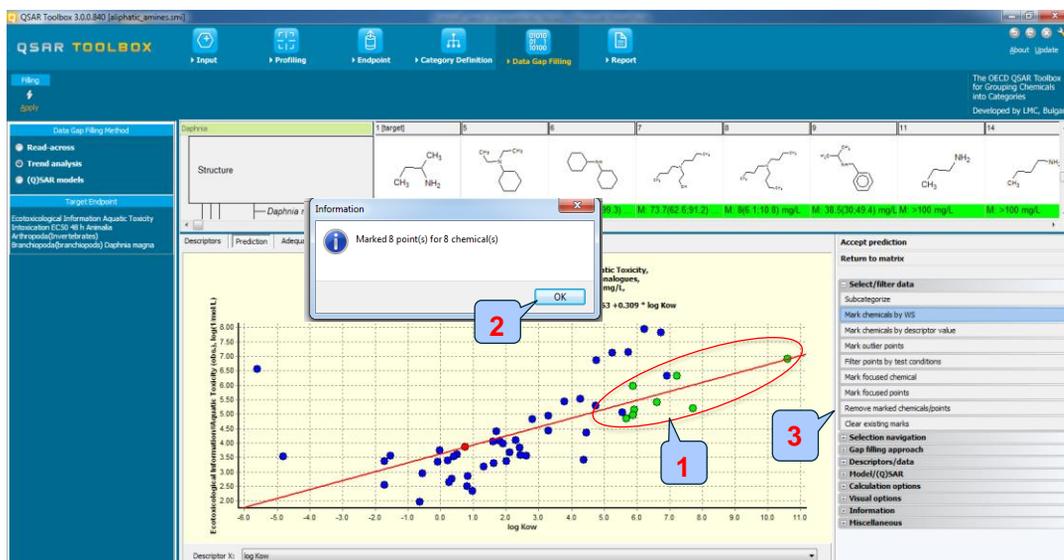
Phase II: Extension Category & Data gap filling Step 3: Data Gap Filling/Subcategorize by Water solubility

The screenshot shows the QSAR Toolbox 3.0.0.840 interface. The 'Data Gap Filling' method is selected. The main workspace displays a table of chemical structures and their predicted toxicity values. A scatter plot shows the relationship between log Kow and Ecotoxicological Information. A 'Select/Filter data' panel is open on the right, with 'Water Solubility (fragments)' selected. Red callout boxes 1, 2, 3, and 4 highlight specific actions: 1. Open Select/Filter data panel; 2. Click Mark chemical by Sw; 3. Select Water solubility (fragments); 4. Click OK.

1. **Open** Select/Filter data panel; 2. **Click** Mark chemical by Sw; 3. **Select** Water solubility (fragments); 4. **Click** OK

QSAR TOOLBOX

Phase II: Extension Category & Data gap filling Step 3: Data Gap Filling/Subcategorize by Water solubility



1. Eight chemicals are marked in green. 2. Click OK. 3. Remove marked chemicals

QSAR TOOLBOX

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/Subcategorize by OFG(nested)

The screenshot shows the QSAR Toolbox interface during the Subcategorization step. The 'Subcategorization' panel on the left lists various methods, with 'Aliphatic Amine, primary' selected as the target. The main window displays a table of chemical structures and their predicted values for Daphnia magna. Below the table is a scatter plot showing the trend analysis prediction of Ecotoxicological Information@Aquatic Toxicity, with the model equation: $\text{Ecotoxicological Information@Aquatic Toxicity} = +3.63 + 0.309 \cdot \log Kow$. The plot shows a positive linear correlation between log Kow and Ecotoxicity. On the right, the 'Accept prediction' panel is visible, with a 'Return to matrix' section containing a list of subcategories and filter options. Three numbered callouts (1, 2, 3) point to specific actions: 1 points to 'Subcategories by WS', 2 points to 'Remove' in the Subcategorization panel, and 3 points to 'Selected 25 (17/52)' in the same panel.

1. Click Subcategorize 2. Select OFG(nested) 3. Remove dissimilar chemicals

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/Subcategorize by Lipinski rules

The screenshot displays the QSAR Toolbox software interface during the 'Subcategorization' phase. The main window shows a graph of Toxicity (log Kow) versus log Kow. A red line represents the predicted toxicity. A pop-up window titled '0 Structures Item: Not bioavailable' lists several chemicals, including CASR 124-00-1, CASR 143-27-1, CASR 2016-43-4, CASR 2575-26-5, CASR 2889-34-3, and CASR 4205-95-7. The interface also includes a sidebar with various methods and a 'Subcategorization' panel on the right with options like 'Select/filter data' and 'Accept prediction'.

Chemicals with very long chain could be removed from the category due to their non-bioavailability.

1. **Select** Lipinski rules
2. **Double click** to see "Non-bioavailable" chemicals
3. **Close** the appeared window
4. **Remove** dissimilar chemicals

QSAR TOOLBOX

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/Prediction result

Structure: CC(C)N, CCN, CCCN, CCCCN, CCCCCN, CCCCCCN, CCCCCN, CCCCCN, CCCCCN, CCCCCN, CCCCCN

Target	11	14	19	20	21	23	28
M	>100 mg/L	>100 mg/L	8.9(7.11) mg/L	9.4(6.9-12) mg/L	1.9(1.5-2.4) mg/L	1.6(1.2-2) mg/L	0.020(0.018-0)

Trend analysis prediction of Ecotoxicological Information: Aquatic Toxicity, making a linear approximation, based on 11 values from 11 analogues.
Observed target value: N/A, Predicted target value: 77.7 mg/L

Model equation: $\text{Ecotoxicological Information: Aquatic Toxicity} = +2.34 + 0.838 * \log Kow$

Descriptor X: log Kow

Prediction result is 77.7 mg/l.

QSAR TOOLBOX

Phase II: Extension Category & Data gap filling Step 3: Data Gap Filling/Accept the Prediction result

The screenshot displays the QSAR Toolbox software interface during the 'Data Gap Filling' phase. The main workspace shows a grid of chemical structures for various log Kow values (1, 14, 19, 20, 21, 23, 26) for the target endpoint 'Daphnia magna'. A central plot area shows a scatter plot of 'Ecotoxicological Information Aquatic Toxicity (Inhibition EC50 48 h Animals)' versus 'log Kow', with a red regression line. A dialog box is open in the center, asking 'The current model is still not saved. Do you want to save it now?' with 'Yes' and 'No' buttons. A right sidebar contains the 'Accept prediction' and 'Return to matrix' buttons. Red callout boxes with numbers 1, 2, and 3 point to the 'Accept prediction' button, the dialog box, and the 'Return to matrix' button respectively.

1. **Click** on Accept prediction. If you want to save the model click Yes, otherwise click No; 2. **Click** No; 3. **Click** Return to the matrix