Outlook

• **Background**
• **Keywords**
• **Objectives**
• **The exercise**
• **Workflow of the exercise**
Background

• This is a step-by-step presentation designed to take you through the workflow of the Toolbox for building a QSAR model for predicting aquatic toxicity.

• By now you have some experience in using the Toolbox so there will be multiple key strokes between screen shots.

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

• Background
• **Keywords**
• Objectives
• The exercise
• Workflow of the exercise
Keywords

TARGET CHEMICAL - chemical of interest

MODULE – a Toolbox module is a section dedicated to specific actions and options (e.g. Profiling)

WORKFLOW – the use, in combination, of the different modules (e.g. prediction workflow: from input to report)

PROFILER - algorithm (rule set) for the identification of specific features of the chemicals. Several types of profilers are available, such as structural (e.g. Organic functional groups), mechanistic (e.g. Protein binding by OECD) and endpoint-specific (e.g. in vitro in vitro mutagenicity (Ames test) alerts by ISS) profilers.

ALERT - the profilers consist of sets of rules or alerts. Each of the rules consists of a set of queries. The queries could be related to the chemical structure, physicochemical properties, experimental data, comparison with the target or list with substances and external queries from other predefined profilers (reference queries).

CATEGORY – “group” of substances sharing same characteristics (e.g. the same functional groups or mode of action). In a typical Toolbox workflow, it consists of the target chemical and its analogues gathered according to the selected profilers

ENDPOINT TREE – Endpoints are structured in a branched scheme, from a broader level (Phys-Chem properties, Environmental Fate and transport, Ecotoxicology, Human health hazard) to a more detailed one (e.g. EC3 in LLNA test under Human health hazard- Skin sensitization)

DATA MATRIX – Table reporting the chemical(s) and data (experimental results, profilers outcomes, predictions). Each chemical is in a different column and each data in a different row

(Q)SAR - (Q)SAR models can be used to fill a data gap if no adequate analogues are found for a target chemical
Outlook

- Background
- Keywords
- **Objectives**
- The exercise
- Workflow of the exercise
This presentation demonstrates building a QSAR model for predicting acute toxicity of aldehydes to *Tetrahymena pyriformis*. The presentation addresses specifically:

- predicting acute toxicity for a target chemical;
- building a QSAR model based on the prediction;
- applying the model to other aldehydes;
- exporting the predictions to a file.
Outlook

• Background
• Keywords
• Objectives
• **The exercise**
• Workflow of the exercise
The Exercise

• This exercise includes the following steps:
  • select a target chemical – Furfural, CAS 98-01-1;
  • extract available experimental results;
  • search for analogues;
  • estimate the target endpoint: 48h-IGC50 for *Tetrahymena pyriformis* by using trend analysis;
  • improve the data set by either:
    • subcategorizing by “Protein binding” mechanisms, or
    • assessing the difference between outliers and the target chemical
  • evaluate and save the model;
  • use the model to display its training set, visualize its applicability domain and perform predictions.
Outlook

• Background
• Keywords
• Objectives
• The exercise

• Workflow of the exercise
Workflow of the exercise

• Remember the Toolbox has 6 modules which are used in a sequential workflow:

  • Input
  • Profiling
  • Data
  • Category Definition
  • Data Gap Filling
  • Report
Outlook

• Background
• Keywords
• Objectives
• The exercise
• **Workflow of the exercise**
  • Input
1. Click on CAS# 2. Enter CAS# 98-01-1; 3. Click Search;
The Toolbox now searches the Toolbox databases and inventories for the presence of the chemical with structure related to the current CAS number. It is displayed as a 2D image. Note it is unselected by default.

1. Mark desired chemical (in case there is only one chemical it is marked by default); 2. Click **OK** to add chemical in data matrix;
Input
Target chemical identity

• Target chemical is displayed on the data matrix.

• To see chemical identification click on the box next to “Structure info” (see next screen shot).
Chemical Input
Target chemical identity
Outlook

• Background
• Keywords
• Objectives
• The exercise
• **Workflow of the exercise**
  • Input
    • Define Target Endpoint
In this exercise we will build a QSAR model to estimate the following endpoint:

Ecotoxicological Information#Aquatic Toxicity#Growth#IGC50#48h#Protozoa#Ciliophora#Ciliatea #Tetrahymena pyriformis

For defining the target endpoint the “Define target endpoint” functionality is used (see next few slides)
Input
Define target endpoint

- Defining of the endpoint allows entering the endpoint of interest e.g. EC3, LC50, gene mutation etc., along with specific metadata information. Based on the metadata, relevancy of the profiles and databases is provided expressed in different highlighting:
  - In green are highlighted the most suitable profilers related to the endpoint and databases including data for the defined target endpoint, while
  - in the orange are colored profilers which are plausible with respect to the defined target endpoint.
Input
Define target endpoint

1. Click Define;
2. Select Aquatic Toxicity;
3. Click Next and consecutively add the following endpoint and metadata (4): Endpoint – IGC50; Effect – Growth; Duration – 48h; Test organism (species): *Tetrahymena pyriformis*; 5. Click Finish
Input
Define target endpoint

The endpoint tree is automatically expanded to the level of the defined endpoint and the row is highlighted in yellow.

⚠️ As mentioned above (slide 19) defining the target endpoint lead to highlighting of relevant profilers and databases (see next slides)
Outlook

• Background
• Keywords
• Objectives
• The exercise
• **Workflow of the exercise**
  • Input
  • Profiling
“Profiling” module refers to the electronic process of retrieving relevant information on the target compound, other than environmental fate, ecotoxicity and toxicity data, which are stored in the Toolbox database.

Available “profilers” includes likely mechanism(s) of action, which could be useful in forming categories that include the target chemical.

“Profilers” are a collection of empirical and mechanism knowledge which could be used to analyse the structural properties of chemicals.

The “profilers” identify the affiliation of the target chemical(s) to preliminary defined categories (functional groups/alerts).

The “Profiling” module contains also observed and simulated metabolisms/transformations, which could be used in combination with the profilers.

The outcome of the profiling determines the most appropriate way to search for analogues, but they are also useful for preliminary screening or prioritization of substances.

The “profilers” are not (Q)SARs, i.e. they are not prediction models themselves;

Based on the “profilers’ relevancy” (determined by the defined target endpoint), the most suitable and plausible ones are getting colour highlighted.
Profiling
Profiling the target chemical

• Select the “Profiling methods” related to the target endpoint

• This selects (a green check mark appears) or deselects (green check disappears) profilers.

• In this case select all green (the most suitable to the target endpoint) profilers – see next slide
Profiling
Profiling the target chemical

1. Go to **Profiling** module
2. Select all suitable (marked in green) profilers
3. Click **Apply** to apply knowledge of the selected profilers to the target chemical
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 (see next screen shot).

• Green rectangles in some result boxes indicate there is more than one profiling result and the field needs to be expanded.
1. Double click on the cell with “Aldehydes (Acute toxicity)” results based on US-EPA Chemical New Chemical Categories to see why the chemical is categorized as aldehyde.

2. Literature information is displayed. The knowledge explained here is used for coding the structural boundaries of the category.

Continued on next slide
1. Structural boundaries of the category - Aldehydes (Acute toxicity); The boundaries which are met are ticked with green.
2. Definition of the SMARTS used for coding the knowledge; Visualization of the common fragment used for coding the knowledge;
3. The target molecule and highlighted (red) part of the molecule meeting the structure boundary.
1) In module Profile, you have profiled the target chemical according to the suitable profilers (green) related to the target endpoint.

2) The target chemical is categorized as “aldehyde” based on predefined Acute aquatic toxicity US-EPA profiler (hereafter called US-EPA) and the two endpoint-specific profilers (Acute aquatic toxicity classification by ECOSAR (hereafter called ECOSAR) and Acute aquatic toxicity MOA by OASIS (hereafter called MOA)).

3) By the endpoint-specific “Acute aquatic toxicity classification by Verhaar” the target is categorized as “Class 3 (unspecific reactivity)”

4) Moreover the target is categorized as “aldehyde” based on Protein binding by OASIS reactiving by Schiff-base formation mechanism.

5) In general the target is classified as “aldehyde”

6) All of the above mentioned profilers could be used for categorization purposes (collecting analogues).

7) In this case US-EPA profiler will be used for categorization purpose (primary grouping).
Outlook

• Background
• Keywords
• Objectives
• The exercise
• **Workflow of the exercise**
  • Input
  • Profiling
  • **Data**
Data

• *Data* module refers to the electronic process of retrieving the environmental fate, ecotoxicity and toxicity data that are stored in the Toolbox databases.

• Data gathering can be executed in a global fashion (i.e. collecting all data of all endpoints) or on a more narrowly defined basis (i.e. collecting data for a single or limited number of endpoints).

• Once the endpoint is selected, the relevant databases are highlighted. Meaning of the colors could be seen within the **Options** (1) by click **Legend** (2).

• In this example, we limit our data gathering to the databases containing aquatic toxicity data for the defined target endpoint (Aquatic OASIS).
1. Go to **Data** module
2. Select the green highlighted database
3. Click **Gather**. 3 data points are collected for the target. A single data point is found for the target endpoint; We will try to reproduce it.
4. Click **OK**
Outlook

• Background
• Keywords
• Objectives
• The exercise
• **Workflow of the exercise**
  • Chemical Input
  • Profiling
  • Data
  • **Category definition**
Category definition
Defining US-EPA category

• As mentioned before, the initial search for analogues is based on structural similarity, of US EPA categorization

• Select US-EPA New Chemical Category

• Click Define (see next screen shot)
1. Go to **Category definition** module; 2. Highlight “US-EPA New Chemical Categories”; 3. Click **Define**; 4. Put a tick in the Strict box (see next screenshot); 5. Click OK to confirm the category **Aldehydes (Acute toxicity)**;
Category definition
Defining US-EPA category strict functionality

• The **Strict** functionality means that the software will group analogues having **ONLY** the categories of the target and will exclude the analogues having any other categories according to the profiler used in the grouping method.

• For example, if the profiling for the target results in *Aldehydes (Acute toxicity)* **ONLY** according to US-EPA category, the group of analogues will include *Aldehydes (Acute toxicity)* **ONLY**. (See next screen shot)
Category definition
Defining US-EPA category strict functionality

The target and analogues have Aldehydes ONLY according to US-EPA category

Defined Category

Analogue 1
Analogue 2
Analogue 3

Analogue 4
Analogue 5
The Toolbox now identifies all chemicals corresponding to *Aldehydes (Acute toxicity)* by US-EPA listed in the databases selected under “Data”.

- 101 analogues including the target chemical are identified; they form a mechanistic category “*Aldehydes (Acute toxicity)*”, which will be used for gap filling.
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). Click OK to read all available data. 175 data points are collected for the list of 101 analogues.
After a message for a number of data collected, the experimental results for the target and analogues are inserted into the matrix.
Outlook

• Background
• Keywords
• Objectives
• The exercise

• **Workflow of the exercise**
  • Chemical Input
  • Profiling
  • Data
  • Category definition
  • **Data gap filling**
Data Gap Filling
Overview

• *Data Gap Filling* module gives access to five different data gap filling tools:
  - Read-across
  - Trend analysis
  - (Q)SAR models
  - Standardized workflow (SW)
  - Automated workflow (AW)

• The most relevant data gap mechanism is used, taking into account the following considerations:
  - *Read-across* is the appropriate data-gap filling method for “qualitative” endpoints like skin sensitisation or mutagenicity for which a limited number of results are possible (e.g. positive, negative, equivocal). Furthermore read-across is recommended for “quantitative endpoints” (e.g., 96h-LC50 for fish) if only a low number of analogues with experimental results are identified.
  - *Trend analysis* is the appropriate data-gap filling method for “quantitative endpoints” (e.g., 96h-LC50 for fish) if a high number of analogues with experimental results are identified.
  - *(Q)SAR models* can be used to fill a data gap if no adequate analogues are found for a target chemical.
  - *Automated and standardized workflows* follow preliminary implemented logic. The AW is not affected by the user activities (proceeding or subsequent), while the SW stops at the each step of the workflows allowing the user to make different selection.

• In this example we will use trend analysis.
1. Go to **Data Gap Filling**; 2. Highlight the data gap corresponding to target endpoint: IGC50, *Tetrahymena pyriformis* under the target chemical; 3. Select **Trend analysis**;
Data Gap Filling
(IGC 50 48h of *T. pyriformis*)
Apply Trend analysis

- A message for possible data inconsistency appears
- It is recommended the log(1/mol/L) scale to be chosen

The resulting plot can be seen on next screen shot
Data Gap Filling
(IGC 50 48h of *T. pyriformis*)
Data Gap Filling
(IGC 50 48h of *T. pyriformis*)
Interpreting dots on the graph

• The resulting plot outlines the experimental results of all analogues (Y axis) according to a descriptor (X axis) with LogKow being the default descriptor (see previous screen shot).

• The **RED** dot represents the predicted value for target chemical.

• The **ORANGE** dot represents the observed data value for the target chemical.

• The **BLUE** dots represent the experimental results available for the analogues.

• The **LIGHT BLUE** dots (see the following screen shots) represent analogues belonging to different subcategories.
In this example, the mechanistic properties of the analogues are consistent.

Subcategorization can be performed based on protein binding mechanisms. This is the second stage of analogue search - requiring the same interaction mechanism.

Acute effects are associated with covalent interaction of chemicals within cell proteins, i.e. with protein binding.

Chemicals with a different protein binding mechanism / reactions compared to the target chemical will be removed.
After the available data has been retrieved, the user can then further subcategorize the results according to the following endpoint-specific subcategorizations:

- Acute aquatic toxicity MOA by OASIS
- Protein binding by OASIS
- Aquatic toxicity classification by ECOSAR

These steps are summarized in the next screen shots.
Data Gap Filling (IGC 50 48h of *T. pyriformis*)
Subcategorization 1: Acute aquatic toxicity MOA by OASIS

1. Click **Select / filter data**, then **Subcategorize**; 2. Select “MOA by OASIS” (Note: the most suitable profilers for subcategorization are again green highlighted); 3. **Click** “Remove selected” to eliminate dissimilar to the target analogues (in this case analogues categorized as “reactive unspecified” based on MOA profiler will be eliminated)
Data Gap Filling (IGC 50 48h of *T. pyriformis*)
Subcategorization 2: Protein binding by OASIS

1. Select "Protein binding by OASIS";
2. Click "Remove selected" to eliminate dissimilar to the target analogues.
Data Gap Filling (IGC 50 48h of *T. pyriformis*)

Subcategorization 3: Aquatic toxicity classification by ECOSAR

1. Select “Aquatic toxicity classification by ECOSAR”;
2. Click “Remove selected” to eliminate the single analogue;
Data Gap Filling (IGC 50 48h of *T. pyriformis*).

Results after subcategorisation.

1. Click “Accept prediction”;
2. Click “Yes” (“No” allows to continue with the subcategorization).
Data Gap Filling
(IGC 50 48h of *T. pyriformis*)
Evaluation of the model

• To assess the model accuracy use:
  • Adequacy (predictions after leave-one-out)
  • Statistics
  • Cumulative frequency
  • Residuals

• See next four screen shots
Data Gap Filling (IGC 50 48h of *T. pyriformis*)
Evaluation of the model - Adequacy

1. Position on the last level of document tree
2. Click “Adequacy”;
1. Click “Cumulative frequency”; The residuals abs (obs-predicted) for 95% of analogues are comparable with the experimental error.
Data Gap Filling (IGC 50 48h of *T. pyriformis*)
Evaluation of the model - Residuals

1. Click "Residuals"
Data Gap Filling (IGC 50 48h of *T. pyriformis*)

Evaluation of the model - Statistics

1. Click "Statistics"
Data Gap Filling (IGC 50 48h of T. pyriformis)  
Results after subcategorisation
Data Gap Filling
(IGC 50 48h of *T. pyriformis*)
Save the derived QSAR model

• To save the new regression model follow these steps:
  • Go to the last row on the Document tree
  • Click on “Model/QSAR”
  • Select Save model
  • Enter the model name and fill editable fields if necessary
  • Click on OK
Data Gap Filling (IGC 50 48h of *T. pyriformis*)

Save the derived QSAR model

1. Click “Model/QSAR”, then “Save model”;
2. Type name of the model and fill the fields in the wizard if necessary (Use Next/Back buttons to navigate within it);
3. Click “Save model”;
4. Click OK.
Outlook

• Background
• Keywords
• Objectives
• The exercise

• **Workflow of the exercise**
  • Input
  • Profiling
  • Data
  • Category definition
  • Data gap filling
  • **QSAR model**
1. Select a non-Gap filling list from the documented tree; 2. Note the accepted prediction will be inserted into data matrix 3. Click "**(Q)SAR**"; 4. The derived QSAR is listed in the panel with Relevant (Q)SAR models.
Data Gap Filling
How to see the derived QSAR?

As seen in the next five screen shots the derived model can be used to:

• Visualize training set of the model;

• Visualize the domain of the model;

• Visualize whether a chemical is in the domain of the model;

• Enter in Data Gap filling;

• Perform predictions for:
  • Selected chemical
  • All chemicals (in the matrix)
  • Chemicals in domain.
Data Gap Filling
Visualisation of the training set

1. Right click on the derived **QSAR model**; 2. Select **Show training set**; 3. Note the experimental data is displayed under CAS# of each chemical; 4. The training set can be saved as *.smi file.
Data Gap Filling
Visualisation of model domain

1. Right click on the derived **QSAR model**; 2. Select “Display Domain”; 3. Note the boundaries of the domain are combined logically; 4. If the chemical answers the query of the domain then the current query is a labelled with **GREEN** tick; 5. Otherwise is labelled with **RED** cross.
Data Gap Filling
Visualisation of whether a chemical is in the domain of the model

1. Highlight the cell of one of the analogues (e.g., chemical # 6 in the data matrix); 2. Click on "(Q)SAR"; 3. A message informs you that the QSAR is applied not on the target chemical. Click Yes; 4. Right click above the model and Left click on Display domain (see next screen shot).
The chemical is an “aldehyde” as required by US-EPA categorization group (boundary 1 on next screen shot).

The chemical is an “aldehyde” as required by Acute aquatic toxicity MOA by OASIS group (boundary 2) and to be not “reactive unspecified” (boundary 3)

It can react with protein by Schiff-base formation (boundary 4) and should not belong to any of the eliminated mechanistic domains according to Protein binding by OASIS (boundary 5):

- Michael addition ($\alpha,\beta$-Aldehydes, Conjugated systems with electron withdrawing groups) (boundary 5)
- SNAr (Activated aryl and heteroaryl compounds) (boundary 5)
- Schiff base formation (Bis aldehydes, Di-substituted $\alpha,\beta$-unsaturated aldehydes and Aromatic carbonyl compounds) (boundary 5)

The chemical should be an “aldehyde” as required by Aquatic toxicity classification by ECOSAR (boundary 6) and not to be “imidazoles” (boundary 7).

Another requirement is Log Kow to be $\geq 0.308$ and $\leq 4.77$ (boundary 8):
Data Gap Filling
Visualisation of whether a chemical is in the domain of the model

The target chemical is out of the model domain due to:
1) Belonging to “Michael addition” mechanism by “Protein binding by OASIS” profiler, which have been eliminated from the domain (negated by logical “NOT”) (boundary 5)
2) The chemical is not an “aldehyde” as requested by ECOSAR profiler (boundary 6).

⚠️ The definitive designation for belonging or not to the domain is the collectible boundary (3) which is red crossed in case of “Out of domain” (green checked in case of “In domain”)
Data Gap Filling
Enter Gap filling

Go to target chemical and call (Q)SAR:
1. Select the model; 2. Click Run; 3. Select Enter Gap filling; 4. Click OK; Then you will be transferred automatically to Gap filling and can operate (not shown);
Data Gap Filling
Perform prediction for chemicals in domain (for selected chemical and all chemicals - analogically)

1. Select the QSAR model; 2. Click Run; 3. Select Predict Chemicals in domain; 4. Click OK;
Data Gap Filling
Perform prediction for chemicals in domain

The process of applying the model is indicated by status bar on the bottom of the window. The predictions are placed on the matrix. Note there are different signs for the origin of the data: M for experimental data, T for result of Trend analysis, Q for prediction from QSAR model.
Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow of the exercise**
  - Input
  - Profiling
  - Data
  - Category definition
  - Data gap filling
    - QSAR model
    - **Export QSAR prediction**
Export QSAR results

• The QSAR predictions for the chemicals in the matrix can be exported into a file

• In the Endpoint tree right click on Tetrahymena pyriformis (for the endpoint IGC50 48h for Tetrahymena pyriformis) and select Export Data matrix from the context menu (see next three screen shots).
1. Right click on the row of endpoint tree associated with predictions from the QSAR model; 2. Select **Export Data matrix** (see next screen shot).
1. The nodes from the tree associated with QSAR predictions which will be exported are selected with check marks; 2. Click Export; 3. Browse to save the file on your PC; 4. Give a name of the file; 5. Click Save; 6. Click OK when the file is exported.
The resulting file in *.csv format can be opened via Microsoft Excel and further analysed.
Outlook

• Background
• Keywords
• Objectives
• The exercise
• **Workflow of the exercise**
  • Input
  • Profiling
  • Data
  • Category definition
  • Data gap filling
  • **Report**
1. Go to **Report** module; 2. Click **QMRF**; 3. Select the name of the user-defined QSAR model; 4. Click **OK**;
1. Navigate through the Wizard to customize the report; 2. Select Create report; 3. Choose QMRF report and then Open (4) to create a PDF format of the report or click Save as if you want to save the file; 5. Choose Training set in order to create a MS Excel file (training set of the QSAR along with their data) or 6. Click Save as;
QMRF report

3.1. Species: *T. pyriformis*, Growth 48h

3.2. Endpoints: IC50

3.3. Comment on the endpoints
Not available

3.4. Endpoint units:
Not available

3.5. Dependent variable:
Not available

3.6. Experimental protocol:
Not available

3.7. Endpoint data quality and variability:
Not available

4. Defining the algorithm (OECD Principle 2)

4.1. Type of model:
Not available

4.2. Training algorithm:
Not available

4.3. Descriptors in the model:
log Kow

4.4. Descriptor selection:
Not available

4.5. Algorithms and descriptor generation:

<table>
<thead>
<tr>
<th>Substance Identity</th>
<th>Training set #1</th>
<th>Training set #2</th>
<th>Training set #3</th>
<th>Training set #4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAS number</td>
<td>87-96-1</td>
<td>112-44-7</td>
<td>112-31-2</td>
<td>613-45-6</td>
</tr>
<tr>
<td>Chemical name</td>
<td>Ethylbutanol</td>
<td>C11-H22-0</td>
<td>Decanal</td>
<td>2,4-DIMETHYLBENZALDEHYDE</td>
</tr>
<tr>
<td>Other identifier</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMILES</td>
<td>CCC(C)(=O)</td>
<td>CCCCCCCCCO</td>
<td>CCCCCCCCCO</td>
<td>CCCCCCCCCO</td>
</tr>
<tr>
<td>Parameters</td>
<td>unit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profiles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Training set data and user gathered

<table>
<thead>
<tr>
<th>Sublevel</th>
<th>Endpoint</th>
<th>Species, duration, test type, type of method, assay, strain, test guideline, year</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute Toxicity</td>
<td>IC50</td>
<td>Tetrathymena pyriformis, micro</td>
<td>3.5</td>
<td>mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>T. pyriformis</em> micro</td>
<td>0.2</td>
<td>mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>T. pyriformis</em> micro</td>
<td>191</td>
<td>mg/L</td>
</tr>
</tbody>
</table>
Congratulations!

- You have used the Toolbox to build a user-defined QSAR model.
- You now know another useful tool in the Toolbox.
- Continue to practice with this and other tools. Soon you will be comfortable dealing with many situations where the Toolbox is useful.