

OECD QSAR Toolbox v.3.4

Step-by-step example of how to build a user-defined profiling scheme

Outlook

- **Background**
- Objectives
- Profiling
- The exercise

Background

- This is a step-by-step presentation designed to provide guidance to users of the Toolbox on how to create their own profiling scheme.

Outlook

- Background
- **Objectives**
- Profiling
- The exercise

Objectives

- **This presentation demonstrates how to build a new profiling scheme including the:**
 - naming of the new scheme
 - building a category by defining its boundaries
 - saving the new profiler

Outlook

- Background
- Objectives
- **Profiling**
- The exercise

Profiling

Overview

- As you are aware “Profiling” refers to the electronic process of retrieving relevant information on a compound which is stored in the Toolbox, other than fate and toxicity data.
- The Toolbox has many predefined profilers but it also allows the user to development new profilers.

Outlook

- Background
- Objectives
- Profiling
- **The exercise**

The Exercise

In this example we build a profiler that identifies chemicals:

- that are aldehydes (rule 1)
- that can react with proteins by “Schiff base formation” (rule 2), but
- that do not react with proteins by “Michael-type nucleophilic addition” (rule 3)
- which also have the fragment C(=O)(O)c1ccccc1 in their structure (rule 4), and
- that have a value of Log KOW between 1 and 7 (rule 5).

The Exercise

Start building a new profiler

We are going to create a new profiling scheme:

- Open the Toolbox.
- Move to the Profiling module
(see next screen shot).

Building of a new profiler

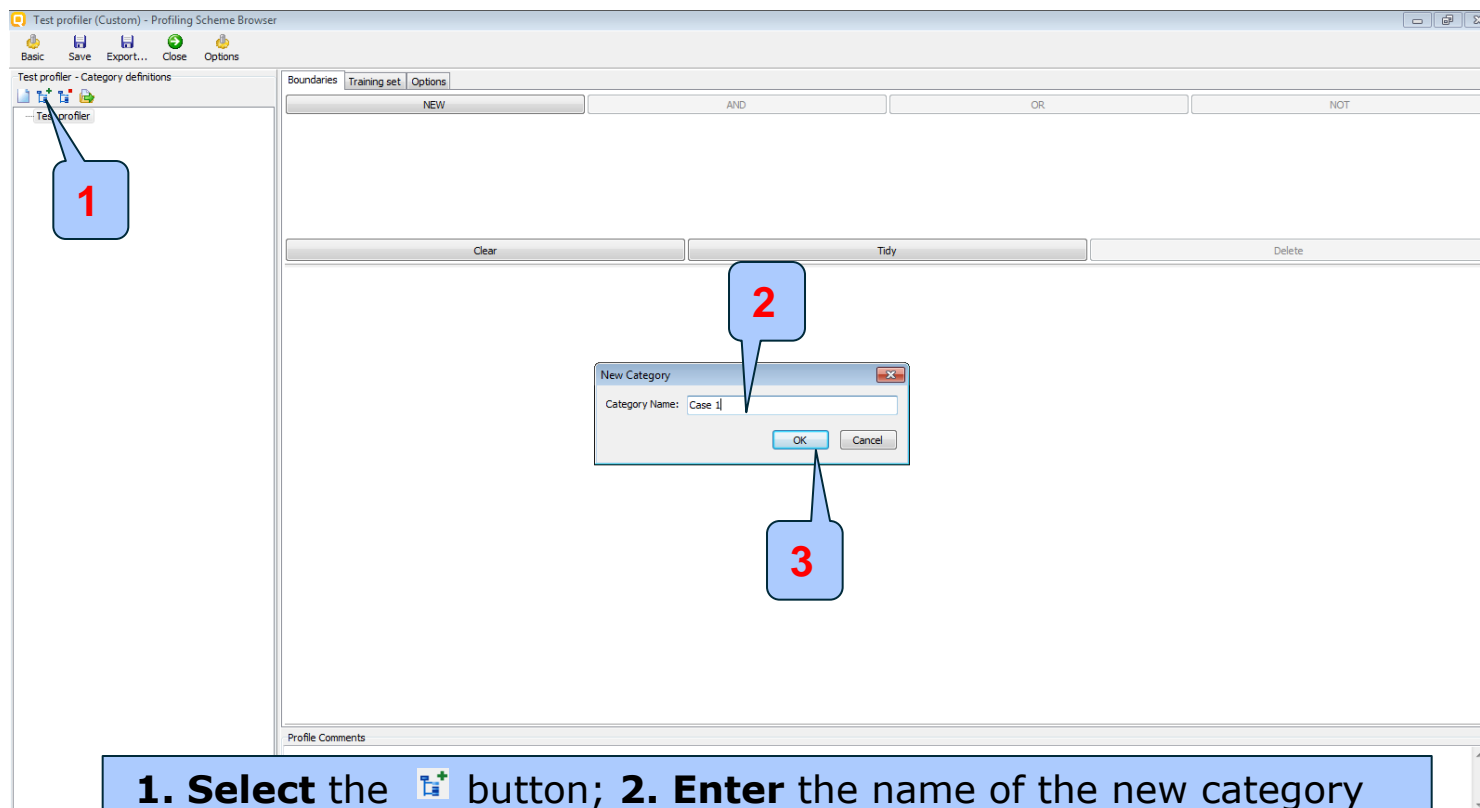
Define the name of the new profiler

The screenshot shows the QSAR Toolbox software interface. The 'Profiling' menu is open, and the 'New Profiling Scheme' dialog box is displayed. The name 'Test Profiler' is entered in the text field, and the 'OK' button is highlighted. Numbered callouts (1-4) indicate the steps: 1. Click Profiling, 2. Click New, 3. Enter the name 'Test Profiler', and 4. Click OK.

1. Select the Profiling mode; 2. Click New; 3. Enter the name of the scheme "Test Profiler"; 4. Click OK.

Building of a new profiler

Add a new category within the profiler



1. Select the  button; **2. Enter** the name of the new category "Case 1"; **3. Click** OK.

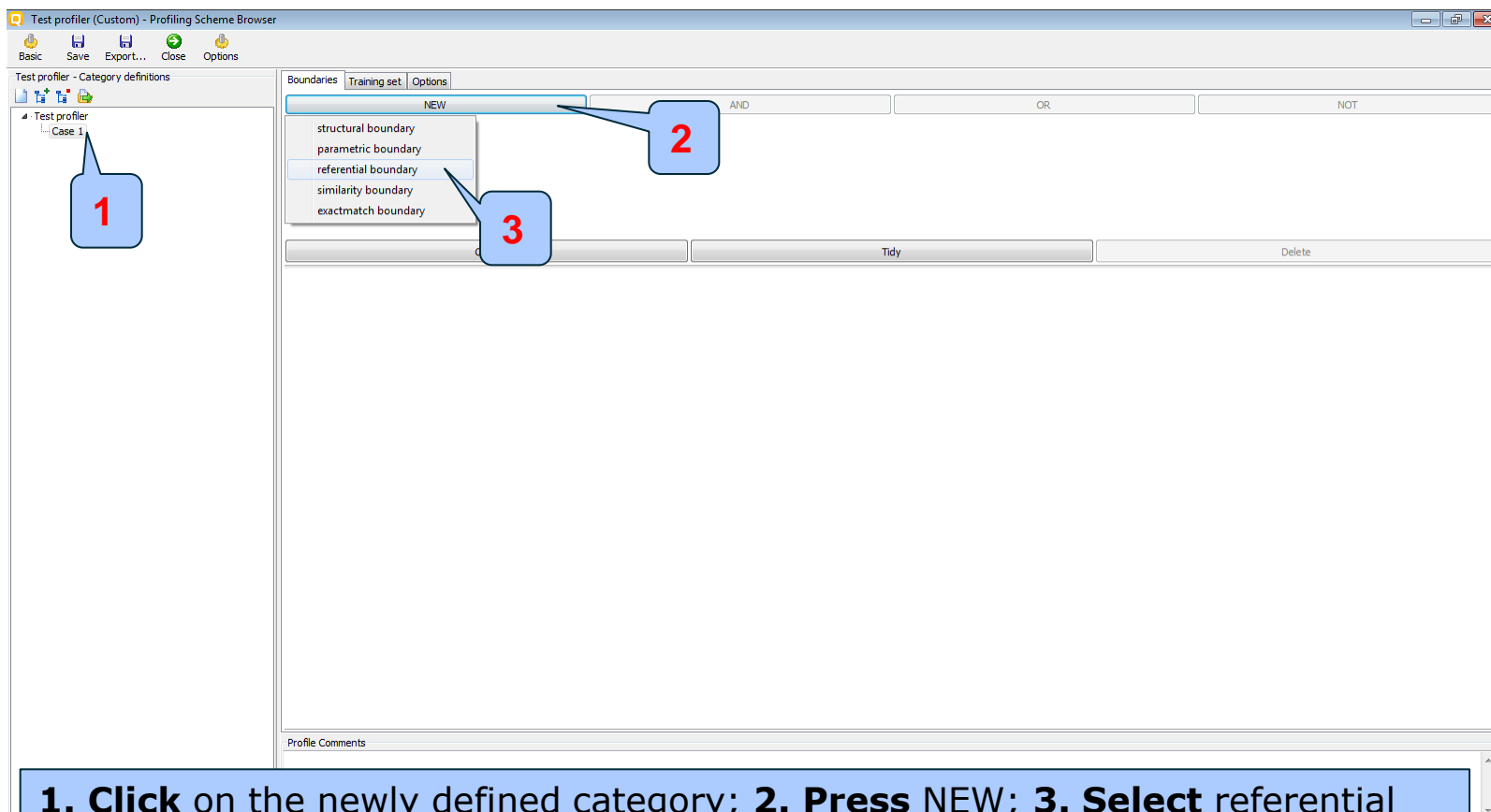
Building of a new profiler

Building the category definition – rule 1

To specify aldehydes, include a referential boundary making use of the predefined category definition "Aldehydes" within the Organic functional groups profiler.

Building of a new profiler

Building the category definition – rule 1



1. Click on the newly defined category; **2. Press** NEW; **3. Select** referential boundary.

Building of a new profiler

Building the category definition – rule 1

The screenshot shows the 'Test profiler (Custom) - Profiling Scheme Browser' window. The left pane shows a tree view under 'Empiric' with 'Organic Functional groups' highlighted by a red circle and a callout box labeled '1'. The central workspace has a 'NEW' button and logical operators 'AND' and 'OR'. Below these are 'Clear' and 'Tidy' buttons. The right pane shows 'Available categories' with 'Aldehyde' selected and a callout box labeled '2'. Below the available categories is an up arrow button with a callout box labeled '3'.

1. From Empiric profilers **highlight** Organic functional groups; **2. Select** Aldehydes from the available categories; **3. Press** up arrow button

Building of a new profiler

Building the category definition – rule 1

The screenshot displays the 'Test profiler (Custom) - Profiling Scheme Browser' application. The main window is divided into several sections:

- Toolbar:** Includes icons for 'Basic', 'Save', 'Export...', 'Close', and 'Options'. Below these are buttons for 'NEW', 'AND', and 'OR'.
- Left Panel:** 'Test profiler - Category definitions' showing a tree structure with 'Test profiler' and 'Case 1'.
- Center Panel:** 'Boundary Options' with a 'Metabolism' tab selected. Below it, 'Profiling schemes' are listed, including 'Predefined', 'General Mechanistic', 'Endpoint Specific', 'Empiric', 'Toxicological', 'Experimental', and 'Custom'. The 'Empiric' sub-tree is expanded, showing categories like 'Chemical elements', 'Groups of elements', 'Lipinski Rule Oasis', 'Organic Functional groups', etc.
- Right Panel:** 'Selected categories' (highlighted with a red circle) and 'Available categories'. The 'Available categories' list includes: Acetal, Acetoxym, Acid anhydride, Acid anhydride, mixed phosphonic, Acridine, Acridone/ Acridinimine, Acrylamide, Acrylate, Acyl halide, Acylal, Acyloin, Alcohol, Aldimine, Aldoxime derivatives, Aliphatic Amine, primary, Aliphatic Amine, secondary, Aliphatic Amine, tertiary.
- Bottom:** 'Multiple Categories' section with radio buttons for 'Strict', 'OR-ed' (selected), and 'AND-ed'. A 'Profile Comments' field is at the very bottom.

Building of a new profiler

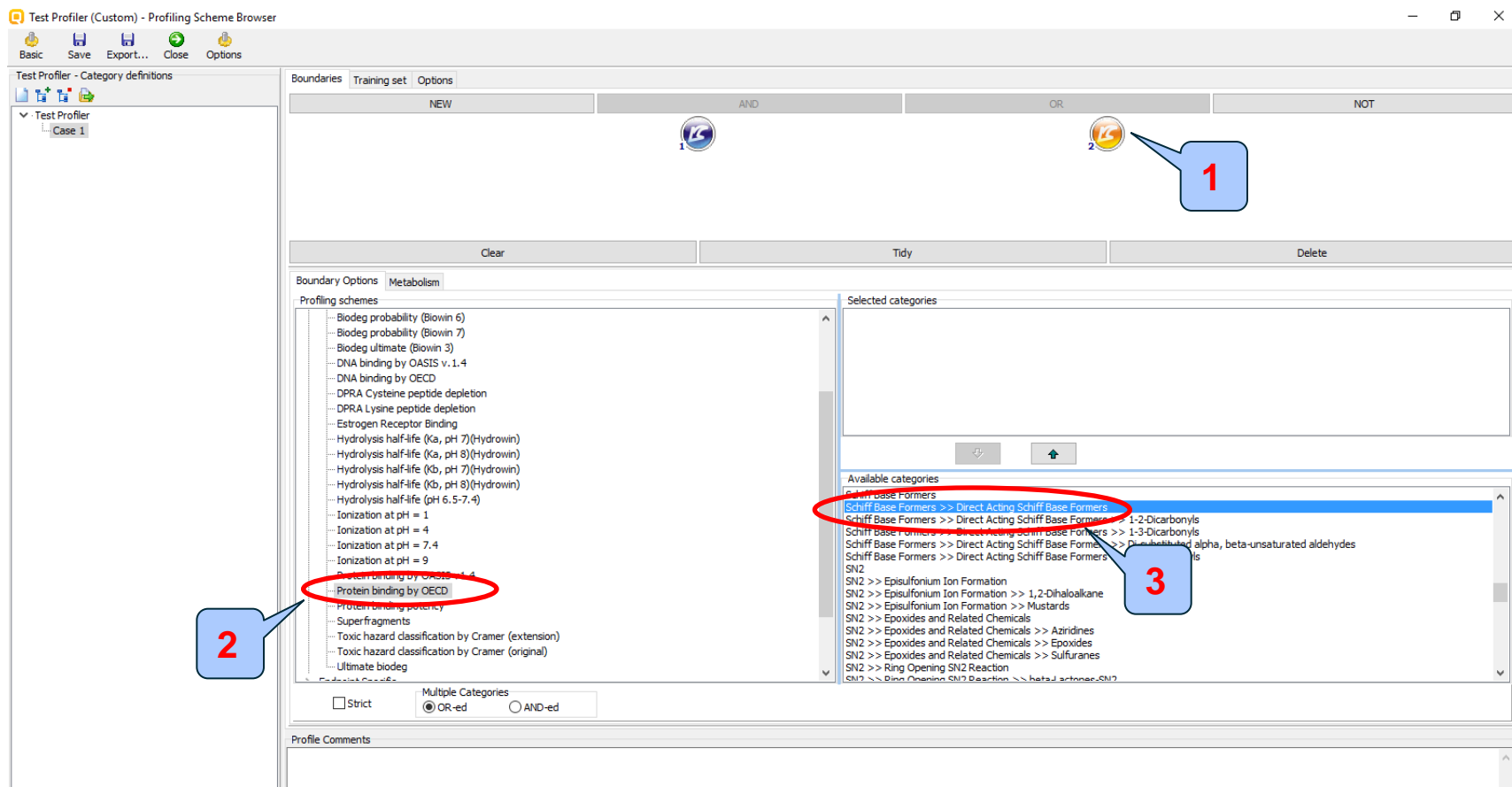
Restriction of the category definition

To restrict the category definition by mechanism, add two additional referential boundaries:

- the first one specifies chemicals that can react with proteins by forming a Schiff base (**rule 2**) according to Protein binding by OECD.
- the other one specifies chemicals that do NOT react with proteins by Michael-type nucleophilic addition (**rule 3**) according to Protein binding by OECD.

Building of a new profiler

Building the category definition – rule 2



1. **Create** new referential boundary; 2. From the General mechanistic profilers **select** Protein binding by OECD; 3. **Select** Schiff base formers>>Direct Acting Schiff Base Formers category from the panel with available categories and move it to the panel Selected categories

Building of a new profiler

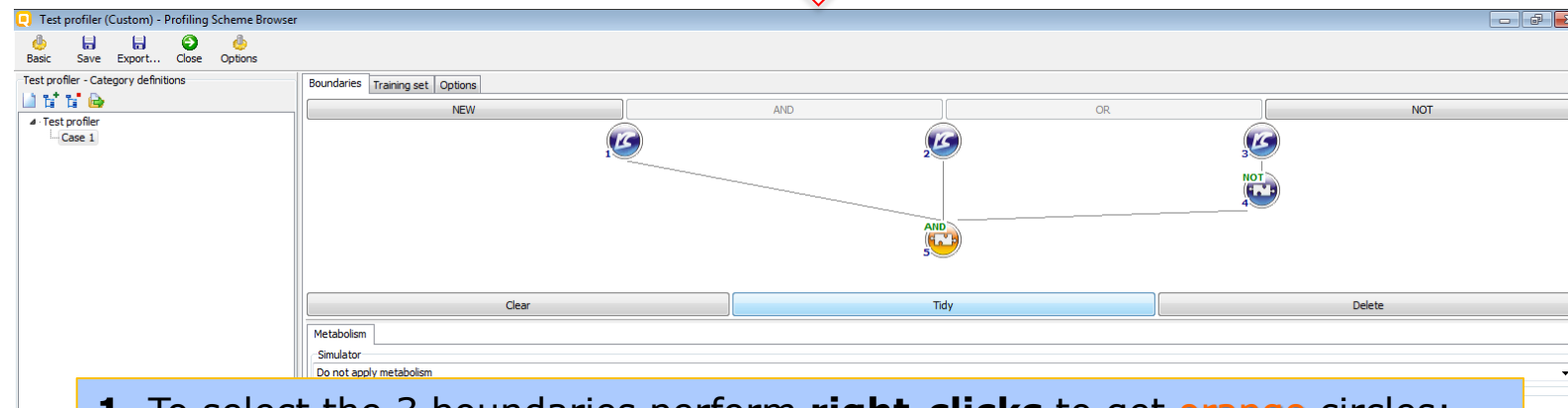
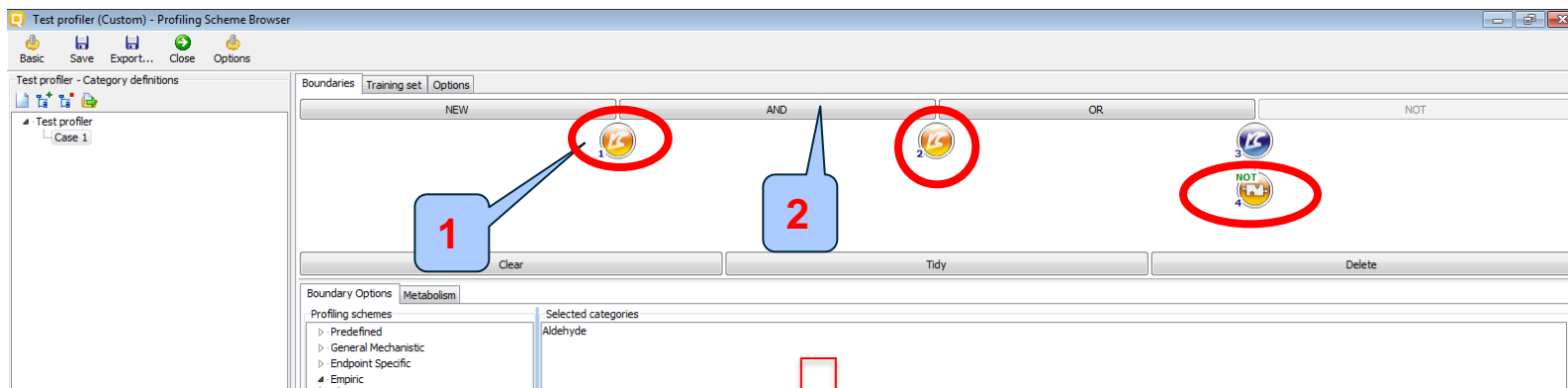
Building the category definition – rule 3

The screenshot shows the 'Test profiler (Custom) - Profiling Scheme Browser' window. The 'NEW' button is highlighted with a blue callout '1'. In the 'Profiling schemes' list, 'Protein binding by OECD' is circled in red with a blue callout '2'. In the 'Selected categories' list, 'Michael addition' is circled in red with a blue callout '3'. The 'NOT' button is highlighted with a blue callout '4'. The 'Available categories' list is visible below, showing a hierarchy of chemical categories.

1. Create new referential boundary; **2. Select** Protein binding by OECD; **3. Select** Michael addition category; **4. to negate** the boundary **click** NOT.

Building of a new profiler

Grouping the referential boundaries.



1. To select the 3 boundaries perform **right-clicks** to get orange circles;
2. Click **AND**.

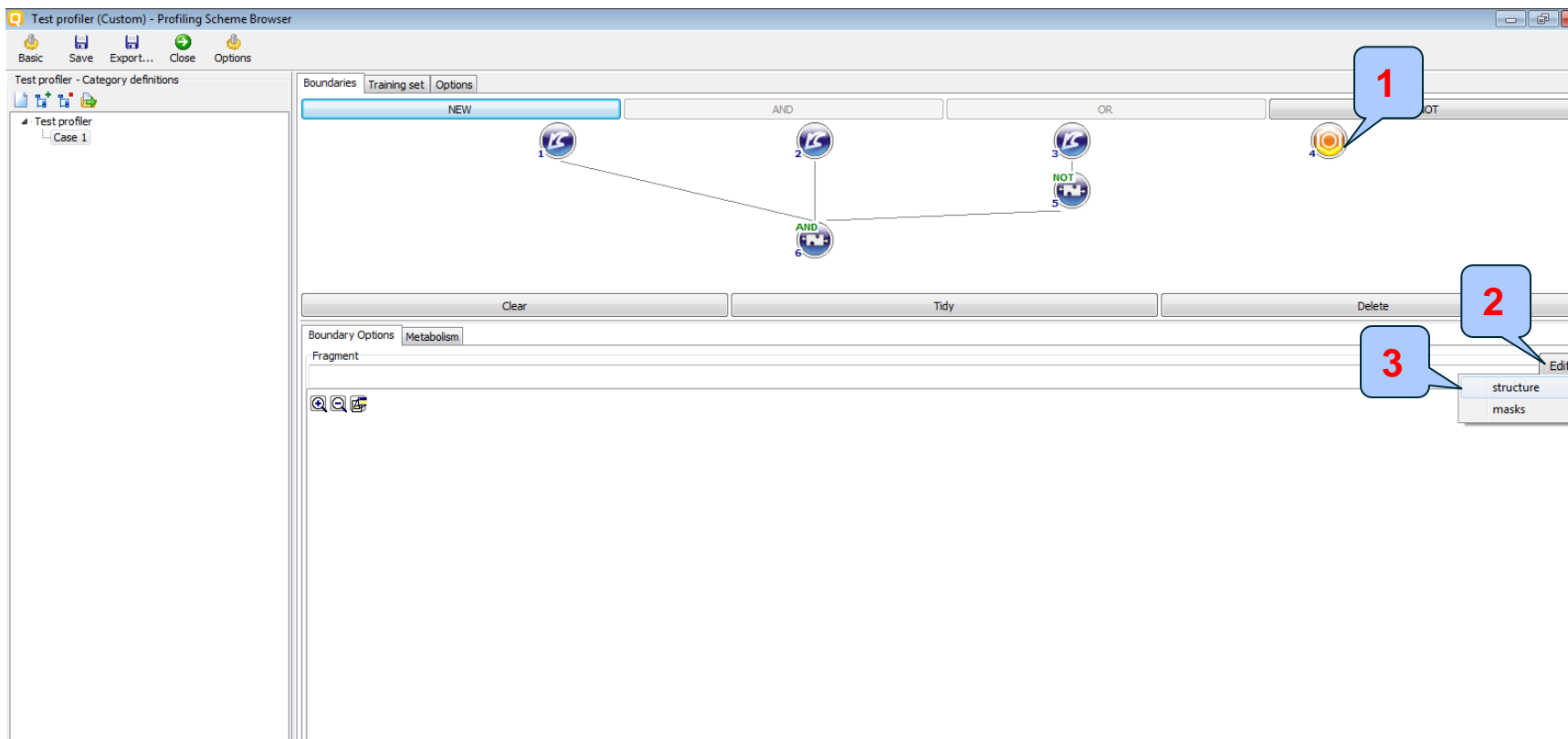
Building of a new profiler

Restriction of the category definition

To complement the category definition add an instruction for ignoring those structures which have the fragment C(=O)(O)c1ccccc1 (where the aldehyde group is deactivated - no binding with protein) - **rule 4**.

Building of a new profiler

Building the category definition – rule 4



1. Create a structural boundary, then **right-click** and changing the circle to **orange** ;
2. **Click** Edit;
3. **Select** Structure and (see next screen shot).

Building of a new profiler

Building the category definition – rule 4

The screenshot shows the 2D Editor window with the following elements:

- 1:** The SMILES string C(=O)c1ccccc1 is entered in the text field and highlighted with a red circle.
- 2:** The 'Draw' button is highlighted with a red circle.
- 3:** The 'OK' button is highlighted with a red circle.

The chemical structure of benzaldehyde is displayed on the right side of the editor.

- 1. Type** the SMILES of the fragment (or draw it in the 2D editor window);
- 2. Click Draw;**
- 3. Click OK.**

Building of a new profiler

Building the category definition – rule 4

1. Right-click over the circle with structural boundary in order to select it (it is orange); then 2. Click NOT to negate the boundary.

Building of a new profiler

Parametric range

To specify the needed parametric range, add a parametric boundary and combine it together with the rest boundaries in an AND query. The parametric range is log Kow between 1 and 7 (rule 5).

Building of a new profiler

Building the category definition – rule 5

Test profiler (Custom) - Profiling Scheme Browser

Test profiler - Category definitions

Boundaries Training set Options

structural boundary W AND OR NOT

parametric boundary 1 2 3 4 5 6 7 8

referential boundary

similarity boundary

exactmatch boundary

Clear Tidy Delete

Boundary Options Metabolism

Parameter Name

- Half-Life (Model Lake)
- Half-Life (Model River)
- Henrys Law Constant (Bond Method)
- Henrys Law Constant (Group Method)
- HIA Multicase
- HOMO Energy
- Hydrolysis half-life (pH 6.5-7.4)
- Ka half-life (pH 7)
- Ka half-life (pH 8)
- Kb half-life (pH 7)
- Kb half-life (pH 8)
- kM
- Koc (Log Kow)
- Koc (MCI)
- Kp (Mackay model)
- Kp (Octanol/air (Koa) model)
- Lipid Solubility
- log BCF max
- Log Koa (Air-water partition coefficient model)
- Log Koa (Henry's law constant model)
- log Kow**
- logP Multicase
- LUMO Energy
- Maximum distance
- Maximum donor delocalizability
- Mean Melting Point
- Melting Point (Arbutov Method)

log Kow

Log Octanol-Water Partition Coefficient (version 1.67 estimate)

Expression

between 1 7

Profile Comments

1. Create the parametric boundary and change the circle to **orange**; **2. From** the list with parameters **select** log Kow; **3. Select** qualifier between and enter the requisite values (1 to 7).

Building of a new profiler

Combine boundaries

Finally, group the combined referential boundaries, the structural restriction and the parametric range in one single **AND** query.

Building of a new profiler

Combine boundaries

Test profiler (Custom) - Profiling Scheme Browser

Basic Save Export... Close Options

Test profiler - Category definitions

- Test profiler
 - Case 1

Boundaries Training set Options

NEW AND OR NOT

Clear Tidy Delete

Metabolism

Simulator

Do not apply metabolism

Process

- Parent
- Metabolites Use parent if none
- All

Match

- Any
- All
- Accumulatively

Ignore inorganic metabolites

Profile Comments

Building of a new profiler

Add a new category within the profiler

The screenshot shows the 'Test profiler (Custom) - Profiling Scheme Browser' window. On the left, under 'Test profiler - Category definitions', 'Case 2' is selected and circled in red, with a callout box labeled '1'. In the main area, the 'NEW' button is highlighted, and a dropdown menu is open showing 'exactmatch boundary' selected, with a callout box labeled '2'. Below the main area are buttons for 'Clear', 'Tidy', and 'Delete', and a 'Boundary Options' section with 'Metabolism' selected.

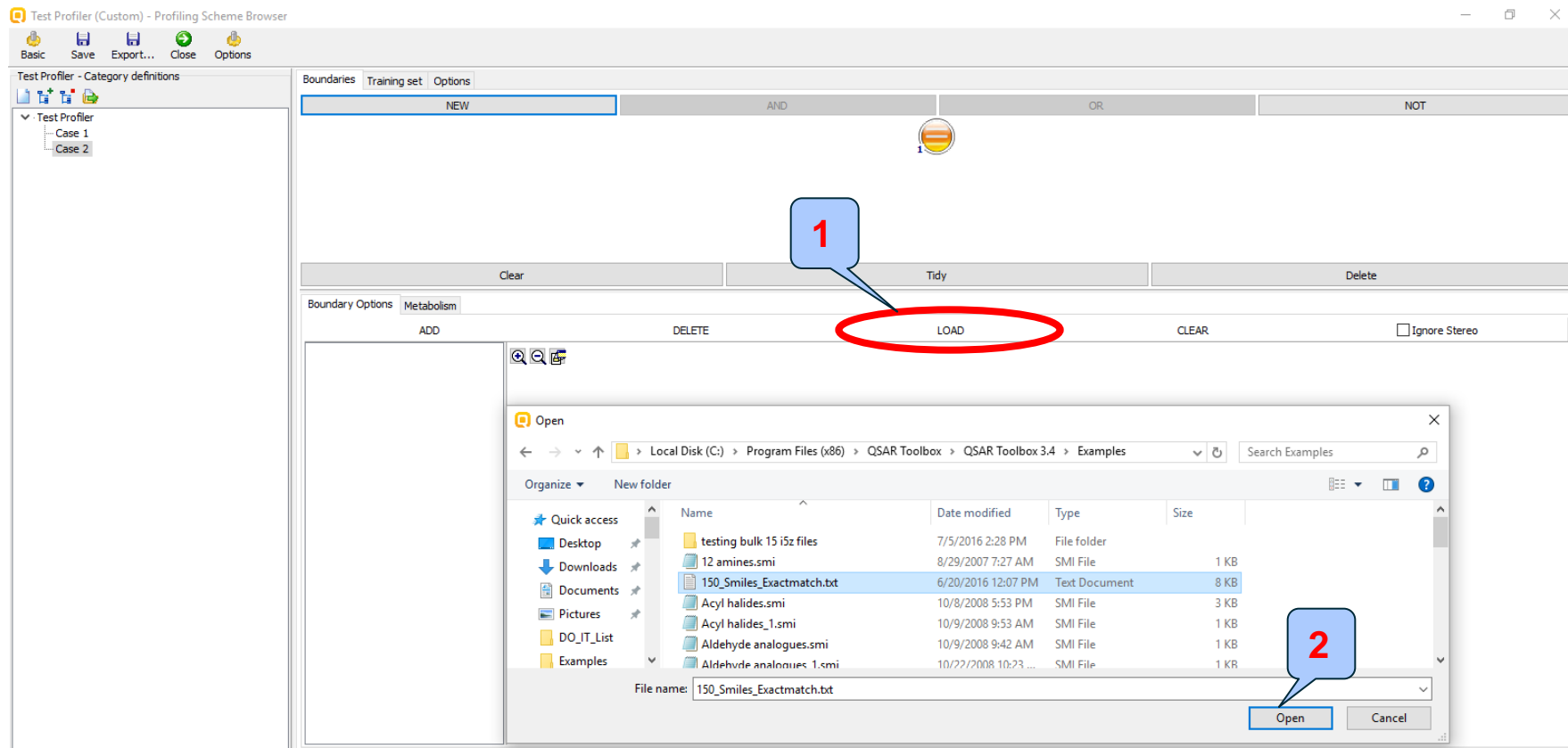
1. Create a new category "Case 2"; then **2.** from the list of possible boundary types **select** exactmatch boundary.

Building of a new profiler

Exactmatch boundary

- The exact match boundary presents search by list of structures (SMILES). If the target chemical belongs to the loaded in boundary list, it fulfills the category definition.
- It is possible to logically combine the exactmatch boundary with boundaries of different type.

Building of a new profiler exactmatch boundary



1. Click Load to add a pre-defined txt or smi file which contains a SMILES*; **2. Click Open.**
* The example file with 150 discrete SMILES could be found in the example folder of TB installation. It is placed in the directory with executable file.

Building of a new profiler exactmatch boundary

The screenshot shows the 'Test profiler (Custom) - Profiling Scheme Browser' application. The main area is divided into several sections:

- Top Bar:** Contains icons for 'Basic', 'Save', 'Export...', 'Close', and 'Options'.
- Left Panel:** 'Test profiler - Category definitions' showing a tree view with 'Test profiler' expanded to show 'Case 1' and 'Case 2'.
- Central Panel:** 'Boundaries' section with tabs for 'NEW', 'AND', 'OR', and 'NOT'. Below these are 'Clear', 'Tidy', and 'Delete' buttons.
- Bottom Panel:** 'Boundary Options' section with a tab for 'Metabolism'. It features an 'ADD' column containing a list of SMILES strings, and buttons for 'DELETE', 'LOAD', and 'CLEAR'. There is also an 'Ignore Stereo' checkbox.
- Chemical Structure:** A 2D chemical structure of acetic acid (CC(=O)O) is shown on the right side of the interface.
- Bottom Bar:** A 'Profile Comments' field.

Building of a new profiler

Final step – save the profiler

The screenshot shows the 'Test Profiler (Custom) - Profiling Scheme Browser' window. The 'Save' button in the top-left toolbar is highlighted with a red box and labeled '1'. A dialog box titled 'Toolbox 3' is open in the center, displaying the message: 'The profiling scheme has been saved to C:\Users\Nina\Documents\QSAR Toolbox\Ver 3.4\UserDir\Test Profiler.cfg', with the 'OK' button highlighted and labeled '2'. A callout '3' points to the close button in the top-right corner of the main window. In the bottom right, the chemical structure of acetic acid (CC(=O)O) is shown with a callout '2' pointing to it.

1. Press the Save button; 2. Click OK to confirm where to save the newly created profiler and 3. Close the window.

New profiler

The new profiler can now be applied to a target chemical or a list of chemicals complying with the newly constructed category definitions (see next screen shot).

New profiler Results

The screenshot displays the QSAR Toolbox software interface. The top navigation bar includes buttons for Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. The Profiling Schemes panel on the left lists various methods, with 'Test Profiler' checked under the 'Custom' category. A red arrow points from this method to the 'Filter endpoint tree...' panel on the right. In this panel, the 'Test Profiler' option is highlighted with a red box. The main workspace shows a chemical structure and a list of endpoints, with 'Case 1' selected under the 'Test Profiler' endpoint.

Congratulation

- You have used several new functions to create a new profiler for use with the Toolbox.
- Continual use of the Toolbox will increase your skills.