

OECD (Q)SAR Toolbox v.4.4.1

Step-by-step example for predicting skin sensitization
accounting for the abiotic activation of the chemicals

Outlook

- **Background**
- Keywords
- Objectives
- The exercise
- Workflow
- Save/Load

Background

- This is a step-by-step presentation designed to take the user through the Toolbox workflow for predicting skin sensitization potential of a target chemical taking into account its abiotic activation.

Outlook

- Background
- **Keywords**
- Objectives
- The exercise
- Workflow
- Save/Load

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

Outlook

- Background
- Keywords
- **Objectives**
- The exercise
- Workflow
- Save/Load

Objectives

This presentation demonstrates a number of functionalities of Toolbox 4.4:

- Define the target chemical and the target endpoint.
- Profiling the target chemical.
- Identifying analogues of the target chemical.
- Filling data gaps of the target chemical by read-across.
- Profiling the target chemical taking into account its (a)biotic activation.
- Identifying analogues of the target based on its (a)biotic activation
- Filling data gaps of the target chemical by read-across when the (a)biotic activation is taken into account.

Outlook

- Background
- Keywords
- Objectives
- **The exercise**
- Workflow
- Save/Load

The Exercise

- In this exercise our target chemical will be **Eugenol [CAS# 97-53-0]**.
- We will predict the Skin sensitization potential of Eugenol.
- Two types of categorization will be applied:
 - Identifying analogues by using well-known categorization group.
 - Identifying analogues based on the autoxidation activation of the target.

Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
- Save/Load

Workflow

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

Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
 - **Input**
 - **Input chemical by CAS#**
 - **Define target endpoint**

Input Overview

- This module provides the user with several means of entering the chemical of interest (i.e. the target chemical).
- Since all subsequent functions are based on chemical structure, the goal here is to make sure the molecular structure assigned to the target chemical is the correct one.

Input

Input chemical(s)

Alternative ways to input chemical(s):

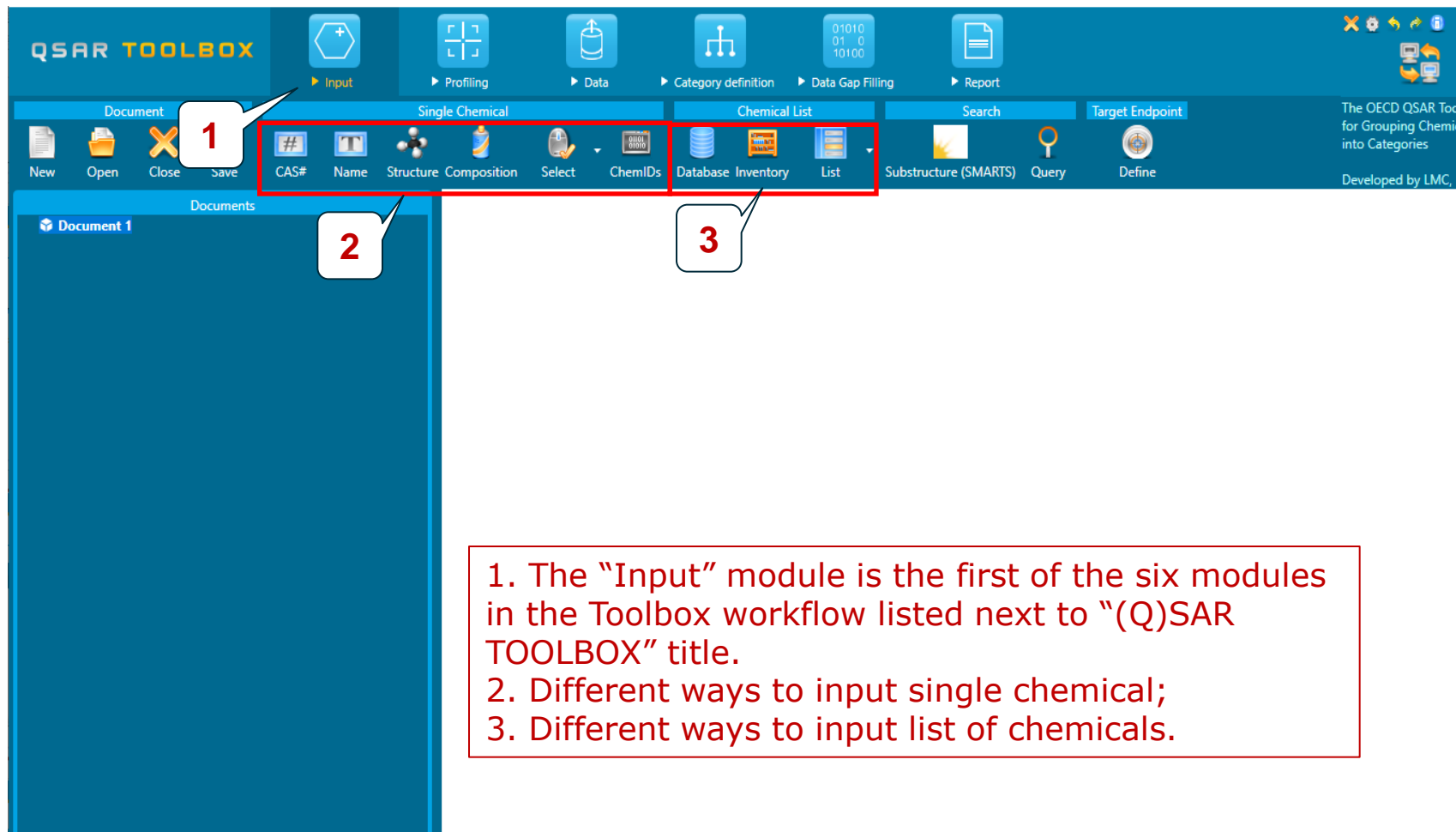
A. Single target chemical

- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- SMILES (simplified molecular information line entry system) notation/InChi
- Drawing chemical structure
- Select from User List/Inventory/Databases

B. Group of chemicals

- User List/Inventory
- Specialized Databases

Input Input Screen



1. The "Input" module is the first of the six modules in the Toolbox workflow listed next to "(Q)SAR TOOLBOX" title.
2. Different ways to input single chemical;
3. Different ways to input list of chemicals.

Input

Input target chemical by CAS#

1. Click on the **CAS#** button;

2. Enter CAS# **97-53-0** in the blank field;

3. Click **Search** button;

4. Click **OK**.

1	CAS	97-53-0
	SMILES	COc1cc(CC=C)ccc1O
	CS Relation	High
	Substance	Mono constituent
	Composition	[checked]
	Name	eugenol (4-allyl-2-methoxyphenol)
	Sources	NICNAS, Canada DSI

Chemical structure of eugenol (4-allyl-2-methoxyphenol):

COc1cc(CC=C)ccc1O

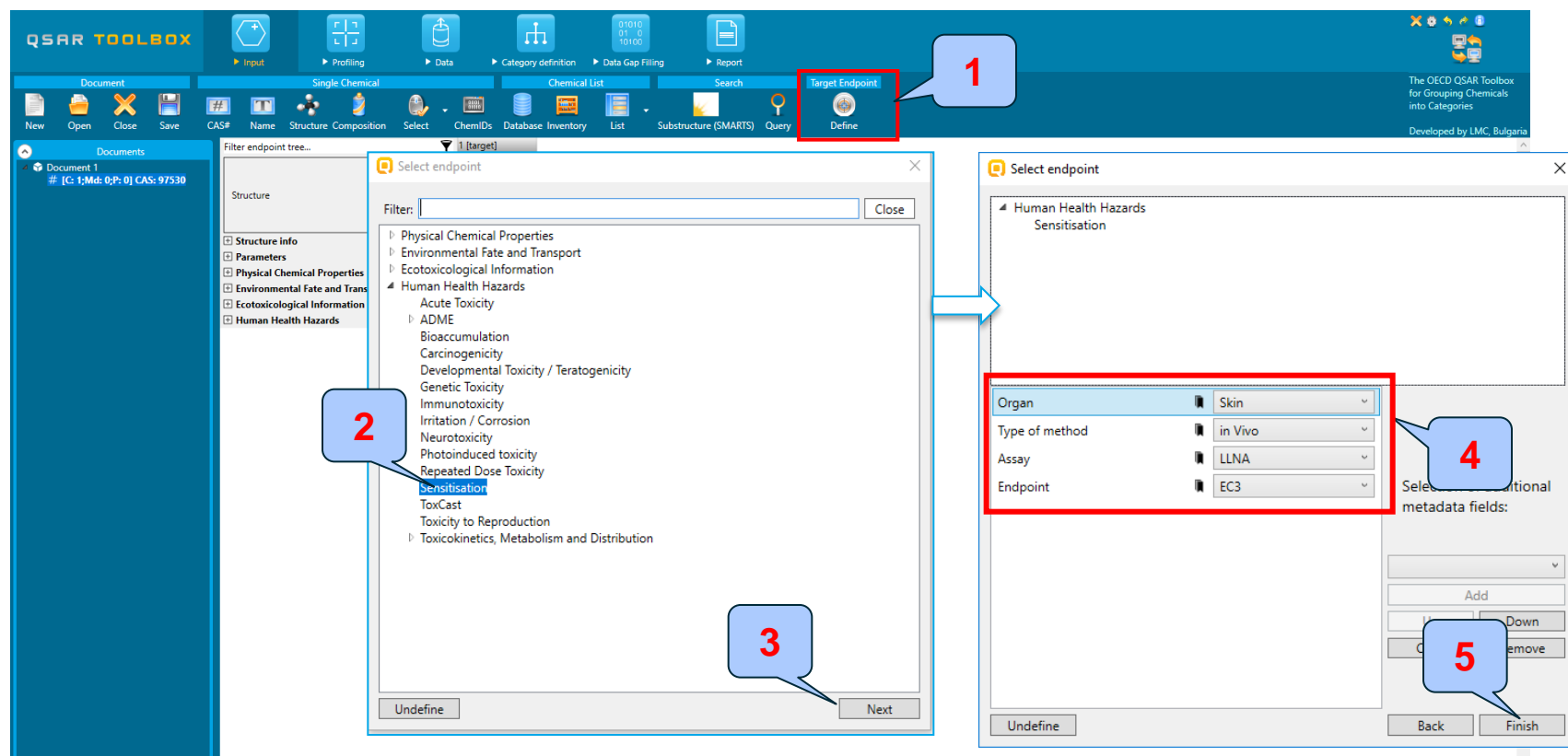
Input

Define target endpoint - overview

- The Define target endpoint functionality allows entering the endpoint of interest e.g., EC3, LC50, gene mutation etc.
- The relevant profiles and databases are become highlighted in color once the targeted endpoint is preliminary defined by this functionality (in green – the most suitable, in orange – the plausible);
- Calculation of alert performance (AP) is only possible if the target endpoint is preliminary defined (will be shown later on);
- There are different ways for defining the target endpoint (via the button from the Input module or by right click from the endpoint tree). For more details press F1 button in order to see the online help.

Input

Define target endpoint

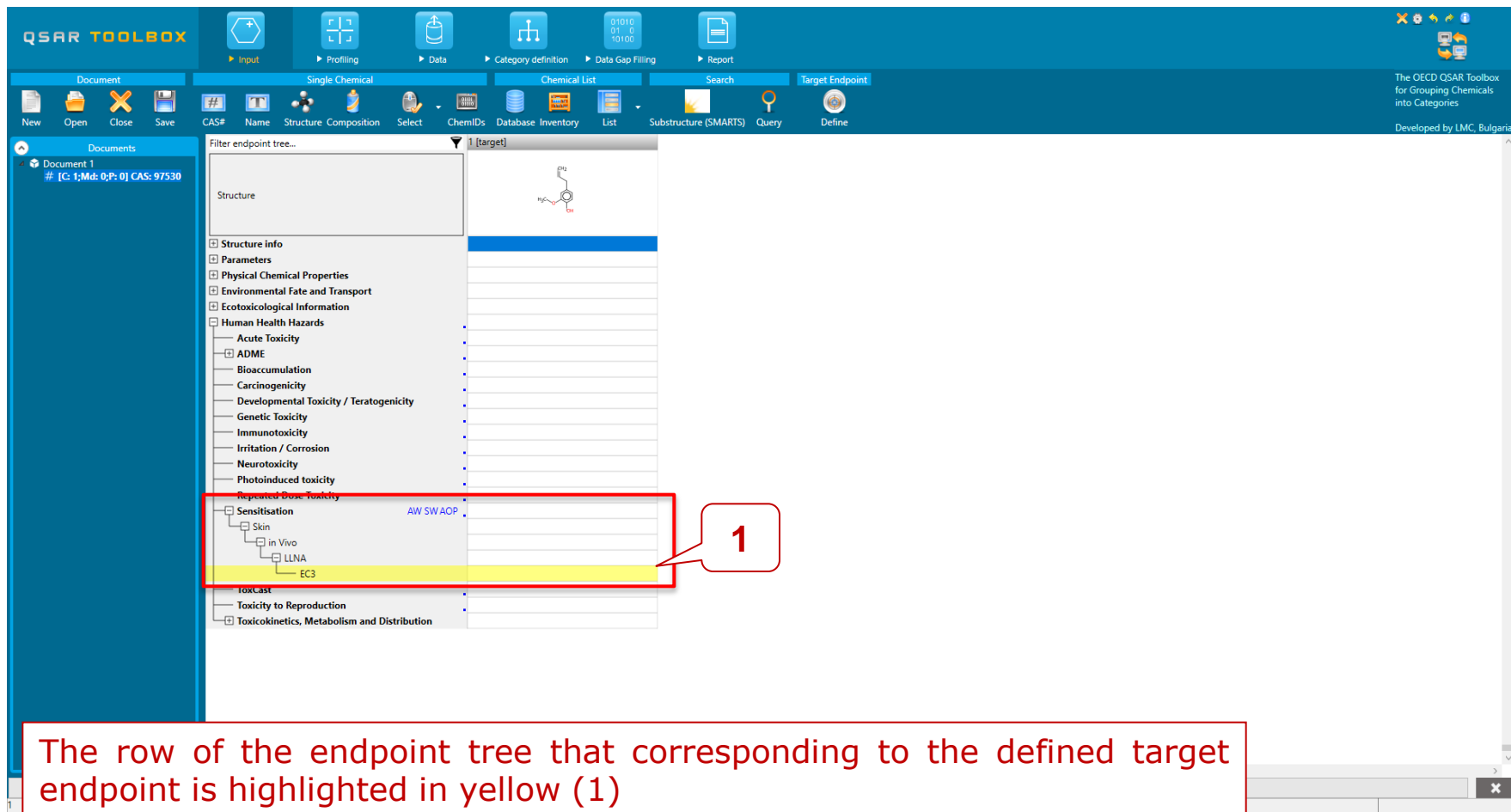


Click on the **Define** button (1). In the new window select the general endpoint "*Sensitization*" (2) and click on **Next** (3).

Specify the endpoint using the drop-down menus as follows: For *Endpoint* select **EC3**; for *Assay* - **LLNA**; for *Type of method* - **in Vivo**; for *Organ* - "**Skin**" (4). Finally click on **Finish** (5).

Input

Define target endpoint



The screenshot displays the QSAR Toolbox software interface. The 'Define' tab is active, showing a 'Filter endpoint tree...' dialog. The tree lists various endpoints, with 'EC3' highlighted in yellow. A red box with the number '1' points to this row. The chemical structure of the target endpoint is shown in the top right of the dialog.

The row of the endpoint tree that corresponding to the defined target endpoint is highlighted in yellow (1)

Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
 - Input
 - **Profiling**

Profiling Overview

- “Profiling” module contains all the knowledge in the system coded in profiling schemes (profilers);
- “Profilers” are a collection of empirical and mechanism knowledge (expertly derived) 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 profiling schemes;
- 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 once are getting colour highlighted*
- For most of the profilers, background information can be retrieved by selection of a profilers (for example, Protein binding by OASIS) and then click on “View” (see the next slide).

*For more details regarding relevancy of the profilers see ppt: Example for predicting skin sensitization taking into account alert performance

Profiling

Profiler's background information

The screenshot shows the QSAR Toolbox Profiling module. The interface is divided into several panels:

- Left Panel (Profiling methods):** Contains a list of profiling methods. Callout 2 points to "Protein binding by OASIS" under the "Suitable" category.
- Top Panel (Profiling Scheme Browser):** Shows a list of categories. Callout 1 points to the "Profiling" button. Callout 3 points to the "View" button. Callout 4 points to the "Aldehydes" category under "Activated Carbonyl compounds".
- Right Panel (Category tree):** Shows a hierarchical tree of categories. Callout 1 points to the "Aldehydes" category.
- Bottom Panel (Query details):** Shows the structure query and SMARTS string. Callout 1 points to the "Structure Query" tab.

Annotations include:

- List with categories:** A red box highlighting the list of categories in the Profiling Scheme Browser.
- Structural boundaries:** A red circle highlighting the "Aldehydes" category in the Category tree.
- Structural fragment:** A red circle highlighting a chemical structure fragment in the Query details panel.

1. Go to **Profiling** module;
 2. Select the profiler "*Protein binding by OASIS*";
 3. Click **View**;
 4. Select the **Aldehydes** category.

Profiling

Profiler's background information

The screenshot displays the QSAR Toolbox Profiling Scheme Browser. On the left, the 'Documents' pane shows 'Document 1' with CAS 97530. Below it, 'Profiling methods' and 'Metabolism/Transformations' are listed with various sub-options. The main 'Categories' pane on the right shows a tree structure under 'Protein binding by OASIS', with 'Aldehydes' highlighted under 'Schiff base formation'. A red box highlights 'Aldehydes' in the list. A callout bubble with the number '1' points to the 'Literature' tab in the top navigation bar. The 'Literature' tab is active, showing detailed information for 'Aldehydes'.

Protein binding by OASIS (General Mechanistic) - Profiling Scheme Browser

Save Scheme | Export Scheme | Save Tests | View Tests | Run All Tests

Filter:

Categories

- Protein binding by OASIS
 - Acylation
 - Ionic interaction
 - Michael addition
 - Nucleophilic addition
 - Radical reactions
 - Schiff base formation
 - Benzoyl Schiff base formation
 - Benzoyl phosphine oxides
 - Direct acting Schiff base formers
 - 1,2-Dicarbonyls and 1,3-Dicarbonyls
 - Di-substituted alpha,beta-unsaturated
 - Schiff base formation with carbonyl compounds
 - Activated Carbonyl compounds
 - Aldehydes**
 - alpha-Ketoesters
 - Aromatic carbonyl compounds
 - Bis aldehydes
 - Schiff base on pyrazolones and pyrazolidinones
 - Pyrazolones and Pyrazolidinones

Mechanistic Domain: Schiff base formation

Mechanistic Alert: Schiff base formation with carbonyl compounds

Structural Alert: Aldehydes

The chemical causes skin sensitization effect as a result of Schiff base formation with aldehydes:

$$\text{R}-\text{C}(=\text{O})-\text{H} \xrightarrow{\text{Pr}-\text{NH}_2} \text{R}-\text{C}(\text{H})=\text{N}-\text{Pr}$$

R = H, alkyl

Aldehydes are highly reactive molecules, many of which are strong sensitizers and their direct conjugation to protein nucleophiles is thought to be responsible. Simple aldehydes react readily with the amino groups of lysine residues on proteins to form imines or Schiff bases.

1. Go to the **Literature** tab in order to see detailed description of highlighted category (in this case "Aldehydes")

Profiling

Profiling the target chemical

- The outcome of the profiling determines the most appropriate way to search for analogues
- An option for coloring the profilers related to a given endpoint is implemented.
- The following profiling schemes are defined as relevant to the **Skin sensitization**:
 - *Protein binding by OASIS* – general mechanistic
 - *Protein binding alerts for skin sensitization according to GHS* – endpoint specific
 - *Protein binding alerts for skin sensitization by OASIS* - endpoint specific

Profiling

Profiling the target chemical

1. Check all profilers highlighted in green;

2. Click **Apply**.

Profiling

Profiling the target chemical

- The actual profiling will take up to several seconds depending on the number and type of profilers selected.
- The profiling result automatically appear as a last level of the endpoint tree (see next screenshot).
- In this case no protein binding alert (PBA) has been identified in the target substance (eugenol).

Profiling

Profiling the target chemical

QSAR TOOLBOX

Profiling Custom profile

Apply View New Delete

Document 1
[C: 1-Md: 0-P: 0] CAS: 97530

Profiling methods 3 Selected

Options f Select All Unselect All Invert About Options

- ☒ Suitable
 - ☒ Protein binding alerts for skin sensitization according to GHS
 - ☒ Protein binding alerts for skin sensitization by OASIS
 - ☒ Protein binding by OASIS
- ☒ Plausible
 - ☐ Aquatic toxicity classification by ECOSAR
 - ☐ Chemical elements
 - ☐ Groups of elements
 - ☐ Keratinocyte gene expression
 - ☐ Lipinski Rule Oass
 - ☐ OECD HPV Chemical Categories
 - ☐ Organic functional groups

Metabolism/Transformations 0 Selected

Options f Select All Unselect All Invert

- ☒ Suitable
 - ☐ Autooxidation simulator
 - ☐ Skin metabolism simulator
- ☒ Plausible
 - ☐ Autooxidation simulator (alkaline medium)
 - ☐ Dissociation simulator
 - ☐ Hydrolysis simulator (neutral)
- ☒ Unclassified
 - ☐ Hydrolysis simulator (acidic)
 - ☐ Hydrolysis simulator (basic)
 - ☐ in vivo Rat metabolism simulator
 - ☐ Microbial metabolism simulator

Filter endpoint tree... 1 [target]

Structure

Structure info

Parameters

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information

Human Health Hazards

- Acute Toxicity
- ADME
- Bioaccumulation
- Carcinogenicity
- Developmental Toxicity / Teratogenicity
- Genetic Toxicity
- Immunotoxicity
- Irritation / Corrosion
- Neurotoxicity
- Photoinduced toxicity
- Repeated Dose Toxicity
- Sensitisation
 - Skin
 - in Vivo
 - LLNA
 - EC3

ToxCast

Toxicity to Reproduction

Toxicokinetics, Metabolism and Distribution

Profiling

- General Mechanistic
 - Protein binding by OASIS
- Endpoint Specific
 - Protein binding alerts for skin sensitization... No alert found
 - Protein binding alerts for skin sensitization... No alert found

Profiling results coming from the selected suitable (green) profilers. No PBAs are identified in the target molecule.

Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - **Data**

Data Overview

- “Data” refers to the electronic process of retrieving the environmental fate, ecotoxicity and toxicity data that are stored in the Toolbox.
- Data gathering can be executed in a global fashion (i.e., collecting all data for all endpoints) or on a more narrowly defined basis (e.g., collecting data for a single or limited number of endpoints).

Data

Case study

- In this example, we limit our data gathering to a single toxicity endpoint (skin sensitization).
- In case of defined endpoint, the databases containing data for this endpoint, will be highlighted in green. The user could select all or just some of them.

Data

Gather data

The screenshot shows the QSAR Toolbox interface. The top toolbar has a 'Data' icon circled in red and labeled with a blue callout '1'. Below the toolbar, the 'Data' tab is selected, and the 'Gather' button is circled in red and labeled with a blue callout '3'. On the left, the 'Databases' list shows 'REACH Skin sensitisation database (normalised)' and 'Skin Sensitization' circled in red, with a blue callout '2' pointing to them. In the center, the 'Filter endpoint tree...' window shows 'Sensitization' selected, with a blue callout '5' pointing to it. On the right, the 'Read data?' dialog box is open, showing 'Choose...' selected under 'Read data?', with a blue callout '4' pointing to it. The 'Sensitization' checkbox is checked, with a blue callout '6' pointing to the 'OK' button.

1. Go to the **Data** module;

2. Select only REACH Skin sensitization database (normalized) and Skin sensitization databases;

3. Click **Gather**;

4. Select to **Choose** endpoint in the pop-up menu;

5. Select **Sensitization**;

6. Confirm by **OK**.

Data

Gather data

- Toxicity information on the target chemical is electronically collected from the selected dataset(s).
- It should be kept in mind that the search for data and analogues is performed only among the chemicals which are listed in the selected databases, which in this example are **Skin sensitization** and **REACH Skin sensitisation database (normalised)**.
- In this example, Positive experimental data is available for the target chemical (see next screen shots)

Data

Available experimental data

The screenshot displays the QSAR TOOLBOX software interface. The top menu bar includes options like Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below this, there are buttons for Data, Import, Export, and Delete. The main window is divided into several panes:

- Documents:** A list of documents, with '2 Selected' indicated.
- Databases:** A list of databases, including 'REACH Skin sensitisation database (normalised)' and 'Skin Sensitization', both of which are checked.
- Filter endpoint tree...:** A tree view showing various endpoints. The 'Sensitisation' endpoint is expanded, showing sub-endpoints like 'Skin', 'in Vivo', 'GPMT', 'HRIPT', 'LLNA', and 'EC3'. The 'EC3' endpoint is highlighted in yellow.
- Target Chemical:** A chemical structure is shown on the right.
- Experimental Data:** A table of experimental data is displayed, showing results for the 'EC3' endpoint. The data includes '1/22' and 'M: 5.4 %'.

A red dashed box highlights the 'Sensitisation' endpoint and its sub-endpoints, with a red arrow pointing to the 'EC3' endpoint. A red text box with an arrow pointing to the 'EC3' endpoint contains the text: "Positive experimental data is available for the target chemical."

Recap

- In the first module (*Input*) we:
 - entered the target chemical by CAS;
 - selected the chemical with High reliability of the CAS/SMILES correlation;
 - defined the target endpoint.
- In the second module (*Profiling*) we:
 - selected the profilers related to the target endpoint (highlighted in green);
 - saw that there is no PBA in target molecule.
- In the third module (*Data*), we:
 - selected two databases containing data for the defined target endpoint (*Skin sensitization* and *REACH Skin sensitisation database (normalised)*)
 - saw that positive skin sensitization data is available for the target chemical.

Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - **Category definition**

Category Definition

Grouping methods

- The different grouping methods allow the user to group chemicals into chemical categories according to different measures of “similarity” so that within a category data gaps can be filled by read-across.

Category Definition

Suitable Categorization/Assessment Phases

Suitable categorization phases:

1. Structure-related profilers
2. Endpoint specific profilers (for sub-categorization)
3. Additional structure-related profilers, if needed to eliminate dissimilar chemicals (to increase the consistency of category) (e.g. chemical elements)

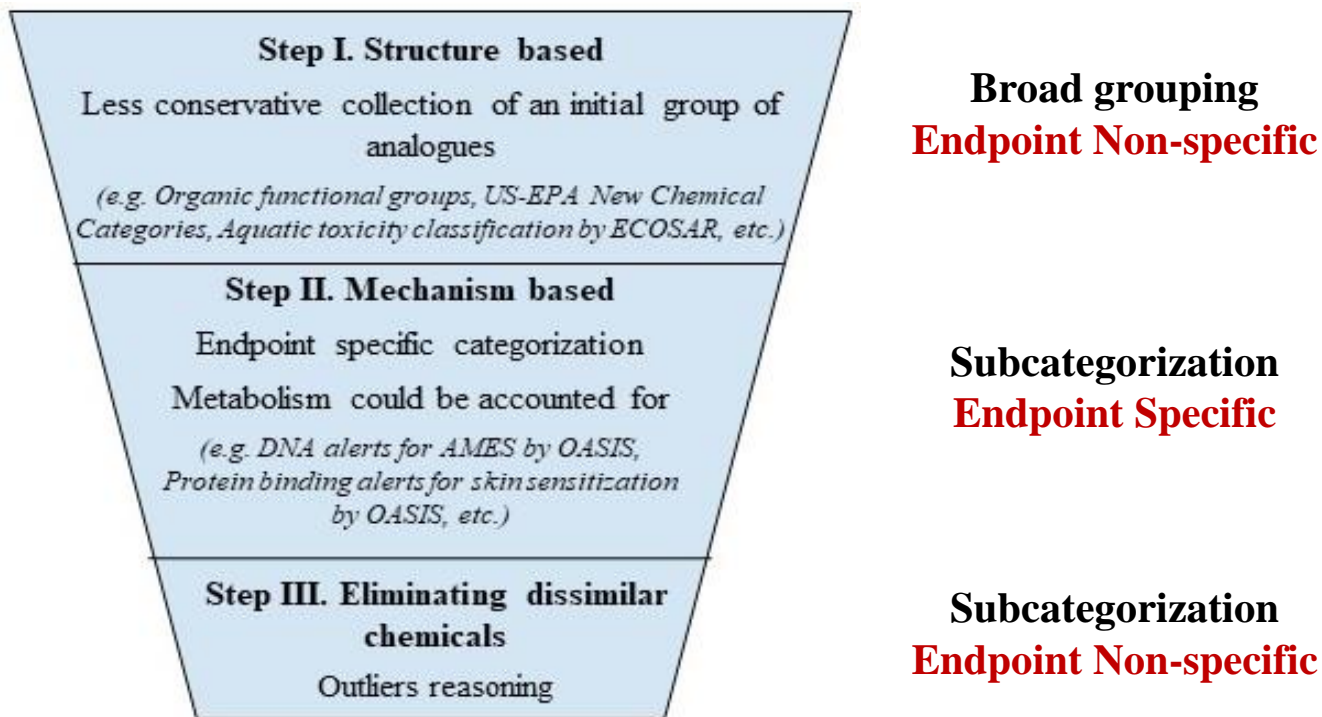
Performing categorization:

1. The categorization phases should be applied successively
2. The application order of the phases depend on the specificity of the data gap filling
3. More categories of same Phase could be used in forming categories
4. Some of the phases could be skipped if consistency of category members is reached

Graphical illustration of the suitable categorization phases is shown on next slide

Category Definition

Suitable Categorization/Assessment Phases



Note: As long as an acceptable level of structural and mechanistic similarity is achieved, it is not mandatory to follow all the stages described in the order given above; they can be executed differently or even skipped.

For example, in case the target chemical or any of its metabolites interact with biomacromolecules via a clearly defined mechanism relevant for the endpoint to predict, Stage I could be skipped and Stage II to be used for the primary categorization step.

Category Definition

Suitable Categorization/Assessment Phases

- In the current example we will define a broad group of analogues using the *Organic functional groups* (OFG) profiler (step I).
- Due to the variety of OFG identifies in the target chemical, only more commonly defined functional groups (i.e. allyl, phenol, ether) will be used for searching analogues.

Category Definition

Define category by OFG

The screenshot illustrates the steps to define a category by Organic Functional Groups (OFG) in the QSAR TOOLBOX. The interface shows the 'Category definition' module selected in the top navigation bar. The 'Organic functional groups' list on the left contains various categories, with 'Alkene moiety' through 'Precursors quinoid compounds' highlighted in blue. The 'Define' button is highlighted with a red box. The 'Grouping options (Organic functional groups)' dialog box is open, showing the 'Target categories' list with 'Allyl', 'Ether moiety', and 'Phenol' selected. The 'Options' section shows 'All categories' and 'Combine profiles' settings. The 'OK' button is highlighted with a red box.

1. Go to **Category definition** module
2. Select **Organic functional groups (OFG)**;
3. Click **Define**
4. Remove all functional groups (highlighted in blue) except *Allyl*, *Ether moiety* and *Phenol* by double click one by one or press CTRL button of the keyboard, mark the undesired groups and click the **Down** button;
5. Finally click **OK**.

Category Definition

Define category by OFG

The screenshot displays the QSAR Toolbox software interface during the 'Category definition' process. The main window shows a tree of 'Organic functional groups' on the left and a grid of chemical structures on the right. Four numbered callouts highlight key steps in the workflow:

- Warning message:** A dialog box asking 'You have selected different from target categories! Do you want to continue?' with 'Yes' and 'No' buttons.
- Grouping results:** A dialog box stating '11 chemical(s) found.' with an 'OK' button.
- Read data?:** A dialog box with 'All endpoints' selected and 'Choose...' as an option, with an 'OK' button.
- Gather data:** A dialog box stating '79 points added across 11 chemicals.' with an 'OK' button.

1. Warning message is displayed informing for changes in the set of profiling categories (OFG in this case) identified in the target chemical. Click **Yes**.
2. 11 chemicals are identified in the selected databases. Click **OK**.
3. Click **OK** to collect all available data.
4. Experimental data results (79 data points) are available for all chemicals in the category (11 chemicals). Click **OK**.

Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - Category definition
 - **Data gap filling**

Data Gap Filling Overview

- “Data Gap Filling” module provides three different data gap filling tools:
 - Read-across
 - Trend analysis
 - (Q)SAR models
- The most relevant data gap mechanism should be used by 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.
- In this example, we will use read-across.

Data Gap Filling

Apply Read-across

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Data Gap Filling', which is highlighted with a red box and a callout '3'. The 'Data Gap Filling' dropdown menu is open, showing 'Read-across' as the selected option. The central data table shows a grid of chemical structures and their associated data. A cell in the table is highlighted with a red box and a callout '2'. The right sidebar shows a 'Possible data inconsistency' dialog box with a callout '4' pointing to the 'OK' button. The dialog box contains metadata information and a list of scale/unit options for skin sensitization.

1. Go to **Data Gap Filling**
 2. Click the cell corresponding to the target endpoint and target chemical;
 3. Go to Data Gap filling and select **Read-across**;
 4. A pop-up window informing about possible data inconsistency appears. Confirm the default scale with **OK**.
- More details about scale definitions is provided on next slide.

Data Gap Filling

Sidebar on the data scales

- The available data in the databases could be represented in different scales depending on the database` donators.
- The skin sensitisation data could be: 1) categorical (for example: positive; negative; weak sensitizer; strong sensitizer, etc.) or 2) numerical (e.g. EC3 values in percentage).
- The main purpose of the scales is to unify all data available in the Toolbox databases for a certain endpoint by making conversions between them.
- The default scale for Skin Sensitisation is "*Skin Sensitisation ECETOC*". It converts all skin data into dichotomous scale: Positive/Negative.

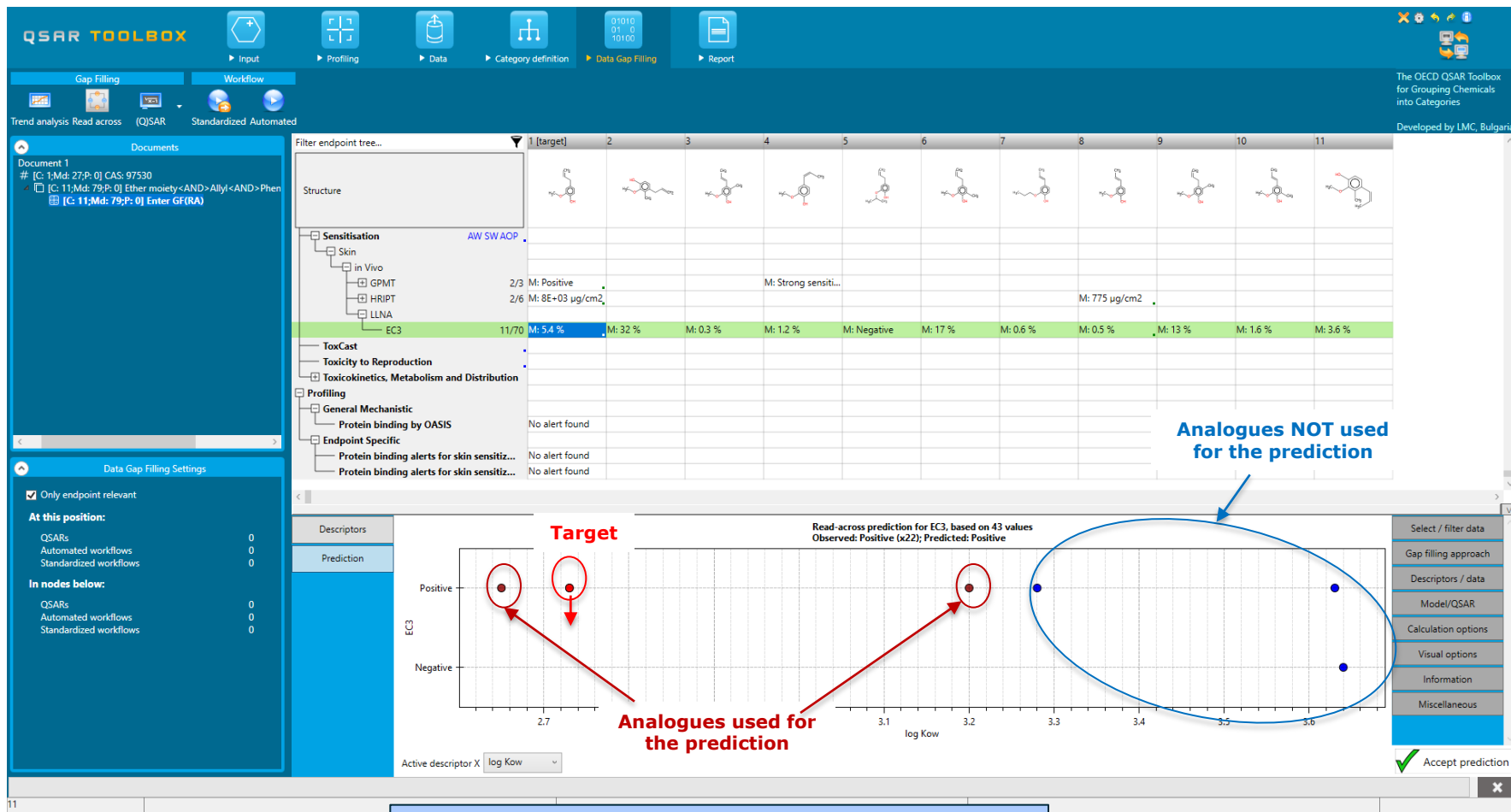
Data Gap Filling

Data Gap Filling graph

- The resulting graph is experimental results of all analogues (Y axis) according to a descriptor (X axis) with the default descriptor of log Kow (see next screen shot).
- The **RED** dot represents predicted result for the target chemical;
- The **BROWN** dots represent the experimental results available for the analogues that are used for the read-across. By default the closest five analogues are used for the prediction;
- The **BLUE** dot represents the experimental results available for the analogues that are not used for read-across.

Note: In some case one brown/blue point could correspond to more than one chemical. This holds for structures (e.g. different isomers) having the same X axis descriptor value (e.g. logKow) and the same Y axis descriptor value (e.g. positive).

Data Gap Filling Read-across



Initial graph without any subcategorization.

Data Gap Filling

Subcategorizations

- The category could be further narrowed-down by using of some of the endpoint-specific or structure-based profilers. This aim to reduce the uncertainty among the category members.
- The following profilers will be used for subcategorization:
 - 1) *Organic functional group* – step I of the categorization process (see slide 36) is repeated in order to eliminate multifunctional analogues
 - 2) *Protein binding alerts for skin sensitization by OASIS*
 - 3) *Protein binding alerts for skin sensitization by OASIS* combined with *Autoxidation simulator*
- These steps are summarized in the next screen shots. Before subcategorizations maximal values are taken into account

Data Gap Filling Subcategorizations

1. Open *Calculation options* > *Data usage*;

2. Choose *Maximal value*;

3. Click **OK.**

The screenshot shows the QSAR Toolbox interface. The top menu bar includes 'Gap Filling', 'Workflow', 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The left sidebar has 'Documents' and 'Data Gap Filling Settings'. The central workspace shows a 'Filter endpoint tree' with a table of endpoints and their predicted values. The 'Read-across prediction' plot shows observed vs. predicted values for EC3. The 'Choose one' dialog box is open, showing the 'Maximal' option selected. The 'Calculation options' panel on the right shows 'Data usage' selected.

Data Gap Filling

Subcategorization 1: Organic functional groups

The screenshot displays the 'Subcategorization' window in the QSAR Toolbox. The interface is divided into several sections:

- Options:** Includes 'Select All', 'Unselect All', 'Invert', and 'About' buttons. The 'Organic functional groups' category is selected under the 'Empiric' section.
- Profiling results of target:** A list of chemical moieties including Alkene moiety, Alkenyl (hetero)arenes, Alkoxy moiety, Alkyl-, alkenyl- and alkynyl (hetero)arenes, Aryl, Ether moiety, Phenol, and Precursors quinoid compounds.
- Profiling results of analogues:** A table showing the distribution of chemical moieties across different categories. The table has columns for '1 [target]', '2', '3', '4', '5', '6', '7', '8', and '9'. The 'M: 5.4 %' value is highlighted for the '1 [target]' column.
- Read-across prediction for EC3, based on 5 values:** A scatter plot showing the relationship between log Kow (x-axis) and EC3 (y-axis). The plot includes a legend for 'Positive' and 'Negative' values. A chemical structure is highlighted as 'Chemical dissimilar to the target with respect to OFG'.
- Select / filter data:** A panel on the right with buttons for 'Subcategorize', 'Mark chemicals by WS', 'Mark chemicals by descriptor value', 'Filter points by test conditions', 'Mark focused chemical', 'Mark focused points', and 'Remove marked data'. The 'Subcategorize' button is circled in red.

Three numbered callouts indicate the workflow steps:

1. Open **Select/filter data>>Subcategorize**;
2. Select **Organic functional groups**.
3. Click **Remove selected** to eliminate the dissimilar chemical.

Data Gap Filling

Subcategorization 2: Protein binding alerts for skin sensitization by OASIS

Protein binding alerts are not identified in the target

Protein binding alerts are not identified in the analogues

No protein binding alerts are identified in target neither in the analogues, which can not explained the positive experimental data found. In this respect metabolism should be taken into account (see next slide).

Data Gap Filling

Subcategorization 3: Protein binding alerts for skin sensitization by OASIS combined with Autoxidation simulator

1 Select *Protein binding alerts for skin sensitization by OASIS*;

2 Select *Autoxidation simulator*;

3 Close the *Subcategorization* window;

4 Accept the prediction.

The metabolites of target chemical and its analogues possess the same protein binding alerts. This could explain the positive experimental data and respectively the positive read-across prediction.

Read-across prediction for EC3, based on 5 values
Observed: Positive (x22); Predicted: Positive

Data Gap Filling

Interpreting Read-across

- In this example the target and all analogues have no protein binding alerts as parents.
- The target and all analogues have the same protein binding alerts when autoxidation is taken into account.
- The latter could explain the positive experimental data of the target and the analogues.

Data Gap Filling

Next actions

- The study continues with a second scenario for predicting the skin sensitization potential of eugenol where a category with analogues will be defined accounting for the abiotic activation of the target.

Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
 - Input
 - **Profiling of (a)biotic products**

Profiling of (a)biotic products

1. Go back to your target by click on the CAS#*;

2. Select all suitable (green) profilers in section Profiling;

3. Select the Autoxidation simulator;

4. Click **Apply**.

5. Click **Yes** to confirm profiling of the generated autoxidation.

Note: Alternatively to going back to the upper level in the document tree, you can open a new document, repeat the steps described in slides 15 and 17 and then to continue with the current example.

Profiling of (a)biotic products

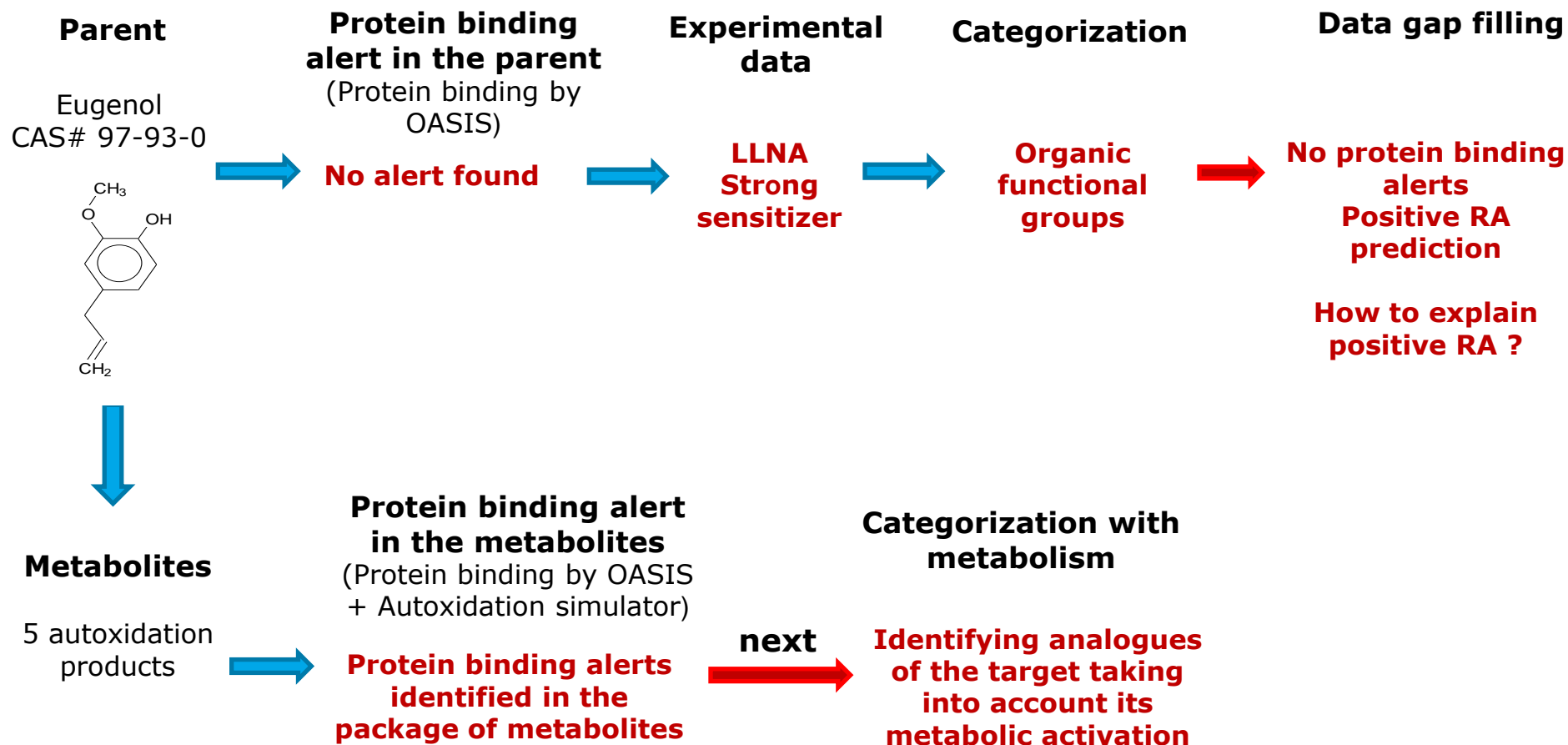
The screenshot displays the QSAR TOOLBOX software interface. The top menu bar includes 'Data', 'Import', 'Export', and 'Delete'. The left sidebar contains sections for 'Gather', 'Document', 'Options', and 'Inventories'. The main workspace shows a 'File' dialog box with five metabolites listed: 'metabolite #1' through 'metabolite #5', each with a chemical structure and 'No CAS number'. A 'Profiling results' window is open, showing a list of chemical reactions and their frequencies. Red arrows and callouts (1 and 2) highlight specific steps in the workflow.

1 Double click the cell with the autoxidation products to see them (5 metabolites are produced);

2 Double click the cell with profiling results identified in the metabolites.

Package of Protein profiling result for the parent and its autoxidation products

Recap



Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling of (a)biotic products
 - **Category definition based on (a)biotic transformations**

Category definition based on (a)biotic transformations

Define category using autoxidation simulator

All available transformation maps in Toolbox – documented and simulated can be used for primary grouping.

1. Go to new document and add CAS#97530;
2. Go to **Category definition** module;
3. Click **Define with metabolism**;
4. Select **Autoxidation simulator**;
5. The appeared button "About" allows you to read short description of the simulator;
6. Click **OK**.

Category definition based on (a)biotic transformations

Grouping options window

Grouping options (Autoxidation simulator)

☒ All queries ☐ At least one

Chemical	Query	Criteria
Parent <chem>CC1=CC=C(C=C1)C(=O)O</chem>	none	No criteria.
Metabolite 1 <chem>CC1=CC=C(C=C1)C(=O)O</chem>	none none Exact match Parametric Profile Similarity	No criteria.
Metabolite 2 <chem>CC1=CC=C(C=C1)C(=O)O</chem>	none	No criteria.
All chemicals		
Parent & Metabolites	none	No criteria.

Alert performance

Scales

Calculate

OK Cancel

Grouping options dialogue appears. It shows all the generated metabolites of the target chemical (use the scroll bar to see them). It has two subsections:

- (1) shows the parent and each of the generated metabolites. This allows defining different criteria for each structure when looking for analogues.
- (2) treats the parent and its metabolites as one, i.e. the criteria is provided for the whole package “parent and metabolites” but not for the individual metabolites.

A drop down menu (3) is available for each of the structures (in the column “Query”) which allow setting the type of criteria for looking for analogues.

Category definition based on (a)biotic transformations

Grouping options

Explanation of each option from the drop down menu is given below:

- **None** – default options; no criteria is set;
- **Exact** – search for analogues which metabolites have the exact structure of the target metabolite; only available for the metabolites and the package “parent + metabolites” but not for the parent chemical;
- **Parametric** – to have specific value or range of variation of defined parameter (a list with all parameters currently available in the Toolbox is provided);
- **Profile** – to have specific category by selected profiler (a list with all profilers is provided);
- **Structural** –allows structural similarity assessment based on the atom centered fragments.

The user can select a profiling, parametric or structural query for both target and its metabolites.

Category definition based on (a)biotic transformations

Grouping options

Grouping options (Autoxidation simulator)

☒ All queries ☐ At least one

Chemical	Query	Criteria
Parent <chem>CC(=O)O</chem>	none	No criteria.
Metabolite 1 <chem>CC(=O)O</chem>	none	No criteria.
Metabolite 2 <chem>CC(=O)O</chem>	none	No criteria.
All chemicals		
Parent & Metabolites	Profile	Profiler: Protein binding alerts for skin sensitization by OASIS Options: Edit

Alert performance
Scales
Calculate

OK Cancel

Options

Target categories
Michael Addition
Michael Addition >> Michael addition on quinoid type compounds
Michael Addition >> Michael addition on quinoid type compounds >> Quinone methide(s)/imines; Quinoid oxime structure; No alert found
Radical reactions
Radical reactions >> Free radical formation
Radical reactions >> Free radical formation >> Hydroperoxides
SN2
SN2 >> Ring opening SN2 reaction
SN2 >> Ring opening SN2 reaction >> Epoxides, Aziridines and Sulfuranes

Options
Down Up Reset Options

All categories
(N/A)
Acylation
Acylation >> (Thio)carbamylation of protein nucleophiles

Combine profiles
☒ AND ☐ OR
☐ Invert result
☐ Strict
☐ Sort results

OK Cancel

1. Select the **Profile** option for the package "parent & metabolites";
2. Select "Protein binding alerts for SS by OASIS" profile;
3. Click **Edit**. The profiling results of the parent structure and its metabolites are shown;
4. Click **OK** to confirm the defined search criteria;
5. Click **OK** in *Grouping options* window to execute the search.

Category definition based on (a)biotic transformations

Collecting data for the identified analogues

The screenshot shows the QSAR Toolbox software interface. The 'Category definition' workflow is active. The 'Filter endpoint tree' is open, displaying a list of endpoints. A 'Read data?' dialog box is shown with 'All endpoints' selected. A second dialog box confirms '31 points added across 5 chemicals.'

1 Click **OK** to collect all data available for them;

2 31 experimental data results are available for the five category members (1 target and 4 analogues). Click **OK**.

Category definition based on (a)biotic transformations

Profiling the identified analogues

1. Go to **Profiling**;

2. Check **Protein binding alerts for skin sensitization by OASIS**

3. Check **Autoxidation simulator**;

4. Click **Apply**;

5. Click **Yes**.

1. Right-click next to the *Protein binding alerts for skin sensitization by OASIS* under the Metabolism/Transformations level (in the grey field). Select **Profile statistic**.
2. Click on the first row to see the four chemicals having the exactly the same profiling results;
3. Close the *Profile statistic* window.

Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling of (a)biotic products
 - Category definition based on (a)biotic transformations
 - **Data Gap Filling**

Data gap filling

Apply Read across

- In the current example we will use the potency scale - *Skin sensitization EC3(ratio)* scale in order to see the final prediction in percentage (%).
- The analogues will be checked for presence of additional protein binding alerts in their parent structures or after activation. The following subcategorizations will be applied for the purpose:
 - *Protein binding alerts for skin sensitization by OASIS*
 - *Protein binding alerts for skin sensitization by OASIS combined with autoxidation simulator*

Data gap filling

Apply Read across

1. Go to **Data Gap Filling** module;

2. Click the cell corresponding to the target chemical and target endpoint;

3. Click **Read-across**;

4. Stay with the scale Skin sensitization EC3 (ratio). Click **OK**.

Data gap filling

Subcategorization 1: Protein binding alerts for skin sensitization by OASIS

1 Open **Select/filter data**;

2 Click **Subcategorize**;

3 Select **Protein binding alerts for skin sensitization by OASIS** profiler.

There are no protein binding alerts identified in the target chemical and its analogues

(3) No alert found

No alert found

Read-across prediction for EC3, based on 3 values
Observed: from 5.4 to 40.9 %; Predicted: 20.7 %

EC3 [%]

30

20

10

0

2.75 2.8 2.85 2.9 2.95 3.0 3.05 3.1 3.15 3.2 3.25 3.3

2.75 2.8 2.85 2.9 2.95 3.0 3.05 3.1 3.15 3.2 3.25 3.3

3.0

3.1

3.2

3.3

3.4

3.5

3.6

3.7

3.8

3.9

4.0

4.1

4.2

4.3

4.4

4.5

4.6

4.7

4.8

4.9

5.0

5.1

5.2

5.3

5.4

5.5

5.6

5.7

5.8

5.9

6.0

6.1

6.2

6.3

6.4

6.5

6.6

6.7

6.8

6.9

7.0

7.1

7.2

7.3

7.4

7.5

7.6

7.7

7.8

7.9

8.0

8.1

8.2

8.3

8.4

8.5

8.6

8.7

8.8

8.9

9.0

9.1

9.2

9.3

9.4

9.5

9.6

9.7

9.8

9.9

10.0

10.1

10.2

10.3

10.4

10.5

10.6

10.7

10.8

10.9

11.0

11.1

11.2

11.3

11.4

11.5

11.6

11.7

11.8

11.9

12.0

12.1

12.2

12.3

12.4

12.5

12.6

12.7

12.8

12.9

13.0

13.1

13.2

13.3

13.4

13.5

13.6

13.7

13.8

13.9

14.0

14.1

14.2

14.3

14.4

14.5

14.6

14.7

14.8

14.9

15.0

15.1

15.2

15.3

15.4

15.5

15.6

15.7

15.8

15.9

16.0

16.1

16.2

16.3

16.4

16.5

16.6

16.7

16.8

16.9

17.0

17.1

17.2

17.3

17.4

17.5

17.6

17.7

17.8

17.9

18.0

18.1

18.2

18.3

18.4

18.5

18.6

18.7

18.8

18.9

19.0

19.1

19.2

19.3

19.4

19.5

19.6

19.7

19.8

19.9

20.0

20.1

20.2

20.3

20.4

20.5

20.6

20.7

20.8

20.9

21.0

21.1

21.2

21.3

21.4

21.5

21.6

21.7

21.8

21.9

22.0

22.1

22.2

22.3

22.4

22.5

22.6

22.7

22.8

22.9

23.0

23.1

23.2

23.3

23.4

23.5

23.6

23.7

23.8

23.9

24.0

24.1

24.2

24.3

24.4

24.5

24.6

24.7

24.8

24.9

25.0

25.1

25.2

25.3

25.4

25.5

25.6

25.7

25.8

25.9

26.0

26.1

26.2

26.3

26.4

26.5

26.6

26.7

26.8

26.9

27.0

27.1

27.2

27.3

27.4

27.5

27.6

27.7

27.8

27.9

28.0

28.1

28.2

28.3

28.4

28.5

28.6

28.7

28.8

28.9

29.0

29.1

29.2

29.3

29.4

29.5

29.6

29.7

29.8

29.9

30.0

30.1

30.2

30.3

30.4

30.5

30.6

30.7

30.8

30.9

31.0

31.1

31.2

31.3

31.4

31.5

31.6

31.7

31.8

31.9

32.0

32.1

32.2

32.3

32.4

32.5

32.6

32.7

32.8

32.9

33.0

33.1

33.2

33.3

33.4

33.5

33.6

33.7

33.8

33.9

34.0

34.1

34.2

34.3

34.4

34.5

34.6

34.7

34.8

34.9

35.0

35.1

35.2

35.3

35.4

35.5

35.6

35.7

35.8

35.9

36.0

36.1

36.2

36.3

36.4

36.5

36.6

36.7

36.8

36.9

37.0

37.1

37.2

37.3

37.4

37.5

37.6

37.7

37.8

37.9

38.0

38.1

38.2

38.3

38.4

38.5

38.6

38.7

38.8

38.9

39.0

39.1

39.2

39.3

39.4

39.5

39.6

39.7

39.8

39.9

40.0

40.1

40.2

40.3

40.4

40.5

40.6

40.7

40.8

40.9

41.0

41.1

41.2

41.3

41.4

41.5

41.6

41.7

41.8

41.9

42.0

42.1

42.2

42.3

42.4

42.5

42.6

42.7

42.8

42.9

43.0

43.1

43.2

43.3

43.4

43.5

43.6

43.7

43.8

43.9

44.0

44.1

44.2

44.3

44.4

44.5

44.6

44.7

44.8

44.9

45.0

45.1

45.2

45.3

45.4

45.5

45.6

45.7

45.8

45.9

46.0

46.1

46.2

46.3

46.4

46.5

46.6

46.7

46.8

46.9

47.0

47.1

47.2

47.3

47.4

47.5

47.6

47.7

47.8

47.9

48.0

48.1

48.2

48.3

48.4

48.5

48.6

48.7

48.8

48.9

49.0

49.1

49.2

49.3

49.4

49.5

49.6

49.7

49.8

49.9

50.0

50.1

50.2

50.3

50.4

50.5

50.6

50.7

50.8

50.9

51.0

51.1

51.2

51.3

51.4

51.5

51.6

51.7

51.8

51.9

52.0

52.1

52.2

52.3

52.4

52.5

52.6

52.7

52.8

52.9

53.0

53.1

53.2

53.3

53.4

53.5

53.6

53.7

53.8

53.9

54.0

54.1

54.2

54.3

54.4

54.5

54.6

54.7

54.8

54.9

55.0

55.1

55.2

55.3

55.4

55.5

55.6

55.7

55.8

55.9

56.0

56.1

56.2

56.3

56.4

56.5

56.6

56.7

56.8

56.9

57.0

57.1

57.2

57.3

57.4

57.5

57.6

57.7

57.8

57.9

58.0

58.1

58.2

58.3

58.4

58.5

58.6

58.7

58.8

58.9

59.0

59.1

59.2

59.3

59.4

59.5

59.6

59.7

59.8

59.9

60.0

60.1

60.2

60.3

60.4

60.5

60.6

60.7

60.8

60.9

61.0

61.1

61.2

61.3

61.4

61.5

61.6

61.7

61.8

61.9

62.0

62.1

62.2

62.3

62.4

62.5

62.6

62.7

62.8

62.9

63.0

63.1

63.2

63.3

63.4

63.5

63.6

63.7

63.8

63.9

64.0

64.1

64.2

64.3

64.4

64.5

64.6

64.7

64.8

64.9

65.0

65.1

65.2

65.3

65.4

65.5

65.6

65.7

65.8

65.9

66.0

66.1

66.2

66.3

66.4

66.5

66.6

66.7

66.8

66.9

67.0

67.1

67.2

67.3

67.4

67.5

67.6

67.7

67.8

67.9

68.0

68.1

68.2

68.3

68.4

68.5

68.6

68.7

68.8

68.9

69.0

69.1

69.2

69.3

69.4

69.5

69.6

69.7

69.8

69.9

70.0

70.1

70.2

70.3

70.4

70.5

70.6

70.7

70.8

70.9

71.0

71.1

71.2

71.3

71.4

71.5

71.6

71.7

71.8

71.9

72.0

72.1

72.2

72.3

72.4

72.5

72.6

72.7

72.8

72.9

73.0

73.1

73.2

73.3

73.4

73.5

73.6

73.7

73.8

73.9

74.0

74.1

74.2

74.3

74.4

74.5

74.6

74.7

74.8

74.9

75.0

75.1

75.2

75.3

75.4

75.5

75.6

75.7

75.8

75.9

76.0

76.1

76.2

76.3

76.4

76.5

76.6

76.7

76.8

76.9

77.0

77.1

77.2

77.3

77.4

77.5

77.6

77.7

77.8

77.9

78.0

78.1

78.2

78.3

78.4

78.5

78.6

78.7

78.8

78.9

79.0

79.1

79.2

79.3

79.4

79.5

79.6

79.7

79.8

79.9

80.0

80.1

80.2

80.3

80.4

80.5

80.6

80.7

80.8

80.9

81.0

81.1

81.2

81.3

81.4

81.5

81.6

81.7

81.8

81.9

82.0

82.1

82.2

82.3

82.4

82.5

82.6

82.7

82.8

82.9

83.0

83.1

83.2

83.3

83.4

83.5

83.6

83.7

83.8

83.9

84.0

84.1

84.2

84.3

84.4

84.5

84.6

84.7

84.8

84.9

85.0

85.1

85.2

85.3

85.4

85.5

85.6

85.7

85.8

85.9

86.0

86.1

86.2

86.3

86.4

86.5

86.6

86.7

86.8

86.9

87.0

87.1

87.2

87.3

87.4

87.5

87.6

87.7

87.8

87.9

88.0

88.1

88.2

88.3

88.4

88.5

88.6

88.7

88.8

88.9

89.0

89.1

89.2

89.3

89.4

89.5

89.6

89.7

89.8

89.9

90.0

90.1

90.2

90.3

90.4

90.5

90.6

90.7

90.8

90.9

91.0

91.1

91.2

91.3

91.4

91.5

91.6

91.7

91.8

91.9

92.0

92.1

92.2

92.3

92.4

92.5

92.6

92.7

92.8

92.9

93.0

93.1

93.2

93.3

93.4

93.5

93.6

93.7

93.8

93.9

94.0

94.1

94.2

94.3

94.4

94.5

94.6

94.7

94.8

94.9

95.0

95.1

95.2

95.3

95.4

95.5

95.6

95.7

95.8

95.9

96.0

96.1

96.2

96.3

96.4

96.5

96.6

96.7

96.8

96.9

97.0

97.1

97.2

97.3

97.4

97.5

97.6

97.7

97.8

97.9

98.0

98.1

98.2

98.3

98.4

98.5

98.6

98.7

98.8

98.9

99.0

99.1

99.2

99.3

99.4

99.5

99.6

99.7

99.8

99.9

100.0

100.1

100.2

100.3

100.4

100.5

100.6

100.7

100.8

100.9

101.0

101.1

101.2

101.3

101.4

101.5

101.6

101.7

101.8

101.9

102.0

102.1

102.2

102.3

102.4

102.5

102.6

102.7

102.8

102.9

103.0

103.1

103.2

103.3

103.4

103.5

103.6

103.7

103.8

103.9

104.0

104.1

104.2

104.3

104.4

104.5

104.6

104.7

104.8

104.9

105.0

105.1

105.2

105.3

105.4

105.5

105.6

105.7

105.8

105.9

106.0

106.1

106.2

106.3

106.4

106.5

106.6

106.7

106.8

106.9

107.0

107.1

107.2

107.3

107.4

107.5

107.6

107.7

107.8

107.9

108.0

108.1

108.2

108.3

108.4

108.5

108.6

108.7

108.8

108.9

109.0

109.1

109.2

109.3

109.4

109.5

109.6

109.7

109.8

109.9

110.0

110.1

110.2

110.3

110.4

110.5

110.6

110.7

110.8

110.9

111.0

111.1

111.2

111.3

111.4

111.5

111.6

111.7

111.8

111.9

112.0

112.1

112.2

112.3

112.4

112.5

112.6

112.7

112.8

112.9

113.0

113.1

113.2

113.3

113.4

113.5

113.6

113.7

113.8

113.9

114.0

114.1

114.2

114.3

114.4

114.5

114.6

114.7

114.8

114.9

115.0

115.1

115.2

115.3

115.4

115.5

115.6

115.7

115.8

115.9

116.0

116.1

116.2

116.3

116.4

116.5

116.6

116.7

116.8

116.9

117.0

117.1

117.2

117.3

117.4

117.5

117.6

117.7

117.8

117.9

118.0

118.1

118.2

118.3

118.4

118.5

118.6

118.7

118.8

118.9

119.0

119.1

119.2

119.3

119.4

119.5

119.6

Data gap filling

Subcategorization 2: Protein binding alerts for skin sensitization by OASIS combined with Autoxidation simulator

1. Select **Protein binding alerts for skin sensitization by OASIS** profiler;

2. Select **Autoxidation simulator**;

3. Close the *Subcategorization* window.

The same protein binding alerts are identified in the target chemical and its analogues when their abiotic activation is taken into account.

Read-across prediction for EC3, based on 3 values
Observed: from 5.4 to 40.9 %; Predicted: 20.7 %

EC3 [%]

log Kow

Select / filter data

Subcategorize

Mark chemicals by WS

Mark chemicals by descriptor value

Filter points by test conditions

Mark focused chemical

Mark focused points

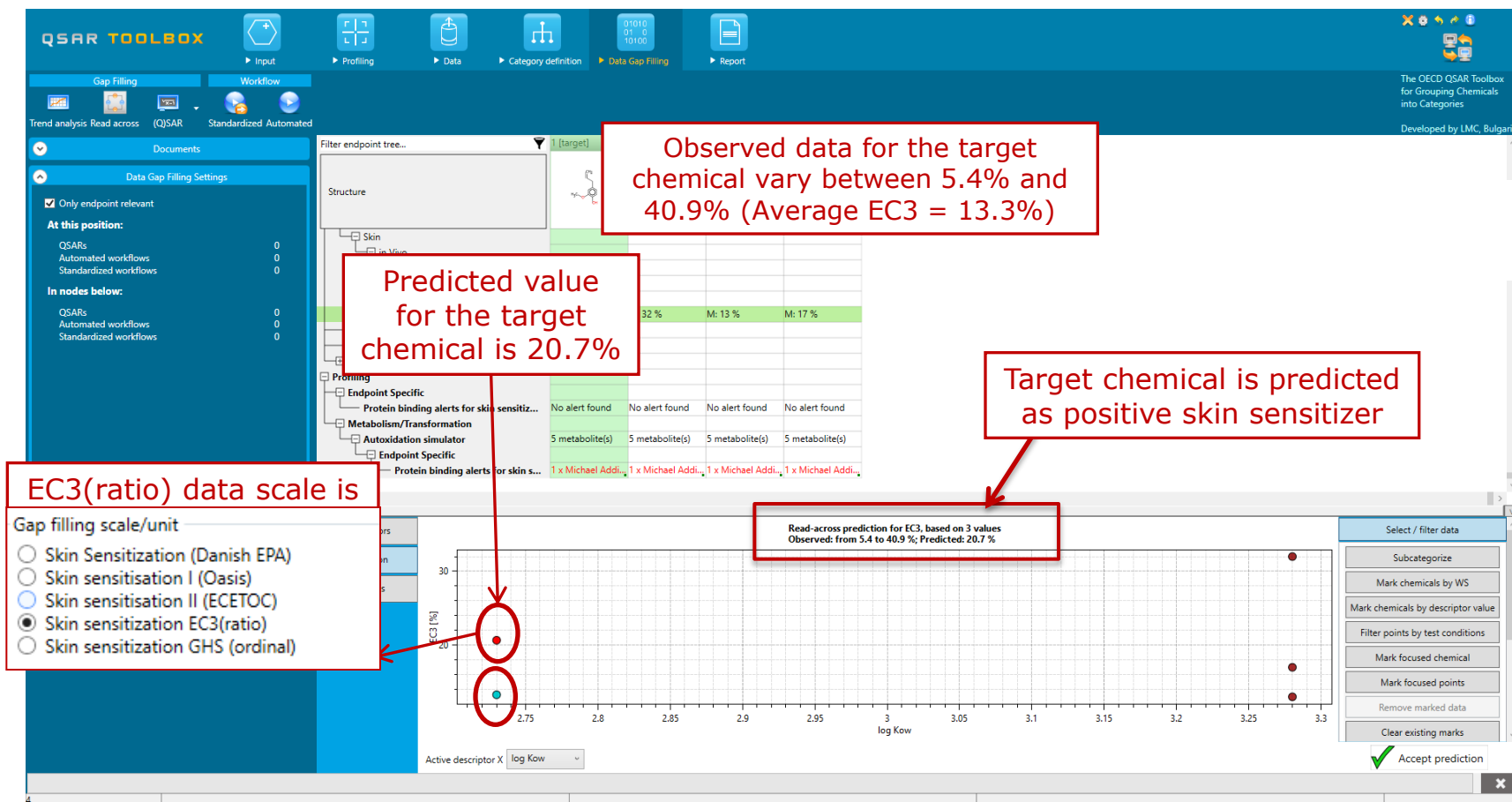
Remove marked data

Clear existing marks

Accept prediction

Data gap filling

Prediction for Skin sensitization, EC3: Overview



Data gap filling

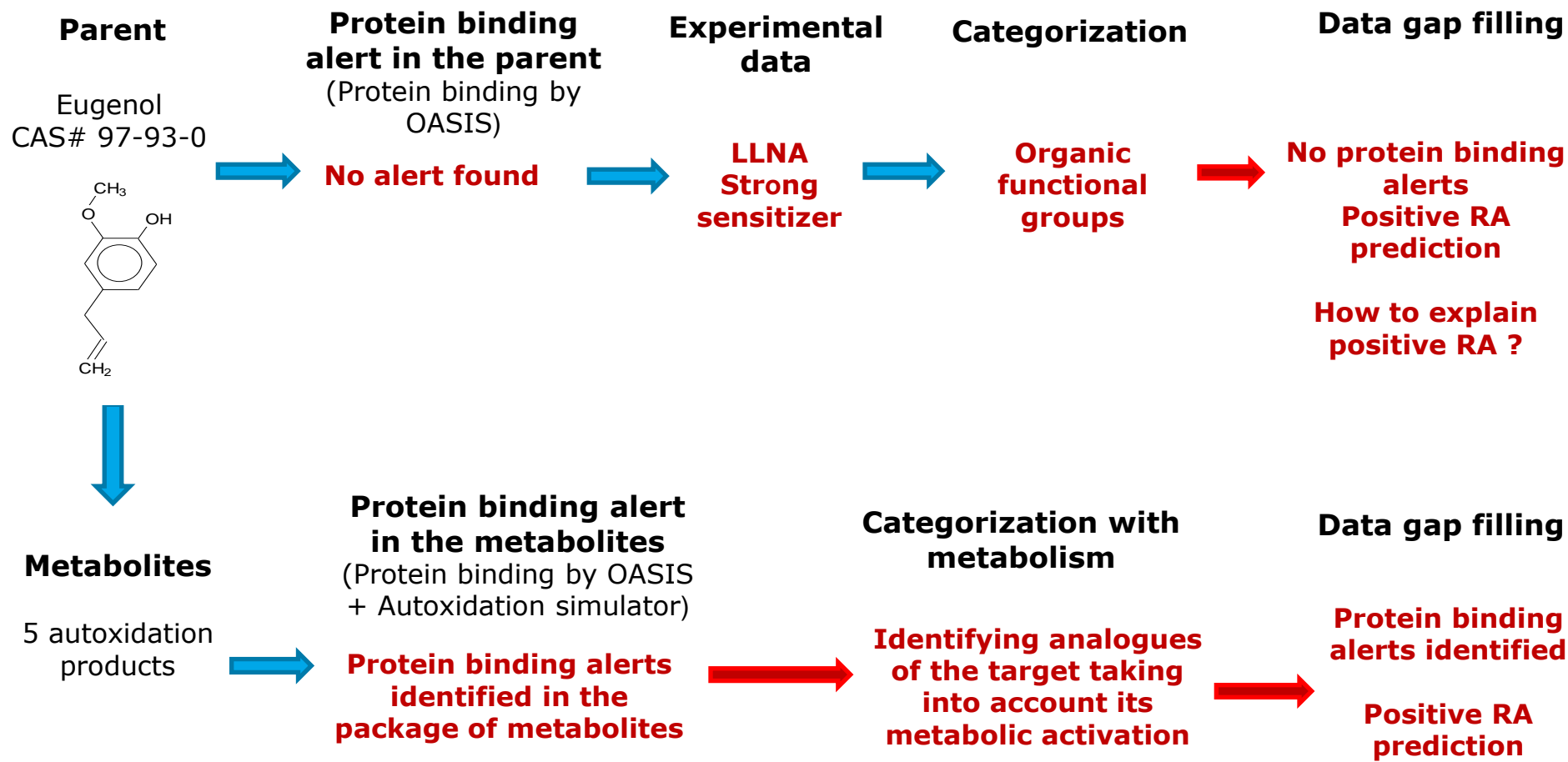
Prediction for Skin sensitization, EC3: Acceptance

The screenshot displays the QSAR Toolbox interface during a Data Gap Filling workflow. A 'Confirm' dialog box is open, warning that the target chemical is out of the parametric domain (log Kow) and asking for confirmation to accept the prediction. A scatter plot shows EC3 [%] on the y-axis (ranging from 20 to 30) versus log Kow on the x-axis (ranging from 2.8 to 3.3). The plot title is 'Read-across prediction for EC3, based on 3 values' and the data is 'Observed: from 5.4 to 40.9 %; Predicted: 20.7 %'. A sidebar on the right contains a list of options, with 'Accept prediction' highlighted. A blue box at the bottom left contains instructions: 1. Click **Accept prediction**; 2. Click **Yes** to confirm the prediction.

1. Click **Accept prediction**;

2. Click **Yes** to confirm the prediction.

Recap



Outlook

- Background
- Keywords
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - Categorization
 - Data gap filling
 - **Report**

Report Overview

- The Report module allows you to generate a report on the predictions performed within the Toolbox.
- This module contains a predefined report template with automatically populated sections as well as manually editable sections, where the users could add some additional/custom information.
- The Prediction report generates three files:
 - *Prediction report* - a PDF file containing information for the target and how the prediction is obtained.
 - *Category report* - a PDF file containing information for the consistency of the final category (target plus used analogues)
 - *Data matrix* - a MS Excel file containing chemicals used for prediction along with their data for selected parameters, profiles and endpoint tree positions.
- The generated reports can then be saved or just opened.

Report

The screenshot displays the QSAR Toolbox Report module. The top navigation bar includes buttons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report (highlighted with a red dashed box and callout 1). The left sidebar contains a document tree with a 'Prediction' button (highlighted with a red box and callout 3). The main area shows a table of results for a chemical structure. The table has columns for target, 1, 2, 3, 4, and 5. The row for 'EC3' is highlighted in yellow, and a cell containing 'M: 5.4 %' is highlighted with a red box and callout 2.

Structure	1 [target]	2	3	4	5
Structure	<chem>CC1=CC=C(C=C1)C(=O)O</chem>	<chem>CC1=CC=C(C=C1)C(=O)O</chem>	<chem>CC1=CC=C(C=C1)C(=O)O</chem>	<chem>CC1=CC=C(C=C1)C(=O)O</chem>	<chem>CC1=CC=C(C=C1)C(=O)O</chem>
Physical Chemical Properties					
Environmental Fate and Transport					
Ecotoxicological Information					
Human Health Hazards					
Acute Toxicity					
ADME					
Bioaccumulation					
Carcinogenicity					
Developmental Toxicity / Teratogenicity					
Genetic Toxicity					
Immunotoxicity					
Irritation / Corrosion					
Neurotoxicity					
Photoinduced toxicity					
Repeated Dose Toxicity					
Sensitisation					
Skin					
in Vivo					
GPMT	1/2 M: Positive				
HRIPT	1/3 M: 8E+03 µg/cm ²				
LLNA					
EC3	4/26 M: 5.4 % M: 5.4 %	M: 32 %	M: 13 %	M: 17 %	
Miscellaneous	1/1 M: Category A				
ToxCast					
Toxicity to Reproduction					
Toxicokinetics, Metabolism and Distribution					
Profiling					
Endpoint Specific					
Protein binding alerts for skin sensitiz...	No alert found	No alert found	No alert found	No alert found	No alert found
Metabolism/Transformation	5 metabolite(s)	37 metabolite(s)	5 metabolite(s)	5 metabolite(s)	5 metabolite(s)
Autoxidation simulator	1 x Michael Addi...	15 x Michael Ad...	1 x Michael Addi...	1 x Michael Addi...	1 x Michael Addi...
Endpoint Specific					
Protein binding alerts for skin s...					

1. Go to the **Report** module;
2. Click on the cell with the prediction result;
3. Click **Prediction**.

Report

Customize report content and appearance

Select which sections to include into report by checking/unchecking the corresponding section box. Rearrange sections order of appearance by using buttons "Move Up" and "Move Down".

Wizard pages

- Customization
 - Customize report
 - Prediction
 - Target and prediction summary
 - Prediction details (I)
 - Prediction details (II)
 - Target profiles
 - Analogues selection details
 - Appendix: Grouping / subcategorization
 - Appendix: Specific report explanations
 - Category
 - Category definition and members
 - Consistency check
 - Options
 - Data matrix
 - Options

Add RAAF scenario

- ☒ **Prediction**
 - ☒ Target and prediction summary
 - ☒ Prediction details (I)
 - ☒ Prediction details (II)
 - ☒ Target profiles
 - ☒ Analogues selection details
 - ☐ Appendix: Grouping / subcategorization
 - ☐ Appendix: Specific report explanations
- ☒ **Category**
 - ☒ Category definition and members
 - ☒ Consistency check
 - ☒ Options
- ☒ **Data matrix**
 - ☒ Options

Remove password protection of the PDF files.
Note: If the protection is removed, this will be specified in the first page of the report.

Back Next Cancel **Create report** Move Up Move Down

Generated report files

The following files were generated.
Select a file to open or save.

Prediction report
Category report
Data matrix

Open Save as

Preview of the generated report files is given on the next slides.

1. Click **Create report** button;
2. Open and/or Save the files

Report

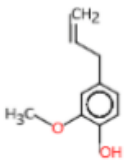
Prediction report

Prediction of EC3 for Eugenol

1 / 7

QSAR Toolbox prediction for single chemical

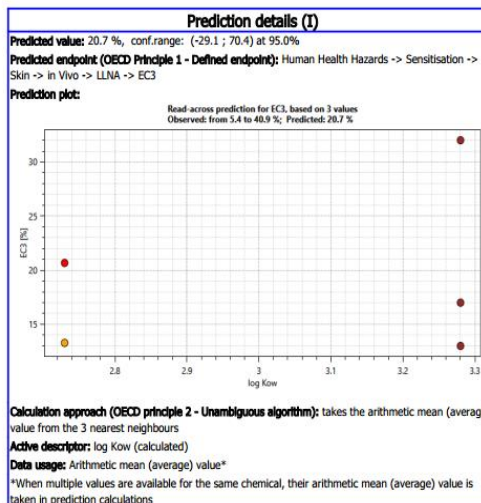
Date: 10 Apr 2020
Author(s):
Contact details:

Target information		
Structural information	Numerical identifiers	Chemical names
SMILES: <chem>COc1cc(CC=C)ccc1O</chem>	CAS#: 97-53-0 Other: EC Number: 2025B91	"eugenol (4-allyl-2-methoxyphenol); eugenol; 4-allyl-2-methoxyphenol; phenol, 2-methoxy-4-[(2-propenyl)]-; 2-methoxy-4-[(prop-2-en-1-yl)]phenol; phenol, 4-allyl-2-methoxy-; 1-allyl-2-methoxy-4-hydroxybenzene; 2-methoxy-4-[(2-propenyl)]phenol; 4-allyl-2-methoxyphenol; p-allylgustacal"
Structure 		[eugenol]eugenol (4-allyl-2-methoxyphenol); 4-allyl-2-methoxyphenol; 4-allyl-2-methoxyphenol; 4-allyl-2-methoxyphenol; 4-allyl-2-methoxyphenol; 4-allyl-2-methoxyphenol; 4-allyl-2-methoxyphenol; 4-allyl-2-methoxyphenol; 4-allyl-2-methoxyphenol; 4-allyl-2-methoxyphenol

Prediction summary
Predicted endpoint: EC3; No effect specified; No species specified; No duration specified; No guideline specified Predicted value: 20.7 (from -29.1 to 70.4) Unit/scale: % Data gap filling method: Read-across analysis Summary: manually editable field Not provided by the user

Prediction of EC3 for Eugenol

2 / 6



Prediction of EC3 for Eugenol

4 / 6

Target profiles	
(OECD principle 5 - Chemical and biological mechanisms)	
Profiles used for grouping/subcategorization	
Using of "Autoxidation simulator" Combined parent and products requirements: No alert found<AND>Michael Addition >> Michael addition on quinoid type compounds >> Quinone methide(s)/imines; Quinone oxime structure; Nitroquinones, Naphthoquinone(s)/imines<AND>Radical reactions >> Free radical formation >> Hydroperoxides<AND>SN2 >> Ring opening SN2 reaction >> Epoxides, Aziridines and Sulfuranes (Protein binding alerts for skin sensitization by OASIS) (primary grouping)	Parent and 5 metabolite(s); Has all of the required categories: Michael Addition, Michael Addition >> Michael addition on quinoid type compounds, Michael Addition >> Michael addition on quinoid type compounds >> Quinone methide(s)/imines; Quinone oxime structure; Nitroquinones, Naphthoquinone(s)/imines, No alert found, Radical reactions, Radical reactions >> Free radical formation >> Hydroperoxides, SN2, SN2 >> Ring opening SN2 reaction >> Epoxides, Aziridines and Sulfuranes
log Kow (calculated): 2.73	

The Prediction report is a file, generated in a PDF format. The Prediction report contains information for:

- The target chemical;
- The prediction
- Profilers used for grouping/ subcategorization
- Analogues selection

The five OECD principles for validation of (Q)SAR models are also given in the report.

Report

Category report

Chemicals category 1 / 10

QSAR Toolbox report for category

1. Category definition

1.1. Category definition

Category name
Not provided by the user

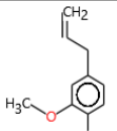
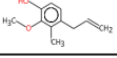
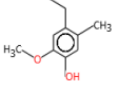
Covered (target) endpoint(s)
- Human Health Hazards/Sensitisation: EC3, LLNA, in Vivo, Skin

Category hypothesis
Not provided by the user

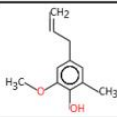
1.2. Category members

Information of category members

Table of category members

#	CAS	Name	SMILES	Structure
1	97-53-0	Eugenol	<chem>COc1cc(C=C)ccc1O</chem>	
2	186743-26-0	3-METHYL_EUGENOL	<chem>COc1c(C)cc(C=C)cc1O</chem>	
3	186743-25-9	5-METHYL_EUGENOL	<chem>COc1cc(C=C)cc(C)cc1O</chem>	

Chemicals category 2 / 10

4	186743-24-8	6-METHYL_EUGENOL	<chem>COc1cc(C=C)cc(C)cc1O</chem>	
---	-------------	------------------	-----------------------------------	---

Ranges for selected physicochemical properties and calculated parameters
Not provided by user

Purity / Impurity
Not provided by the user

1.3. Profiles/Metabolisms

List of profiles/metabolisms

Profiles used for grouping/subcategorization:

- Using of "Autoxidation simulator" Combined parent and products requirements: No alert found<AND>Michael Addition >> Michael addition on qinoid type compounds >> Quinone methide(s)/imines; Quinone oxime structure; Nitroquinones, Naphthoquinone(s)/imines<AND>Radical reactions >> Free radical formation >> Hydroperoxides<AND>SN2 >> Ring opening SN2 reaction >> Epoxides, Aziridines and Sulfuranes (Protein binding alerts for skin sensitization by OASIS) (primary grouping)

2. Consistency check

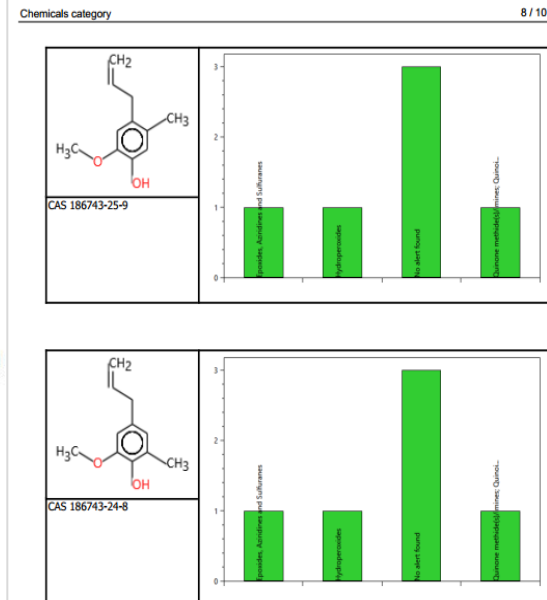
2.1. Physicochemical similarity

Physicochemical similarity based on calculated parameters
Physicochemical similarity based on experimental data
Not available

Comments on physicochemical similarity
Not provided by the user

2.2. Structural similarity

Structural similarity




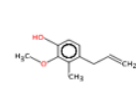
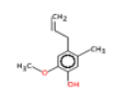
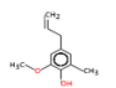
The Category report is a file, generated in a PDF format.
The Category report* contains information for:

- Category members (i.e. target and analogues used for the prediction);
- Physicochemical similarity between the category members
- Structural similarity between the category members
- Mechanistic similarity between the category members
- Additional data related to the target endpoint

*** Note:** In the current example some of the Category report sections will be empty. They will be automatically populated if the *Category elements* are preliminary applied. See more details in *Tutorial 27: Category elements for assessing Category consistency*.

Report

Data matrix report

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Target chemical		Neighbour #1		Neighbour #2		Neighbour #3									
2	Substance ID information															
3	Structure															
4	CAS number		97-53-0		186743-26-0		186743-25-9		186743-24-8							
5	Chemical name		Eugenol		3-METHYL_EUGENOL		5-METHYL_EUGENOL		6-METHYL_EUGENOL							
6	Other identifier															
7	SMILES		COc1cc(CC=C)ccc1O		COc1c(O)ccc(CC=C)c1C		COc1cc(CC=C)c(C)cc1O		COc1cc(CC=C)cc(C)c1O							
8																
9																
10	Profiling information used for grouping/subcategorization															
11	Profiles															
12	Profiles used for grouping/subcategorization															
13	Using of "Autoxidation simulator" Combined parent and products requirements: No alert found<AND>Michael Addition >> Michael addition on quinoid type compounds >> Quinone methide(s)/imines; Quinoid oxime structure; Nitroquinones, Naphthoquinone(s)/imines<AND>Radical reactions >> Free radical formation >> Hydroperoxides<AND>SN2 >> Ring opening SN2 reaction >> Epoxides, Aziridines and Sulfuranes (Protein binding alerts for skin sensitization by OASIS) (primary grouping)		Parent and 5 metabolite(s); Has all of the required categories: Michael Addition, Michael Addition >> Michael addition on quinoid type compounds, Michael Addition >> Michael addition on quinoid type compounds >> Quinone methide(s)/imines; Quinoid oxime structure; Nitroquinones, Naphthoquinone(s)/imines, No alert found, Radical		Parent and 5 metabolite(s); Has all of the required categories: Michael Addition, Michael Addition >> Michael addition on quinoid type compounds, Michael Addition >> Michael addition on quinoid type compounds >> Quinone methide(s)/imines; Quinoid oxime structure; Nitroquinones,		Parent and 5 metabolite(s); Has all of the required categories: Michael Addition, Michael Addition >> Michael addition on quinoid type compounds, Michael Addition >> Michael addition on quinoid type compounds >> Quinone methide(s)/imines; Quinoid oxime structure; Nitroquinones,		Parent and 5 metabolite(s); Has all of the required categories: Michael Addition, Michael Addition >> Michael addition on quinoid type compounds, Michael Addition >> Michael addition on quinoid type compounds >> Quinone methide(s)/imines; Quinoid oxime structure; Nitroquinones,		Parent and 5 metabolite(s); Has all of the required categories: Michael Addition, Michael Addition >> Michael addition on quinoid type compounds, Michael Addition >> Michael addition on quinoid type compounds >> Quinone methide(s)/imines; Quinoid oxime structure; Nitroquinones,					
14	Endpoint Specific															
15	Protein binding alerts for skin sensitization by OASIS		No alert found		No alert found		No alert found		No alert found							
16																
17	Measured and predicted data															
18	Data used for prediction															
19	sublevel	endpoint	value	unit	species, duration, test type, type of method, assay, strain, test guideline, year, reference	value	unit	species, duration, test type, type of method, assay, strain, test guideline, year, reference	value	unit	species, duration, test type, type of method, assay, strain, test guideline, year, reference	value	unit	species, duration, test type, type of method, assay, strain, test guideline, year, reference	value	unit
20	Sensitisation	EC3			EC3 in Vivo	32	%	EC3 in Vivo	13	%	EC3 in Vivo	17	%	EC3 in Vivo		
21	Human Health Hazards#Sensitisation															
22	sublevel	endpoint	value	unit	species, duration, test type, type of method, assay, strain, test guideline, year, reference	value	unit	species, duration, test type, type of method, assay, strain, test guideline, year, reference	value	unit	species, duration, test type, type of method, assay, strain, test guideline, year, reference	value	unit	species, duration, test type, type of method, assay, strain, test guideline, year, reference	value	unit
23	Sensitisation	CD54	Positive		in Vitro			in Vitro			in Vitro			in Vitro		
24	Sensitisation	CD54	Positive		Dendritic Cell Activity (h-CLAT)			Dendritic Cell Activity (h-CLAT)			Dendritic Cell Activity (h-CLAT)			Dendritic Cell Activity (h-CLAT)		
25	Sensitisation	CD86	Positive		in Vitro			in Vitro			in Vitro			in Vitro		

Outlook

- Background
- Keywords
- Objectives
- The exercise
- Workflow
- **Save/Load**

Saving the prediction result

Overview

- The Save/Load functionalities allow storing/restoring of Toolbox documents including loaded chemicals, experimental data, profiles, predictions, etc.
- These functionalities are implemented based on saving the sequence of actions performed in a Toolbox document and later executing of these actions in the same sequence of steps in order to get the same result(s).
- Saving/Loading of a Toolbox file is shown on the next screenshots.

Saving the prediction result

1. Go back to the Input module;

2. Click **Save** button;

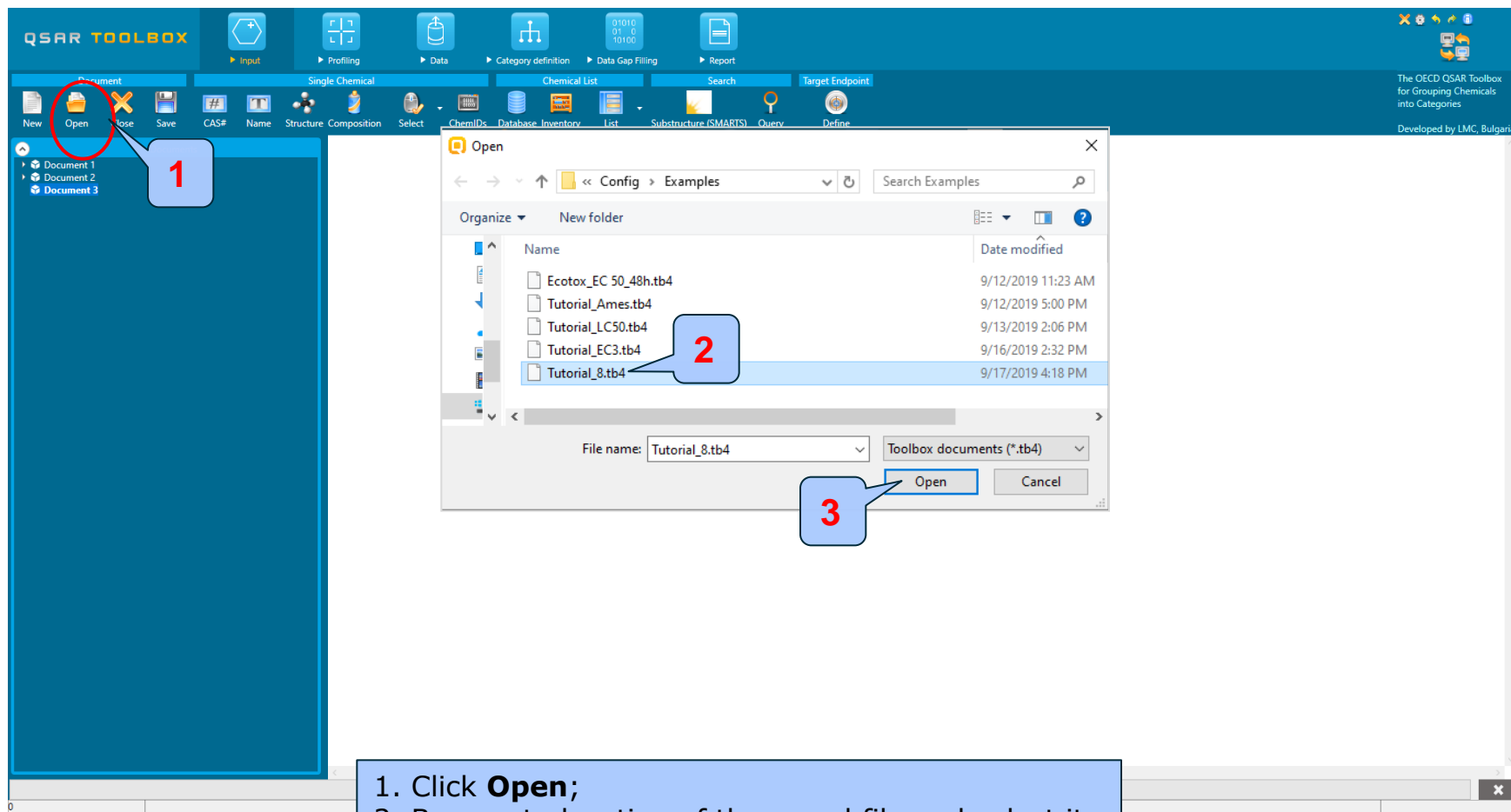
3. Click **Yes**;

4. Define name of the file;

5. Click on the **Save** button;

6. Click **OK**.

Open the saved file



1. Click **Open**;
2. Browse to location of the saved file and select it;
3. Click **Open**