

OECD QSAR Toolbox v.4.1

Example for predicting Skin Sensitization of
mixture

Outlook

- **Background**
- Objectives
- The exercise
- Workflow
- Save the prediction

Background

This is a step-by-step presentation designed to take the user of the Toolbox through the workflow for prediction skin sensitization of mixture.

Outlook

- Background
- **Objectives**
- The exercise
- Workflow
- Save the prediction

Objectives

- **This presentation reviews a number of functionalities of the Toolbox:**
 - 2D editor for defining Mixture components
 - Filling data gaps by Independent mode approach

Outlook

- Background
- Objectives
- **The exercise**
- Workflow
- Save the prediction

Exercise

- In this exercise we will predict skin sensitization of mixture, which is the “target” chemical,
- Investigate the mode of action for each components of the mixture,
- Gather available experimental data for target chemical,
- Investigate skin sensitization of non-tested component,
- Applying read across for non-tested component, and
- Predict skin sensitization potential of mixture based on experimental data of tested compounds and predicted data of non-tested one.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
- Save the prediction

Workflow

- **The Toolbox has six modules which are used in a sequential workflow:**
 - Chemical Input
 - Profiling
 - Endpoints
 - Category Definition
 - Filling Data Gaps
 - Report

Outlook

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

Chemical Input Overview

- This module provides the user with several means of entering the chemical of interest or 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.

Chemical Input

Ways of Entering a mixture

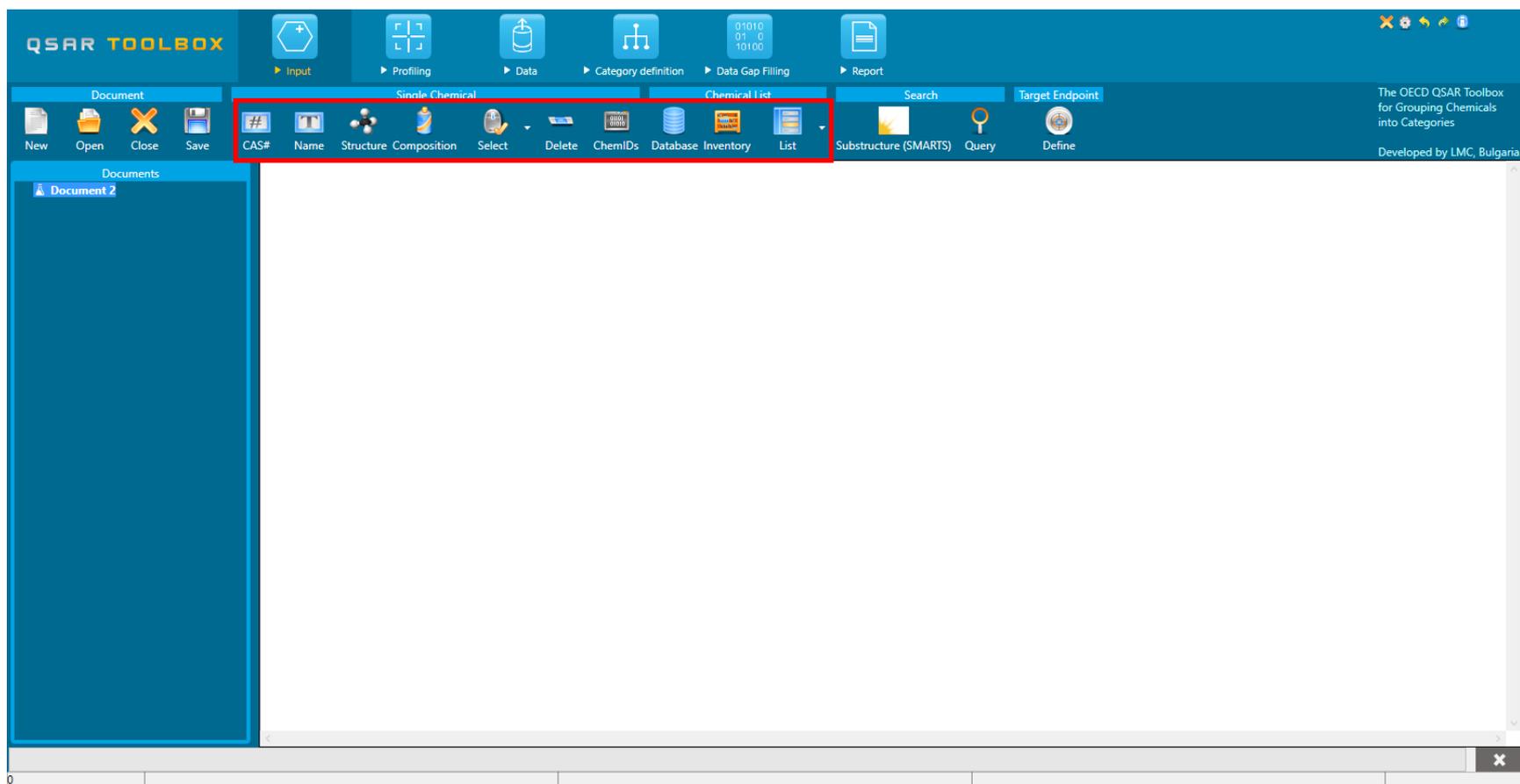
User alternatives for defining mixtures with known compositions:

- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- SMILES (simplified molecular information line entry system) notation/InChi
- Drawing mixture constituents and defining their quantities
- Select from User List/Inventory/Databases
- Load file with mixture

Getting Started

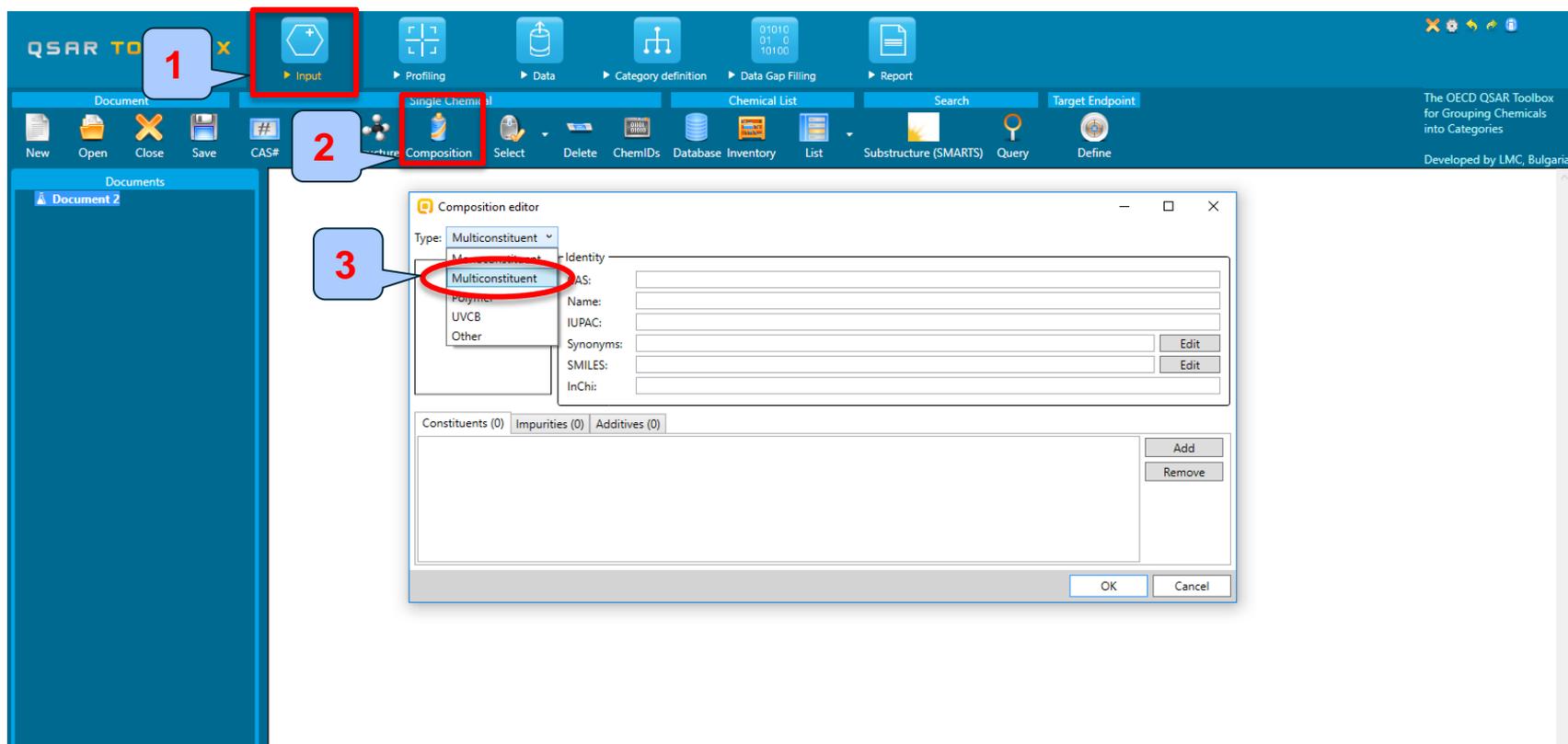
- Open the Toolbox.
- The six modules in the workflow are seen listed next to “QSAR TOOLBOX”.
- Click on “Input” (see next screen shot)

Chemical Input Input Screen



Chemical input

Input mixture



1. Select **Input**
2. Click on **Composition**
3. From composition editor select **Multiconstituent**

Chemical input

Input mixture

The screenshot shows the 'Composition editor' window. On the left, under 'Constituents (1)', there is a list of chemical structures. One structure is highlighted with a red box and a blue callout labeled '1'. This structure is also shown in the main 2D editor window, which is highlighted with a blue callout labeled '2'. The 2D editor window has an 'OK' button at the bottom, highlighted with a blue callout labeled '3'. Below the 2D editor, there are fields for 'Synonyms', 'SMILES', and 'InChI', each with an 'Edit' button. At the bottom of the main window, there are 'Add' and 'Remove' buttons, and a 'Concentration' section with 'Typical concentration' and 'Concentration range' fields, each with a 'Family' dropdown set to 'Mass' and a 'Unit' dropdown.

1. Click on **Edit** to open the 2D editor,
2. **Draw structure** of the first component (#1),
3. Confirm by clicking **OK**.

Chemical input

Input mixture

The screenshot displays the 'Composition editor' window with a list of components. The 'Add' button is highlighted with a red box and labeled '1'. Below the list, the 'Remove' button is labeled '2'. The '2D Editor' window is open, showing a chemical structure of a cyclohexane ring with a carbonyl group attached to another cyclohexane ring, labeled '4'. The 'Edit' button is labeled '3', and the 'OK' button is labeled '5'. The 'Composition editor' window also shows fields for 'Identity' and 'CAS', and a 'Concentration' section with a 'Typical concentration' field.

1. Click on **Add**,
2. Scroll down to add the second component (#2),
3. Click **Edit** to open the 2D editor,
4. Draw structure of the second component,
5. Confirm by OK.

Use the same procedure to add as much components as needed (see next slide).

Chemical input

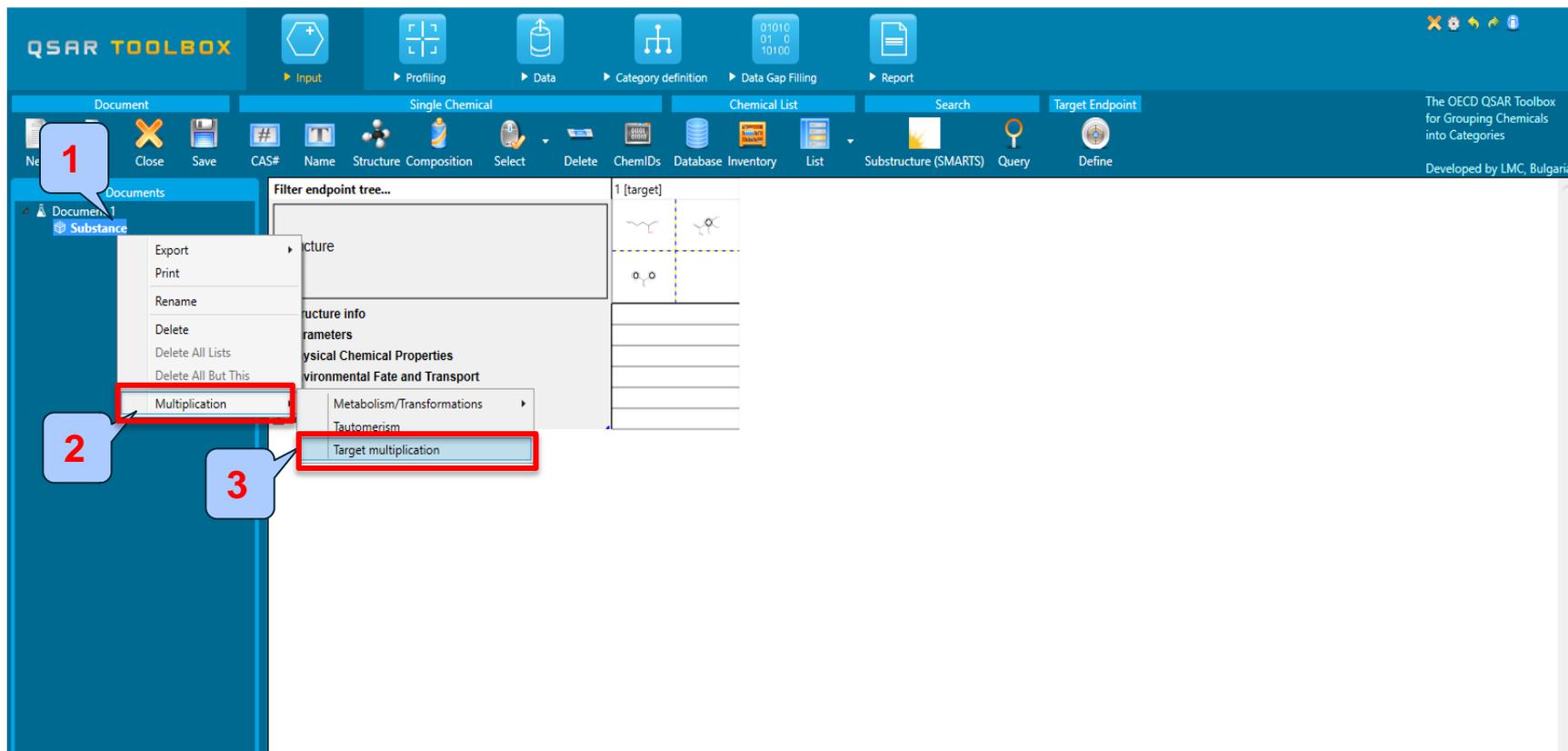
Input mixture

The screenshot displays the QSAR Toolbox software interface. The main window is titled 'Composition editor' and is set to 'Multiconstituent' type. It features a 'Identity' section with input fields for CAS, Name, IUPAC, Synonyms, SMILES, and InChI, each with an 'Edit' button. Below this are sections for 'Constituents (0)', 'Impurities (0)', and 'Additives (0)', each with 'Add' and 'Remove' buttons. At the bottom, there are three separate chemical structure editors, each with a 'Rectangle' dropdown and a toolbar. The first editor shows 2,4,6-trichlorobenzaldehyde, the second shows benzophenone, and the third shows 1-propanol. The main interface includes a menu bar with options like Document, Single Chemical, Chemical List, Search, and Target Endpoint, and a toolbar with icons for New, Open, Close, Save, and various chemical operations.

- There are three components in this exercise,
- Close the Composition editor by clicking **OK** and return to data matrix.

Chemical input

Input mixture



1. Make right mouse click on **Substance**,
2. Select **Multiplication**,
3. Click on **Target multiplication**.

Chemical Input

Target chemical identity

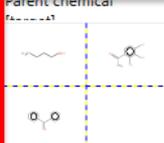
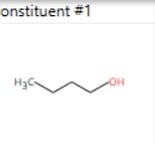
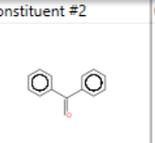
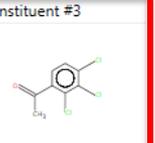
- The already drawn target structures automatically appear on the data matrix.
- Note that no CAS number or name is associated with this chemical.
- This means the target chemical is not listed in the chemical inventories/databases available in Toolbox (see next slide).

Chemical Input

Target chemical identity

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes options like Document, Single Chemical, Chemical List, Search, and Target Endpoint. Below the menu is a toolbar with icons for New, Open, Close, Save, CAS#, Name, Structure, Composition, Select, ChemIDs, Database, Inventory, List, Substructure (SMARTS), Query, and Define. The main workspace is divided into several panels:

- Documents:** A tree view on the left showing a hierarchy: Document 1 > Substance > Composition list > Constituent #1, #2, and #3.
- Filter endpoint tree...:** A central panel with a search box and a list of endpoints: Structure info, Parameters, Physical Chemical Properties, Environmental Fate and Transport, Ecotoxicological Information, and Human Health Hazards.
- Chemical List:** A table on the right showing the chemical identity of the constituents. The first row is highlighted with a red border.

Parent chemical	Constituent #1	Constituent #2	Constituent #3
	 <chem>CCCCO</chem>	 <chem>O=C(c1ccccc1)c2ccccc2</chem>	 <chem>CC(=O)c1cc(Cl)cc(Cl)c1</chem>

All three constituents of the mixture can be treated as individual substances.

Outlook

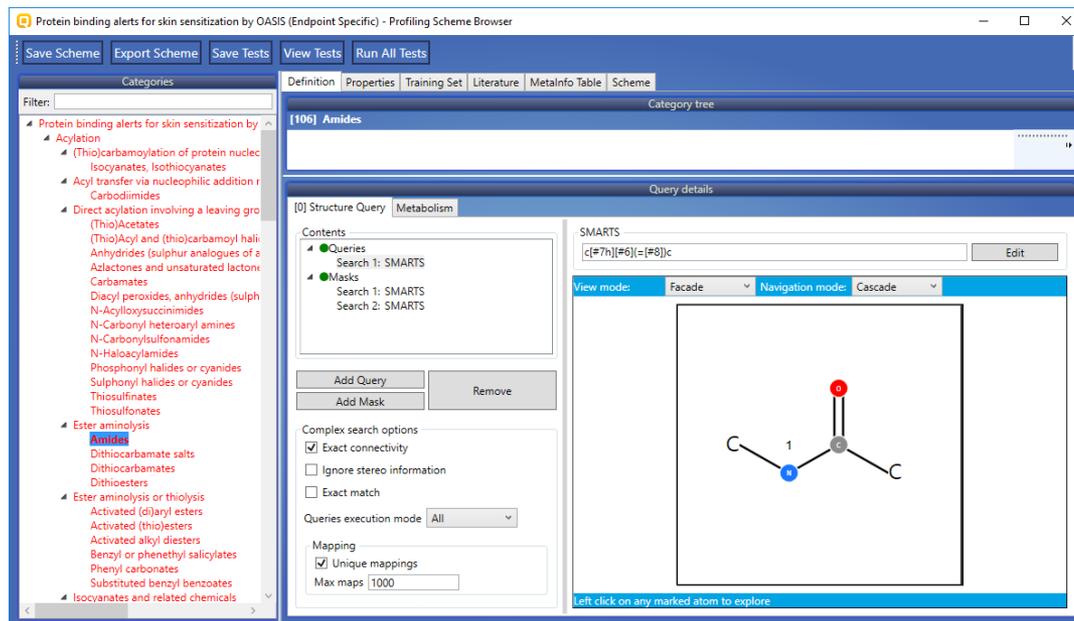
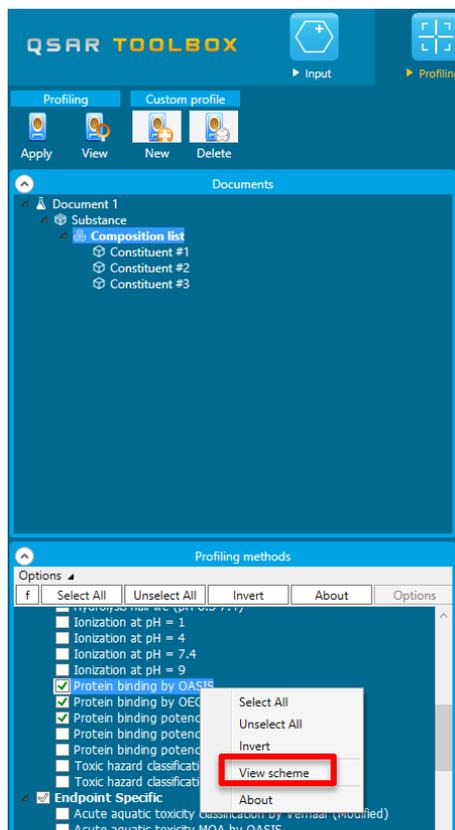
- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - **Profiling**

Profiling Overview

- “Profiling” 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 information includes likely mechanism(s) of action, as well as observed or simulated metabolites.
- For most of the profilers, background information can be retrieved by highlighting one of the profilers (for example, Protein binding by OASIS and clicking on “View scheme” (see next screen shots)).

Profiling

Side-Bar to Profiling



Make right mouse click on selected profiler followed by "View scheme" to see defined structural boundaries associated with Amides in **Protein binding by OASIS** profile.

Profiling Side-Bar to Profiling

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

Categories: Filter: []

- Protein binding by OASIS
 - Acylation
 - (Thio)carbamoylation of protein nucleophiles
 - Isothiocyanates, Isocyanates
 - Acyl transfer via nucleophilic addition reaction
 - Carbodiimides
 - Direct acylation involving a leaving group
 - (Thio)Acetates
 - (Thio)Acyl and (thio)carbamoyl halides and cyanides
 - Anhydrides (sulphur analogues of anhydrides)
 - Azlactones and unsaturated lactone derivatives
 - Carbamates
 - Diacyl peroxides, anhydrides (sulphur analogues of dia...
 - N-Acylloxysuccinimides
 - N-Carbonyl heteroaryl amines
 - N-Carbonylsulfonamides
 - N-Haloacylamides
 - Phosphonyl halides or cyanides
 - Sulphonyl halides or cyanides
 - Thiosulfonates
 - Thio-sulfonates
 - Ester aminolysis
 - Amides** (1)
 - Amide salts
 - Amides
 - Ester aminolysis or thiolysis
 - Carbonyl esters
 - Activated (thio)esters
 - Activated alkyl diesters
 - Benzyl or phenethyl salicylates
 - Phenyl carbonates
 - Substituted benzyl benzoates
 - Isocyanates and related chemicals
 - Ketenes
 - Ring opening acylation
 - Active cyclic agents
 - beta-Lactams
 - Cyclopropanones
 - Thio-lactones
 - Ionic interaction
 - Electrostatic interaction of tetraalkylammonium ion with pro...
 - Tetraalkylammonium ions

Definition | Properties | Training Set | **Literature** (2) | MetaInfo Table | Scheme

Mechanistic Domain: Acylation

Mechanistic Alert: Ester aminolysis

Structural Alert: Amide (3) → **Mechanistic alert**

The chemical is a strong sensitizer as a result of **Amide aminolysis**:

Textual description (3)

Ester aminolysis proceeds in two stages as an addition-elimination process in which a nucleophilic reactant bonds to the electrophilic carbonyl carbon atom, to create a tetrahedral intermediate. This tetrahedral intermediate then undergoes an elimination to yield the products. In this two-stage mechanism the carbonyl carbon atom undergoes a hybridization change from sp^2 to sp^3 and back again.

A key feature of the first stage is the site at which the starting ester is protonated. Protonation of the carbonyl oxygen makes the carbonyl group more susceptible to nucleophilic attack, i.e. electrophilic enough for attack by protein.

The protein NH_2 or SH groups are good nucleophiles and attack the carbonyl carbon.

Electron withdrawing groups, such $-CN$, $-NO_2$, $-CF_3$, $-SO_3H$, facilitate the occurrence of the reaction.

1. Click on **Amides**,
2. Select **Literature** to see mechanistic justification.

Profiling

Profiling the target chemical

- Select the “Profiling methods” related to the target endpoint by clicking on the box next to the profilers name.
- This selects (a green check mark appears) or deselects (green check disappears) profilers.
- For this example, select the following profilers relevant to the Skin sensitization (see next screenshot):
 - Protein binding by OASIS – general mechanistic
 - Protein binding by OECD – general mechanistic
 - Protein binding potency – general mechanistic
 - Protein binding alerts for skin sensitization by OASIS – endpoint specific

Profiling

Side-Bar to Profiling

The screenshot shows the QSAR Toolbox Profiling interface. The top toolbar has a 'Profiling' button highlighted with a red box and a callout '2'. Below the toolbar, the 'Profiling methods' panel on the left has three items checked: 'Protein binding by OASIS', 'Protein binding by OECD', and 'Protein binding potency', with a callout '1'. The main table displays results for three constituents. A red box highlights a 'Schiff base formation' alert for Constituent #3.

Filter endpoint tree...	Constituent #1	Constituent #2	Constituent #3
Structure	<chem>CCCCO</chem>	<chem>c1ccc(cc1)C(=O)c2ccccc2</chem>	<chem>CC1=CC=C(C=C1)C(=O)C</chem>
Structure info			
Parameters			
Physical Chemical Properties			
Environmental Fate and Transport			
Ecotoxicological Information			
Human Health Hazards			
Profile			
General Mechanistic			
Protein binding by OASIS	No alert found	No alert found	Schiff base formation Schiff base formation >> Schiff base formation with carbonyl compounds Schiff base formation >> Schiff base formation with carbonyl compounds >> Arom SNAr SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compo SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compo
Protein binding by OECD	No alert found	No alert found	No alert found
Protein binding potency	Not possible to classify	Not possible to classify	Not possible to classify according to these rules (GSH)
Endpoint Specific			
Protein binding alerts for skin sensitization ...	No alert found	No alert found	Schiff base formation Schiff base formation >> Schiff base formation with carbonyl compounds Schiff base formation >> Schiff base formation with carbonyl compounds >> Arom SNAr SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compo SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compo
Metabolism/Transformations			
Observed rat liver metabolism with quantitati ...	0 metabolites	0 metabolites	0 metabolites

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

Outlook

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

Data

- “Data” refers to the electronic process of retrieving the environmental fate, eco-toxicity and toxicity data that are residing in the Toolbox.
- Data gathering can be executed in a global fashion (i.e. collecting all data of all endpoints) or on a more narrowly defined basis (e.g. collecting data for a single or limited number of endpoints).
- In this example, we limit our data gathering to the common Skin endpoints from databases associated with Skin Sensitization endpoint.

Data

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Data', 'Import', and 'Export'. The 'Data' menu is open, showing 'Gather', 'Import', 'IUCLID6', and 'IUCLID6'. A red circle highlights the 'Gather' button, with a callout '2' pointing to it. Below the menu, the 'Documents' panel shows a tree view with 'Document 1' expanded to 'Substance' and 'Composition list', which includes 'Constituent #1', 'Constituent #2', and 'Constituent #3'. The 'Databases' panel shows a list of databases with 'Skin Sensitization' and 'Skin sensitization ECETO6' selected, circled in red with a callout '1'. The main window displays a 'Filter endpoint tree...' on the left, a 'Parent Chemical' structure in the center, and a table of constituents on the right. The table has columns for 'Parent Chemical', 'Constituent #1', 'Constituent #2', and 'Constituent #3'. The 'Sensitisation' endpoint is highlighted in blue in the filter tree, with 'AW SW AOP' written next to it.

1. Select databases related to the target endpoint;
2. Click on **Gather**

Data

Process of collecting data

The screenshot displays the QSAR Toolbox software interface. The top navigation bar includes icons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below this, the 'Data' tab is active, showing options for Gather, Import, IUCLID, and IUCLID6. The main workspace is divided into several panels:

- Documents:** A tree view showing 'Document 1' containing a 'Substance' with a 'Composition list' of three constituents.
- Databases:** A list of toxicity databases with checkboxes. 'Skin Sensitization' and 'Skin sensitization ECETOC' are checked.
- Filter endpoint tree...:** A hierarchical tree of toxicity endpoints. 'Sensitisation' is expanded to show 'Skin' and 'in Vivo' (LLNA, EC3, Miscellaneous). 'ToxCast' is also visible.
- Table:** A table with columns for 'Parent Chemical', 'Constituent #1', 'Constituent #2', and 'Constituent #3'. The table contains two rows of data, both with '(1/1)' in the first column. The second row has 'M: Negative' in the second and fourth columns, which is circled in red.
- Dialog Box:** A small window in the center-right displays the message '2 points added across 2 chemicals.' with an 'OK' button.

Recap

- You have entered the mixture with defined components.
- You have profiled the target chemical mixture and found no protein binding alerts for two of the mixture constituents. The third constituent has positive protein binding alerts and could elicit skin sensitization effect.
- Negative experimental data has been found for two of mixture components. No experimental data has been found for the third constituent.
- The constituent without experimental data and positive protein binding alert has been used for further read across analysis. Then, all of the available data – experimental and predicted will be used for skin sensitization prediction of the mixture.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - **Read across prediction of constituent without data**
 - **Focus constituent without experimental data**

Read across prediction of constituent without data

Focus constituent

The screenshot displays the QSAR Toolbox interface. The top navigation bar includes icons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below this, there are sub-menus for Data, Import, and Export. The main workspace is divided into several panels:

- Documents:** Shows a tree view under 'Document 1' with 'Substance' and 'Composition list' containing 'Constituent #1', 'Constituent #2', and 'Constituent #3'.
- Databases:** A list of databases is shown, with 'REACH Skin sensitisation database (normalised)' highlighted.
- Filter endpoint tree...:** A hierarchical tree of endpoints is displayed, including 'Acute Toxicity', 'Bioaccumulation', 'Sensitisation', and 'ToxCast'.
- Data Table:** A table with columns for 'Parent chemical', 'Constituent #1', 'Constituent #2', and 'Constituent #3'. A row is highlighted in blue, and a context menu is open over it with 'Focus' selected.

1. Right mouse click on the chemical without experimental data,

2. Select **Focus**

This constituent is selected for further read-across prediction

1. Right mouse click on the chemical without experimental data,
2. Select **Focus**

Read across prediction of constituent without data

Focus constituent

The screenshot displays the QSAR Toolbox software interface. The top navigation bar includes sections for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below this, there are buttons for Gather, Import, IUCLID, and IUCLID6. The main workspace is divided into several panels:

- Documents:** A tree view on the left showing a hierarchy: Document 1 > Substance > Composition list > Constituent #1 (highlighted in blue).
- Filter endpoint tree...:** A central panel with a 'Structure' section containing a chemical structure diagram of a benzene ring with various substituents. A red callout bubble with the number '1' points to this structure.
- Structure info:** A list of expandable sections including Parameters, Physical Chemical Properties, Environmental Fate and Transport, Ecotoxicological Information, and Human Health Hazards.
- Databases:** A panel on the bottom left with 'Options' (f, Select All, Unselect All, Invert) and a list of databases with checkboxes. 'Skin Sensitization' and 'Skin sensitization ECETOC' are checked.
- Inventories:** A panel at the very bottom.

A red-bordered text box on the right side of the interface contains the following text:

This focused component appeared in separate data matrix

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - **Read across prediction of constituent without data**
 - Focus constituent without experimental data
 - **Define category**

Category Definition Overview

- This module provides the user with several means of grouping chemicals into a toxicologically meaningful category that includes the target molecule.
- This is the critical step in the workflow.
- Several options are available in the Toolbox to assist the user in refining the 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”.
- Detailed information about grouping chemical (Chapter 4) could be found in document “Manual for Getting started” published on OECD website:

<http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm>

Basic guidance for category formation and assessment

Suitable categorization phases:

1. Structure-related profilers.
2. Endpoint specific profilers (for sub-cat).
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 suitable categorization phases is shown on next slide

Suitable Categorization/Assessment Phases

Phase I. Structure based

- US EPA Categorization
- OECD Categorization
- Organic functional group
- Structural similarity
- ECOSAR

**Broad grouping
Endpoint Non-specific**

Repeating Phase I due to Multifunctionality of chemicals

Phase II. Mechanism based

- DNA binding mechanism
- Protein binding mechanism
- Genotoxicity/carcinogenicity
- Cramer rules
- Verhaar rule
- Skin/eye irritation corrosion rules

**Subcategorization
Endpoint Specific**

Metabolism accounted for

Phase III. Eliminating dissimilar chemicals

**Apply Phase I – for structural dissimilarity
Filter by test conditions – for Biological dissimilarity**

**Subcategorization
Endpoint Specific**

Read across prediction of constituent without data

Forming category for studied endpoint

Suitable Categorization/Assessment Phases

Phase I. Structure based

- US EPA Categorization
- OECD Categorization
- Organic functional group
- Structural similarity
- ECOSAR

Repeating Phase I due to Multifunctionality of chemicals

**Broad grouping
Endpoint Non-specific**

Phase I categorization in Toolbox

It is not recommended to use "Neutral organic" * as phase I

46 analogues are identified as Aryl halides by OFG

*Neutral organic category include chemicals having different functionalities as alcohols, ketones, ethers etc. In this respect the basic principle that structurally similar chemicals may elicit similar effects would not be preserved, because Neutral organic mixed many different functionalities

Read across prediction of constituent without data

Forming category for studied endpoint

- Based on the above recommendations and classifications from structurally similar profilers the OFG is used as initial categorization group
- Refinement of the initial group is based on endpoint-specific protein binding profiler:
 - Protein binding alerts for skin sensitization.

Category definition is a tool for grouping chemicals, which allows to group chemicals based on different measures of “similarity”. For more details see tutorials posted on LMC and OECD website:

<http://www.oecd.org/env/ehs/risk-assessment/theoecdqsartoolbox.htm>

<http://superhosting.oasis-lmc.org/products/software/toolbox/toolbox-support.aspx>

See next slides

Read across prediction of constituent without data

Define category by OFG

Based on above recommendations the OFG is used as an initial group (phase I)

1. Select Organic functional groups (OFG) and click on **Define**
2. Combination of three organic functional groups has been applied.
3. Two chemicals have been found.

Read across prediction of constituent without data

Define category by OFG

Except Aryl halide all other groups are removed

To make a category based on Aryl halides only:

1. Select Organic functional groups,
2. Click **Define**,
3. Use Control from keyboard for subsequent selection of Aryl and Ketone,
4. Click **Down** to remove Aryl and Ketone; Aryl halides remain only,
5. Click **OK** to confirm.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - **Read across prediction of constituent without data**
 - **Focus constituent without experimental data**
 - **Define category**
 - **Gather data for analogues**

Read across prediction of constituent without data

Gather data for analogues chemicals

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

Documents

- Document 1
 - Substance
 - Composition list
 - Constituent #1
 - Organic functional groups
 - Constituent #2
 - Constituent #3

Organic functional groups

Options

- Select All
- Unselect All
- Invert

Empiric

- Chemical elements
- Groups of elements
- Lipinski Rule Oasis
- Organic functional groups
- Organic functional groups (nested)
- Organic functional groups (US EPA)
- Organic functional groups, Norbert Hal
- Structure similarity
- Tautomers unstable

Toxicological

- Repeated dose (HESS)

Custom

- Example Prioritization Scheme (PBT)

Filter endpoint tree...

1 [target]

Structure

- Structure info
- Parameters
- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
- Profile

Grouping results

72 chemicals found.

OK

1

Read data?

All endpoints Choose...

from Tautomers

OK Cancel

1

Gather data

103 points added across 71 chemicals.

OK

2

1. Click **OK**,

2. Gather data

Read across prediction of constituent without data

Gather data for analogues chemicals

The screenshot displays the QSAR Toolbox interface. On the left, there is a 'Documents' tree showing a hierarchy: Document 1 > Substance > Composition list > Constituent #1 > Organic functional groups. Below this is a 'Databases' section with a search bar and a list of databases, including 'Skin sensitization ECETOQ' which is checked. The main area is a data matrix with columns numbered 1 to 8. Column 1 is labeled '[target]' and contains chemical structures. The other columns (2-8) also contain chemical structures. A red box highlights the text 'Experimental data for the identified analogues appear on data matrix'. A red oval highlights the 'in Vivo' row in the matrix, which shows 'M: Positive' for columns 2, 3, 4, 5, 6, 7, and 8, and 'M: Negative' for column 7. The 'in Vivo' row is also associated with '(71/102)' in the 'Skin' category.

Endpoint	1 [target]	2	3	4	5	6	7	8
Structure								
Structure info								
Parameters								
Physical								
Environment								
Ecotoxicology								
Human Health								
Acute								
Bioaccumulation								
Carcinogenicity								
Developmental Toxicity / Teratogenicity								
Genetic Toxicity								
Immunotoxicity								
Irritation / Corrosion								
Neurotoxicity								
Photoinduced toxicity								
Repeated Dose Toxicity								
Sensitisation								
- Respiratory Tract (1/1)								
- Skin (71/102)								
- in Vivo		M: Positive	M: Negative	M: Positive	M: Positive	M: Positive	M: Negative	M: Positive
- ToxCast								
- Toxicity to Reproduction								
- Toxicokinetics, Metabolism and Distribution								
Profile								

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - **Read across prediction of constituent without data**
 - **Focus constituent without experimental data**
 - **Define category**
 - **Gather data for analogues**
 - **Apply read across**

Read across prediction of constituent without data

Apply read across

QSAR TOOLBOX

Input Profiling Data Category definition Data Gap Filling Report

Gap Filling Workflow

Trend analysis Read across (QSAR) Standardized Automated

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

Documents

- Document 1
 - Substance
 - Composition list
 - Constituent #1
 - Organic functional groups
 - Enter GF(RA) with 60 chemicals, 85 data points
 - Enter GF(RA) with 40 chemicals, 50 data points
 - Constituent #2
 - Constituent #3

Filter endpoint tree...

Structure

LLNA (39/50)
EC3 (3/5)
Miscellaneous

ToxCast
Toxicity to Reproduction
Toxicokinetics, Metabolism and Distribution

Profile

1 [target]	4	5	7	9	15	
	M: Positive	M: Positive	M: Negative	M: Positive	M: Positive	M: N

Data Gap Filling Settings

Only endpoint relevant
 Only chemical relevant

At this position:

Select a cell with a rigid (bold) path

Automated workflows 1
Standardized workflows 1

Descriptors

Prediction

Read-across prediction for EC3, based on 6 values
Predicted: Positive

Active descriptor X log Kow

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

Accept prediction

Read across prediction of constituent without data

Subcategorization

- The initial category could be refined by subcategorizing the analogues according to the “Protein binding alerts for skin sensitization by OASIS” profiler.
- These steps are summarized in the next screen shots.

Read across prediction of constituent without data

Subcategorization by Protein binding alert for SS

The screenshot displays the QSAR Toolbox interface during a subcategorization process. The 'Subcategorization' dialog box is open, showing various options for filtering chemicals. In the 'Options' list, 'Protein binding alerts for skin sensitization by OASIS' is selected and circled in red, with a callout '2' pointing to it. The 'Adjust options' panel on the right shows 'Differ from target by' set to 'At least one category', also circled in red with callout '3'. The 'Data matrix' shows chemical structures in columns 23-28, with 'M: Positive' labels under columns 24 and 27. The 'EC3' plot shows a positive prediction for the target, circled in red with callout '4'. The 'Subcategorize' button is highlighted in the 'Select / filter' menu, circled in red with callout '1'.

Analogs have been found to be positive.
 Predicted SS effect of the target is clearly positive

1. Select filter data/Subcategorize
2. Select **Protein binding alerts for skin sensitization by OASIS**.
3. **Remove selected** to eliminate dissimilar chemicals.
4. **Accept prediction** to return to data matrix.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Read across prediction of constituent without data
 - **Filling data gap for skin sensitization of mixture**

Filling data gap for skin sensitization of mixture

Applying Independent Mode of Action

The screenshot shows the QSAR Toolbox interface with the 'Data Gap Filling' module active. A dialog box titled 'Possible data inconsistency' is open, displaying a list of assay and endpoint options. Under the 'Gap filling scale/unit' section, 'Skin sensitization II (ECETOC)' is selected. The dialog also shows 'converted data' and 'Data 3/3; Chemicals 3/3'. The 'OK' button in the dialog is highlighted with a red circle and callout 4. In the background, the 'Independent MOA' button is circled with a red circle and callout 2. The 'Skin sensitization II (ECETOC)' option in the dialog is circled with a red circle and callout 3. The 'OK' button in the dialog is also circled with a red circle and callout 1. A red arrow points from the 'Independent MOA' button to the 'Skin sensitization II (ECETOC)' option.

1. Click on the cell corresponding to the Skin Sensitization;
2. Select **Independent MOA**,
3. Select **Skin sensitization II(ECETOC)**,
4. Click **OK**.

Filling data gap for skin sensitization of mixture

Applying Independent Mode of Action

The screenshot displays the QSAR Toolbox interface during a 'Data Gap Filling' workflow. The main window shows a 'Filter endpoint tree...' with 'Sensitisation' selected, specifically 'Skin in Vivo'. Below this, a table shows the predicted outcomes for four chemical constituents (1-4) based on their log Kow values. The predicted outcomes are: M: Negative, M: Negative, and R: Positive. A scatter plot at the bottom shows the empirical calculation of A B C, EC3 based on 3 values, with a predicted positive outcome. The plot shows log Kow on the x-axis (ranging from 1 to 3.5) and A B C, EC3 on the y-axis (ranging from Negative to Positive). Three data points are plotted, with the highest log Kow point being predicted as positive. A green checkmark and 'Accept prediction' button are visible at the bottom right of the plot area.

Read across is applied for the mixture (assuming Independent Mode of Action)
 "Maximal" approximation type is set by default for categorical endpoints (worst case scenario).

Filling data gap for skin sensitization of mixture

Applying Independent Mode of Action

The screenshot displays the QSAR Toolbox software interface during a 'Data Gap Filling' session. The main workspace shows a 'Filter endpoint tree...' on the left and a 'Choose one' dialog box in the center. The dialog box lists various prediction modes, with 'Maximal' selected and highlighted by a red box and callout '3'. Below the dialog box is a scatter plot with 'log Kow' on the x-axis and 'A, B, C, EC3' on the y-axis. The plot shows several data points, with one point highlighted by callout '1'. On the right side of the interface, a sidebar contains several options, with 'Data usage' highlighted by a red box and callout '2'. At the bottom right, there is a 'Accept prediction' button with a green checkmark and callout '4'. A red text box in the upper right of the workspace states: 'Based on the positive skin sensitization value for one of the mixture components the prediction for the mixture is positive'. The top menu bar includes options like 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The left sidebar shows 'Documents' and 'Data Gap Filling Settings' with checkboxes for 'Only endpoint relevant' and 'Only chemical relevant'.

Consecutive steps:

1. Calculation options;
2. Data usage;
3. Maximal data;
4. Accept prediction.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Read across prediction of constituent without data
 - Filling data gap for skin sensitization of mixture
 - **Generating report for mixture**

Report

- Remember the report module allows user to generate a report on the predictions performed with the Toolbox.
- This module contains predefined report templates as well as a template editor which allows users to provide modifications.
- The report can be printed or saved in different formats.
- Generating the report is shown on next screenshots.

Report

The screenshot shows the QSAR Toolbox software interface. The top toolbar includes icons for 'input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The 'Reports' menu is open, showing 'Prediction Data' and 'Report' options. A 'Generated report files' dialog box is displayed, listing files like 'Prediction for mixture Prediction report Data matrix' and 'Individual prediction #1 Prediction report Data matrix'. A 'Wizard pages' dialog box is also visible, showing options for 'Prediction for mixture' and 'Customize report content and appearance'. A table at the bottom shows prediction results for 'IMOA: Positive'.

Filter endpoint tree...	Parent chemical [target]	Constituent #1	Constituent #2	Con...
Structure	<chem>CC1=CC=CC=C1</chem>	<chem>CC1=CC=CC=C1</chem>	<chem>CC1=CC=CC=C1</chem>	
Structure info				
Parameters				
Physical Chemical Properties				
Environmental Fate and Transport				
Ecotoxicological Information				
Human Health Hazards				
Acute Toxicity				
Bioaccumulation				
Carcinogenicity				
Developmental Toxicity / Teratogenicity				
Genetic Toxicity				
Immunotoxicity				
Irritation / Corrosion				
Neurotoxicity				
Photoinduced toxicity				
Repeated Dose Toxicity				
Sensitisation				
Skin				
in Vivo				
ToxCast				

1. Select **Report**
2. Click on the cell corresponding to Skin Sensitization/ in Vivo for mixture
3. **Prediction**
4. **Create report**
5. **Open** Prediction report

Report

Prediction of A B C, EC3 for mixture

1 / 6

QSAR Toolbox prediction for multicomponent substance

Based on observed and predicted data for mixture components

Date: 2 Aug 2017

Author(s):

Contact details:

Target information		
<p>Structural information</p> <p>SMILES: <chem>CCCCO.CC(=O)c1ccc(Cl)c(Cl)c1Cl.O=C(c1ccc(cc1)c1ccccc1</chem></p> <p>Structure</p> <div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <chem>H3C-CH2-CH2-CH2-CH2-OH</chem>  </div> <div style="text-align: center;">  </div> </div> <hr style="border-top: 1px dashed gray;"/> <div style="text-align: center;">  </div>	<p>Numerical identifiers</p> <p>EC#: N/A CAS#: Invalid CAS number: 0-00-0 Other: N/A</p>	<p>Chemical names</p>

Toolbox report for mixture

Outlook

- Background
- Objectives
- The exercise
- Workflow
- **Save the prediction result**

Saving the prediction result

- Saving functionality allows storing/restoring the current state of Toolbox documents including loaded chemicals, experimental data, profiles, predictions, etc.
- This functionality is implemented based on saving the sequence of actions that led to the current state of the Toolbox document and later executing these actions in the same sequence in order to get the same result(s).
- Saving the file with TB prediction is illustrated on next screenshot.

Saving the prediction result

The screenshot displays the QSAR Toolbox software interface during the 'Report' step. The main window shows a 'Wizard pages' dialog with a list of report sections: Prediction for mixture, Customize report, Target and prediction summary, Prediction details, Target profiles, Mixture components used for prediction, Data for mixture components, and Appendix: Data. A 'Generated report files' dialog is open, listing the following files: Prediction for mixture, Prediction report, Data matrix, and Individual prediction #1. A 'Save As' dialog is open, showing the file name 'Prediction report.pdf' and the 'Save' button highlighted with a red '3'. The 'Save as type' is set to 'Pdf files (*.pdf)'. The 'Save' button is also highlighted with a red '1'.

1. Click on **Save as** button;
2. Define path and name of the pdf file;
3. Click **Save** button.