

OECD QSAR Toolbox v.3.4

Predicting skin sensitization potential of
3,4-dinitrophenol taking into account
tautomerism

Outlook

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

Background

- This is a step-by-step presentation designed to take the user of the Toolbox through the workflow for filling data gap for skin sensitization taking into account tautomerism of target chemical.

Outlook

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

Objectives

This presentation demonstrates a number of functionalities of the Toolbox:

- Providing tautomeric set of target chemical
- Identify analogues for a set of tautomers
- Retrieve experimental results available for those analogues
- Fill data gaps by read across
- Save the prediction

Outlook

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

The Exercise

- In this exercise we will predict the skin sensitization potential for **(3,4-dinitrophenol) [CAS 577-71-9]**
- Set of simulated tautomers for the target chemical will be provided
- This prediction will be accomplished by collecting a set of similar analogues for set of target and its tautomers
- The initial category will be defined by Protein binding by OASIS v1.4
- Data gap will be filled by read-across

Outlook

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

Workflow

- **As you know the Toolbox has 6 modules which are typically used in sequence:**
 - Chemical Input
 - Profiling
 - Endpoint
 - Category Definition
 - Data Gap Filling
 - Reporting

Outlook

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

Chemical Input

There are two ways for simulating tautomers of chemicals

- During the process of entering the structure into the system (scenario 1)
- Simulating tautomersim of already entered structure (scenario 2)

In order to accelerate the workflow all databases are preliminary tautomerized, calculated and profiled. The results are stored in the database.

See next screen shots

Chemical Input

Search by CAS# in tautomerized databases (scenario 1)

The screenshot shows the QSAR Toolbox interface with a search dialog box open. The dialog box is titled 'Search by CAS #' and contains the following elements:

- 1:** A callout pointing to the 'CAS#' button in the main toolbar.
- 2:** A callout pointing to the text input field containing '577-71-9'.
- 3:** A callout pointing to the checked 'Tautomeric sets' checkbox.
- 4:** A callout pointing to the 'Search' button.
- 5:** A callout pointing to the 'OK' button.

The search results table in the dialog box is as follows:

Selected	CAS	Depiction	Names	CAS/Name	2D	AS/2D
1. Yes	577-71-9	Oc1ccc(N	<chem>Oc1ccc(N(=O)=O)cc1</chem>	1:: High 1:: Ar 2:: B 3:: E 4:: E 5:: G 6:: P 7:: R 8:: S 2:: Low 1:: F	1:: High 1:: E 2:: Ar 3:: R 4:: P 5:: E 6:: G 7:: B 8:: S 2:: Low 1:: F	High

Tautomeric set functionality allows searching for tautomeric forms of target chemical in previously tautomerized databases

- | | |
|----------------------------------|--------------------------|
| 1. Click on CAS# | 2. Enter 577-71-9 |
| 3. Select Tautomeric sets | 4. Click Search |
| | 5. OK |

Chemical Input

Target chemical identity (scenario 1)

The screenshot displays the QSAR Toolbox interface. In the 'Documents' panel on the left, a chemical structure is highlighted with a red circle and a blue callout box containing the number '1'. The main workspace shows the chemical structure of the target, also circled in red. A dialog box titled 'Tautomeric forms' is open, displaying five different representations of the chemical structure, each labeled with 'CAS# 577-71-9'. A red bracket underlines these five structures, with the text 'Tautomeric forms' centered below them. The dialog box includes 'Save to smi', 'Search', and 'OK' buttons.

Target with its tautomeric forms are identified and loaded on a data matrix.
 1. **Double click** over the target structure displays target and its tautomeric forms

Chemical Input

Multiplication a tautomeric set of already defined target (scenario 2)

The screenshot shows the QSAR Toolbox interface. In the 'Documents' panel on the left, the SMILES string Oc1ccc(N(=O)=O)c(N(=O)=O)c1 is selected (1). A right-click context menu is open over this entry, with 'Multiplication-Tautomerism' selected (2). On the right, a list of generated tautomers is shown, including the original SMILES and several variations (3).

- Select** the SMILES of the target chemical perform right click on it and then
- Select** Multiplication-Tautomerism
- Generated** tautomers appear in tree like form

Chemical Input

Implementation of Modeling modes:

- **Component Mode All** – all tautomers are analyzed as a package

The screenshot shows the QSAR Toolbox interface. The 'Component Mode' dropdown menu is highlighted with a red box, and the 'All' option is selected. The main window displays a chemical structure (a benzene ring with a hydroxyl group and a nitro group) and its associated data fields, including 'Substance Identity', 'Physical Chemical Properties', 'Environmental Fate and Transport', 'Ecotoxicological Information', and 'Human Health Hazards'. The 'Filter endpoint tree...' field contains the text '1 [target]' and '5 [T]'. The 'Structure' field contains the chemical structure.

- **Component Mode Single** – each tautomer is analyzed individually

The screenshot shows the QSAR Toolbox interface. The 'Component Mode' dropdown menu is highlighted with a red box, and the 'Single' option is selected. The main window displays a table of chemical structures representing different tautomeric forms. The table has six columns, each with a header: '1 [target]', '2 [target,tautomer]', '3 [target,tautomer]', '4 [target,tautomer]', '5 [target,tautomer]', and '6 [target,tautomer]'. The first column contains the chemical structure of the target, and the subsequent columns contain its tautomeric forms. A red oval highlights the entire table area.

Different modes for visualization of tautomeric sets.
A package of target and its tautomeric forms are used in further read across.

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.

Profiling

Side-Bar to Profiling

- For most of the profilers, background information can be retrieved by highlighting one of the profilers (for example, Protein binding by OASIS v1.4 and clicking on “View” (see next screen shot)).

Profiling Side-Bar to Profiling

2

1

1. Highlight the profiler
2. Click View
3. Select "Aldehydes"

Structural boundaries

Structural fragment

Protein binding by OASIS v1.4 (General Mechanistic) - Profiling Scheme Browser

Protein binding by OASIS v1.4 - Category definitions

- ROS generation and direct attack of hydroxyl radical to the C8 position of nucleoside bases
 - Heterocyclic Aromatic Amines
- ROS generation and protein carbonylation
 - Bipyridium Herbicides
- 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 aldehydes
 - Schiff base formation with carbonyl compounds
 - Acetylated Carbonyl compounds
 - Aldehydes**
 - alpha-ketoesters
 - Aromatic carbonyl compounds
 - Schiff base on pyrazolones and pyrazolidinones
 - Pyrazolones and Pyrazolidinones
- SE reaction (CYP450-activated heterocyclic amines)
 - Direct attack of arylnitrenium cation to the C8 position of nucleoside bases
 - Heterocyclic Aromatic Amines
- SN1
 - Carbenium ion formation (enzymatic)
 - Carbenium ion
 - DNA and protein alkylation via direct attack at carbonyl carbon atom and N-alkyl-N-nitrosocarbamates
 - DNA and protein alkylation via the formation of alkyl diazonium ion

Boundary Options Training set Options

NEW AND OR NOT

Clear Tidy Delete

Boundary Options Metabolism

Fragment

<m:c1(C(=O)c(=O)cccc(=O)1;m><m:c1(C(=O)c(O(=O)c(=O)c(C)cc1O(=O)1;m><m:c1(C(=O)1 Edit

Masked Fragments

Profiling

Profiling the set of target and tautomers

- For this example, the following profilers relevant to **skin sensitization** are used (see next screenshot):
 - Protein binding by OASIS v1.4
 - Protein binding by OECD
 - Protein binding potency
 - Protein binding for skin sensitization by OASIS v1.4
- 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.

Profiling

Profiling the set of target and tautomers
Profile statistic

The screenshot displays the QSAR Toolbox interface. On the left, the 'Profiling methods' panel is visible, with 'Protein binding by OASIS v1.4' selected. The 'Structure' panel shows a chemical structure of a target compound. The 'Filter endpoint tree...' panel lists various endpoints, with 'Protein binding alerts for skin sensitization by OASIS v1.4' selected. The 'Profile statistic' window is open, showing a table of categories and counts, a bar chart, and a 2D chemical structure representation. The table shows the following data:

#	Category	Count	%
1	Michael Addition+Michael Addition >> Michael Addition >> Michael addition on conjugated systems with electron withdrawing group >> alpha,beta-Carbonyl compounds with polarized double bonds +Michael Addition >> Michael addition on conjugated systems with electron withdrawing group >> Conjugated systems with electron withdrawing group	1	16.67
2	Michael Addition+Michael Addition >> Michael Addition >> Michael addition on conjugated systems with electron withdrawing group >> alpha,beta-Carbonyl compounds with polarized double bonds +Michael Addition >> Michael addition on conjugated systems with electron withdrawing group	1	16.67
3	Michael Addition+Michael Addition >> Michael Addition >> Michael addition on conjugated systems with electron withdrawing group >> alpha,beta-Carbonyl compounds with polarized double bonds +Michael Addition >> Michael addition on conjugated systems with electron withdrawing group	1	16.67
4	Michael Addition+Michael Addition >> Michael Addition >> Michael addition on conjugated systems with electron withdrawing group >> alpha,beta-Carbonyl compounds with polarized double bonds +Michael Addition >> Michael addition on conjugated systems with electron withdrawing group	1	16.67
5	Michael Addition+Michael Addition >> Michael Addition >> Michael addition on conjugated systems with electron withdrawing group >> alpha,beta-Carbonyl compounds with polarized double bonds +Michael Addition >> Michael addition on conjugated systems with electron withdrawing group	1	16.67
6	No alert found	1	16.67

The bar chart shows the distribution of target and its tautomers across the categories. The 2D representation shows the chemical structure of the target compound. Red annotations highlight the 2D representations and the number of tautomers in a category bin.

Four tautomeric forms of the target chemical have Protein binding alerts for skin sensitization: "Michael addition/Michael addition on conjugated systems with electron withdrawing group"

Profiling

Profiling the set of target and tautomers

Profile statistic

Profile statistic

Select profiles: Report Stop

Stack mode: Stacked to 100% Group by category

Profiles in use: Protein binding by OECD

#	Category	Count	%
1	Michael addition	5	83.33
2	Michael addition >> Polarisated Alkenes	5	83.33
3	Michael addition >> Polarisated Alkenes >> P...	4	66.67
4	Michael addition >> Polarisated Alkenes >> P...	4	66.67
5	No alert found	2	33.33

Bar chart: Protein binding by OECD

Count: 5, 4, 4, 2

Categories: Michael addition, Michael addition >> Polarisated Alkenes, Michael addition >> Polarisated Alkenes >> P..., No alert found

Michael addition >> Polarisated Alkenes >> Polarisated alkene - ketones

4 CAS# 577-71-9 (S) [T]

5 CAS# 577-71-9

6 CAS# 577-71-9

7 CAS# 577-71-9

Four of the tautomeric forms of the target have positive protein binding alert "Michael addition/Polarised alkenes"

Distribution of target and its tautomers across Protein binding by OECD

Profiling

Profiling the set of target and tautomers

Profile statistic

#	Category	Count	%
1	Extremely reactive (GSH)+Extremely reactiv	1	16.67
2	Extremely reactive (GSH)+Extremely reactiv	1	16.67
3	Moderately reactive (GSH)+Moderately react	1	16.67
4	Not possible to classify according to these r	3	50.00

Three of the five tautomers are "Reactive" by Protein binding potency

Protein binding potency categories of tautomers in the tautomeric set

Outlook

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

Endpoint

- “Endpoint” refer to the electronic process of retrieving the environmental fate, ecotoxicity and toxicity data that are stored in the Toolbox database.
- 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 sensitization from databases containing skin sensitization data
- Data for target chemical and its tautomeric forms is extracted from selected databases if available

Endpoint

- For this example, the following database are relevant to the skin sensitization(see next screen shot):
 - Skin sensitization
 - Skin sensitization ECETOC

Endpoint Gather data

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Endpoint' menu is open, showing options like 'Gather', 'Import', 'Export', 'Delete', and 'Tautomerize'. The 'Gather' button is circled in red and labeled with a red '3'. On the left, the 'Databases' list is expanded under 'Human Health Hazards', labeled with a red '1'. Several databases are checked with green boxes, labeled with a red '2'. The 'Structure' field on the right shows a chemical structure of a target molecule.

1. **Expand** the Human Health Hazards section
2. **Select** databases related to the target endpoint by adding a green check in the box before the database name.
3. **Click** Gather

Endpoint Gather data

The screenshot displays the QSAR Toolbox software interface. The top navigation bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. Below this, a secondary bar contains 'Gather', 'Import', 'Export', 'Delete', and 'Tautomerize'. The 'Gather' button is circled in red. On the left, the 'Databases' panel shows a tree view where 'Human Health Hazards' is expanded, and 'Skin sensitization' and 'Skin sensitization ECETOC' are checked and circled in red. The main workspace shows a 'Filter endpoint tree...' panel with a chemical structure of a substituted benzene ring. Below the tree, a data matrix entry '(1/1) M: Positive' is circled in red. A red box on the right contains the text: 'Positive experimental data appears on datamatrix'.

Endpoint Gather data

The screenshot displays the QSAR Toolbox interface with the 'Endpoint' tab selected. The main window shows a table of results for a chemical target. The table has columns for different target definitions: '1 [target]', '2 [target,tautomer]', '3 [target,tautomer]', '4 [target,tautomer]', '5 [target,tautomer]', and '6 [target,tautomer]'. The rows represent various endpoints, including 'Structure', 'Human Health Hazards', 'Profile', and 'Endpoint Specific'. A red circle highlights the 'M. Positive' result in the 'Human Health Hazards' section for target 1. A red box contains the following text:

Positive experimental data available for target cannot be explained with absence of protein binding alert. Apparently tautomeric analysis is needed to explain the conflict between the positive data and absence of protein binding alerts

Recap

- You have entered the target chemical
- You have simulate the tautomeric forms of target chemical
- You have profiled the tautomeric set of the target and identified no protein binding alert for the target. However, four tautomers have positive protein binding alerts
- You have gather data for chemical and its tautomeric forms and found positive experimental data for target.
- It is needed to verify the experimental data by searching for analogs having same functionalities
- Now you are ready to continue with next step of the workflow "Category definition".

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - **Category definition**

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**

Category Definition

Grouping methods

- For this example, the specific endpoint classification of target and its tautomers is identified by Organic Functional Group. Consistency of the category member is reached and phase I could be skipped (point 4 from performing categorization, slide #33).
- For this example initial group of analogues presented as tautomeric sets is identified by Organic Functional Group profiler
- Software search analogues presented as tautomeric sets having same protein binding distribution as those of the target tautomeric set

Category definition is a tool for grouping chemicals. 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>

Category Definition

Defining OFG

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Category Definition' menu is active, showing options like 'Define', 'Define with metabolism', 'Subcategorize', 'Combine', 'Clustering', 'Delete', and 'Delete All'. The 'Define' button is circled in red and labeled with a blue callout '2'. On the left, the 'Empiric' section is expanded, and 'Organic Functional groups' is highlighted with a blue callout '1'. A dialog box titled 'Organic Functional groups' is open, showing a list of functional groups: Allyl, Aryl, Cycto conjugated system, Cycloalkene, Cycloketone, Nitro alphatic, Nitro alphatic conjugated, Nitrobenzene, Phenol, and Quinoid compounds. The 'Aryl' group is circled in red and labeled with a blue callout '3'. The 'OK' button in the dialog is also circled in red.

1. **Highlight** the "Organic Functional Group"
2. **Click** Define
3. **Click** OK to confirm the defined categories for the tautomeric set

Category Definition

Defining OFG

The screenshot shows the QSAR Toolbox software interface. The main window is titled 'Category Definition' and displays a 'Filter endpoint tree...' on the left. The tree is expanded to show 'Organic Functional groups' under 'Human Health Hazards'. A 'Define category name' dialog box is open, showing 'Category name (6 chemicals)' and 'Quinoid compounds (Organic Functional groups)'. A blue callout box with the number '1' points to the 'OK' button. A text box at the bottom explains that the software identifies six chemical tautomers with the same protein binding alerts as the target set.

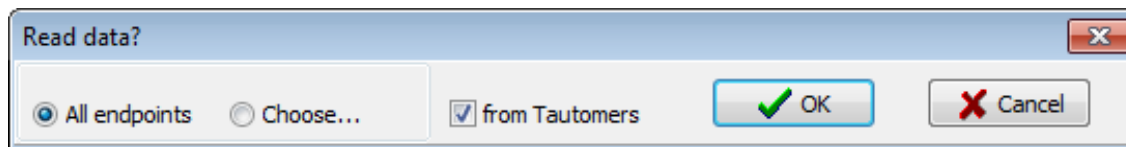
The software identify six chemical (presented as tautomeric set) having same protein binding alerts as the target set

1. **Select OK**

Category Definition

Read data for Analogues

- The Toolbox automatically request the user to select the endpoint that should be retrieved.
- The user can either select the specific endpoint or by default choose to retrieve data on all endpoints (see below).



- In this example, since only databases that contain information for human health hazards endpoints are selected, both options give same results.
- As the Toolbox must search the database, this may take some time.

Category Definition

Read data for Analogues

The screenshot shows the QSAR Toolbox software interface. The main window is titled 'Category Definition' and displays a 'Filter endpoint tree...' on the left and a grid of chemical structures on the right. A dialog box titled 'Read data?' is open in the center, with the following options:

- All endpoints
- Choose...
- from Tautomers
-
-

A blue callout box with the number '1' points to the 'OK' button. A red box highlights the text 'Read data for tautomeric sets'.

1. Click OK

Recap

- You have identified a category of analogues presented as tautomeric sets having same distribution of functional groups as the target tautomeric set
- The available experimental results for these 6 analogues have been collected from the selected databases (Skin sensitization and Skin sensitization ECETOC)
- But before the user can proceed with the “Filling Data Gap” module, he/she should navigate through the endpoint tree and find the specific gap that will be filled

Category Definition

Navigation through the endpoint tree

- The user can navigate through the data tree by closing or opening the nodes of the endpoint tree.
- The data tree is extensive but logically constructed; it can be mastered with a practice.
- In this example, the “EC3” is the target endpoint.
- You can navigate through the endpoint tree: Double-click on the node next to **Human Health Hazards** then effect **Sensitisation**, followed by **Skin**, type of method **In Vivo** and assay **LLNA** and finally **EC3** (see next screen shots)

Recap

- You have now retrieved the available skin sensitisation data for the four analogues represented by their tautomeric forms.
- You have identified the target endpoint of “Sensitization /Skin/In vivo/LLNA/EC3”.
- You are ready to fill in the data gap, so click on “Data Gap Filling” (see next screen shot).

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Category definition
 - **Data Gap Filling**

Data Gap Filling

Apply Read across analysis

The screenshot shows the QSAR Toolbox interface during a Data Gap Filling session. The top navigation bar includes tabs for Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. The left sidebar shows the 'Data Gap Filling Method' menu with 'Read-across' selected. The main workspace displays a 'Filter endpoint tree...' on the left and a table of target chemicals on the right. A 'Possible data inconsistency' dialog box is open, showing 'Skin sensitisation I (Oasis) (3 points)' and 'Skin sensitization EC3(ratio) (3 points)' selected. Red callouts and circles highlight key steps: 1. Highlighting the 'EC3' endpoint box in the table; 2. Selecting 'Read-across' in the sidebar; 3. Clicking the 'Apply' button in the top bar; 4. Clicking 'OK' in the dialog box.

1. **Highlight** the data endpoint box corresponding to "EC3" under the target chemical.
2. **Select** Read-across
3. **Click** Apply
4. **OK**

Data Gap Filling

Subcategorisation by Protein binding alerts for skin sensitization by OASIS v1.4

The screenshot shows the QSAR Toolbox software interface. The top navigation bar includes tabs for Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. The main workspace displays a table of chemical structures with their predicted target values (e.g., M. Positive, M. Negative) and a scatter plot of log Kow values. A red box highlights the 'Select/filter data' button in the right sidebar, and callout boxes 1, 2, and 3 point to specific actions: 1. Select/filter data, 2. Select protein binding alerts, and 3. Click Remove.

1. **Select** Select/Filter data and then **Subcategorization**
2. **Select** Protein binding alerts for skin sensitization by OASIS v1.4
3. **Click** Remove

Data Gap Filling

Result of Subcategorisation

The screenshot displays the QSAR Toolbox interface during the 'Data Gap Filling' process. The top navigation bar includes 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The left sidebar lists various grouping methods, with 'Endpoint Specific' selected. The main workspace shows a table of chemical structures and their predicted EC3 values. Below this, a scatter plot shows the distribution of log Kow values for 'Positive' and 'Negative' protein binding alerts. A yellow text box highlights that tautomeric sets of analogues have the same distribution of protein binding alerts as the target set.

Structure	EC3	Mode
1 [target]	(4/4)	M: Positive
2	(1/1)	M: Positive
4	(1/1)	M: Positive
6	(1/1)	M: Positive

Read across prediction of EC3, taking the highest mode from the nearest 5 neighbours, based on 3 values from 3 neighbour chemicals, Observed target value: 'Positive', Predicted target value: 'Positive'

The tautomeric sets of analogues have same distribution of Protein binding alerts as the target set

Data Gap Filling

Interpreting Read-across

- In this example, all results of the analogues are consistent; all the analogues are Skin sensitizers.
- The tautomeric sets of analogues have same distribution of Protein binding alerts as the target set.
- The same positive sensitising potential is therefore predicted for the target chemical.
- Accept the prediction by **clicking** “Accept prediction” (see next screen shot).

Data Gap Filling Result of Read-across

The screenshot displays the QSAR Toolbox interface during a read-across prediction. The main workspace shows a scatter plot of EC3 (obs.) versus log Kow. The plot title reads: "Read across prediction of EC3, taking the highest mode from the nearest 5 neighbours, based on 3 values from 3 neighbour chemicals, Observed target value: 'Positive', Predicted target value: 'Positive'". The plot shows a single red data point at approximately (-0.8, Positive). An information dialog box is open, stating "The current prediction was accepted" with an "OK" button. The right sidebar contains a list of options, with "Accept prediction" and "Return to matrix" buttons circled in red. A callout box with the number "1" points to these buttons. Another callout box with the number "2" points to the "OK" button in the dialog box. A third callout box with the number "3" points to the "Subcategorize" button in the sidebar.

1. **Select** Accept prediction
2. **Click OK** and then 3. **Click** Return to matrix

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Category definition
 - Data Gap Filling
 - **Report**

Report

The screenshot shows the QSAR Toolbox software interface. The top navigation bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. Below this is a 'Filter endpoint tree...' panel on the left with a tree view of various endpoints. A central table displays chemical structures and their predicted values for six different targets. A context menu is open over the 'Report' prediction in the table, with callouts '1' and '2' indicating the steps: '1. Select prediction' and '2. Right Click and Select Report'.

Structure	1 [target]	2	3	4	5	6
<chem>O=C1C=CC(=O)N1</chem>	[5] [7]	[11] [7]	[5] [7]	[11] [7]	[12] [7]	[11] [7]
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 Chemico						
In Vitro						
In Vivo						
GPMT						
LLNA						
EC3						
Undefined Assay						
ToxCast						
Toxicity to Reproduction						
Toxicokinetics, Metabolism and Dist						
	M: Positive R: Positive	IUCLID5	negative	M: Positive	M: Positive	M: Positive

1. Select prediction
2. Right Click and Select Report

Report

The screenshot displays the QSAR Toolbox software interface. The main window shows a report titled "QSAR Toolbox prediction based on read-across" for the "Prediction of EC3 for 3,4-dinitrophenol". A callout box with the number "1" points to the "Summary" section of the report, which is circled in red. The summary text describes the prediction of toxicity based on read-across from category members. Below the text is a table with two columns: "Endpoint(s)" and "Descriptor(s)".

Endpoint(s)	Descriptor(s)
Human Health Hazards; Sensitisation	log Kow

1. Summary information for the prediction of tautomeric set

Report

1 j. Predicted value (model result):
Positive

QSAR Toolbox 3.4.0.17
Database version: 3.8.8/3.1.2

QSAR TOOLBOX TPRF v.3.4.1.34101

Prediction of EC3 for 3,4-dinitrophenol 8 / 47

2 k. Predicted value (comments):
NOT provided by the user

4.3. Applicability domain (OECD Principle 3):
The target set of tautomers FALLS within applicability domain
(see Section 3.1.b for detailed description of the domain)

1 2 3 4 5
6 7
8

AND AND AND

1. Predicted value 2. Applicability domain

Report

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The left sidebar shows 'Available data to report' (Predictions, QSARs, Categories) and 'Available report templates' (Standard and Custom). The main window displays a prediction report for 3,4-dinitrophenol. The report title is 'Prediction of EC3 for 3,4-dinitrophenol' and the section is 'APPENDIX 7 - Chemical components'. A red oval highlights the title, and a red box highlights the text 'used in prediction'.

Prediction of EC3 for 3,4-dinitrophenol 20 / 47
Appendix 7 - Chemical components

QSAR Toolbox prediction based on read-across

Prediction of EC3 for 3,4-dinitrophenol

APPENDIX 7 - Chemical components

Tautomer No.1 of target chemical: **used in prediction**

- 1. CAS number:**
577-71-9
- 2. Other regulatory numbers:**
Not reported
- 3. Chemical name(s):**
Not available

Additional Appendix 7 list tautomers of target and analogue chemicals used in read-across
 Also an information about which tautomer is used in the RA prediction is provided.

Outlook

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

Saving the prediction result

- This functionality allow storing/restoring the current state of Toolbox documents including loaded chemicals, experimental data, profiles, predictions etc, on the same computer. The 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/Loading the file with TB prediction is shown on next screenshots

Saving the prediction result

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Report' button is highlighted with a callout box labeled '1'. Below the menu bar, there are buttons for 'Create', 'Print', 'Close', 'Save as', 'Register', 'Unregister', 'Update', 'Clone', and 'Design'. The left sidebar shows 'Available data to report' and 'Available report templates'. The main window displays a list of chemical categories, including 'phenol' and 'components'. A 'Save As' dialog box is open, showing the file path 'Data (D:) > Backup > TOOLBOX 4.0 > Toolbox 3.4 > 5.07.2016'. The file name is 'Tutorial 15.tbw' and the save type is 'Toolbox work file (*.tbw)'. Callout boxes labeled '2' and '3' point to the file name field and the 'Save' button, respectively.

1. Click on Save button; **2. Define** name of the file; **3. Click** Save button

Open saved file

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Input' menu is open, showing 'Open' as the selected option. A 'Select file' dialog box is open, showing a file tree with 'Tutorial 16.tbw' selected. The 'File name' field contains 'Tutorial 16.tbw' and the file type is set to 'Toolbox work file (*.tbw)'. The 'Open' button is highlighted. The background shows a chemical structure and a table of results.

4	5	6
<chem>Nc1ccc(O)cc1</chem>	<chem>Nc1ccc(O)cc1</chem>	<chem>Nc1ccc(O)cc1</chem>
M: Positive	M: Positive	M: Positive
Michael addition > ...	Michael addition > ...	Michael addition > ...
Michael addition > ...	Michael addition > ...	Michael addition > ...
Michael addition > ...	Michael addition > ...	Michael addition > ...
Michael addition > ...	Michael addition > ...	Michael addition > ...
No alert found	Michael addition	Michael addition
Michael addition > ...	Michael addition > ...	Michael addition > ...
Michael addition > ...	Michael addition > ...	Michael addition > ...
Michael addition > ...	Michael addition > ...	Michael addition > ...
Michael addition > ...	Michael addition > ...	Michael addition > ...
No alert found	No alert found	No alert found
Extremely reactive...	Extremely reactive...	Extremely reactive...

Once the file has been saved **1. Go** to Input; **2. Click** Open; **3. Find** and **select file**; **4. Click** Open