

Computational approaches to support chemical hazard identification: application of the QSAR Toolbox software to case studies

Chiara Laura BATTISTELLI

Environment and Health Department-

Istituto Superiore di Sanità

Content

1. Introduction to *in silico* methods, focusing on QSAR approaches (such as models, category/grouping, read across, TTC)
2. Introduction to the QSAR Toolbox software:
 - ✓ general introduction (main functionalities)
 - ✓ step by step, with practical use of the tool
3. Practical exercise on case studies with QSAR Toolbox software (in working groups)

Tuesday 17.10.2023 – Marotta Room

9:00 – 10:00	Q&A Case studies of the previous day (Trainers: Dr. Emanuela Testai, Dr. Maura Manganelli, Dr. Simona Scardala Scardala)
10.00 – 11.30	Introduction to <i>in silico</i> methods: QSAR, read across, TTC, and the related tools (QSAR Toolbox) characterization (Trainers: Dr. Chiara Battistelli, Dr. Cecilia Bossa, Dr. Olga Tcheremenskaia)
11.30 – 13.00	Presentation of tool functioning (Trainers: Dr. Chiara Battistelli, Dr. Cecilia Bossa, Dr. Olga Tcheremenskaia)
13.00 – 14:30	Lunch
14.30 – 16:30	Practical exercise on case studies, in working groups with QSAR Toolbox software (Trainers: Dr. Chiara Battistelli, Dr. Cecilia Bossa, Dr. Olga Tcheremenskaia)
16.30 – 17:00	Discussion of exercises

Toxicological studies (for hazard identification)



Animal testing

in vivo assays

Alternative methods*

(Non animal models)

Experimental models

(in vitro, ex vivo):

cells, organs, tissues or enzymes



Non testing models (computational or in silico)

- Speed-up number of evaluated chemicals
- Save of money and time
- Animal welfare



* According the 3Rs principles: *Refinement, Reduction, Replacement* (Russel and Burch, 1960)

(Q)SAR approaches

(Q)SAR

(Quantitative) Structure-Activity Relationship



- ❖ Methodologies based on the concept that a property (such as toxicity) is a function of chemical structure
- ❖ Used to predict (physical chemical or (eco)toxicological) properties, on the basis of chemical structure (Known)
 - **QSAR models:** quantitative structure–activity relationship models. Mathematical equation linking the biological activity to chemical structure, identified by **molecular descriptors** (i.e. physical chemical or other molecular properties)
 - **SAR models:** qualitative structure–activity relationship models (i.e. Structural Alerts)
 - **Chemical grouping and categories:** approach considering more than one chemical at the same time, sharing similar characteristics and physical chemical or (eco)toxicological properties or following a trend, as results of structural similarity**.
 - **Read-Across:** technique used for filling data gaps by predicting endpoint information for one (or more) chemical(s) (*target chemical*), using data for the same endpoint from one or more similar substances (*source chemicals*).

Good-quality data

Threshold of Toxicological Concern (TTC): applied to derive a threshold for exposure, below which a toxic effect on human health by the compound is not expected; based on databases on general toxicity

(*) **Structural alerts (SAs):** functional groups or structural features associated with a potential reactivity for a defined endpoint

(**) **Chemical similarity** is not limited to structural similarity but should consider factors that drive a given toxicity and how these can be linked back to chemical properties or features, e.g. reactivity."

(Q)SAR models

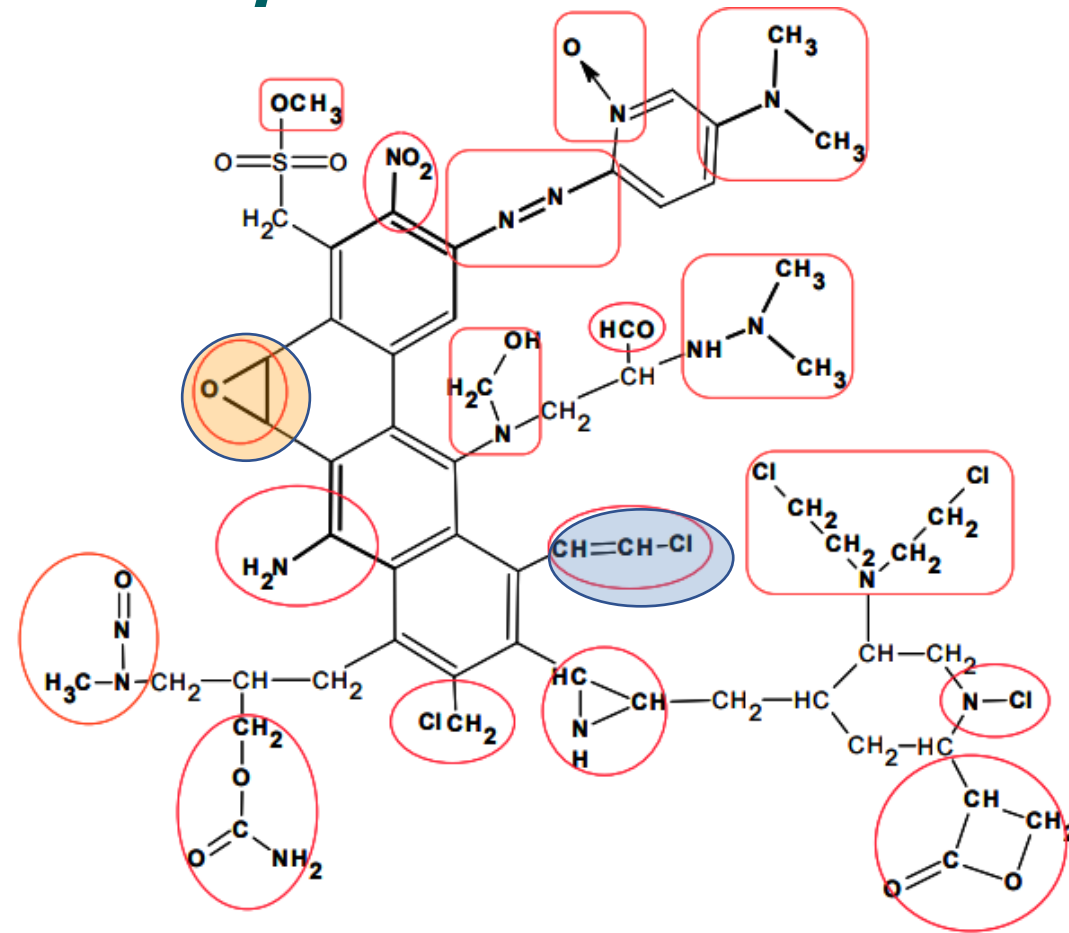
- **Local models:** built on congeneric set of chemicals, i.e., chemicals with similar structure with the same mechanism of action; generally more effective, but with a narrower domain of applicability (strictly related to the set of chemicals used for the model definition)
- **Global models:** often implemented in software tools, ensure a broader applicability, to more than one chemical class.
 - ✓ **Statistical-based:** use machine learning techniques to associate structural features and chemicals activity. These models are data-driven, without expert supervision.
 - ✓ Negative predictions are more accurate.
 - ✓ **Rule-based (expert system):** recognition and codification of functional groups or structural features associated with a potential reactivity for a defined endpoint (**Structural alerts, SA**). SA are rarely defined by an applicability domain (absence of SAs, is not absence of toxicity, but a lack of knowledge)
 - ✓ **Hybrid:** integrate both expert knowledge and statistically derived rules, trying to overcome disadvantages of both approaches.

Structural alerts (SAs): functional groups or structural features associated with a potential reactivity for a defined endpoint

SAR models: the case of SAs for genotoxicity

1. Genotoxic carcinogens: DNA reactivity, direct or after metabolic activation
2. The basis of the alerts: Miller¹ theory, electrophilic species are able to react with nucleophilic sites in DNA
3. Ashby² compiled a list of SA (based on experimental data), which if present in a molecule, give it potential reactivity:
 - Direct acyls (acyl halides, β -lactones..)
 - Direct alkylants (epoxides and aziridines, lactones, nitrites, α - β unsaturated carbonyls, simple aldehydes, quinones..);
 - Indirect alkylating agents (mono-halogen alkene, hydrazines...).
 - Intercalants and forming adducts (PAHs, aromatic hydrocarbons...)
 - Amino aryls that form adducts (aromatic amines, amides, ..)

1. Miller and Miller, 1981; 2. Ashby et al, 1988

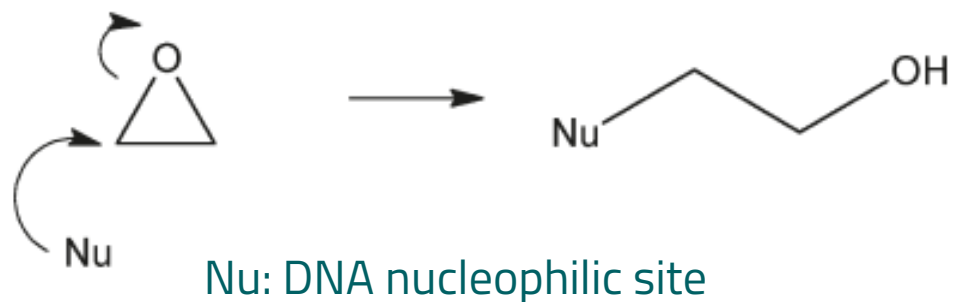


An effective representation → the polycarcinogen (imaginary molecule)

Each SA codes for a chemical class → specific mechanism of action

Mechanism of direct and indirect alkylants

Epoxide: Direct alkylating agent



**CHEMICAL
REVIEWS**

[dx.doi.org/10.1021/cr100222q](https://doi.org/10.1021/cr100222q) | *Chem. Rev.* 2011, 111, 2507–2536

REVIEW

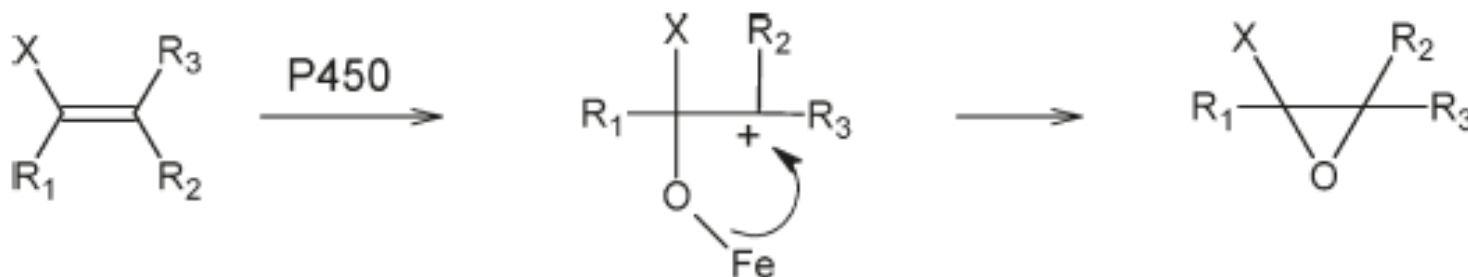
pubs.acs.org/CR

Mechanisms of Chemical Carcinogenicity and Mutagenicity: A Review with Implications for Predictive Toxicology

Romualdo Benigni* and Cecilia Bossa

Istituto Superiore di Sanita', Environment and Health Department, Viale Regina Elena, 299 00161 Rome, Italy

Mono halogen alkene: indirect alkylant



Grouping of substances and Read-across

	Chemical 1	Chemical 2	Chemical 3	Chemical 4
Structure	xxxxxxxx	xxxxxxxx	xxxxxxxx	xxxxxxxx
Property 1	● → ○	● → ○	● → ○	● → ○
Property 2	● → ○	○ ← ●	○ ← ●	○ ← ●
Property 3	○ ← ●	○ ← ●	○ ← ●	○ ← ●
Activity 1	● → ○	● → ○	● → ○	● → ○
Activity 2	● → ○	○ ← ●	○ ← ●	○ ← ●
Activity 3	○ ← ●	○ ← ●	○ ← ●	○ ← ●

● Existing data point ○ Missing data point

SAR/Read-across

Interpolation

Extrapolation

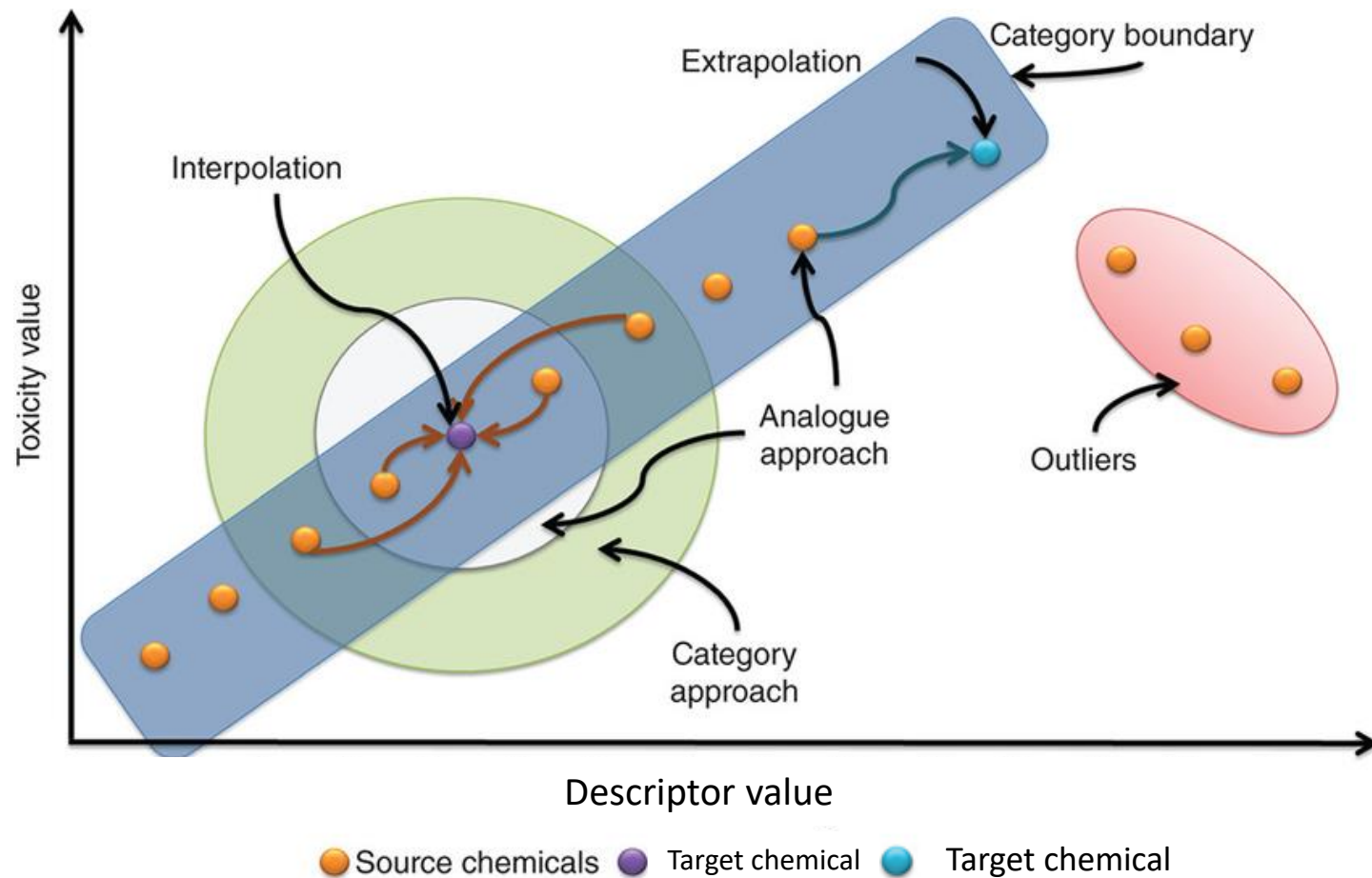
SAR/Read-across

Interpolation

Extrapolation

- Substances that have similar properties may be considered as a **group**, or **category**
- If we have data for one, or more chemicals (**source chemicals**), they can be read-across to fill the data gaps for substances with no data (**target substance**)
- Structural similarity is a pre-requisite (for any grouping and read-across approach).
- Similarities may be due to different factors such as:
 - common functional group;
 - common precursors and/or common breakdown products
 - constant pattern in the changing of the potency of the properties across the group

Grouping /category and Read-across



Category approach: chemicals sharing similar characteristics and (eco)toxicological properties are grouped together. The assessment concerns the category as a whole because data gaps may exist for different category members and different endpoints.

Analogue approach: assessment of one specific chemical, using for the prediction experimental data from others (one or more) similar substances

Trend analysis is a method of predicting toxicity of a chemical by analyzing toxicity trends (increase, decrease, or constant) of tested chemicals.

Source chemicals: substances with experimental data, considered similar to the target

Target substance: substance to predict (with no experimental data)

Raies and Bajic, 2016, doi: 10.1002/wcms.1240

Threshold of Toxicological Concern (TTC)

- Pragmatic approach used in risk assessment, in the absence of chemical-specific toxicity data
- Used to derive a threshold for exposure, below which a toxic effect on human health is not expected
- Based on a general toxicity database (for oral exposure)*
- Used in some regulatory framework (such as EMA, FDA, EFSA, ECHA, SCCS, ...) if exposure is low
- TTC is not strictly a QSAR approach, but a non testing methods which uses structure-activity relationships

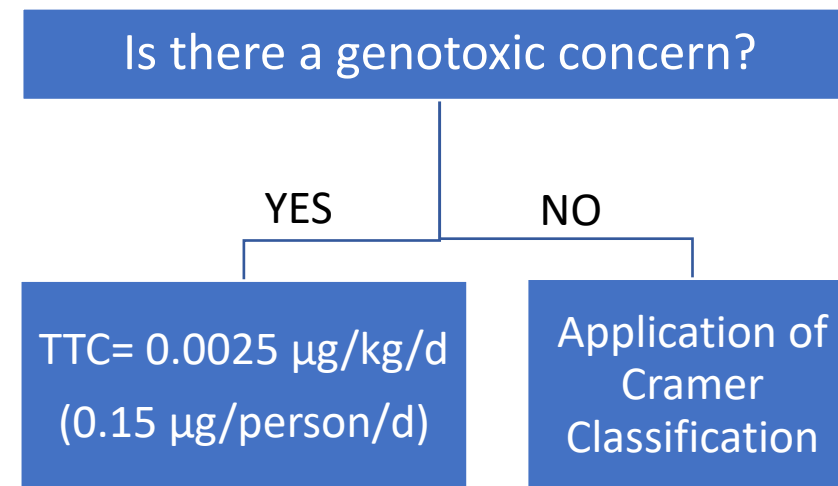
Domain of applicability:

- Applicable to organic compounds with known chemical structure
- Not applicable (or recommended) to high potency carcinogens and highly bioaccumulating substances, inorganics, metals, metal containing compounds, polymers, proteins, nanomaterials...

*Munro et al 2008; Kroes et al 2004; Batke et al 2021

TTC: procedure

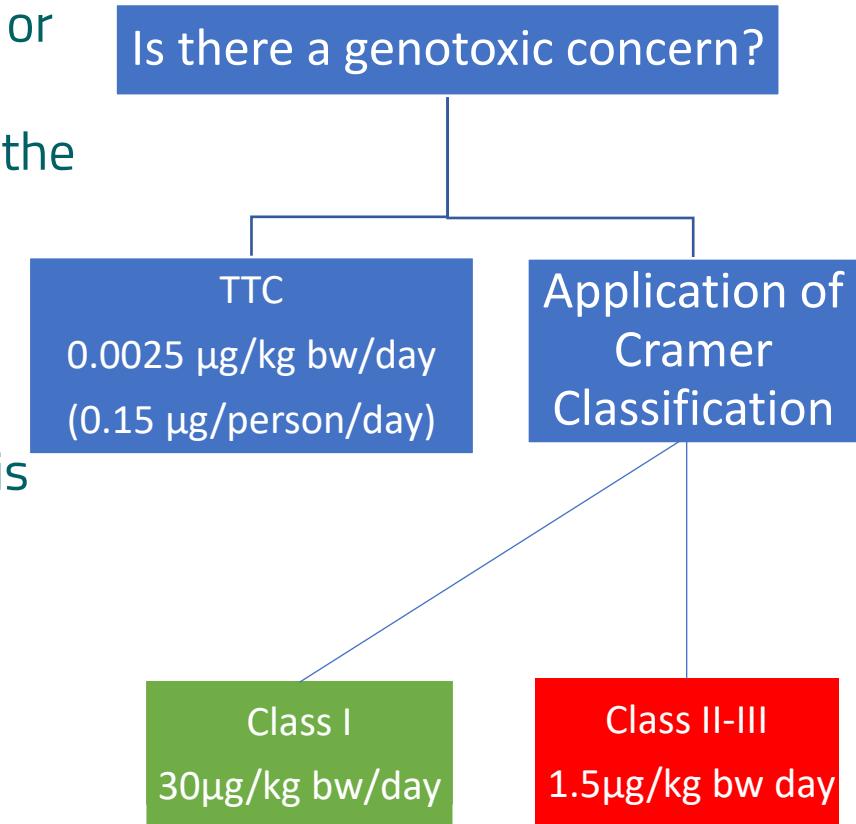
- 1) The first step is to assess the genotoxicity by evaluating all available data, and/or applying (Q)SAR approaches in a WoE (Weight of Evidence Approach: integration of different source of data, experimental and estimated data, with an expert supervision)
- 2) For the potential DNA-reactive mutagens the TTC value of $0.0025 \mu\text{g/kg/d}$ ($0.15 \mu\text{g/person/day}$) is considered
- 3) If No genotoxic concern is assigned, to the substance, the Cramer classification scheme is applied



Munro et al 1996; Kroes et al 2004; Batke et al 2021

Cramer classification scheme

- Decision tree is used to categorize non genotoxic chemicals
 - The original Cramer decision tree consists of “YES” or “NO” questions or rules (including extension and modification respect the original scheme)
 - The answer to each question leads to a final Cramer classification for the chemical in one of three classes:
 - ✓ Class I – low toxicity
 - ✓ Class II – intermediate toxicity
 - ✓ Class III – high toxicity
- } together
- Once the Cramer class is determined, a corresponding TTC threshold is chosen: if the chemical is below the TTC threshold, the toxic effect on human health is not expected
 - **QSAR Toolbox** can be used to assign the Cramer class (original and extended version)
 - It is a crucial step because the interpretation of each rule may vary
 - ✓ expert judgment is often needed



*Cramer et al., 1978; Munro et al 1996

In which regulatory framework (Q)SAR can be used?



Pesticides: Panel on
Plant Protection
Products and Their
Residues

APPROVED: 12 March 2019
doi:10.2903/sp.efsa.2019.EN-1598

Evaluation of the applicability of existing (Q)SAR models for predicting the genotoxicity of pesticides and similarity analysis related with genotoxicity of pesticides for facilitating of grouping and read across

Romualdo Benigni*, Chiara Laura Battistelli**, Cecilia Bossa**, Alessandro Giuliani**, Elena Fioravanzo***, Arianna Bassan****, Mojca Fuart Gatnik***, James Rathman****, Chihai Yang***** and Olga Tcheremenskaia**

REACH (EU Regulation on the **R**egistration, **E**valuation and **A**uthorisation (restriction) of **C**hemicals). In silico models can be used for:

- Risk assessment
- Classification
- Prioritization



Pharmaceuticals:
Assessment and control
of DNA reactive (mutagenic)
impurities in pharmaceuticals

Cosmetics:
Regulation (EC) No
1223/2009 SCCS
safety evaluation of
cosmetic ingredients
(SCCS1647/22)



Scientific Committee on Consumer Safety

SCCS

THE SCCS NOTES OF GUIDANCE FOR THE TESTING OF
COSMETIC INGREDIENTS AND THEIR SAFETY EVALUATION

12TH REVISION



The SCCS adopted this guidance document
by written procedure on 15 May 2023

REACH: EU Regulation on the **R**egistration, **E**valuation and **A**uthorisation (restriction) of **C**hemicals

- Industries are responsible for the **safe use of the chemicals** they produce or import
- **Hazard identification** of substances is the starting point: all the information are needed to be reported in registration dossiers (IUCLID format).
- The higher the tonnage, the more information is required.
- **Reach** requires **vertebrate** testing as a **last resort** and provides and promotes:
 - ✓ **Data sharing**
 - ✓ **Use of alternative methods, such as (Q)SAR, Read-Across e chemical categories**Annex XI reports criteria for (Q)SAR regulatory acceptance (reliable results, fit for purpose, adequate justified)

To increase the use of in silico methods, improving their regulatory acceptance, ECHA and OECD have funded the QSAR Toolbox software

QSAR Toolbox

- ❖ The QSAR Toolbox is a free software application that supports (eco)toxicologists in performing **reproducible** and **transparent** chemical hazard assessment using non-animal methods.
- ❖ ECHA e OECD are co-owners (Ver 4.6), developed (IT) and managed by «Laboratory of Mathematical Chemistry» (LMC)
- ❖ Website: www.qsartoolbox.org
- ❖ Resources: tutorials, manuals, webinar, helpdesk (on line support), forum (public discussion), ontologies
<https://qsartoolbox.org/support/>

Version 4.6, 2022

QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

Loading resources...

In collaboration with

OECD
BETTER POLICIES FOR BETTER LIVES

ECHA
EUROPEAN CHEMICALS AGENCY

Developed by

LMC
Laboratory of Mathematical Chemistry

WHAT IS THE QSAR TOOLBOX?

The Toolbox is a free software application that supports reproducible and transparent chemical hazard assessment. It offers functionalities for retrieving experimental data, simulating metabolism and profiling properties of chemicals. These information and tools can be used to find structurally and mechanistically defined analogues and chemical categories, which can serve as sources for read-across and trend analysis for data gap filling.

$$\text{O}_3 + 2 \text{B}(\text{NO}_3)_3 + 6 \text{HBr}$$

DOWNLOAD LATEST VERSION

QSAR Toolbox: management and contributors

- ✓ ECHA/OECD: co-owner and co-developer
- ✓ LMC: developer, under the OECD and ECHA umbrella
- ✓ LMC/ECHA/OECD: **"QSAR Toolbox Coordination Group"**
- ✓ **"QSAR Toolbox Management Group"**: coordination group + experts from industry, authorities , NGOs
 - ✓ Discuss, approve and test software developments
- ✓ Third parties contribute with data, profiler and experience
- ✓ Excellent collaboration between such different entities, such as academic, regulatory and industrial

Supporters or donors (data, profilers and experience)

- | | |
|-------------------------------|------------------------|
| ▪ OECD | ▪ EFSA |
| ▪ ECHA | ▪ Univ of California |
| ▪ LMC | ▪ Cefic |
| ▪ EC | ▪ OASIS |
| ▪ EURL ECVAM | ▪ L'Oreal |
| ▪ US EPA | ▪ DuPont |
| ▪ Environment Canada | ▪ Givaudan |
| ▪ Health Canada | ▪ Dow chemicals |
| ▪ NITE Japan | ▪ BASF |
| ▪ NIES Japan | ▪ ExxonMobil |
| ▪ Danish EPA | ▪ 3M |
| ▪ UBA Germany | ▪ Firmenich SV |
| ▪ NICNAS Australia | ▪ SRC, Syracuse |
| ▪ DEWNA Australia | ▪ Unilever |
| ▪ ISS Italy | ▪ Multicase |
| ▪ Fraunhofer Germany | ▪ ChemAxon |
| ▪ BfR Germany | ▪ ECETOC |
| ▪ Ministry of the Envir Japan | ▪ Mario Negri (Milano) |
| ▪ MHLW Japan | ▪ |
| ▪ INERIS | |

QSAR Toolbox: helps reduce animal testing

- ❖ **Prevent duplication of animal tests:** when high quality data are found, there is no need to duplicate the test.
- ❖ **Intelligent testing strategies:** by forming categories and identifying data gaps, informed testing strategies can be designed to optimize costs and number of animals required.
- ❖ **Predict toxicity using a category approach:** the Toolbox results can be used for data-gap filling and as supporting evidence for read-across cases.
- ❖ **Sustainable development and green chemistry:** the toxicity of substances can be predicted even before they are produced, facilitating sustainable product development and green chemistry

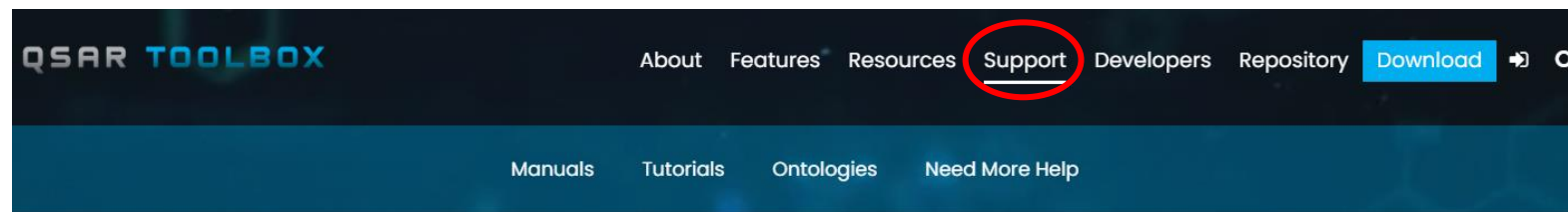
QSAR Toolbox: what is the users expertise?

- **IT Technical Skills:** decreased in recent versions, because of the simplified User Interface o Automatic workflow
- **Expertise in Toxicology:** in risk assessment procedure and in the endpoints to predict (e.g. data quality assessment)
- **Expertise in organic and computational chemistry:** to justify and interpret similarity between analogues, to evaluate the QSAR prediction,...

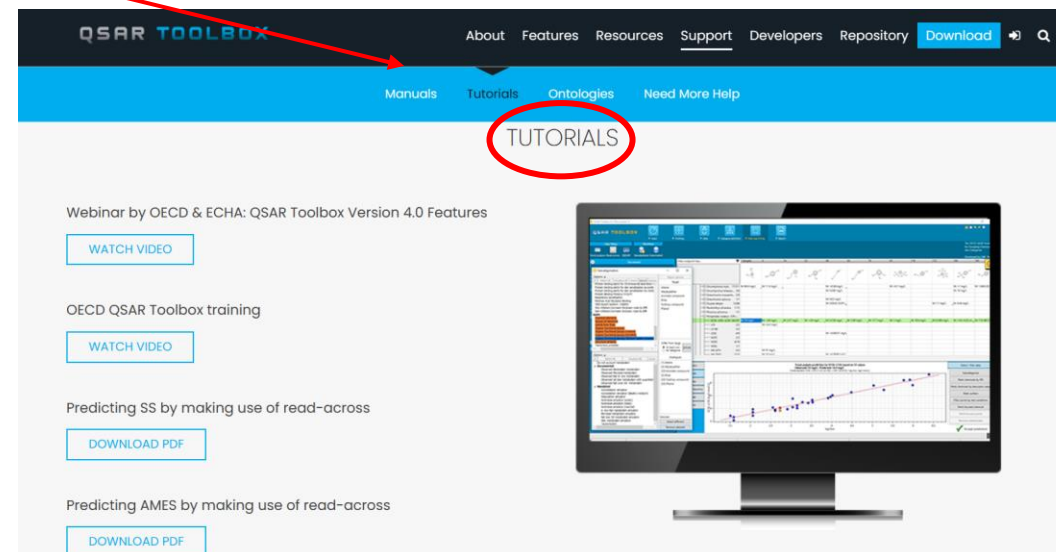
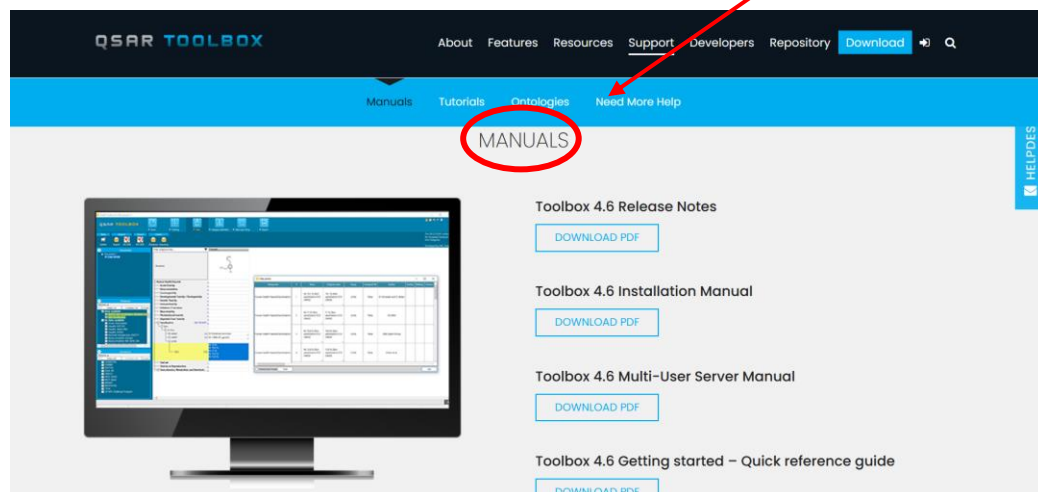
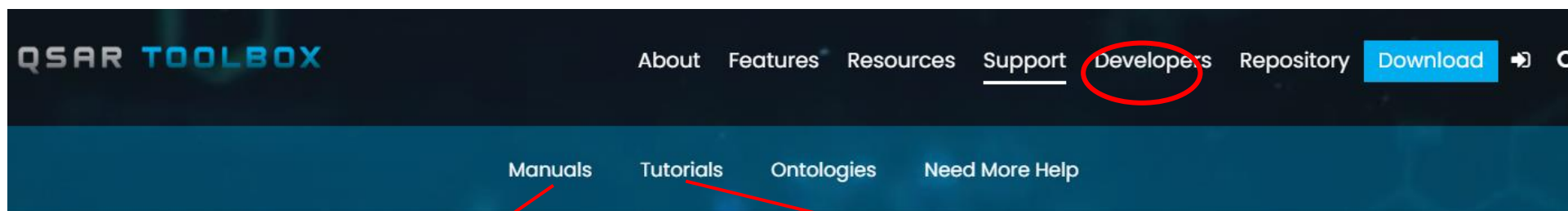
Support section

<https://qsartoolbox.org/support/>

Manuals (installation and user manuals), tutorials (training), ontologies (Controlled toxicological vocabularies and interrelations), more help (helpdesk)...

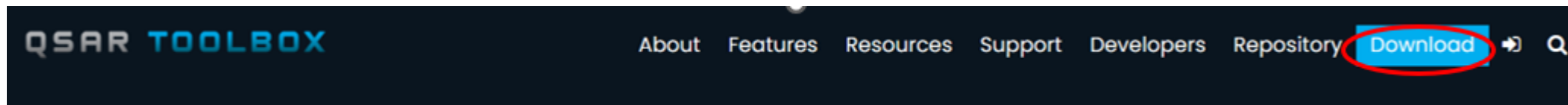


QSAR Toolbox: support



<https://qsartoolbox.org/>

QSAR Toolbox: download



INSTALLATION PACKAGE

QSAR Toolbox Installer v.4.6

DOWNLOAD INSTALLER

RELEASE NOTES

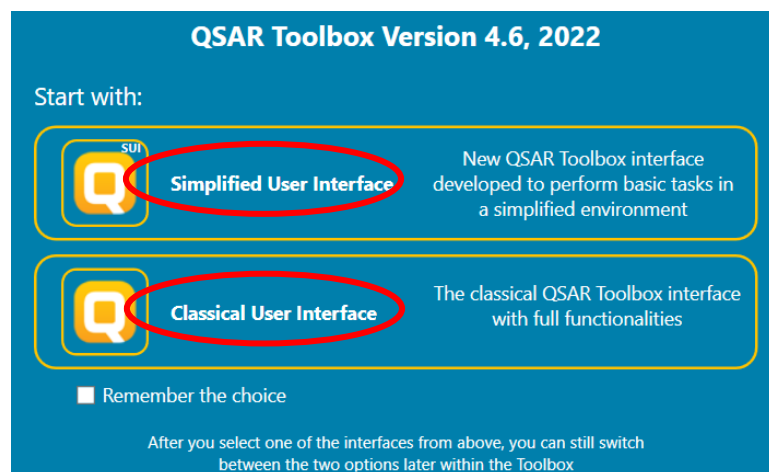
<https://qsartoolbox.org/>

Registration to the web account is required

QSAR Toolbox: interfaces

- ✓ **Simplified user interface***: Easy to use but includes only simple functionalities*
- ✓ **Classical user interface***: Main interface. Includes all functionalities but needs training.
- ✓ **Web client**: Latest interface. Runs on all operating systems

*Only runs on Windows



Modules on principal workflow

Functionality refers to the selection of the first row

Data matrix

Structure info	
Additional Ids	
CAS Number	122-04-3
CAS-SMILES relation	High
Chemical name(s)	4-Nitrobenzoyl chloride
Composition	
Molecular formula	C7H4ClNO3
Predefined substance type	Mono constituent
SMILES	[O-][N+](=O)c1ccc(cc1)C(Cl)=O
Parameters	
Physical Chemical Properties	1/1 M: 75 °C
Environmental Fate and Transport	
Ecotoxicological Information	1/3 M: 5 ppm

QSAR Toolbox: Simplified user interface (SUI)

With the simplified user interface the user can:

- Collect data
- Apply profiling
- Find analogues

The screenshot displays the 'Simplified User Interface' window. It features a top navigation bar with a 'Define goal' button. The main area is divided into three panels: 'Input chemical', 'Select your goal', and 'Help panel'. The 'Input chemical' panel on the left contains a text box with the instruction 'Please start by defining an Input chemical.' and a 'Define' button. Below it is a 'Target Endpoint' section with a dropdown menu set to 'Select (optional)'. The central 'Select your goal' panel offers three main options: 'Collect data' (with 'All data' and 'Specific data' buttons), 'Apply profiling' (with 'All profilers', 'Specific profilers', and 'With metabolism' buttons), and 'Find analogues' (with 'By functional groups' and 'By other criteria' buttons). The 'Help panel' on the right provides detailed instructions for each step, explaining how to input a chemical, select a target endpoint, and choose specific data, profilers, or criteria.

Define goal

Input chemical

Please start by defining an Input chemical.

Define

Target Endpoint

Select (optional)

Select your goal

Collect data

All data Specific data

Apply profiling

All profilers Specific profilers With metabolism

Find analogues

By functional groups By other criteria

Help panel

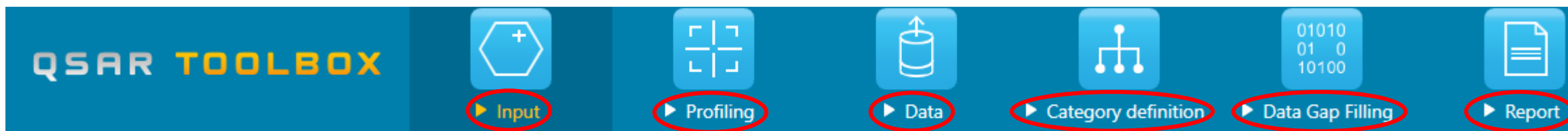
The simplified user interface facilitates the execution of conceptually simple tasks within the Toolbox.

Input a chemical and, if relevant, a target endpoint. Then, choose one of the goals and follow the instructions.

Collect data
"All data" – collect all data available for the target chemical from all databases
"Specific data" – requires selection of endpoints for which data to be collected. The desired endpoint could be either defined as a target endpoint or selected from the endpoint tree position. Based on this selection the relevant databases are displayed.

Apply profiling
"All profilers" – applies all profilers for the target chemical
"Specific profilers" – requires selection of profilers for application from the list with all profilers. If the target endpoint is defined, this selection is controlled by the defined target endpoint (e.g. for a defined target endpoint related to EC3/LLNA/Skin sensitization, only the relevant to this endpoint profilers will be selected).
"With metabolism" – requires selection of profilers and metabolic simulators which to be applied in combination. If target endpoint is defined, this selection is controlled by

QSAR Toolbox: key functionalities



The Toolbox consists of a logical and sequential workflow, with the following modules:

- 1. Input:** starting point and provides different ways to specify the identity of the target substance and the property under consideration.
- 2. Profiling:** contains the knowledge coded in profiling schemes (**profilers**). The profilers identify the affiliation of the target chemical(s) to categories (functional groups/alerts), and include **observed** and **simulated metabolism** and transformation
- 3. Data:** include all the data of TB databases. Data can contain chemical information (CAS, name, SMILES), experimental data and supporting information (metadata)
- 4. Category definition:** used to group chemicals into a toxicological **category**, according to structural or mechanistic similarity, to be used in read-across or trend analysis
- 5. Data Gap Filling:** used to fill a data gap using data from analogues with **trend analysis, read-across or existing QSAR models**.
- 6. Report:** produce a report for prediction, export the chemicals on the data matrix and related information

Modules do not necessarily have to be used sequentially

QSAR Toolbox: keywords

- ✓ **Target chemical:** chemical of interest
- ✓ **Module:** section dedicated to a specific action and option (6 modules: input, profiling,...)
- ✓ **Workflow:** the use, in combination, of the different modules (prediction workflow, from input to report)
- ✓ **Profiler:** algorithm (rules set) for the identification of specific features of the chemicals: structural (i.e. *organic functional groups*), general mechanistic (i.e. *protein binding by OECD*), endpoint specific (i.e. *in vitro mutagenicity alerts by ISS*)
- ✓ **Category:** group of substances sharing the same characteristics (e.g. same functional group or mode of action). In a typical Toolbox workflow it consists of the target chemicals and its analogues gathered according to the selected profilers
- ✓ **Endpoint tree:** branched tree scheme, from a broader level (Phys-Chem prop, Environmental fate and transport, Ecotoxicology, Human Health) to a more detailed one (e.g. *in vitro* or *in vivo* assay, species and other metadata)
- ✓ **Data matrix:** table reporting the chemical(s) and data (experimental results, profilers outcomes, predictions). Each chemicals is in a different column, each data in a different row

QSAR Toolbox

2. Modules and 3. Workflow

1. Target chemical
2. Module
3. Workflow
4. Profiler
5. Document tree
6. Category
7. Endpoint tree
8. Data matrix

The screenshot shows the QSAR Toolbox interface with several components highlighted by red boxes and numbered annotations:

- 1. Target chemical:** Points to the chemical structure of 1,3-dinitrobenzene in the 'Structure' field.
- 2. Module:** Points to the 'Input' module icon in the top toolbar.
- 3. Workflow:** Points to the 'Profiling' module icon in the top toolbar.
- 4. Profiler:** Points to the 'Profiling methods' panel on the left, showing selected methods like 'DNA binding by OASIS' and 'DNA binding by OECD'.
- 5. Documents tree:** Points to the 'Documents' panel on the left, showing a list of documents including '[C: 4;Md: 60;P: 0] Acyl halides<AND>Nitro'.
- 6. Category:** Points to the 'Filter endpoint tree...' panel on the right, showing a tree structure for 'Genetic Toxicity'.
- 7. Endpoint tree:** Points to the 'Genetic Toxicity' section in the 'Filter endpoint tree...' panel.
- 8. Data matrix:** Points to the table on the right showing results for four target chemicals across various endpoints.

8. Data matrix:

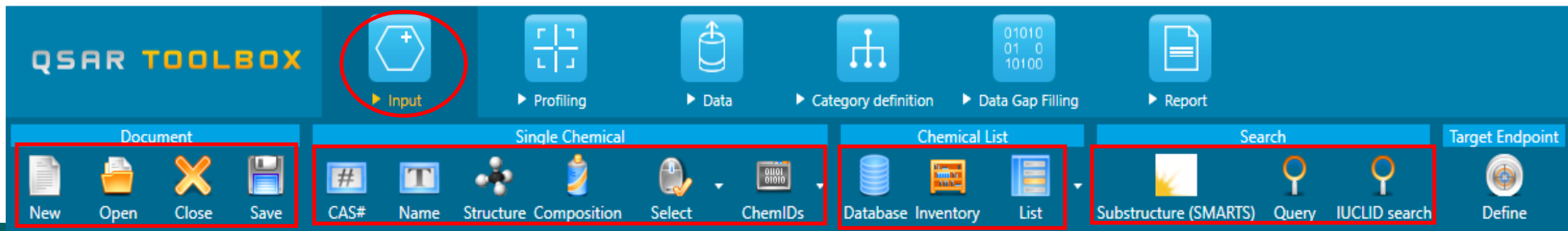
	1 [target]	2	3	4
Structure				
Physical Chemical Properties	3/4 M: 75 °C	M: 74 °C		M: 36 °C
Environmental Fate and Transport				
Ecotoxicological Information	1/3 M: 5 ppm			
Human Health Hazards				
Acute Toxicity	2/2 M: 5,6E+03 mg/kg			M: 3,91E+03 mg...
ADME				
Bioaccumulation				
Carcinogenicity				
Developmental Toxicity / Teratogenicity				
Genetic Toxicity				
in Vitro				
Bacterial Reverse Mutation Assay (e.g....)				
Gene mutation				
Salmonella typhimurium				
No S9 Info	4/4 M: Positive	M: Positive	M: Positive	M: Positive
With S9	4/13 M: Equivocal	M: Positive	M: Negative	M: Equivocal
Without S9	4/27 M: Negative	M: Negative	M: Equivocal	M: Negative

OECD QSAR Toolbox: input

The users are able to:

- ✓ **Open a new or already saved document**
- ✓ **Close or save the current document**
- ✓ **Load a single target chemical** – by CAS, Name, Structure, SMILES* or drawing (including mixtures), select from a file
- ✓ **Load a list of chemicals** – database, inventory, custom file
- ✓ **Customized search** – searching chemicals and/or data within the Toolbox databases. One or more than one criteria combined with logical operators (AND, OR, NOT) could be used.
- ✓ **Search in IUCLID databases** – searching chemicals within the IUCLID databases imported in Toolbox. One or more than one criteria for the composition (e.g. impurity, additive) of the searchable substances could be defined.

*SMILES notation: string representation of a molecule (Simplified Molecular Input Line Entry Specification)



QSAR Toolbox: input from the CAS

Target chemical:
4-nitrobenzoyl chloride (CAS 122-04-3)

1. Enter CAS N

2. Click search button

3. Press ok

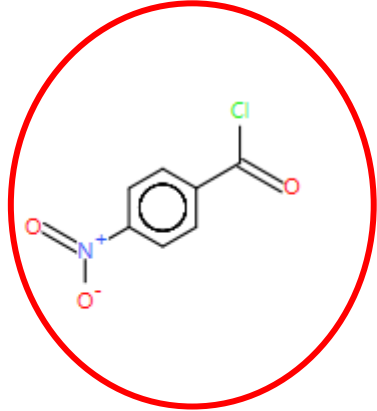
Search by CAS #

122043 Search

Select All Unselect All Invert Selection Selected 1 of 1

1	CAS	122-04-3
	SMILES	[O-][N+](=O)c1ccc(cc1)C(Cl)=O
	CS Relation	High
	Substance	Mono constituent
	Identity	Sources:15
	Name	4-Nitrobenzoyl chloride;Benzoyl chl...
	Sources	AIIC Canada DSL DSSTOX

✓



QSAR TOOLBOX

Input Profiling Data Category definition Data Gap Filling Report

Document Single Chemical Chemical List Search Target Endpoint

New Open Close Save CAS# Name Structure Composition Select ChemIDs Database Inventory List Substructure (SMARTS) Query IUCLID search Define

QSAR Toolbox: input (reliability of the target identity)

CAS-SMILES relation indicates the reliability of the target identifier

- **High**: This label is assigned if the chemical belongs to *at least one high quality data source* (database or inventory)
- **Moderate**: The moderate label is assigned if the chemical belongs to *three or more sources with unknown quality* (marked with "Distribute to QA").
- **Low**: This label is assigned if the chemical belongs to *less than three, but at least one source with unknown quality* ("Distribute to QA").

QSAR TOOLBOX

Input Profiling Data Category definition Data Gap Filling Report

Document Single Chemical Chemical List Search Target Endpoint

New Open Close Save CAS# Name Structure Composition Select ChemIDs Database Inventory List Substructure (SMARTS) Query Define

Documents

Document 1
[C: 1;Md: 0;P: 0] CAS: 122043

Structure

Filter endpoint tree... 1 [target]

Structure info

Additional Ids
CAS Number
CAS-SMILES relation
Chemical name(s)
Composition
Molecular formula
Predefined substance type
SMILES

Parameters

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information

Human Health Hazards

EC Number:2045174
122-04-3
High
4-Nitrobenzoyl chloride
C7H4ClNO3
Mono constituent
[O-][N+](=O)c1ccc(cc1)C(Cl)=O

CAS-SMILES relationship show "High" relation for the target: good quality

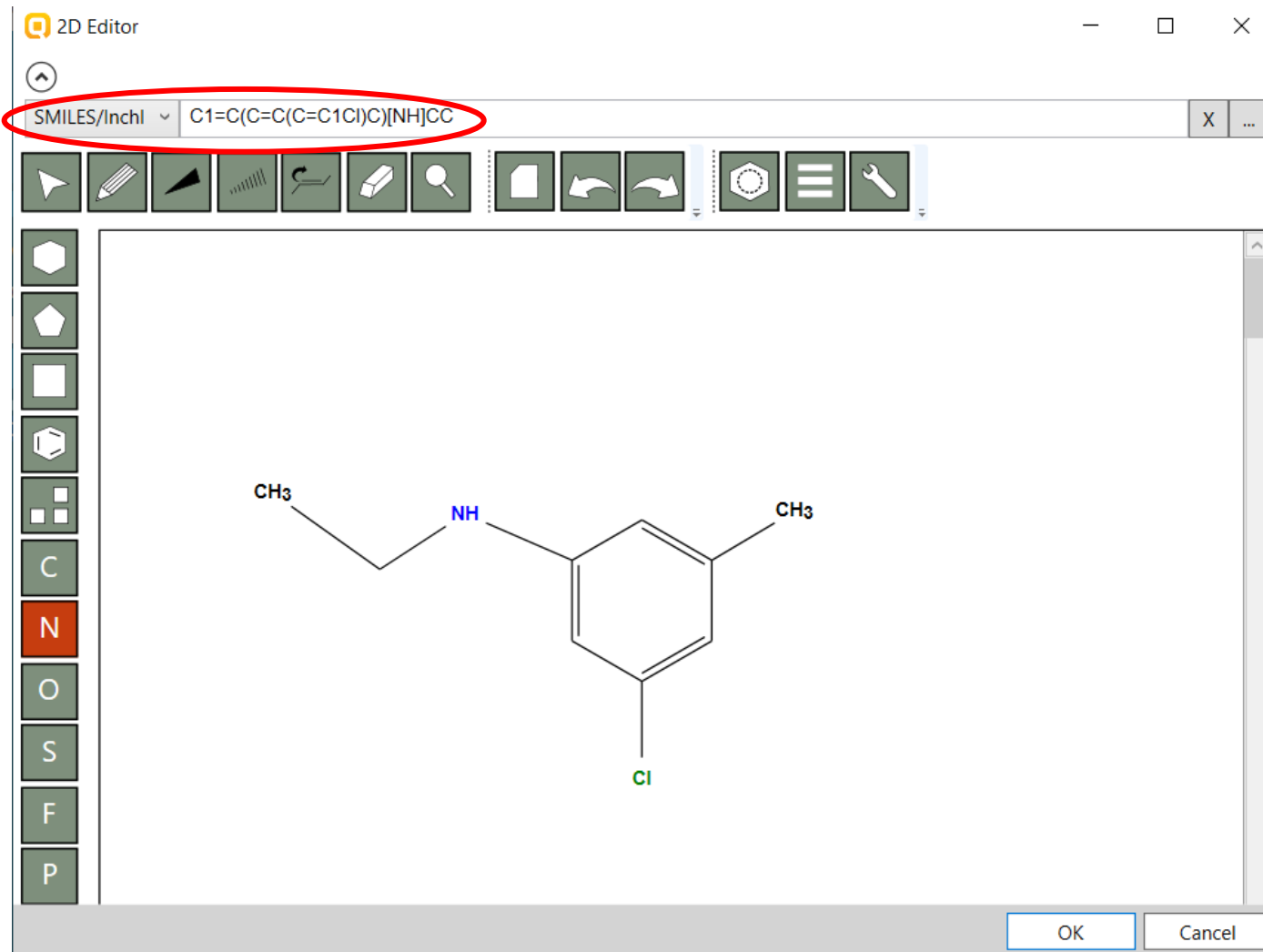
QSAR Toolbox: input

- ❖ The user can search the chemicals based on:
 - ❖ CAS
 - ❖ name
 - ❖ drawing the 2D Structure (2D editor)
 - ❖ pasting or drawing the SMILES or InChI
 - ❖ A given property (selecting DB)
 - ❖ Subfragment (using SMART)
 - ❖ Profilig results

SMILES: Simplified Molecular Input Line Entry System

InChI: International Chemical Identifier

SMART: SMiles ARbitrary Target Specification, extension of SMILES code



QSAR Toolbox: Input (Define target endpoint 1)

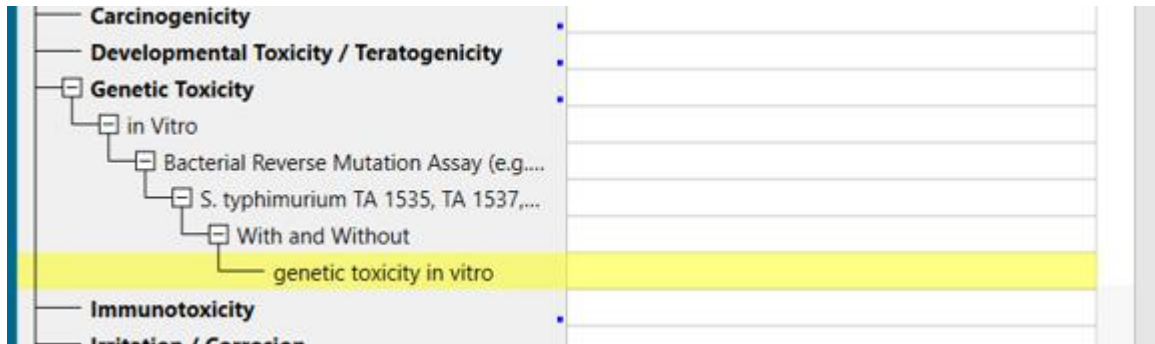
Endpoint of interest can be specified, during the input

The most relevant profilers and databases will be highlighted with different colors.

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The 'Input' menu is open, showing options like 'New', 'Open', 'Close', 'Save', 'CAS#', 'Name', 'Structure', 'Composition', 'Select', 'ChemIDs', 'Database', 'Inventory', 'List', 'Substructure (SMARTS)', and 'Query'. The 'Define' button is highlighted with a red box and labeled '1. Click «Define» button'. The 'Select endpoint' dialog box is open, showing a list of endpoints under 'Human Health Hazards'. The 'Genetic Toxicity' endpoint is selected and highlighted with a red box, labeled '2. Select «Genetic Toxicity»'. The 'Next' button is highlighted with a red box and labeled '3. Click «Next» button'. The 'Structure' panel shows a chemical structure of a nitrobenzene derivative.

OECD QSAR Toolbox: Input (Define target endpoint 2)

1. First click on Endpoint and select the endpoint from the drop-down menu (e.g. "Genetic toxicity *in vitro*")
2. Next select Type of method (e.g. "*In Vitro*")
3. Then consecutively select other metadata: Select "Strain", "Metabolic activation", "Test organ (species)"
4. Finally click on Finish



Decision tree is expanded and the row is yellow

2. Type of method

3. Select other metadata

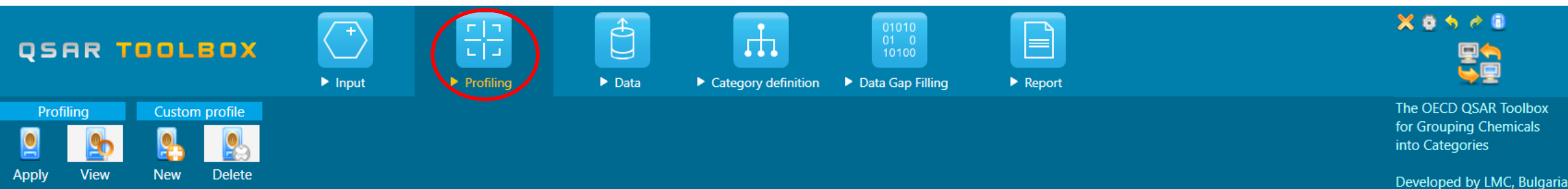
1. Select «Endpoint»

4. click on Finish

QSAR Toolbox: Profiling

- **Profiling**: contains the **structural** or **mechanistic** knowledge of the target, coded in profiling schemes (**profilers**), but not (experimental, (eco)toxicity) data!
- The **profilers** identify the affiliation of the target chemical(s) to categories (**functional groups/alerts**). Mechanistic justification for the identified alerts is provided.
- The outcome of the profiling determines the most appropriate way to search for analogues, also useful for screening or prioritization of substances
- The “Profiling” module contains also **observed and simulated metabolisms/transformations**, which could be used in combination with the profiling schemes
- The most relevant profilers will be highlighted with different colors, based on the endpoint definition (green or orange)

The profiler results are not a predictions and should not be used as such!



QSAR Toolbox: Defining the endpoint

➤ Coloring of the databases



Database

Database contains experimental data for the target (selected) endpoint

➤ Coloring of the profiling schemes/metabolic simulators



Suitable profiler/metabolic simulator

Profiler/metabolic simulator is suitable to be used for the target (selected) endpoint



Plausible profiler/metabolic simulator

The profiler/metabolic simulator could be used for the target (selected) endpoint

➤ Colors in the Data matrix



Target endpoint

The row corresponding to the target endpoint is highlighted in yellow on the data matrix

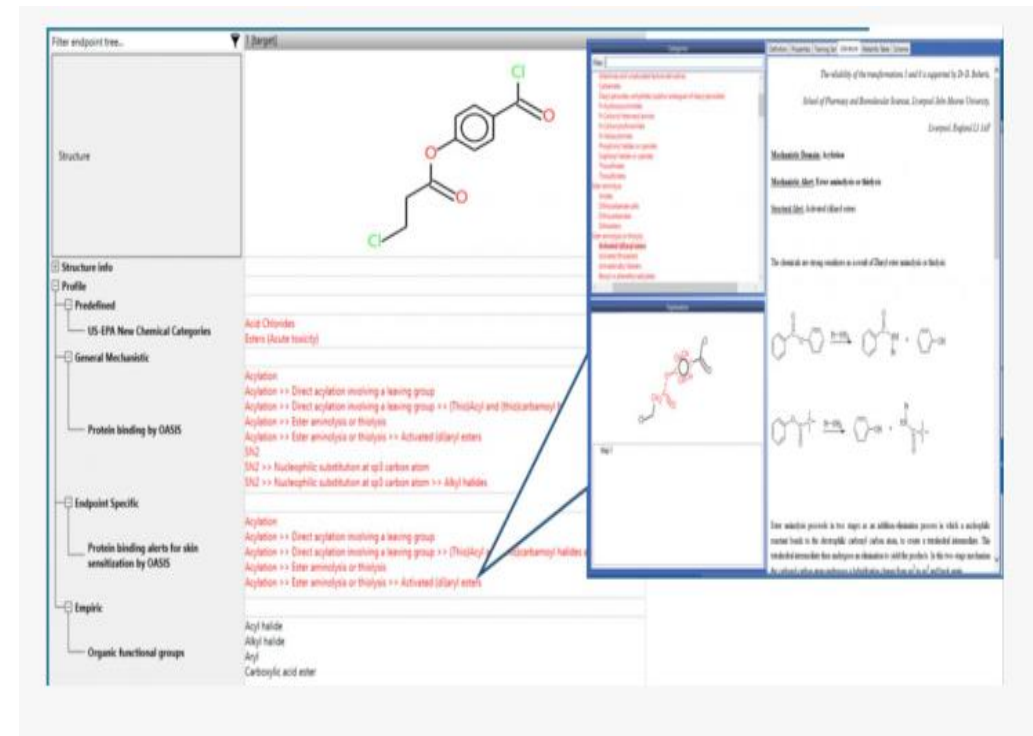


Supporting endpoint


The row corresponding to supporting endpoint is highlighted in pale yellow on the data matrix


OECD QSAR Toolbox: Profiling


- **Structural profilers** support the identification of structurally similar substances.
- **Mechanistic profilers** provide an understanding of the mode of action, which is key to predict the activity of substances or to form categories based on the mode of action.
- The results of the profilers in the Toolbox, include descriptions and references to scientific papers to explain the outcome





QSAR TOOLBOX


 Input


 Profiling


 Data


 Category definition


 Data Gap Filling

 Report

 Apply

 View

 New

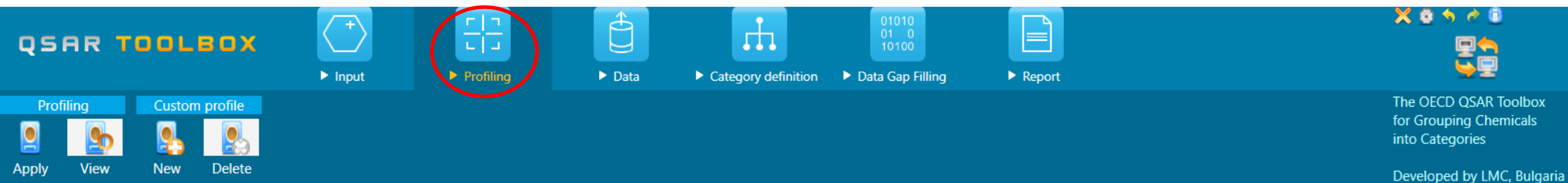
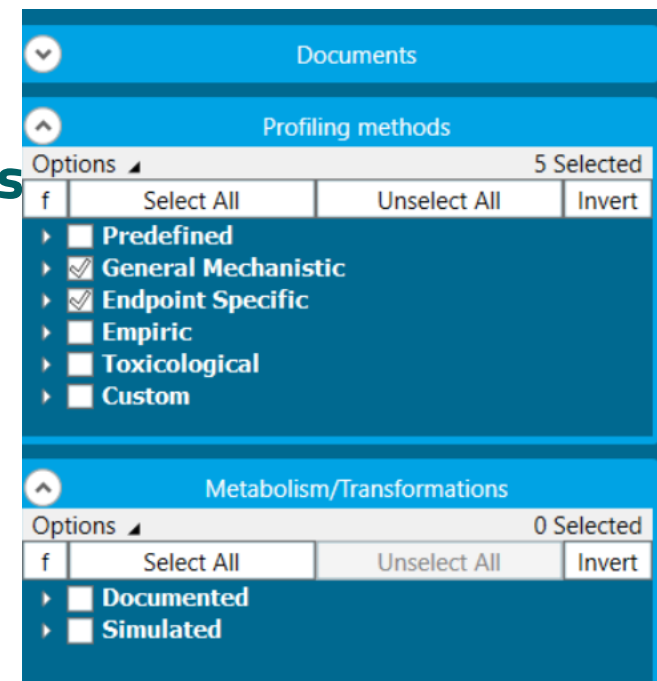
 Delete

The OECD QSAR Toolbox
for Grouping Chemicals
into Categories

Developed by LMC, Bulgaria

OECD QSAR Toolbox: Profiling

- **Predefined:** include standard categorization schemes e.g. OECD HPV Chemical categories
- **General mechanistic:** consist of rules of **general chemical characteristics** based on published or expert knowledge (e.g. DNA binding by OECD, Estrogen Receptor...)
- **Endpoint specific:** consist of **specific endpoint** based on published or expert knowledge e.g. *in vitro* mutagenicity (Ames test) by ISS
- **Empiric:** e.g. organic functional group US EPA, groups of elements
- **Toxicological:** includes only one profiler, the repeated dose HESS
- **Custom:** schemes based on the user's knowledge. Example Prioritization Scheme (PBT)



QSAR Toolbox: Profiling

1. Go to Profiling module
2. Select the name of the profiler
3. Select About
4. Close before proceeding

Relevance
Green: more relevant
profilers

Orange: plausible profilers,
related in some way to the
endpoint

Not highlighted: profiler
with no relation

The screenshot shows the QSAR Toolbox Profiling module. The top toolbar has icons for Input, Profiling (highlighted with a red circle and labeled '1. Click profiling'), Data, Category definition, Data Gap Filling, and Report. Below the toolbar, there are tabs for Profiling and Custom profile. The Profiling tab is active, showing a list of documents and a list of profiling methods. The 'in vitro mutagenicity (Ames test) alerts by ISS' profiler is selected (highlighted in green and labeled '2. Select the name of the profiler'). The 'About' button for this profiler is highlighted with a red box and labeled '3. Select About'. The 'About' dialog box is open, showing the name, short description, disclaimer, donor, author, website, and details of the profiler. The 'Close' button in the dialog box is highlighted with a red box and labeled '4. Close'.

1. Click profiling

2. Select the name of the profiler

3. Select About

4. Close

Documents

- Document 1
 - # [C: 1;Md: 0;P: 0] CAS: 122043
 - # [C: 1;Md: 0;P: 0] CAS: 122043

Profiling methods

Options 4 Selected

f Select All Unselect All Invert About Options

Plausible

- ☐ Aquatic toxicity classification by ECOSAR
- ☐ Chemical elements
- ☒ DNA alerts for AMES, CA and MNT
- ☐ DNA binding by OASIS
- ☐ DNA binding by OECD
- ☐ DNA binding by OECD
- ☒ in vitro mutagenicity (Ames test) alerts by ISS
- ☐ in vivo mutagenicity (micro nucleus) alerts by ISS
- ☐ Lipinski Rule Oasis
- ☐ OECD HPV Chemical Categories
- ☐ Organic functional groups
- ☐ Organic functional groups (US EPA)
- ☐ Organic functional groups (US EPA)
- ☐ Organic functional groups, Norbert Haider (checkmol)

Metabolism/Transformations

Options 0 Selected

f Select All Unselect All Invert

Plausible

- ☐ Dissociation simulator
- ☐ Hydrolysis simulator (neutral)
- ☐ Observed Mammalian metabolism
- ☐ Observed Rat Liver S9 metabolism
- ☐ Rat liver S9 metabolism simulator

Filter endpoint tree...

Structure

info

Parameters

- ☒ Physical Chemical Properties
- ☒ Environmental Fate and Transport
- ☒ Ecotoxicological Information
- ☒ Human Health Hazards
- ☒ Acute Toxicity
- ☒ Bioaccumulation
- ☒ Carcinogenicity
- ☒ Developmental Toxicity / Teratogenicity
- ☒ Genetic Toxicity
 - ☒ in Vitro
 - ☒ Bacterial Reverse Mutation Assay (e.g....
 - ☒ S. typhimurium TA 1535, TA 1537,...
 - ☒ With and Without
 - genetic toxicity in vitro
- ☒ Immunotoxicity
- ☒ Irritation / Corrosion

About

Name

in vitro mutagenicity (Ames test) alerts by ISS

Short Description

This profiler is based on the Mutagenicity/Carcinogenicity module of the software Toxtree. It works as a decision tree for estimating in vitro (Ames test) mutagenicity, based on a list of 30 structural alerts (SAs). The SAs for mutagenicity are molecular functional groups or substructures known to be linked to the mutagenic activity of chemicals. As one or more SAs embedded in a molecular structure are recognised, the system flags the potential mutagenicity of the chemical. The present list of SAs is a subset of the original Toxtree list, obtained by eliminating the SAs for nongenotoxic carcinogenicity.

Disclaimer

The structural boundaries used to define the chemical classes (e.g. "Alcohol" – chemical class from "Organic functional group" profiler) or alerting groups responsible for the binding with biological macromolecules (e.g. "Aldehydes" – structural alert for protein binding), represent structural functionalities in the molecule which could be used for building chemical categories for subsequent data gap filling. They are not recommended to be used directly for prediction purposes (as SARs).

Donator(s)

Institute for Health and Consumer Protection, Joint Research Centre - European Commission, Ispra, Italy; Istituto Superiore di Sanità (ISS), Rome, Italy

Author(s)

Romualdo Benigni, Cecilia Bossa

Website

Details

Name	Value
Scheme type	Linear

QSAR Toolbox: how to select the profiler?

Several approaches can be taken to identify the profilers to be applied:

- 1) choose the most appropriate ones from mechanistic and/or structural profilers
- 2) choose the relevance for an endpoint (e.g., in vitro mutagenicity alerts by ISS), also with the help of orange/green colors
- 3) choose from the list in blue (unguided)

In the matrix the result of applying the profilers is reported

The screenshot shows the QSAR Toolbox interface with several annotations:

- 1. Click profiling:** A red circle highlights the 'Profiling' icon in the top toolbar.
- 2. Select the name of the profilers:** A red circle highlights the 'in vitro mutagenicity (Ames test) alerts by ISS' option in the 'Profiling methods' list.
- 3. Click apply:** A red circle highlights the 'Apply' button in the 'Documents' section.

The interface also displays a chemical structure of 4-chloro-2-nitrobenzoic acid, a list of endpoints (Genetic Toxicity, Immunotoxicity, etc.), and a table of results for the selected profiler.

Endpoint	Result
Acyl halides (Genotox)	
Nitro-aromatic (Genotox)	
Radical	
Radical >> Radical mechanism via ROS formation (indirect)	
Acyl halides	
Nitro-aromatic	
Acyl halides	

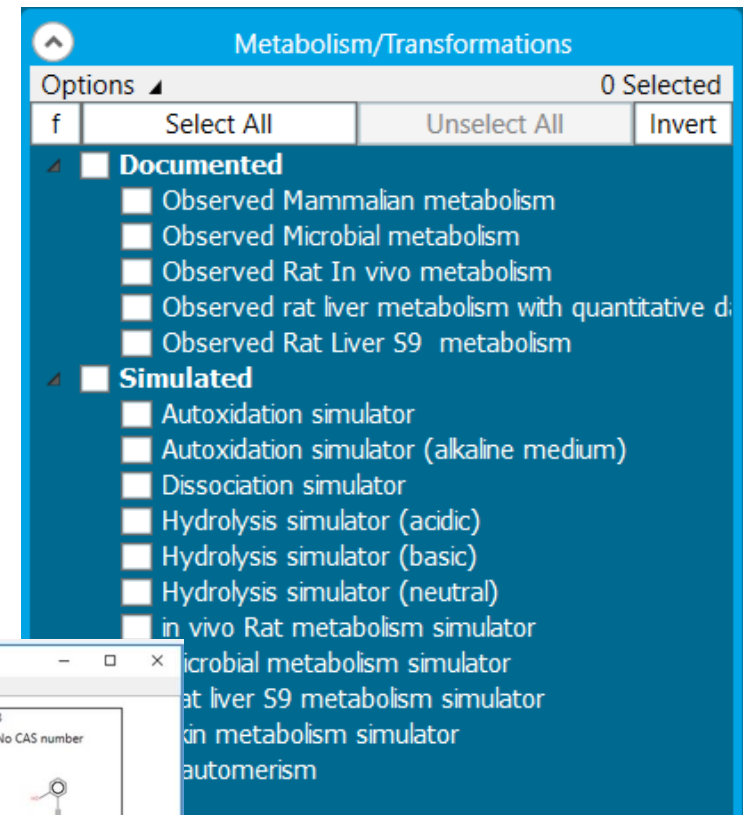
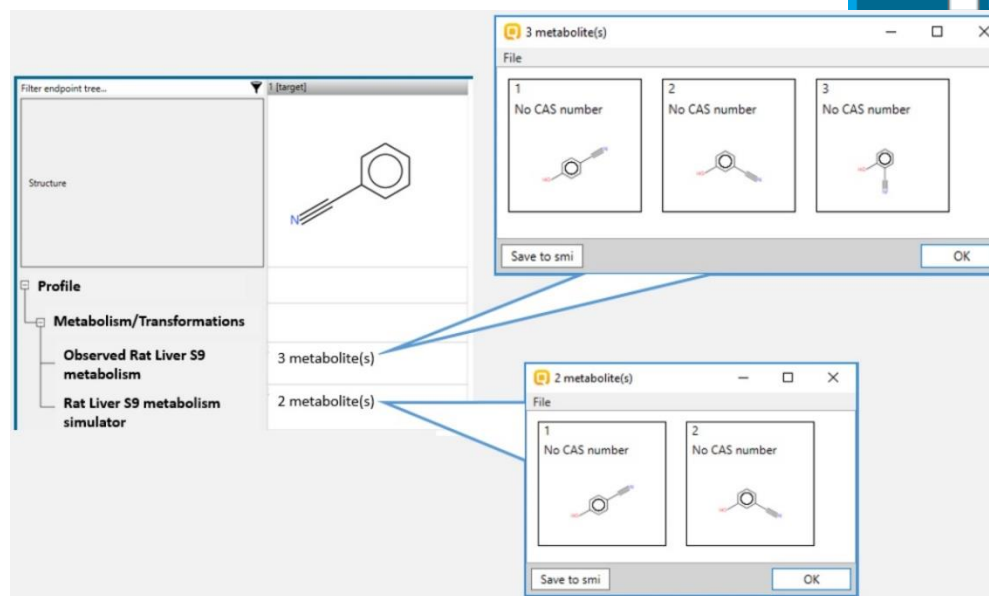
QSAR Toolbox: Profiling, metabolism/transformation

Toolbox can take metabolism into account:

- 5 Databases with observed metabolism (experimental)
- 11 Metabolism simulators (biotic and abiotic)

A screening to all the metabolites for their potential hazard

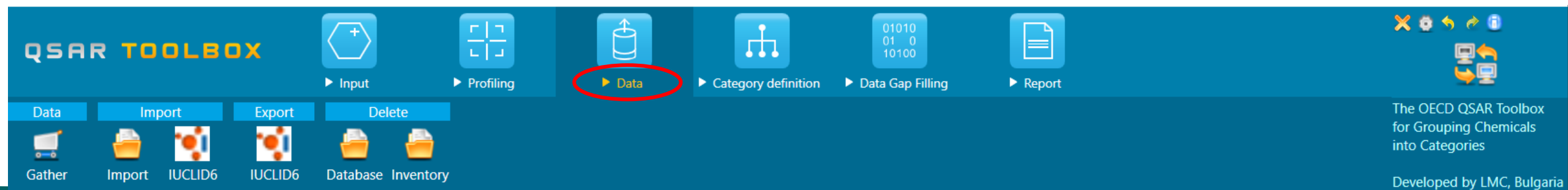
- Applying the profilers
- Searching through Toolbox databases



QSAR Toolbox: Data

- No need to perform new experimental studies if the information is already (publicly) available
- Toolbox includes experimental data on chemicals from 62 **databases**, including 100,000 chemicals and over 3 million data points
- Experimental data from analogues can also be used for read-across predictions
- **Inventories**: chemicals with ID information, and no experimental data
- Data quality cannot be guaranteed, as data curation is the responsibility of the donor

Database content	Chemicals	Data points
Physical chemical properties	50 642	239 949
Environmental fate and transport	15 356	171 861
Ecotoxicological information	23 137	1 349 467
Human health hazards	45 904	1 263 995
Total number	135 039	3 025 272



QSAR Toolbox: Data

Database can be selected for:

- Search analogs for data gap filling
- Search data for one or more substances

The image shows the QSAR Toolbox software interface with several annotations and a data summary window.

Annotations:

- Gather data from the selected database(s) for the chemicals on the data matrix** (points to the 'Gather' button).
- Import custom database/inventory or Import data from Toolbox to IUCLID** (points to the 'Import' button).
- Export data from IUCLID to Toolbox** (points to the 'Export' button).
- Delete custom database(s)/inventory(ies)** (points to the 'Delete' button).
- Databases** (points to the 'Databases' list).
- Inventories** (points to the 'Inventories' list).

Software Interface:

- Documents:** Shows a list of documents.
- Databases:** A list of databases with checkboxes for selection. The 'Data' button is circled in red.
- Options:** A dropdown menu for selecting options.
- Filter endpoint tree...** A tree view showing various endpoints and their associated data.
- Structure:** A chemical structure diagram of a target molecule.
- Structure info:** A table showing parameters and their values.

Structure info Table:

Parameters	Value
Physical Chemical Properties	1/31 M: 4.24
Environmental Fate and Transport	1/5 M: 50 %
Ecotoxicological Information	
- Aquatic Toxicity	AW SW 1/96 M: 0.22
- Sediment toxicity	
- Terrestrial Toxicity	1/55 M: 0.62
Human Health Hazards	
- Acute Toxicity	1/11 M: =1.9
- ADME	1/2 M: 22.6
- Bioaccumulation	
- Carcinogenicity	1/10 M: Equivocal
- Developmental Toxicity / Teratogenicity	1/1 M: =250 mg/kg bdwt
- Genetic Toxicity	1/133 M: Positive
- Immunotoxicity	
- Irritation / Corrosion	1/10 M: Category II
- Neurotoxicity	
- Photoinduced toxicity	
- Repeated Dose Toxicity	1/17 M: =67 mg/kg bdwt/d
- Sensitisation	AW SW AOP 1/50 M: <241 µM
- ToxCast	1/24 M: 0.00619 mg/L
- Toxicity to Reproduction	1/8 M: =100 mg/kg bdwt/d
- Toxicokinetics, Metabolism and Distributi...	

Data Summary Window:

453 points added across 1 chemicals.

QSAR Toolbox: Data and metadata

- 1) **Data** are shown in the matrix
- 2) Click on data to check **metadata** (data that describes data)
- 3) Click on the link to open the **website**

2. Metadata

Filter endpoint tree... 1 [target]

Structure

H2C=O

- Mouse C3H/10T1/2 1/1 M: Positive
- Rat 1/87 M: 1,35 mg/kg bdwt/d
- Syrian hamster embryo cells (SHE) 1/1 M: Positive
- Undefined Test organisms (species) 1/2 M: Positive
- Developmental Toxicity / Teratogeni... 1/34 M: >9,4 mg/kg bdwt/d
- Genetic Toxicity
 - in Vitro
 - Bacterial Reverse Mutation Assay... 1/41 M: Equivocal
 - Comet Assay 1/18 M: positive
 - in Vitro Mammalian Cell Micronucl... 1/9 M: positive
 - in Vitro Mammalian Chromosome... 1/6 M: Positive
 - Mammalian Cell Gene Mutation A... 1/7 M: Positive
 - Other 1/9 M: negative
 - Other Test Type 1/18 M: negative
 - Sister Chromatid Exchange Assay 1/12 M: positive
 - Undefined Test Type 1/2 M: Positive
 - in Vivo 1/68 M: Equivocal
- Immunotoxicity 1/9 M: 0,2 ppm
- Irritation / Corrosion 1/33 M: Category 1 (irreversible effect...
- Neurotoxicity 1/18 M: 0,1 ppm
- Photoinduced toxicity
- Repeated Dose Toxicity 1/1120 M: 10 mg/kg bdwt/d
- Sensitisation AW SW AOP 1/169 M: 63,2 µM

1. Data in the matrix

Data points

Datapoints	Incider	Title	Type of genotoxicity	Type of method	URL	Year
Test);Gene mutation;Other Test organisms (species);With S9;Undefined Strain		for 250 chemicals.		in Vitro	dossier/15858/7/7/2/?documentUUID=26357348-891d-4e36-8f7c-db9e5b48cb8b	1983
Human Health Hazards;Genetic Toxicity;in Vitro;Bacterial Reverse Mutation Assay (e.g. Ames Test);Gene mutation;Other Test organisms (species);With S9;Undefined Strain		Salmonella mutagenicity test results for 250 chemicals.		in Vitro	https://echa.europa.eu/registration-dossier/-/registered-dossier/15858/7/7/2/?documentUUID=26357348-891d-4e36-8f7c-db9e5b48cb8b	1983
Human Health Hazards;Genetic Toxicity;in Vitro;Bacterial Reverse Mutation Assay (e.g. Ames Test);Gene mutation;Other Test organisms (species);Without S9;Undefined Strain		Salmonella mutagenicity test results for 250 chemicals.				

Salmonella mutagenicity test results for 250 chemicals.

EC number: 200-001-8 CAS number: 50-00-0

Formaldehyde

EC number: 200-001-8 CAS number: 50-00-0

General information

- Toxicokinetics, metabolism and distribution
- Acute Toxicity
- Irritation / corrosion
- Sensitisation
- Repeated dose toxicity
- Genetic toxicity
 - Endpoint summary
 - Genetic toxicity: in vitro
 - Genetic toxicity: in vivo
- Carcinogenicity
- Toxicity to reproduction
- Specific investigations
- Exposure related observations in humans
- Toxic effects on livestock and pets
- Additional toxicological data

Classification & Labelling & PBT assessment

Manufacture, use & exposure

Physical & Chemical properties

Environmental fate & pathways

Ecotoxicological information

Toxicological information

Genetic toxicity: in vitro

Currently viewing: 001 Key | Experimental result

Administrative data Data source Materials and methods Results and discussion Applicant's summary and conclusion

Administrative data

Endpoint: in vitro DNA damage and/or repair study

Type of information: experimental study

Adequacy of study: key study

Reliability: 2 (reliable with restrictions)

Rationale for reliability incl. deficiencies: study well documented, meets generally accepted scientific principles, acceptable for assessment

Data source

Reference

3. Website

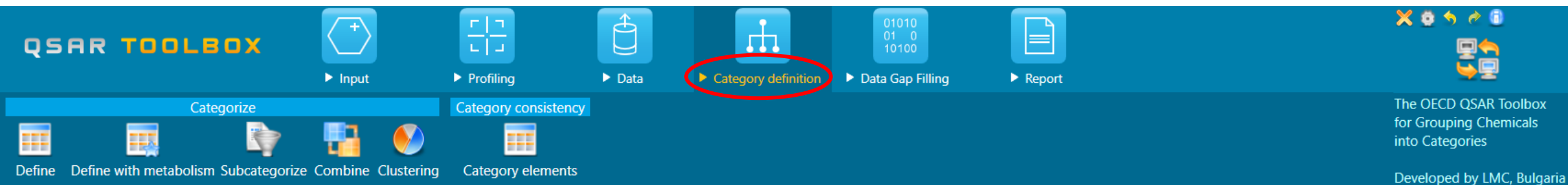
OECD QSAR Toolbox: Category definition

Module for category definition of a target molecule, can be assigned according to:

- structural similarity (e.g. functional groups)
- mechanistic similarity (e.g. protein binding)
- predefined categories (e.g. OECD HPV Chemical categories)

Substances are grouped through a selected profiler («Profiler»)

Substances and data are searched thorough the selected («Data»)



QSAR Toolbox: Category definition

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', and 'Data'. The 'Category' menu is open, showing options: 'Define', 'Define with metabolism', 'Subcategorize', 'Combine', 'Clustering', and 'Category consistency'. Below this, a list of chemical categories is visible, including 'OECD HPV Chemical Categories', 'Substance type', 'US-EPA New Chemical Categories', 'General Mechanistic', '(AOT)Protein binding by OASIS v1', 'Biodeg BioHC half-life (Biowin)', 'Biodegradation primary (Biowin 4)', 'Biodegradation probability (Biowin 1)', 'Biodegradation probability (Biowin 2)', 'Biodegradation probability (Biowin 5)', 'Biodegradation probability (Biowin 6)', 'Biodegradation probability (Biowin 7)', 'Biodegradation ultimate (Biowin 3)', 'DNA binding by OASIS', 'DNA binding by OECD', 'Estrogen Receptor Binding', 'Hydrolysis half-life (Ka, pH 7)(Hydrowin)', 'Hydrolysis half-life (Ka, pH 8)(Hydrowin)', 'Hydrolysis half-life (Kb, pH 7)(Hydrowin)', 'Hydrolysis half-life (Kb, pH 8)(Hydrowin)', 'Hydrolysis half-life (pH 6.5-7.4)', and 'Ionization at pH = 1'.

Define a category based on the target` alerts OR Define a category with metabolism based on the target`s metabolites alerts

Subcategorize the defined category

Combine two or more chemical lists into one

Cluster a chemical list according to the knowledge in a selected profiler

Apply the Category elements in order to check the consistency of your category with respect to a defined endpoint

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

QSAR Toolbox: Analogues and categories

Data from analogues can be used to predict the property of a target substance using read-across or trend analysis.

Toolbox can take into account structural, mechanistic and metabolic aspects, to find toxicologically relevant analogues with data on the property of interest.

A workflow recommended by ECHA includes different phases: first a categorization structural based, then an endpoint specific, finally a sub categorization with an expert judgment.

**Categorization structural based
(non specific endpoint)**

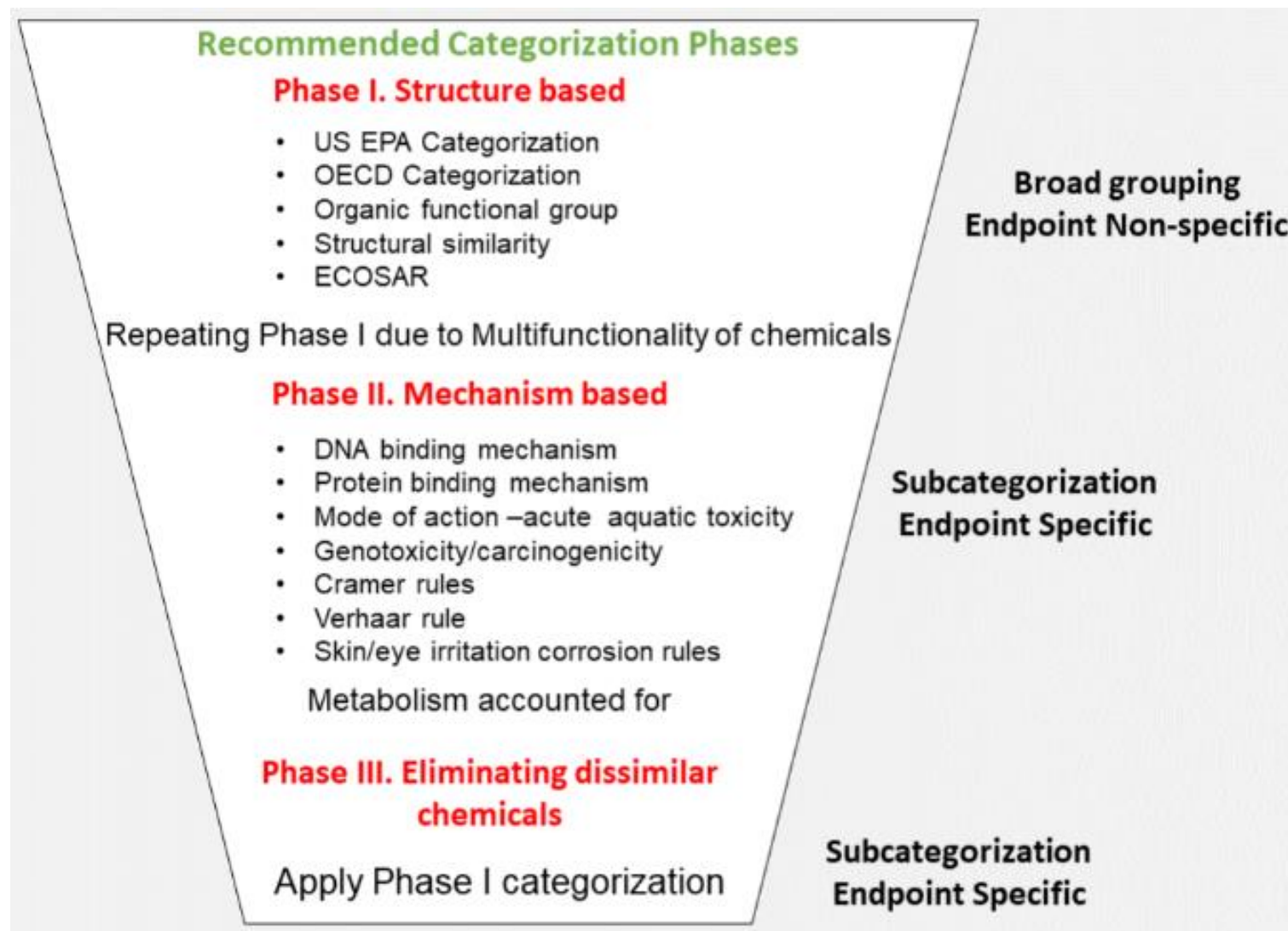


**Sub-categorization
endpoint specific**



**Sub-categorization
Expert judgment**

QSAR Toolbox: ECHA recommended categorization



Expert judgment

QSAR Toolbox: Category definition

Relevance of the profiler in the «Category definition» can be:

- ❖ **Green**, for relevant profiler
- ❖ **Orange**, for plausible profilers
- ❖ With no colour, for profilers with no connection

Metabolism profiler can be applied to take into account in the category definition

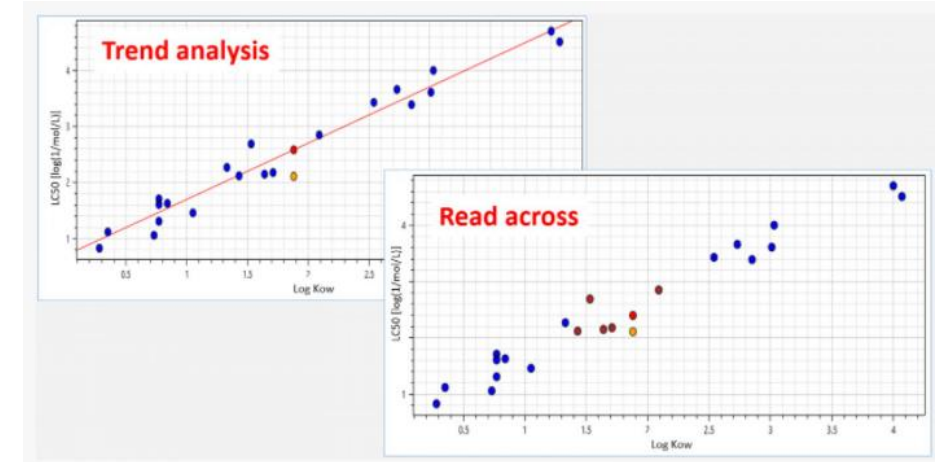
Analogues can be searched according these criteria:

- Parents and metabolites with the same profiler
- Parents and metabolites with a defined profiler
- Common metabolites
- Metabolites similar to a specific compound

QSAR Toolbox: Data Gap Filling

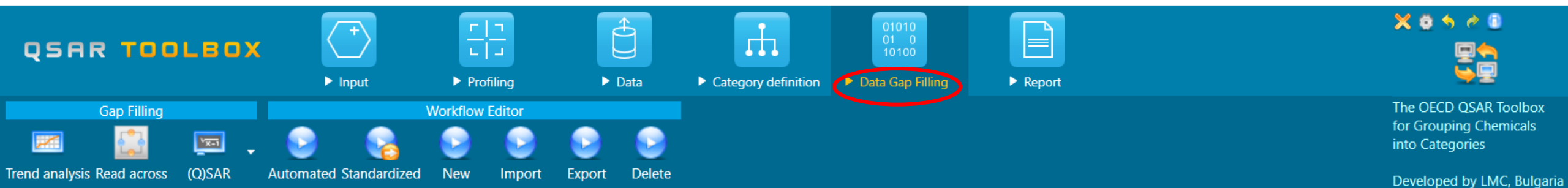
Prediction of the data for the target chemical:

- **Read-across:** using some of the experimental data for some of the closest analogs, identified with the Toolbox
- **Trend analysis:** using all the data from the analogs to derive the regression equation, used to derive the target data
- **QSAR models:** EPIsuite, ECOSAR, DK QSAR database



Trend analysis: method of predicting toxicity of a chemical by analyzing toxicity trends of tested chemicals

Read-across: technique of predicting toxicity of the target chemical using data from source chemicals



QSAR Toolbox: Data Gap Filling

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling' (highlighted with a red circle), and 'Report'. Below the menu bar, there are three main sections: 'Gap Filling' (containing 'Trend analysis' and 'Read across'), '(Q)SAR' (containing 'Automated', 'Standardized', 'New', 'Import', 'Export', and 'Delete'), and 'Workflow Editor'. Three yellow callout boxes provide instructions: 1. 'Apply Trend analysis or Read across approach in order to fill a data gap' (pointing to the 'Trend analysis' and 'Read across' buttons). 2. 'Use some of the available (Q)SAR models or create your custom (Q)SAR models' (pointing to the '(Q)SAR' section). 3. 'Run some of the available Automated or Semi-automated workflows for predefined endpoints or manage your custom workflows - create New, Export/Import, Delete' (pointing to the 'Automated', 'Standardized', 'New', 'Import', 'Export', and 'Delete' buttons). The bottom left panel shows a summary of the current state: 'Only endpoint relevant' is checked. 'At this position:' shows 0 QSARs, 0 Automated workflows, and 0 Standardized workflows. 'In nodes below:' shows 0 QSARs, 0 Automated workflows, and 0 Standardized workflows.

QSAR TOOLBOX

Input Profiling Data Category definition **Data Gap Filling** Report

Gap Filling (Q)SAR Workflow Editor

Trend analysis Read across (Q)SAR Automated Standardized New Import Export Delete

Documents

Document 1

Apply Trend analysis or Read across approach in order to fill a data gap

Use some of the available (Q)SAR models or create your custom (Q)SAR models

Run some of the available Automated or Semi-automated workflows for predefined endpoints or manage your custom workflows - create New, Export/Import, Delete

☒ Only endpoint relevant

At this position:

QSARs	0
Automated workflows	0
Standardized workflows	0

In nodes below:

QSARs	0
Automated workflows	0
Standardized workflows	0

The OECD QSAR Toolbox for Grouping Chemicals into Categories

Developed by LMC, Bulgaria

QSAR Toolbox: Data Gap Filling

- **Relevance** of the profiler used in the sub-categorization, to remove analogues from the prediction is highlighted by colors
- **Workflow**
 - ✓ Automated (fully decided by the TB)
 - ✓ Standardized (partial user selection)
 - ✓ Manual (fully decided by the user)

The screenshot displays the QSAR Toolbox interface with the 'Data Gap Filling' workflow selected. The top navigation bar includes 'Input', 'Profiling', 'Data', 'Category definition', and 'Data Gap Filling'. The 'Data Gap Filling' section is active, showing a 'Filter endpoint tree...' on the left and a 'Structure' panel on the right. The 'Structure' panel displays the chemical structure of 4-chlorobenzonitrile (N#Cc1ccc(Cl)cc1). The 'Filter endpoint tree...' lists various endpoints, including 'Human Health Hazards', 'Acute Toxicity', 'ADME', 'Bioaccumulation', 'Carcinogenicity', 'Developmental Toxicity / Teratogenicity', and 'Genetic Toxicity'. The 'Genetic Toxicity' endpoint is selected, showing a table of results for 'Bacterial Reverse Mutation Assay (e.g., Salmonella typhimurium)'.

The screenshot displays the QSAR Toolbox interface with the 'Data Gap Filling' workflow selected. The top navigation bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The 'Data Gap Filling' section is active, showing a 'Filter endpoint tree...' on the left and a 'Structure' panel on the right. The 'Structure' panel displays the chemical structure of 4-chlorobenzonitrile (N#Cc1ccc(Cl)cc1). The 'Filter endpoint tree...' lists various endpoints, including 'Human Health Hazards', 'Acute Toxicity', 'ADME', 'Bioaccumulation', 'Carcinogenicity', 'Developmental Toxicity / Teratogenicity', and 'Genetic Toxicity'. The 'Genetic Toxicity' endpoint is selected, showing a table of results for 'Bacterial Reverse Mutation Assay (e.g., Salmonella typhimurium)'.

OECD QSAR Toolbox: Report

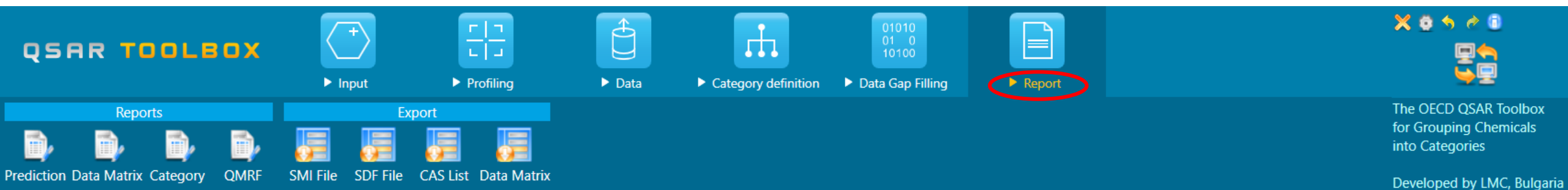
The report module creates a document after accepting the prediction in the previous module. Three different reports can be created:

1. Prediction report
2. Category report
3. Data matrix

All these reports provide the basis for justifying the reliability of the prediction that needs to be critically reviewed by the users and can be manually completed for comments and explanations

In this module the user can:

- ✓ Choose which section to include in the final report, and in what order
- ✓ Include information and data from analogs in the report
- ✓ Enter comments and interpretation of results in the editable fields



OECD QSAR Toolbox: Report

The screenshot shows the OECD QSAR Toolbox interface. The top navigation bar includes icons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. The Report icon is circled in red. Below the navigation bar, there are two main sections: Reports and Export. The Reports section contains icons for Prediction, Data Matrix, Category, and QMRF. The Export section contains icons for SMI File, SDF File, CAS List, and Data Matrix. A Documents bar is located below these sections. Two yellow callout boxes provide additional information:

Create a report for:

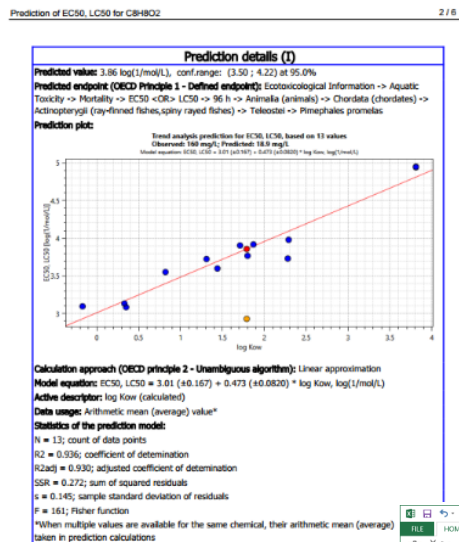
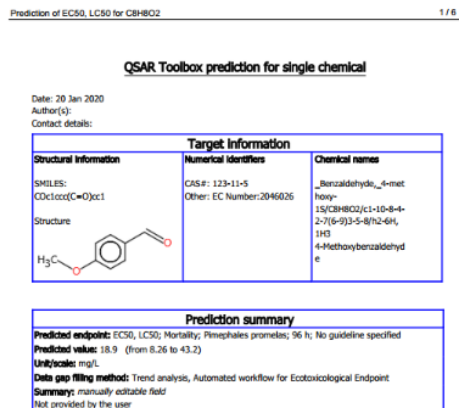
- **Prediction** obtained in TB
- Current information in the **Data matrix**
- consistency of the current chemical **Category**
- an available (Q)SAR model – a library of **QMRFs** is available

Export information for the chemicals (and data) available in the data matrix

QMRF: QSAR model report format
A harmonised template for summarising and reporting key information on QSAR models, including information on model validity

QSAR Toolbox: report

1. Prediction report



3. Data matrix

QSAR Toolbox 4.4 (RC)
Database version: 4.4 (RC) QSAR TOOLBOX TPRF v4.4 (RC)

QSAR Toolbox 4.4 (RC)
Database version: 4.4 (RC) QSAR TOOLBOX

Data matrix_20_17_25_21 - Excel

FILE	HOME	INSERT	PAGE LAYOUT	FORMULAS	DATA	REVIEW	VIEW	Sign in
Data matrix_20_17_25_21 - Excel								
E31								
A B C D E F G H I J K L M N O P Q R S								
1 Substance identity								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								
36								
37								
38								
39								
40								
41								
42								
43								
44								
45								
46								
47								
48								
49								
50								
51								
52								
53								
54								
55								
56								
57								
58								
59								
60								
61								
62								
63								
64								
65								
66								
67								
68								
69								
70								
71								
72								
73								
74								
75								
76								
77								
78								
79								
80								
81								
82								
83								
84								
85								
86								
87								
88								
89								
90								
91								
92								
93								
94								
95								
96								
97								
98								
99								
100								

2. Category report

Chemicals category 1 / 8

QSAR Toolbox report for category

1. Category definition

1.1. Category definition

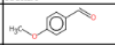
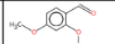
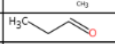
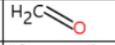
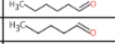
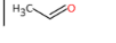

Category name
Not provided by the user manually editable field

Covered (target) endpoint(s)
Ecotoxicological Information/Aquatic Toxicity: Pimephales promelas, Actinopterygii (ray-finned fishes, spiny rayed fishes), Chordata (chordates), Animals (animals), EC50 <OR> LC50, Mortality, Duration=96 h manually editable field




Category hypothesis
Not provided by the user manually editable field

1.2. Category members

Information of category members

#	CAS	Name	SMILES	Structure
1	123-11-5	CBH02	<chem>COc1ccc(C=O)cc1</chem>	
2	613-45-6	2,4-DIMETHOXYBENZALDEHYDE	<chem>COc1cc(C=O)cc(OC)c1</chem>	
3	123-38-6	Propenal	<chem>CCC=O</chem>	
4	50-00-0	CHO	<chem>C=O</chem>	
5	66-25-1	Hexanal	<chem>CCCCCC=O</chem>	
6	110-62-3	CCCCC=O	<chem>CCCCC=O</chem>	
7	75-07-0	Hydral	<chem>CC=O</chem>	

Chemicals category 2 / 8

8	61096-84-2	4-O-methoxy-m- anibenzaldehyde	<chem>COc1ccc(C=O)cc1OC</chem>	
9	4408-86-0	2,4,5- Trimethoxybenzaldehyde	<chem>COc1cc(C=O)c(OC)c(OC)c1OC</chem>	
10	10031-62-0	4- ethoxybenzaldehyde	<chem>CCOC1=CC=C(C=O)C=C1</chem>	

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

Purity / Impurity
Not provided by the user manually editable field

1.3. Profiles/Metabolisms

List of profiles/metabolisms

Profiles used for grouping/subcategorization:

- Aldehydes (Acute toxicity) (US-EPA New Chemical Categories) (primary grouping)
- Substance type (subcategorization)
- Acute aquatic toxicity HGA by OASIS (subcategorization)
- US-EPA New Chemical Categories (subcategorization)
- Aquatic toxicity classification by ECOSAR (subcategorization)
- Organic functional groups, Norbert Haider (checkmol) (subcategorization)
- Organic functional groups (subcategorization)

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 manually editable field

QSAR Toolbox 4.4 (RC)
Database version: 4.4 (RC) QSAR TOOLBOX TPRF v4.4 (RC)

User can customize the report, inserting comments and interpreting the data in the editable fields

QSAR Toolbox: Take home messages

- ❖ **Free software application:** ECHA, OECD and LMC, contribution from donors of data, profilers, models
- ❖ Supports (eco)toxicologists in performing **reproducible** and **transparent** chemical hazard assessment using non-animal methods
- ❖ Source of **QSAR models, existing data, metabolic and mechanistic information** can be used to fill data gaps
- ❖ The **profiler results** are **not predictions**, but they can contribute to find suitable analogues with data to build read-across, or they can be used in a WoE approach
- ❖ Simplified user interface and automatic workflow can help user, together with support section (manuals, forum, help-desk) in the website www.qsartoolbox.org
- ❖ TB should not be considered a Black Box, in which a user uncritically enter an input and get an output, without expert supervision
- ❖ Specific expertise is required e.g., in Toxicology/computational and organic chemistry/IT skills: expert judgement is required in critical step such as for the identification of the analogues, the relevance and reliability of the data

→ **Practical exercises with QSAR Toolbox**

Application of TTC approach for a 'data poor' chemical

Hands-on exercise with the QSAR Toolbox

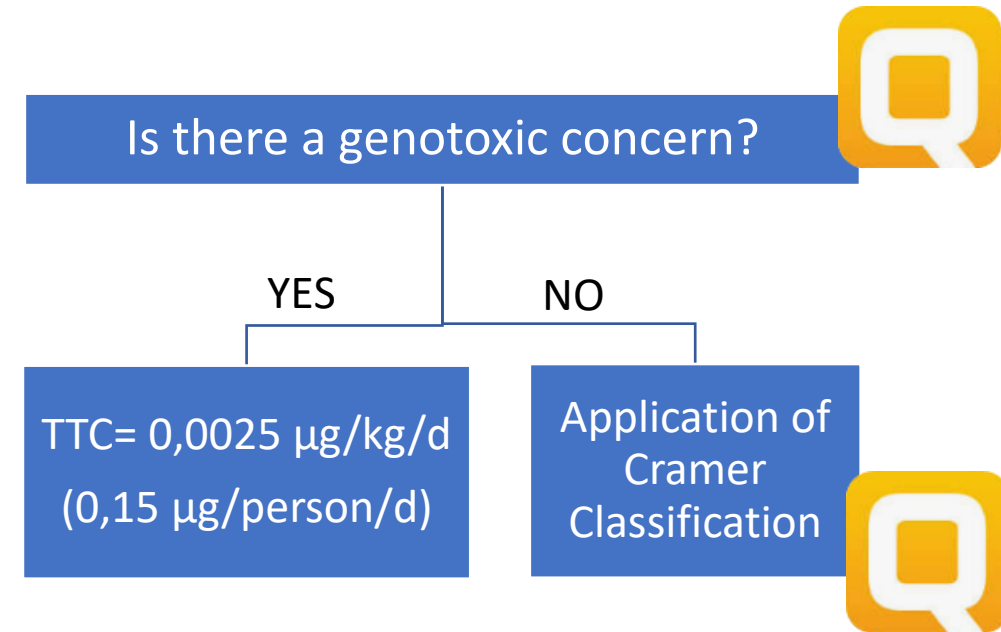
Cecilia BOSSA

Environment and Health Department-Istituto Superiore di Sanità

Practical exercises with QSAR Toolbox


Application of TTC approach for a 'data poor' chemical

- 1) The first step is the **genotoxicity** assessment
 - For the potential DNA-reactive mutagens the TTC value of 0,0025 $\mu\text{g}/\text{kg}/\text{d}$ (0.15 $\mu\text{g}/\text{person}/\text{day}$) is considered
 - If No genotoxic concern is assigned, to the substance, the Cramer classification scheme is applied




n-hexanal







Leibniz Institute for
Food Systems Biology
at the Technical University of Munich



Water
TOP
COST ACTION CA18225



COST
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY



Funded by the
European Union

WaterTOP COST Action (CA18225)

for Water Taste and Odor Problems

About

Publication

Flavor Wheel

CSV

PDF

Print

Column visibility ▾

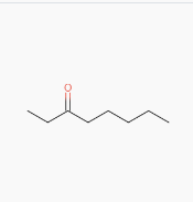
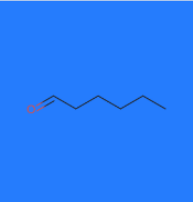
Show

10 ▾

 entries

Search:

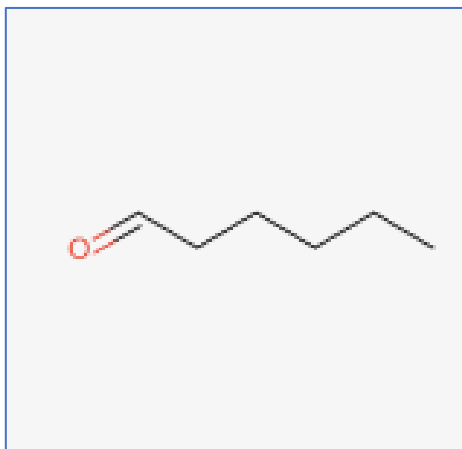
hexanal

Compound	Structure	MW	Quality	OT	Links	Source
<div>Search Compound</div>	<div>Structure Search Tool</div>	<div>Min</div> <div>Max</div>	<div>Search Quality</div>	<div>Min</div> <div>Max</div>	<div>Search Links</div>	<div>Search Source</div>
3-Octanone		128.21	mouldy (details)	39 µg/L(details)	PubChem:246728 FSBI-DB	
Hexanal		100.16	grassy, fatty, lettuce heart (details)	4.5 µg/L(details)	PubChem:6184 FSBI-DB	algae (details)

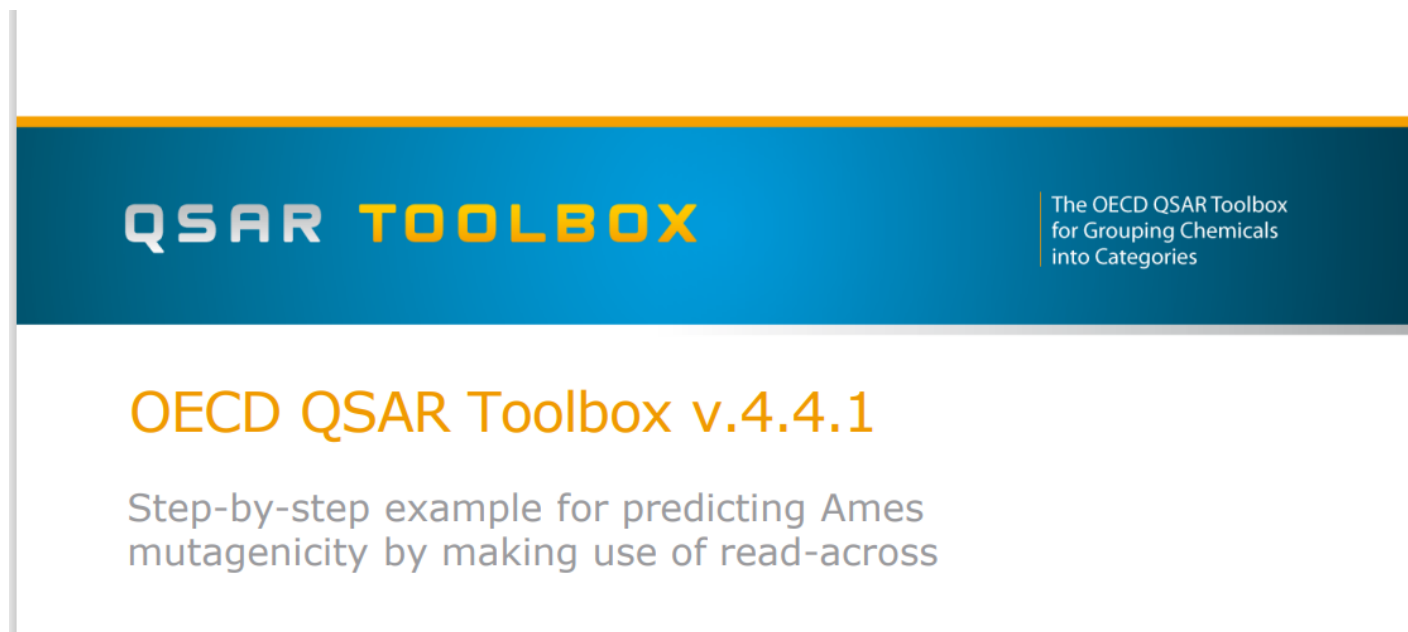
OT=4,5 µg/L



QSAR Toolbox: mutagenicity prediction by Read Across approach



n-hexanal



https://qsartoolbox.org/wp-content/uploads/2020/04/Tutorial_2_Predicting-AMES-by-making-use-of-read-across.pdf

