

## OECD (Q)SAR Toolbox v.4.4.1

Examples illustrating customized  
search (Query Tool) in Toolbox

# Outlook

- **Background**
- Objectives
- Overview of Query tool
- Query tool window
- The exercise
- Workflow process
- Save Query tool searches

## Background

- This is a step-by-step presentation designed to take the user through the functionalities of the Query tool (QT) engine implemented in Toolbox

# Outlook

- Background
- **Objectives**
- Overview of Query tool
- Query tool window
- The exercise
- Workflow process
- Save QT searches

# Objectives

**In this presentation we will use functionalities of the QT to identify:**

- 1) chemicals having specific structural fragments in their molecules
- 2) chemicals answering specific structural and parametric criteria
- 3) chemicals answering specific combinations of data, structural and parametric criteria

# Outlook

- Background
- Objectives
- **Overview of Query tool**
- Query tool window
- The exercise
- Workflow process
- Save QT searches

# Overview of Query tool

**Goal:** Query tool allows searching for chemicals by structure, sub fragments, phys-chem properties and experimental data

- **Chemical identifier search**

- CAS
- Name
- Molecular structures

- **Parameter search**

- Calculated 2D parameters
- Calculated 3D parameters

- **Data search**

- Search for data and metadata within imported databases

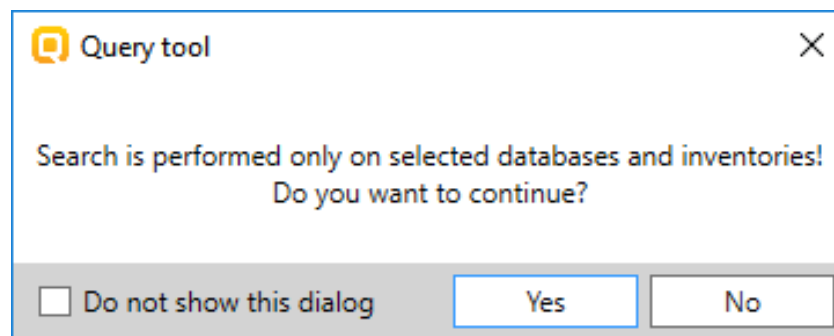
- **Extended search**

- Profiling
- Substructure
- Similarity

# Overview of Query tool

## Prerequisites

- The Query tool functionality searches for single structures matching desired criteria.
- The Query tool functionality searches for chemicals within the selected databases and inventories only

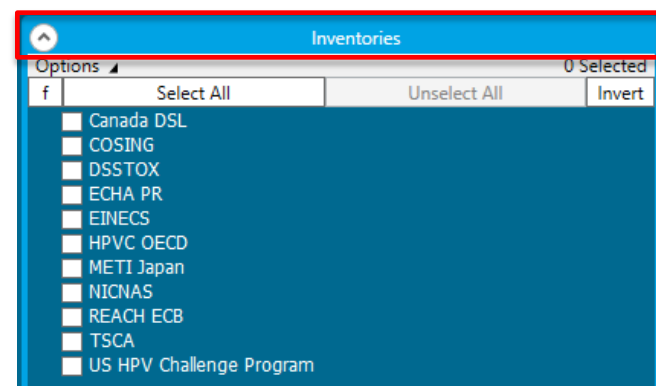
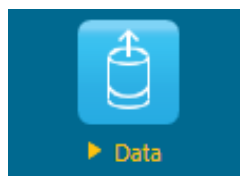
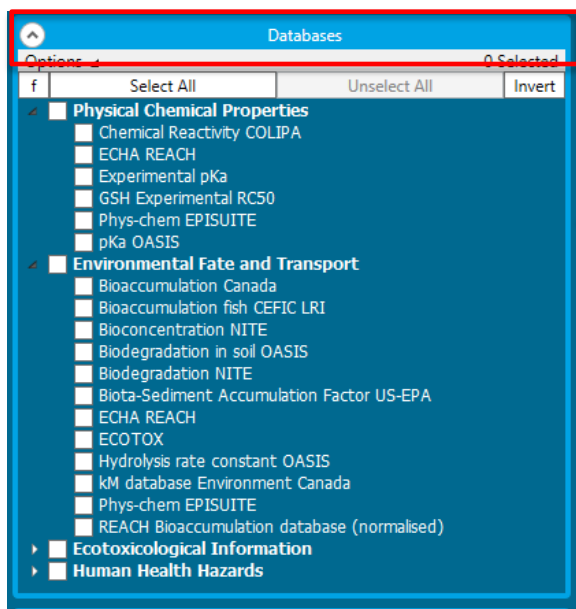




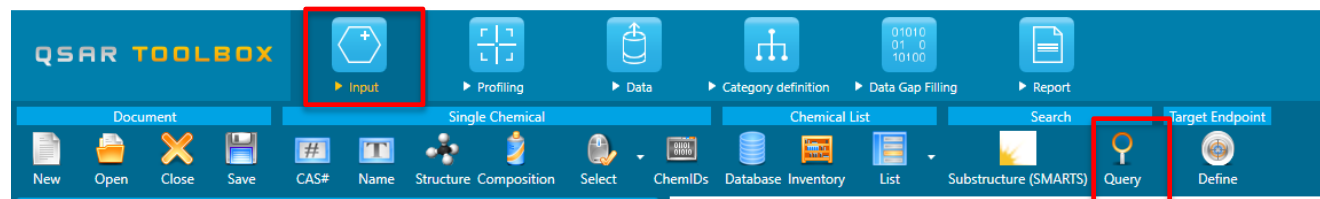
# Overview of Query tool

## Procedure for defining query

**Step 1:** Select databases/inventories of interest in the *Data* module



**Step 2:** Click *Query* button



# Overview of Query tool

## Procedure for defining query

### **Step 3:** Specify criteria for searching structures

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

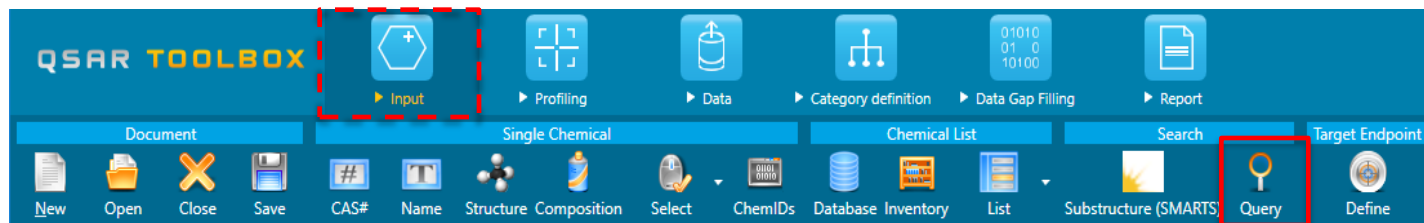
- Top Bar:** Includes a 'Search' icon and a 'Clear All' button.
- Parameter Selection:** A list of parameters is shown, with 'Vapor Pressure (Antoine method)' selected and highlighted in blue.
- Expression Configuration:**
  - Expression:** A dropdown menu is set to '>=' and a text box contains '1E+03'.
  - Origin:** A dropdown menu is set to 'Pressure'.
  - Unit:** A list of units is shown, with 'MPa' selected (indicated by a radio button).
- Logic Diagram:** A visual representation of the query logic is shown on the right. It features a central 'OR' node connected to two input nodes (represented by icons). The diagram is divided into horizontal sections labeled 0, 1, 2, 3, and 4.
- Buttons:** At the bottom, there are buttons for 'Update', 'Add', 'Save', 'Load', 'Execute', and 'Close'.

# Outlook

- Background
- Objectives
- Overview of Query tool
- **Query tool window**
- The exercise
- Workflow process
- Save QT searches

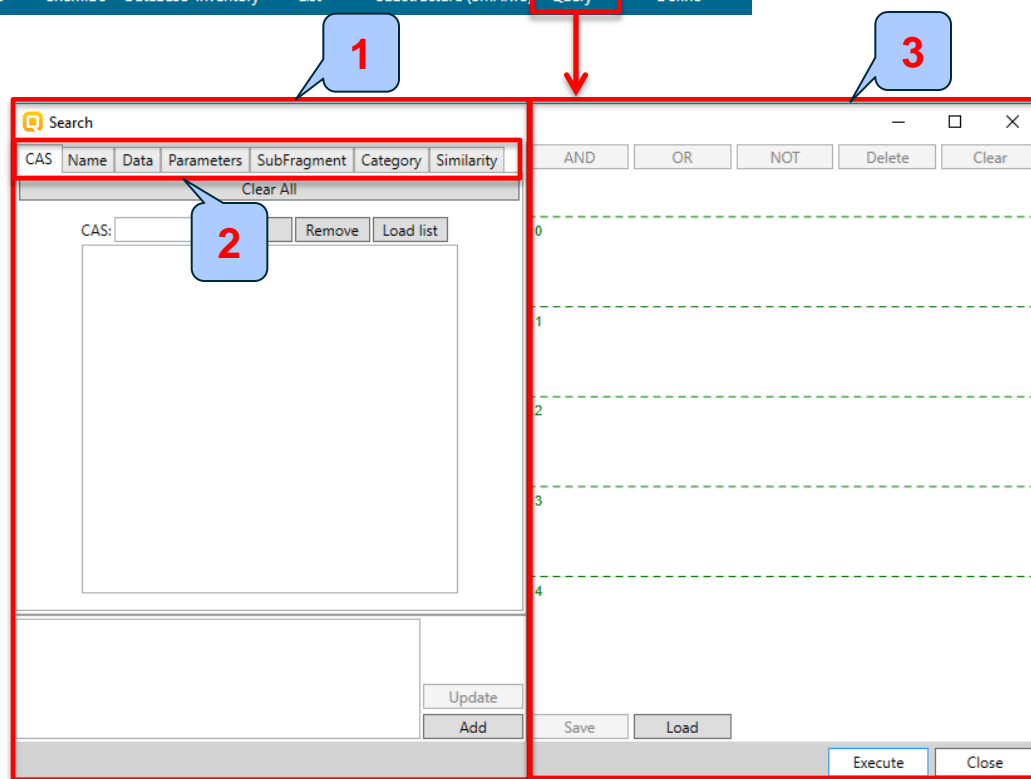
# Query tool window

- The Query tool is easily accessible on the Toolbox input panel



The main components of Query tool window includes:

- Query edit panel (1)
  - search group panel (2)
- Query tree logic panel (3)



# Outlook

- Background
- Objectives
- Overview of Query tool
- Query tool window
- **The exercise**
- Workflow process
- Save QT searches

## The Exercise

- In this exercise we will demonstrate the following queries:
  - Subfragment search for:
    - *Substituted diphenyl amine (Example 1)*
    - *Aliphatic halogens (Example 2)*
  - Combination of queries with Environmental Data (BCF) and 3D parameters (*Dmax*) (*Example 3*)
  - Combination of queries with predefined category (Aldehydes) and ecotox data ( $LC_{50} < 1\text{mg/l}$ ) (*Example 4*)
  - Combination of queries with positive Ames; positive Carcinogenicity data and Subfragment search (*Epoxides*) (*Example 5*)
  - Combination of Skin sensitization data (EC3) and predefined category (*Aldehydes*) queries (*Example 6*)

# Outlook

- Background
- Objectives
- Overview of Query tool
- Query tool window
- The exercise
- **Workflow process**
- Save QT searches

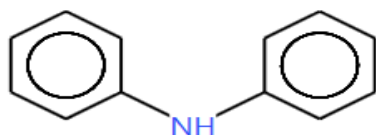
# Subfragment search for identifying chemicals

## *Substituted diphenyl amine*

### Example 1

Search for structures that meet the following structural requirements:

- *Substituted diphenyl amine*



N(c1ccccc1)c1ccccc1

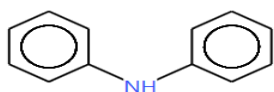


# Subfragment search for identifying chemicals

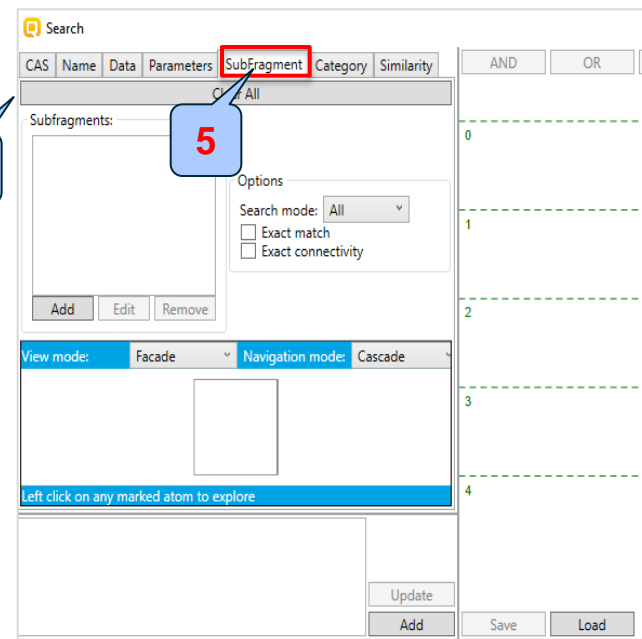
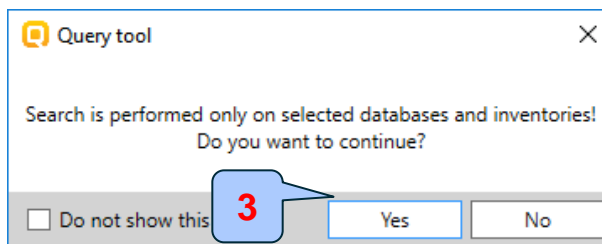
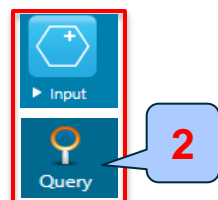
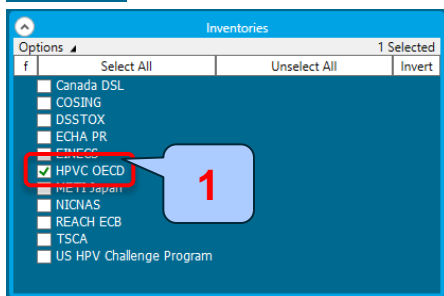
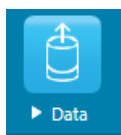
## Procedure for defining

### Example 1

*Substituted diphenyl amine*



N(c1ccccc1)c1ccccc1



QT procedure starts with selecting databases/inventories used for searching the desired criteria:

1. **Select** HPVC OECD inventory located under **Data** module. No databases have been selected in this exercise;
2. Click **Query** button located under **Input** module; The message informs the user that search will be performed on selected databases or inventories only
3. Click **Yes**;
4. Query panel appears;
5. Select **SubFragment** panel;

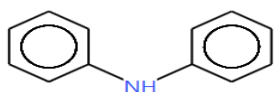
Continued on the next slide

# Subfragment search for identifying chemicals

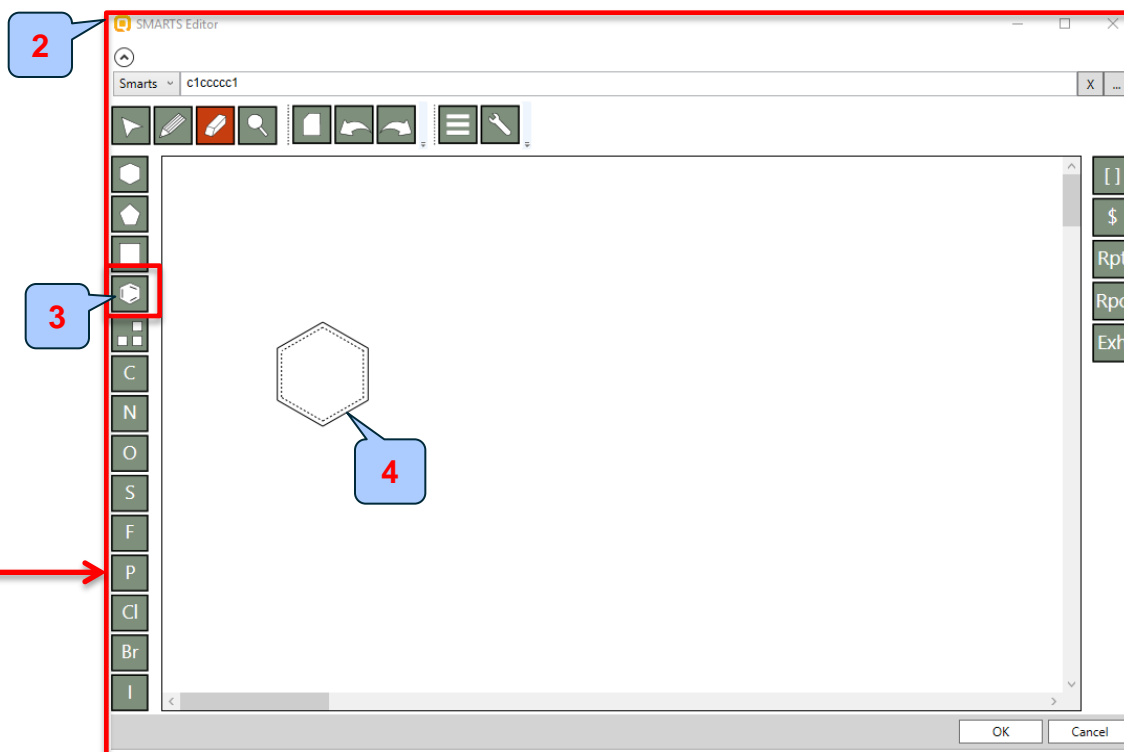
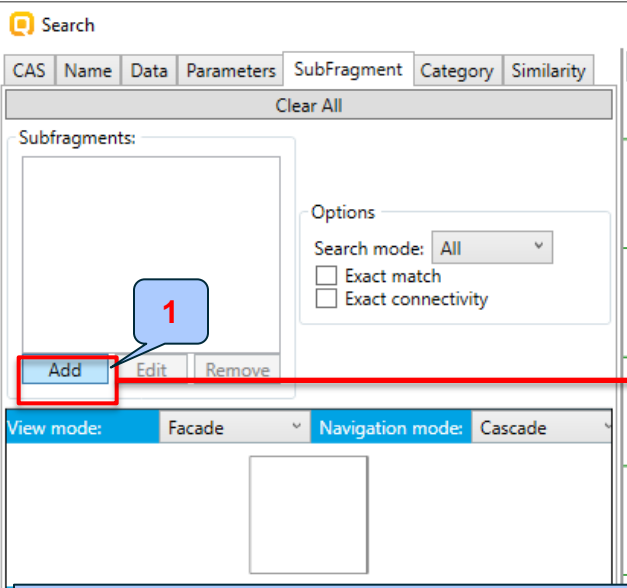
## Procedure for defining

### Example 1

*Substituted diphenyl amine*



N(c1ccccc1)c1ccccc1



1. Click **Add** button;
2. The SMARTS Editor window appears;
3. Perform **left click** on the benzene ring from the Template panel
4. **Left click** on the blank plot in order to put benzene fragment;

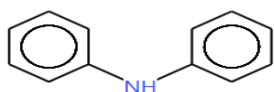
Continued on the next slide

# Subfragment search for identifying chemicals

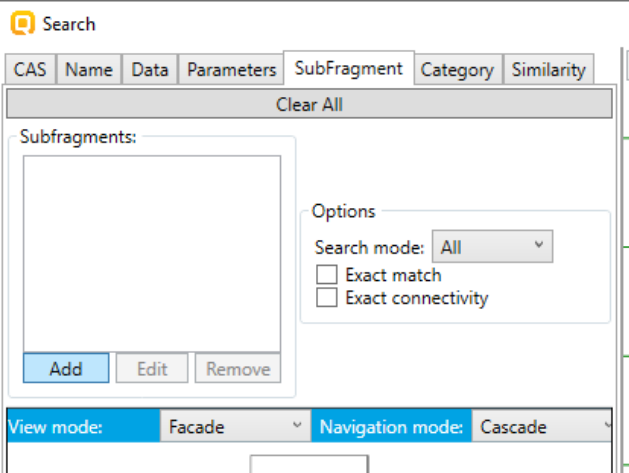
## *Procedure for defining*

### Example 1

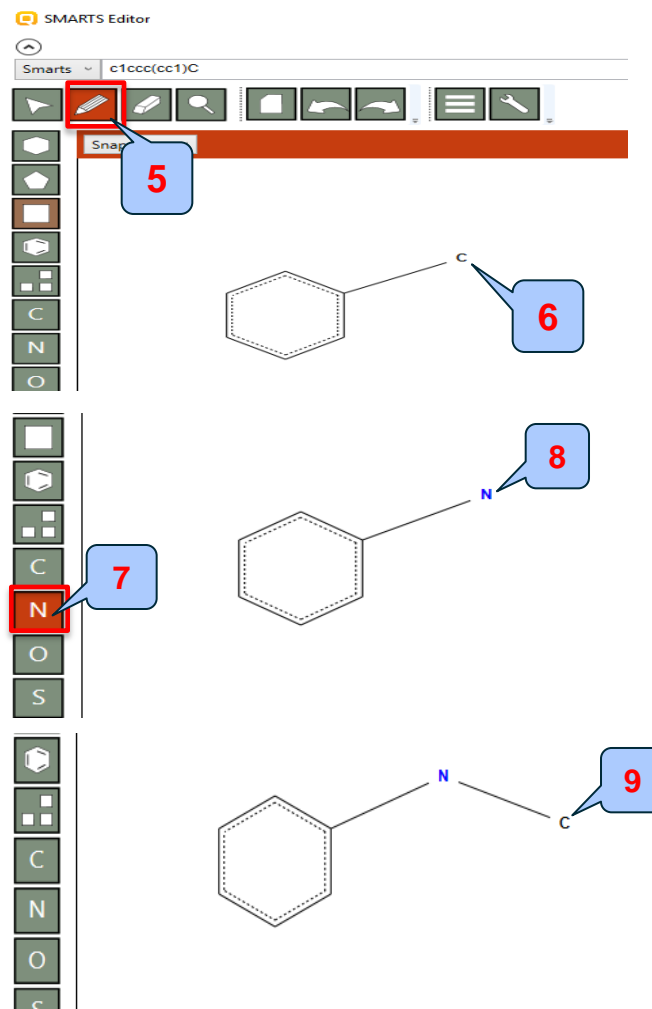
*Substituted diphenyl amine*



N(c1ccccc1)c1ccccc1



5. Click the **pencil** button to draw a single bond;
6. Click near to one of the C atom from the benzene ring to draw a single bond;
7. Click **N** atom from the templates;
8. Put the selected N atom over the C atom from the single bond;
9. Repeat step 5 and draw a single bond to the N atom;



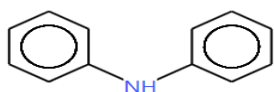
Continued on the next slide

# Subfragment search for identifying chemicals

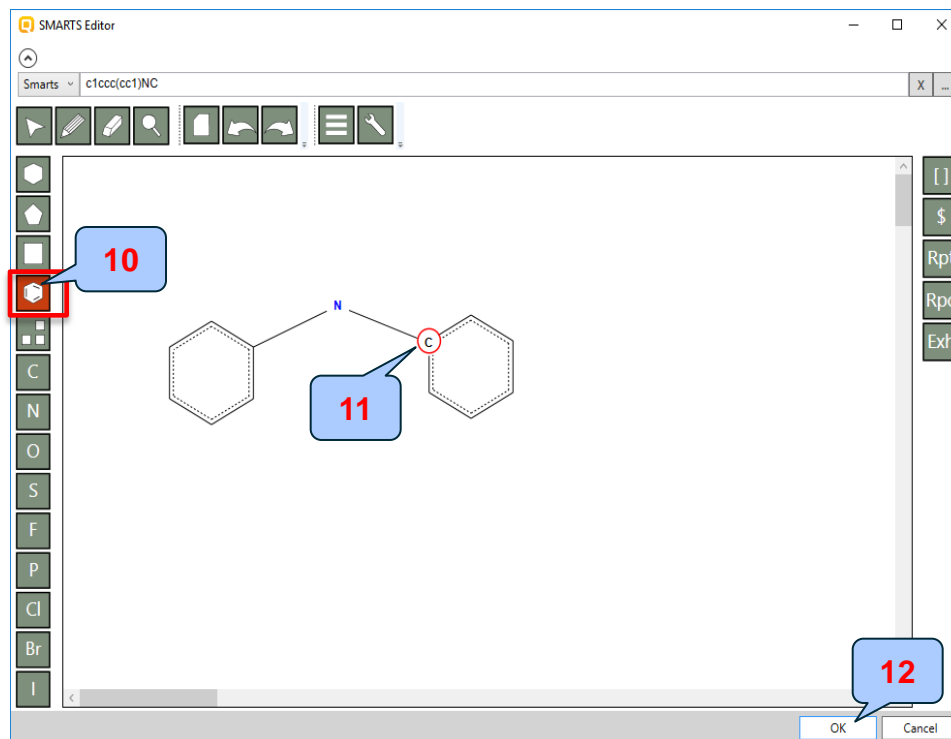
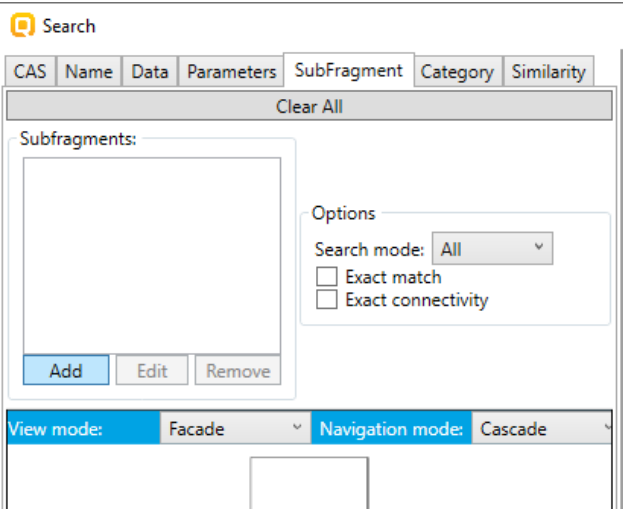
## *Procedure for defining*

### Example 1

*Substituted diphenyl amine*



N(c1ccccc1)c1ccccc1



10. Left click over benzene ring from the template in order to take benzene ring
11. Left click over the C atom in order to define second benzene ring to the NH group
12. Click **OK**.

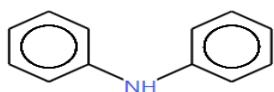
Continued on the next slide

# Subfragment search for identifying chemicals

## Procedure for defining

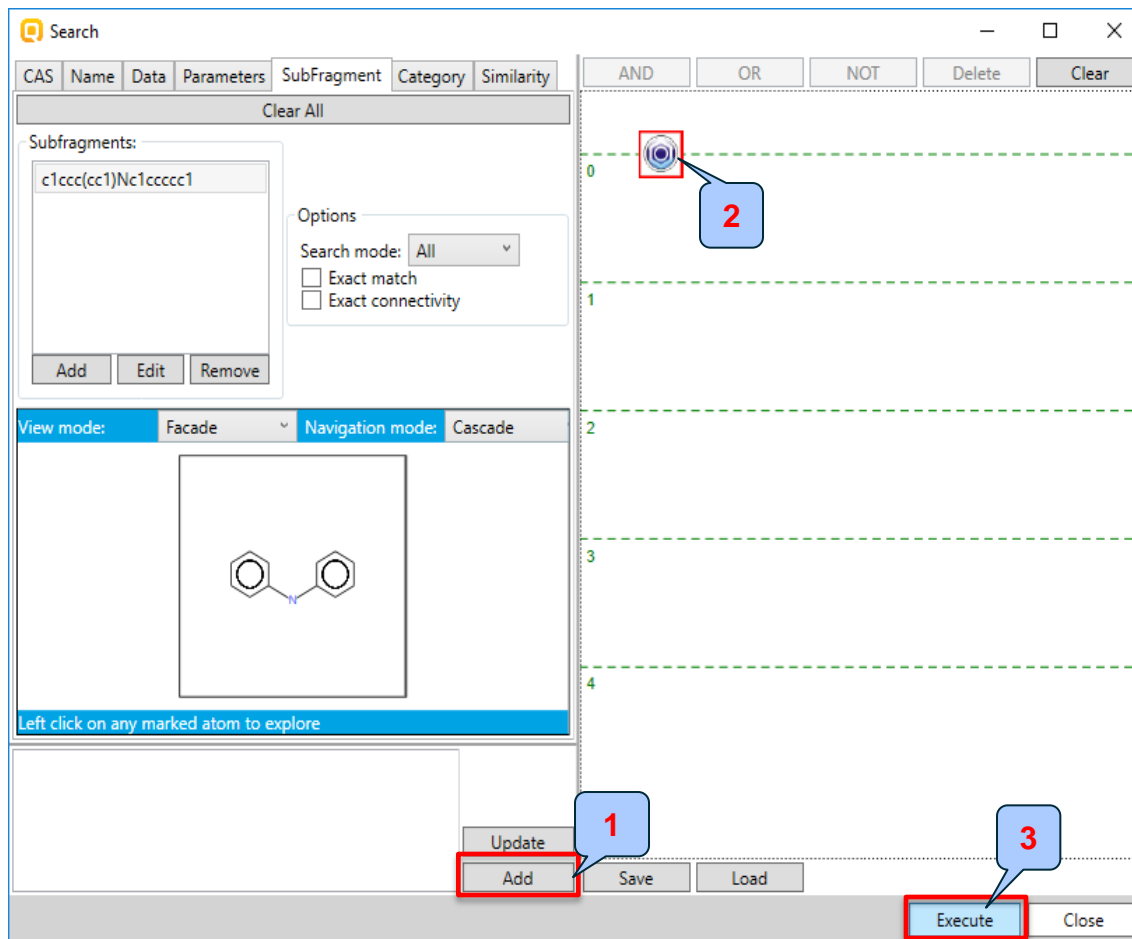
### Example 1

#### Substituted diphenyl amine



N(c1ccccc1)c1ccccc1

1. Click **Add** button;
2. The defined query appears on logic panel;
- Double click over the query or click
3. **Execute** button to execute it.



# Subfragment search for identifying chemicals

## *Procedure for defining*

### Example 1

#### *Substituted diphenyl amine*

The screenshot shows the QSAR Toolbox interface. The 'Search' tab is active, and the 'Query' tool is used to define search criteria. A red box highlights the search results table, which lists 21 chemicals matching the criteria. Below the table, a red arrow points from a chemical structure (a diphenyl amine) to a more complex substituted diphenyl amine structure, illustrating the structural criteria used in the search.

| EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        | EC Number        |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 36431-22-8       | 1324-76-1        | 68442-68-2       | 70571-81-2       | 81-77-6          | 41317-15-1       | 29512-49-0       | 101-67-7         | 101-72-4         | 298-46-4         | 90-30-2          | 836-30-6         | 74-31-7          | ...              | ...              |
| High             | ...              | ...              | ...              | ...              | ...              | ...              | ...              | ...              | ...              | ...              | ...              | ...              | ...              | ...              |
| Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent | Mono constituent |

Structural criteria

1. The Query tool (QT) identified 21 chemicals matching the desired criteria

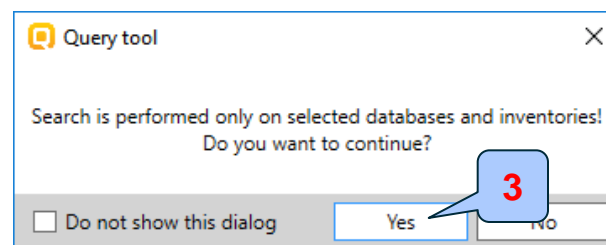
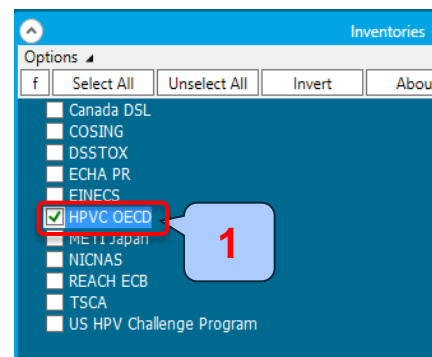
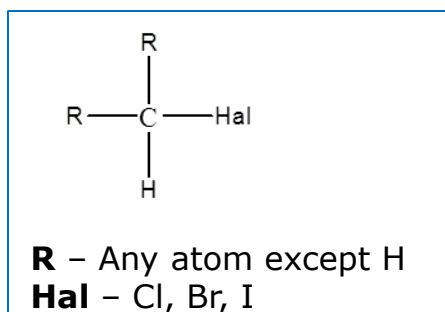
# Subfragment search for identifying chemicals

## *Aliphatic halogens*

### Example 2

Search for structures that meet the structural requirements:

- Aliphatic halogens*



1. Select the **HPVC OECD** inventory (no database has been selected in this case, only HPVC OECD) located under *Data* section;
2. Click **Query** button under **Input** module;
3. Click **Yes**

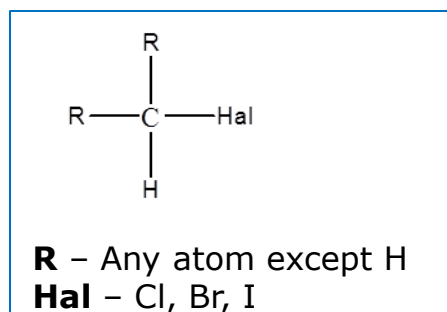
# Subfragment search for identifying chemicals

## *Aliphatic halogens*

### Example 2

Search for structures that meet the structural requirements:

- Aliphatic halogens*



- HPVC OECD** inventory has been selected
- Search for discrete chemicals

- Query** panel appears;
- Go to **SubFragment** panel;
- Click **"Add"** button;
- "SMARTS Editor" window appears;

The screenshot shows the QSAR Toolbox interface. The 'Query' panel is active, and the 'SubFragment' tab is selected. The 'Add' button is highlighted. The 'SMARTS Editor' window is open, showing the 'Rectangle' selection tool. Numbered callouts 1 through 4 indicate the steps: 1. Query panel, 2. SubFragment panel, 3. Add button, 4. SMARTS Editor window.



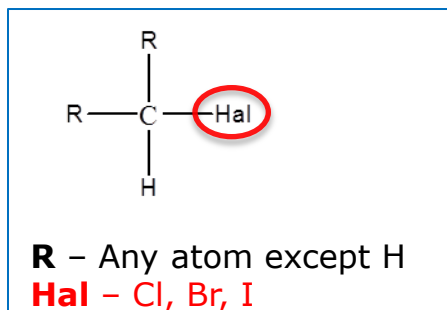
# Subfragment search for identifying chemicals

## *Aliphatic halogens*

### Example 2

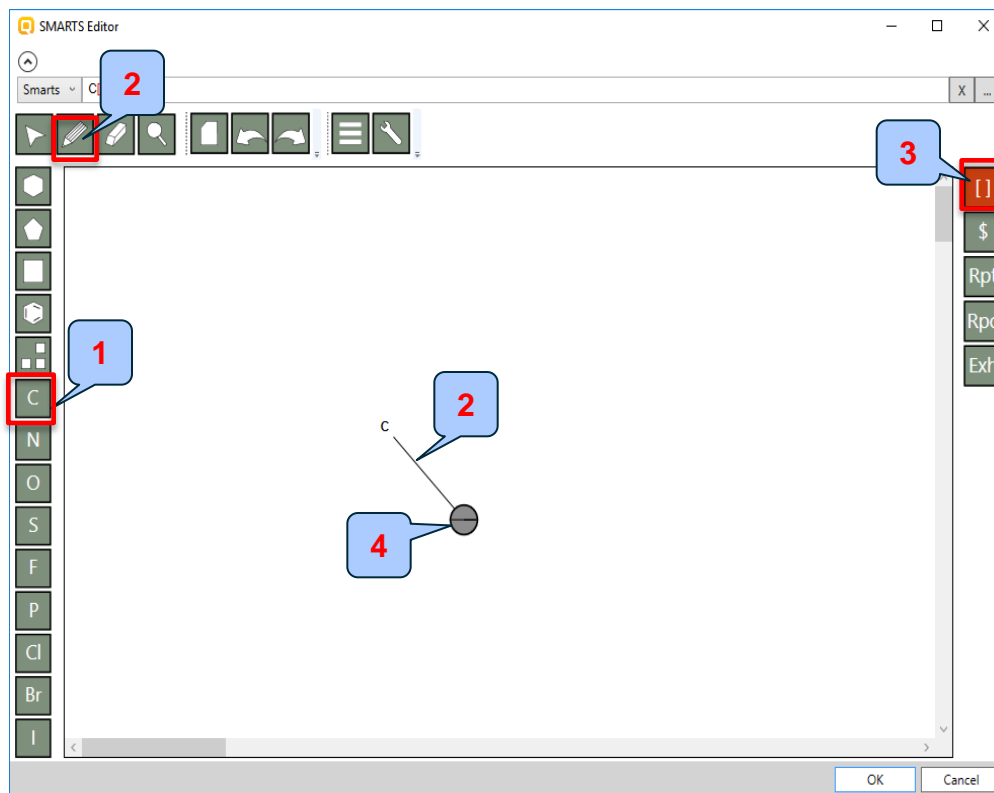
Search for structures that meet the structural requirements:

- Aliphatic halogens*



- Define **Enumeration** fragment with three members (halogen atoms)

1. Left click the **C** symbol from the template and then left click in the drawing panel;
2. Left click the **pencil** button and draw a single bond;
3. Select the "[ ]" (enumeration) button;
4. Click over one of the carbon atoms;



Continued on the next slide

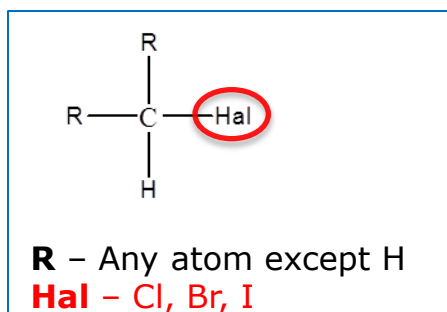
# Subfragment search for identifying chemicals

## *Aliphatic halogens*

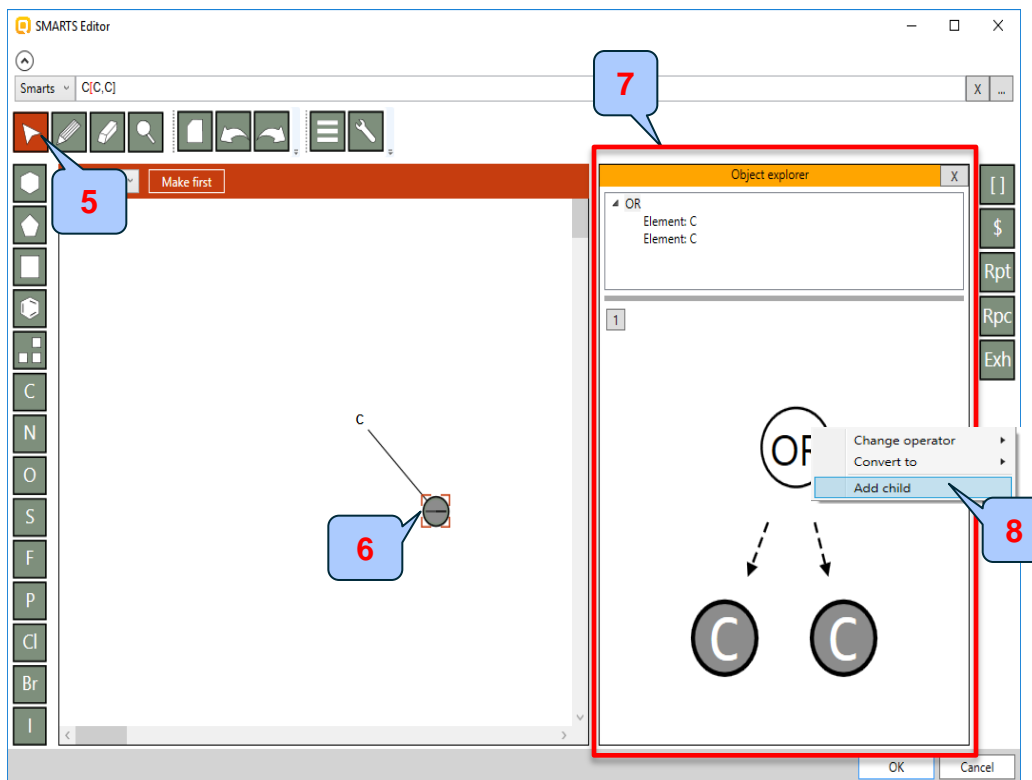
### Example 2

Search for structures that meet the structural requirements:

- Aliphatic halogens*



- Define **Enumeration** fragment with three members (halogen atoms)



- Click the **selection tool**;
- Click the new object;
- Object explorer* panel appears;
- Right click over the **OR** expression node and select **Add child**;

Continued on the next slide

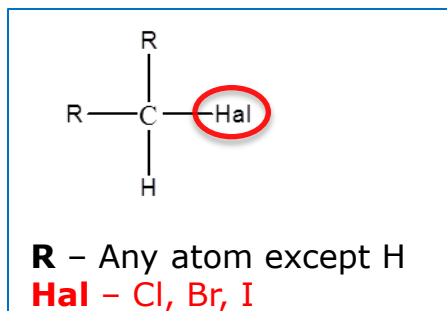
# Subfragment search for identifying chemicals

## Aliphatic halogens

### Example 2

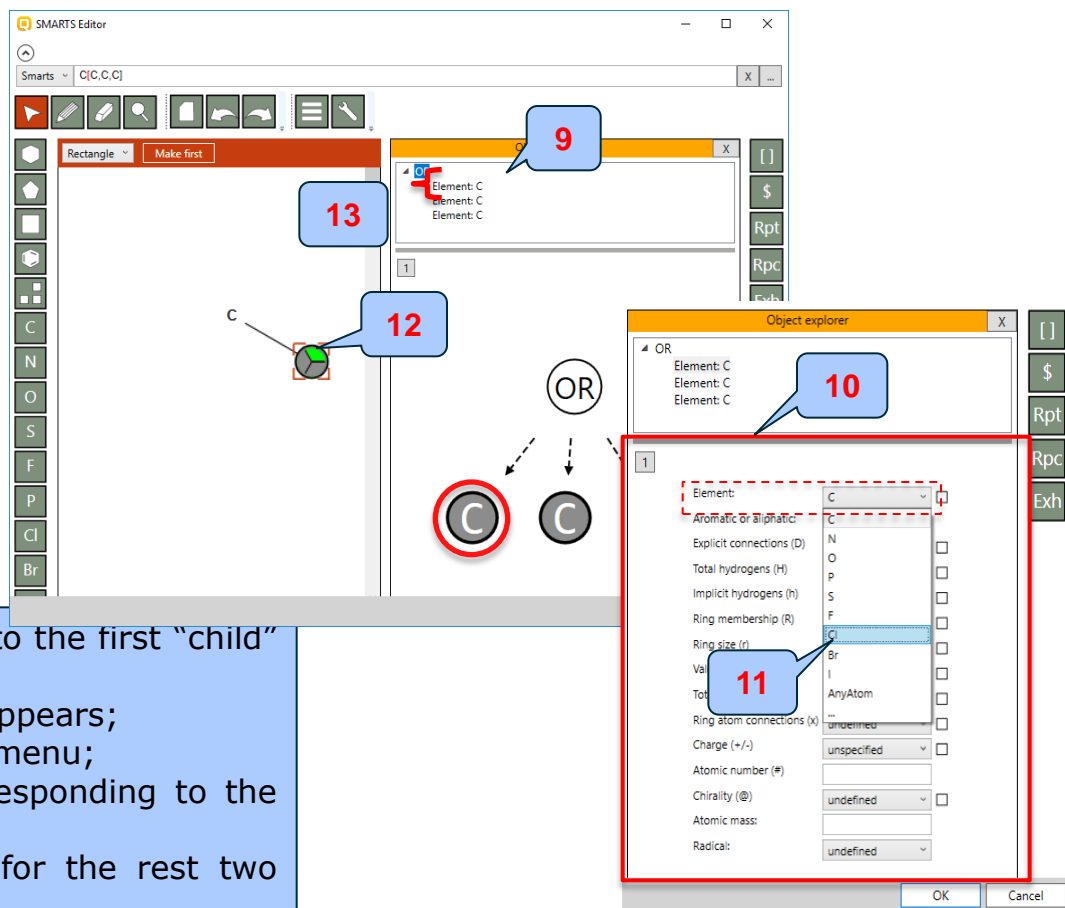
Search for structures that meet the structural requirements:

- Aliphatic halogens



- Define **Enumeration** fragment with three members (halogen atoms)

- Click the first element which corresponds to the first “child” query;
- The panel with the atom characteristics appears;
- Select “**Cl**” from the *Element* drop-down menu;
- The part of the SMARTS fragment corresponding to the first element have been colored in green.
- Select “**Br**” and “**I**” in the same way for the rest two elements , successively.



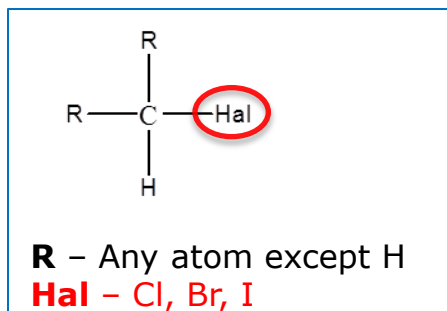
# Subfragment search for identifying chemicals

## *Aliphatic halogens*

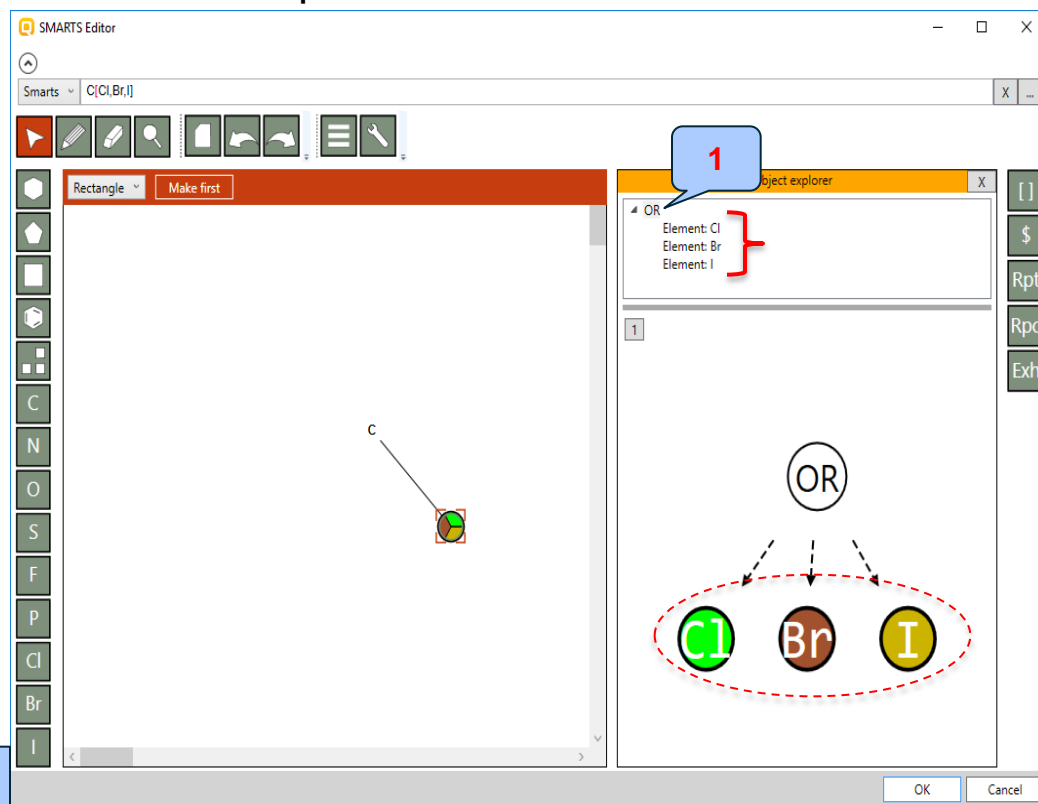
### Example 2

Search for structures that meet the structural requirements:

- Aliphatic halogens*



- Define **Enumeration** fragment with three members (halogen atoms)



**Hal** substituent is ready when all elements are selected. Click on OR (1). Now we can continue with defining of the rest substituents of the carbon atom (see the general structure above).

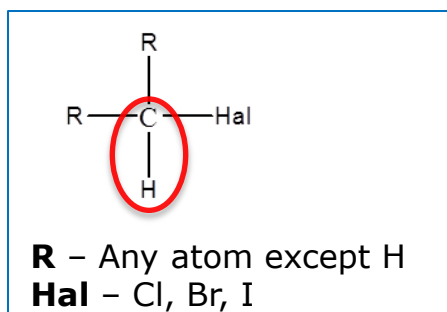
# Subfragment search for identifying chemicals

## Aliphatic halogens

### Example 2

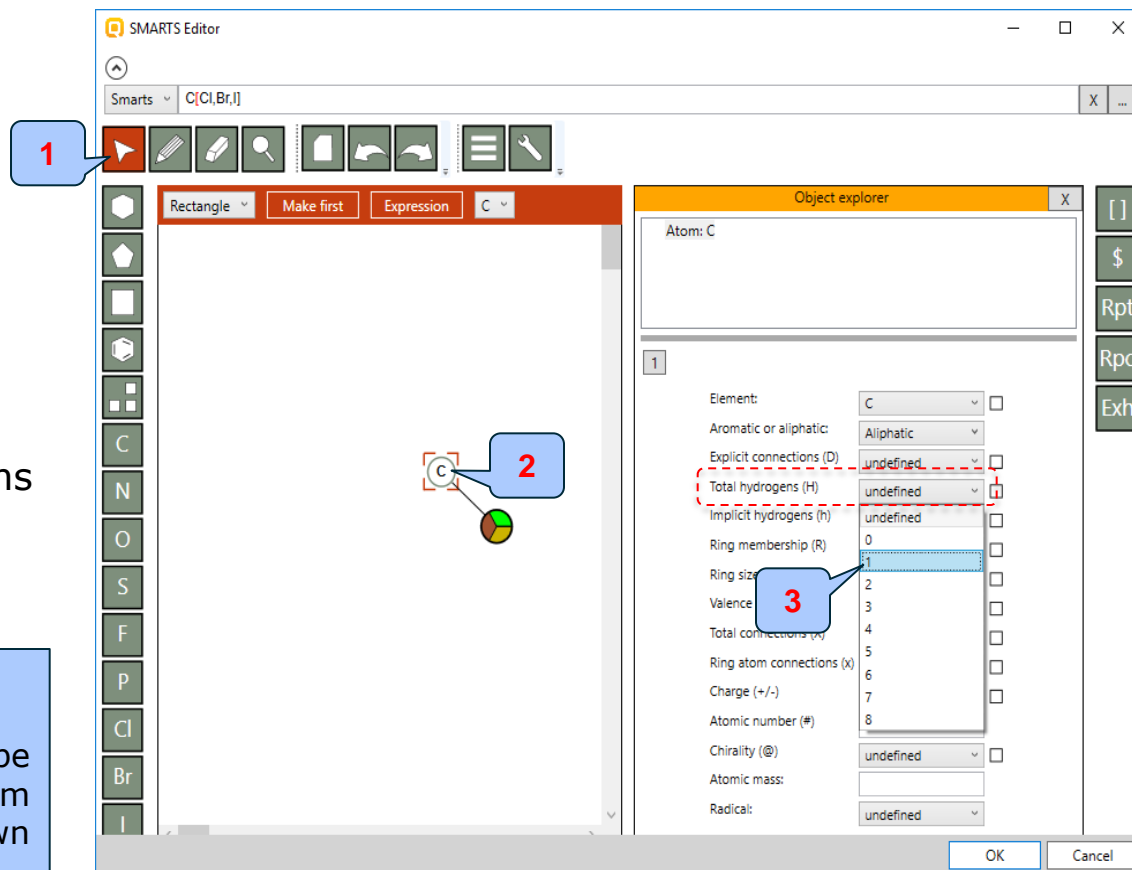
Search for structures that meet the structural requirements:

- Aliphatic halogens



- Define the number of hydrogens bonded to **C atom**

- Click **Selection tool**;
- Click **C atom**;
- Define this carbon atom to be bonded with total one H atom from the *Total hydrogens* drop-down menu



SMARTS Editor

Smarts: C[Cl,Br,I]

Object explorer

Atom: C

1

Element: C

Aromatic or aliphatic: Aliphatic

Explicit connections (D): undefined

Total hydrogens (H): 1

Implicit hydrogens (h): undefined

Ring membership (R): 0

Ring size: 2

Valence: 3

Total connections (x): 4

Ring atom connections (x): 5

Charge (+/-): 7

Atomic number (#): 8

Chirality (@): undefined

Atomic mass:

Radical:

OK Cancel

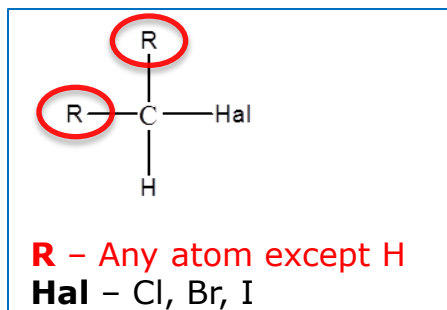
# Subfragment search for identifying chemicals

## *Aliphatic halogens*

### Example 2

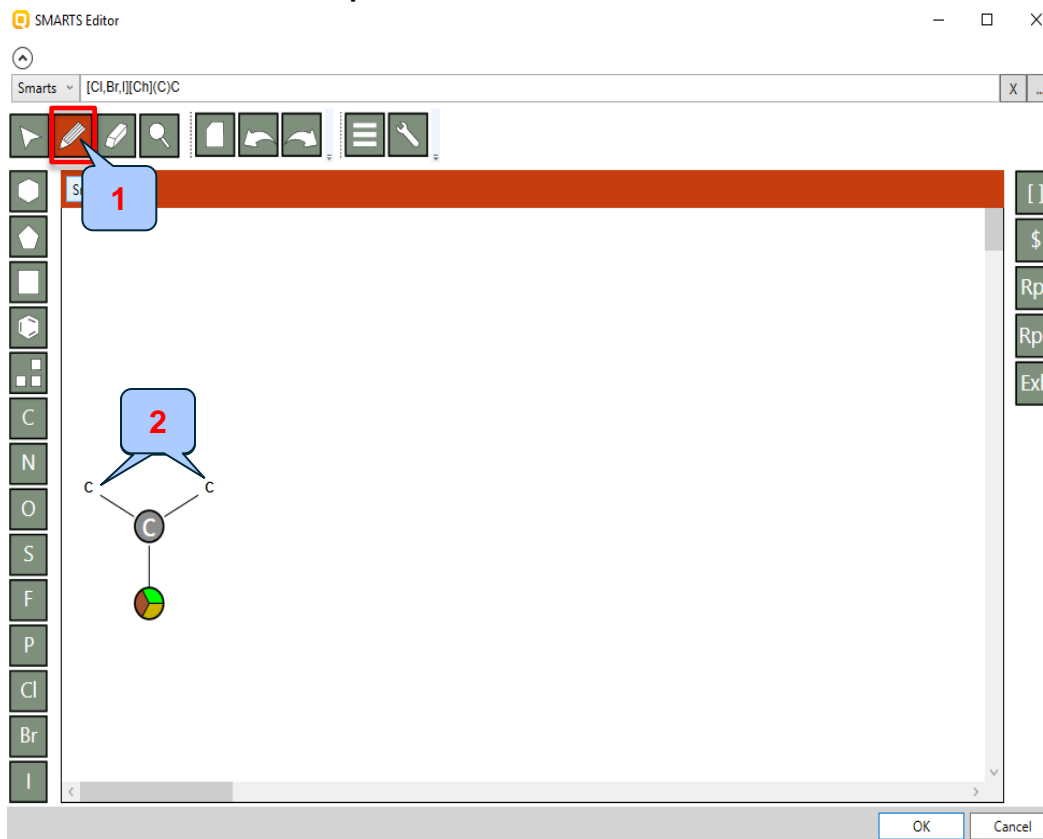
Search for structures that meet the structural requirements:

- *Aliphatic halogens*



- Define fragment R including any type atom except H atom

1. Select the **pencil** button;
2. Draw two **single bonds** to the C atom;



Continued on the next slide

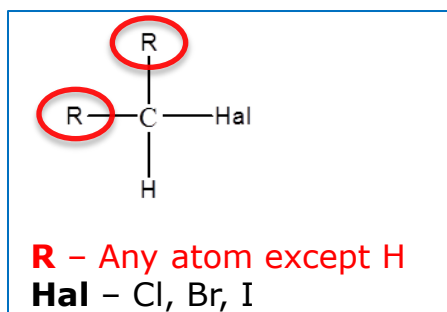
# Subfragment search for identifying chemicals

## Aliphatic halogens

### Example 2

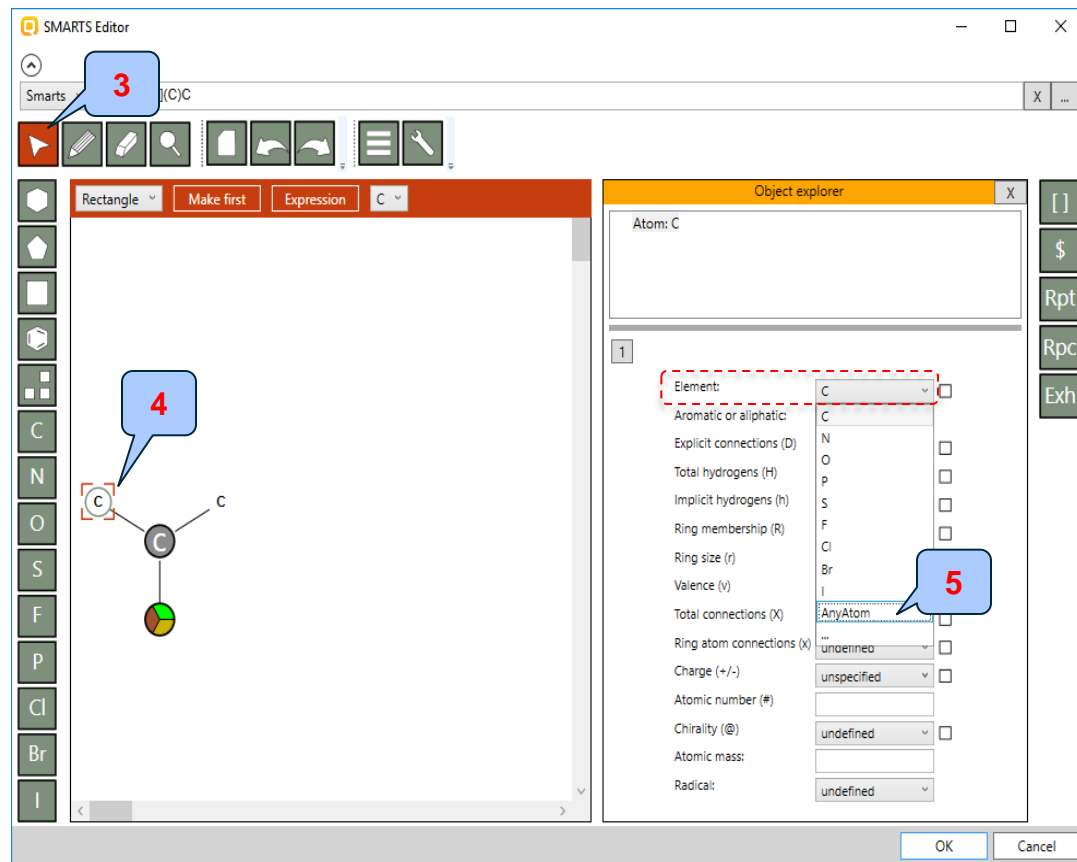
Search for structures that meet the structural requirements:

- *Aliphatic halogens*



- Define fragment R including any type atom except H atom

3. Click **Selection tool**;
4. Select C atom;
5. Select "**Any atom**" from the *Element* drop-down menu;



Continued on the next slide

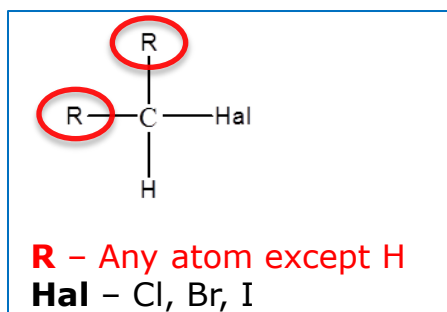
# Subfragment search for identifying chemicals

## Aliphatic halogens

### Example 2

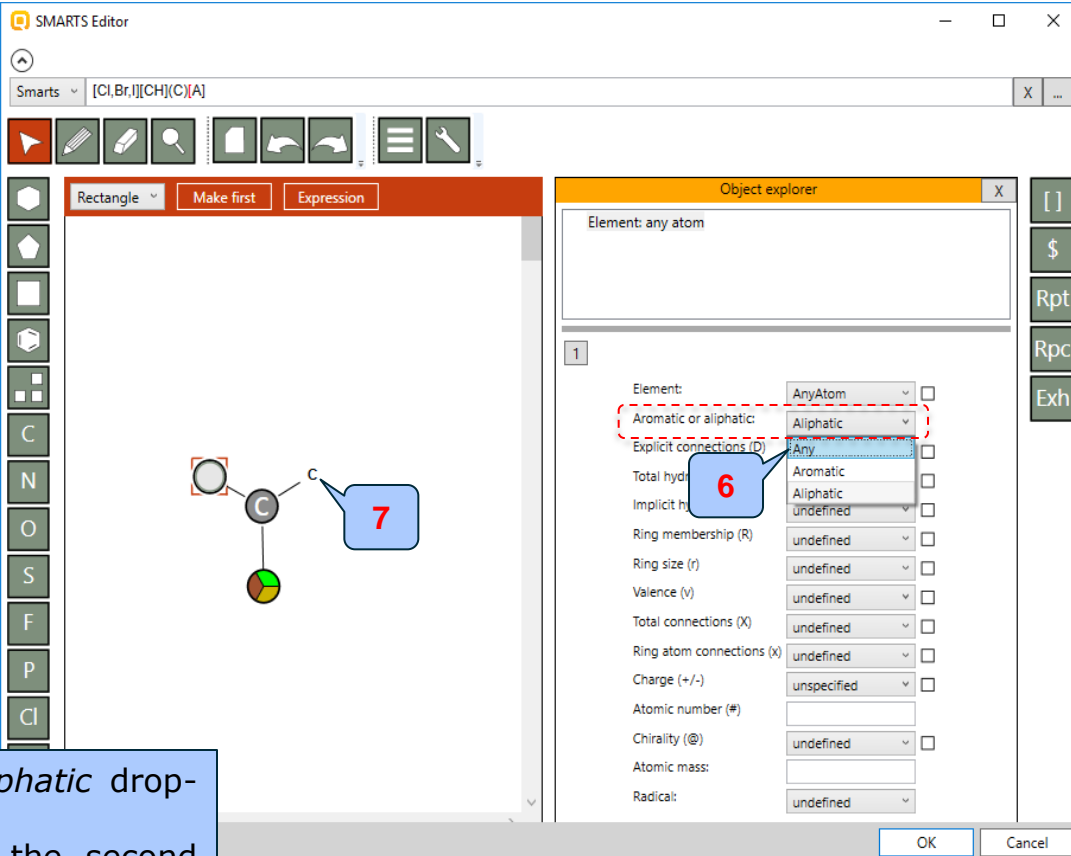
Search for structures that meet the structural requirements:

- Aliphatic halogens



- Define fragment R including any type atom except H atom

6. Select **Any** from the *Aromatic or aliphatic* drop-down menu;
7. Repeat the steps from 3 to 6 for the second carbon atom



SMARTS Editor

Smarts: [Cl,Br,I][CH](C)[A]

Object explorer

Element: any atom

1

Element: AnyAtom ☐

Aromatic or aliphatic: Aliphatic ☐

Explicit connections (D): Any ☐

Total hydrogens (H): Any ☐

Implicit hydrogens (I): Any ☐

Ring membership (R): undefined ☐

Ring size (r): undefined ☐

Valence (v): undefined ☐

Total connections (X): undefined ☐

Ring atom connections (x): undefined ☐

Charge (+/-): unspecified ☐

Atomic number (#):

Chirality (@): undefined ☐

Atomic mass:

Radical: undefined ☐

OK Cancel



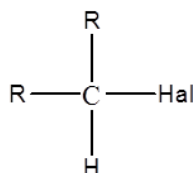
# Subfragment search for identifying chemicals

## *Aliphatic halogens*

### Example 2

Search for structures that meet the structural requirements:

- Aliphatic halogens*



**R** – Any atom except H

**Hal** – Cl, Br, I

- Define skeleton of target compound

When all substituents of the carbon atom are defined (Hal, H, R) click on **OK** (1).

SMARTS Editor

Smarts [Cl,Br,I][CH]([\*])[\*]

Rectangle Make first Expression C

Object explorer

Atom: AnyAtom

1

|                           |             |                          |
|---------------------------|-------------|--------------------------|
| Element:                  | AnyAtom     | <input type="checkbox"/> |
| Aromatic or aliphatic:    | Any         | <input type="checkbox"/> |
| Explicit connections (D)  | undefined   | <input type="checkbox"/> |
| Total hydrogens (H)       | undefined   | <input type="checkbox"/> |
| Implicit hydrogens (h)    | undefined   | <input type="checkbox"/> |
| Ring membership (R)       | undefined   | <input type="checkbox"/> |
| Ring size (r)             | undefined   | <input type="checkbox"/> |
| Valence (v)               | undefined   | <input type="checkbox"/> |
| Total connections (X)     | undefined   | <input type="checkbox"/> |
| Ring atom connections (x) | undefined   | <input type="checkbox"/> |
| Charge (+/-)              | unspecified | <input type="checkbox"/> |
| Atomic number (#)         |             | <input type="checkbox"/> |
| Chirality (@)             |             | <input type="checkbox"/> |
| Atomic mass:              |             | <input type="checkbox"/> |

1

OK Cancel

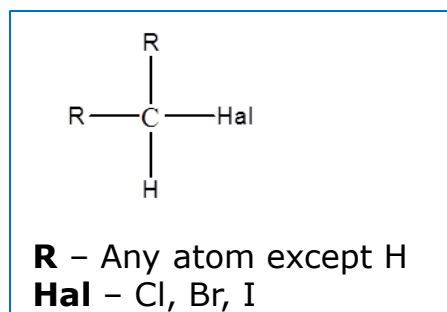
# Subfragment search for identifying chemicals

## *Aliphatic halogens*

### Example 2

Search for structures that meet the structural requirements:

- Aliphatic halogens*



- Define skeleton of target compound

1. The defined general structure of the aliphatic halogens appears on the separate plot;
2. Click **Add** button;
3. Double click the query to perform the search

# Subfragment search for identifying chemicals

## Aliphatic halogens

### Example 2

- Aliphatic halogens

1

| 1                                | 2                   | 3                   | 4                     | 5                  | 6                 | 7                   | 8                    | 9                    | 10                  | 11                 | 12                   | 13                  | 14      |
|----------------------------------|---------------------|---------------------|-----------------------|--------------------|-------------------|---------------------|----------------------|----------------------|---------------------|--------------------|----------------------|---------------------|---------|
|                                  |                     |                     |                       |                    |                   |                     |                      |                      |                     |                    |                      |                     |         |
| EC Number:2536...                | EC Number:2008...   | EC Number:2011...   |                       |                    | EC Number:2120... | EC Number:2446...   |                      | EC Number:2008...    | EC Number:2416...   | EC Number:2018...  | EC Number:2451...    | EC Nur              |         |
| 121776-57-6                      | 37764-25-3          | 75-34-3             | 79-00-5               | 68527-01-5         | 25637-99-4        | 760-23-6            | 21850-44-2           | 3194-55-6            | 75-45-6             | 17639-93-9         | 88-66-4              | 22591-21-5          | 98-87-  |
| Low                              | High                | High                | High                  | Moderate           | Moderate          | High                | High                 | Moderate             | High                | High               | High                 | High                | High    |
| Ethanone, 2,2-di...              | 2,2-dichloro-N,N... | 1,1-Dichloroetha... | 1,1,2-Trichloroeth... | Alkenes, C12-30... | Cyclododecane...  | "3,4-dichloro-1-... | "benzene, 1,1'-(1... | 1,2,5,6,9,10-Hexa... | chloro(difluoro)... | methyl 2-chloro... | 1-chloro-2-(dichl... | 1,1-Dichloro-3,3... | (dichlo |
| C11H13Cl2NO3                     | C8H11Cl2NO          | C2H4Cl2             | C2H3Cl3               | C12H22BrCl         | C12H18Br6         | C4H6Cl2             | C21H20Br8O2          | C12H18Br6            | CHClF2              | C4H7ClO2           | C7H5Cl3              | C6H10Cl2O           | C7H6C   |
| Mono constituent                 | Mono constituent    | Mono constituent    | Mono constituent      | UVCB               | Mono constituent  | Mono constituent    | Mono constituent     | Mono constituent     | Mono constituent    | Mono constituent   | Mono constituent     | Mono constituent    | Mono    |
| CC1(C)OC(CN1C...CIC(C)C(=O)NC... | CC(C)Cl             | ClCC(C)Cl           | CCCC(C)C(Br)CC...     | BrC1CC(Br)CC(Br... | ClCC(C)C=C        | CC(C)(c1cc(Br)cf... | BrC1CCC(Br)C(Br...   | FC(F)Cl              | COC(=O)C(C)Cl       | ClC(C)C1CCCC1Cl    | CC(C)C(C(=O)C...     | ClC(C)Cl            |         |

1. The Query tool (QT) identified 40 chemicals matching the desired criteria.

## Subfragment search for identifying chemicals

*BCF and (Dmax)*

### Example 3

Search for structures that meet the structural requirements:

- Endpoint is BCF <2000 L/kg
- 3D parameter – minimum value of Diameter maximum >17 Å

# Subfragment search for identifying chemicals

## BCF and (*D*<sub>max</sub>)

### Example 3

- Structures search criteria
  - Endpoint is BCF < 2000 L/kg
  - 3D parameter – minimum value of Diameter maximum >17Å\*

The screenshot shows the QSAR Toolbox Query tool interface. The 'Databases' panel on the left (1) has 'Environmental Fate and Transport' selected, with 'Bioaccumulation Canada', 'Bioaccumulation fish CEFIC LRI', and 'Bioconcentration NITE' checked. The 'Query tool' dialog (3) shows 'Do you want to continue?' with 'No' selected (4). The 'Search' window (5) has the 'Clear' button highlighted. The 'Data' panel (6) is active, showing the 'Endpoint definition' (7) with 'Environmental Fate and Transport' and 'Bioaccumulation: aquatic' selected, and 'BCF' checked. The 'Metadata' panel (8) shows 'BCFss lipid' selected (9). The 'Data' panel (10) shows 'Mean value:' set to '<' (8) and '2000' (9), with 'Unit' set to 'Specific...' and 'L/kg' (10). The 'Add' button (11) is highlighted. The 'Execute' button is also visible.

1. Select databases including BCF data (Bioaccumulation Canada; Bioaccumulation fish CEFIC LRI and Bioconcentration NITE). No inventory has been selected in this exercise (keep them unselected);
2. Click **Query** button;
3. Confirm that the searches will be performed on selected databases and inventories;
4. Click **No** on the message asking to restore source (HPVC OECD inventory) used in the previous example.
5. Click **Clear** button to clean the desk from previous exercise;
6. Go to **Data** panel;
7. Expand the endpoint tree and select BCF;
8. Specify qualifier "<" and add 2000 in the *Mean value* field (9);
10. Select scale "**Specific volume**" and select the appropriate unit (**L/kg**);
11. Click **Add** button;

# Subfragment search for identifying chemicals

## BCF and (*D*max)

### Example 3

- Structures search criteria
  - Endpoint is  $BCF < 2000 \text{ L/kg}$
  - 3D parameter – minimum value of Diameter maximum  $> 17 \text{ \AA}$

- Click on **Parameters** panel;
- Select **Diameter maximum** from the list;
- Select qualifier " $\geq$ " from the drop down list and add "**17**" in the *Expression* field;
- Unit of the *D*max parameter is "A", it is selected by default;
- Click **Add** button.

The screenshot shows the 'Search' window in the QSAR Toolbox. The 'Parameters' tab is selected, indicated by a red box and callout 1. In the 'Parameter name' list, 'Diameter maximum' is selected, highlighted by a red box and callout 2. Below this, the 'Diameter maximum' section shows the 'Expression' field with a dropdown set to '>' and the value '17', highlighted by a red box and callout 3. The 'unit' dropdown is set to 'A' (Angstroms), highlighted by a red box and callout 4. At the bottom, the 'Add' button is highlighted by a red box and callout 5. The right side of the window shows a logical expression builder with 'AND', 'OR', and 'NOT' operators, and a 'Clear' button. The bottom of the window has 'Save', 'Load', 'Execute', and 'Close' buttons.

# Subfragment search for identifying chemicals

## BCF and (*D*<sub>max</sub>)

### Example 3

- Structures search criteria
  - Endpoint is BCF < 2000 L/kg
  - 3D parameter – minimum value of Diameter maximum >17 Å

Both queries should be linked together by logical **AND**. For this purpose both queries should be selected first. How to do this:

- Right click** over the first query to select it (the selected boundary should become orange colored);



Selected query



Not selected query

- When the both queries are selected, **click "AND"** button;
- The two queries are combined by logical "AND";
- Double click **"AND"** query or click **"Execute"** button to execute the search;

The figure consists of three screenshots of the QSAR Toolbox Search interface, illustrating the steps to combine two queries using the AND operator.

- Top Screenshot:** Shows two separate search queries. The first query is selected, indicated by an orange border around its icon (labeled 1). The second query is not selected, indicated by a blue border around its icon (labeled 2).
- Middle Screenshot:** Shows the two queries combined by the logical "AND" operator. The "AND" button is highlighted with a red box (labeled 3).
- Bottom Screenshot:** Shows the final result of the search. The "AND" query is highlighted with a red box (labeled 4), and the "Execute" button is also highlighted with a red box.

# Subfragment search for identifying chemicals

## *BCF and ( $D_{\max}$ )*

### Example 3

The screenshot displays the QSAR Toolbox interface. The top menu bar includes options like Document, Single Chemical, Chemical List, Search, and Target Endpoint. The left sidebar shows a list of documents, with the first document named '1 [C: 47; Md: 0; P: 0] Query'. A red box highlights a table of 13 chemical structures. A blue callout with the number '1' points to the document name in the sidebar.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|----|----|----|----|
|   |   |   |   |   |   |   |   |   |    |    |    |    |

The Query tool (QT) identified 47 chemicals matching the desired criteria. The list of identified chemicals appear automatically on data matrix. Number of found chemicals is marked in the name of the document (1). In order to check the correctness of the performed query search, the user should calculate  $D_{\max}$  for the list of chemicals and to gather BCF data for them. The last two steps are presented on the next two slides.



# Subfragment search for identifying chemicals *BCF* and (*Dmax*)

## Example 3

**Step I**

**Step II**

**Step 1:** Expand the **Parameters** node of the endpoint tree;

**Step 2:** Expand the **3D** node;

**Step 3:** Right click over the row corresponding to the **Diameter maximum** parameter;

**Step 4:** Select **Calculate Diameter maximum for all chemicals** from the appeared menu;

**Step 5:** Once the values are calculated right click over the Diameter maximum row and select **Sort** values by **Ascending** order.

# Subfragment search for identifying chemicals *BCF and ( $D_{max}$ )*

## Example 3

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

2. The databases are already selected, click **Gather** data;

3. Data appear on data matrix; Click **OK**;

4. All 47 chemicals have BCF data less than 2000 L/kg (log 3.3 L/kg);

## Subfragment search for identifying chemicals

*LC 50 and predefined category (Aldehydes)*

### Example 4

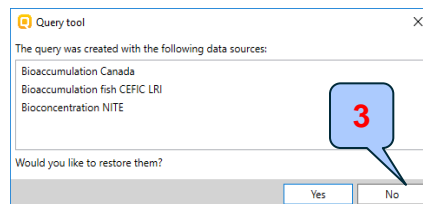
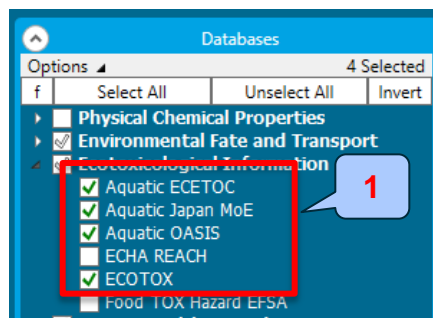
- Structures search criteria
  - Endpoint: LC 50 <1 mg/l
  - Fish: *P.promelas*
  - Effect: *Mortality*
  - Predefined category:  
*aldehydes*

# Subfragment search for identifying chemicals

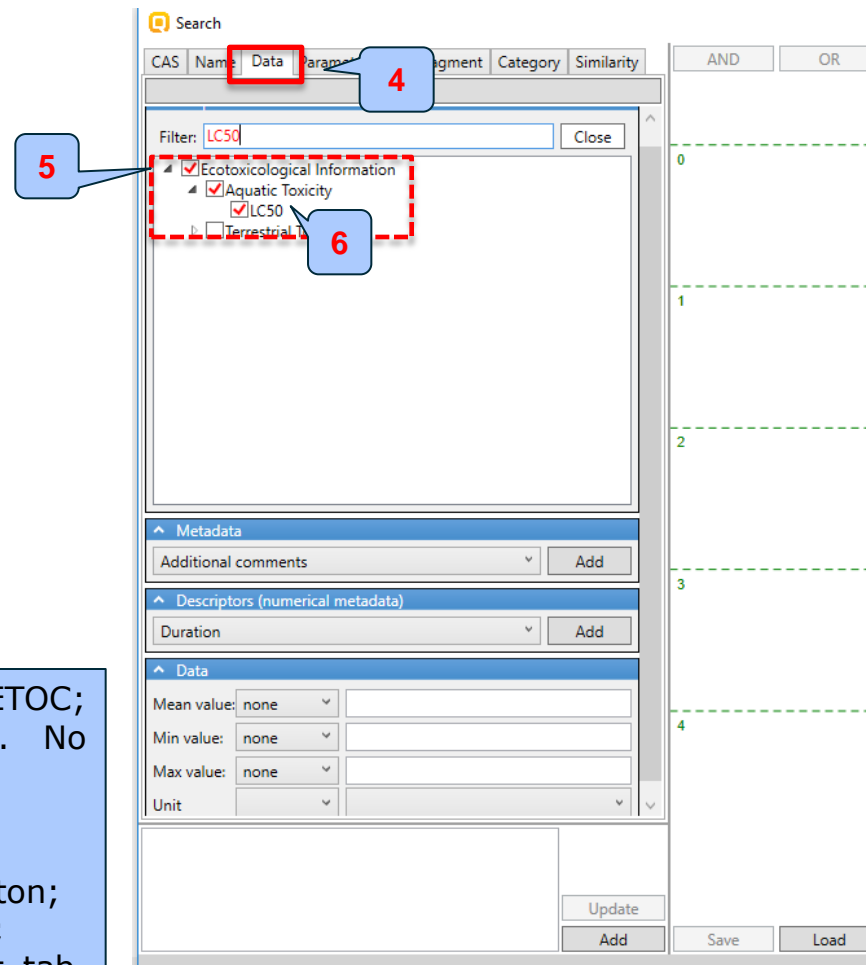
## LC 50 and predefined category (Aldehydes)

### Example 4

- Structures search criteria
  - Endpoint: LC 50 <1 mg/l
  - Fish: *P.promelas*
  - Effect: *Mortality*
  - Predefined category: *aldehydes*



1. Select databases related to LC50 data (Aquatic ECETOC; Aquatic Japan MoE; Aquatic OASIS; ECOTOX). No inventories has been selected in this case;
2. Click **Query** tool button;
3. Select **"No"**;
4. Go to **Data** panel, but before that click on **Clear** button;
5. Type in filter the name of searched endpoint (**LC50**);
6. Select **Aquatic Toxicity** and expand the endpoint tab, check **LC50**



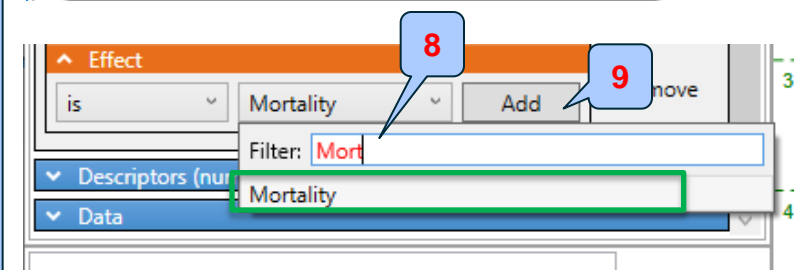
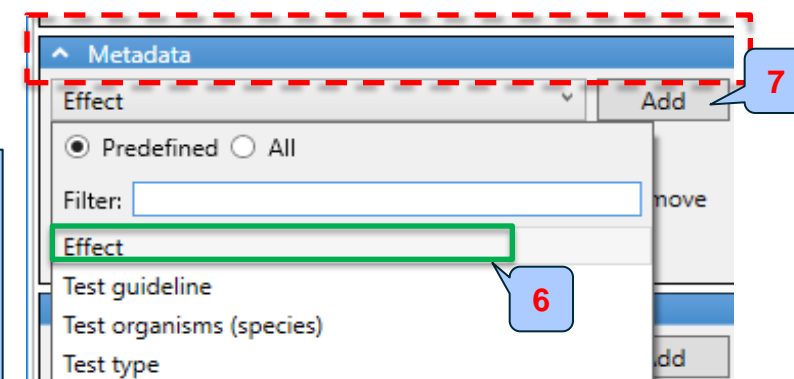
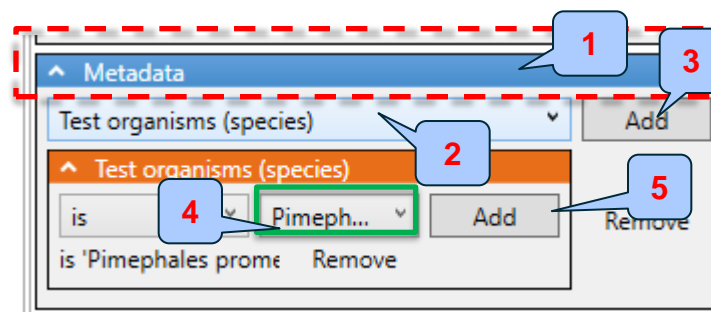
# Subfragment search for identifying chemicals

## LC 50 and predefined category (Aldehydes)

### Example 4

- Structures search criteria
  - Endpoint: LC 50 <1 mg/l
  - Fish: *P.promelas*
  - Effect: *Mortality*
  - Predefined category: *aldehydes*

1. Open **Metadata** field
2. Select "**Test organism (species)**" from the drop-down menu;
3. Click **Add**;
4. Select ***Pimephales promelas*** from the drop-down menu;
5. Click **Add**;
6. Go back to **Metadata** field and select **Effect** from the drop-down menu;
7. Click **Add**;
8. Use filter to find "**Mortality**" from the appeared drop-down menu;
9. Click **Add**



# Subfragment search for identifying chemicals

## LC 50 and predefined category (Aldehydes)

### Example 4

- Structures search criteria
  - Endpoint: LC 50 <1 mg/l
  - Fish: *P.promelas*
  - Effect: *Mortality*
  - Predefined category: *aldehydes*

The screenshot shows the 'Search' window in the QSAR Toolbox. The 'Data' tab is selected. The 'Data' panel is expanded, showing 'Mean value' set to '<' and 'Unit' set to 'mg/L'. The 'Filter' dropdown is open, showing 'Mass concentration' selected. The 'Add' button is visible at the bottom right.

Numbered callouts indicate the steps:

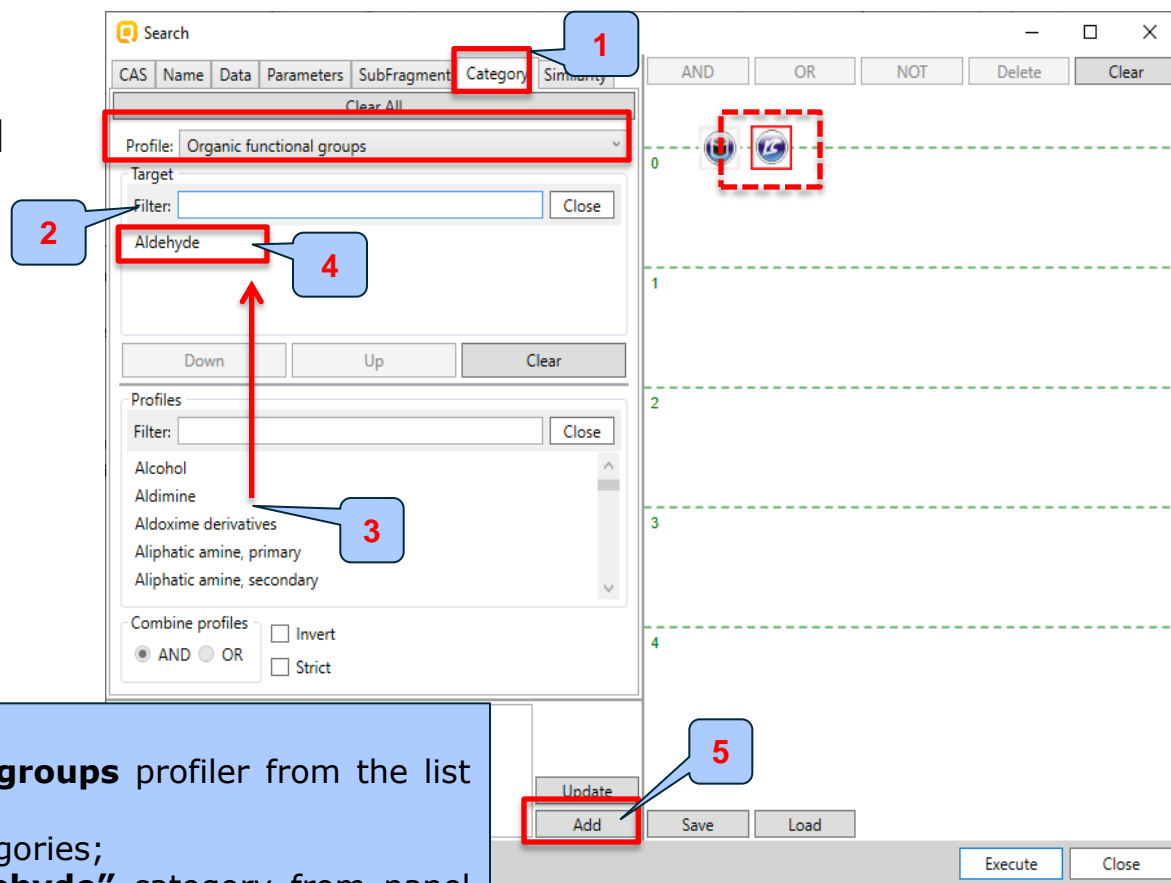
- Open **Data** panel;
- Open qualifiers from *Mean value* section and select "<";
- Type "1" in the blank field;
- From the *Unit* section use Filter to find "**Mass concentration**";
- Select "**mg/l**" from the list with units;
- Click **Add**;

# Subfragment search for identifying chemicals

## LC 50 and predefined category (Aldehydes)

### Example 4

- Structures search criteria
  - Endpoint: LC 50 <1 mg/l
  - Fish: *P.promelas*
  - Effect: *Mortality*
  - Predefined category: *aldehydes*



1. Open **Category** panel;
2. Use *Filter* to find **Organic functional groups** profiler from the list with profilers and select it;
3. Find "**Aldehyde**" from the list with categories;
4. Click "**Up**" to move the selected "**Aldehyde**" category from panel "*Profilers*" to the panel "*Target*"
5. The selected category appears in the panel "*Target*" and finally click **Add** button. The query with predefined category appears on the logic panel.

# Subfragment search for identifying chemicals

## LC 50 and predefined category (Aldehydes)

### Example 4

- Structures search criteria
  - Endpoint: LC 50 <1 mg/l
  - Fish: *P.promelas*
  - Effect: *Mortality*
  - Predefined category: *aldehydes*

The screenshot shows the QSAR Toolbox Search interface. The 'SubFragment' tab is selected. The 'Profile' is set to 'Organic functional groups' and the 'Target' is 'Aldehyde'. The 'Parameters' tab is also visible, showing 'LC 50 <1 mg/l'. The 'AND' button is highlighted with a red box and a blue callout labeled '2'. The 'Execute' button is highlighted with a red box and a blue callout labeled '4'. The 'Execute' button is also highlighted with a red box and a blue callout labeled '4'.

The both queries should be combined by logical "AND". Follow the steps explained on slide # 39

1. Right click over the both queries to highlight them;
2. Click **AND** button;
3. The two queries are combined by logical "AND";
4. Click **Execute** button in order to execute the Query



# Subfragment search for identifying chemicals

## *LC 50 and predefined category (Aldehydes)*

### Example 4

The QSAR TOOLBOX interface displays a subfragment search for identifying chemicals. The 'Filter endpoint tree' is open, and a red box highlights the 'Structure' node. A blue callout with the number '1' points to the 'Structure' node. The 'Chemical List' tab is active, displaying a table with 6 columns (1-6) and 6 rows of chemical structures. The 'Documents' panel on the left shows a list of documents, with 'Document 1' selected. The 'Query' tool (QT) is used to identify 6 chemicals matching the criteria.

The Query tool (QT) identified 6 chemicals matching the desired criteria. The list of identified chemicals appear automatically on the data matrix. Number of found chemicals is marked in the name of the node (1). In order to check the correctness of the performed query search the user should gather ecotox data and profile according to OFG Profiler. The last two steps are presented on the next two slides.

# Subfragment search for identifying chemicals

## LC 50 and predefined category (Aldehydes)

### Example 4

1. Go to the **Data** module

2. Databases are already selected, click gather data;

3. Click **OK** to read all data;

4. Confirm the number of gathered data.

5. Data appear on data matrix answering the following criteria:

5a. LC50 < 1mg/l    5b. Effect: Mortality    5c. Species: *P.promelas*

5d. All 6 chemicals are categorized as "Aldehydes" according to OFG profiler

| Chemical | LC50               | Category |
|----------|--------------------|----------|
| 1        | M: 0.366 mg/L      | Aldehyde |
| 2        | M: 0.3 mg/L        | Aldehyde |
| 3        | M: 0.76 (0.65÷0.9) | Aldehyde |
| 4        | M: 0.845 mg/L      | Aldehyde |
| 5        | M: 0.767 mg/L      | Aldehyde |
| 6        | M: 0.465 mg/L      | Aldehyde |

## Subfragment search for identifying chemicals

*Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment*

### Example 5

- Structures search criteria

- Endpoint: Gene mutation
- Species: *S.thyphimurium*
- Data: *Positive*

*AND*

- Endpoint: *Summary carcinogenicity*
- Species: *Rat*
- Route: *gavage*
- Data: *Positive*
- Predefined category: *Epoxides*

# Subfragment search for identifying chemicals

## Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment

### Example 5

#### •Structures search criteria

- Endpoint: Gene mutation
- Species: *S.thyphimurium*
- Data: *Positive*

AND

- Endpoint: *Summary carcinogenicity*
- Species: *Rat*
- Route: *gavage*
- Data: *Positive*
- Predefined category: *Epoxides*

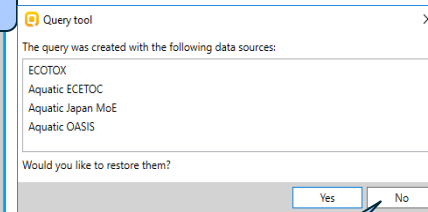
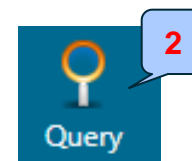
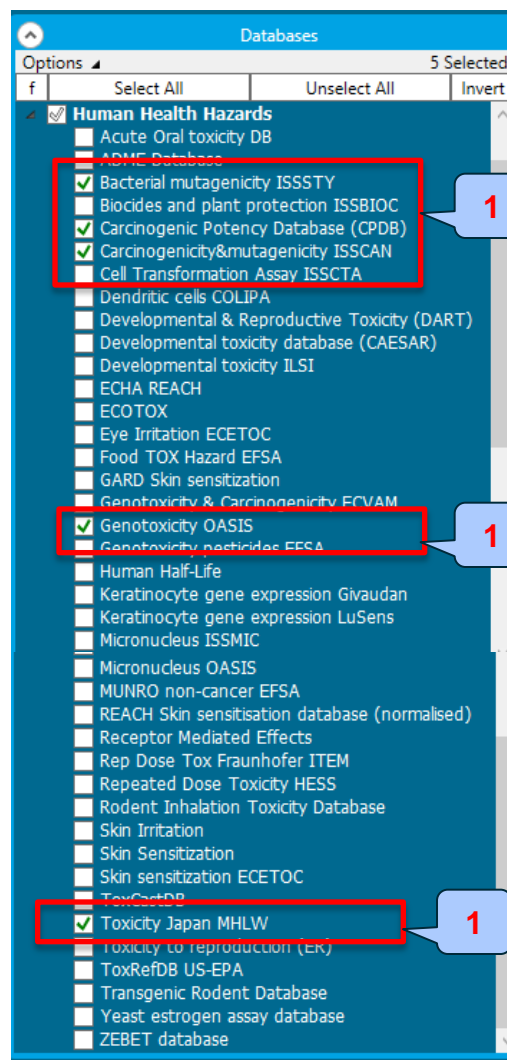
Before application of Query tool, the user should select databases which includes required data. Inventories are not selected in this particular case.

1. Select the following databases:

- Bacterial mutagenicity ISSSTY
- Carcinogenicity Potency Database (CPDB)
- Carcinogenicity&mutagenicity ISSCAN
- Genotoxicity OASIS
- Toxicity Japan MHLW

2. Click **Query** button;

3. Select **No** to not restore the databases used in the previous example.



# Subfragment search for identifying chemicals

## Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment

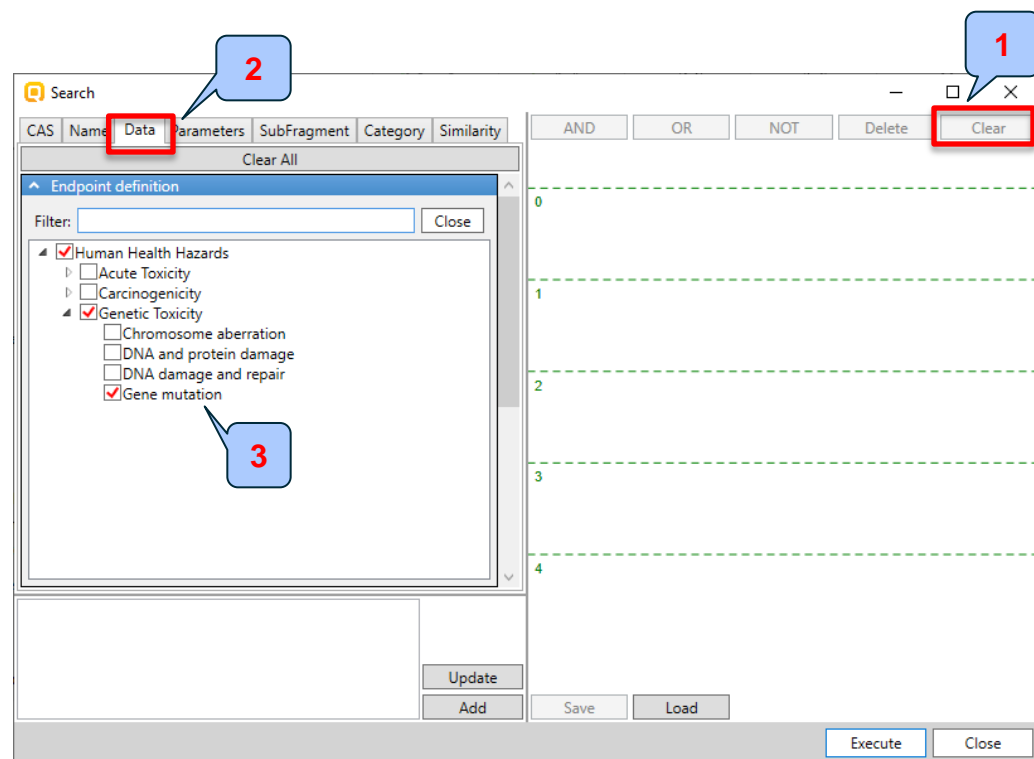
### Example 5

#### Structures search criteria

- Endpoint: Gene mutation
- Species: *S.thyphimurium*
- Data: *Positive*

AND

- Endpoint: *Summary carcinogenicity*
- Species: *Rat*
- Route: *gavage*
- Data: *Positive*
- Predefined category: *Epoxides*



1. First click on **Clear** button
2. Click on the **Data** panel;
3. Open Human health hazard>>Genetic Toxicity and select **Gene mutation**;

# Subfragment search for identifying chemicals

## Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment

### Example 5

#### • Structures search criteria

- Endpoint: Gene mutation
- Species: *S.thyphimurium*
- Data: *Positive*
- Endpoint: *Summary carcinogenicity*
- Species: *Rat*
- Route: gavage
- Predefined category: *Epoxides*

1. Open Metadata panel;
2. Select **Test organism (species)** from the pop-up menu;
3. Click "Add";
4. In a newly appeared field select **S.thyhimurium** from the pop-up-menu
5. Click **Add**;
6. Open **Data** panel;
7. Select **Gene mutation I** scale (belonging to the Mutagenicity family) from the *Units* drop-down menu;
8. Select **Positive** from the list with scale members;
9. Click **Add** button;

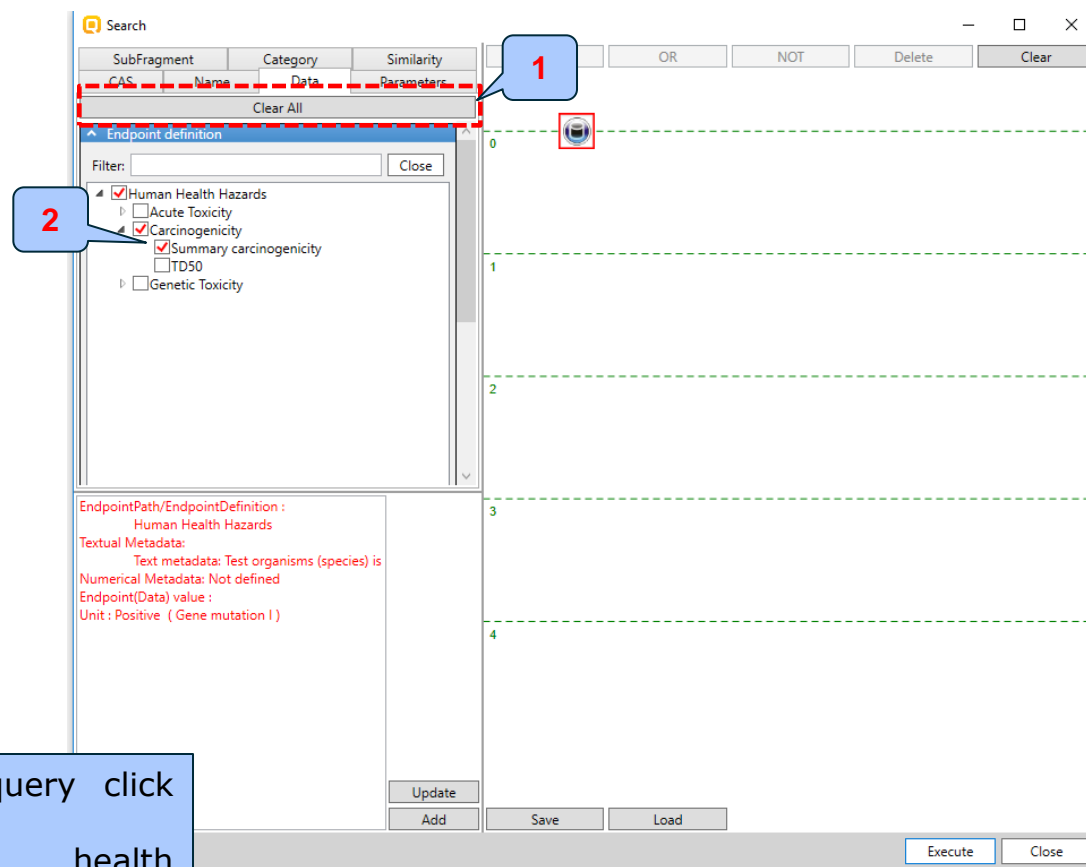
The screenshot shows the 'Search' window of the QSAR Toolbox. The interface is divided into several panels: 'Metadata', 'Descriptors (numerical metadata)', and 'Data'. The 'Metadata' panel is expanded, showing a list of search criteria. The 'Data' panel is also expanded, showing a list of search criteria. The 'Test organisms (species)' field is highlighted with a red dashed box and a callout '2'. The 'Test organisms (species)' field is also highlighted with a red dashed box and a callout '4'. The 'Test organisms (species)' field is also highlighted with a red dashed box and a callout '5'. The 'Test organisms (species)' field is also highlighted with a red dashed box and a callout '6'. The 'Test organisms (species)' field is also highlighted with a red dashed box and a callout '7'. The 'Test organisms (species)' field is also highlighted with a red dashed box and a callout '8'. The 'Test organisms (species)' field is also highlighted with a red dashed box and a callout '9'.

# Subfragment search for identifying chemicals

## *Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment*

### Example 5

- Structures search criteria
  - Endpoint: Gene mutation
  - Species: *S.thyphimurium*
  - Data: *Positive*
  - Endpoint: *Summary carcinogenicity*
  - Species: *Rat*
  - Route: gavage
  - Predefined category: *Epoxides*



- Before defining the second query click **Clear All** button;
- Open path: Human health hazard >> Carcinogenicity and select *Summary carcinogenicity* node

# Subfragment search for identifying chemicals

## Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment

### Example 5

- Structures search criteria
  - Endpoint: Gene mutation
  - Species: *S.thyphimurium*
  - Data: *Positive*
  - Endpoint: *Summary carcinogenicity*
  - Species: *Rat*
  - Route: *gavage*
  - Predefined category: *Epoxides*

1. Open *Metadata* panel;

2. Select **Route of administration** from pop-up menu;

3. Click **Add** button;

4. Select **Gavage** from the list. Click **Add** button;

5. Go back to *Metadata* panel and select **Test organism (species)** from the pop-up menu. Click **Add** button;

6. Select **Rat** from the list. Click **Add** button;

7. Open *Data* panel

8. Select scale **Carcinogenicity I (ISSCAN)**;

9. Select **Positive**;

10. Click **Add** button;

11. The query appears on the logic panel.



# Subfragment search for identifying chemicals

## Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment

### Example 5

#### Structures search criteria

- Endpoint: Gene mutation
- Species: *S.thyphimurium*
- Data: *Positive*
- Endpoint: *Summary carcinogenicity*
- Species: *Rat*
- Route: gavage
- Predefined category: *Epoxides*

1. Click **Category** panel;
2. Select **Carcinogenicity (genotox and nongenotox) alerts by ISS**;
3. Select category **Epoxides and aziridines (Genotox)**;
4. Click **Up** to move the selected category to the panel Target;
5. Click **Add** button;
6. The query appears on the logic panel;

The screenshot shows the 'Search' window in the QSAR Toolbox. The 'Category' panel is active, displaying a list of predefined categories. The 'Target' panel is empty. The 'Profiles' panel shows a list of profiles, including 'Epoxides and aziridines (Genotox)'. The 'Add' button is highlighted. The 'Logic' panel shows the resulting query: 'Carcinogenicity (genotox and nongenotox) alerts by ISS AND Epoxides and aziridines (Genotox)'. The 'Execute' button is visible at the bottom right.

Numbered callouts indicate the steps:

1. Click **Category** panel;
2. Select **Carcinogenicity (genotox and nongenotox) alerts by ISS**;
3. Select category **Epoxides and aziridines (Genotox)**;
4. Click **Up** to move the selected category to the panel Target;
5. Click **Add** button;
6. The query appears on the logic panel;

# Subfragment search for identifying chemicals

## *Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment*

### Example 5

#### •Structures search criteria

- Endpoint: Gene mutation
- Species: *S.thyphimurium*
- Data: *Positive*
- Endpoint: *Summary carcinogenicity*
- Species: *Rat*
- Route: gavage
- Predefined category: *Epoxides*

The screenshot shows the 'Search' window in the QSAR Toolbox. The 'SubFragment' tab is active. The 'Profile' is set to 'Carcinogenicity (genotox and nongenotox) alerts by ISS'. The 'Target' is 'Epoxides and aziridines (Genotox)'. The 'Filter' is empty. The 'Profiles' list includes: (N/A), (Poly) Halogenated Cycloalkanes (Nongenotox), 1,3-Benzodioxoles (Nongenotox), Acyl halides (Genotox), Aliphatic azo and azoxy (Genotox), Aliphatic halogens (Genotox), Aliphatic N-nitro group (Genotox), and Alkenylbenzenes (Genotox). The 'Combine profiles' section has 'AND' selected. The 'Execute' button is highlighted with a blue callout labeled '4'. A red box highlights the three queries and the 'AND' button, with callouts '1', '2', and '3' indicating the steps: right-clicking the queries, clicking the 'AND' button, and keeping 'AND' selected.

1. Right click over the three queries (see slide # 39);
2. Click **AND** button;
3. The three queries are combined by logical "**AND**" Keep the "**AND**" selected (clicked);
4. Click **Execute** button;

# Subfragment search for identifying chemicals

*Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment*

## Example 5

The screenshot displays the QSAR Toolbox software interface. The 'Filter endpoint tree...' window is open, showing a list of 6 chemicals. The table below shows the chemical structures and their corresponding data.

| Chemical                         | 1 | 2 | 3 | 4 | 5 | 6 |
|----------------------------------|---|---|---|---|---|---|
| Structure                        |   |   |   |   |   |   |
| Structure info                   |   |   |   |   |   |   |
| Parameters                       |   |   |   |   |   |   |
| Physical Chemical Properties     |   |   |   |   |   |   |
| Environmental Fate and Transport |   |   |   |   |   |   |
| Ecotoxicological Information     |   |   |   |   |   |   |
| Human Health Hazards             |   |   |   |   |   |   |

The Query tool (QT) identified 6 chemicals matching the desired criteria. In order to check the correctness of the query the user should gather experimental data for Ames and Carcinogenicity and profile the identified chemicals according to applied Carcinogenicity profiler. The last two steps are presented on the next two slides.

# Subfragment search for identifying chemicals

## Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment

### Example 5

**1** Click **Data** in the top menu bar.

**2** Click **Gather** in the left sidebar.

**3** Click **OK** in the **Read data?** dialog box.

**4a** Summary carcinogenicity (gavage, rat): Positive

**4b** Ames mutagenicity (S. thuphimurium): Positive

178 points added across 6 chemicals.

**4** Click **OK** in the status bar.

1. Go to the **Data** module;
  2. Databases are already selected, click **Gather** data;
  3. Click **OK**;
  4. Click **OK**;
- Data appear on data matrix answering the following criteria:
- 4a. Summary carcinogenicity (gavage, rat): Positive
  - 4b. Ames mutagenicity (S. thuphimurium): Positive

# Subfragment search for identifying chemicals

## *Chemicals with Ames and Carcinogenicity positive data including Epoxide fragment*

### Example 5

1. Go to **Profiling**;

2. Click **Unselect All**;

3. Use filter and select *Carcinogenicity (genotox and nongenotox) alerts by ISS*;

4. Click **Apply**;

5. All 6 chemicals have "Epoxides and Aziridines" fragment within its structures.

## Subfragment search for identifying chemicals

*Chemicals with Skin sensitization positive data including Aldehyde fragment*

### Example 6

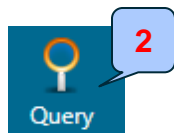
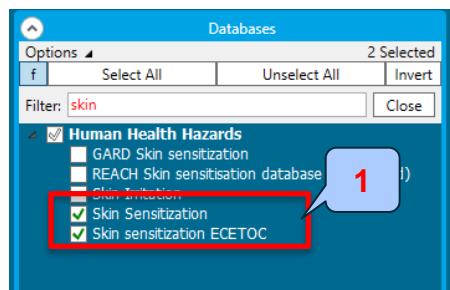
- Structures search criteria
  - Endpoint: *EC3*
  - Data: *Positive*
  - *SubFragment: Aldehyde*

# Subfragment search for identifying chemicals

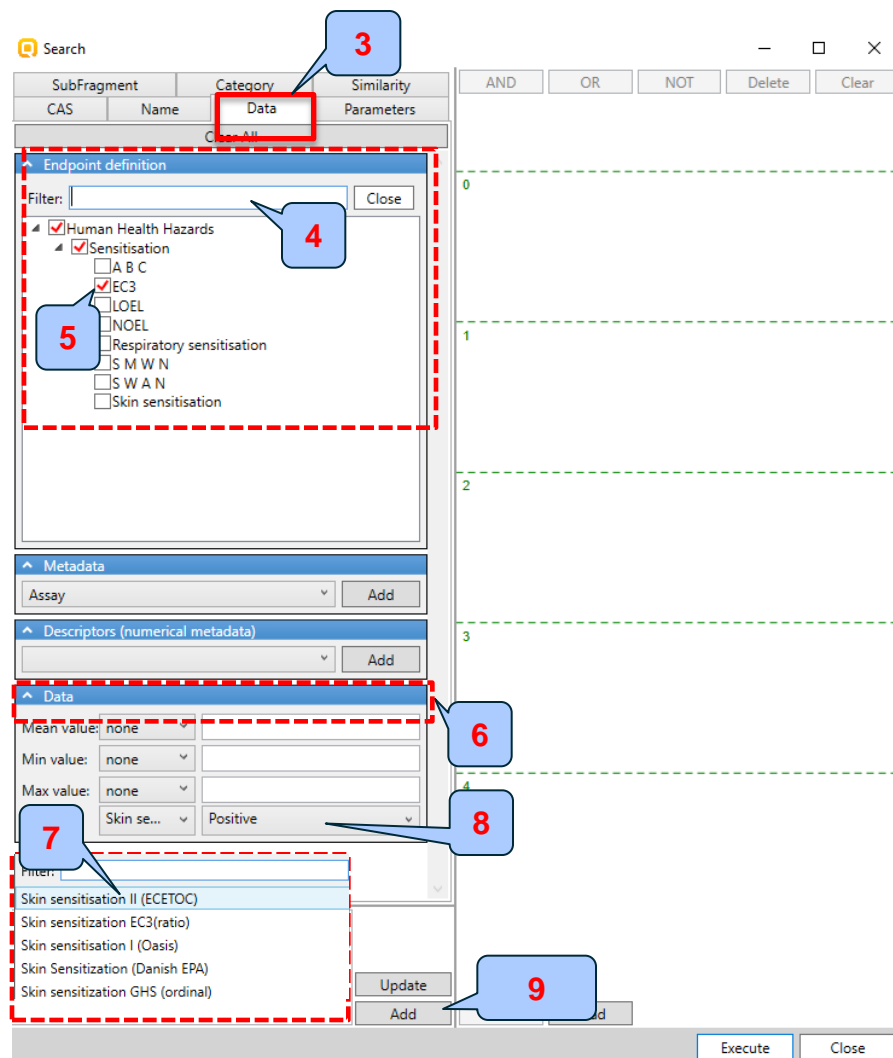
*Chemicals with Skin sensitization positive data including Aldehyde fragment*

## Example 6

- Structures search criteria
  - Endpoint: *EC3*
  - Data: *Positive*
  - SubFragment: Aldehydes



- Select databases including skin sensitization data. No inventories has been selected in this case;
- Click **Query tool** button and select **No** to confirm that the databases in previous example will not be restored;
- Click on **Data** panel;
- Use filter to find searched endpoint – EC3;
- Check **EC3**;
- Go to **Data** panel;
- Select scale **Skin sensitization II (ECETOC)**;
- Select **Positive**;
- Click **Add** button;



# Subfragment search for identifying chemicals

*Chemicals with Skin sensitization positive data including Aldehyde fragment*

## Example 6

- SubFragment: Aldehydes

1. Click on **Subfragment** panel

2. Click **Add**

3. Draw the fragment using the pencil and atoms available in the template panel (for more details see the refreshing slides 24-28). Keep in mind that:

3a. The central C atom should be "aliphatic"

3b. The central C atom should have one H atom;

4. Click **OK**.



# Subfragment search for identifying chemicals

*Chemicals with Skin sensitization positive data including Aldehyde fragment*

## Example 6

- Structures search criteria
  - Endpoint: *EC3*
  - Data: *Positive*
  - SubFragment: Aldehydes

- The drawn fragment appears in a new panel
- Click **Add** button;
- The query appears on the logic panel.

The screenshot displays the 'Search' window in the QSAR Toolbox. The 'SubFragment' tab is active, showing a list of subfragments with 'O=[CH]C' entered. A red box highlights this entry, with a callout '1' pointing to it. Below the list are 'Add', 'Edit', and 'Remove' buttons. To the right, the 'Options' section shows 'Search mode: All' and checkboxes for 'Exact match' and 'Exact connectivity'. A large red box highlights the 'View mode: Facade' and 'Navigation mode: Cascade' section, which contains a chemical structure editor showing the drawn aldehyde fragment (O=CH-C). A callout '2' points to the 'Add' button at the bottom right. On the right side, the logic panel shows a query tree with a callout '3' pointing to the 'AND' operator. At the bottom, there are buttons for 'Update', 'Add', 'Save', 'Load', 'Execute', and 'Close'.

# Subfragment search for identifying chemicals

*Chemicals with Skin sensitization positive data including Aldehyde fragment*

## Example 6

- Structures search criteria
  - Endpoint: *EC3*
  - Data: *Positive*
  - SubFragment: Aldehydes

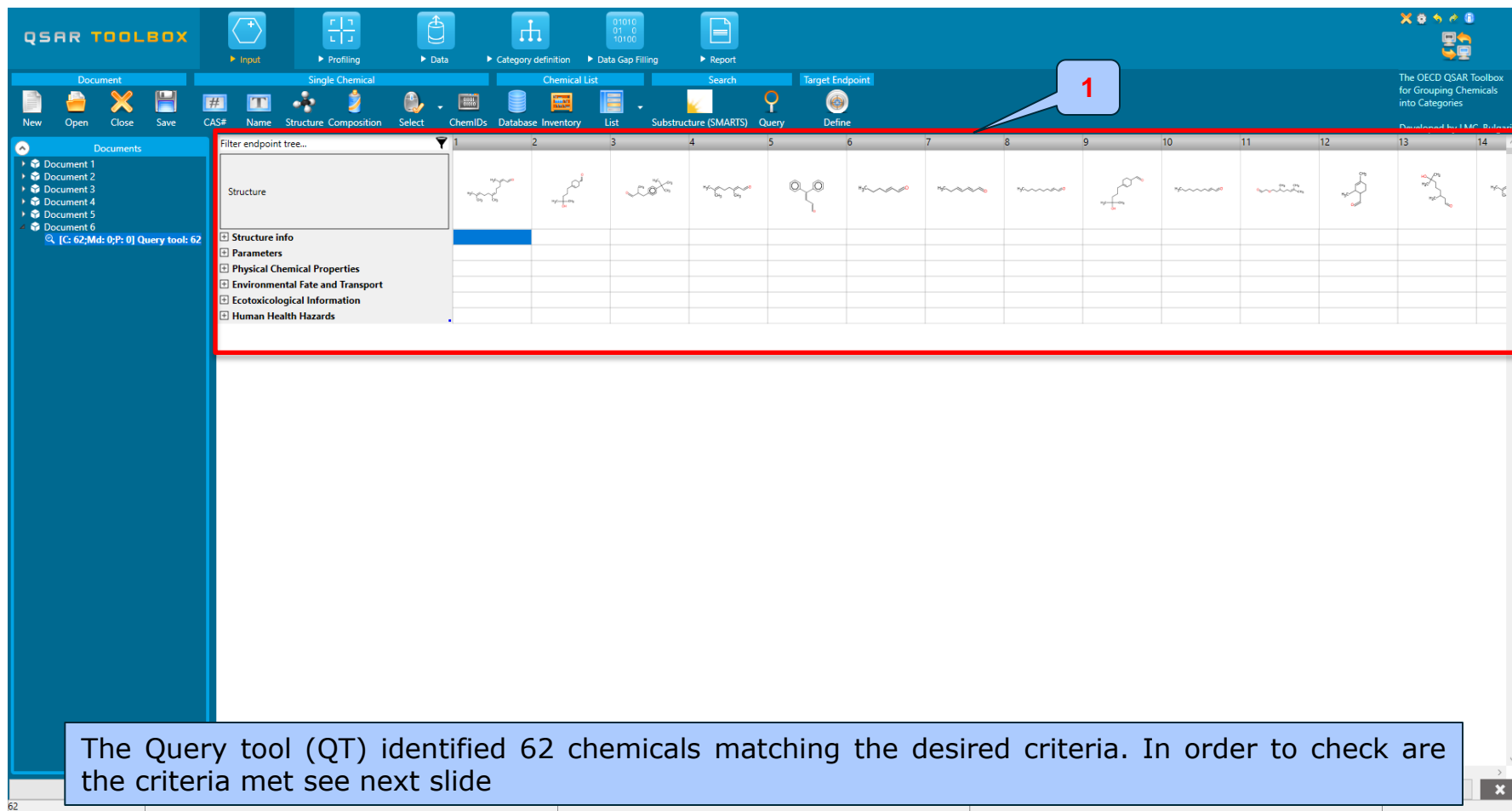
The screenshot shows the 'Search' window in the QSAR Toolbox. The 'Subfragments' list contains 'O=[CH]C'. The 'Options' section shows 'Search mode: All' and 'Exact match' and 'Exact connectivity' are unchecked. The 'View mode' is set to 'Facade' and 'Navigation mode' is 'Cascade'. The main window displays a chemical structure of an aldehyde fragment (O=CH-C). The right side of the window shows a logical query builder with a red box around the 'AND' button and a red box around the 'Execute' button. Numbered callouts 1 through 4 indicate the steps: 1. Right click over the queries in order to select them; 2. Click **AND** button; 3. The two queries are combined together by logical "AND"; 4. Click **Execute** button or double click *And*.

- Right click over the queries in order to select them;
- Click **AND** button;
- The two queries are combined together by logical "AND";
- Click **Execute** button or double click *And*;

# Subfragment search for identifying chemicals

*Chemicals with Skin sensitization positive data including Aldehyde fragment*

## Example 6



The screenshot displays the QSAR Toolbox software interface. The top menu bar includes options like Document, Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below the menu bar is a toolbar with icons for various functions. On the left, a sidebar shows a list of documents (Document 1 to Document 6) and a search query: [C: 62; Md: 0; P: 0] Query tool: 62. The main workspace is divided into two sections. The top section, labeled 'Filter endpoint tree...', contains a list of criteria: Structure, Structure info, Parameters, Physical Chemical Properties, Environmental Fate and Transport, Ecotoxicological Information, and Human Health Hazards. The bottom section is a table with 14 columns, each containing a chemical structure. A red box highlights the 'Filter endpoint tree...' section and the first row of the table. A callout bubble with the number '1' points to the 'Filter endpoint tree...' dropdown.

The Query tool (QT) identified 62 chemicals matching the desired criteria. In order to check are the criteria met see next slide

# Subfragment search for identifying chemicals

## *Chemicals with Skin sensitization positive data including Aldehyde fragment*

### Example 6

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

2. Databases are already selected, click **Gather** data;

3. Click **OK**;

4. Click **OK**;

5. Data appear on data matrix answering the following criteria: *EC3 to be positive (i.e. EC3<50%)*. If OFG profiler is applied to the identified chemicals it is seen that they all are categorized as aldehydes

250 points added across 62 chemicals.

Chemicals are sorted by "Descending order"

| Chemical | EC3    | M: 12 % | M: 17 % | M: 8.8 % | M: 6.5 % | M: 0.6 % | M: 3.78 % | M: 4 % | M: 2.5 % | M: 17.1 % | M: 2.5 % | M: 28.3 % | M: 3.25 % | M: 18 % |
|----------|--------|---------|---------|----------|----------|----------|-----------|--------|----------|-----------|----------|-----------|-----------|---------|
| EC3      | 62/201 |         |         |          |          |          |           |        |          |           |          |           |           |         |

# Outlook

- Background
- Objectives
- Overview of Query tool
- Query tool window
- The exercise
- Workflow process
- **Save QT searches**

## Saving the prediction result

- This functionality allows 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 the next screenshots

# Saving the QT requests

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

2. Click **Save** button;

3. Select **Yes** to confirm the operation;

4. Define name of the file;

5. Click **Save** the file;

6. A message informs that the document is saved successfully, click **Ok**.

# Open saved file

Once the file has been saved:

1. Click **Open** button within the Input module;
2. Find and select the file;
3. Click **Open**;
4. A message informs that the document is opened successfully, click **OK**.