

OECD QSAR Toolbox

Tutorial on SMARTS structures search

Outlook

- **Background**
- Aims
- Overview of SMARTS language
- SMARTS Editor window
- The exercise
- Workflow process

Background

- This is a step-by-step presentation designed to take the Toolbox user through the SMARTS Editor functionalities for substructure search;
- The SMARTS substructure search is useful when the user does not have a specific chemical in mind but rather wants to find a group of chemicals containing the same structural characteristics.

Outlook

- Background
- **Aims**
- Overview of SMARTS language
- SMARTS Editor window
- The exercise
- Workflow process

Aims

- To introduce to the Toolbox user the SMARTS Editor;
- To familiarize the user with the new query fragments for substructure search.

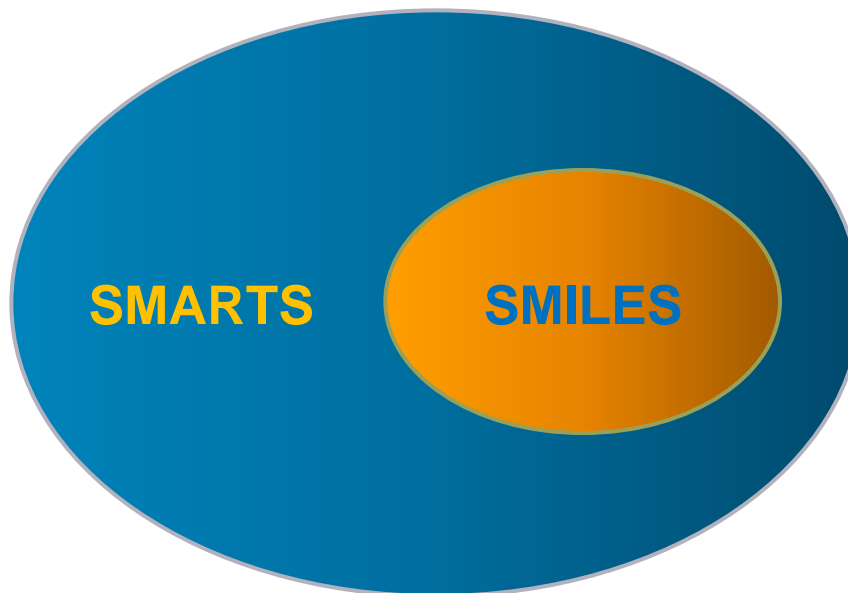
Outlook

- Background
- Aims
- **Overview of SMARTS language**
- SMARTS Editor window
- The exercise
- Workflow process

Overview of SMARTS language

SMARTS vs. SMILES

- SMARTS (**SM**iles **AR**bitrary **T**arget **S**pecification) - A language for describing molecular patterns used for fragment matching;
- SMARTS allows you to specify substructures using rules that are straightforward extensions of SMILES*

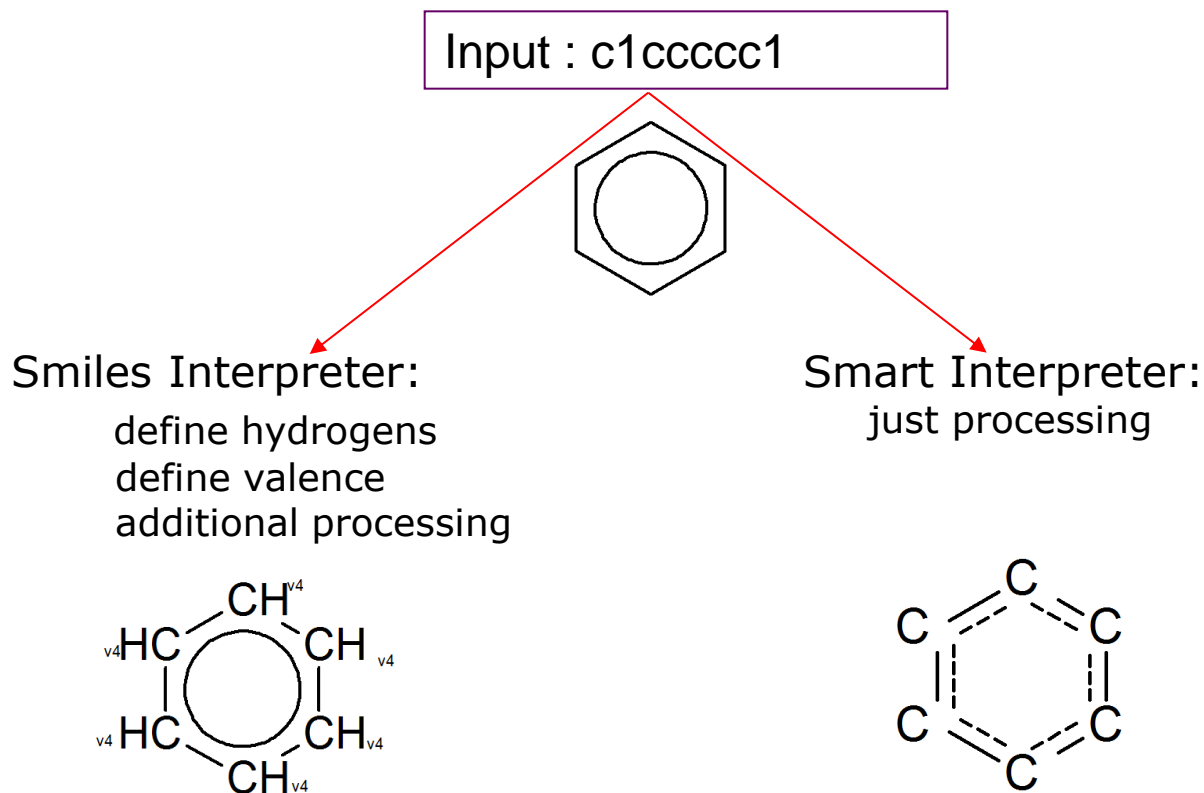


*<http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>

Overview of SMARTS language

SMARTS vs. SMILES

Any valid SMILES is valid SMARTS.



Outlook

- Background
- Aims
- Overview of SMARTS language
- **SMARTS Editor window**
- The exercise
- Workflow process

SMARTS Editor

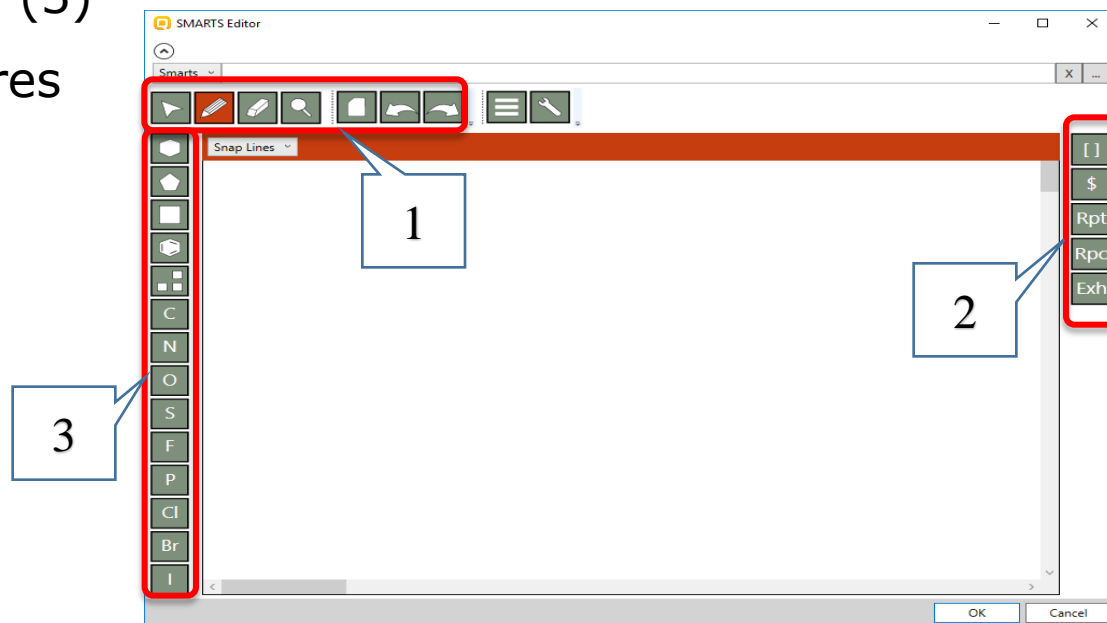
Overview

- The new editor is developed to be compatible with the SMARTS language;
- Editor fragments of the previous Toolbox versions are updated and extended to facilitate the users in writing SMARTS;
- The SMARTS Editor can be used for substructure search as well as for creating of custom profiles.

SMARTS Editor window

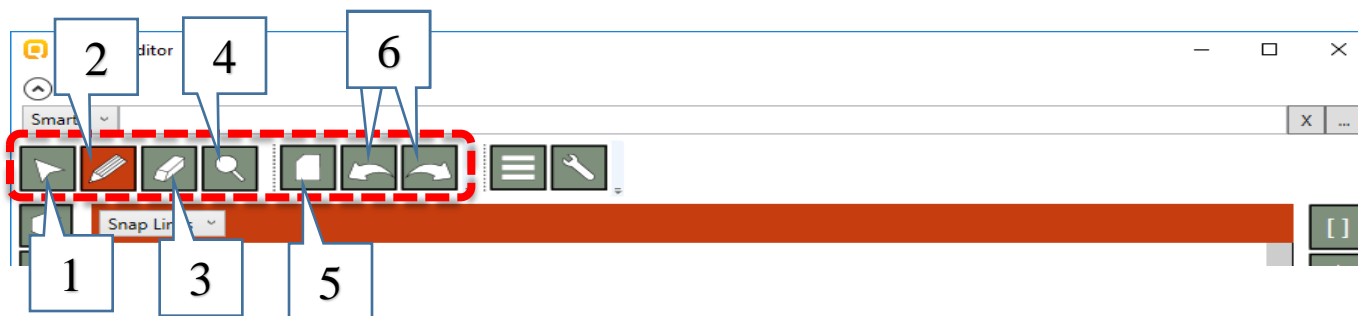
The SMARTS Editor window includes:

- Action toolbar (1)
- Fragments toolbar (2)
- Auxiliary toolbar (3)
- Additional features



SMARTS Editor window

Action toolbar



The **Action toolbar** includes:

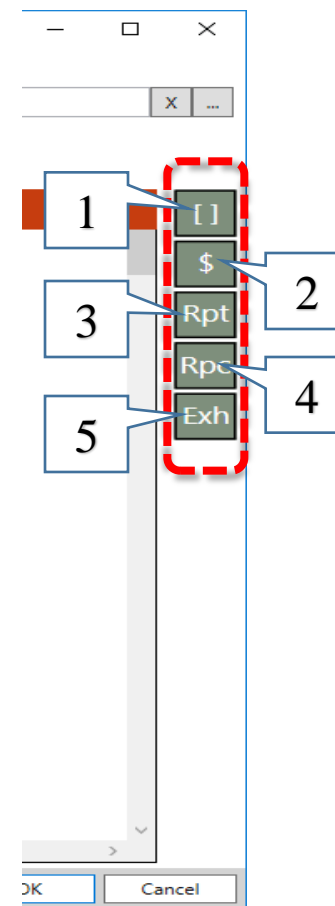
- **Selection tool** (1) – allows to define atom characteristics, to mark or rotate a structure, to start new activity, etc;
- **Pencil** button (2) – allows drawing of hydrocarbon structures (scaffolds);
- **Eraser** button (3) – removes selected bonds or other parts of a drawn structure;
- **Magnifying glass** button (4) – allows to zoom in (zoom out) with the mouse scroller;
- **Blank page** (5) – clears all in the drawing pane;
- **Arrows** buttons (6) – allow to undo any action or redo right after undo.

SMARTS Editor window

Fragments toolbar

The **Fragments toolbar** includes:

- "[]" button (enumeration) (1): allows enumeration of atoms which could be bonded to a given atom in the molecule;
- "\$" button (recursive SMARTS expression) (2): allows to describe the surrounding of an atom. For example it could be used to describe N -atom, which is not part of nitro group;
- "Rpt" button (repeat fragment) (3): allows to specify how many times an atom or a fragment could be repeated in the molecule;
- "Rpc" button (replace fragment) (4): gives an option to replace a given atom in the molecule with different chemical elements;
- "Exh" buttons (exhaust fragment) (5): allows only those atoms or fragments listed in the exhaust fragment to be bonded to a given atom.

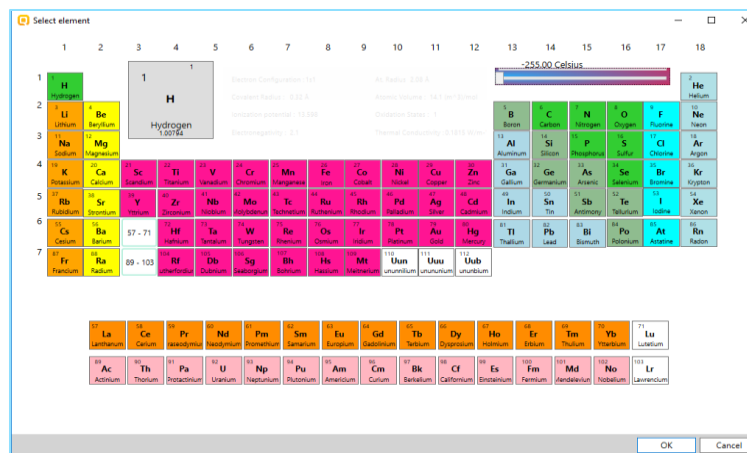
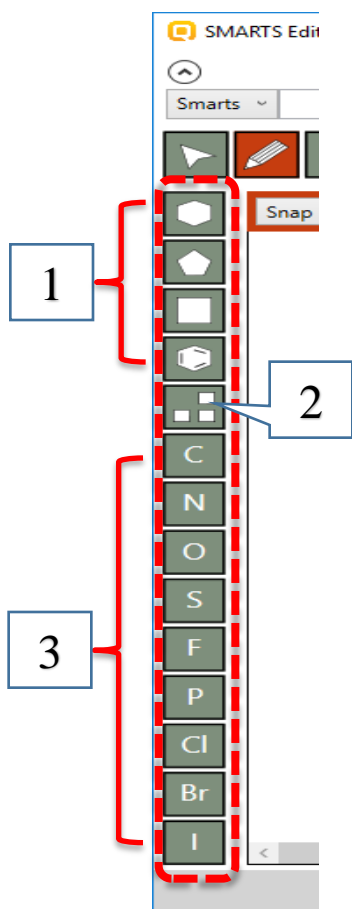


SMARTS Editor window

Auxiliary toolbar

The **Auxiliary toolbar** includes:

1. Some of the frequently used hydrocarbon structures, which can be used as scaffolds;
2. Periodic table – allows to select any chemical element. The new developed periodic table is more colorful and informative. It also allows the user to see the physical state of the elements at different temperature.



3. Some of the frequently used atoms.

SMARTS Editor window

Additional features

Some additional features are:

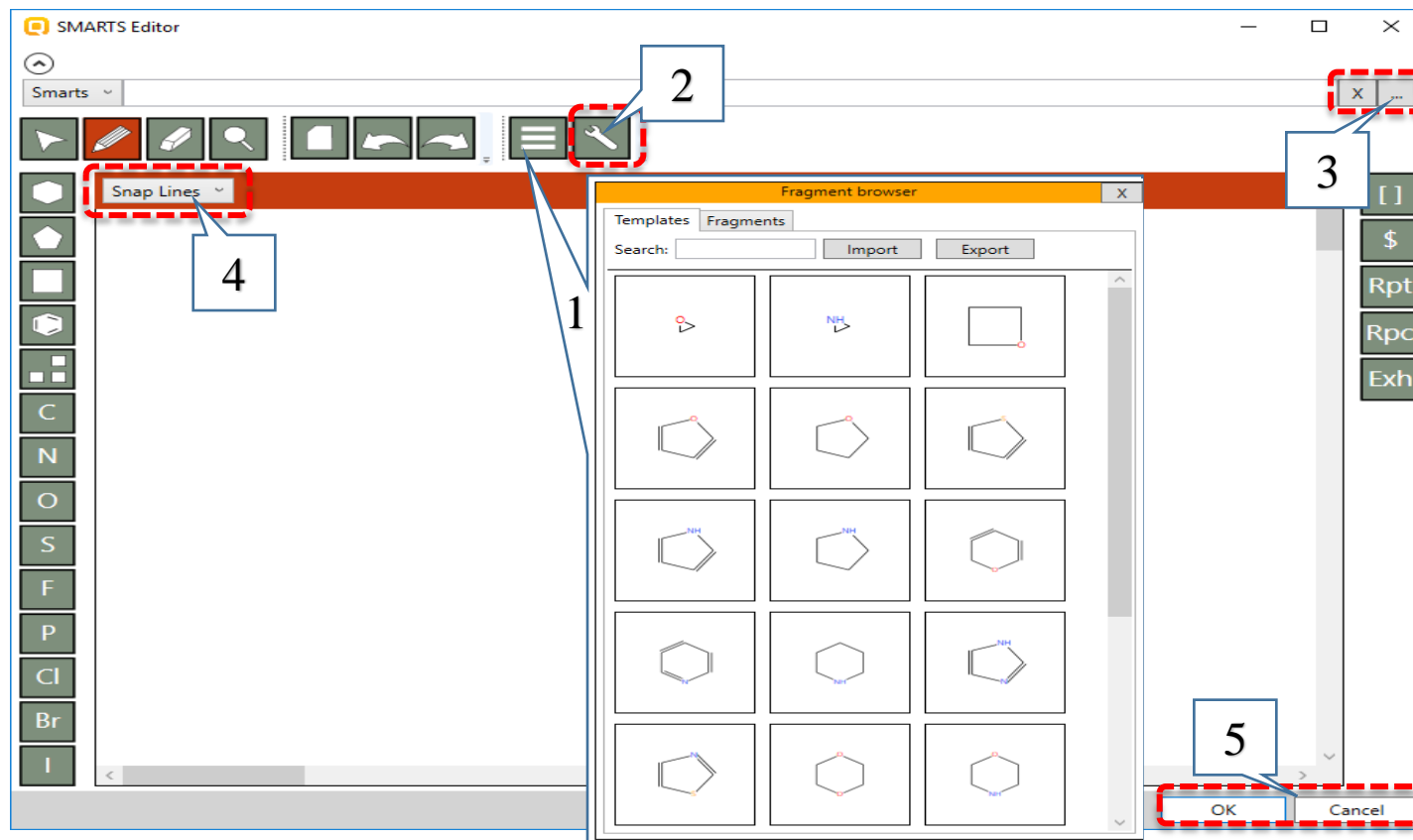
- **Buttons**
 1. Templates button - opens fragment browser, which contains list of templates of cyclic structures and list of fragments. The user can also add their structures and to make own fragment library;
 2. Wrench button – corrects all angles in the structure;
 3. Buttons allowing to delete or insert SMARTS;
 4. Red bar – includes some different additional secondary options;
 5. Buttons allowing to execute or cancel the search;
- **Object explorer**

See on the next two slides.

SMARTS Editor window

Additional features

- Buttons

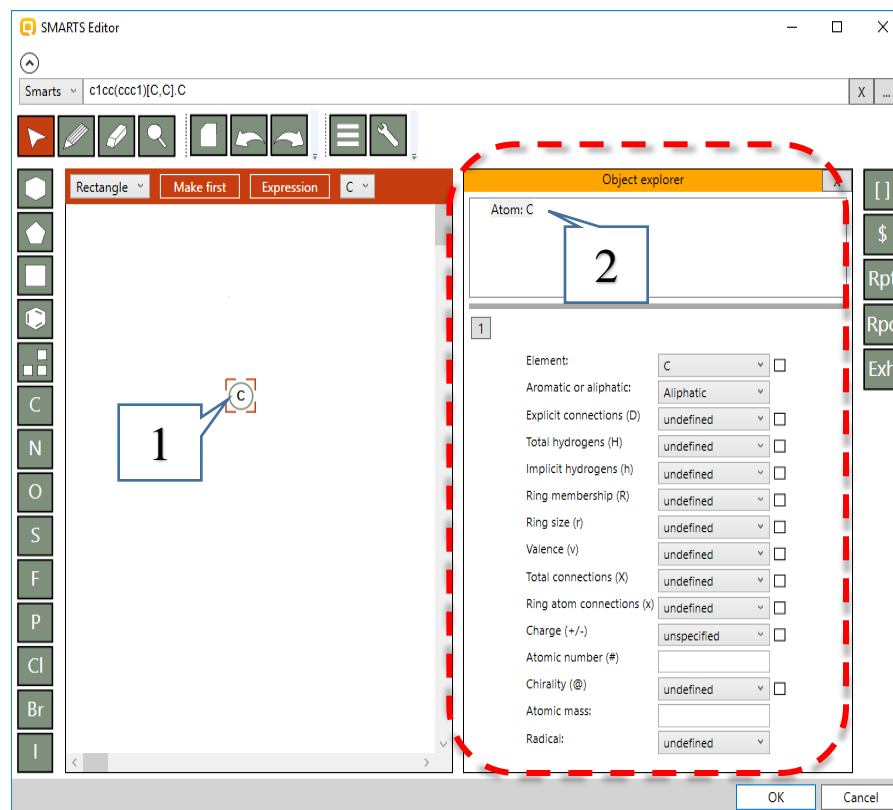


SMARTS Editor window

Additional features

- **Object explorer**

1. The *Object explorer* shows what information stays behind an atom or a fragment by click on the corresponding object in the drawing pane;
2. In the *Object explorer* the user can specify details for the selected object.



Outlook

- Background
- Aims
- Overview of SMARTS language
- SMARTS Editor window
- **The exercise**
- Workflow process

The exercise

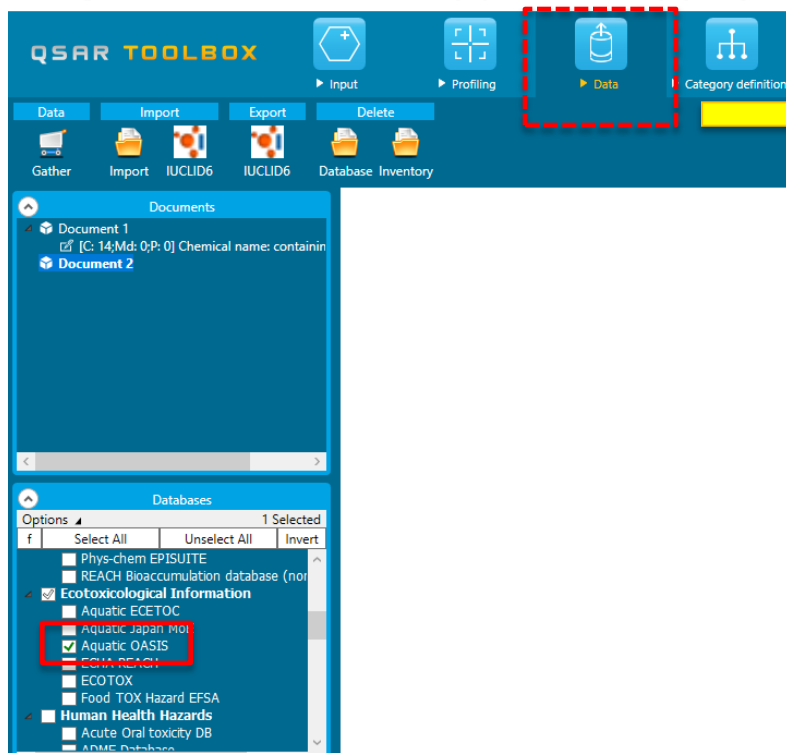
- In this exercise we will demonstrate how to construct different queries for structure searching;
- We will use each of the new fragments for the current purposes;
- *Aquatic OASIS* database will be used.

Outlook

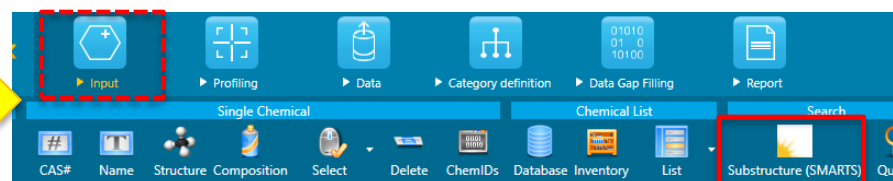
- Background
- Aims
- Overview of SMARTS language
- SMARTS Editor window
- The exercise
- **Workflow process**

Workflow process Overview

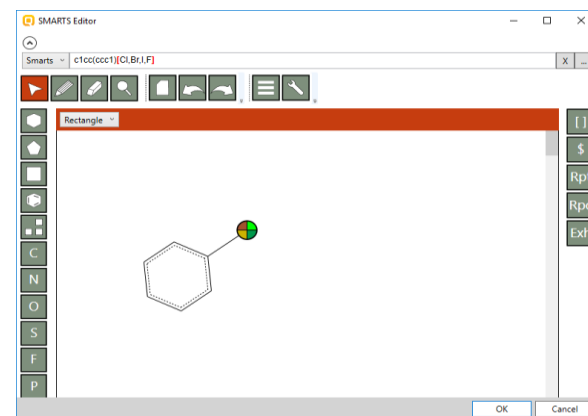
Step 1: Select database(s)/inventory(ies)



Step 2: Click on **Substructure (SMARTS)**



Step 3: Specify criteria for searching structures

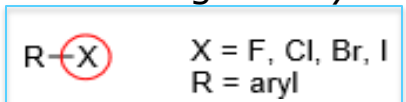


We will use **Aquatic OASIS** database for all examples. Searching criteria will be different only.

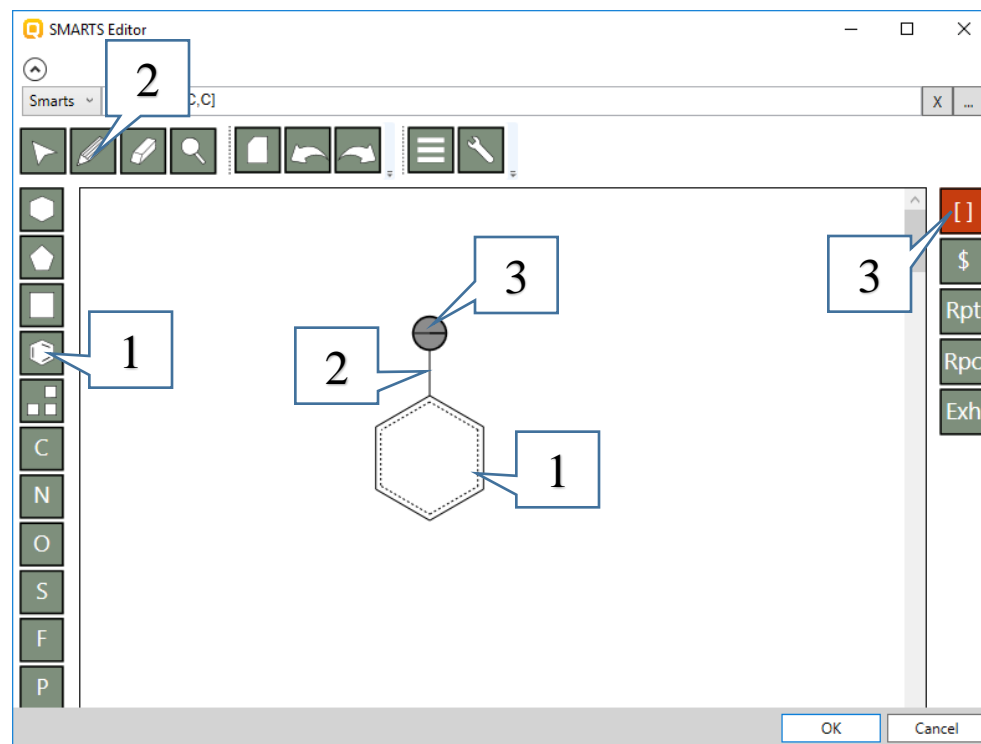
Workflow process

Structures search

Example 1: Searching for *Aryl halides*



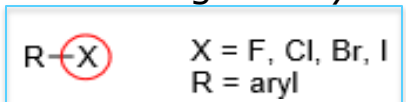
1. Click on the **benzene** template and then click in the drawing pane;
2. Click on the **pencil** button and then draw a bond;
3. Click on the “[]” (enumeration) button and then click over the new carbon atom.



Workflow process

Structures search

Example 1: Searching for *Aryl halides*



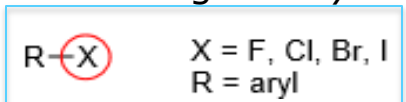
1. Click on the **selection tool** and then click on the new object;
2. *Object explorer* appears;
3. Right click over the “**OR**” expression node and select **Add child**. Repeat this step again.

The screenshot shows the QSAR Toolbox software interface. The main workspace displays a chemical structure (a benzene ring) with a selection tool (arrow icon) and a new object (hexagon) added. The 'Object explorer' panel on the right shows a tree structure with an 'OR' node and two 'Element: C' children. A context menu is open over the 'OR' node, showing options like 'Change operator', 'Convert to', and 'Add child'. Numbered callouts 1, 2, and 3 indicate the steps: 1. Clicking the selection tool (arrow icon), 2. Clicking the new object (hexagon), and 3. Right-clicking the 'OR' node and selecting 'Add child'.

Workflow process

Structures search

Example 1: Searching for Aryl halides



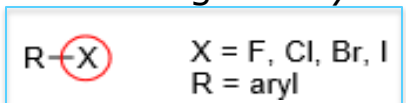
1. Click on the first element which corresponds to the first "child" query. The panel with the atom characteristics appears;
2. Select **Cl** from the *Element* drop-down menu;
3. The part of the SMARTS fragment corresponding to the first element have colored.
4. Click on the next elements selecting **Br**, **I**, **F** successively.

The top screenshot shows the QSAR Toolbox interface with the SMARTS fragment c1ccccc1[C]. A callout '1' points to the 'C' atom in the 'OR' list. The bottom screenshot shows the 'Element' drop-down menu open, with 'Cl' selected. A callout '2' points to the 'Cl' option in the menu. A callout '3' points to the highlighted 'C' atom in the structure, and a callout '4' points to the 'Element: Cl' entry in the 'OR' list.

Workflow process

Structures search

Example 1: Searching for *Aryl halides*



When all elements are selected (1) click on **OK** (2) to execute the search.

The screenshot shows the SMARTS Editor window. The Smarts input field contains the query c1cc(ccc1)[Cl,Br,I,F]. The main canvas displays a benzene ring with a multi-colored atom icon at the top position. The Object explorer on the right shows a tree structure with a red bracket labeled '1' encompassing the 'Element' category, which lists Cl, Br, I, and F. Below this, a list of search criteria is shown, with 'Element' set to 'F'. A red bracket labeled '2' points to the 'OK' button at the bottom right of the Object explorer panel.

Workflow process

Structures search

Example 1: Searching for *Aryl halides*

415 chemicals are found in the selected database. Click on Select all button (1) and then **OK** (2) to load all of them on the data matrix. You can see that all of the chemicals fulfill the criteria.

1

2

QSAR TOOLBOX

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

Filter endpoint tree...

Structure

Structure info

Parameters

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information

Human Health Hazards

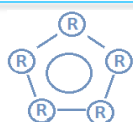
Depiction

CAS # 2008-58-4
2,6-Dichlorobenzamide
Benzamide, 2,6-dichloro-

Workflow process

Structures search

Example 2: Searching for *Five-membered heteroaromatic compounds*



R = C (1 to 4), S, N, O

1. Use the pentane template;
2. Click on one of the carbons;
3. Specify the atom to be **AnyAtom** and **Aromatic** from the drop-down menus;
4. Click on a bond;
5. Change the bond to **Aromatic**.

Repeat these steps for each of the atoms and bonds in the structure.

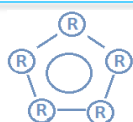
The screenshot illustrates the workflow steps for creating a SMARTS query in the SMARTS Editor:

- Step 1:** The 'Object explorer' panel shows the 'Element' dropdown menu set to 'AnyAtom'.
- Step 2:** The 'Object explorer' panel shows the 'Aromatic or aliphatic' dropdown menu set to 'Aromatic'.
- Step 3:** The 'Object explorer' panel shows the 'Explicit connections (D)' dropdown menu set to 'undefined'.
- Step 4:** The 'Object explorer' panel shows the 'Total hydrogens (H)' dropdown menu set to 'undefined'.
- Step 5:** The 'Change bond' panel shows the 'Bond' dropdown menu set to 'Aromatic'.

Workflow process

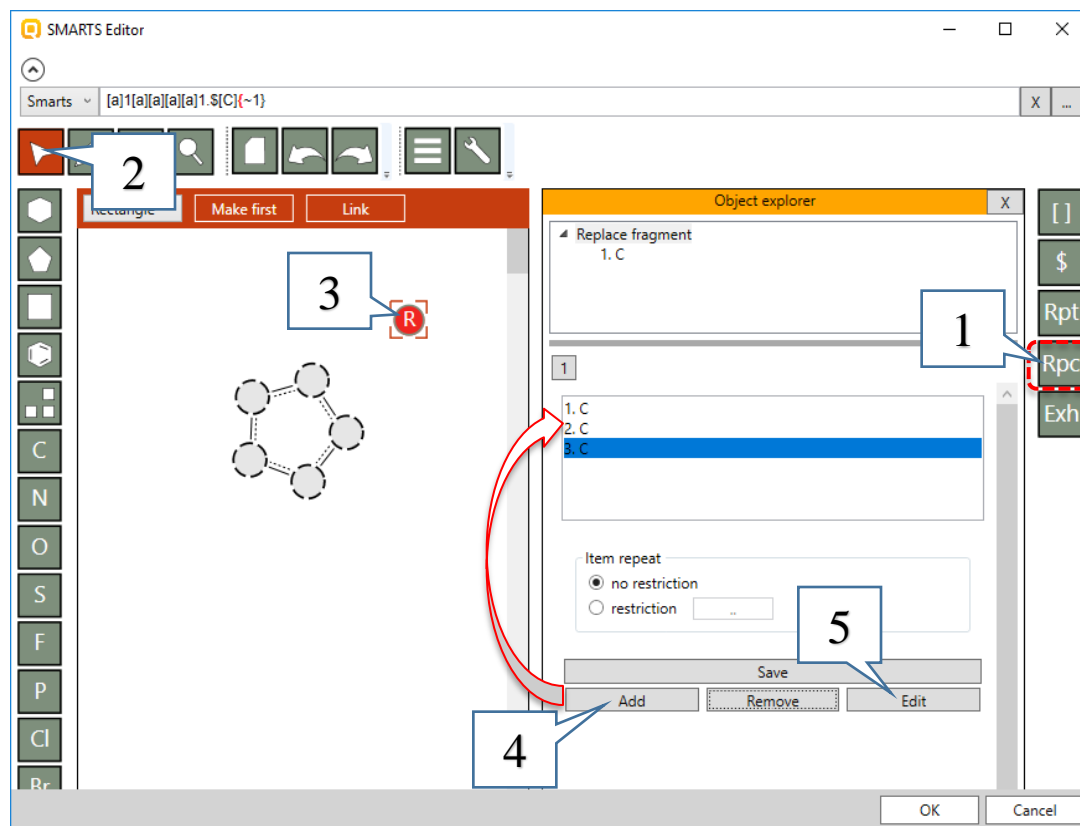
Structures search

Example 2: Searching for *Five-membered heteroaromatic compounds*



R = C (1 to 4), S, N, O

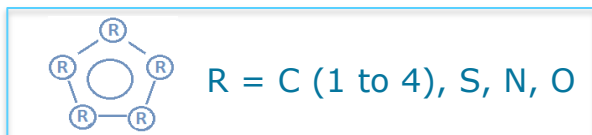
1. Click on the **Rpc** (replace) fragment and paste it in the drawing pane;
2. Click on the **selection tool**;
3. Click on the **red R** (Rpc) fragment;
4. Click on **Add** button three times to include three elements. The elements are aliphatic carbons by default.
5. Click on **Edit** button.



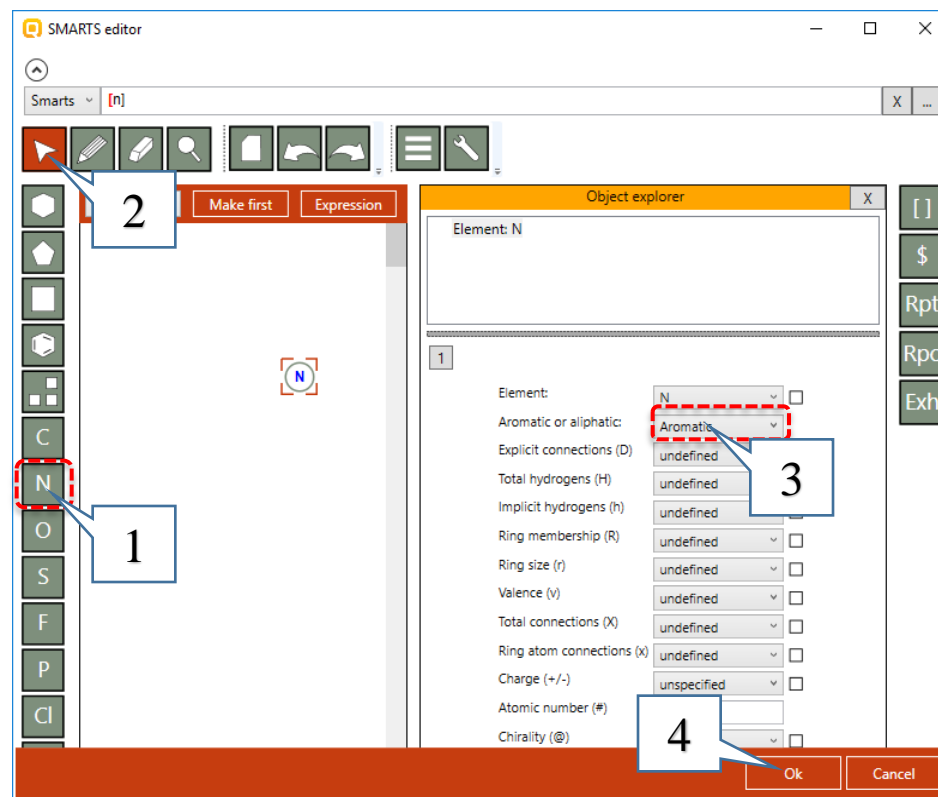
Workflow process

Structures search

Example 2: Searching for *Five-membered heteroaromatic compounds*



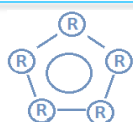
1. Select **N** symbol and put it over the C atom;
2. Click on **selection tool**;
3. Define the nitrogen atom to be aromatic;
4. Click **OK** to confirm



Workflow process

Structures search

Example 2: Searching for *Five-membered heteroaromatic compounds*



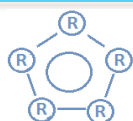
R = C (1 to 4), S, N, O

Change the rest aliphatic carbons to aromatic sulphur and oxygen in the same way;

Workflow process

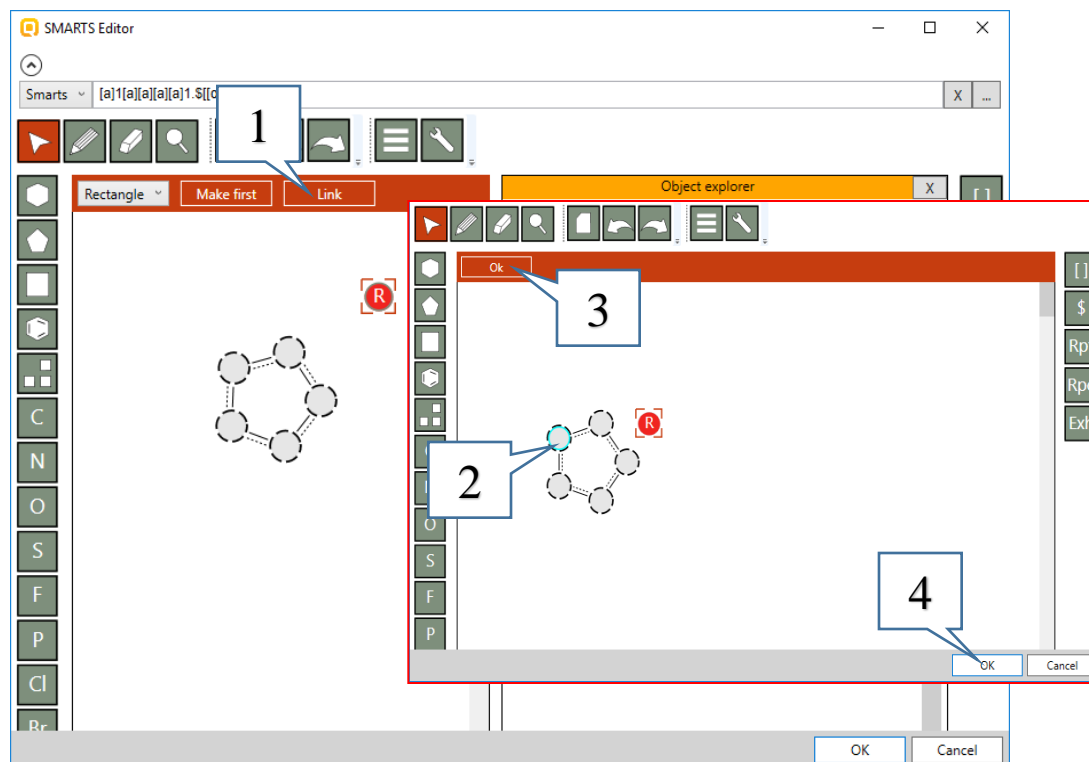
Structures search

Example 2: Searching for *Five-membered heteroaromatic compounds*



R = C (1 to 4), S, N, O

1. Click on **Link** button to show the atoms, which will be replaced;
 2. Click over one of the atoms;
 3. Press **OK** to confirm the criteria;
- In this way we say that we want the found structures to have at least one heteroaromatic atom;
4. Execute the search by the main **OK** button



Workflow process

Structures search

Example 2: Searching for *Five-membered heteroaromatic compounds*

60 chemicals are found in the selected database. Click on **Select All** button (1) then **OK** (2) to load all of them on the data matrix. You can see that all of the chemicals fulfill the criteria.

1 Select chemicals

CAS	SMILES	CS Relation	Substance	Composition	Name	Sources
19335-11-6	<chem>Nc1ccc2[nH]ncc2c1</chem>	High	Mono constituent		1H-indazol-5-amine; 5-Aminoindazol	DSSTOX ECHA PR FINPFS
2034-22-2	<chem>BrC1nc(Br)c(Br)[nH]1</chem>	High	Mono constituent		1H-Imidazole, 2,4,5-trithromo-, 2,4,5-	DSSTOX ECHA PR
625-86-5	<chem>Cc1ccc(C)o1</chem>	High	Mono constituent		2,5-Dimethylfuran, Furan, 2,5-dimeth	NICNAS DSSTOX

2 OK

QSAR TOOLBOX

Document: Single Chemical, Chemical List, Search, Target Endpoint

Filter endpoint tree...

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

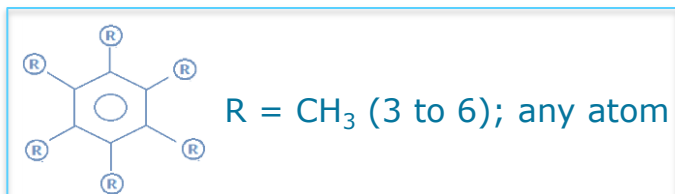
Depiction

CAS # 19335-11-6
1H-indazol-5-amine
5-Aminoindazole
indazol-5-amine

Workflow process

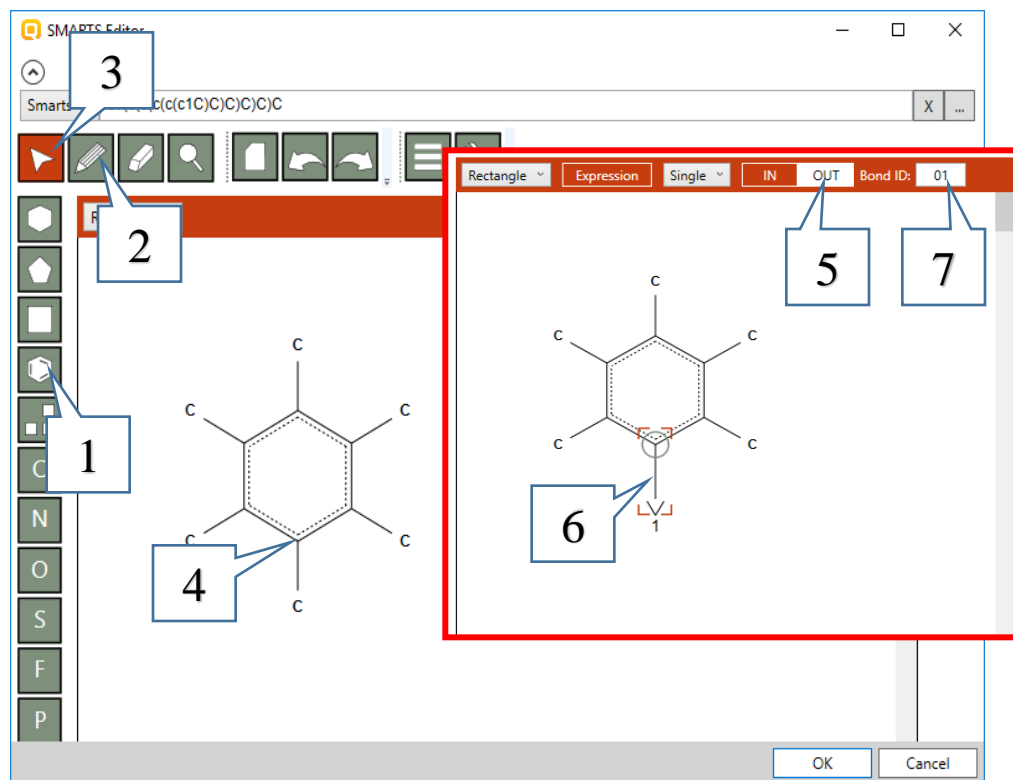
Structures search

Example 3: Searching for *aryl compounds with three or more methyl groups*



1. Select **benzene** template and put it in the drawing pane;
2. Select the **pencil** and draw bonds to each of the benzene atoms;
3. Click on **selection tool**;
4. Click over a bond;
5. Select “**Out**” of the new appeared buttons in the red bar;
6. Click over the same bond again;
7. Write “1” in the *Bond ID* field.

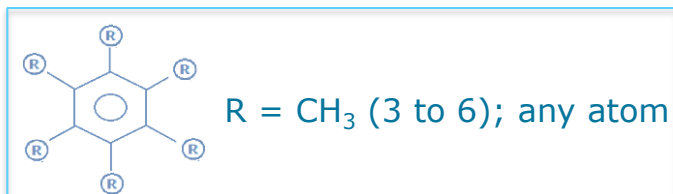
Repeat the steps above (from 4 to 7) for each of the bonds.



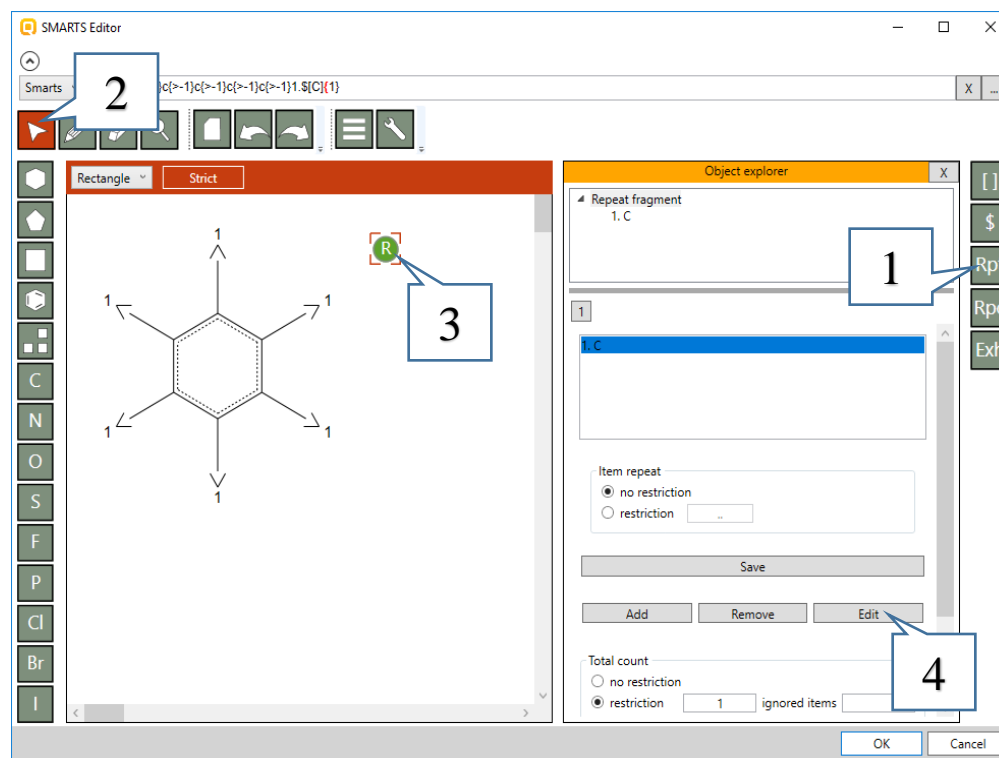
Workflow process

Structures search

Example 3: Searching for *aryl compounds with three or more methyl groups*



1. Select **Rpt** (repeat) fragment and put it in the drawing pane;
2. Click on the **selection tool**;
3. Click over the **green R** (Rpt);
4. Click on the **Edit** button.

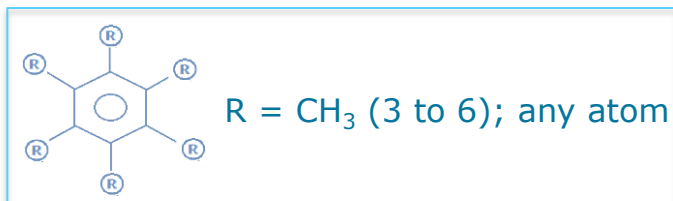


The screenshot shows the SMARTS Editor window. The main drawing pane displays a benzene ring with six '1' labels at the attachment points. A green 'R' (Rpt) fragment is being added to the ring, indicated by a red box and a callout '3'. The 'Object explorer' on the right shows the 'Repeat fragment' list with '1. C' selected. The 'Item repeat' section has 'no restriction' selected. The 'Total count' section has 'restriction' selected with a value of '1'. The 'Edit' button is highlighted with a callout '4'. The 'Smarts' input field at the top contains the query: c{>-1}c{>-1}c{>-1}c{>-1}c{>-1}1.[C]([1]). The 'Selection tool' icon is highlighted with a callout '2'.

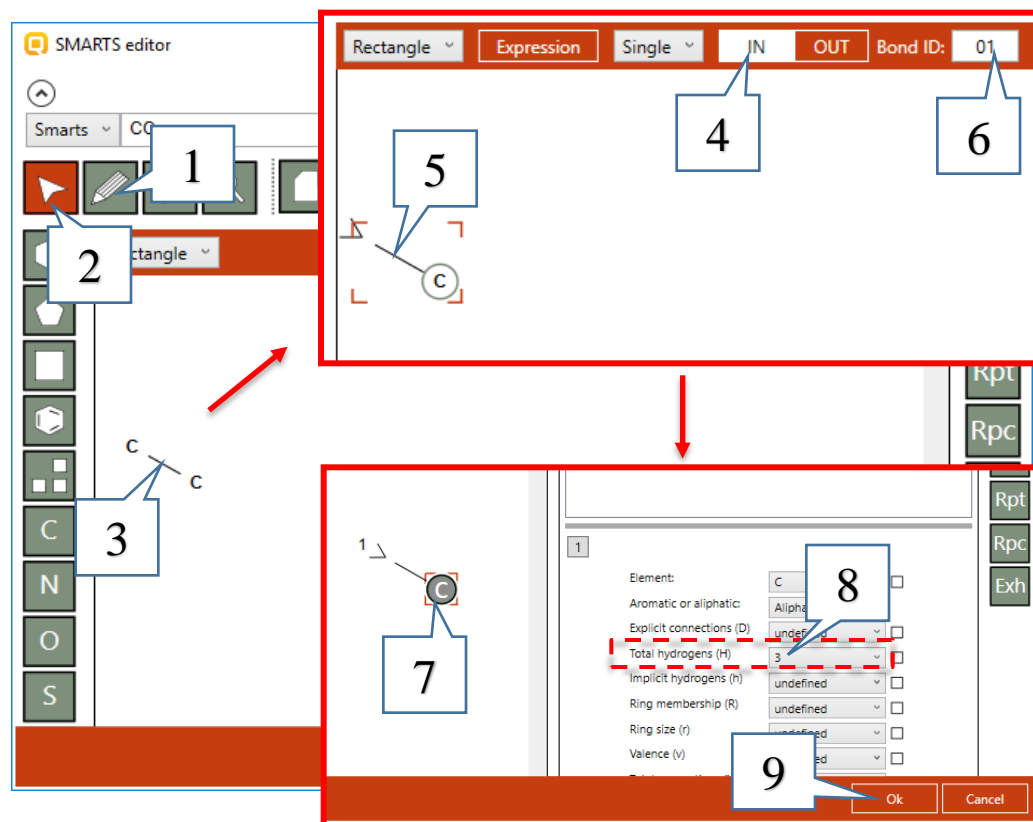
Workflow process

Structures search

Example 3: Searching for *aryl compounds with three or more methyl groups*



1. Select the **pencil** and draw a bond;
2. Click on the **selection tool**;
3. Click over the bond;
4. Select "In" of the new appeared buttons in the red bar;
5. Click over the same bond again;
6. Write "1" in the *Bond ID* field.
7. Click over the carbon atom;
8. Define total hydrogens to be three;
9. Confirm with **OK**.



The screenshot illustrates the workflow for defining a methyl group in the SMARTS editor. The interface includes a toolbar with a pencil icon (1) and a selection tool icon (2). A red bar at the top contains buttons for bond types: Rectangle, Expression, Single, IN (4), and OUT. The Bond ID field (6) is set to 01. The main workspace shows a carbon atom (C) with a bond (5) being drawn. A zoomed-in view of the workspace shows the carbon atom (7) and the bond (5). A zoomed-in view of the properties panel shows the definition of the bond (8) with the following settings:

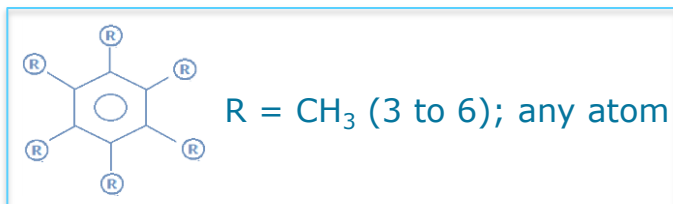
Property	Value
Element	C
Aromatic or aliphatic	Aliphatic
Explicit connections (D)	undefined
Total hydrogens (H)	3
Implicit hydrogens (h)	undefined
Ring membership (R)	undefined
Ring size (r)	undefined
Valence (v)	undefined

The OK button (9) is at the bottom right of the properties panel.

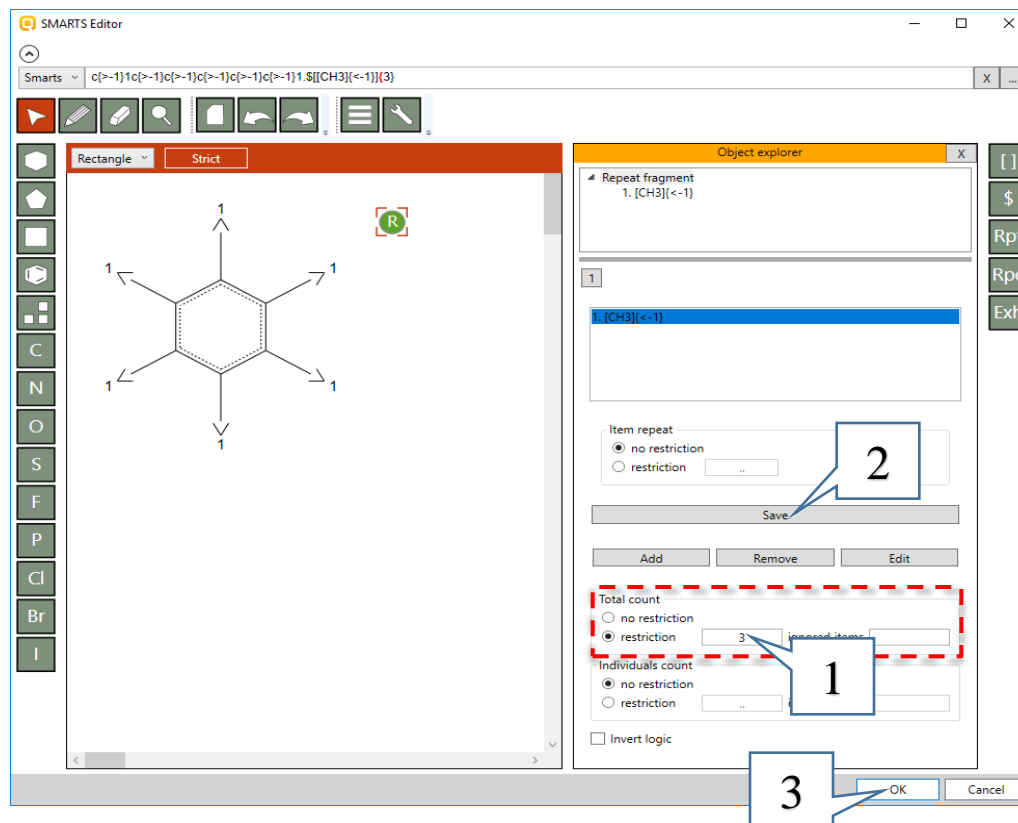
Workflow process

Structures search

Example 3: Searching for *aryl compounds with three or more methyl groups*



1. Type **3** in the restriction field for the **Total count**. Structures with at least three methyl groups will be searched;
2. Click on **Save**;
3. Execute the search by click on **OK**



The screenshot shows the SMARTS Editor window with the following components:

- Smarts field:** Contains the query c[>-1]1c[>-1]c[>-1]c[>-1]c[>-1]c[>-1]1.\$[CH3][<-1]](3).
- Object explorer:** Shows a 'Repeat fragment' list with '1. [CH3][<-1]'.
- Item repeat:** Radio buttons for 'no restriction' (selected) and 'restriction'.
- Total count:** Radio buttons for 'no restriction' and 'restriction' (selected). The 'restriction' field is set to '3'.
- Individuals count:** Radio buttons for 'no restriction' (selected) and 'restriction'.
- Invert logic:** A checkbox that is currently unchecked.
- Buttons:** 'Save', 'Add', 'Remove', 'Edit', 'OK', and 'Cancel'.

Numbered callouts indicate the workflow steps:

1. Points to the 'restriction' field in the 'Total count' section, where the value '3' is entered.
2. Points to the 'Save' button.
3. Points to the 'OK' button.

Workflow process

Structures search

Example 3: Searching for *aryl compounds with three or more methyl groups*

15 chemicals (1) are found in the selected database. Click on **OK** (2) to load all of them on the data matrix. You can see that all of the chemicals fulfill the criteria.

The screenshot illustrates the workflow process for searching aryl compounds with three or more methyl groups using the QSAR Toolbox. The process is divided into two main steps:

- Step 1: Select chemicals** (indicated by a red dashed box and a callout '1'). This step involves searching the database for chemicals that match the criteria. The search results are displayed in a list, showing the CAS number, SMILES string, CS Relation, Substance, Composition, Name, and Sources for each chemical. The list shows 15 chemicals, with the first three highlighted. The chemical names are: "2,3,5-trimethylhydroquinone", "1,4-B", "2,3,6-trimethylphenol", "2,3,6-Trimethyl", "1,3,5-Trimethylbenzene;Benzene, 1,3".
- Step 2: Load chemicals** (indicated by a callout '2'). This step involves clicking the **OK** button to load all 15 chemicals into the data matrix.

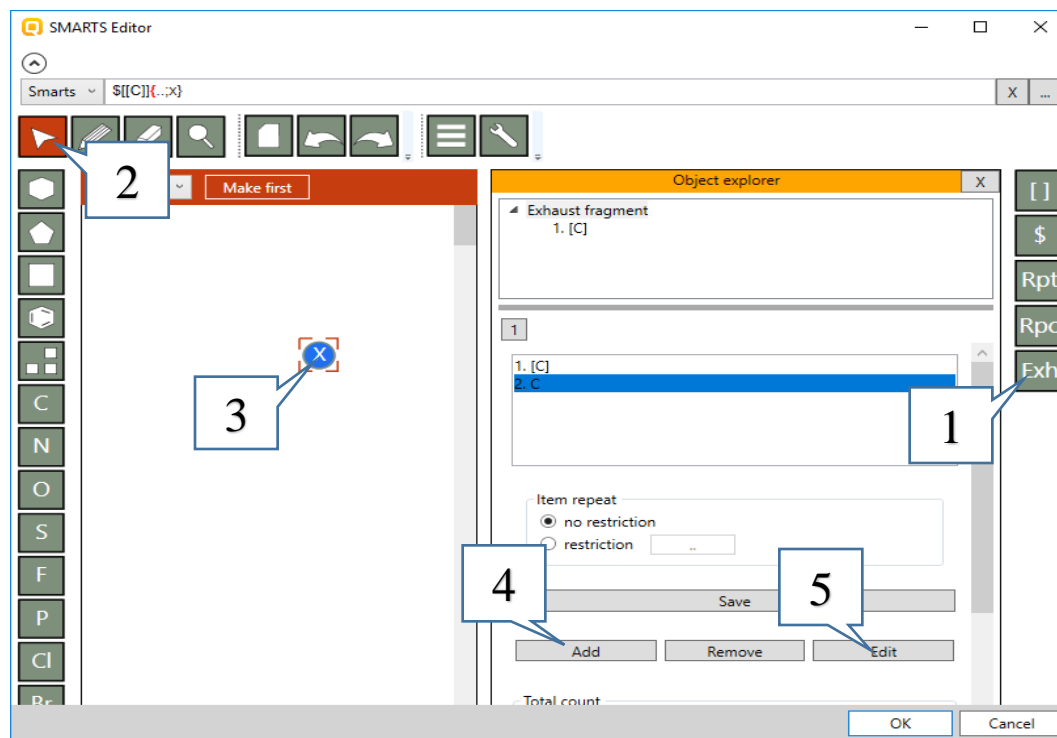
The QSAR Toolbox interface shows the **Documents** panel with a list of documents. The **Structure** panel shows the chemical structure of the selected chemical, with the methyl groups highlighted in red circles. The **Depiction** panel shows the chemical structure of the selected chemical, with the methyl groups highlighted in red circles.

Workflow process

Structures search

Example 4: Searching for hydrocarbon structures with 3 to 6 atoms and at least one hydroxyl group

1. Select **Exh** (exhaust) fragment and put it in the drawing pane;
2. Click on **selection tool**;
3. Click over the **X** (Exh);
4. Click on **Add** button;
5. Click on **Edit** button.

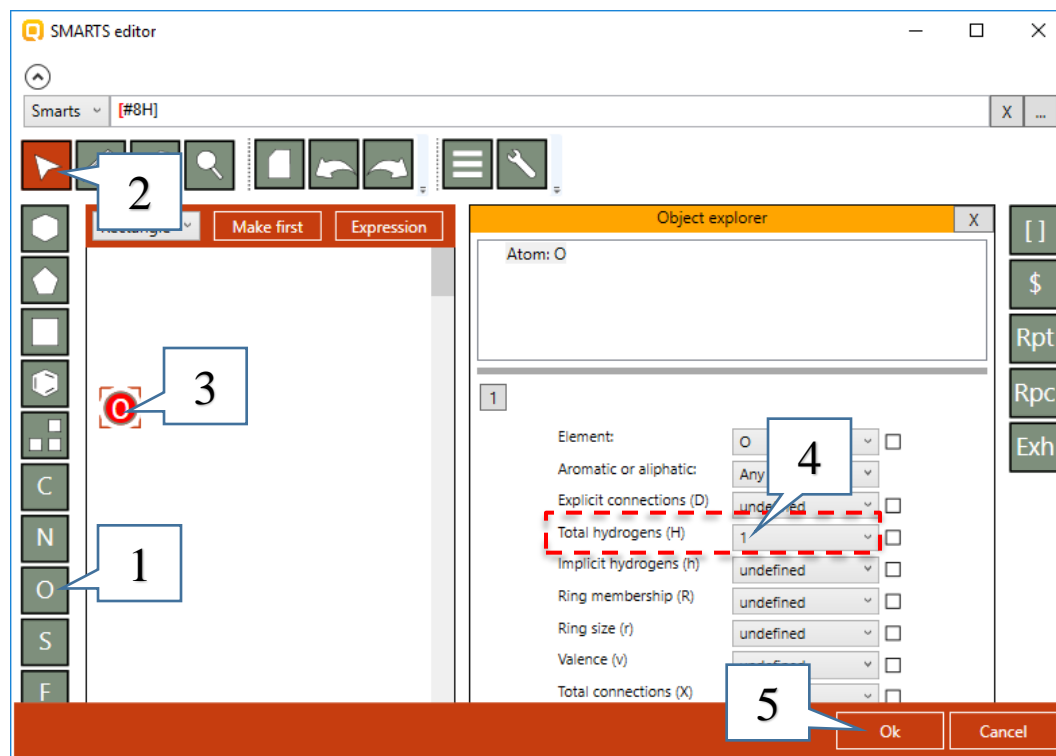


Workflow process

Structures search

Example 4: Searching for hydrocarbon structures with 3 to 6 atoms and at least one hydroxyl group

1. Select oxygen symbol (**O**) and put it over the carbon atom;
2. Click on **selection tool**;
3. Click over the **oxygen**;
4. Define number of *Total hydrogens (H)* to be one;
5. Confirm with **OK**.

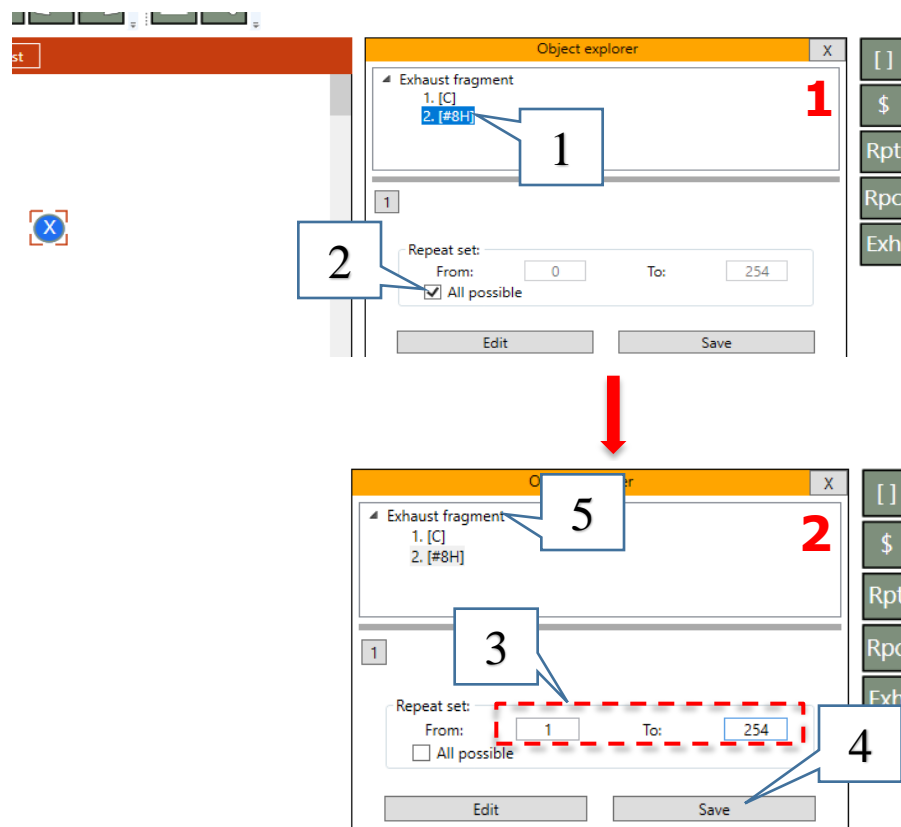


Workflow process

Structures search

Example 4: Searching for hydrocarbon structures with 3 to 6 atoms and at least one hydroxyl group

1. Click on the second item (hydroxyl group) in the exhaust fragment;
2. Uncheck *All possible*;
3. Define range from 1 to 254 (or to present at least one time in the structure);
4. Click on **Save**;
5. Go back to the *Exhaust fragment*.

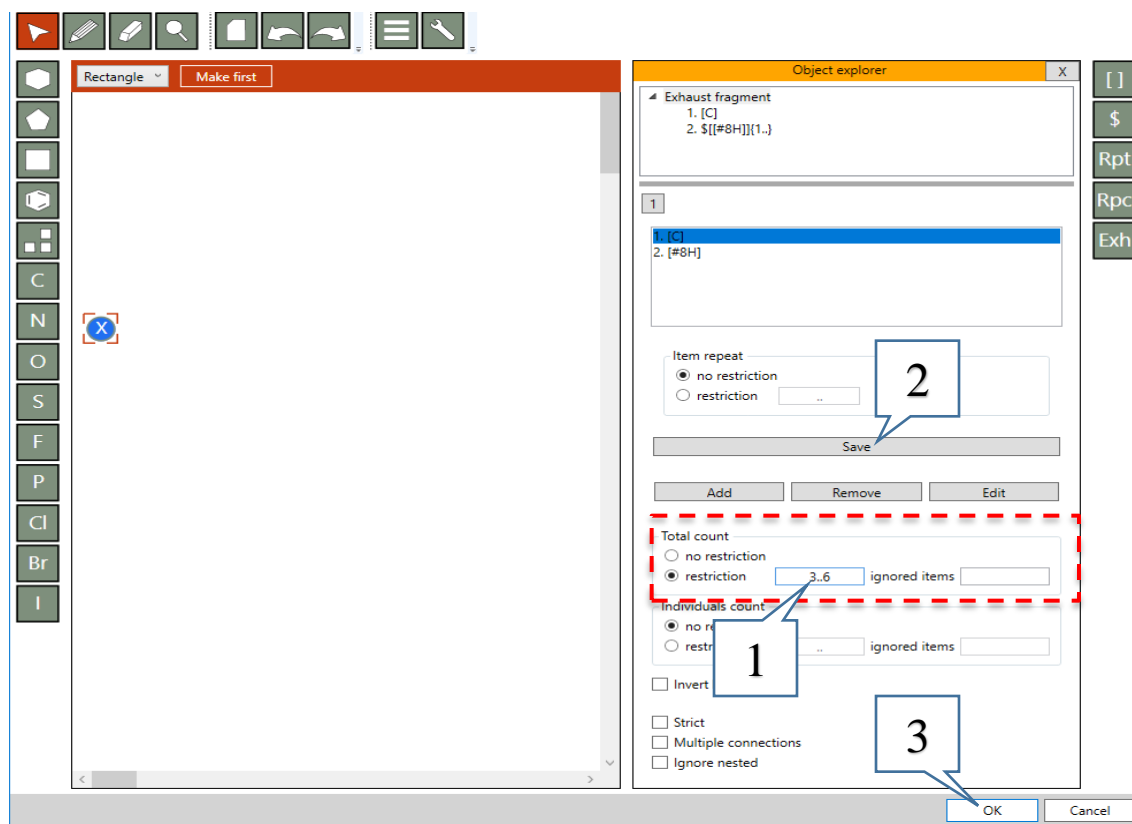


Workflow process

Structures search

Example 4: Searching for hydrocarbon structures with 3 to 6 atoms and at least one hydroxyl group

1. Set total restriction of the atoms to be from **3** to **6**;
2. Click on **Save**;
3. Execute the search by **OK**



Workflow process

Structures search

Example 4: Searching for hydrocarbon structures with 3 to 6 atoms and at least one hydroxyl group

52 chemicals (1) are found in the selected database. Click on **Select All** (1) button and then **OK** (2) to load all of them on the data matrix. You can see that all of the chemicals fulfill the criteria.

1 Select chemicals

Select All Unselect All Invert Selection Selected 52 of 52

1 CAS 5390-04-5
SMILES OCCCCC#C
CS Relation High
Substance Mono constituent
Composition
Name 4-pentyn-1-ol; Pent-4-yn-1-ol
Sources DISTOX ECHA PR EINECS

2 CAS 71-36-3
SMILES OCCCCO
CS Relation High
Substance Mono constituent
Composition
Name *1-butanol; butanol, 1-; butanol; butan-1-ol
Sources NICHAS CAS#s DISTOX

3 CAS 10229-10-4
SMILES C#CCCCO
CS Relation High
Substance Mono constituent
Composition
Name 3-Pentyn-1-ol; pent-3-yn-1-ol
Sources DISTOX ECHA PR EINECS

2 OK Cancel

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: 14;Md: 0;P: 0] Chemical name: conta
- Document 2 [C: 415;Md: 0;P: 0] Search by SMARTS
- Document 3 [C: 60;Md: 0;P: 0] Search by SMARTS
- Document 4 [C: 15;Md: 0;P: 0] Search by SMARTS
- Document 5 [C: 52;Md: 0;P: 0] Search by SMARTS

Filter endpoint tree...

Structure

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

Depiction

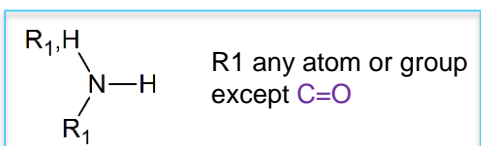
CAS # 10229-10-4
3-Pentyn-1-ol
pent-3-yn-1-ol

CC#CCCCO

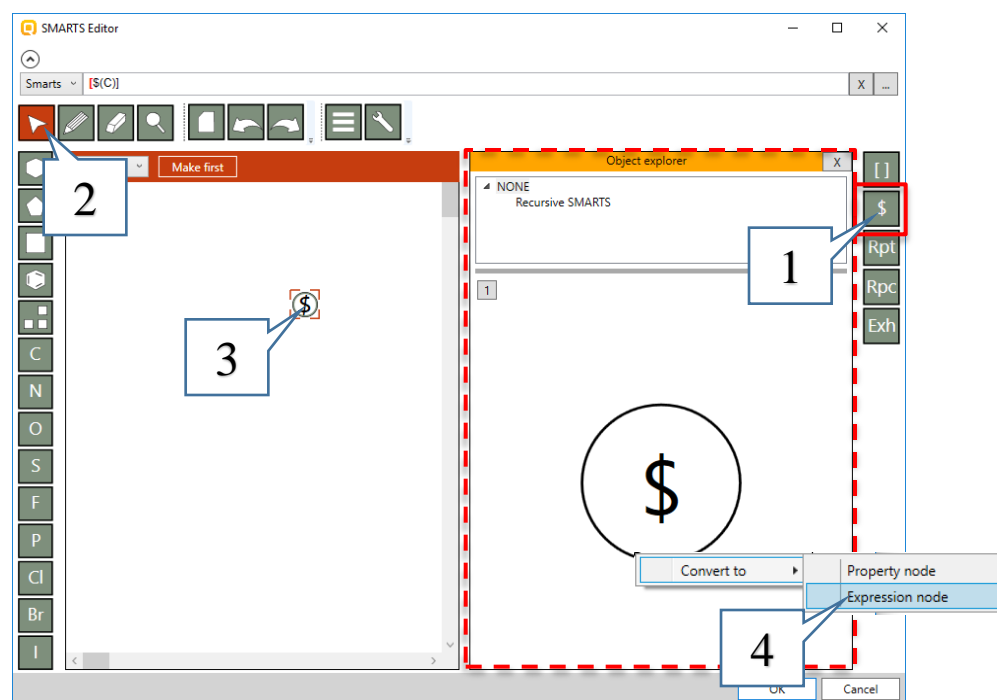
Workflow process

Structures search

Example 5: Searching for chemicals having primary or secondary amine group, but not amide



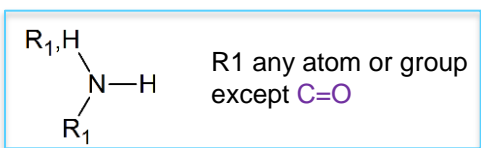
1. Select \$ (recursive SMARTS) fragment and put it in the drawing pane;
2. Click on the **selection tool**;
3. Click over the \$ fragment;
4. Right click over the \$ in the *Object explorer* panel and select **Convert to > Expression node**



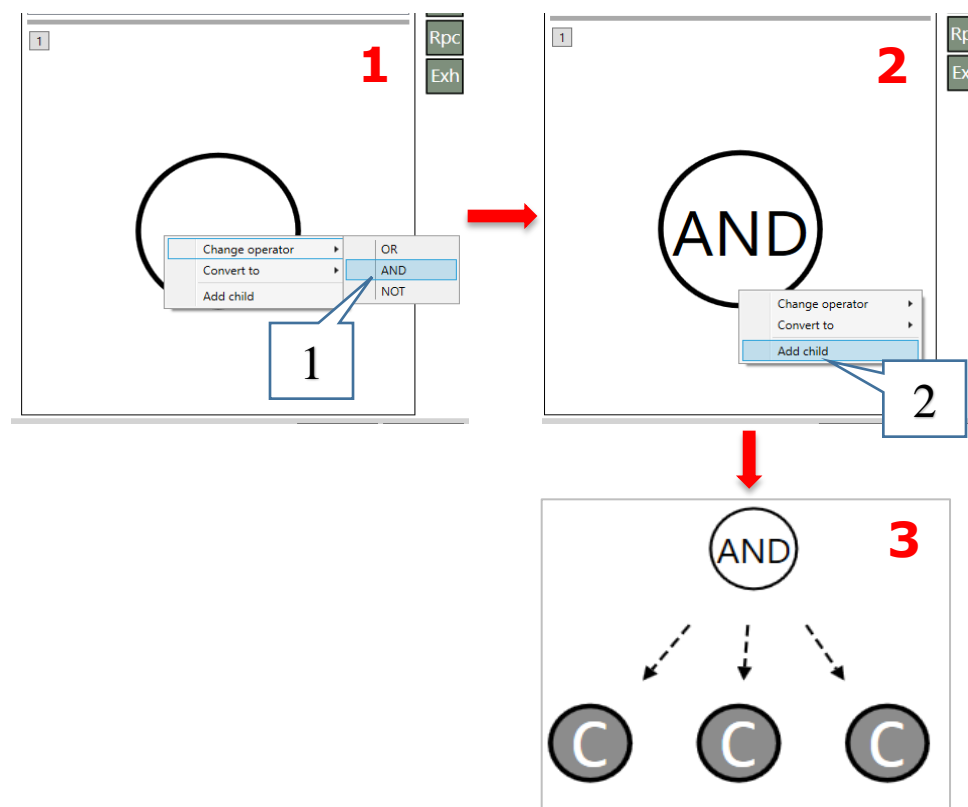
Workflow process

Structures search

Example 5: Searching for chemicals having primary or secondary amine group, but not amide



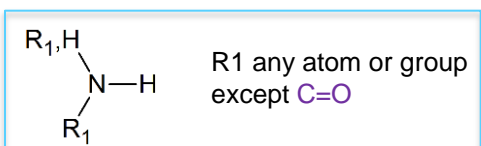
1. Right click over the ball and select **Change operator > AND**
2. Right click over the AND operator and select **Add child**.
3. Repeat the 2nd step three times to obtain three "child" queries.



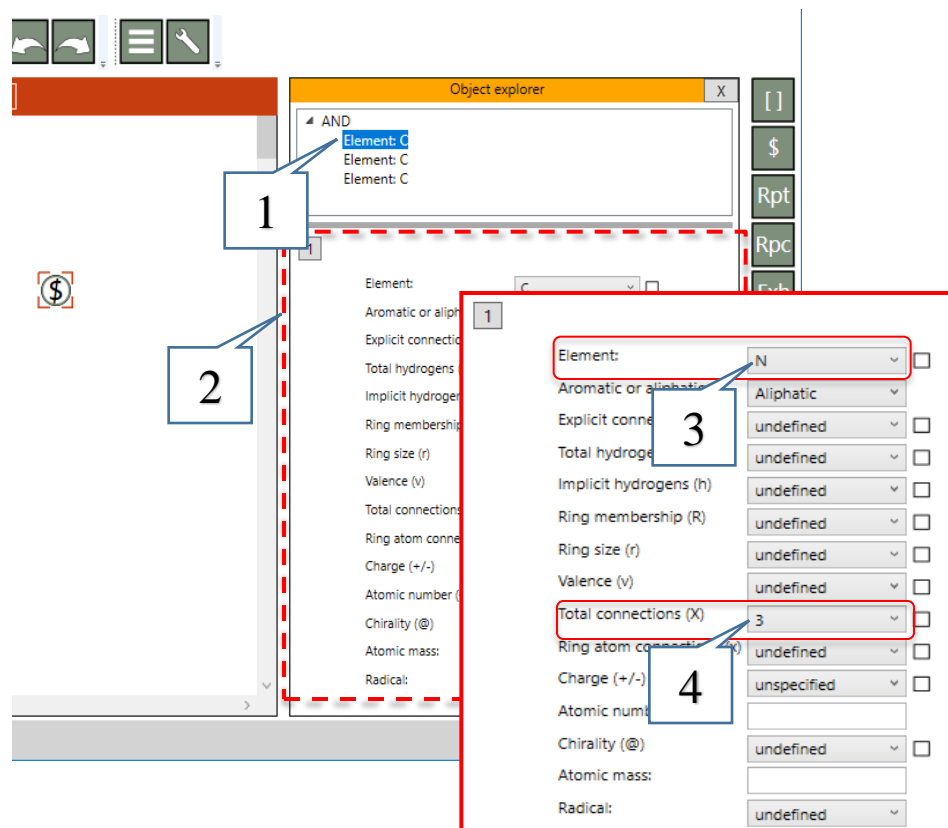
Workflow process

Structures search

Example 5: Searching for chemicals having primary or secondary amine group, but not amide



1. Click on the first item of “AND”;
2. Atom characteristics window appears;
3. Select *Element* to be nitrogen from the drop-down menus;
4. Select *Total connections* to be three from the drop-down menu.

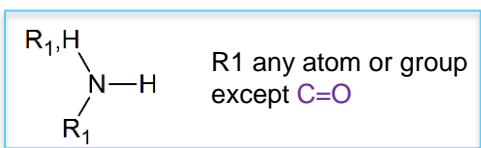


The screenshot shows the QSAR Toolbox interface. The 'Object explorer' window is open, showing a search condition: 'AND Element: C Element: C Element: C'. The 'Atom characteristics' window is open, showing a list of atom characteristics. The 'Element' dropdown is set to 'N' (Nitrogen) and the 'Total connections (X)' dropdown is set to '3'. The 'Atom characteristics' window is highlighted with a red dashed box.

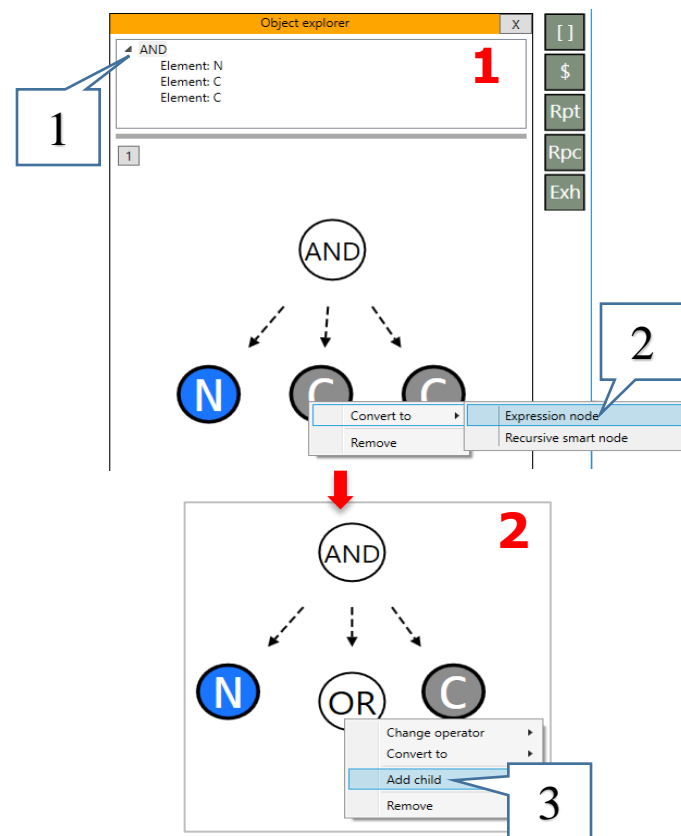
Workflow process

Structures search

Example 5: Searching for chemicals having primary or secondary amine group, but not amide



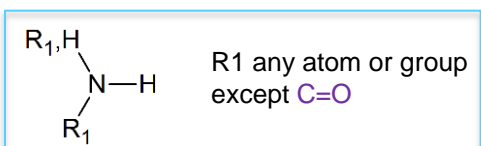
1. Go back to the “AND”;
2. Right click over the second query and select **Convert to > Expression node**;
3. Right click over the **OR** and select **Add child**. Add two child queries.



Workflow process

Structures search

Example 5: Searching for chemicals having primary or secondary amine group, but not amide



1. Click on the first item of “OR”
2. Define primary amine group by select the *Element* to be **N** and *Total hydrogens* to be **1**;
3. Click on the second item of “OR”
4. Define secondary amine group by select the *Element* to be **N** and *Total hydrogens* to be **2**;
5. Go back to the **AND** root.

The screenshot illustrates the search workflow in the QSAR Toolbox. It shows a hierarchical tree structure with 'AND' and 'OR' nodes. Two search criteria are defined:

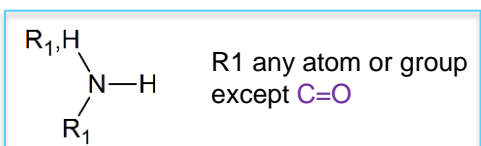
- Criterion 1 (Primary amine):** Element: N, Total hydrogens (H): 1.
- Criterion 2 (Secondary amine):** Element: N, Total hydrogens (H): 2.

The interface includes dropdown menus for various chemical properties and checkboxes for selection. The 'OR' node is highlighted, indicating the search for chemicals having either primary or secondary amine groups.

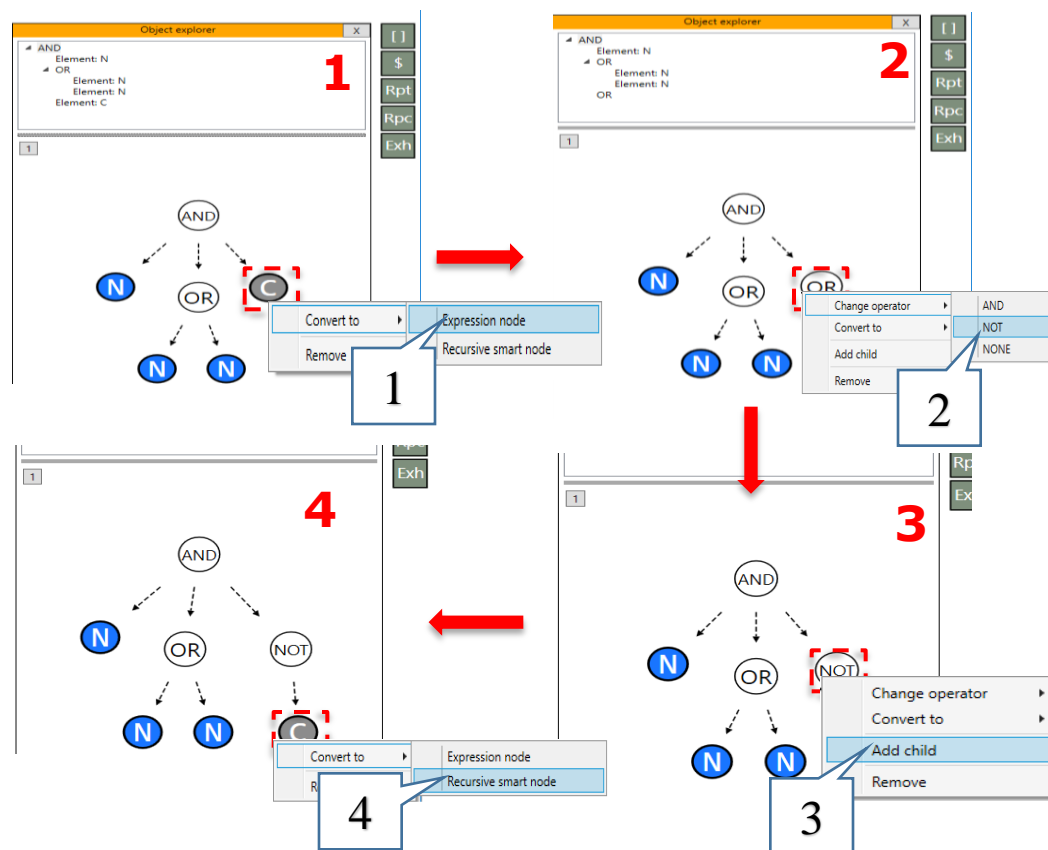
Workflow process

Structures search

Example 5: Searching for chemicals having primary or secondary amine group, but not amide



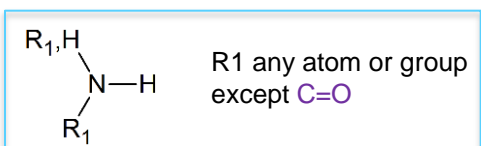
1. Right click on the last query and select **Convert to > Expression node**;
2. Right click on the new OR expression node and select **Change operator > NOT**;
3. Right click on the NOT query and select **Add child**;
4. Right click on the new child query and select **Convert to > Recursive smart node**.



Workflow process

Structures search

Example 5: Searching for chemicals having primary or secondary amine group, but not amide



1. Click on the **Recursive SMARTS** label. New SMARTS Editor window appears;
2. Draw amide group (NC=O);
3. Click over the nitrogen (N) of the amide group;
4. Select **Make first**;
5. Confirm with **OK**.

The screenshot shows the SMARTS Editor window with the query [N](C(=O)H)H and the Object Explorer window showing the query tree. The workflow steps are numbered 1 through 5:

1. Click on the **Recursive SMARTS** label in the Object Explorer.
2. Draw the amide group (NC=O) in the SMARTS Editor.
3. Click over the nitrogen (N) of the amide group in the SMARTS Editor.
4. Select **Make first** in the SMARTS Editor.
5. Confirm with **OK** in the SMARTS Editor.

The Object Explorer shows the query tree structure:

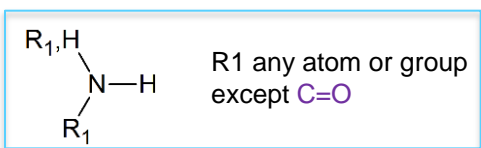
```

graph TD
    AND((AND)) --> N1((N))
    AND --> OR1((OR))
    AND --> NOT1((NOT))
    OR1 --> N2((N))
    OR1 --> N3((N))
    NOT1 --> DOLLAR((\$))
  
```

Workflow process

Structures search

Example 5: Searching for chemicals having primary or secondary amine group, but not amide

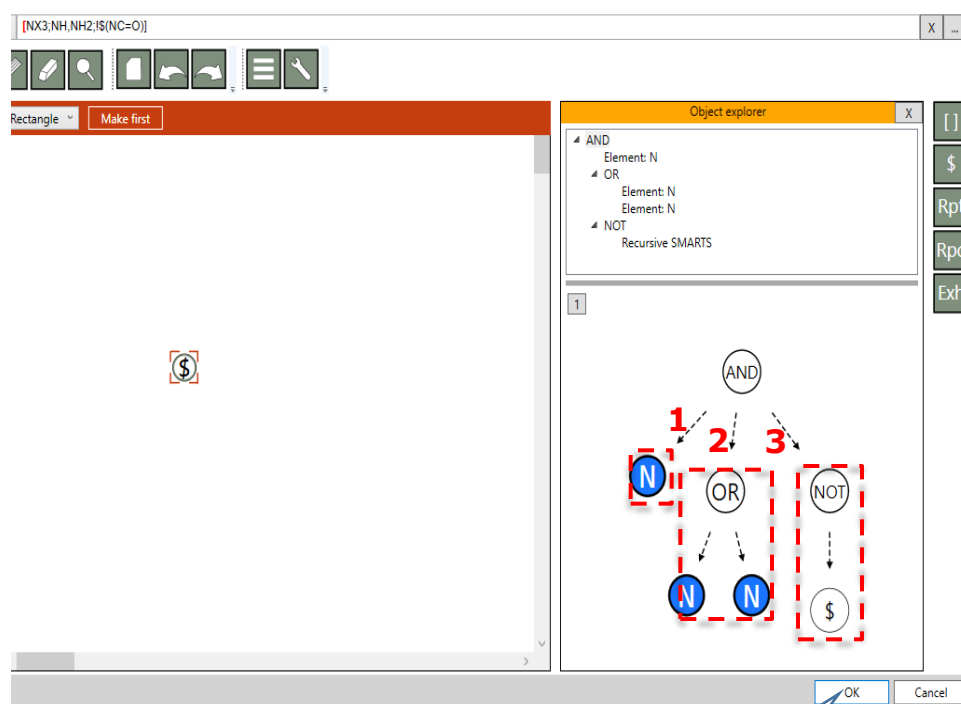


Now we are ready to execute the search.

Recap:

1. We defined that we will search for chemicals with nitrogen atom with three connections to other atoms;
2. We defined that the nitrogen to be part of primary or secondary amine functional group;
3. We defined that the nitrogen to be not part of amide functional group.

Click on **OK** (1) to execute the search.

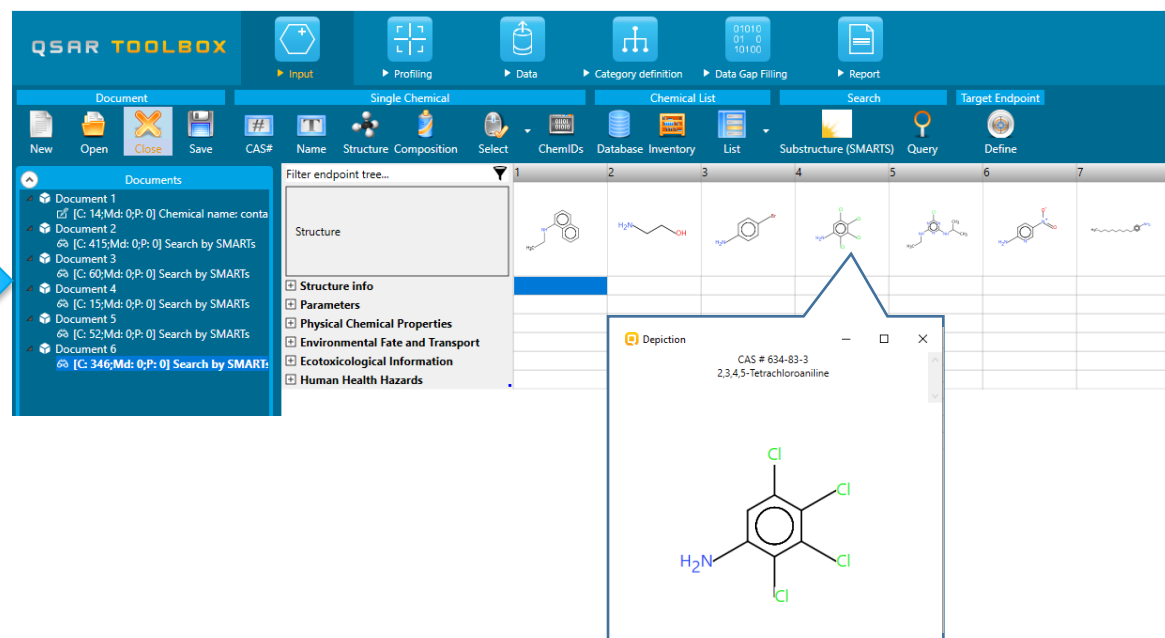
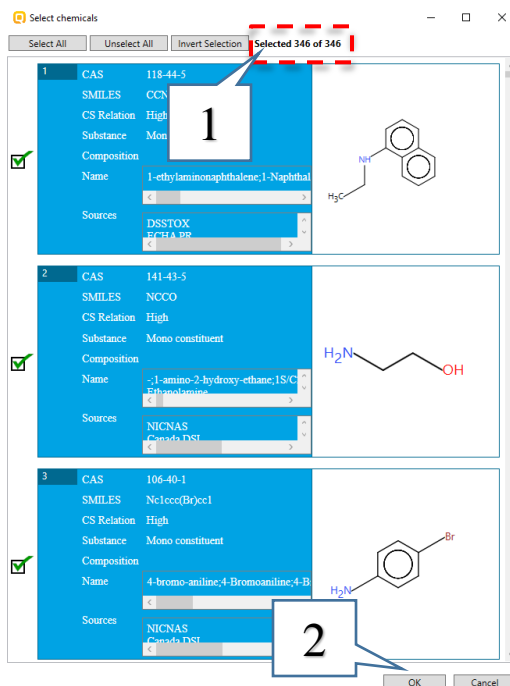


Workflow process

Structures search

Example 5: Searching for chemicals having primary or secondary amine group, but not amide

346 chemicals (1) are found in the selected database. Click on **Select All** button and then click on **OK** (2) to load all of them on the data matrix. You can see that all of the chemicals fulfill the criteria.



Workflow process

Overview

- Five examples illustrating searching for structures by SMARTS editor fragments were represented:
 - **[]** – Searching for Aryl halides;
 - **Rpc** – Searching for Five-membered heteroaromatic compounds;
 - **Rpt** – Searching for aryl compounds with three or more methyl groups;
 - **Exh** – Searching for hydrocarbon structures with 3 to 6 atoms and at least one hydroxyl group;
 - **\$** - Searching for chemicals having primary or secondary amine group, but not amide

Congratulation

- You have completed the tutorial on the substructure search by SMARTS editor;
- You have now been introduced to each of the general TB SMARTS fragments;
- Note proficiency comes with practice.