

OECD QSAR Toolbox v.4.1

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.

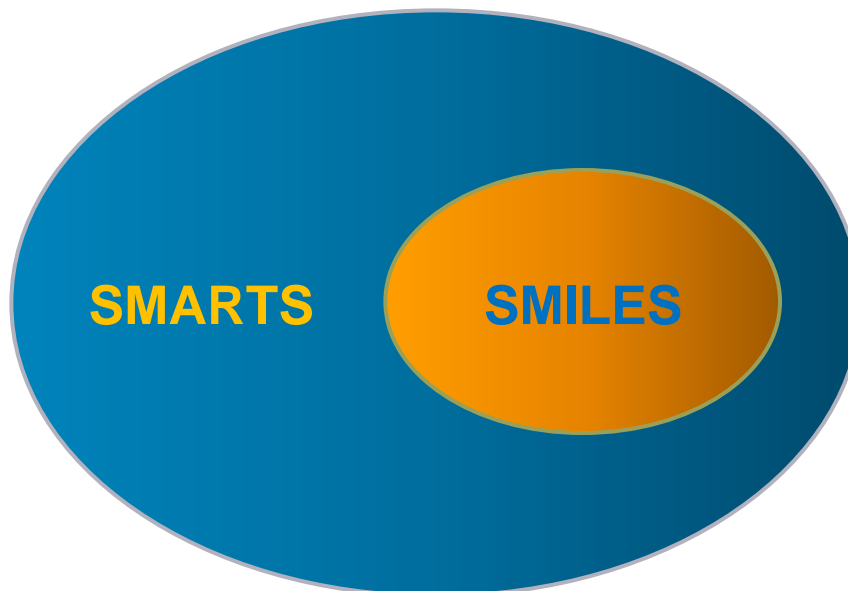
Outlook

- 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*

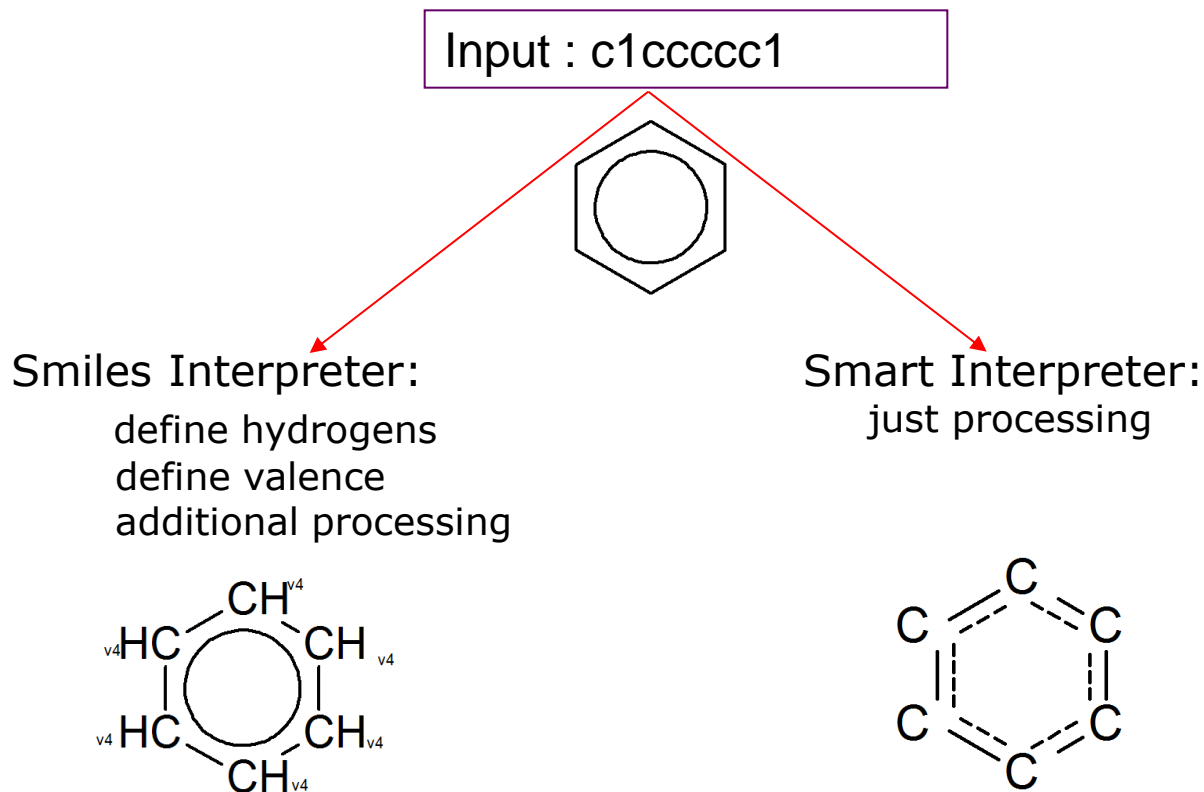


*<http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>

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

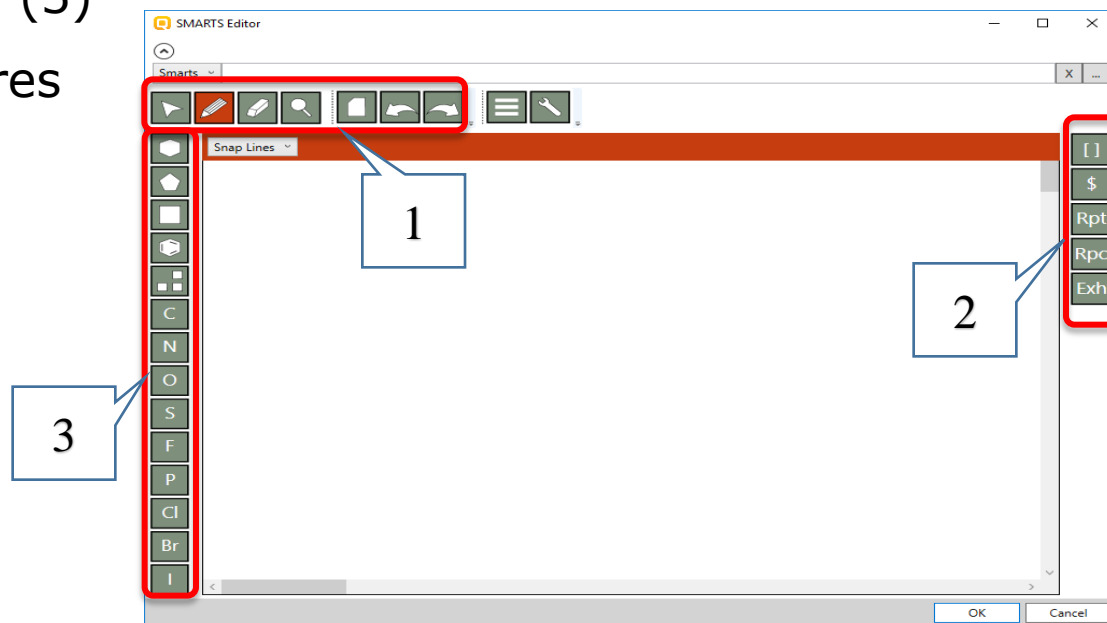
Overview

- 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

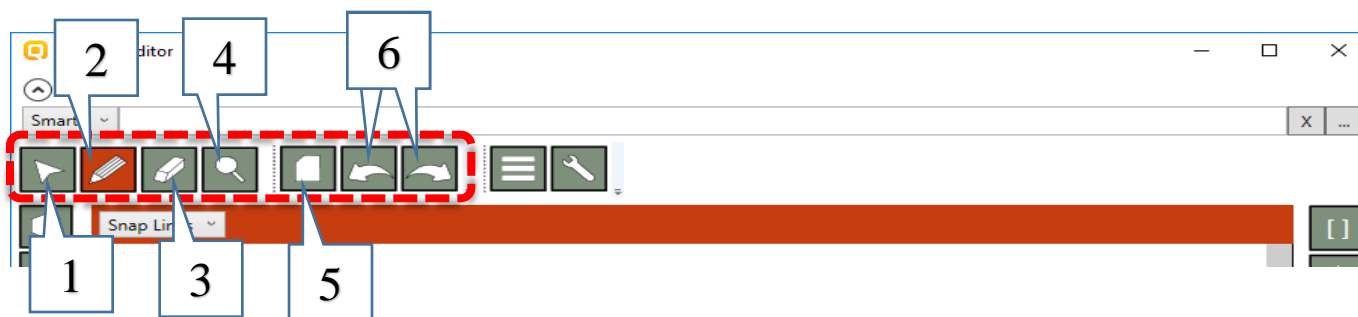
The SMARTS Editor window includes:

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



SMARTS Editor window

Action toolbar



The **Action toolbar** includes:

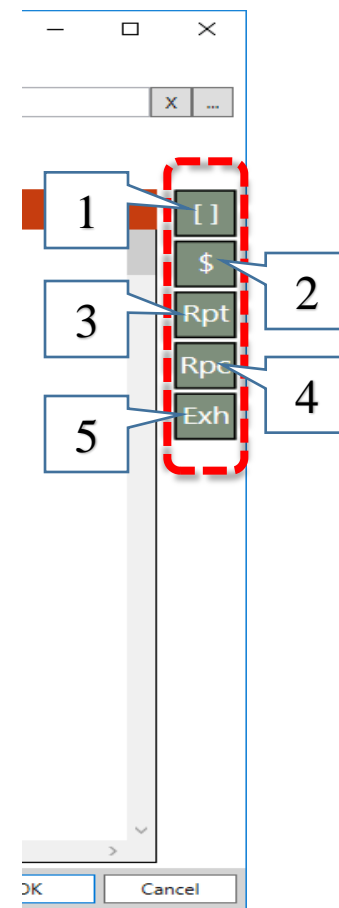
- **Selection tool** (1) – allows to define atom characteristics, to mark or rotate a structure, to start new activity, etc;
- **Pencil** button (2) – allows drawing of hydrocarbon structures (scaffolds);
- **Eraser** button (3) – removes selected bonds or other parts of a drawn structure;
- **Magnifying 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.

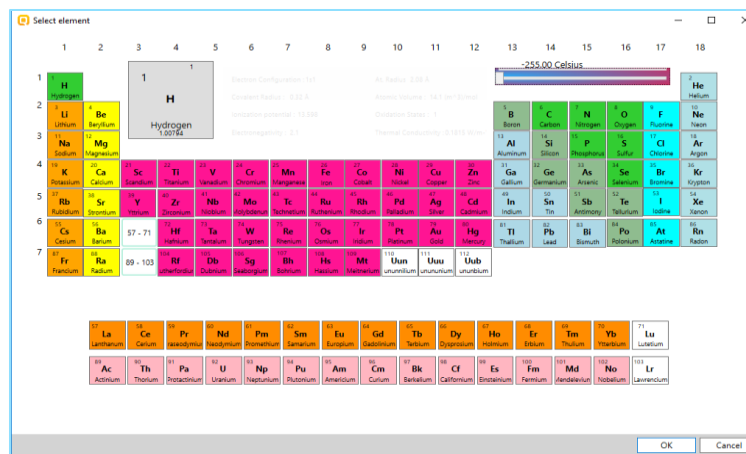
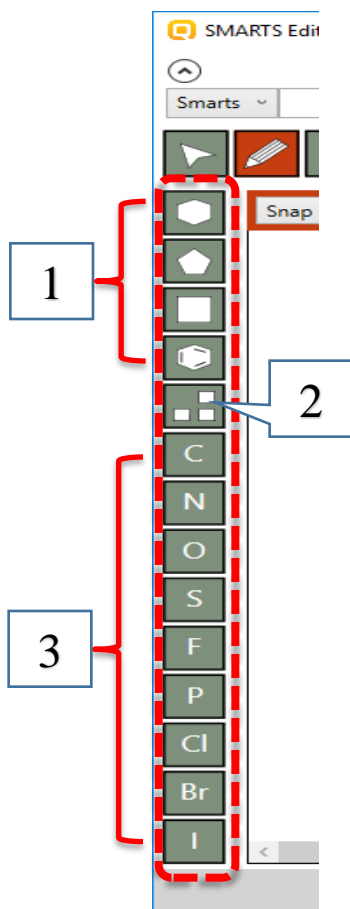


SMARTS Editor window

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;

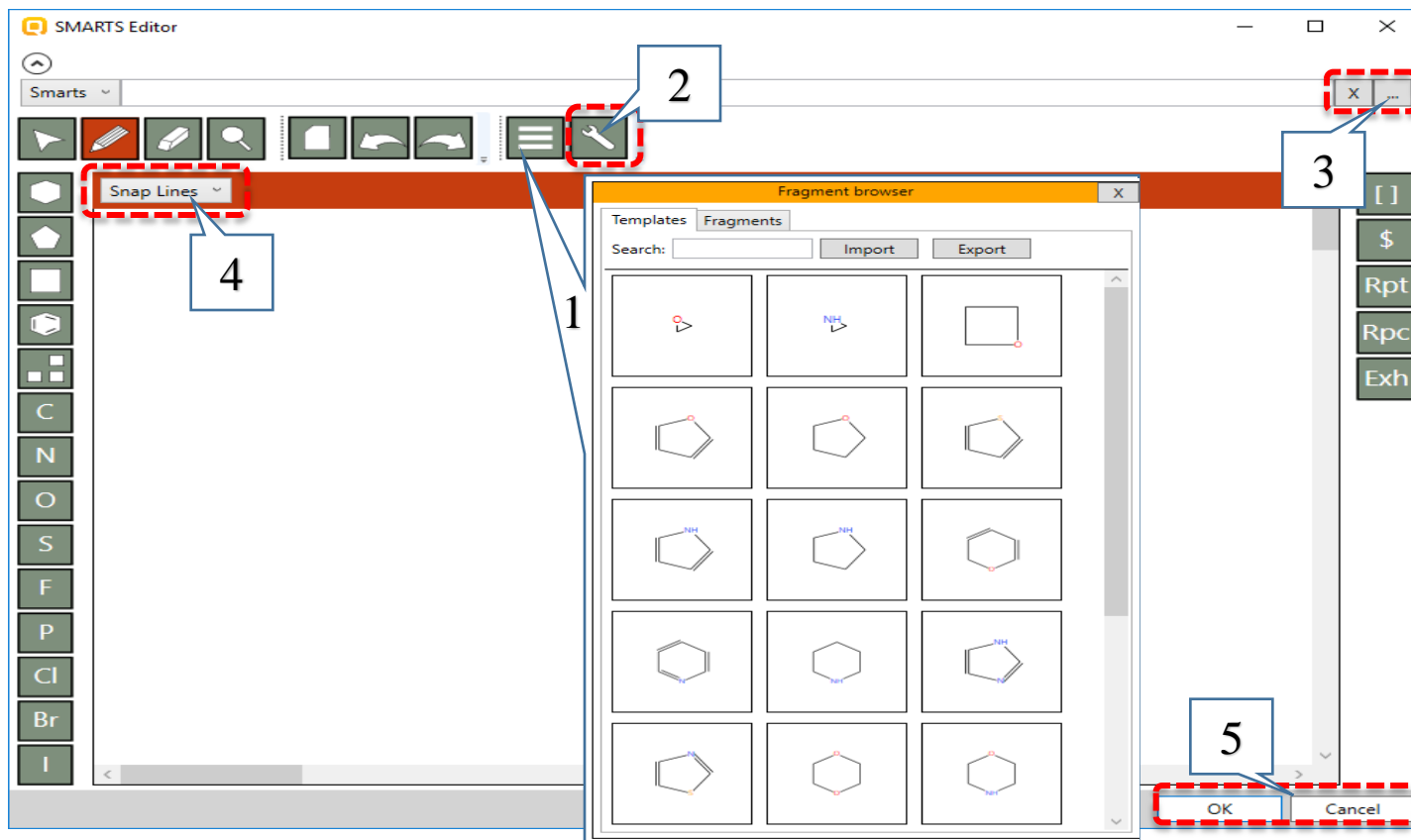
- **Object explorer**

See on the next two slides.

SMARTS Editor window

Additional features

- Buttons

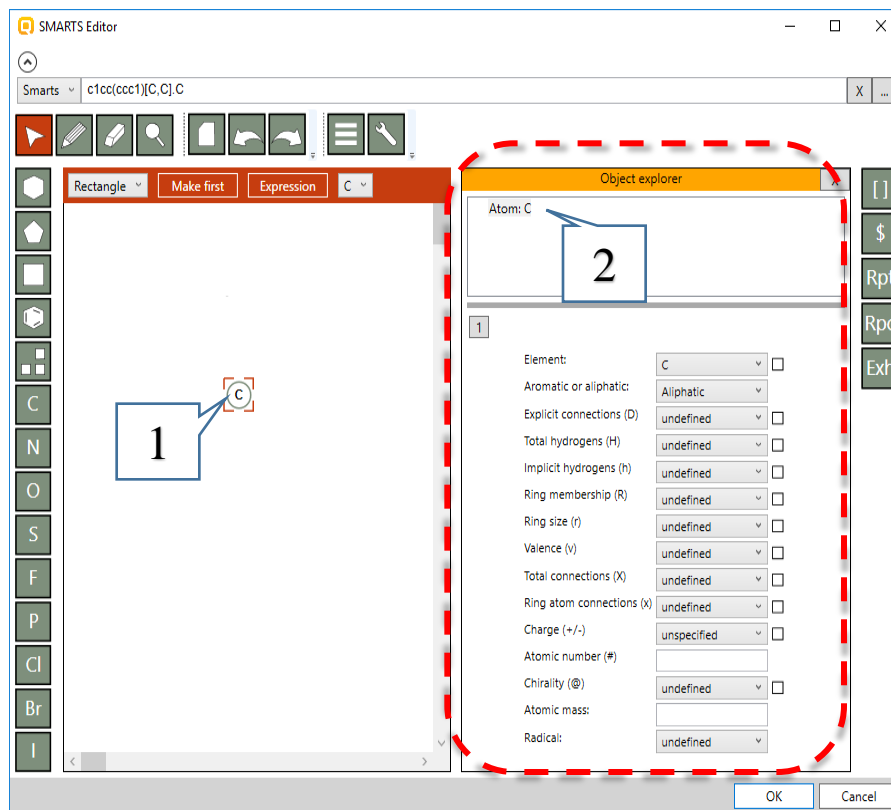


SMARTS Editor window

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

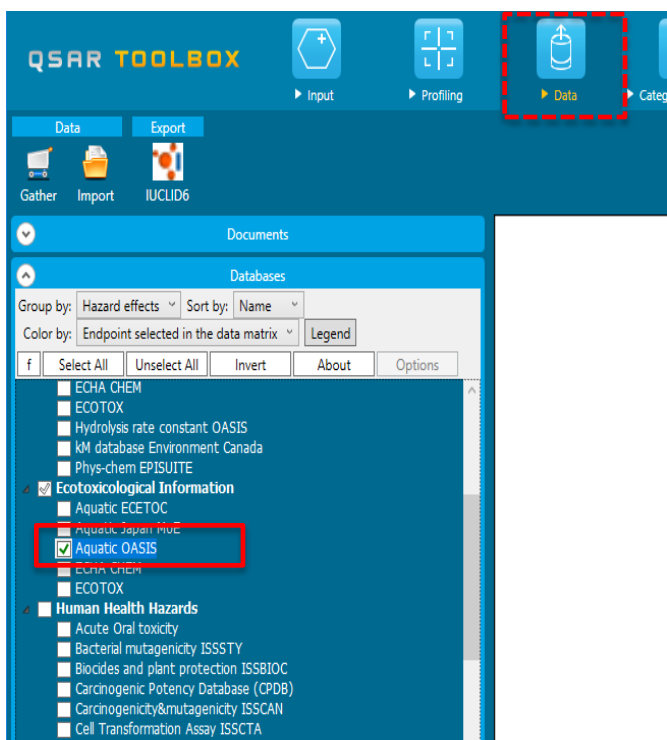
- In this exercise we will demonstrate how to construct different queries for structure searching;
- We will use each of the new fragments for the current purposes;
- *Aquatic OASIS* database will be used.

Outlook

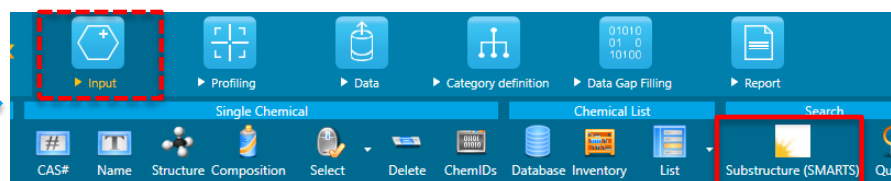
- Background
- Aims
- Overview of SMARTS language
- SMARTS Editor window
- The exercise
- **Workflow process**

Workflow process Overview

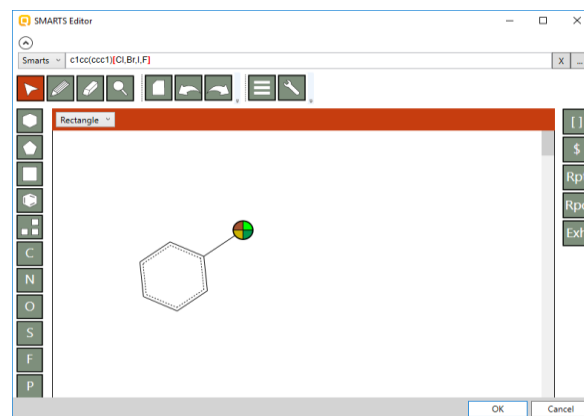
Step 1: Select database(s)/inventory(ies)



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



Step 3: Specify criteria for searching structures

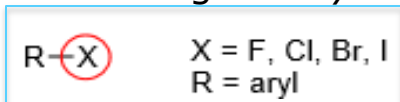


We will use **Aquatic OASIS** database for all examples. Searching criteria will be different only.

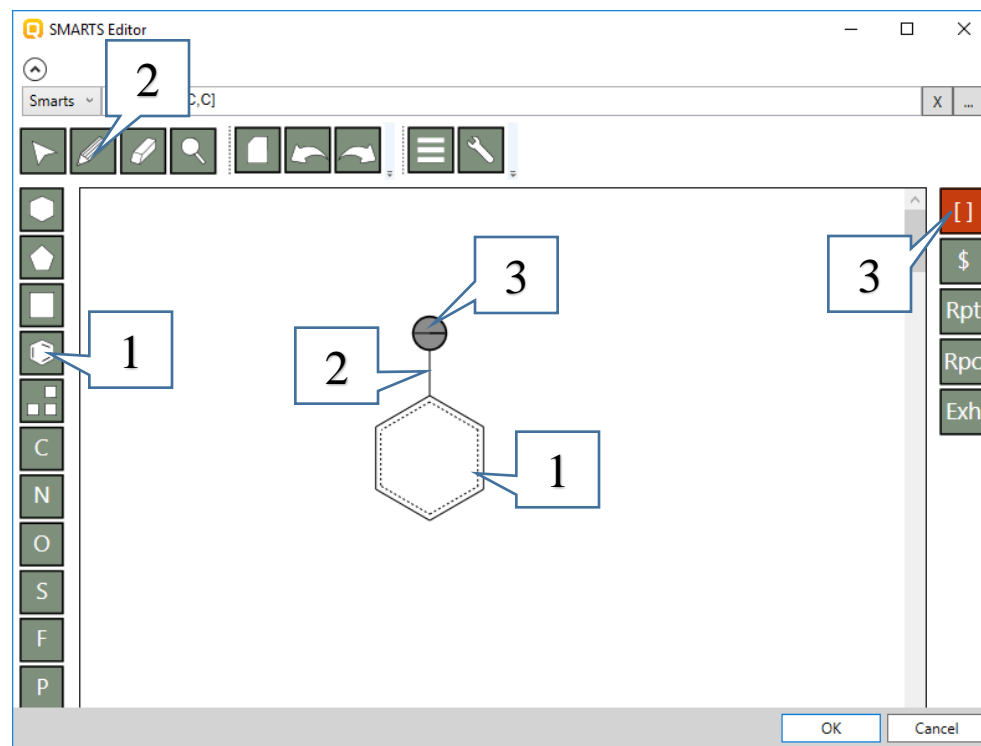
Workflow process

Structures search

Example 1: Searching for *Aryl halides*



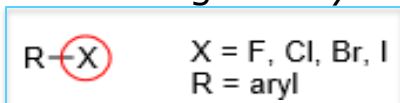
1. Click on the **benzene** template and then click in the drawing pane;
2. Click on the **pencil** button and then draw a bond;
3. Click on the “[]” (enumeration) button and then click over the new carbon atom.



Workflow process

Structures search

Example 1: Searching for *Aryl halides*

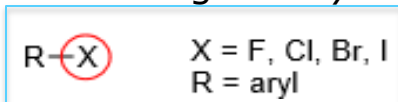


1. Click on the **selection tool** and then click on the new object;
2. *Object explorer* appears;
3. Right click over the “**OR**” expression node and select **Add child**. Repeat this step again.

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 figure consists of two screenshots of the QSAR Toolbox software interface, illustrating the workflow for searching for aryl halides.

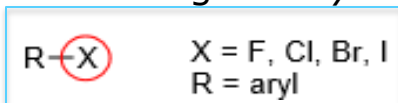
Top Screenshot: Shows the initial search. The SMARTS string is c1ccccc1[C,C,C,C]. A benzene ring structure is displayed with a red box around the top carbon atom. A panel on the right shows a search tree with "Element C" selected (labeled 1). Below it, four "Element C" nodes are shown, with the first one highlighted by a red dashed box.

Bottom Screenshot: Shows the selection of a specific element. The SMARTS string is updated to c1cc(ccc1)[Cl,C,C,C]. The benzene ring structure is shown with a red box around the top carbon atom (labeled 3). A panel on the right shows the search tree with "Element Cl" selected (labeled 4). A drop-down menu is open, showing the "Element" list with "Cl" selected (labeled 2). The "Element" list includes C, N, O, P, S, F, Br, I, and AnyAtom.

Workflow process

Structures search

Example 1: Searching for *Aryl halides*



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

SMARTS Editor

Smarts: c1cc(ccc1)[Cl,Br,I,F]

Object explorer

OR

- Element: Cl
- Element: Br
- Element: I
- Element: F

1

1

- Element: F
- Aromatic or aliphatic: Aliphatic
- Explicit connections (D): undefined
- Total hydrogens (H): undefined
- Implicit hydrogens (h): undefined
- 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:

2

OK Cancel

Workflow process

Structures search

Example 1: Searching for *Aryl halides*

415 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.

1

ID	CAS	SMILES	CS Relation	Name
1	1836-...	[O-]Jc(C)C...	High	4-NITROPHENYL 2,4,6-T... 1,3,5-trichloro-2-(4-nitroph... 2,4,6-Trichlorohenvl-4-Ni...
2	1746-81-2	CON(C)C(=O)Nc1ccc(Cl)cc1	High	3-(4-chlorophenyl)-1-methio... monolinuron Monolinuron (ISO)
3	1712-70-5	CC(=C)c1ccc(Cl)cc1	High	4-Isopropenyl-chlorobenzene 1-chloro-4-(prop-1-en-2-yl)... 4-chloro- α -methylstyrene
4	1570-65-6	Cc1cc(C)cc(Cl)c1O	High	2,4-Dichloro-6-methyl-phe... 2,4-dichloro-6-methylphenol

2

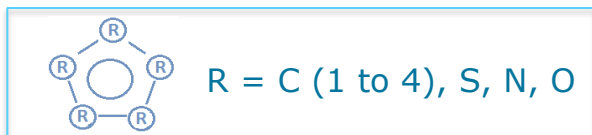


CON(C)C(=O)Nc1ccc(Cl)cc1

Workflow process

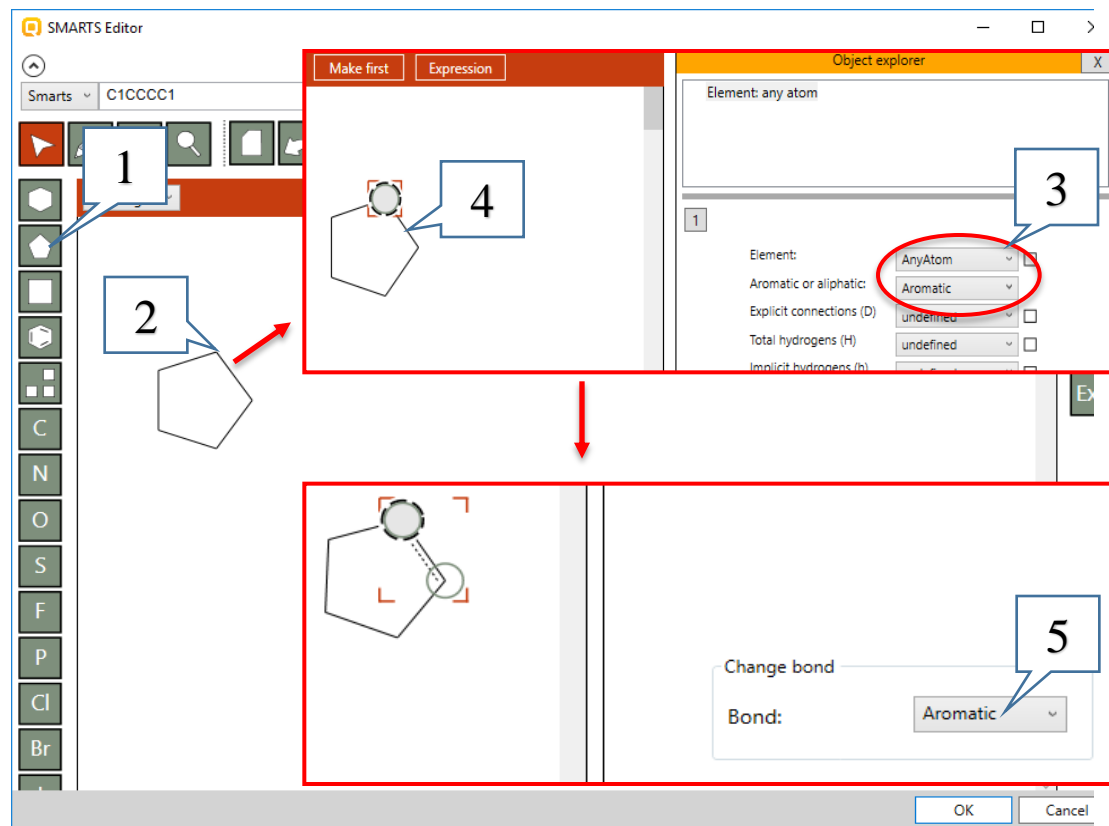
Structures search

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



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.



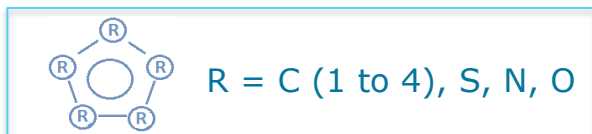
The screenshot shows the SMARTS Editor interface with the following workflow steps highlighted:

- 1:** The 'Make first' button in the toolbar is clicked.
- 2:** A pentane ring template is selected from the toolbar.
- 3:** In the 'Object explorer' panel, the 'Element' dropdown is set to 'AnyAtom' and the 'Aromatic or aliphatic' dropdown is set to 'Aromatic'.
- 4:** A carbon atom in the pentane structure is selected.
- 5:** A bond is selected, and the 'Change bond' dialog box is used to set the bond type to 'Aromatic'.

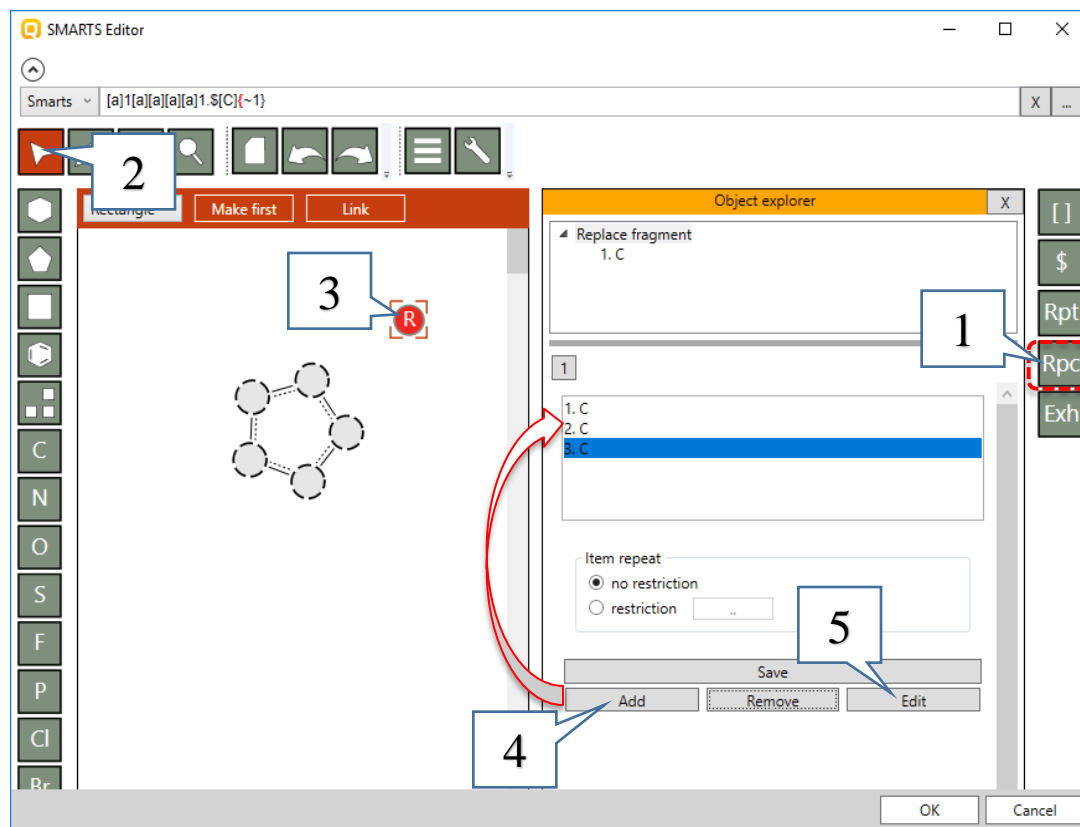
Workflow process

Structures search

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



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.



SMARTS Editor

Smarts [a]1[a][a][a][a]1.[C]{~1}

Object explorer

Replace fragment

1. C

2. C

3. C

Item repeat

no restriction

restriction ..

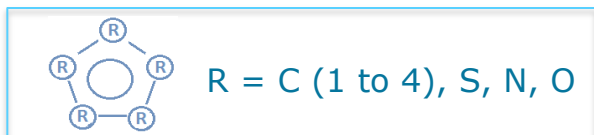
Add Remove Edit

OK Cancel

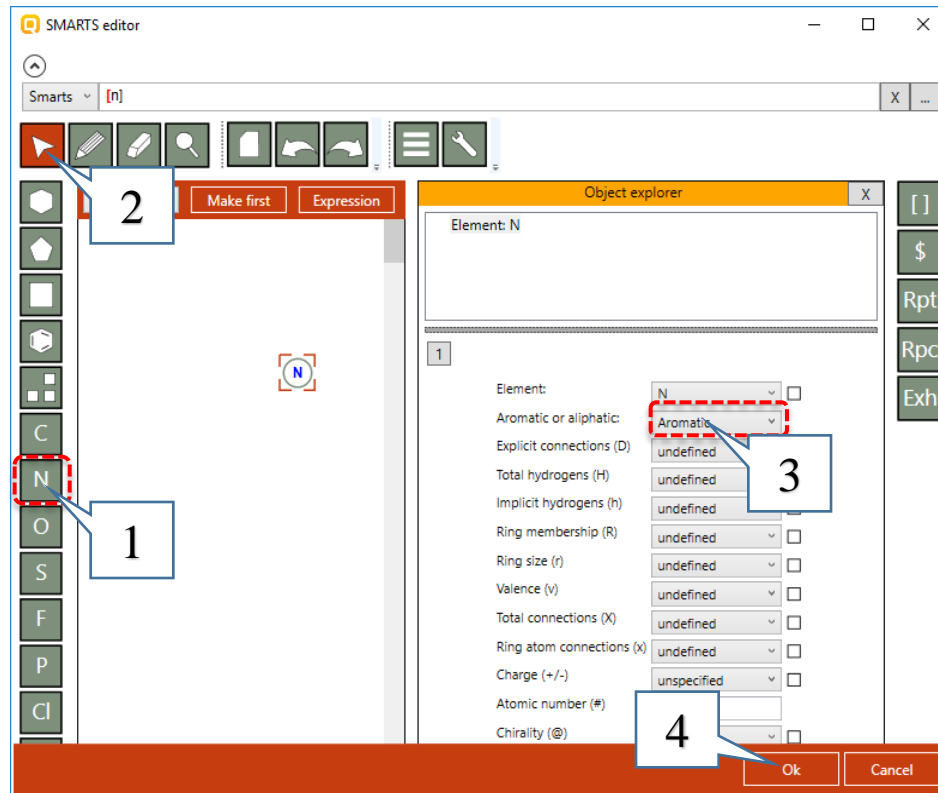
Workflow process

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



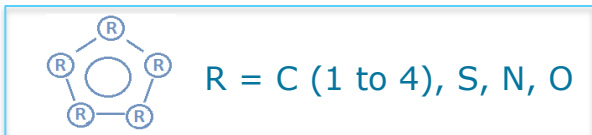
The screenshot shows the SMARTS editor window with the following elements:

- Step 1:** The 'N' symbol in the left-hand element palette is highlighted with a red dashed box.
- Step 2:** A mouse cursor is over the 'selection tool' icon in the top toolbar.
- Step 3:** In the 'Object explorer' panel, the 'Aromatic' dropdown menu for the 'Element' property is open and highlighted with a red dashed box.
- Step 4:** The 'Ok' button at the bottom right of the Object explorer panel is highlighted with a red dashed box.

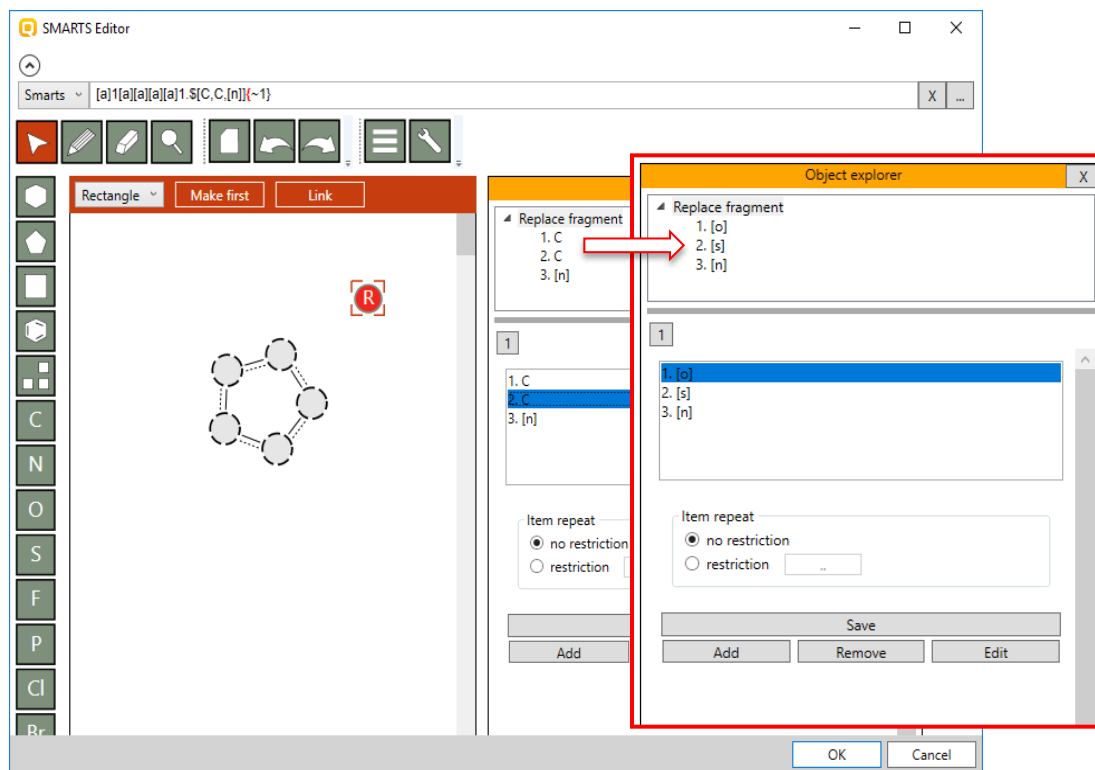
Workflow process

Structures search

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



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



SMARTS Editor

Smarts [a]1[a][a][a][a]1.S[C,C,[n]]{~1}

Object explorer

Replace fragment

- 1. [o]
- 2. [s]
- 3. [n]

Item repeat

no restriction

restriction

Save

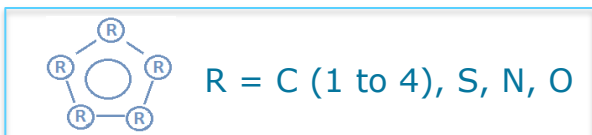
Add Remove Edit

OK Cancel

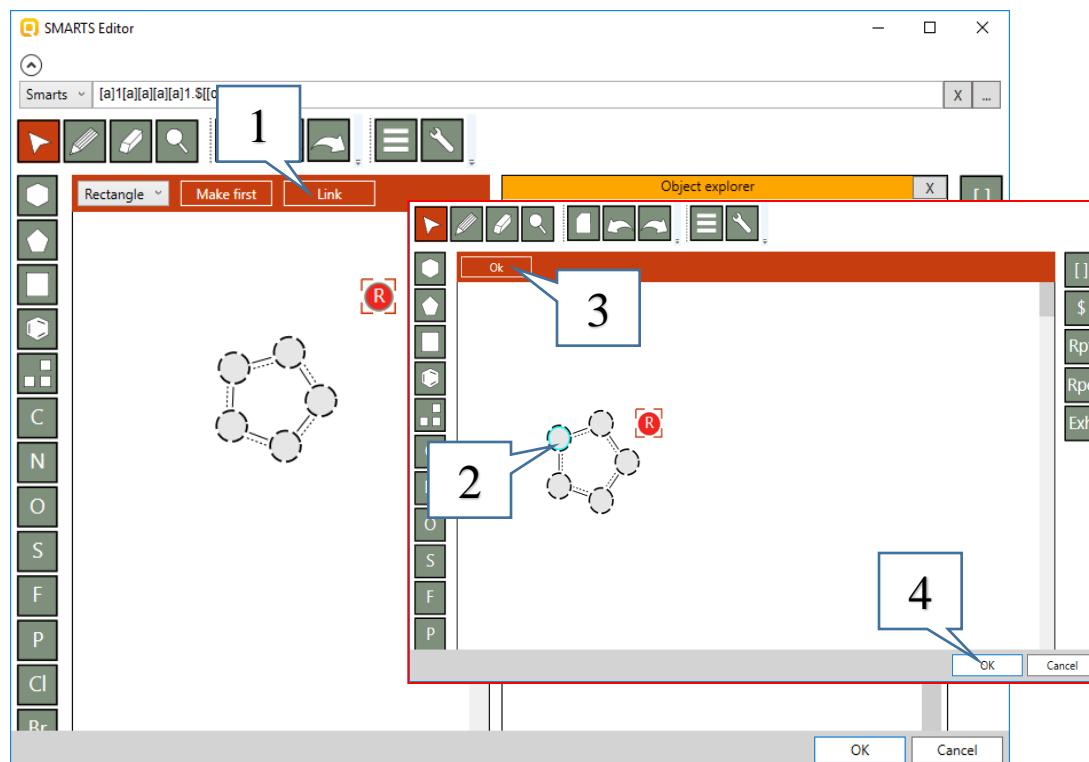
Workflow process

Structures search

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



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



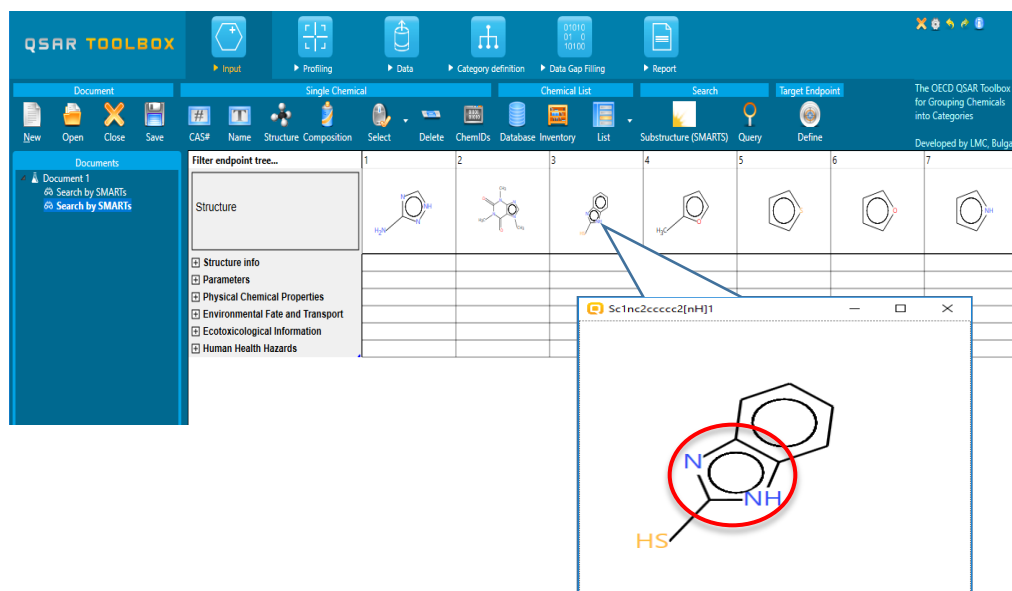
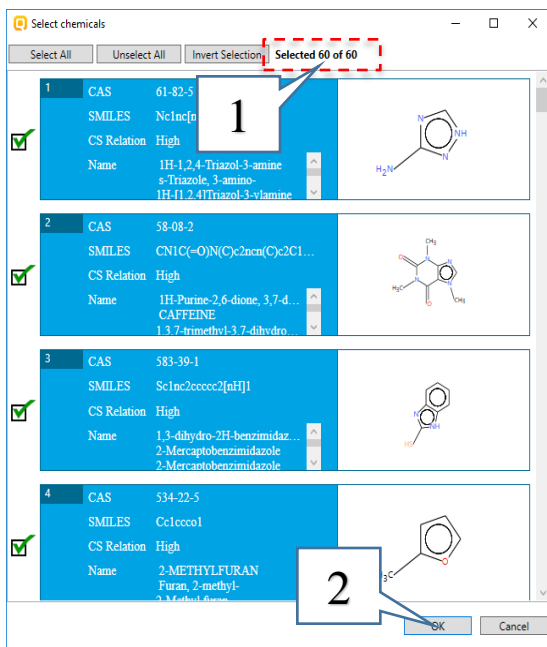
The screenshot shows the SMARTS Editor window. The top toolbar contains a 'Link' button (1). The main workspace displays a five-membered ring structure (2). A dialog box is open, showing the criteria for the search (3). The main 'OK' button is highlighted (4).

Workflow process

Structures search

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

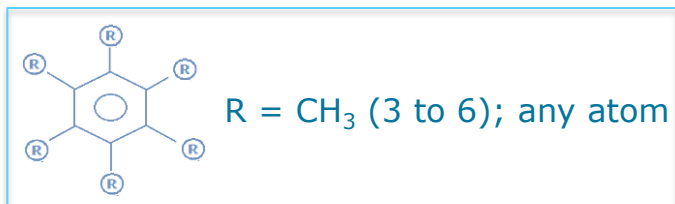
60 chemicals (1) are found in the selected database. Click on **OK** (2) to load all of them on the data matrix. You can see that all of the chemicals fulfill the criteria.



Workflow process

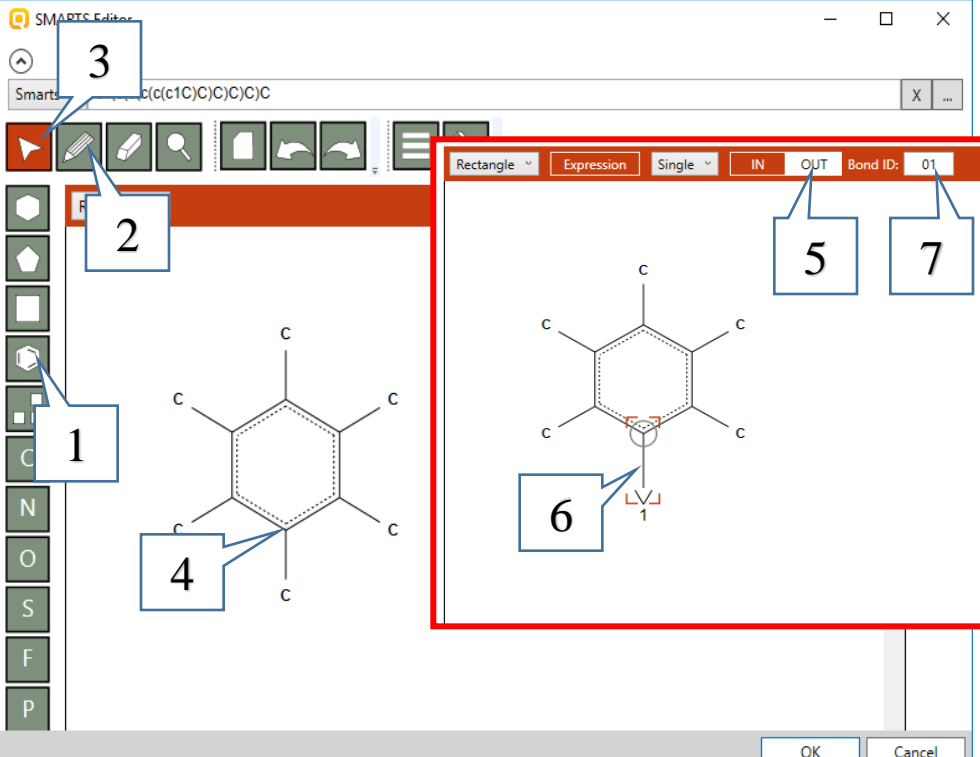
Structures search

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.

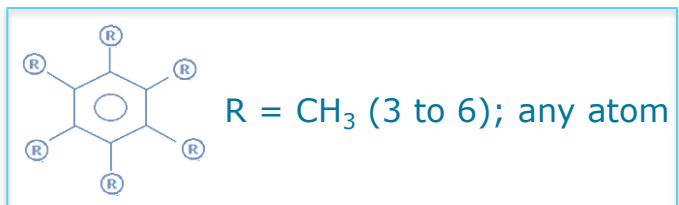
Repeat the steps above (from 4 to 7) for each of the bonds.



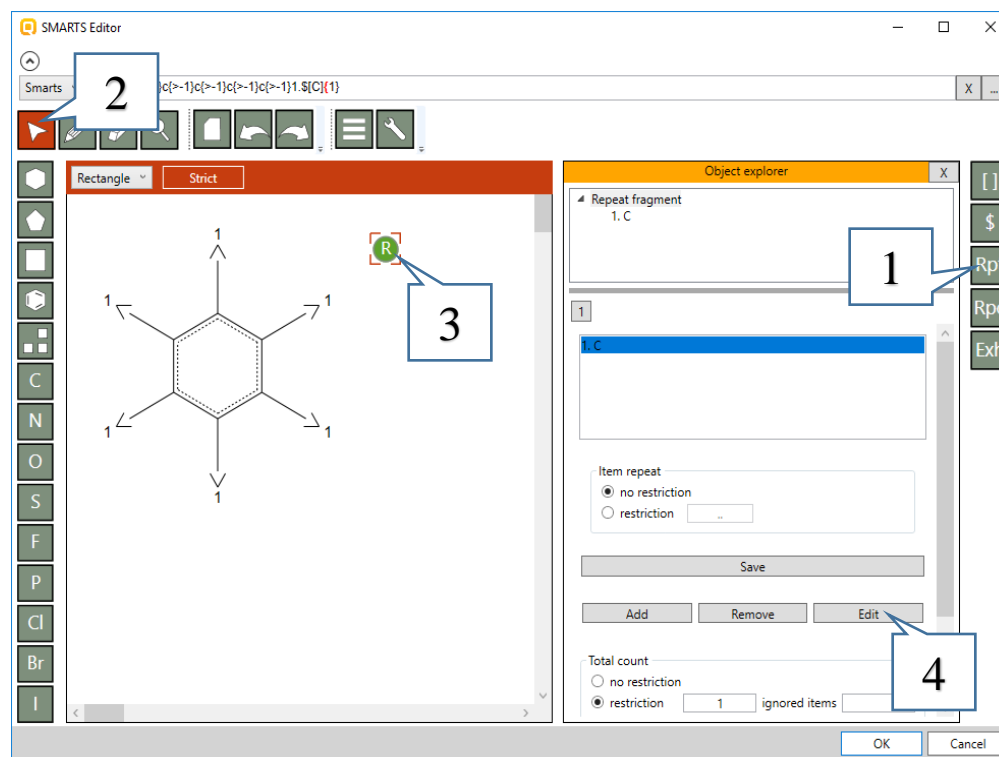
Workflow process

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.



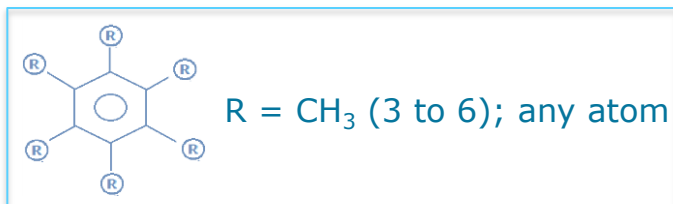
The screenshot shows the SMARTS Editor window with the following elements:

- Smarts input field:** Contains the query [c]>[1][c]>[1][c]>[1][c]>[1][c]>[1].S[C][1]. A callout box labeled '2' points to the selection tool icon in the toolbar.
- Drawing pane:** Shows a benzene ring with six '1' labels at the attachment points. A green 'R' fragment is highlighted with a red box, and a callout box labeled '3' points to it.
- Object explorer:** Shows a tree view with 'Repeat fragment' and '1. C'. A callout box labeled '1' points to the 'Rpt' button on the right-hand side.
- Configuration panel:** Shows 'Item repeat' with 'no restriction' selected and 'restriction' set to '1'. Below it, 'Total count' has 'restriction' selected and '1' entered. A callout box labeled '4' points to the 'Edit' button.

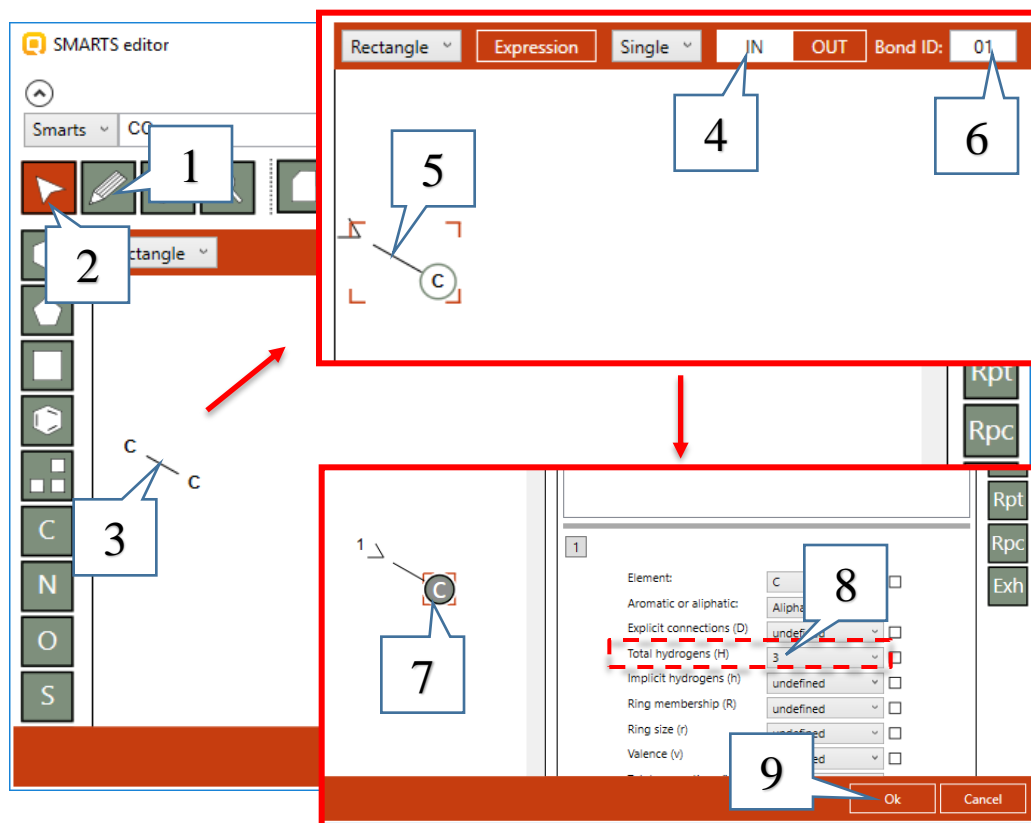
Workflow process

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**.



SMARTS editor

Rectangle Expression Single IN OUT Bond ID: 01

Smarts CO

Rectangle

C N O S

1 2 3 4 5 6 7 8 9

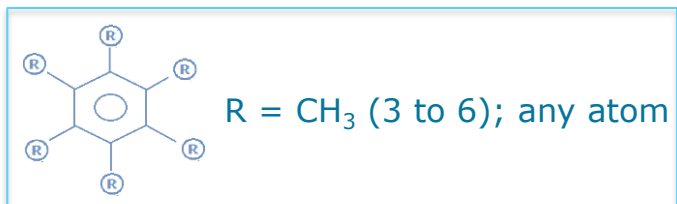
Element: C
 Aromatic or aliphatic: Aliphatic
 Explicit connections (D): undefined
 Total hydrogens (H): 3
 Implicit hydrogens (h): undefined
 Ring membership (R): undefined
 Ring size (r): undefined
 Valence (v): 4

Ok Cancel

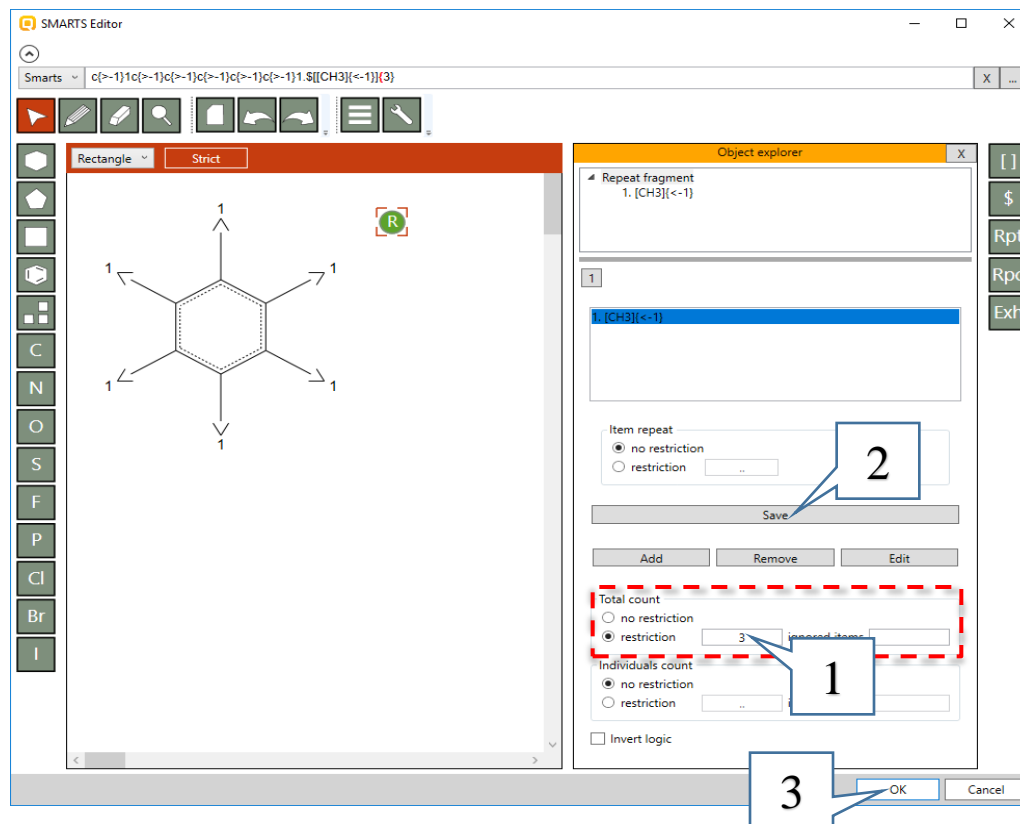
Workflow process

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**



SMARTS Editor

Smarts c[>-1]1c[>-1]c[>-1]c[>-1]c[>-1]c[>-1]1.[CH3][<-1]{3}

Rectangle Strict

Object explorer

Repeat fragment

1. [CH3][<-1]

Item repeat

no restriction

restriction

Save

Add Remove Edit

Total count

no restriction

restriction 3

Individuals count

no restriction

restriction

Invert logic

OK Cancel

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.

Select chemicals

Select All Unselect All Invert Selection Selected 15 of 15

1	CAS	527-54-8	SMILES	<chem>Cc1cc(O)c(C)c1</chem>	
	CS Relation	High	Name	3,4,5-trimethylphenol 3,4,5-trimethylphenol Phenol, 3,4,5-trimethyl-	
2	CAS	527-53-7	SMILES	<chem>Cc1cc(C)c(C)c1</chem>	
	CS Relation	High	Name	1,2,3,5-TETRAMETHYL... 1,2,3,5-tetramethylbenzene 1,2,3,5-tetramethylbenzene	
3	CAS	488-23-3	SMILES	<chem>Cc1ccc(C)c(C)c1</chem>	
	CS Relation	High	Name	1,2,3,4-TETRAMETHYL... 1,2,3,4-tetramethylbenzene 1,2,3,4-tetramethylbenzene	
4	CAS	108-67-8	SMILES	<chem>Cc1cc(C)cc1</chem>	
	CS Relation	High	Name	1,3,5-TRIMETHYLBENZ... Benzene, 1,3,5-trimethyl- 1,3,5-Trimethylbenzene	

OK Cancel



QSAR TOOLBOX

Document Single Chemical Chemical List Search Target Endpoint

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

Documents

- Document 1
- Search by SMARTS
- Search by SMARTS
- Search by SMARTS

Filter endpoint tree...

Structure

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

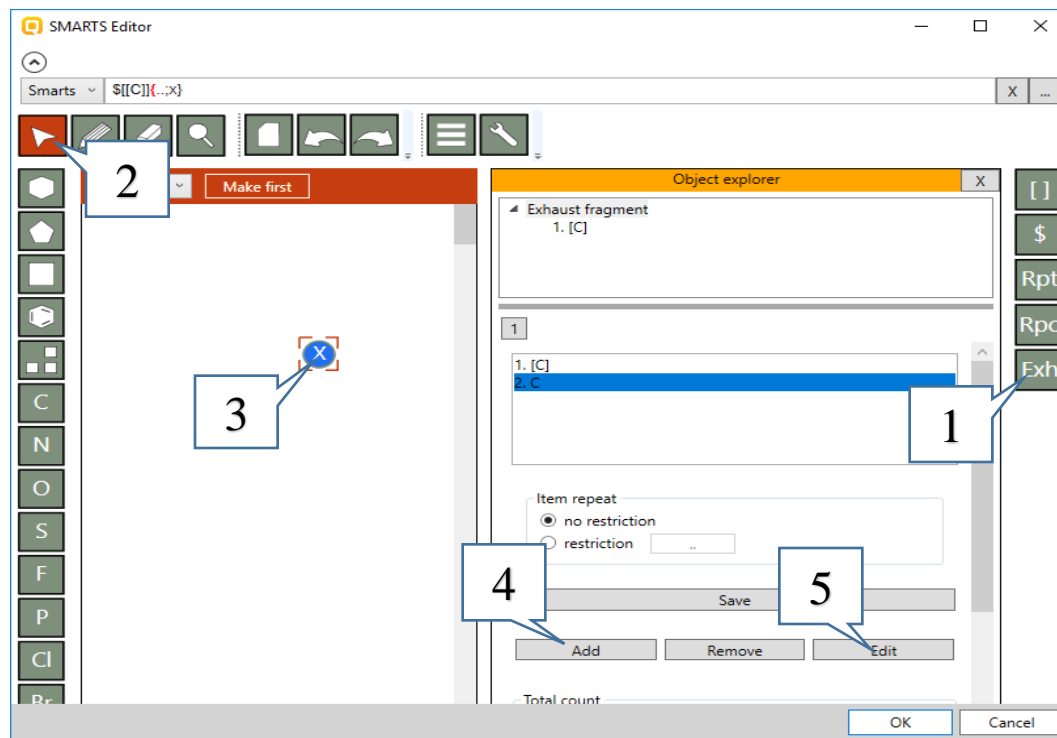
Cc1cc(O)cc(C)c1C

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.

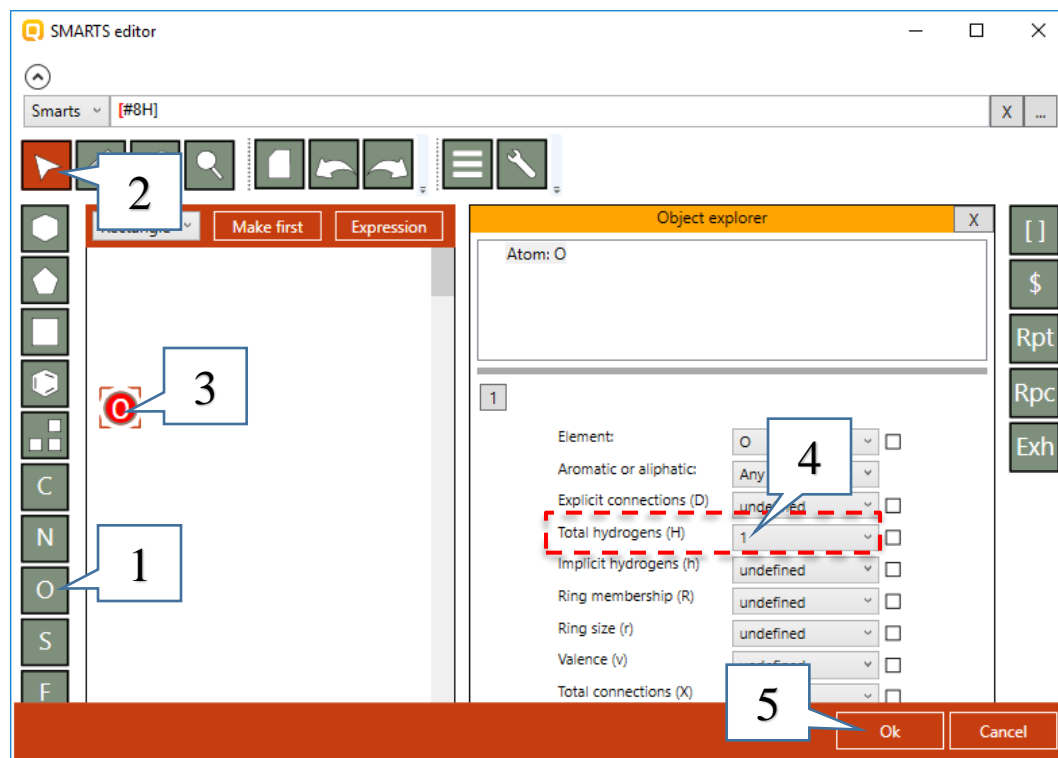


Workflow process

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*.

The image shows two screenshots of the 'Object explorer' window in a software interface, illustrating the workflow steps for searching for hydrocarbon structures with 3 to 6 atoms and at least one hydroxyl group.

Top Screenshot: The 'Object explorer' window shows an 'Exhaust fragment' with two items: '1. [C]' and '2. [#8H]'. A blue box labeled '1' points to item '2. [#8H]'. Below the fragment list, the 'Repeat set:' section shows 'From: 0' and 'To: 254', with the 'All possible' checkbox checked. A blue box labeled '2' points to the 'All possible' checkbox. A red arrow points down to the bottom screenshot.

Bottom Screenshot: The 'Object explorer' window shows the same 'Exhaust fragment' with '1. [C]' and '2. [#8H]'. A blue box labeled '5' points to the 'Exhaust fragment' header. The 'Repeat set:' section shows 'From: 1' and 'To: 254', with the 'All possible' checkbox unchecked. A red dashed box highlights the 'From: 1' and 'To: 254' fields, with a blue box labeled '3' pointing to it. A blue box labeled '4' points to the 'Save' button. A red '2' is visible in the top right corner of the window.

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**

The screenshot displays the QSAR Toolbox software interface. The main workspace is empty, with a toolbar at the top. On the left, there is a vertical toolbar with icons for various functions. The 'Object explorer' panel on the right shows the search criteria: 'Exhaust fragment' with '1. [C]' and '2. \$[[#8H]](1..)'. The 'Total count' section is highlighted with a red dashed box, showing 'restriction' selected with a range of '3..6'. The 'Individuals count' section shows 'no restriction' selected. The 'Save' button is highlighted with a blue callout box labeled '2', and the 'OK' button is highlighted with a blue callout box labeled '3'. A blue callout box labeled '1' points to the 'restriction' radio button in the 'Individuals count' section.

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 **OK** (2) to load all of them on the data matrix. You can see that all of the chemicals fulfill the criteria.

Select chemicals

Select All Unselect All Invert Selection Selected: 52 of 52

1	CAS	2117-11-5	SMILES	CC(O)CC#C	
2	CAS	764-01-2	SMILES	CC#CCO	
3	CAS	137-32-6	SMILES	CCC(C)CO	
4	CAS	107-21-1	SMILES	OCCO	

OK Cancel



QSAR TOOLBOX

Document Single Chemical Chemical List Search Target Endpoint

Filter endpoint tree...

Structure

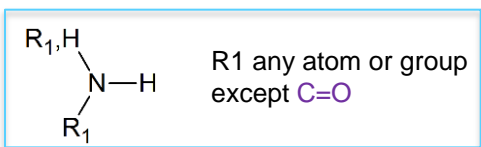
Structure info
Parameters
Physical Chemical Properties
Environmental Fate and Transport
Ecotoxicological Information
Human Health Hazards

ccc(c)co

Workflow process

Structures search

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



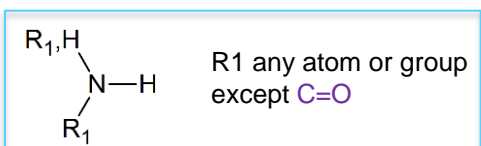
1. Select \$ (recursive SMARTS) fragment and put it in the drawing pane;
2. Click on the **selection tool**;
3. Click over the \$ fragment;
4. Right click over the \$ in the *Object explorer* panel and select **Convert to > Expression node**

The screenshot shows the SMARTS Editor interface. The drawing pane contains a '\$' symbol. The Object Explorer panel on the right shows a tree structure with '\$' selected. A context menu is open over the '\$' in the Object Explorer, with 'Convert to > Expression node' highlighted. Numbered callouts 1 through 4 indicate the sequence of actions described in the text.

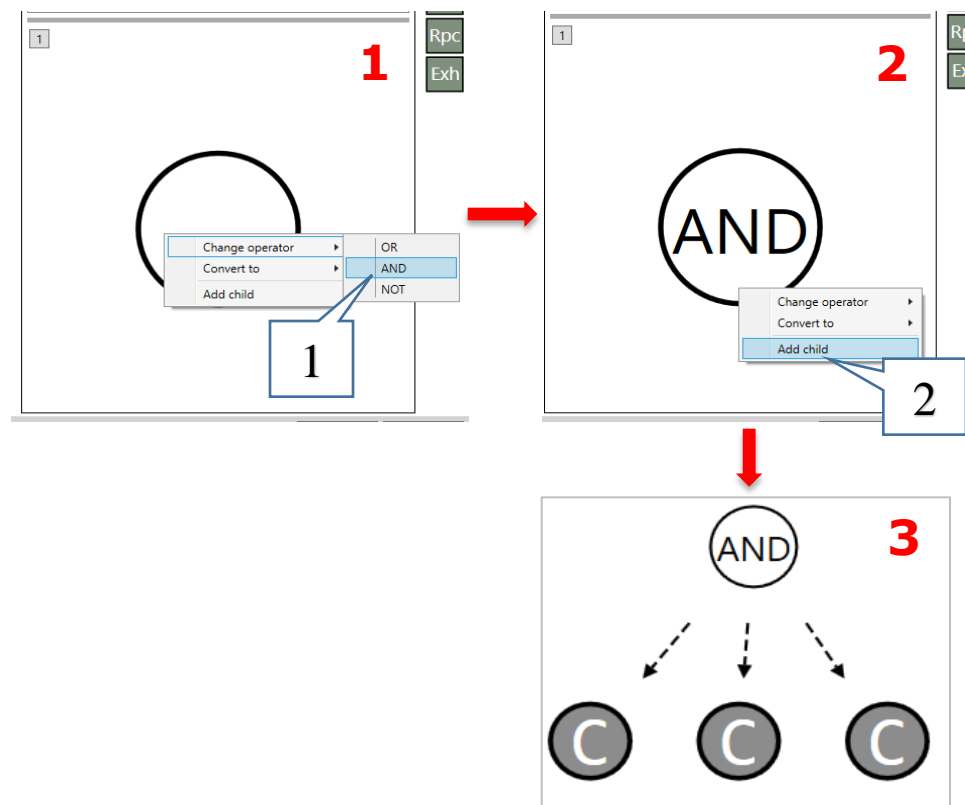
Workflow process

Structures search

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



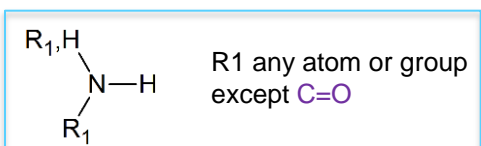
1. Right click over the ball and select **Change operator > AND**
2. Right click over the AND operator and select **Add child**.
3. Repeat the 2nd step three times to obtain three "child" queries.



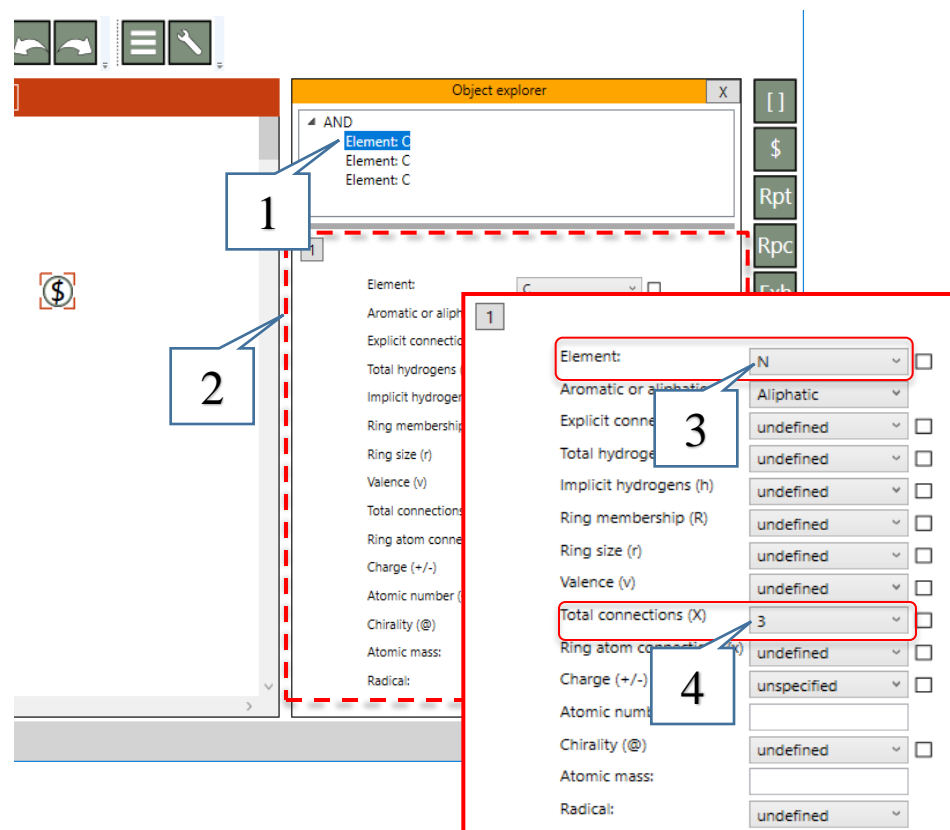
Workflow process

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.



Object explorer

AND

Element: C

Element: C

Element: C

Element:

Aromatic or aliphatic

Explicit connections

Total hydrogens

Implicit hydrogens

Ring membership

Ring size (r)

Valence (v)

Total connections

Ring atom connections

Charge (+/-)

Atomic number

Chirality (@)

Atomic mass

Radical

Element: N

Aromatic or aliphatic: Aliphatic

Explicit connections: undefined

Total hydrogens: undefined

Implicit hydrogens (h): undefined

Ring membership (R): undefined

Ring size (r): undefined

Valence (v): undefined

Total connections (X): 3

Ring atom connections (X): undefined

Charge (+/-): unspecified

Atomic number:

Chirality (@): undefined

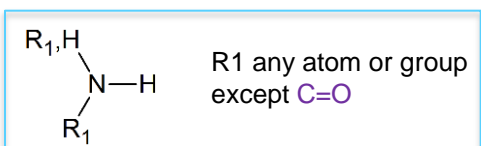
Atomic mass:

Radical: undefined

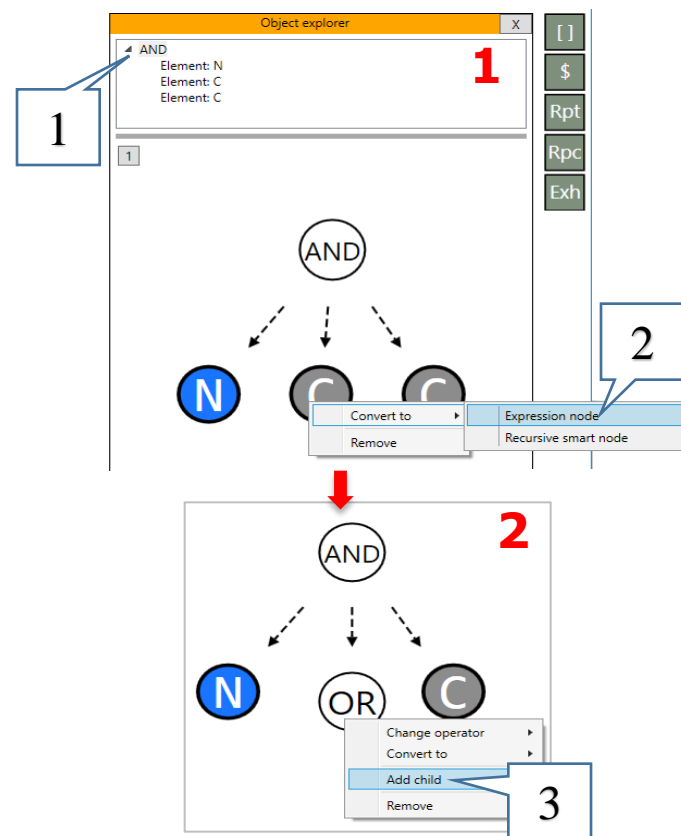
Workflow process

Structures search

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



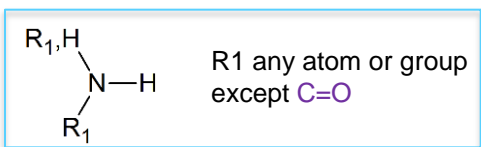
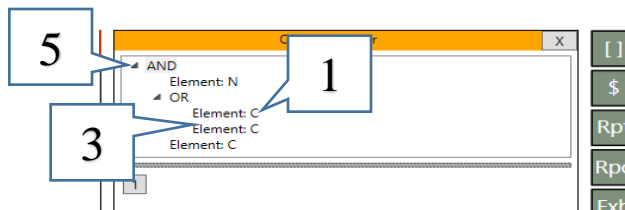
1. Go back to the "AND";
2. Right click over the second query and select **Convert to > Expression node**;
3. Right click over the **OR** and select **Add child**. Add two child queries.



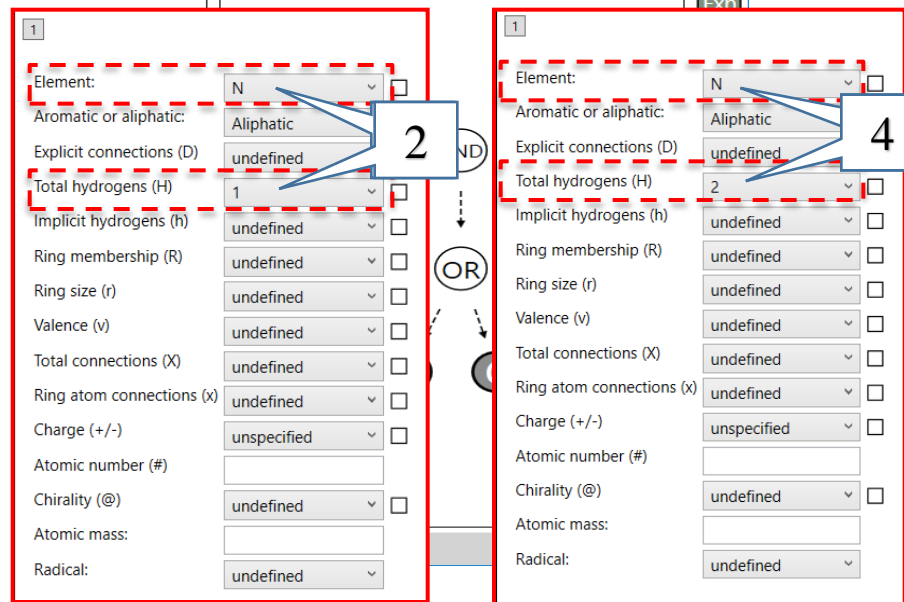
Workflow process

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.

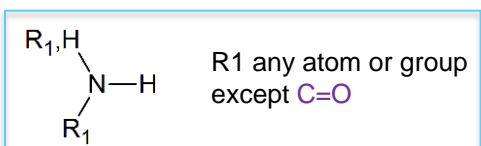


Property	Primary Amine (H=1)	Secondary Amine (H=2)
Element	N	N
Aromatic or aliphatic	Aliphatic	Aliphatic
Explicit connections (D)	undefined	undefined
Total hydrogens (H)	1	2
Implicit hydrogens (h)	undefined	undefined
Ring membership (R)	undefined	undefined
Ring size (r)	undefined	undefined
Valence (v)	undefined	undefined
Total connections (X)	undefined	undefined
Ring atom connections (x)	undefined	undefined
Charge (+/-)	unspecified	unspecified
Atomic number (#)		
Chirality (@)	undefined	undefined
Atomic mass:		
Radical:	undefined	undefined

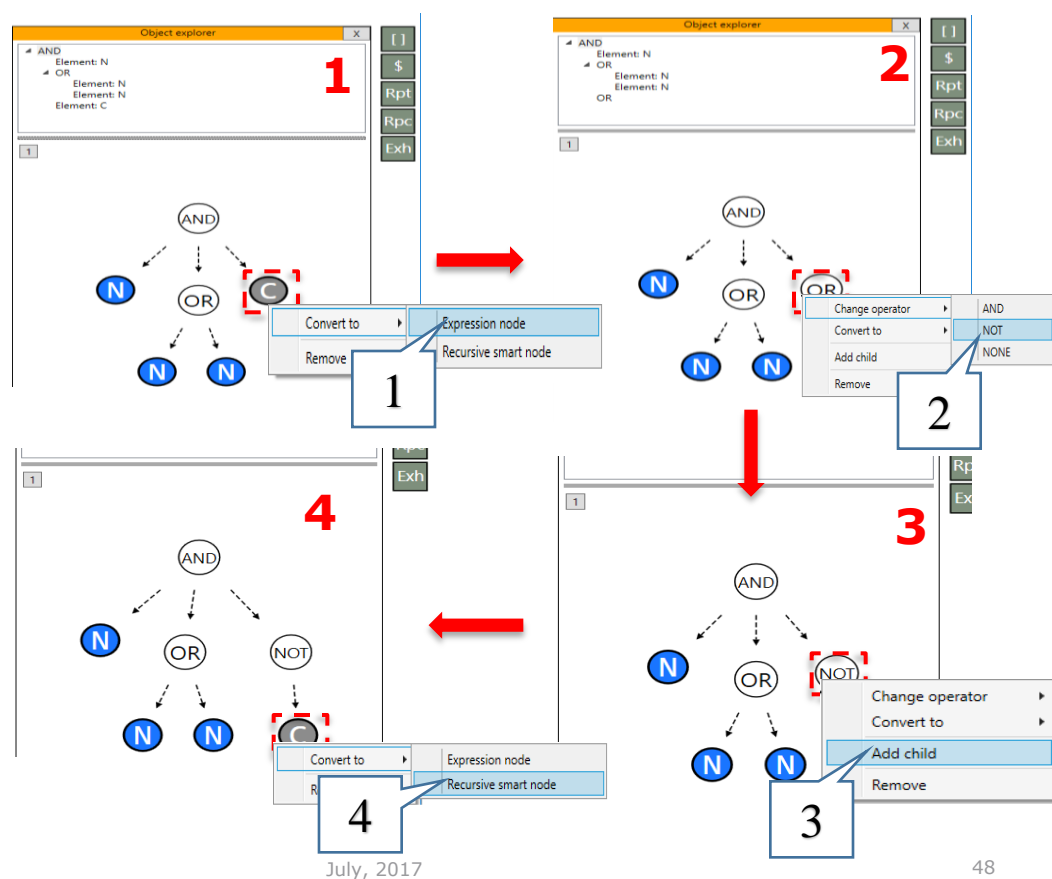
Workflow process

Structures search

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



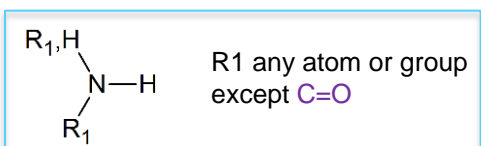
1. Right click on the last query and select **Convert to > Expression node**;
2. Right click on the new **OR** expression node and select **Change operator > NOT**;
3. Right click on the **NOT** query and select **Add child**;
4. Right click on the new child query and select **Convert to > Recursive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



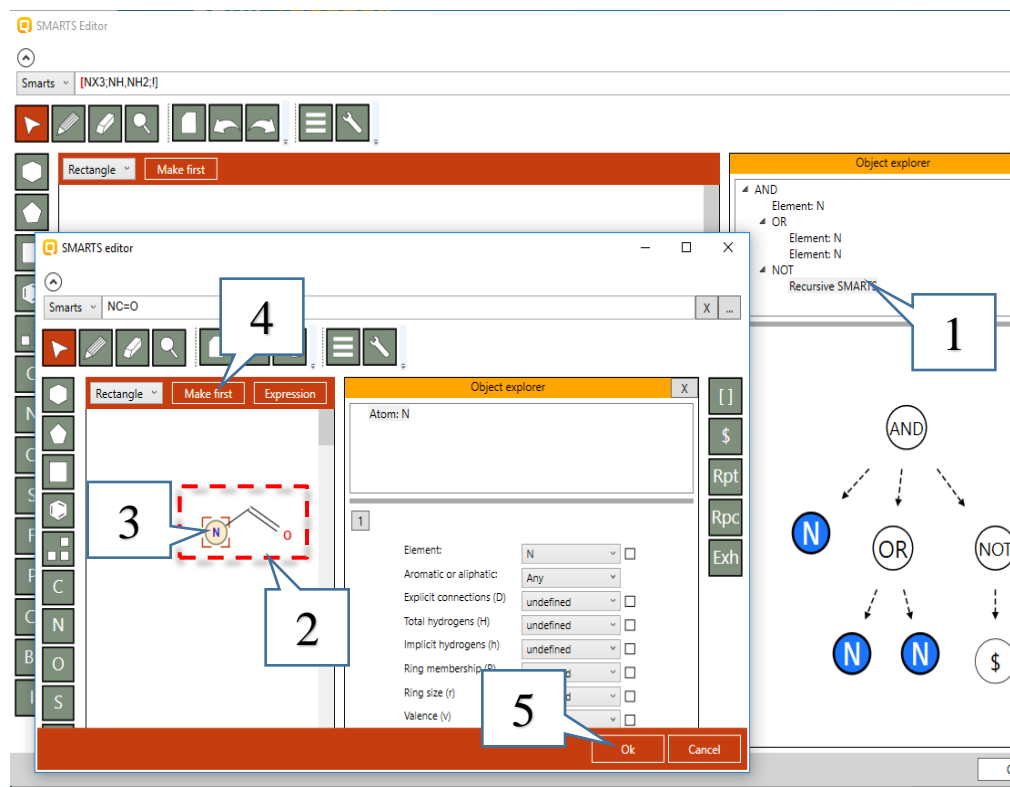
Workflow process

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**.



SMARTS Editor

Smarts: [NX3.NH.NH2;1]

Rectangle Make first

SMARTS editor

Smarts: NC=O

Rectangle Make first Expression

Object explorer

Atom: N

Element: N

Aromatic or aliphatic: Any

Explicit connections (D): undefined

Total hydrogens (H): undefined

Implicit hydrogens (h): undefined

Ring membership (R):

Ring size (r):

Valence (v):

Object explorer

- AND
 - Element: N
- OR
 - Element: N
 - Element: N
- NOT
 - Recursive SMARTS

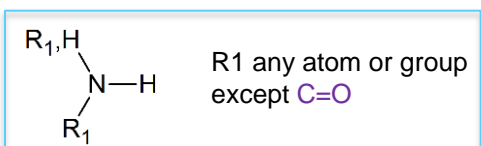
AND(N, OR(N, N), NOT(\$))

OK Cancel

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.

[NX3,NH,NH2,!(NC=O)]

Object explorer

- AND
 - Element: N
 - OR
 - Element: N
 - Element: N
 - NOT
 - Recursive SMARTS

1

OK Cancel

1

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 **OK** (2) to load all of them on the data matrix. You can see that all of the chemicals fulfill the criteria.

Select chemicals

Select All Unselect All Invert Selection Selected 346 of 346

1	CAS	18358-63-9	SMILES	CN1ccc(cc)X	CS Relation	Low	Name	methyl-4-meth...
2	CAS	2243-47-2	SMILES	Nc1ccc(c1)c1ccccc1	CS Relation	Low	Name	3-aminobiphenyl 3-aminobiphenyl_ (3-phenylam...
3	CAS	827-23-6	SMILES	Nc1c(Br)cc(Br)cc1[N+](=O)[O-]	CS Relation	Moderate	Name	2,4-dibromo-6-nitroaniline 2,4 Dibromo-6-nitroaniline 2,4 Dibromo-6-nitroanilinef
4	CAS	627-35-0	SMILES	CCNC	CS Relation	Moderate	Name	n-methylpropylamine

OK Cancel



QSAR TOOLBOX

Document Single Chemical Chemical List Search Target Endpoint

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

Documents

- Document 1
 - Search by SMARTS
 - Search by SMARTS
 - Search by SMARTS
 - Search by SMARTS

Filter endpoint tree...

Structure

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

Nc1c(Br)cc(Cl)cc1Br

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.