

# Virtual Reaction Design for Chemists

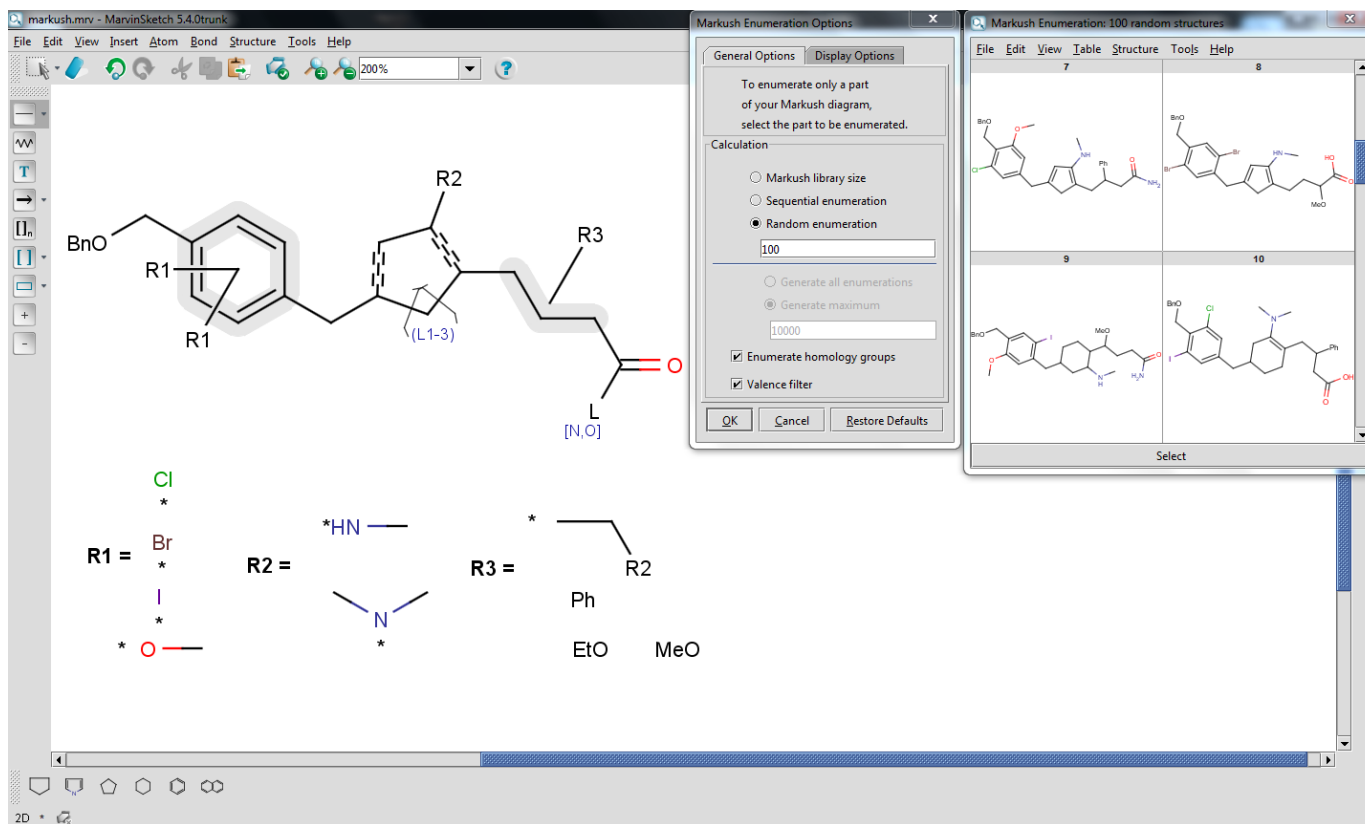
György Pirok



**ChemAxon**  
*Solutions for Cheminformatics*

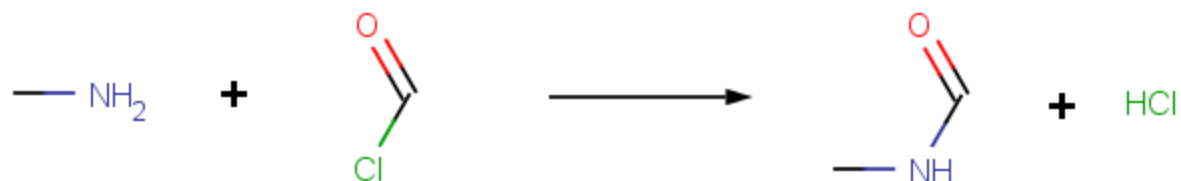
# Compound enumeration

Some enumeration tools like ChemAxon's Markush Enumeration plugin generate molecule libraries from a scaffold containing variable ligands. Functional groups are not transformed, synthesis is not in focus.

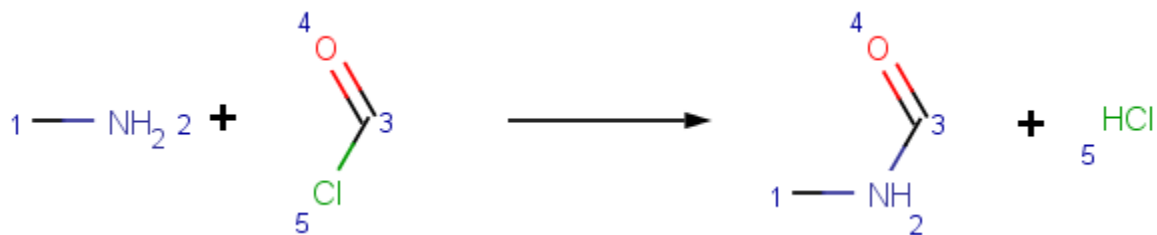


# Generic reactions

Virtual reaction tools transform reactants to products according to generic transformation schemes.

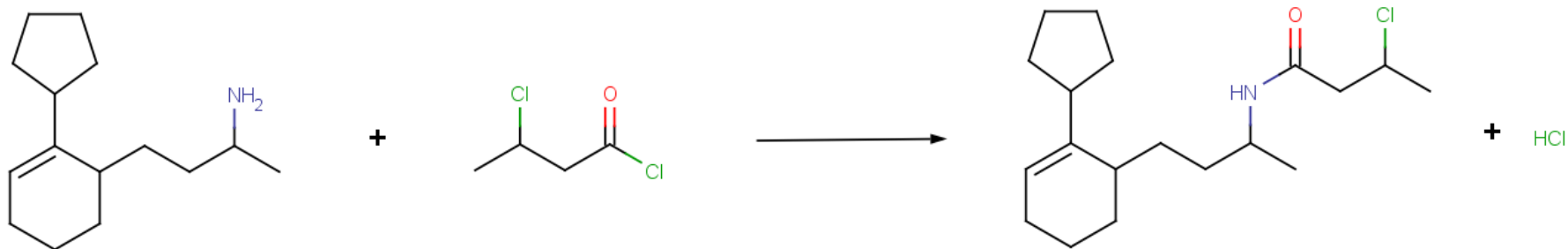
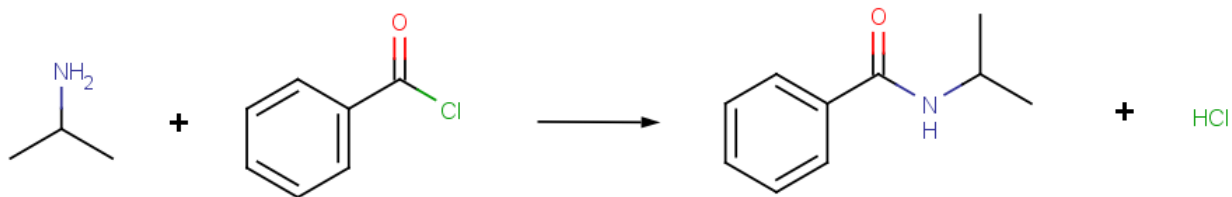


In these schemes, map numbers identify the corresponding atoms on the two sides of the reaction arrow unambiguously. If no maps are assigned, automatic mapping is performed.



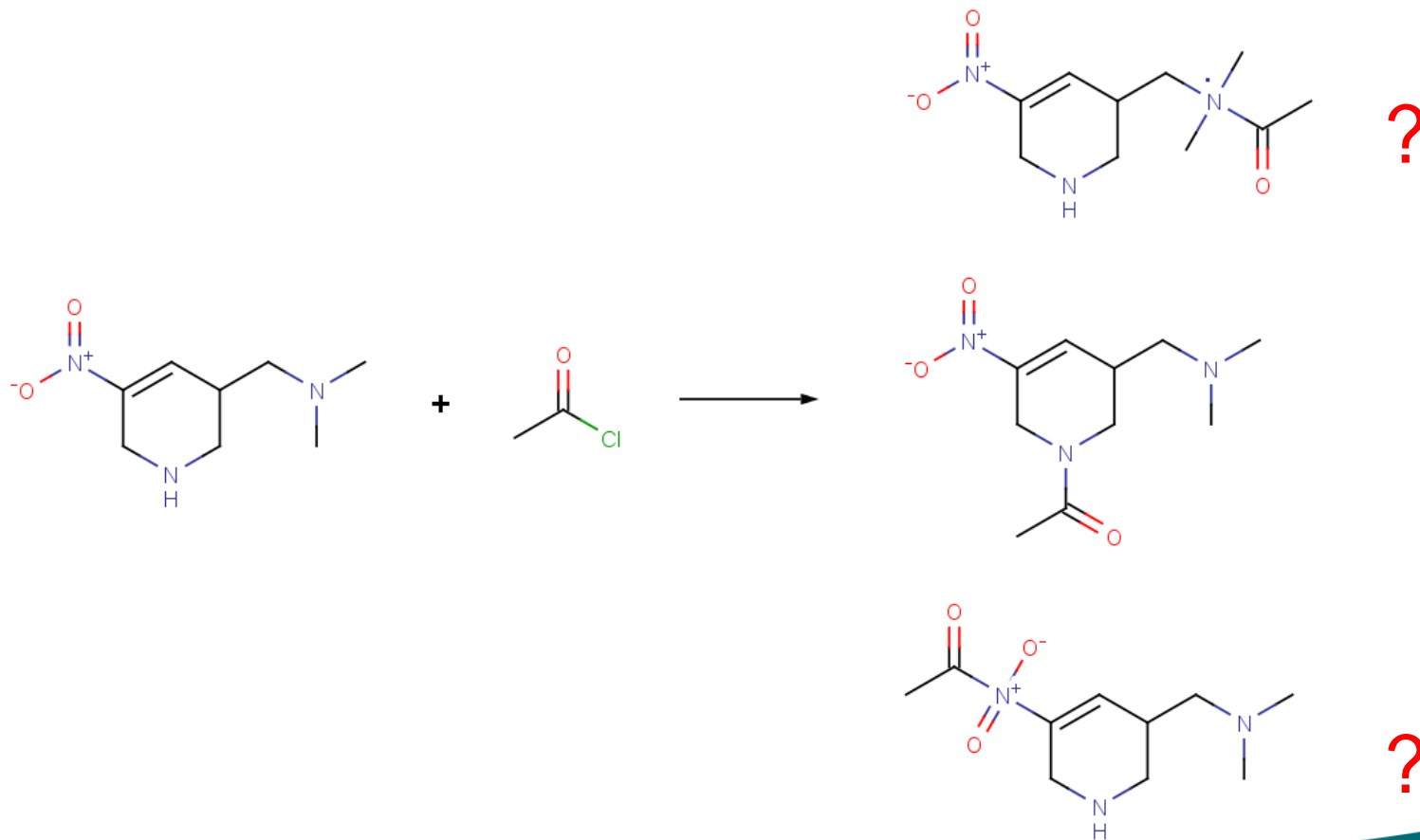
# Virtual reactions

In case of simple “click” reactions and carefully preselected input molecules the products can be synthetically feasible.



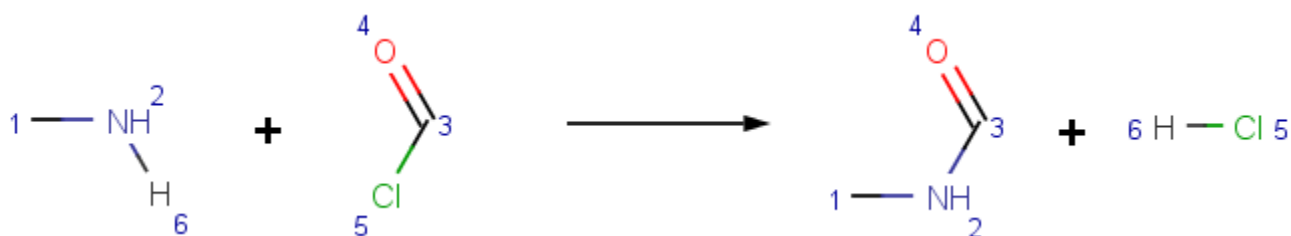
# Typical problems I.

Some problems may arise, however, even in the most basic cases. Depending on the reaction engine, various invalid molecules can appear in the output.

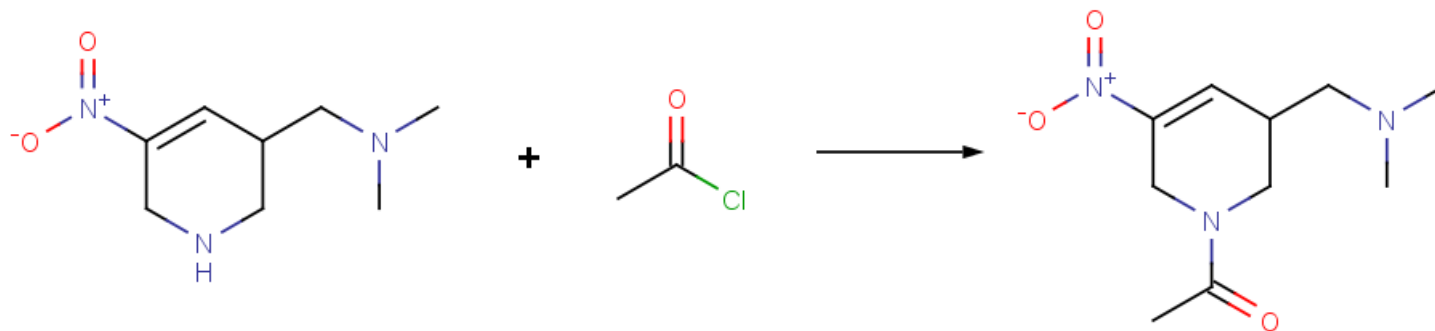


# Avoiding structural errors

Each reactant of the generic reaction scheme is used as substructure query, so never forget to specify the required hydrogens explicitly.



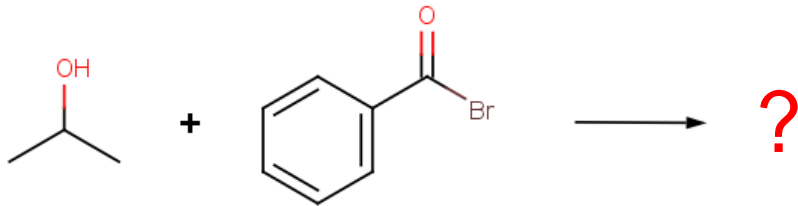
Most of the valence errors disappear after adding the “moving” hydrogens to the generic reaction scheme.



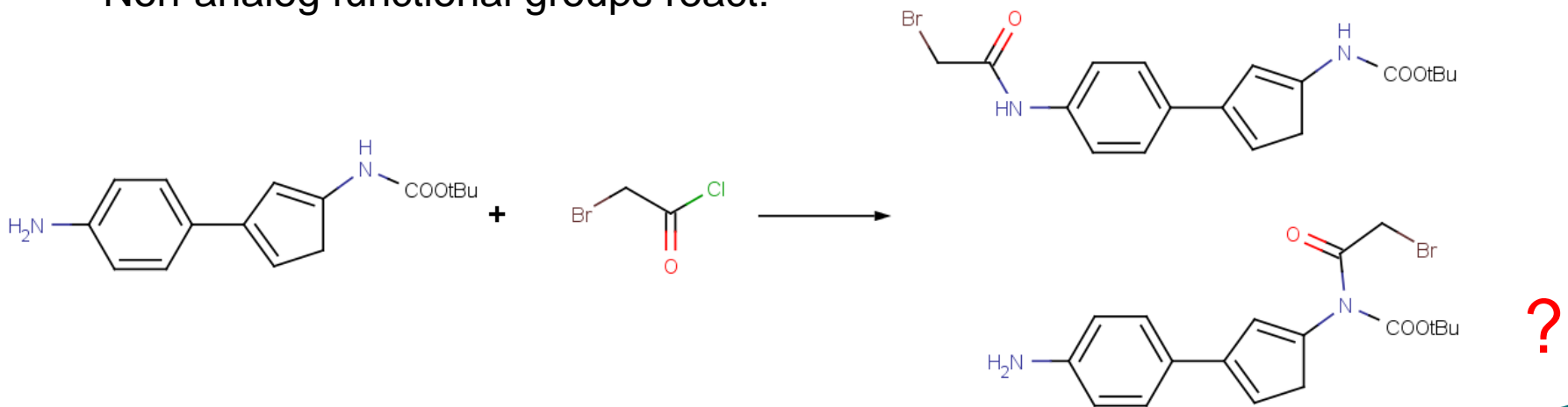
# Typical problems II.

Further issues can be identified even if the resulting compounds are chemically correct.

Analog functional groups do not react:

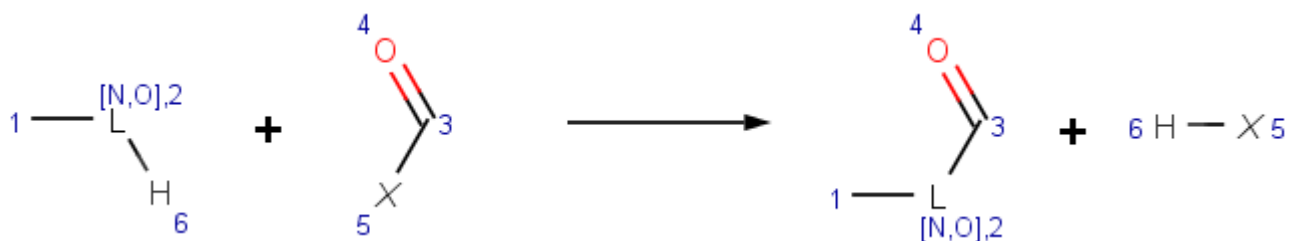


Non-analog functional groups react:

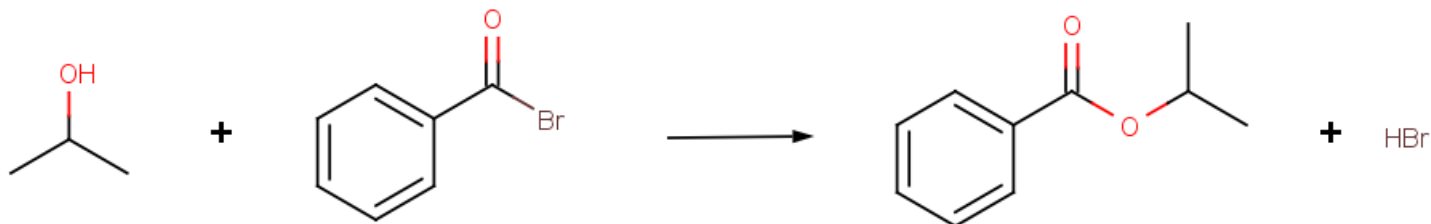


# Making reactions more generic

List atoms and generic atoms can help to extend the reaction scheme to support additional nucleophile functionalities.

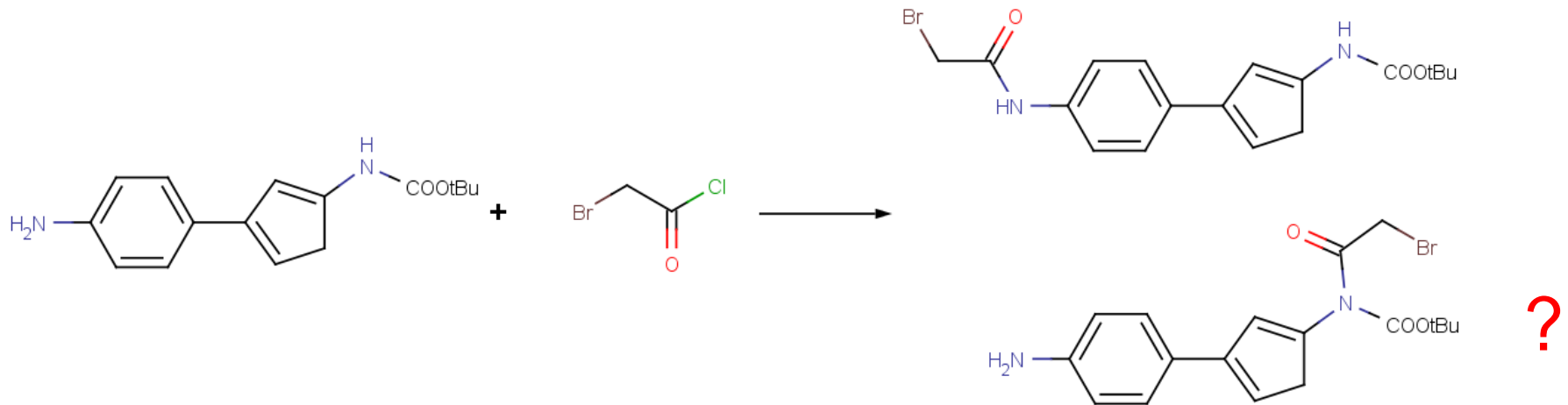


The same reaction scheme will acylate both amines and alcohols with any acyl halides.



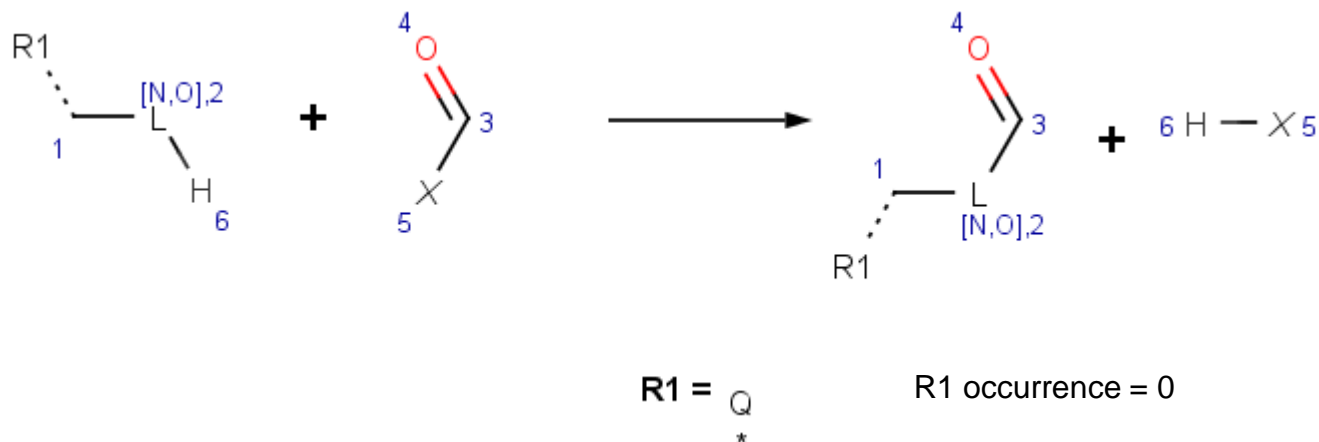
# Typical problems III.

A more difficult problem, however, is that often non-analog functional groups react.



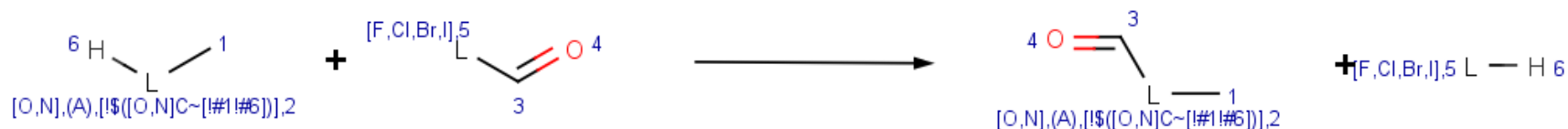
# Making generic reactions more specific I.

To restrict the reaction to amines and alcohols we need to specify somehow, that the hetero group is neither an amide nitrogen nor the OH of a carboxylic acid functionality. One option is to use R-groups with occurrence set to zero, but it is not too intuitive. We indicate something on the structure that should not be there.



# Making generic reactions more specific II.

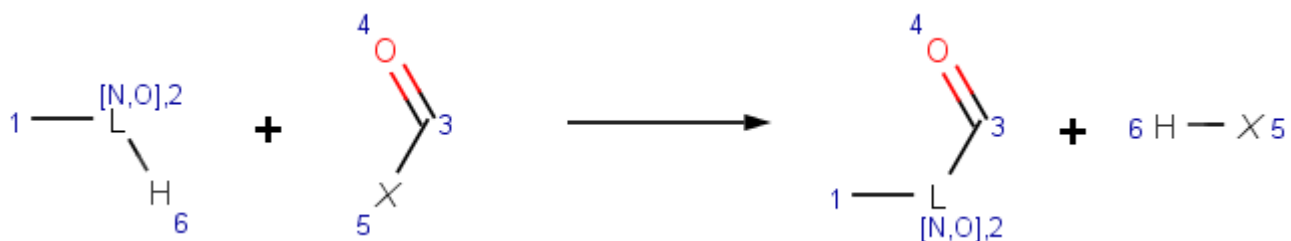
SMARTS expressions provide a very flexible but hard to understand way to restrict the structural neighborhood of the reaction sites. We can specify that our reaction site is a nitrogen or oxygen NOT connected to a carbon that is connected to another hetero atom.



```
$ react -r
'[H:6][O,N;!$([O,N]C~[!#1!#6]):2][#6:1].[F,Cl,Br,I:5][C:3]=[O:4]>>[#6:1][O,N;!$([O,N]C~[!#1!#6]):2][C:3]=[O:4].[F,Cl,Br,I:5][H:6]'
'CC(C)(C)OC(=O)NC1=CC(=CC1)C1=CC=C(N)C=C1'
'ClC(=O)CBr'
CC(C)(C)OC(=O)NC1=CC(=CC1)C1=CC=C(NC(=O)CBr)C=C1
Cl
```

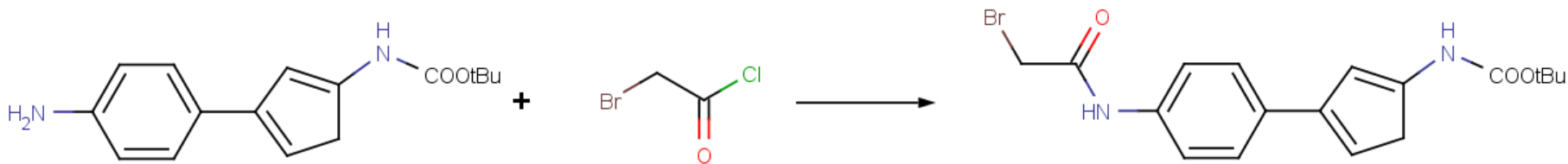
# Making generic reactions more specific III.

The chemical names of the functional groups provide a more understandable way for the precise specification of the reaction sites.



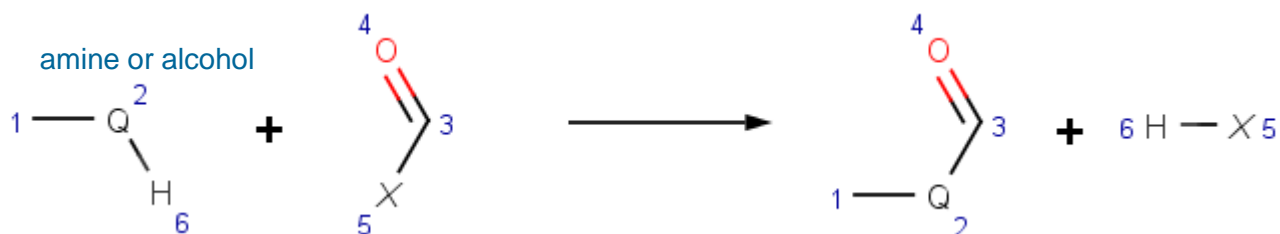
**REACTIVITY:** `match(ratom(2), amine, 2)` or `match(ratom(2), alcohol, 2)`

Functional groups can be referred by names in Chemical Terms expressions within the reactivity rule of the reaction. The rules like reactivity are interpreted by ChemAxon's Reactor engine.



# Making generic reactions more specific IV.

Future versions of the ChemAxon toolkit might support functional group names inlined as a simple and intuitive solution.



# Reactor

Reactor is ChemAxon's virtual reaction engine for the *selective* transformation of starting compounds. Reactor or the reactions do not know chemistry. But chemists can create reaction schemes that will produce synthetically feasible products. Quick demo...

The screenshot displays the Reactor 5.3.2 application window. On the left is a sidebar with navigation options: "Select Reaction", "Specify Reactants", "Set Reactor Options", and "Run Reactor". The main area is titled "Select Reaction" and contains a list of 12 reaction types, with "66. Friedel-Crafts acylation" selected. Below the list is a chemical reaction scheme showing benzene reacting with acetyl chloride to form acetophenone and HCl. The interface includes a "Reaction File" path, "Open", "New", and "Edit" buttons, and a bottom navigation bar with "Back", "Next", "Finish", and "Cancel" buttons.

Reactor 5.3.2

Select Reaction  
Select a reaction from the list or open a reaction file.

Reaction File: C:\ChemAxon\Sources\chem\chemaxon\_reaction\_library.mrv [Open] [New] [Edit]

- 57. Eschenmoser methylation (Mannich like reaction)
- 58. Eschweiler-Clarke amine methylation
- 59. Feist-Benary furan synthesis
- 60. Finnegan tetrazole synthesis
- 61. Fischer indole synthesis
- 62. Fischer oxazole synthesis from cyanohydrins
- 63. Fischer oxazole synthesis from alpha-hydroxyamides
- 64. Forster diazo synthesis
- 65. Freund-Gustavson cyclopropane synthesis
- 66. Friedel-Crafts acylation
- 67. Friedlander quinine synthesis
- 68. Fritsch-Buttenberg-Wiechell acetylene synthesis (rearrangement)
- 69. Fukuyama coupling
- 70. Gabriel primary amine synthesis
- 71. Gewald 2-aminofurane synthesis
- 72. Gewald 2-aminopyrrole synthesis
- 73. Gewald 2-aminothiophene synthesis
- 74. Goldberg coupling (Ullmann type reaction)
- 75. Grignard compound addition to carbonyl compounds
- 76. Grignard reagent formation
- 77. Guaresky-Thorpe pyridone synthesis

Reverse direction

c1ccccc1.CC(=O)Cl>>CC(=O)c1ccccc1.Cl

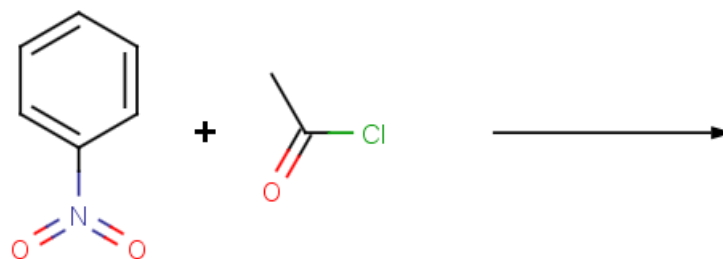
Online Help  
License Manager  
Hide Sidebar

Press 'Next' to specify reactants

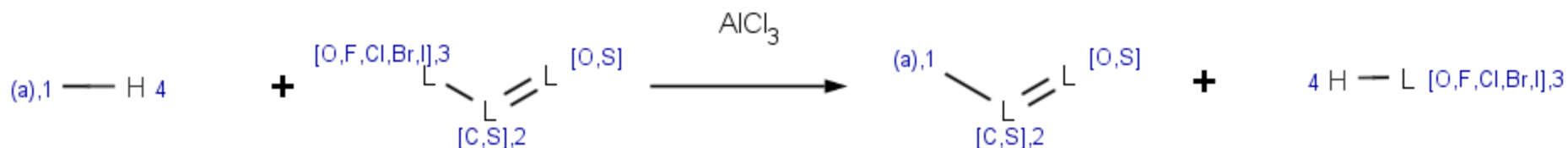
< Back Next > Finish Cancel

# Non-structural conditions - Reactivity

Some functional groups like the nitro group deactivate the aromatic ring for Friedel-Crafts reaction, so deactivated rings cannot be acylated in this way.



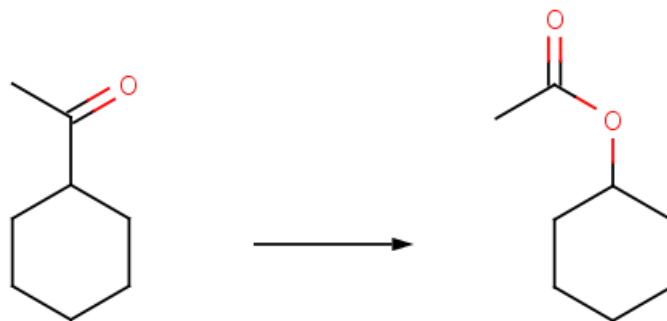
Neither R-groups, nor SMARTS offer physicochemical property conditions, but the reaction rules do. Property calculations of the reactivity rule allow the consideration of activation/deactivation effects.



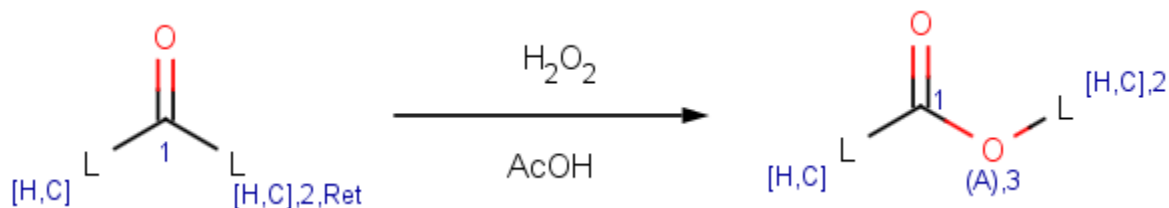
**REACTIVITY:** `charge` (ratom(1), "aromaticsystem") <= -0.2

# Non-structural conditions - Selectivity

The Baeyer-Villiger oxidation selectively takes place between the carbonyl and the more positively charged carbon atom.



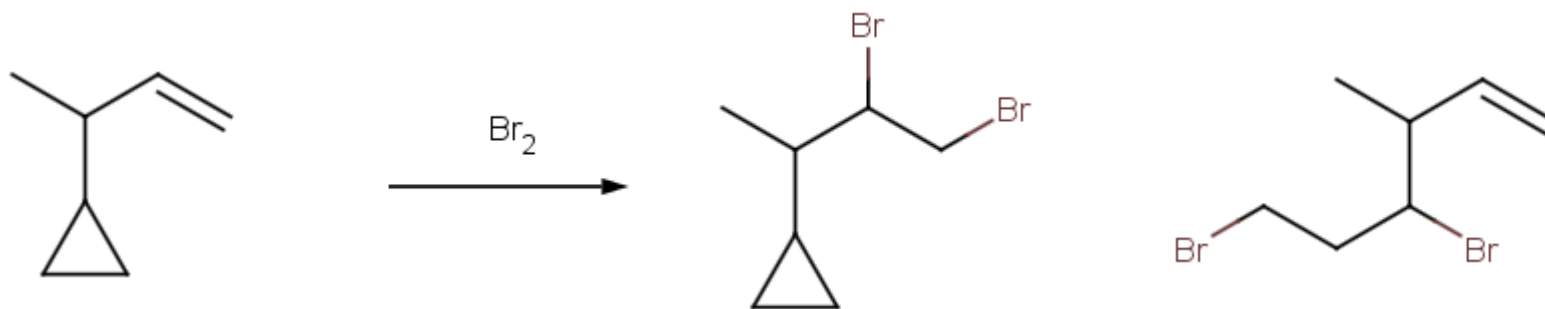
The active sites are sorted by values calculated from selectivity rule expression. The ones on the top lead to the main products.



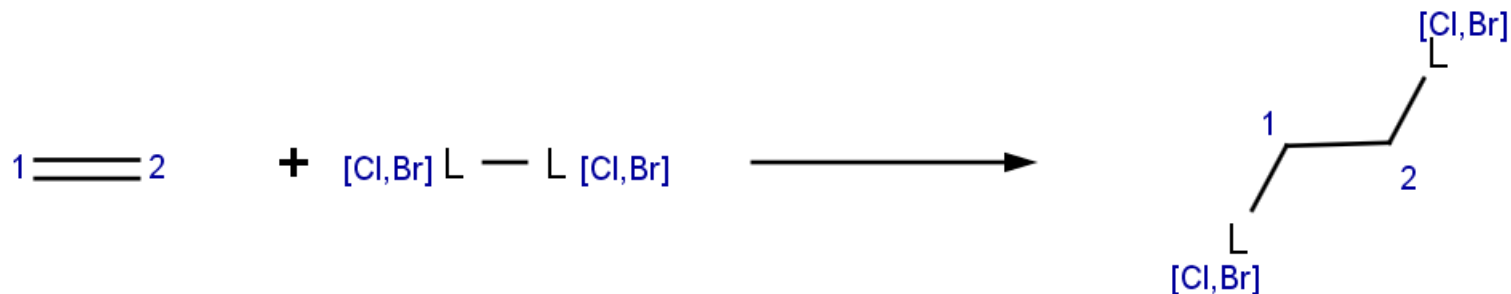
**SELECTIVITY:** `charge (ratom(2), "sigma")`

# The exclude rule

Cyclopropyl ligands cause ring opening side reactions during the halogenation of olefins.

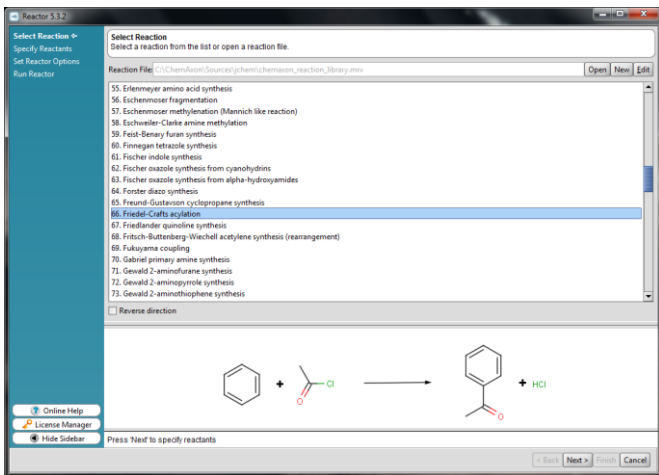


Some reactants contain sensitive functionalities causing side reactions. These input molecules can be identified by the exclude rule and can be excluded from the reaction.

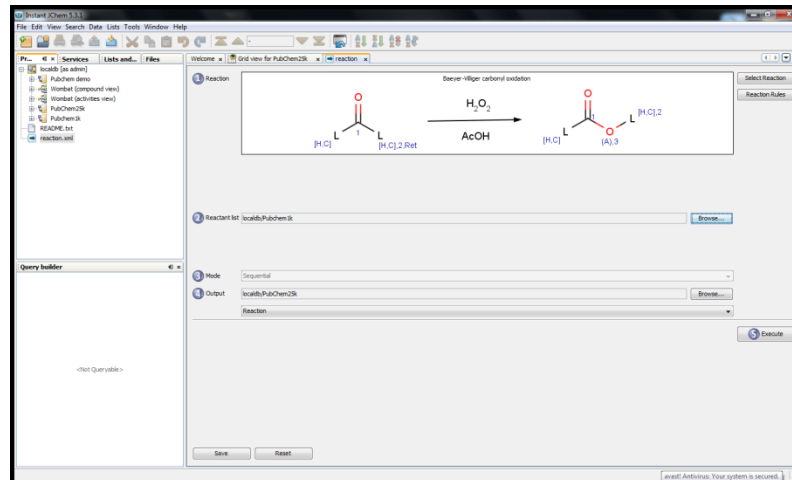


**EXCLUDE:** `smallestRingSize`(reactant(0)) = 3

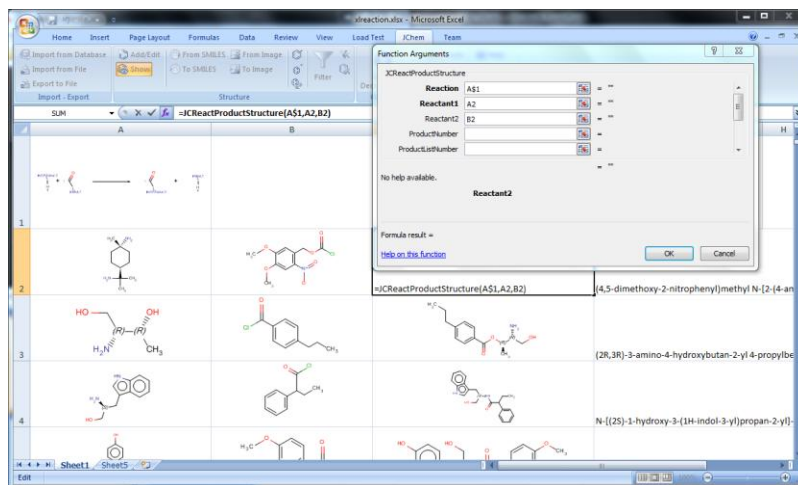
# Reactor Interfaces I.



Reactor Wizard

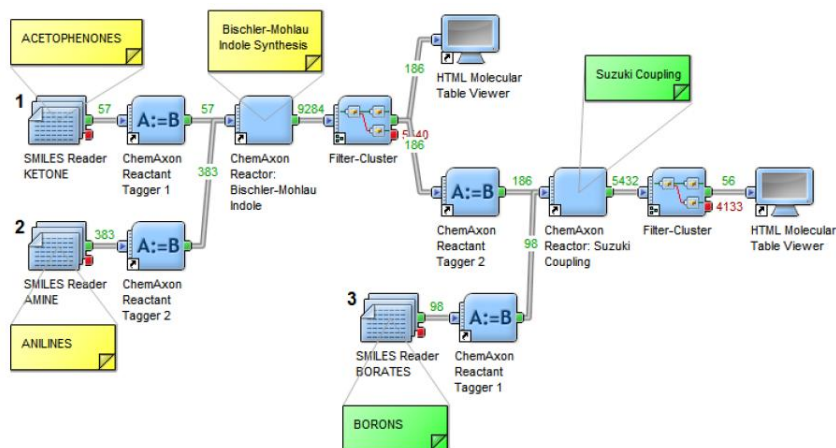


Instant JChem

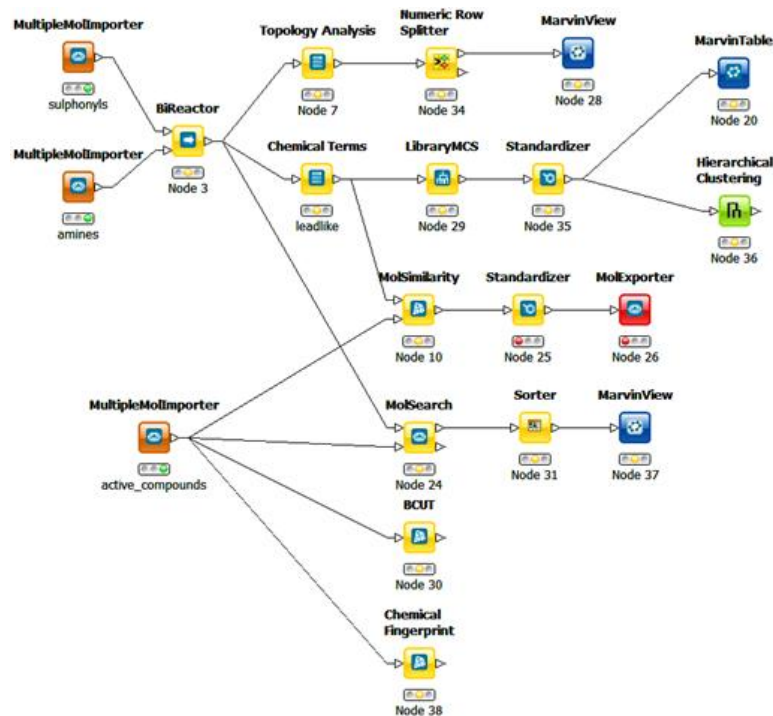


JChem for Excel

# Reactor Interfaces II.



Pipeline Pilot



Knime

```
<react xmlns="http://webservice.jchem.chemaxon">
  <reactionString>
    [C1:3][C:1]=[O:2].[H:5][N:4][#6:6]>>[#6:6][N:4][C:1]=[O:2].[C1:3][H:5]
  </reactionString>
  <reactants>ClC(=O)c1ccccc1</reactants>
  <reactants>CC(C)N</reactants>
  <options>outFormat:smarts</options>
</react>
```

JChem Web Services

# Reactor Interfaces III.

```
SELECT
jc_react(' [C:1] (=O:2) [C1:3] . [H:99] [N:4] ([H:100]) [C:0]>>[C:1] (=O:2) [N:4] ([H:100]
) [C:0] . [C1:3] [H:99]', 'ClC(=O)c1ccccc1.CC(C)N', '') products FROM DUAL;
```

## JChem Cartridge

```
$ react -r
' [H:6] [O,N;!$([O,N]C~[!#1!#6]):2] [#6:1] . [F,Cl,Br,I:5] [C:3]=[O:4]>>[#6:1] [O,N;!$([O
,N]C~[!#1!#6]):2] [C:3]=[O:4] . [F,Cl,Br,I:5] [H:6] '
'CC(C)(C)OC(=O)NC1=CC(=CC1)C1=CC=C(N)C=C1' 'ClC(=O)CBr'
CC(C)(C)OC(=O)NC1=CC(=CC1)C1=CC=C(NC(=O)CBr)C=C1
Cl
```

## react command line tool

```
// create Reactor
Reactor reactor = new Reactor();

// set the reaction ('rxmol' is the reaction molecule)
// reaction rules are read from RDF/MRV tags
reactor.setReaction(rxmol);

// set the reactants
reactor.setReactants(reactants);

// get the results
Molecule[] products;
while ((products = reactor.react()) != null) {
    // export result molecules
    for (Molecule product : products) {
        exporter.write(product);
    }
}
```

## Java and .NET API

# Reactor Features

---

- Chemically feasible reactions
- Support for quick SMIRKS, RXN based enumerations
- Not a black box, reactions can be modified, new reactions can be designed (Reaction Editor)
- High performance
- All property calculations are included for the use within rules
- ChemAxon's reaction library is included
- Available for developers as Java, .NET, web service, cartridge functions
- Available for users in Wizard, JChem4XL, Instant JChem

# Limitations

---

- Some calculations need improvements
- Some calculations can be slow for combichem reactions (i.e. those based on conformers)
- Some calculations are not available yet (i.e. HOMO-LUMO)
- Multistep reaction schemes are not supported yet, so no calculations are directly available for transition states or intermediates
- Sometimes we have no idea what properties influence the outcome of a reaction
- Some special but important stereo selectivities like prochirality are not interpreted by cheminformatics systems

# Future directions

---

- New execution mode in Reactor wizard select products manually
- Improving current property calculations relevant to organic reactions
- Developing new, fast algorithms for the calculation of computationally demanding properties (i.e. steric hindrance)
- Extending the current set of reaction stereo features with new ones like prochirality
- Reviewing the current reactions in the ChemAxon reaction library
- Adding new, practically important reactions to the library in a systematic way
- Validating the reaction library with experimental data

# Acknowledgements

---



Nóra Máté



Zsolt Mohácsi



Jenő Varga



Attila Szabó



István Rábel



Anna Tomin