UNIVERSITÉ DE STRASBOURG

ChémInformatique STRASBOURG

Chimie de la Matière Complexe (CMC)

Institute of Organic Chemistry NAS of Ukraine

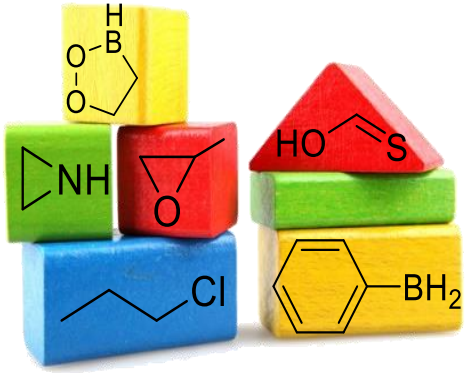
Enamine

SynthI: a New Toolkit for Synthon-Based Library Design and Building Blocks Analysis

Yuliana Zabolotna

D.M.Volochnyuk, S. V.Ryabukhin, K. Gavrylenko, D. Horvath, O. Klimchuk, O. Oksiuta, G. Marcou, A. Varnek

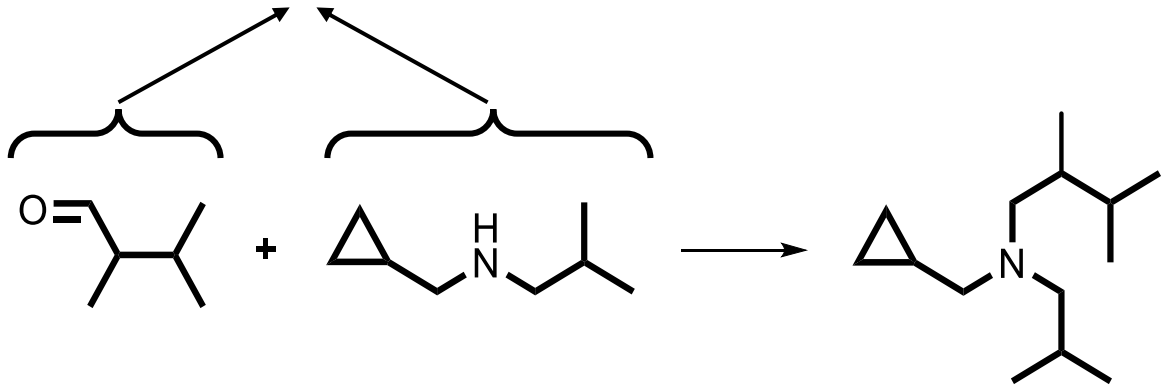
Building Blocks in Medicinal Chemistry



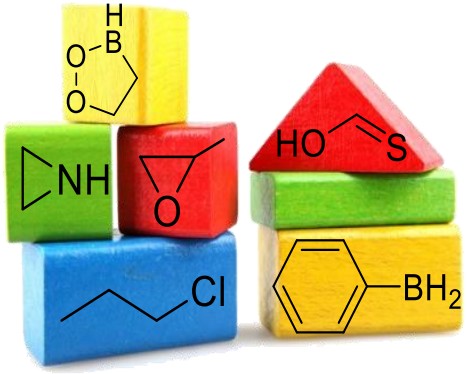
Building blocks (BB) – molecules that possess at least one reactive functional group and thus can be used as reagents for synthesis and as a building material for libraries design.

Building blocs (BBs) or reagents

Reaction product



Building Blocks in Medicinal Chemistry



Building blocks (BB) – molecules that possess at least one reactive functional group and thus can be used as reagents for synthesis and as a building material for libraries design.

Usage of the medicinally relevant BB can significantly improve the quality of the designed compounds

Up to date, there are only a few publications dedicated to the analysis of the BB used in medicinal chemistry

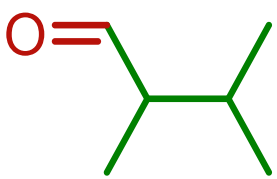
BBs analysis challenges

Synthetic chemists: "How?"

- functional groups that define BB's reactivity

This is an aldehyde which can participate in:

- imine formation
- alkylation
- C-C coupling with Li, Mg, Zn organics



Medicinal chemists: "What?"

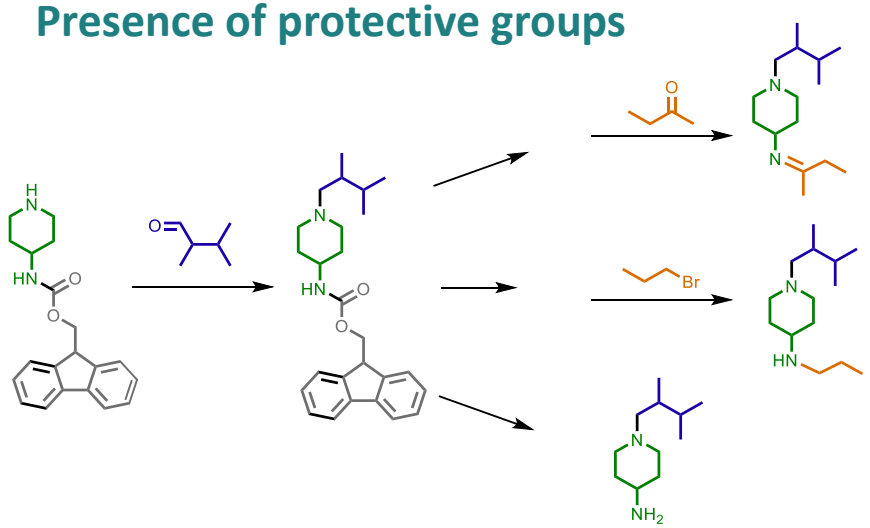
- structural fragments introduced to the final molecule

This is a reagent that can introduce 2,3-dimethylbutane moiety

Both aspects needs to be considered in BBs analysis!

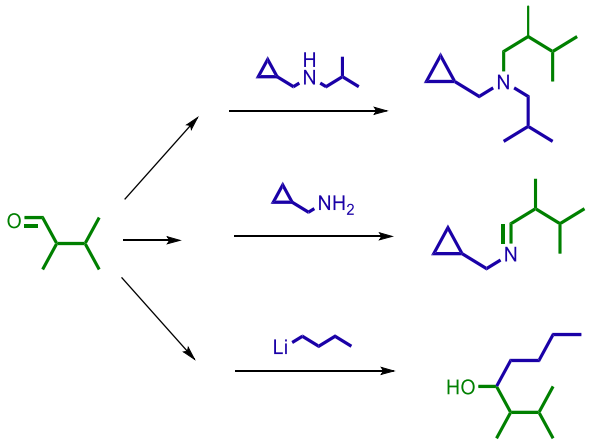
BBs analysis challenges

Presence of protective groups



- Functional groups in BB can be protected
- Protective groups can be larger than BB itself

Versatile reaction behavior



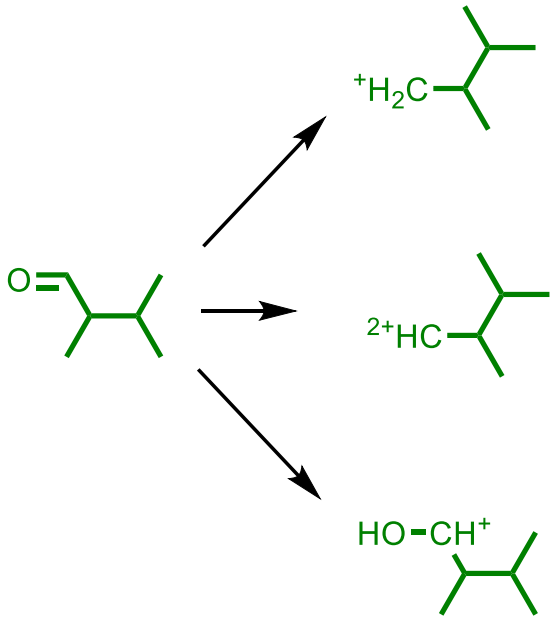
Depending on reaction partner the same BB can:

- behave differently
- contribute different fragments to the final molecule

Concept of Synthons

“Synthon” – is a structural increment, introduced by the BB into the reaction product

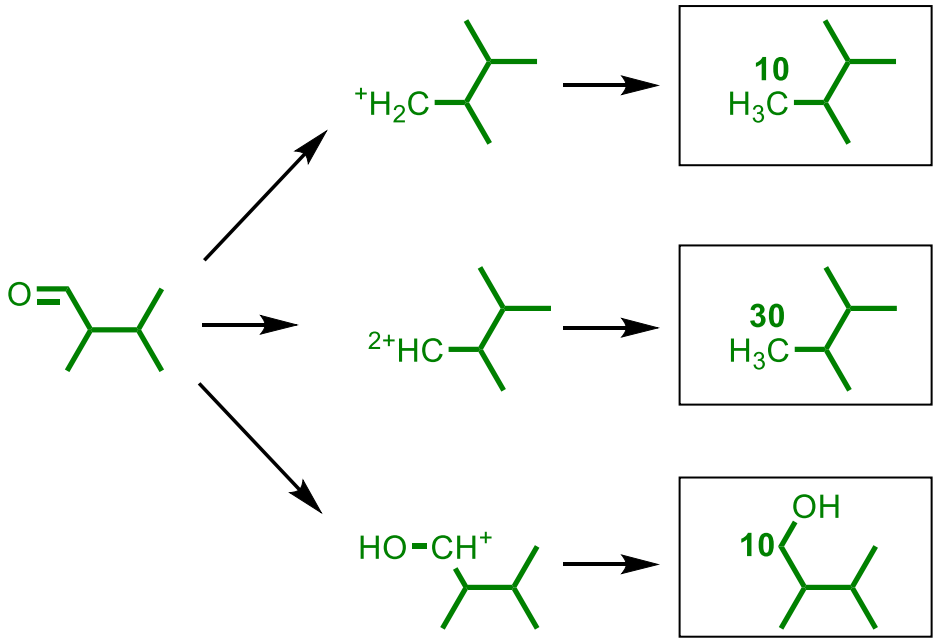
Synthon = BB – Leaving Group



Concept of Synthons

"Synthon" – is a structural increment, introduced by the BB into the reaction product

Synthon = BB – Leaving Group + Reaction Centre Label



[CH3:10]C(C)C(C)C

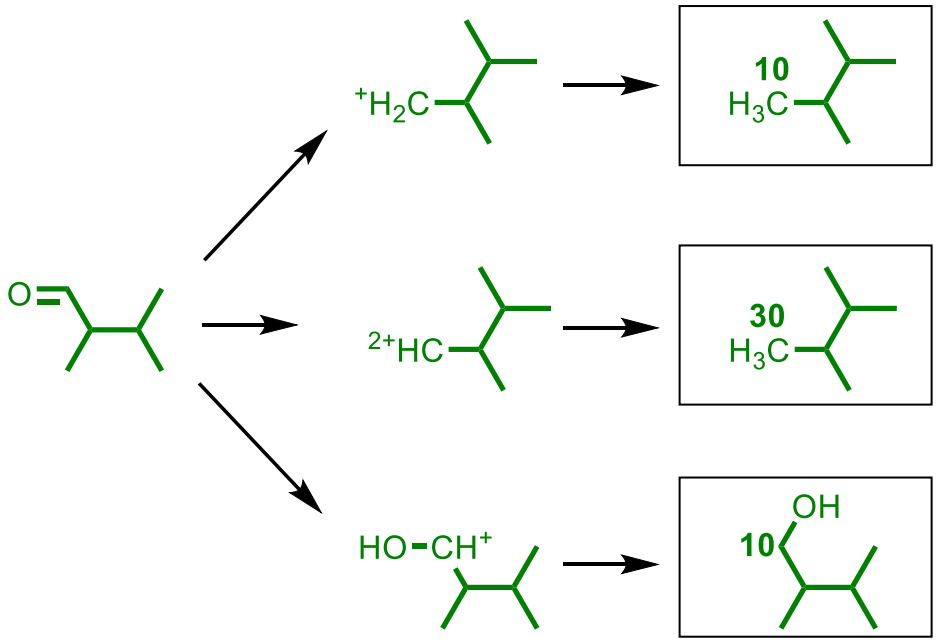
[CH3:30]C(C)C(C)C

CC(C)C(C)[CH2:10]O

Concept of Synthons

"Synthon" – is a structural increment, introduced by the BB into the reaction product

Synthon = BB – Leaving Group + Reaction Centre Label



- neutral fragments
- have valid valences

PhysChem properties of synthons can be calculated

Syntons can be mapped on GTM

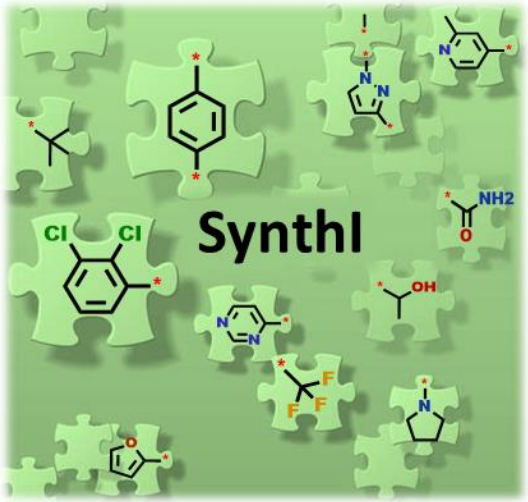
Synthons Interpreter (SynthI)

SynthI-Classification

SynthI-BB

BBs classification

Classifies and generates synthons for library of BBs



SynthI-Fragmentation

Fragments library of molecules into synthons using selected list of reactions

Knowledge-based generalized toolkit empowering library design and BBs analysis

SynthI-Enumeration

Enumerates compound library from user-defined synthons using selected reactions

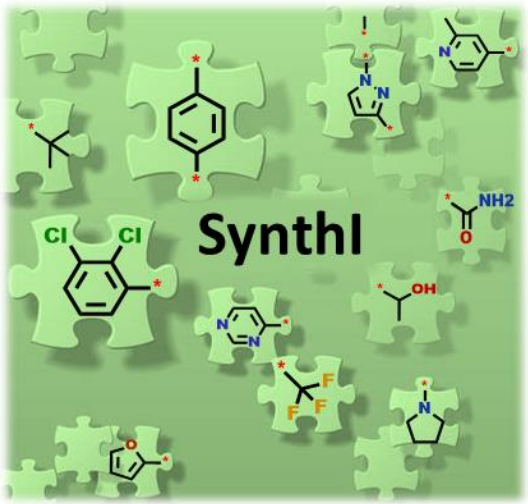
https://github.com/Laboratoire-de-Chemoinformatique/SynthI

Synthons Interpreter (SynthI)

SynthI-Classification

BBs classification

SynthI-BB

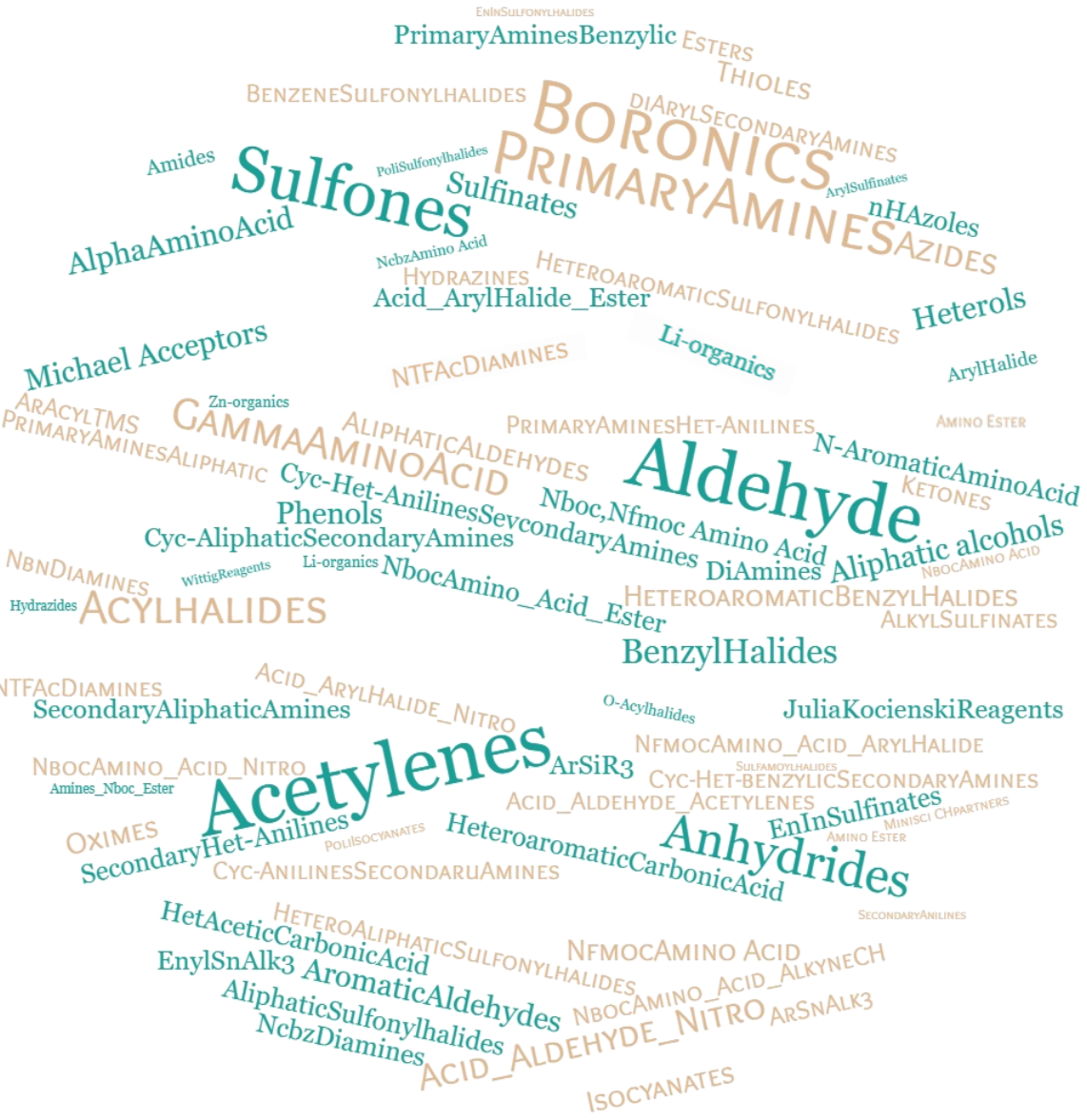


SynthI-Fragmentation

Knowledge-based generalized toolkit empowering library design for organic chemistry and BBs analysis

SynthI-Enumeration

Building Blocks Classification



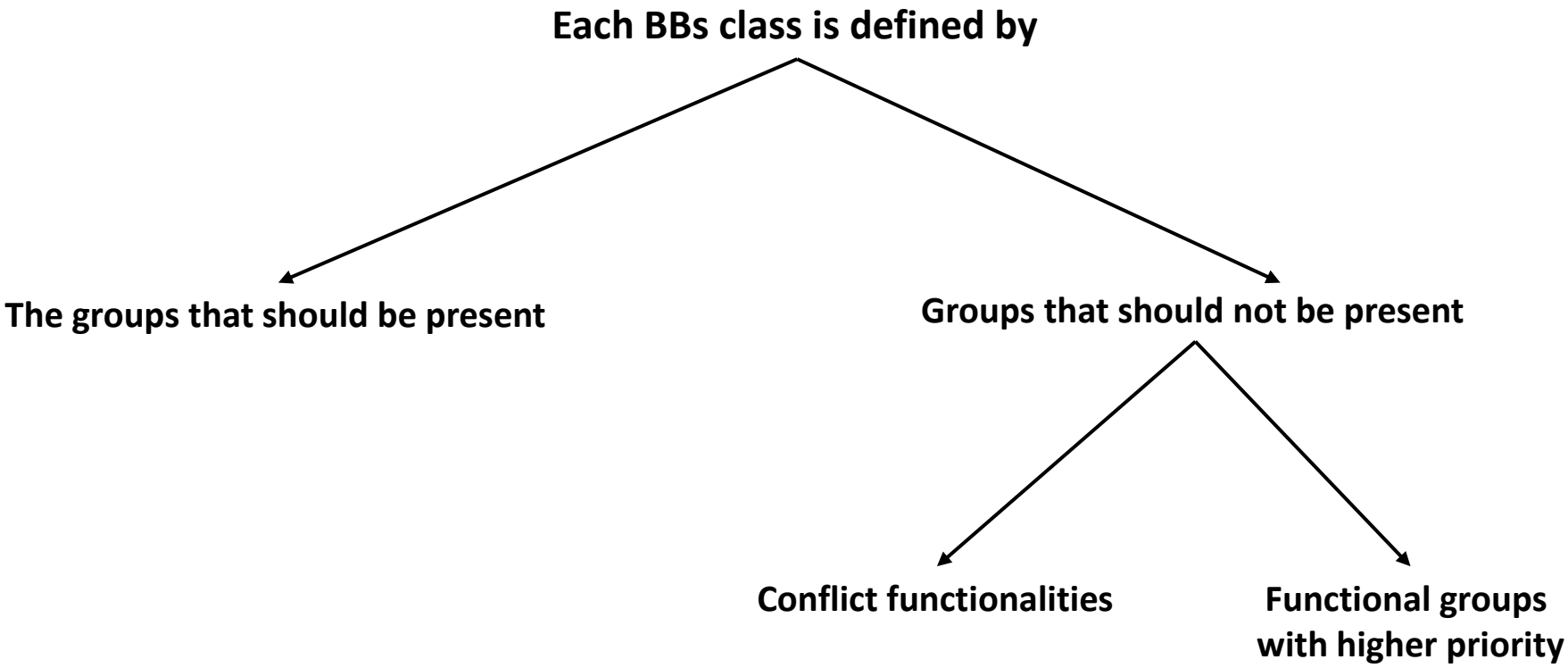
- >200 SMARTS specification for BBs classification
- ≈150 BB classes (100 mono-, 28 bi- and 19 trifunctional)

Institute of Organic Chemistry NAS of Ukraine

Enamine

Years of experience of synthetic chemists

Building Blocks Classification



Building Blocks Classification: case study

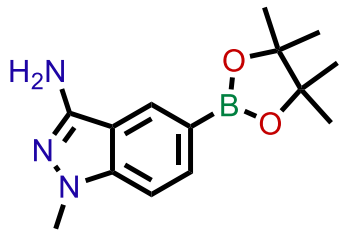
The molecule is not considered as *primary amine* if R contains:

I. Conflict functionalities

- Second amino group
- Hydroxylamines
- (Thio)Carbonyls
- Iso(thio)cyanates
- Metal-organics
- Acylators
- Alkylators
- Oxiranes
- Imines

II. FG with higher priority than NH_2

- Thiols
- Element Organics
- Boronics*

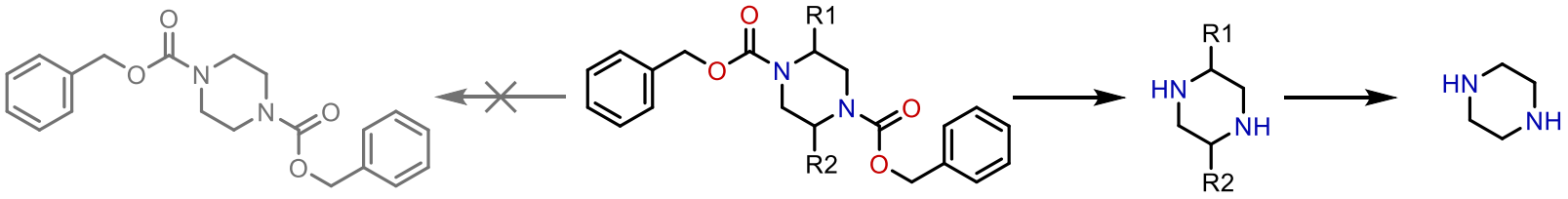


*A chemist would select this BB because of the boronic acid functionality (e.g. for Suzuki or Chan-Lam coupling) rather than for the amine group

*Goldberg, F.W., et al., Drug Discovery Today, 2015. **20**(1): p. 11-17.

Scaffold analysis of BBs

Scaffold generation for BBs with ring-containing protective groups



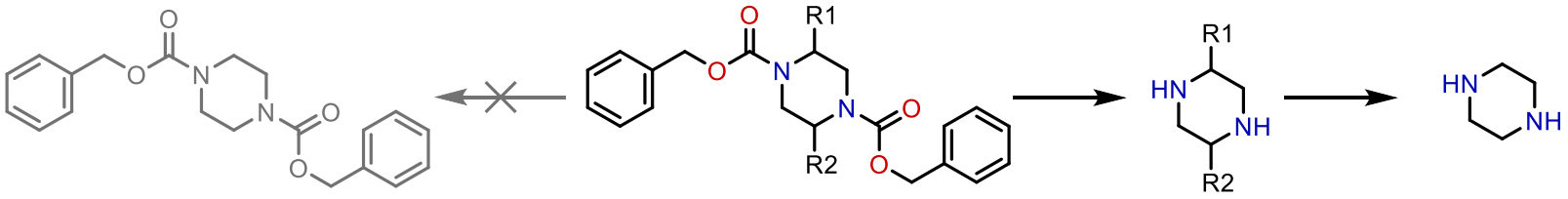
Incorrect Bemis-Murcko Scaffold

Protected BB

Correct Bemis-Murcko Scaffold

Scaffold analysis of BBs

Scaffold generation for BBs with ring-containing protective groups

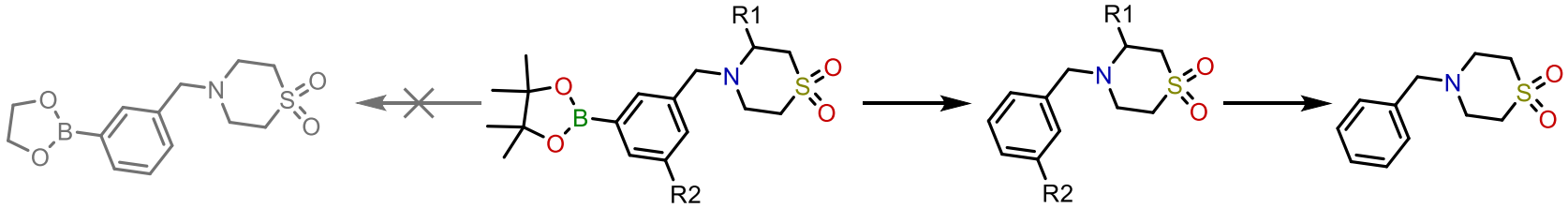


Incorrect Bemis-Murcko Scaffold

Protected BB

Correct Bemis-Murcko Scaffold

Scaffold generation for BBs with ring-containing leaving groups



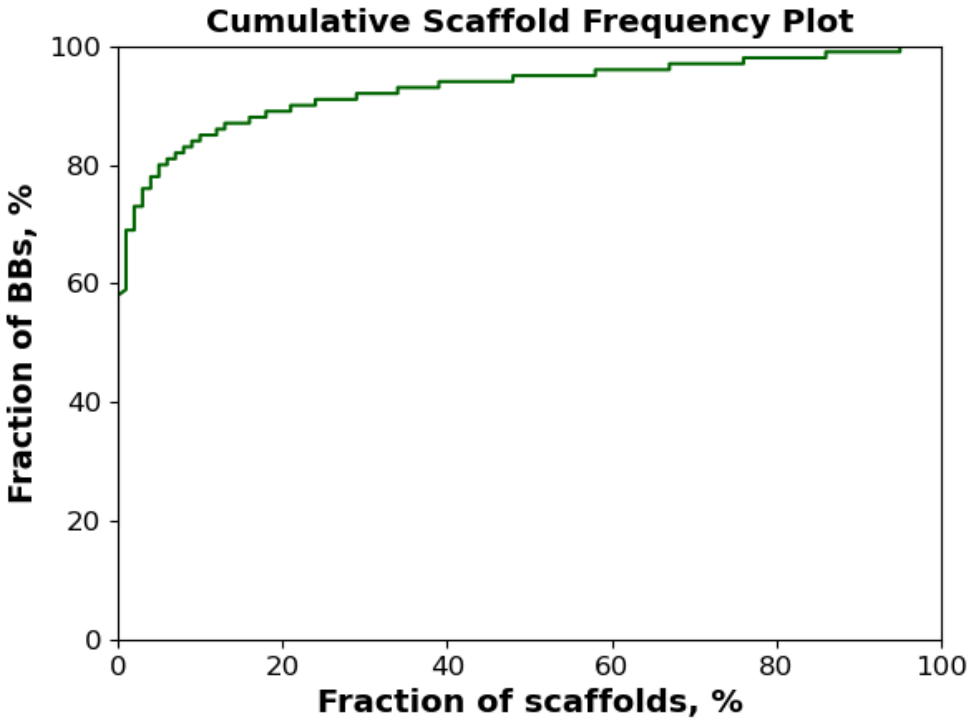
Incorrect Bemis-Murcko Scaffold

BB with cyclic leaving group

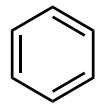
Correct Bemis-Murcko Scaffold

Scaffold analysis of BBs

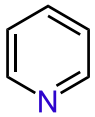
SynthI allows performing **scaffold analysis** including generation of the **scaffold frequency plots**



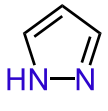
The most popular scaffolds



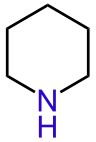
35 414 BBs



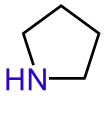
8 633 BBs



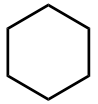
5 004 BBs



3 611 BBs



2 456 BBs



2 313 BBs

2 285 BBs

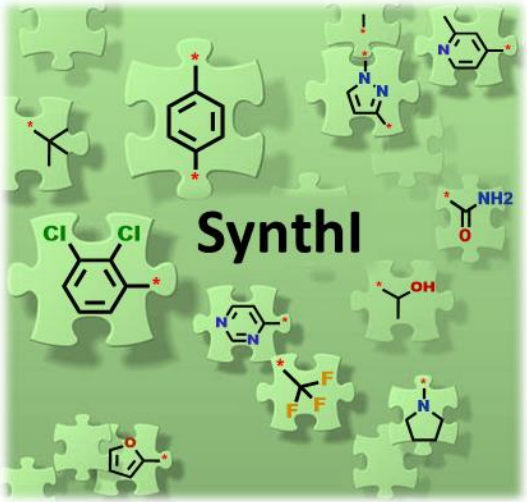
2 150 BBs

Synthons Interpreter (SynthI)

SynthI-Classification

SynthI-BB

Generates synthons for library of BBs



SynthI-Fragmentation

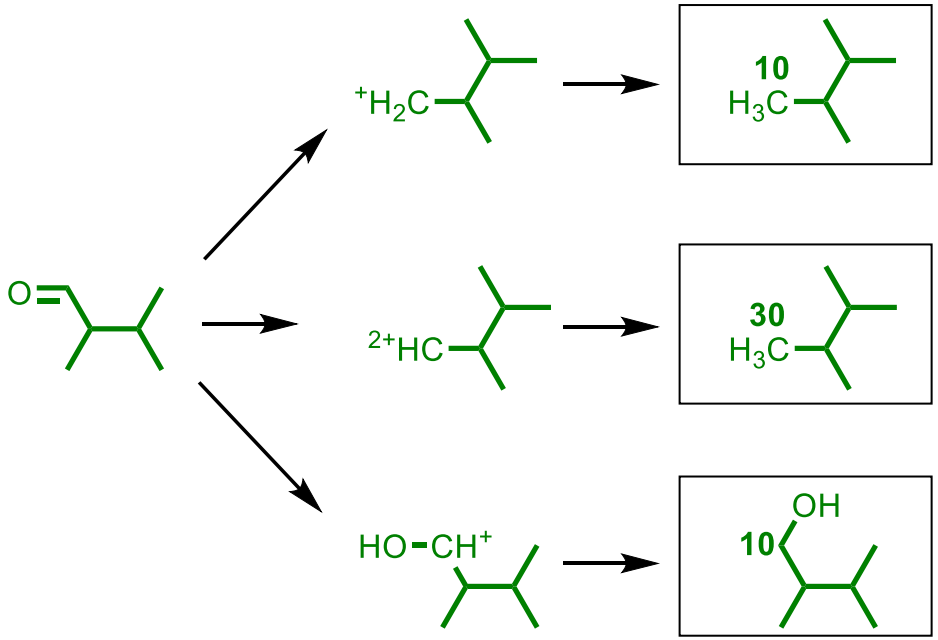
Knowledge-based generalized toolkit empowering library design for organic chemistry and BBs analysis

SynthI-Enumeration

SynthI-BBs

Generates the list of synthons for the given BB depending on the class assigned to this BB on the previous step

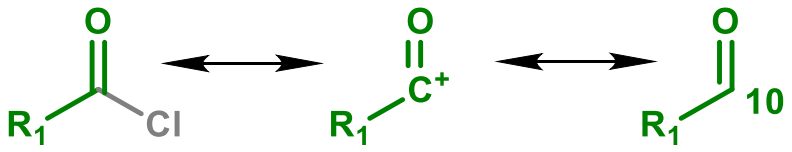
Synthon = BB – Leaving Group + Reaction Centre Label



9 types of labels

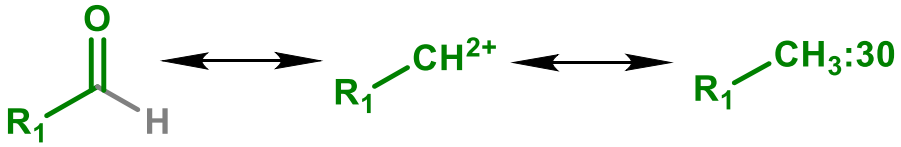
Synthons classes and labels system: electrophilic synthons

- **Electrophilic synthons**



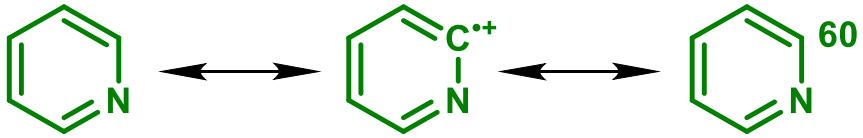
N-, O-acylation

- **Bivalent Electrophilic synthons**



Imines formation, olefination

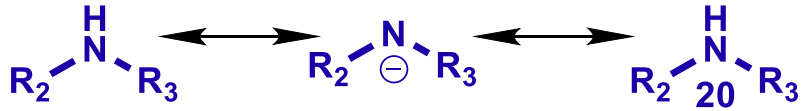
- **Electrophilic radical synthons**



Minisci Reaction

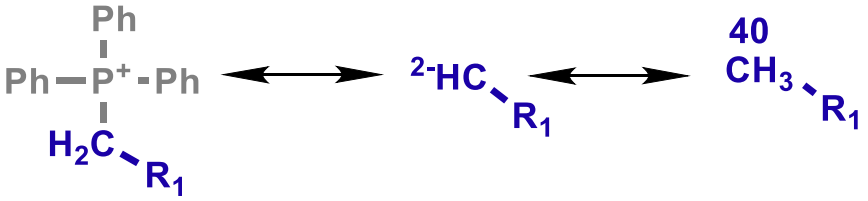
Synthons classes and labels system: nucleophilic synthons

- **Nucleophilic synthons**



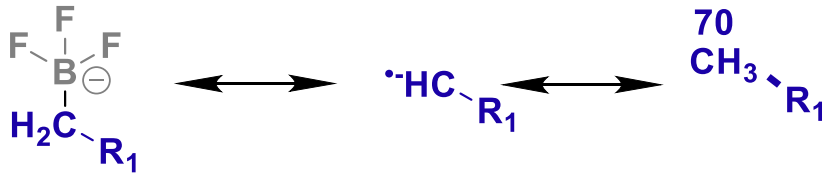
N-acylation, -alkylation, -arylation

- **Bivalent Nucleophilic synthons**



Wittig-type olefination

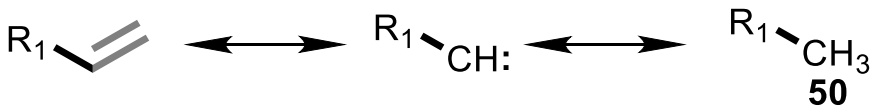
- **Nucleophilic radical synthons**



Minisci and Giese Reaction

Synthons classes and labels system: special synthons classes

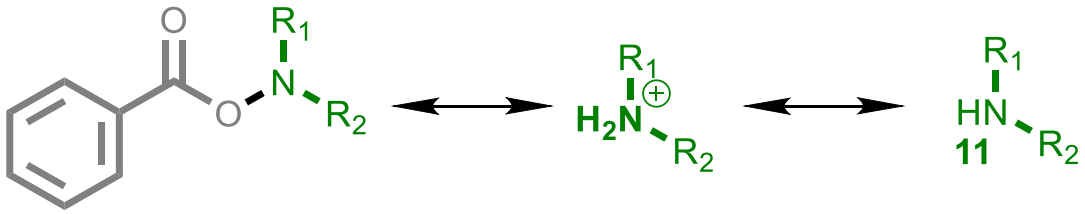
- **Neutral biradical synthons**



Alkene Metathesis

- **Umpolung N-electrophile synthons**

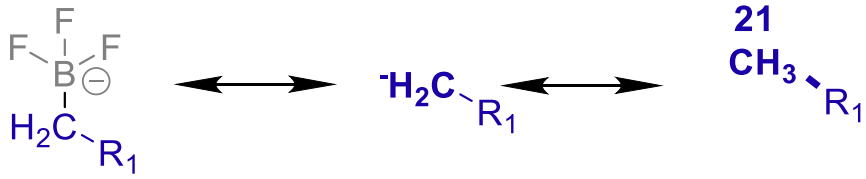
Only from acylated hydroxylamines
$BzO\text{-}NR_1R_2$



Umpoled amination

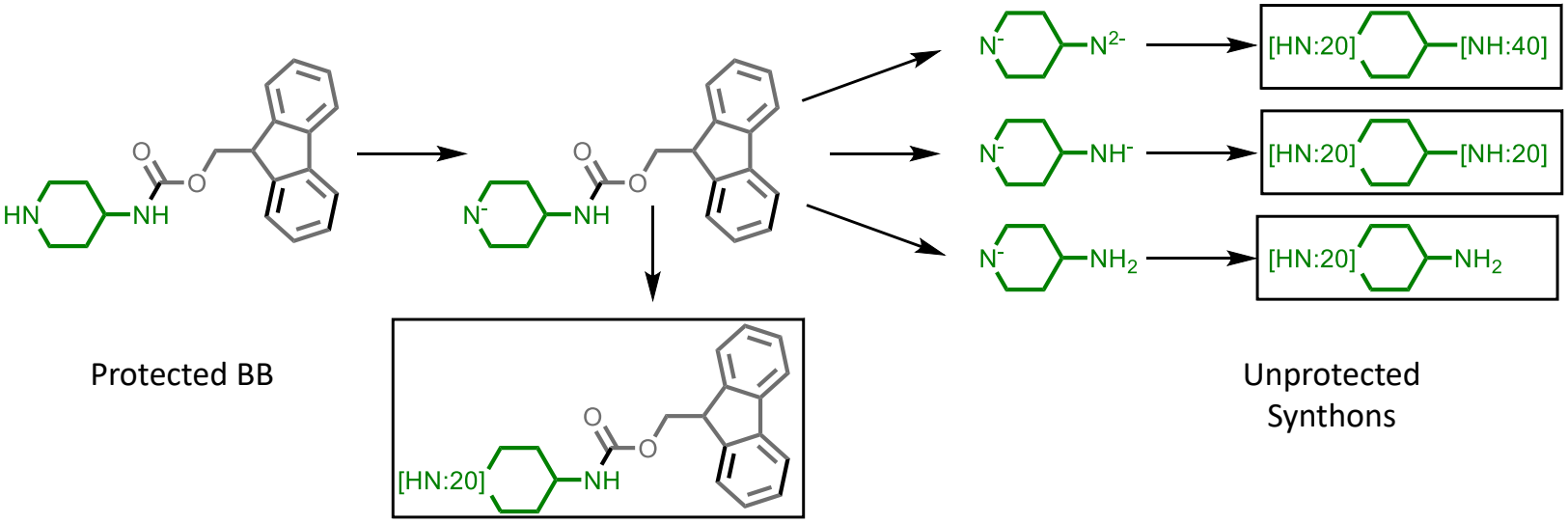
- **Boronic-derived C-nucleophile synthons**

Only BF_3 and MIDA boronates



Chamn-Lam C-O and C-N couplings
Suzuki couplings

Protected BBs



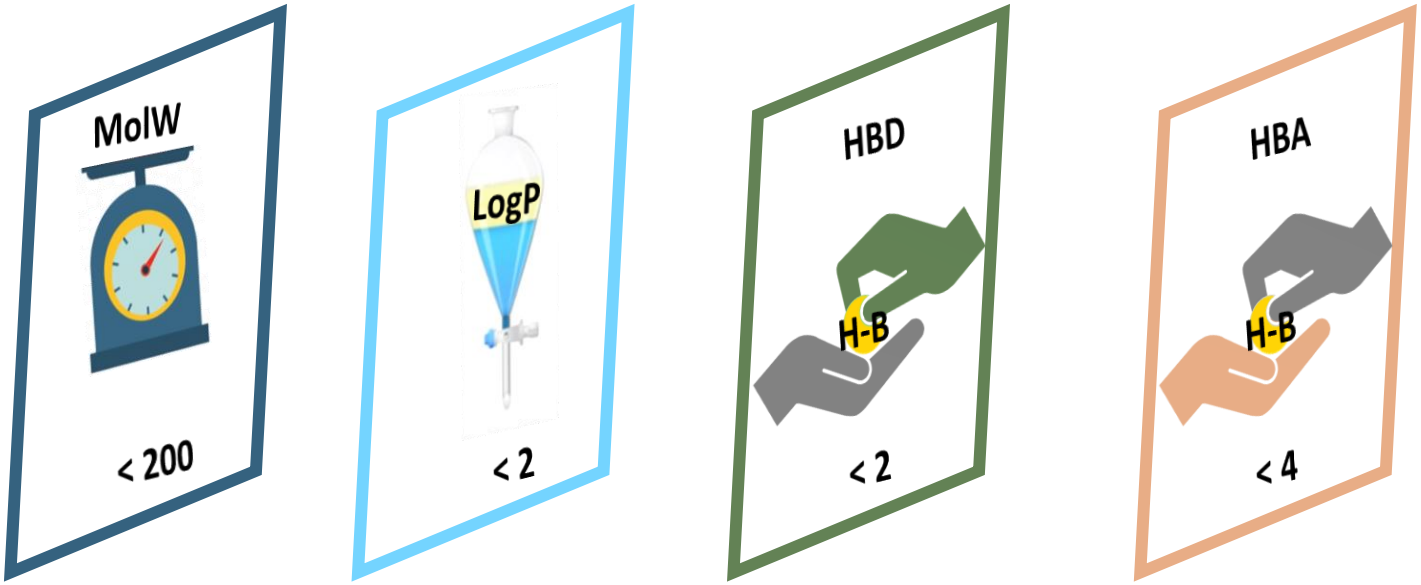
Protected BB

Protected Synthon

Unprotected Synthons

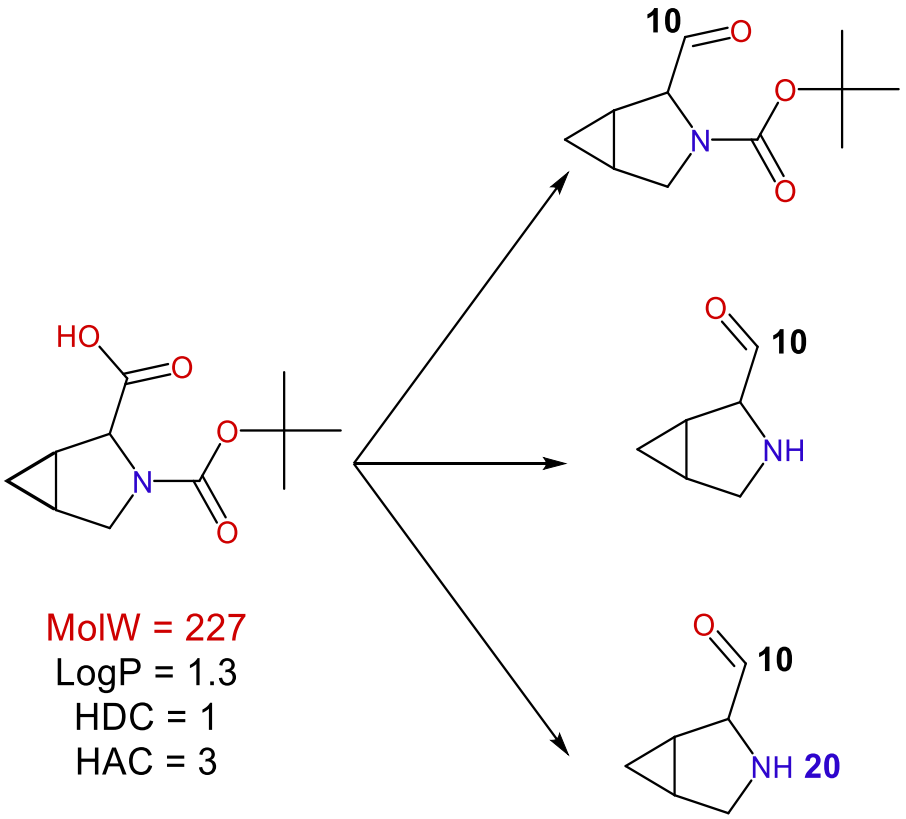
Rule of two filtration

Ro2 filtration is a simple way of BBs prioritization for designing compounds with physical properties that are suitable for oral administration



Should be applied not to the BBs, but to the increments that will be introduced to the molecule by BB - **synthons**

Rule of two filtration

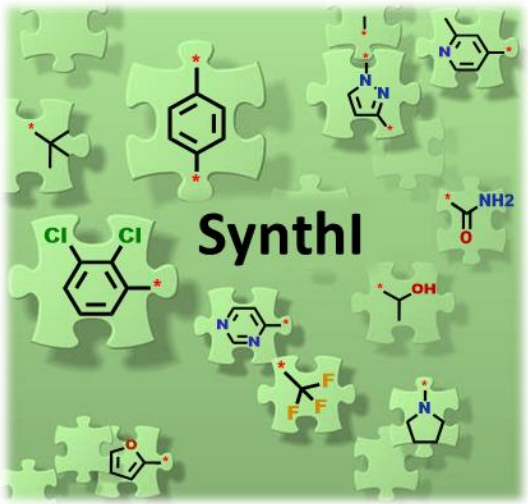


10 - Electrophilic RC 20 - Nucleophilic RC

Synthons Interpreter (SynthI)

SynthI-Classification

SynthI-BB



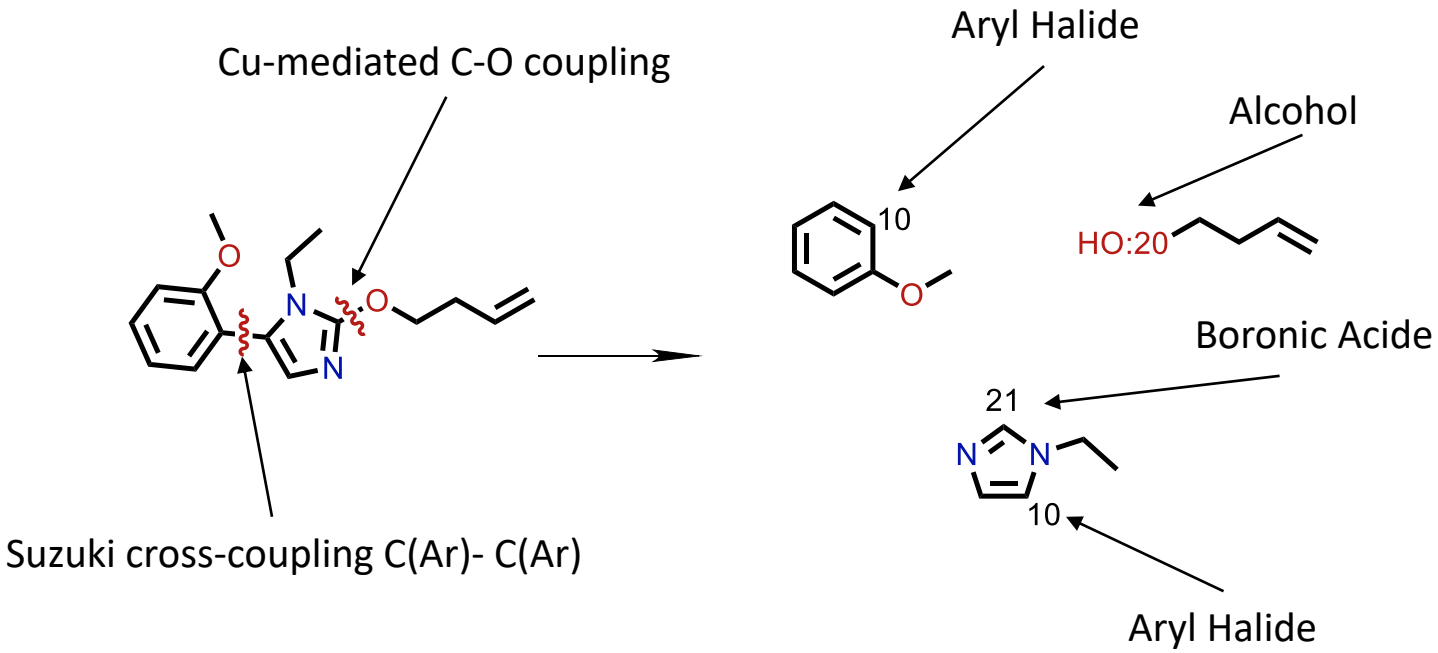
SynthI-Fragmentation

Fragments library of molecules into synthons using selected list of reactions

Knowledge-based generalized toolkit empowering library design for organic chemistry and BBs analysis

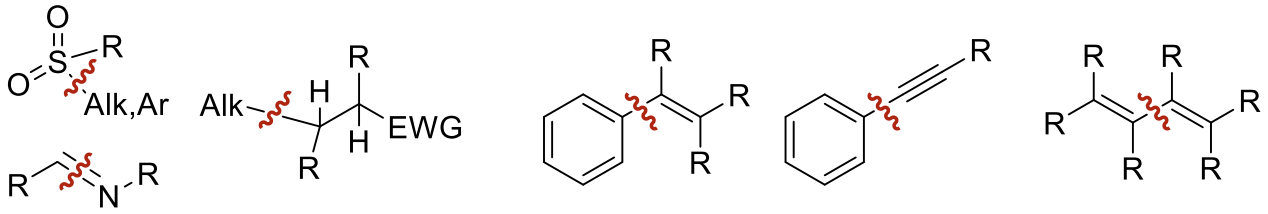
SynthI-Enumeration

SynthI-Fragmentation

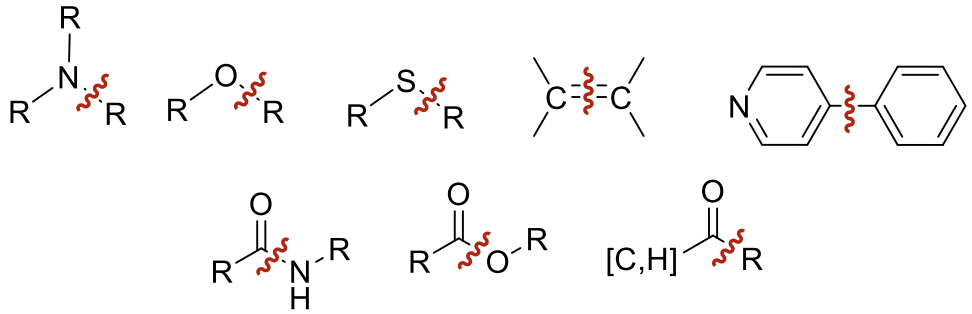


SynthI vs BRICS

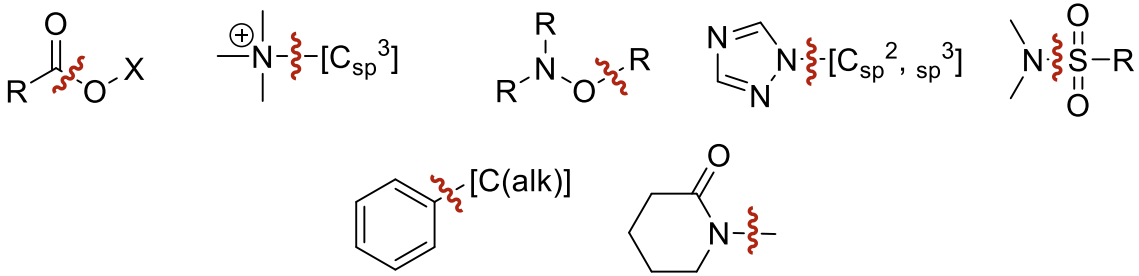
Bonds that are not cleaved at all in BRICS



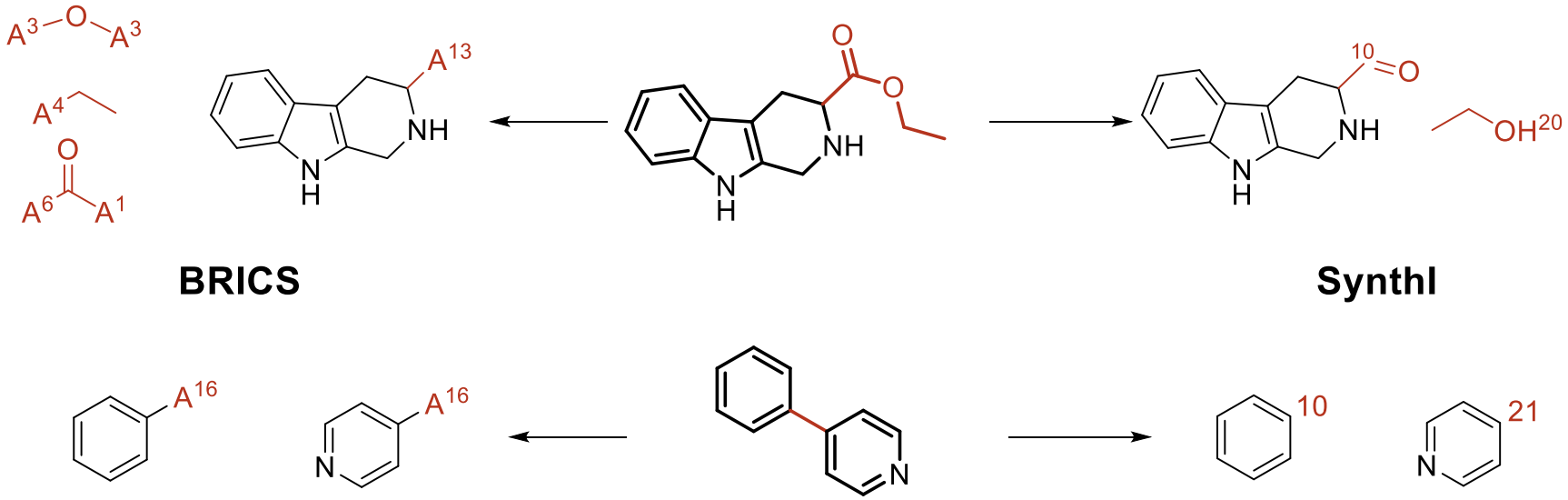
Bonds that are cleaved in BRICS but in different manner



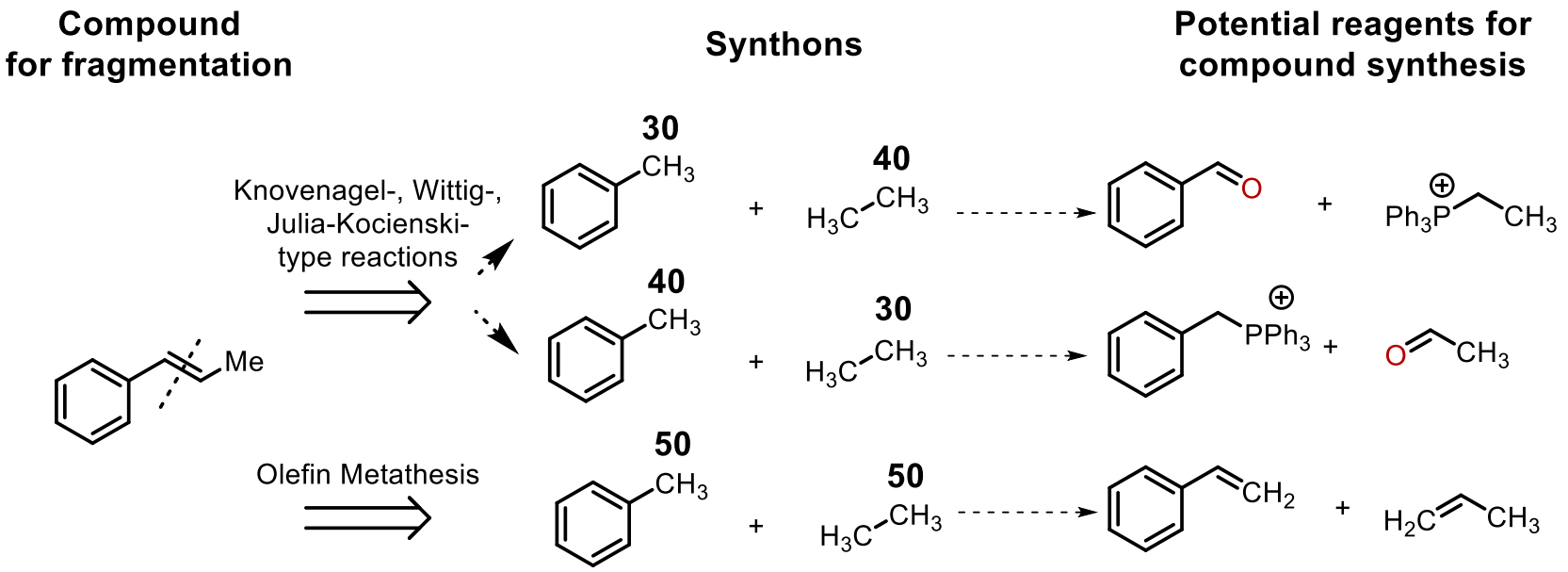
Bonds that are cleaved in the same way in SynthI and BRICS



SynthI vs BRICS: different fragmentations



Different mechanisms of the same bond formation/disconnection in SynthI

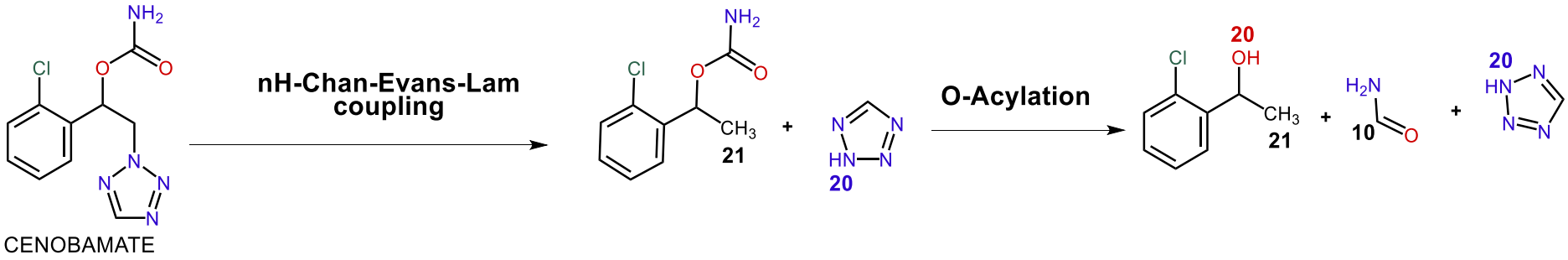


30 - Bivalent electrophilic RC

40 - Bivalent nucleophilic RC

50 - Bivalent neutral RC

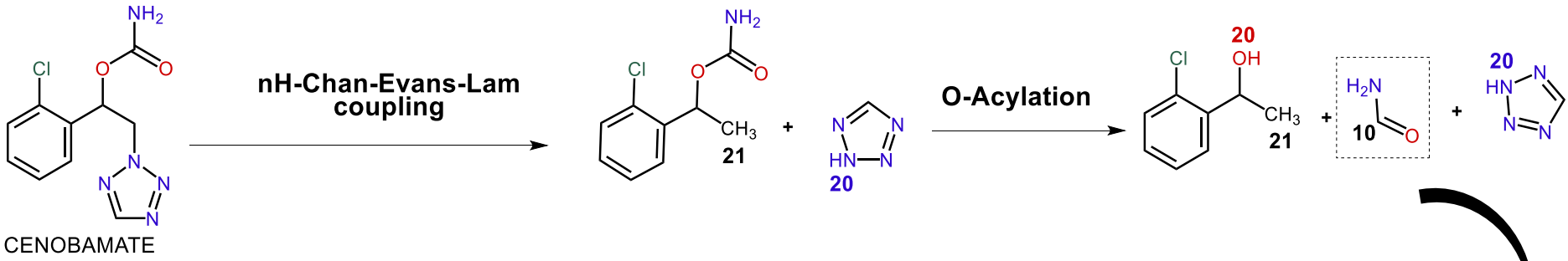
Selection of the optimal synthetic path



10 - Electrophilic RC 20 - Nucleophilic RC 21 - Boronics-derived nucleophilic RC

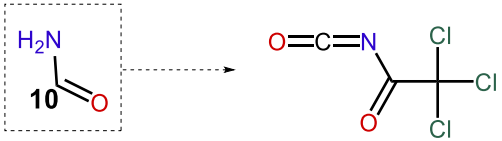
Selection of the optimal synthetic path

$$\textbf{Availability rate} = \frac{\sum \textbf{heavy atoms coming from available synthons}}{\textbf{Total number of heavy atoms in a molecule}} * 100\%$$



10 - Electrophilic RC 20 - Nucleophilic RC 21 - Boronics-derived nucleophilic RC

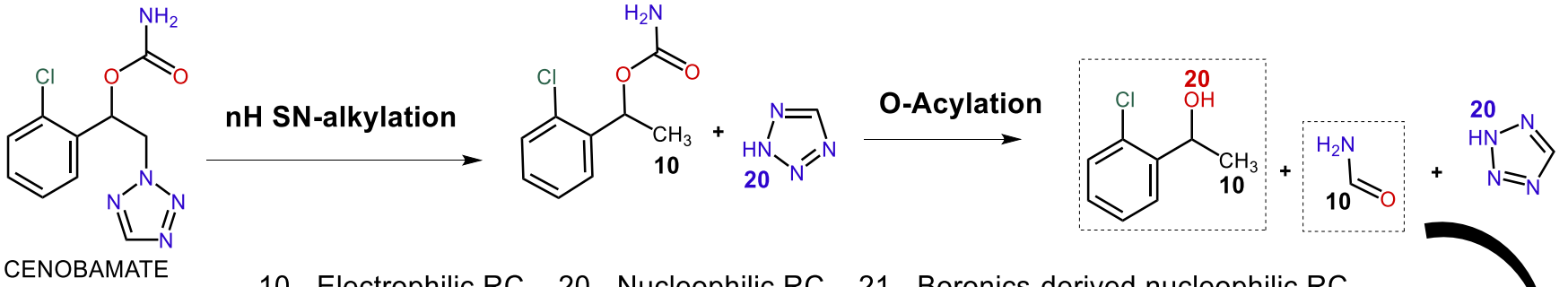
$$\textbf{Availability rate} = \frac{3}{18} * 100\% \approx 17\%$$



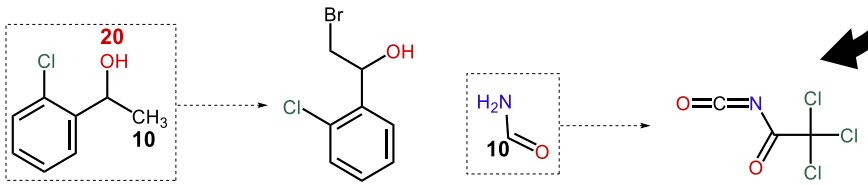
EN300-50197

Selection of the optimal synthetic path

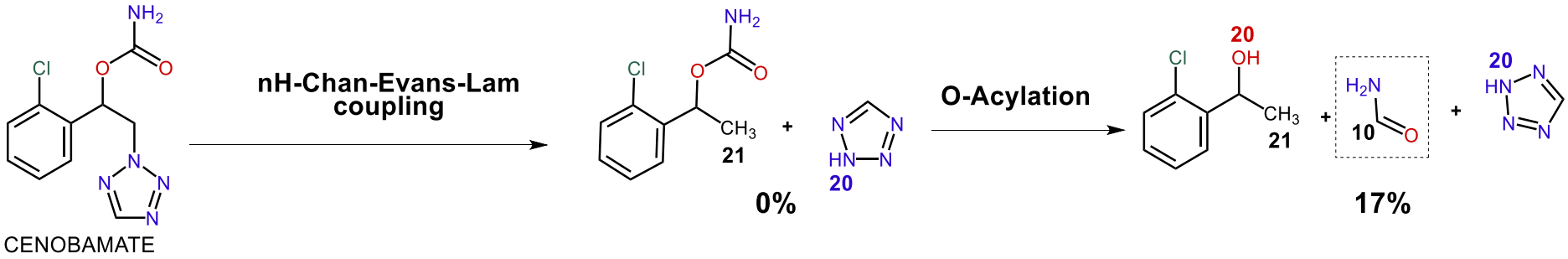
$$\textbf{Availability rate} = \frac{\sum \textbf{heavy atoms coming from available synthons}}{\textbf{Total number of heavy atoms in a molecule}} * 100\%$$

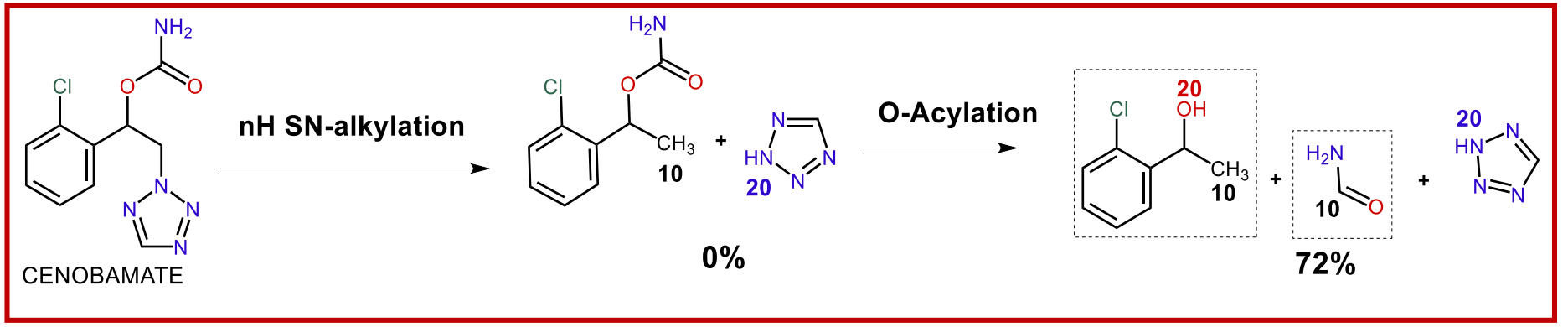


$$\textbf{Availability rate} = \frac{10 + 3}{18} * 100\% \approx 72\%$$



Selection of the optimal synthetic path

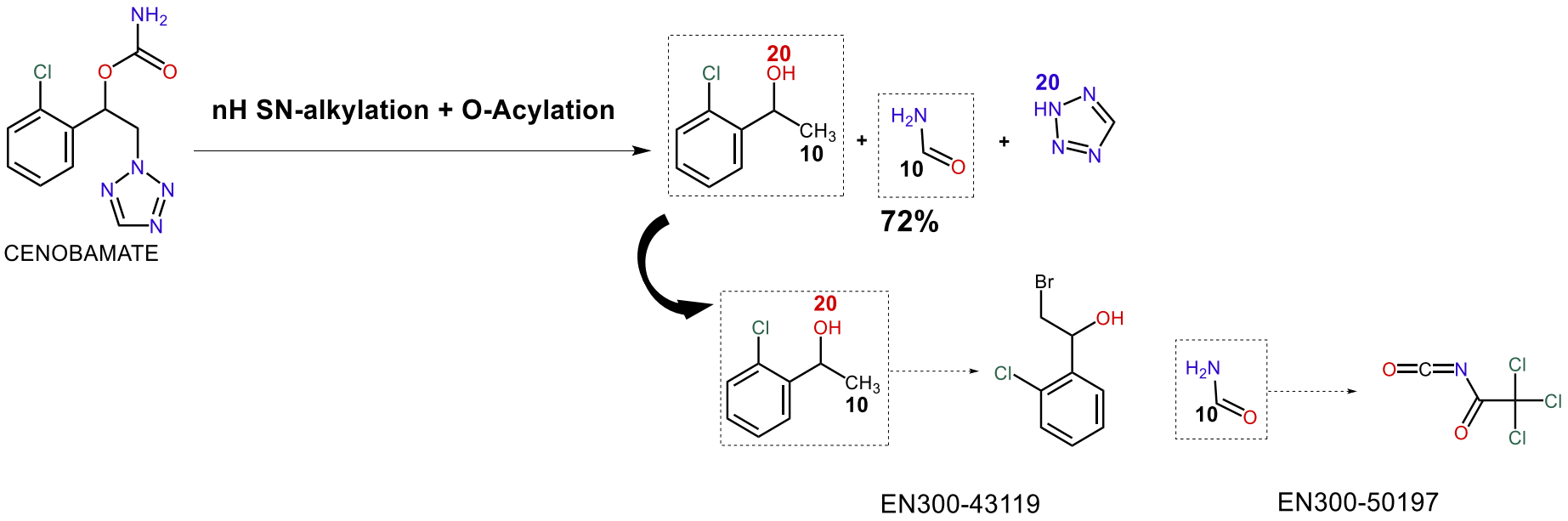




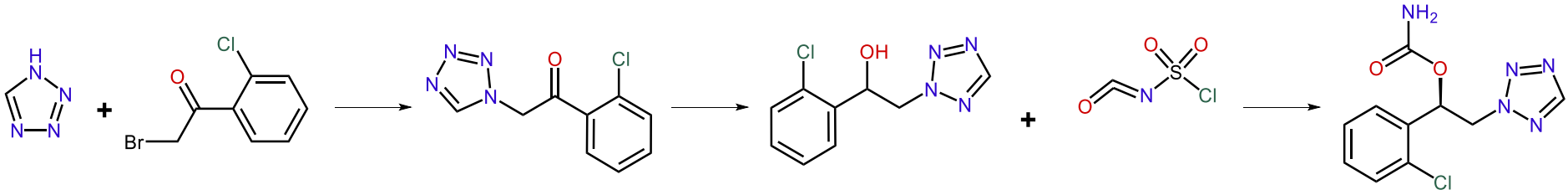
10 - Electrophilic RC 20 - Nucleophilic RC 21 - Boronics-derived nucleophilic RC

SynthI vs experimental synthesis

SynthI selected path



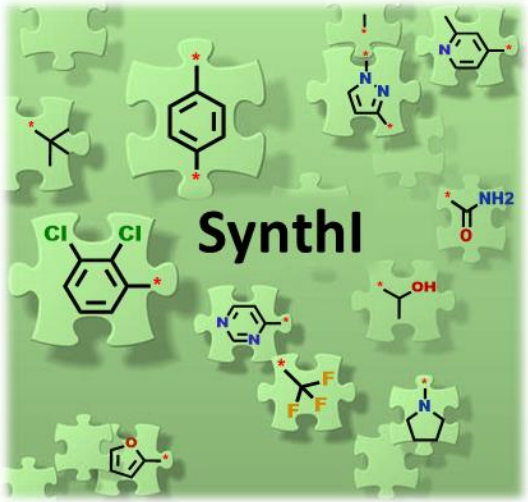
Experimantal synthetic path reported in literature



Synthons Interpreter (SynthI)

SynthI-Classification

SynthI-BB



SynthI-Fragmentation

Knowledge-based generalized toolkit empowering library design for organic chemistry and BBs analysis

SynthI-Enumeration

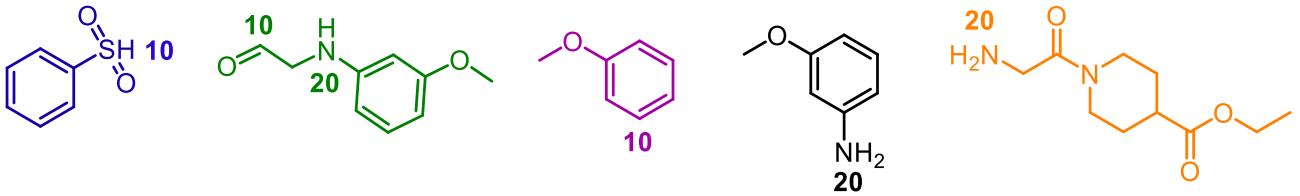
Enumerates compound library from user-defined synthons using selected reactions

SynthI-Enumeration

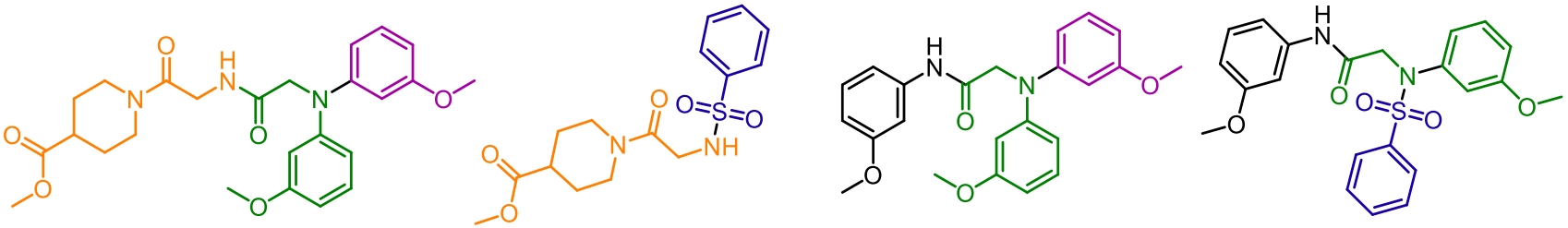
BB library

SynthI-BBs

Collection of synthons



Examples of enumerated compounds

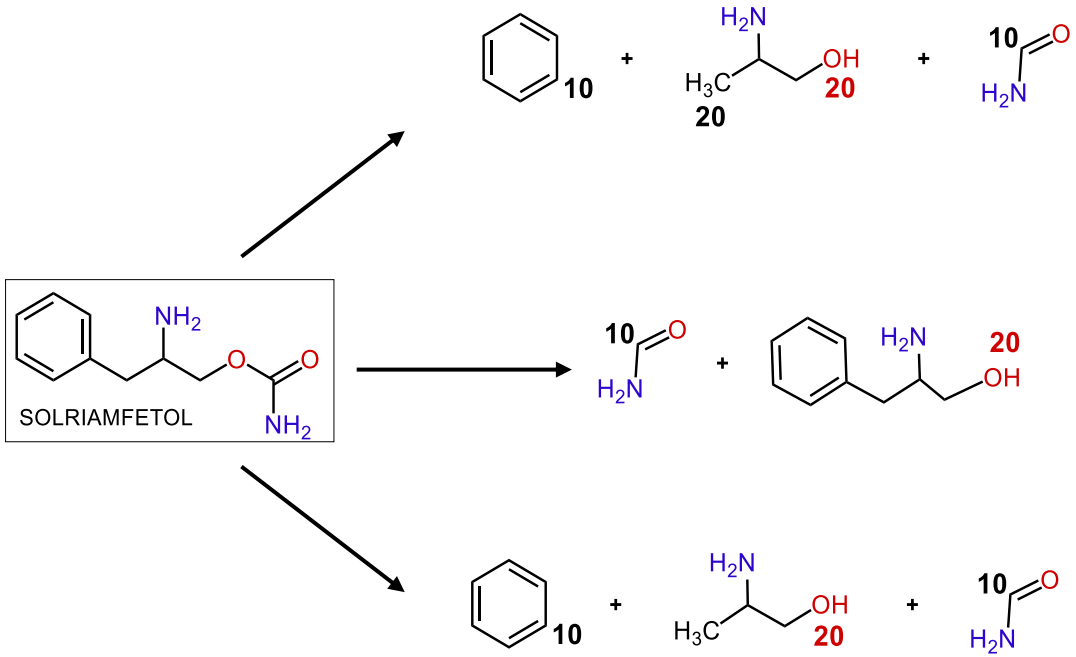


10 - Electrophilic RC

20 - Nucleophilic RC

Analogues library generation with SynthI

Step 1 – fragmentation of the reference molecule

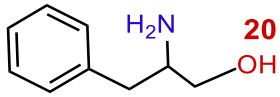


10 - Electrophilic RC 20 - Nucleophilic RC 21 - Boronics-derived nucleophilic RC

Analogues library generation with SynthI

Step 2 – selection of the analogues for each synthon of each path

(source -> library of available synthons)

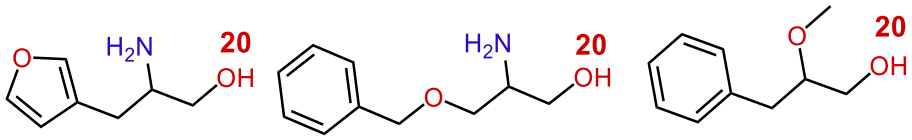


20 - Nucleophilic RC

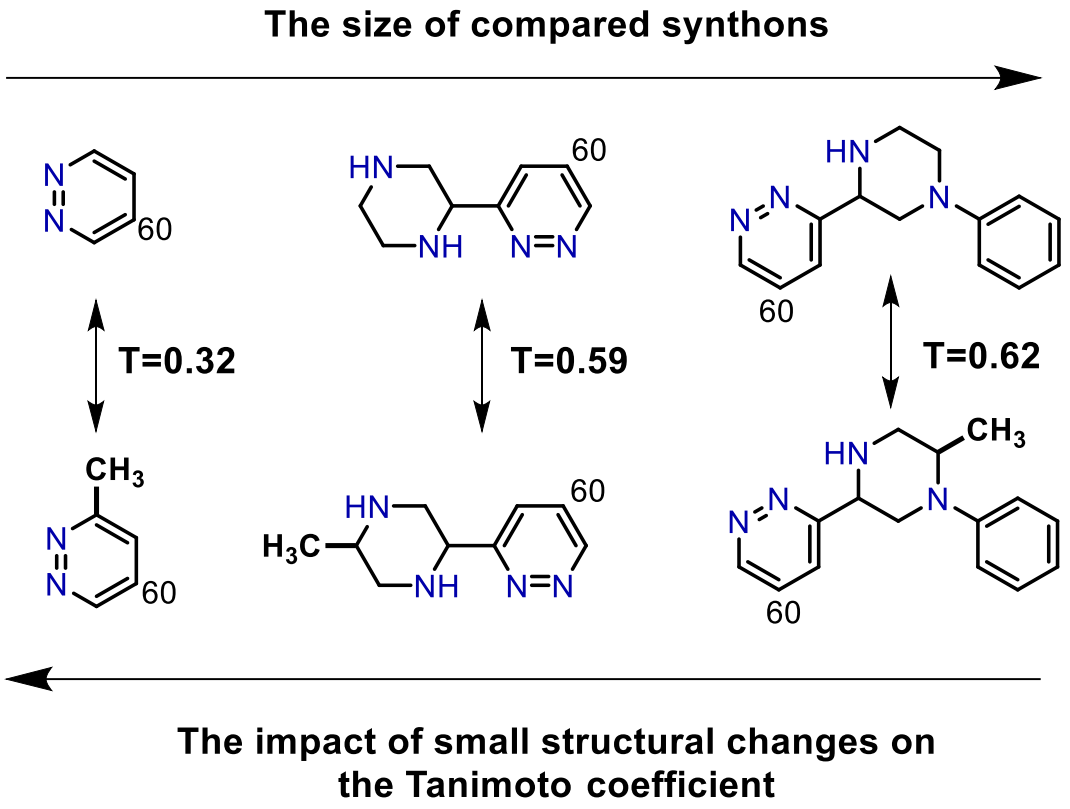
Primary constraints:

- Number of RCs
- Type of RC
- Number of cycles

Morgan Fingerprints Similarity (Tanimoto>=0.5)

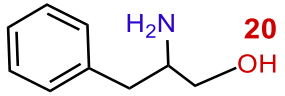


Similarity in synthons space



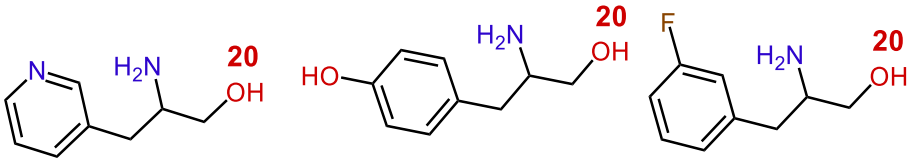
Analogues library generation with SynthI

Step 2 – selection of the analogues for each synthon of each path



20 - Nucleophilic RC

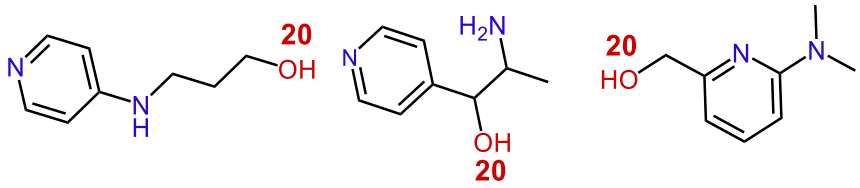
Positional Analogues Scanning (PAS) approach



Addition or removal of
CH_3, F, NH_2, OH,

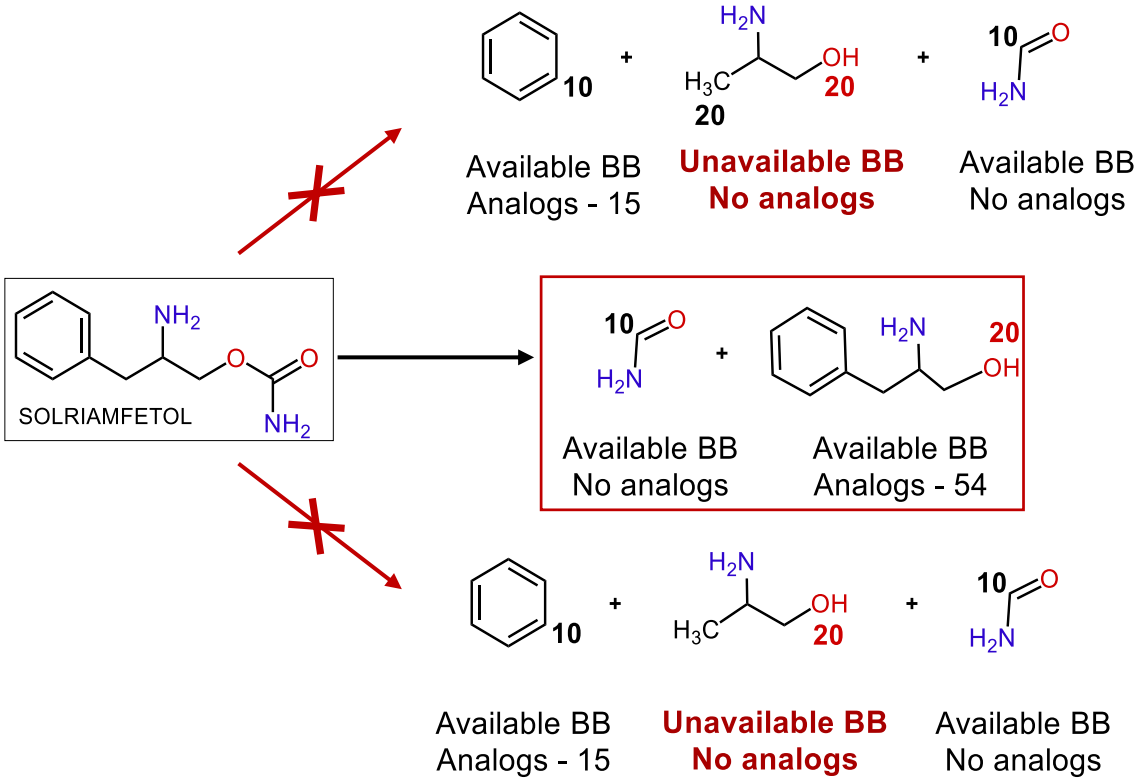
C_{Ar} -> N_{Ar} replacement + Isomeric rearrangement

C_{Ar}→N_{Ar} or N_{Ar}→C_{Ar} replacements combined with isomeric rearengents



Analogues library generation with SynthI

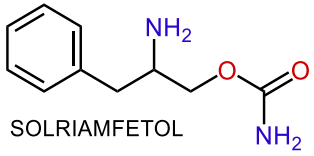
Step 3 – selection of the fragmentation scheme(s) for analogues distribution



10 - Electrophilic RC 20 - Nucleophilic RC 21 - Boronics-derived nucleophilic RC

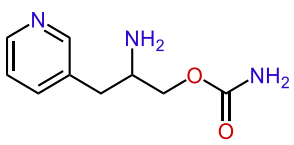
Analogues library generation with SynthI

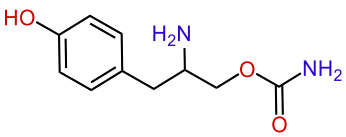
Step 4 – enumeration of the analogues of reference compound

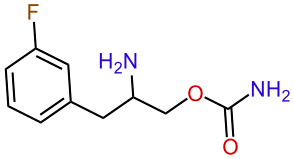


SOLRIAMFETOL

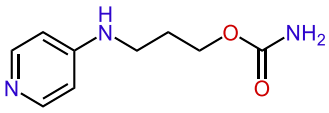
a) PAS approach

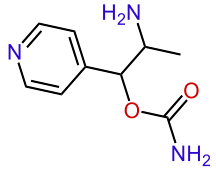


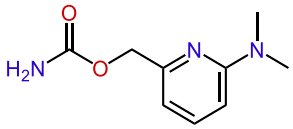




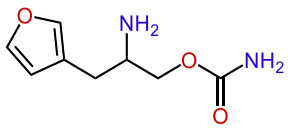
b) C_{Ar}->N_{Ar} replacement + Isomeric rearrangements

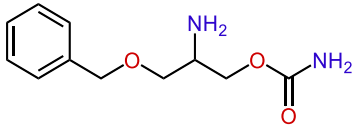


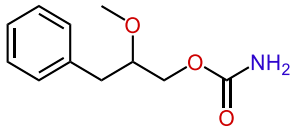




c) Morgan Fingerprints Similarity (Tanimoto>=0.5)

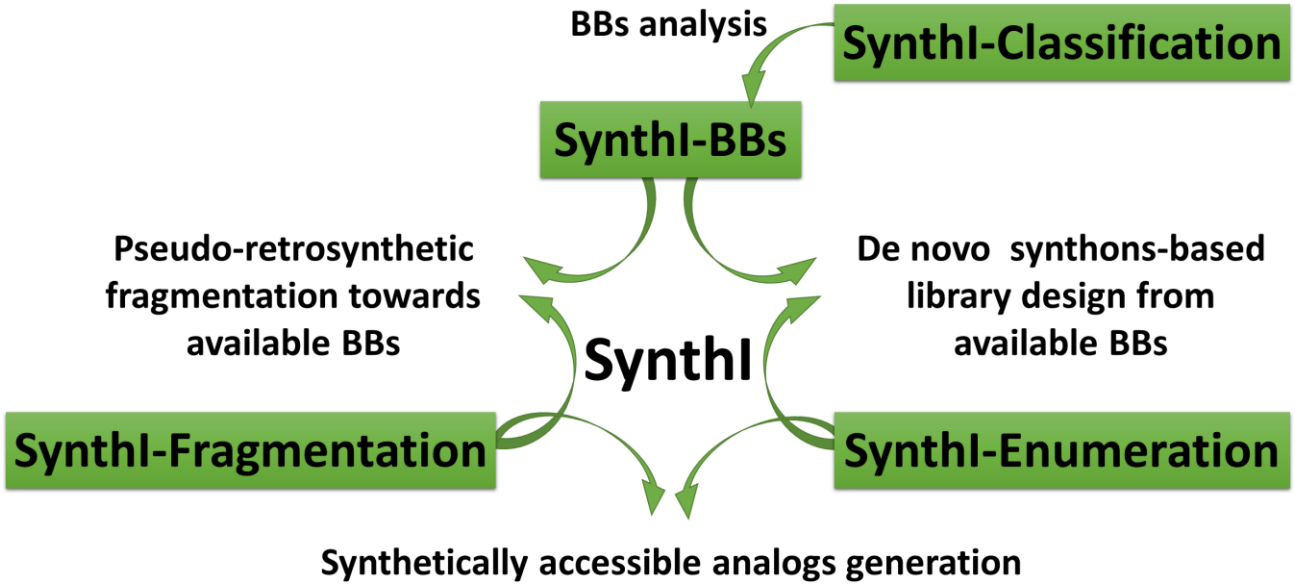






Conclusions

SynthI is an open-source tool for synthons generation, BBs analysis and library design (https://github.com/Laboratoire-de-Chemoinformatique/SynthI).



Zabolotna, Y. et al. SynthI: a new open-source tool for synthon-based library design. ChemRxiv. ***Cambridge: Cambridge Open Engage*****. 2021 (Accepted in JCIM)**

UNIVERSITÉ DE STRASBOURG

ChémInformatique STRASBOURG

Chimie de la Matière Complexe (CMC)

Thank you!

Supervisors:

Prof. Alexandre VARNEK

Dr. Dragos HORVATH

Dr. Gilles MARCOU

Dr. Olga KLIMCHUK

External collaborators (medicinal chemistry expertise):

Prof. Dmitrii VOLOCHNYUK; Sergey V. RYABUKHIN; Kostiantyn GAVRYLENKO; Oleksandr OKSIUTA