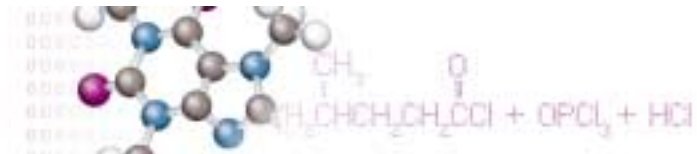




-7.5829	-1.8843	-0.8096	C	0.00
-8.8889	0.1947	-0.3218	O	0.00
10.2795	2.5700	0.7193	C	0.00
-8.5296	-1.8624	-0.8250	H	0.00
-7.5171	-2.3313	0.2621	H	0.00
-7.0177	-1.8782	-1.2207	H	0.00



Novel Approach to Retrosynthesis Automatic Generation of Transform Libraries

V. Eigner-Pitto, J. Eiblmaier, H. Kraut, H. Saller and P. Loew.
InfoChem GmbH, Landsberger Strasse 408, Munich, 81241, Germany

Why Retrosynthesis?

Is there a shorter synthesis path to my target?



How can I synthesize this compound?



Can we increase the yield of this process by choosing another synthesis strategy?





Retrosynthesis: Expected Features

Main Challenge: Automatic generation of chemically meaningful precursors and their reasonable validation

1. Evaluation of Alternative Synthesis Pathways

Sorting criteria are necessary for broad representations with many expanding paths:

- Yield
- Type of transformation (e.g. ring-closures or construction reactions preferred, etc.)

2. Support of Multistep Synthesis

Necessary to guide the synthesis from the target to all existing precursors that can be obtained from commercially available substances

3. Organized Visualization

Allows a broad view over the different strategies. A tree view, i.e., gives an organized representation of the synthesis paths to the target



Synthesis-Planning Tools

1. Use of known compounds and published reactions

i.e. MDL (*CrossFire Commander*), Symyx (*Reaction Pathfinder*), InfoChem (*STS*)

- ✓ Handle multistep synthesis
- ✓ Order in synthesis paths - Sorting criteria
- ✓ Shows synthesis options in your database (tree representation)
- ✗ Structure and corresponding reaction(s) must be in database

2. „Intellectual“ retrosynthesis systems

- ✓ New molecules (not in the database)
- ✓ Alternative routes to known molecules
- ✓ Reaction transformation / Synthesis strategy to specify
- ✓ Transform based



„Intellectual“ Retrosynthesis Approaches

2a. Transforms compiled manually from literature

- ✓ accurate, highly specific
- ✗ tedious, expensive, restrictive

2b. Transforms based on enumerated bond changes

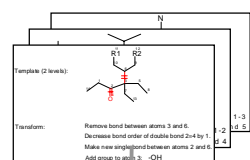
- ✓ fully automated
- ✗ unspecific, combinatorial explosion

Examples:

- LHASA, SECS / CASP ... (Corey, Wipke)
- WODCA, EROS ... (Gasteiger)
- SYNGEN ... (Hendrickson)
- AIPHOS ... (Sasaki)
-

RetroSynthesis Concept Overview

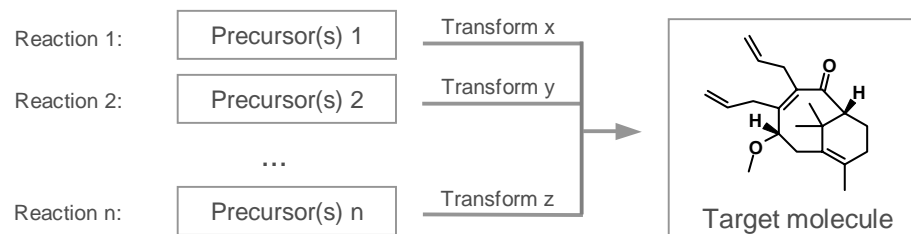
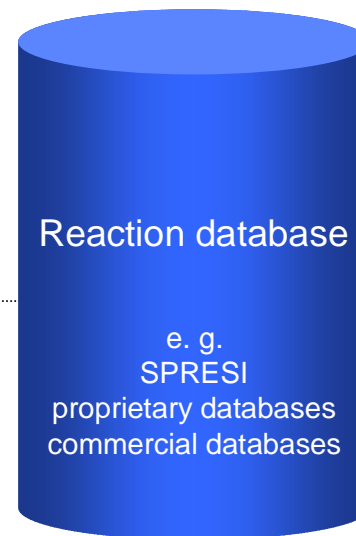
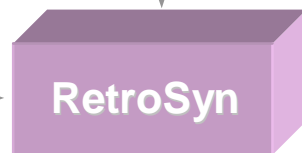
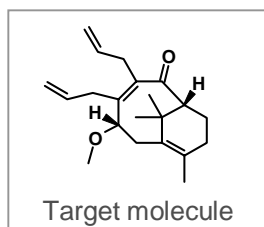
1. Pre-processing



Automatic transform extraction (ICMAP/CLASSIFY)



2. RetroSynthesis

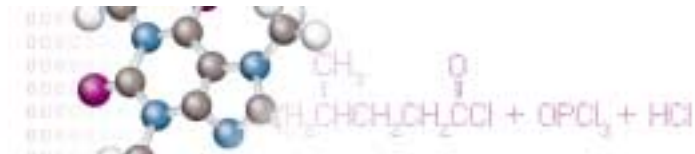


Lookup stored examples

Transform x
 Transform y
 Transform z



-7.5829	-1.8843	-0.8999	C
-8.8889	-0.1947	-0.3218	O
10.2795	2.5708	0.7393	C
-8.5096	-1.8624	-0.8250	H
-7.5171	-2.3313	0.2621	H
-7.0177	-1.8782	-1.2207	H



Reaction Database Requirements

1. Correct Reaction Mapping

Description of the transformation occurring during the reaction

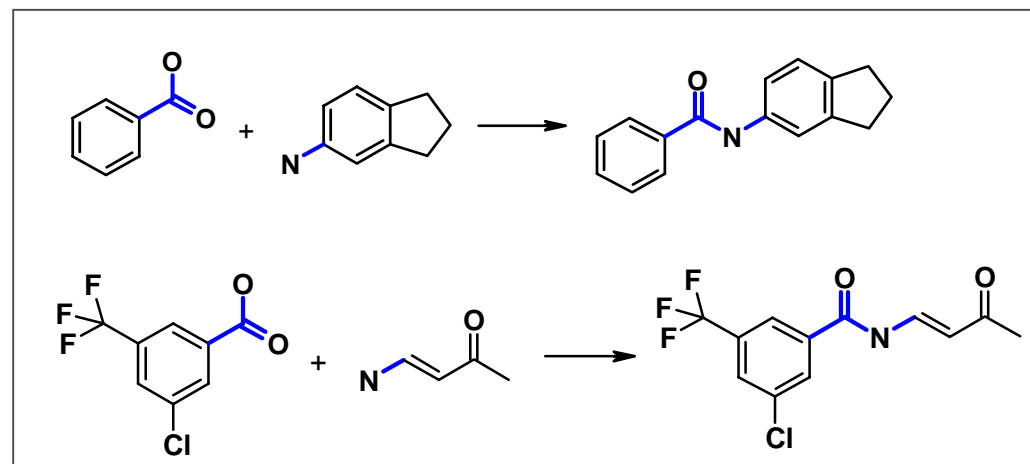
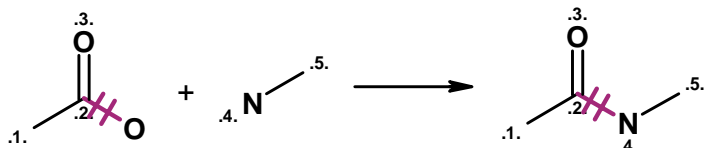
⇒ Determination of reaction centers

2. Reaction Classification

Automatic identification of the structural environment around the reacting centers

⇒ Determination of unique identifiers (ClassCodes)

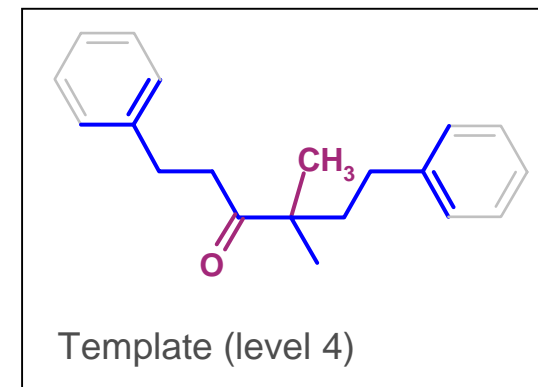
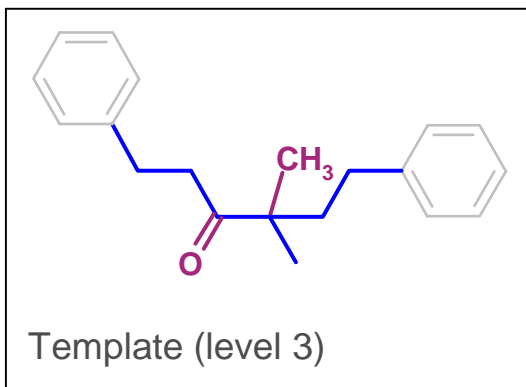
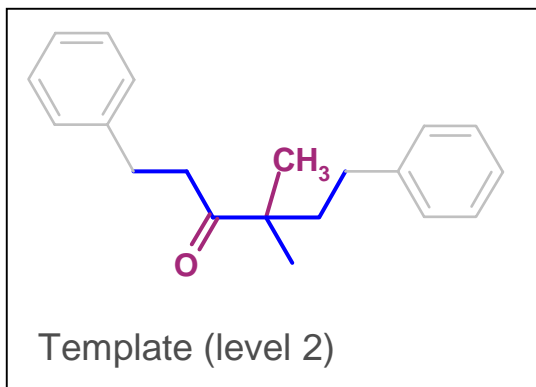
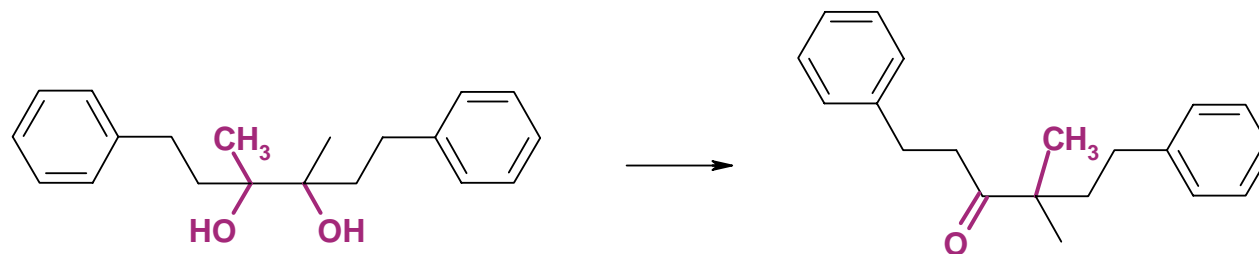
Reaction Type **Amination**:



Definition of Template (Retron)

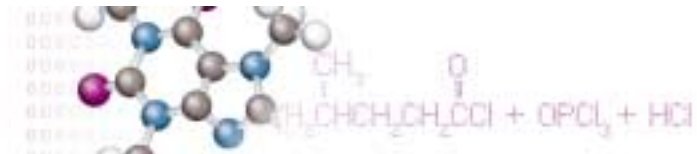
SPRESI Reaction Reg-No.: 3179546

Kita Yasuyuki, Yoshida Yutaka, Mihara Sachiko, Fang Dai-Fei, Higuchi Kazuhiro, Furukawa Akihiro, Fujioka Hiromichi
 Efficient pinacol rearrangement mediated by trimethyl orthoformate
 Tetrahedron Lett., 38 (1997) 48, 8315-8318



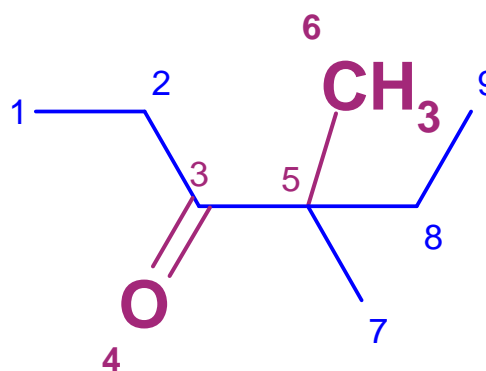


-7.5829	-1.8843	-0.8098	C	0.00
-8.8889	0.1947	-0.3218	O	0.00
10.2795	2.5708	0.7193	C	0.00
-8.5096	-1.8624	-0.8250	H	0.00
-7.5171	-2.3313	0.2621	H	0.00
-7.0177	-1.8782	-1.2207	H	0.00



Definition of Transform

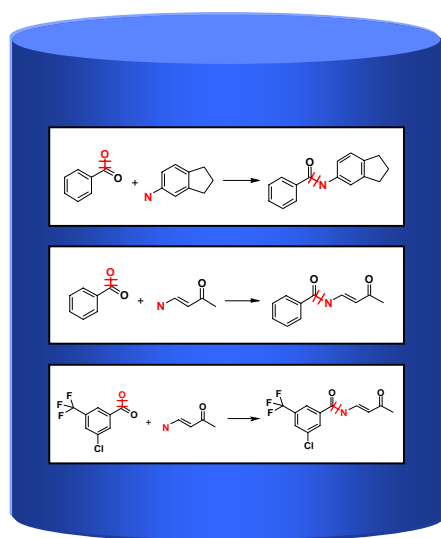
Template:



- Indications:**
- Remove bond between atom 5 and 6
 - Decrease bond order of double bond 3=4 by 1
 - Make new single bond between atoms 3 and 6
 - Add group to atom 5: -OH

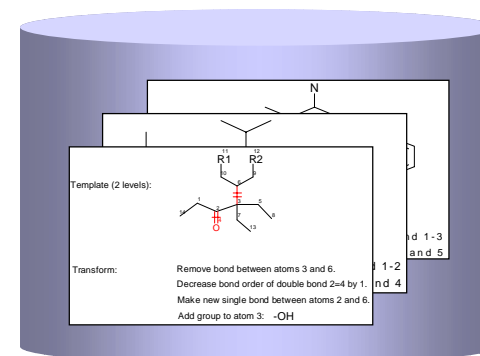
Pre-processing Step (Transform Library Building)

Reaction Database

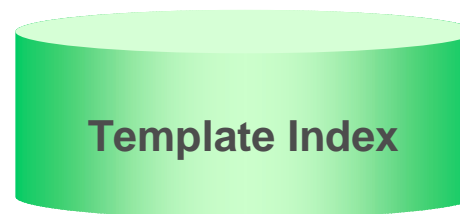


Transform Generator

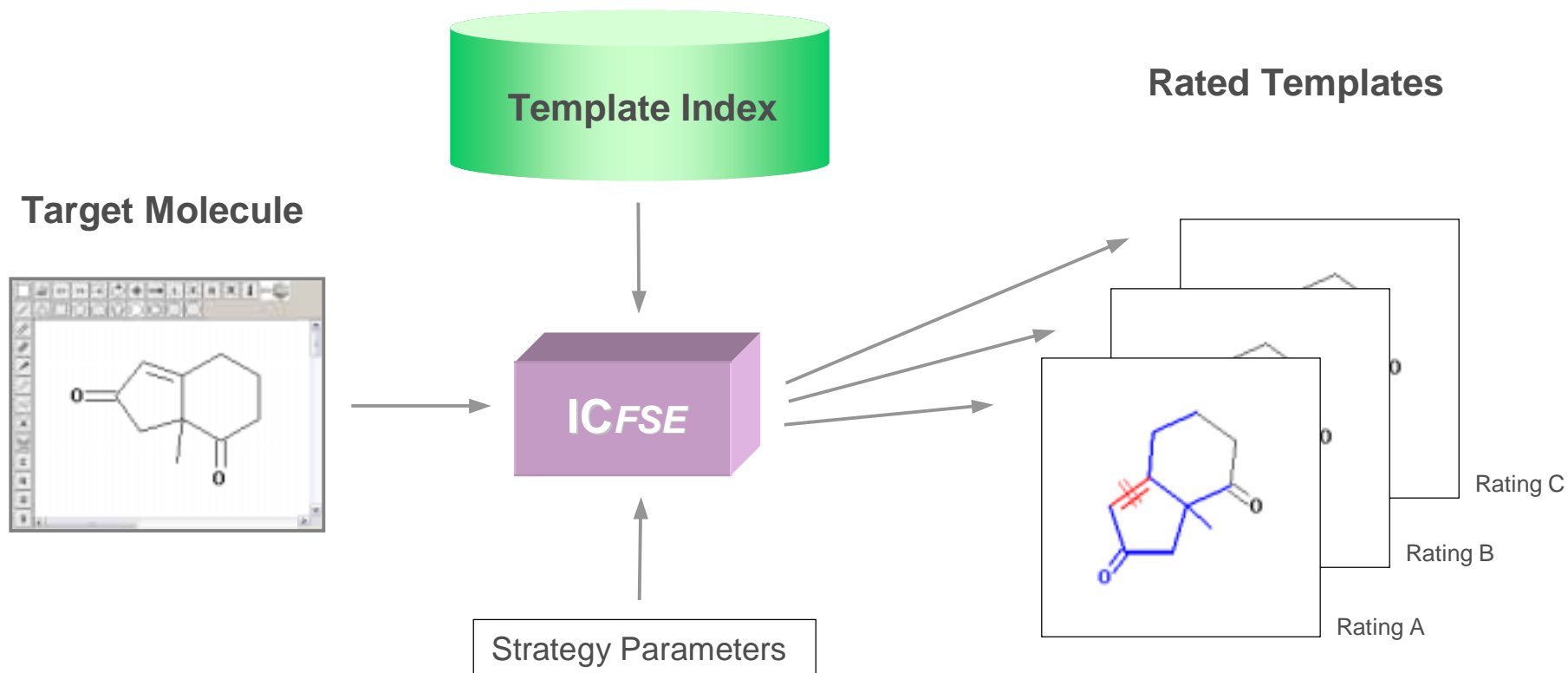
Transform Library



Generation of substructure index



Templates Search





Strategy Parameters and Precursors Rating

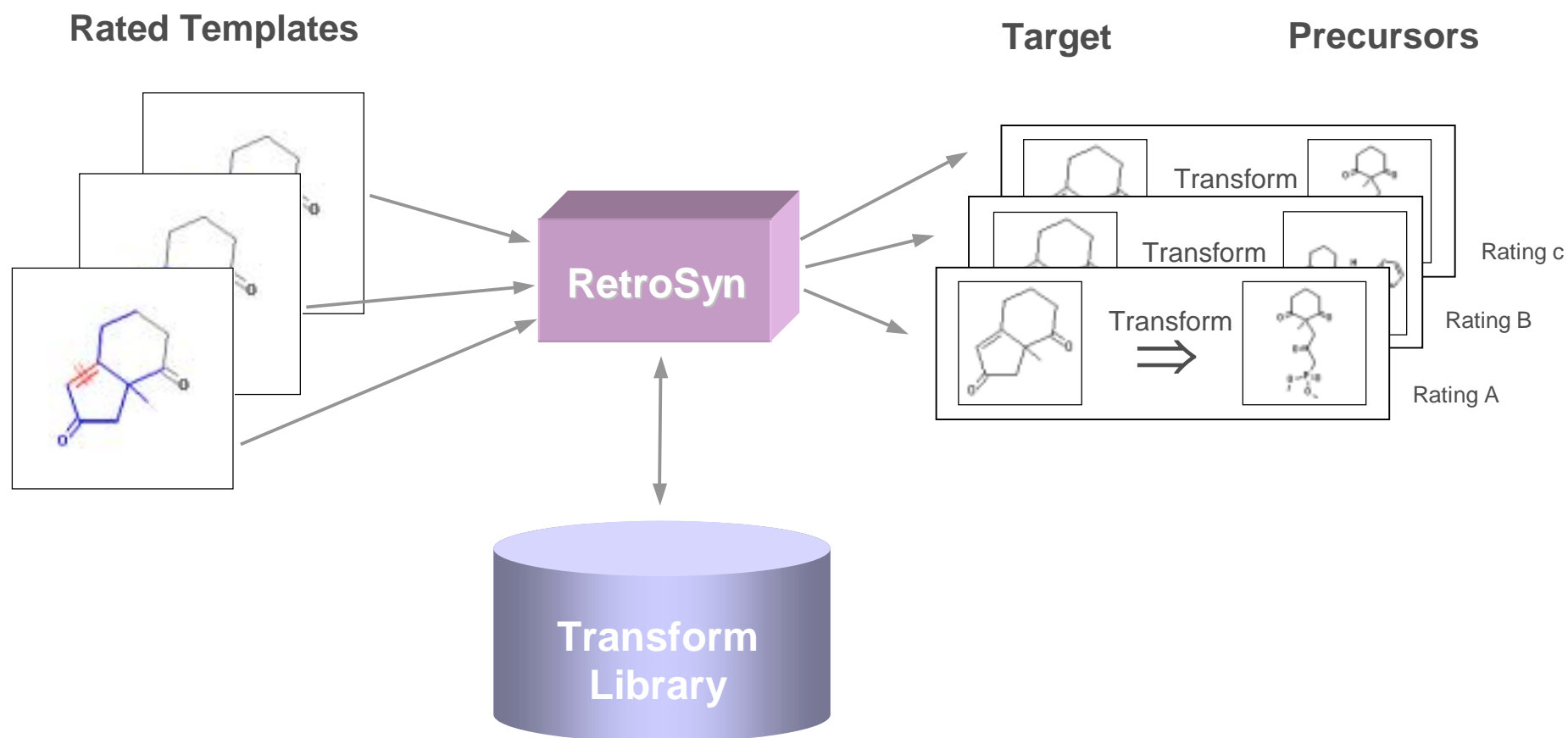
Strategy parameters for precursor-rating consider the following criteria:

- formation of a new carbon-carbon bond
- bond change within a functional group
- formation of a new ring/aromatic bond
- formation of a new bond between two symmetric fragments
- formation of a new bond near the center of the molecule
- number of generated precursors

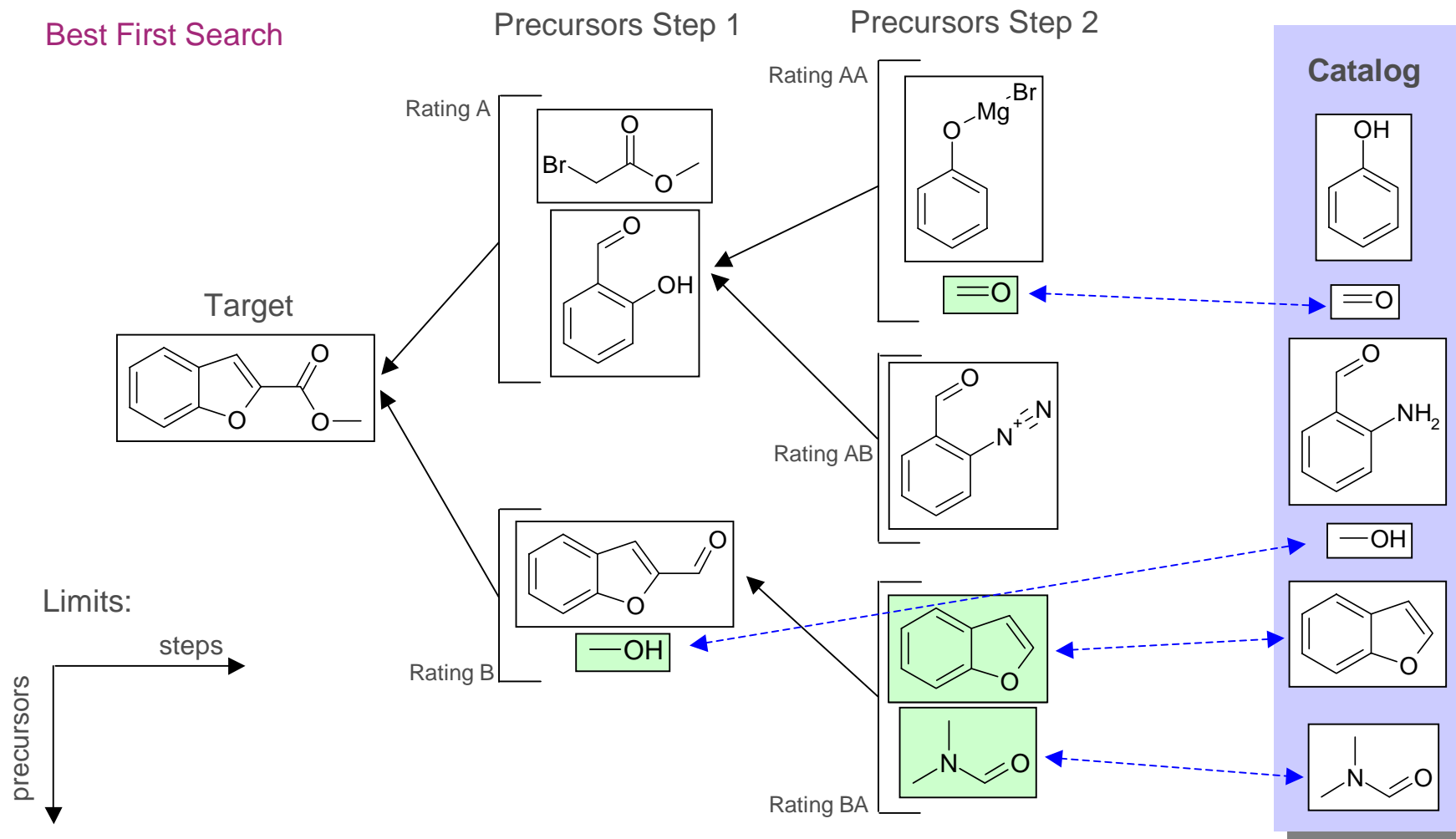
These criteria are weighted depending on the chosen configuration:

- sets of pre-defined strategy parameters are available
- sets may be defined from user

Precursors Generation from Target



Multistep Synthesis





-7.5829	-1.8843	-0.8996	C	0.00
-8.8889	-0.1947	-0.3218	O	0.00
10.2795	2.5708	0.7393	C	0.00
-8.5096	-1.8624	-0.8250	H	0.00
-7.5171	-2.3313	0.2621	H	0.00
-7.0177	-1.8782	-1.2207	H	0.00



Prototype: Query Form

SPRESIweb -- InfoChem GmbH -- Microsoft Internet Explorer

File Edit View Favorites Tools Help

SPRESIweb 2.3

Home | New Query | History | Help | Logout

You are logged in as **Valentina** via **password** Status: InfoChem GmbH

SEARCH ...

- Molecules
- Reactions
- References
- Retrosynthesis

STS ...

- New session
- Open session
- Delete session

QUERY ...

- New
- Save
- Open
- Delete

HIT LIST ...

- Save
- Open
- Delete
- Import
- Export
- Download statistic

NEW RETROSYNTHESIS QUERY

Strategy:

Library:

Start with:

Number of precursors per target:

CHOOSE EDITOR:

- ICEDIT
- ISIS/Draw
- ChemDraw

Show Quickhelp



Reaction Databases Used in Prototype

- SPRESI Synthesis**

2.5 mio reactions from SPRESI
reactions with 50% minimum yield
- Name Reactions**

140.000 reactions from SPRESI
all examples for a large set of Name Reactions
- EROS**

68.000 reactions from J. Wileys
„Encyclopedia of Reagents for Organic Synthesis“
- SOS**

120.000 reactions from Thiemes
„Science of Synthesis“
- CAC**

6.600 reactions from Springers
„Comprehensive Asymmetric Catalysis“



-7.5829	-1.8843	-0.8998	C
-8.8889	0.1947	-0.3218	O
-10.2795	2.5708	0.7393	C
-8.5090	-1.8624	-0.0250	H
-7.5171	-2.3313	0.2621	H
-7.0177	-1.8782	-1.2257	H



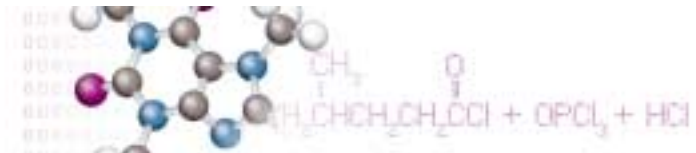
17 / 20

Prototype: Suggested Precursors

The screenshot displays the SPRESiweb interface in a Microsoft Internet Explorer browser window. The left sidebar, titled "PRECURSORS", lists three reaction suggestions for the synthesis of a furan derivative. The second suggestion is selected, showing a list of reaction IDs: 1710501, 1710502, 1710503, 1710504 (highlighted with a mouse cursor), 1710505, 1710506, 1710507, 1710508, and 9655. The main content area shows the reaction details for reaction ID 1710504, including the chemical structures of the reactants (2-chloro-3-methylfuran and piperidine) and the product (2-(3-methylfuran-2-yl)propanoate piperidine salt). The reaction is identified as a Schotten-Baumann reaction with a 95% yield. A reference is provided: SHAFIEE A.; SAVABI F.; REZVANI A.; FARRKHSIAR M.; KHOYI A.; SYNTHESIS UND LOKALANAESTHETISCHE WIRKUNG VON BENZO-FURANDERIVATEN. J. PHARM. SCI., 1978, 67, NO 1, 125-127.



-7.5829	-1.8843	-0.8096	C
-8.8829	-0.1947	-0.3218	O
10.2795	2.5708	0.7393	C
-8.5096	-1.8624	-0.8250	H
-7.5171	-2.3313	0.2621	H
-7.0177	-1.8782	-1.2207	H



Prototype: Synthesis Tree

The screenshot shows a web browser window titled "http://192.168.0.39 - SPRESI 2 Search Page - Microsoft Internet Explorer". The browser displays a synthesis tree with a red box highlighting a specific node. Below the tree, a chemical reaction is shown, also highlighted with a red box. The reaction involves the synthesis of a complex heterocyclic molecule from a starting material and a reagent.

Chemical reaction shown:

CC1=C(C(=O)Cl)OC2=CC=CC=C12 + CO >> CC1=C(C(=O)OC)OC2=CC=CC=C12

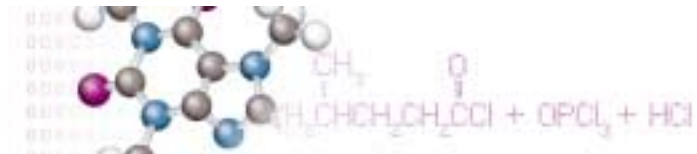


Conclusions

- Very large pool of data available (SPRESI 4 million reactions)
- Automatic transform generation from reaction databases
- Suggestions based on real (analogous) reactions
- Verification of suggested synthesis reactions, through link to reaction database / literature
- Availability-check of precursors in catalogs
- System will be integrated in a retrieval system providing common search tools



-7.5829	-1.8843	-0.8098	C
-8.8889	0.1947	-0.3218	O
10.2795	2.5708	0.7193	C
-8.5096	-1.8628	-0.8250	H
-7.5171	-2.3313	0.2621	H
-7.0177	-1.8782	-1.2207	H



Thank you!

InfoChem GmbH:

www.infochem.de

www.spresi.de

info@infochem.de