Computer-aided synthesis design (CASD) and forward reaction prediction tools for both idea generation in new synthesis route planning and for de novo molecule design

V. Eigner-Pitto, F. Huerta*, M. Hutchings, H. Saller, P. Loew
InfoChem GmbH, Landsberger Strasse 408, Munich, 81241, Germany
*Chemnotia AB, Södertällje 151 36, Sweden
Outline

• Setting the scene
  o CASD and InfoChem
  o New motivations

• Retrosynthesis tool: IC\textit{SYNTH}
  o Workflow overview
  o Examples
    ▪ Twistane
    ▪ Oxaspiroketone

• Forward reaction prediction tool: IC\textit{FRP}
  o Workflow overview
  o Application fields
  o Interfaces and examples

• Conclusions
Outline

- Setting the scene
  - CASD and InfoChem
  - New motivations

- Retrosynthesis tool: IC\textit{SYNTH}
  - Workflow overview
  - Examples
    - Twistane
    - Oxaspiroketone

- Forward reaction prediction tool: IC\textit{FRP}
  - Workflow overview
  - Application fields
  - Interfaces and examples

- Conclusions
CASD and InfoChem Over Time

- 1960: InfoChem acquires rights for the SPRESI data
- 1980: Development of ICMAP and CLASSIFY
- 1990: Launch of Synthesis Tree Search (CDROM)
- 1992: Comprehensive synthesis planning using multiple reaction search algorithms
- 2006: Cooperation with AstraZeneca Process Development
- 2007: Defined as "business critical" in process development department
- 2011: Roll-out for 50 scientists at AZ
- 2012: Defined as "business critical" in process development department

Comprehensive synthesis planning using multiple reaction search algorithms
Valentina Eigner-Pitto¹, Josef Eiblmaier¹, Hans Kraut¹, Heinz Saller¹, Peter Loew¹, and Guenter Grethe²

SimBioSys Inc
Introducing Route Designer v1.0
A Peter Johnson¹, Zsolt Zsoldos², Aniko Simon², Darryl Reid², James Law², Yang Liu², Sing Yoong Khew², and Howard Y. Ando³
New Motivations

Extensive reaction databases
- e.g. SPRESI (InfoChem): 4 M reactions

Algorithms / software
- Atom-atom-mapping
- Reaction centre i/d
- Reaction classification
- Fast search engine

Computer hardware
- Powerful
- Inexpensive
- Networked
- Secure
Outline

• Setting the scene
  o CASD and InfoChem
  o New motivations

• Retrosynthesis tool: IC\textit{SYNTH}
  o Workflow overview
  o Examples
    ▪ Twistane
    ▪ Oxaspiroketone

• Forward reaction prediction tool: IC\textit{FRP}
  o Workflow overview
  o Application fields
  o Interfaces and examples

• Conclusions
ICSYNTH: Workflow Overview

1. Pre-processing

2. Synthesis Planning

Target molecule

Transform library

Automatic transform extraction (ICMAP/CLASSIFY)

Reaction database

Lookup stored examples

Precursor(s) 1

Precursor(s) 2

Precursor(s) n

Transform x

Transform y

Transform n

Target molecule

Transform x

Transform y

Transform n
Examples

Publication in Organic Process Research and Development:

Route Design in the 21st Century: The ICSYNTH Software Tool as an Idea Generator for Synthesis Prediction

Twistane
- Since 50 years an interesting problem in synthesis
- Synthesis teaching topic in many organic chemistry courses
- Routinely used during development
- Not really interesting from the process development point of view

Oxaspiroketone
- Problem for AZ Scale-Up Lab in Södertälje
- Late intermediate to a series of potential Alzheimer's treatments
- Medicinal Chemistry synthesis path not suitable for scale-up

Bøgevig ‡, Federsel†, Huerta ‡, Hutchings §, Kraut †, Löw §, Oppawsky §, Rein †, and Saller §
†Chemical Development, AstraZeneca R&D, Macclesfield, SK10 2NA Cheshire, U.K., ‡Chemnotia AB, 151 36 Södertälje, Sweden, §InfoChem GmbH, D-81241 München, Germany
Example: Twistane
Example: Twistane

- 3 of the first 5 suggestions in step 1 correspond to key intermediates in published synthesis
  - reassuring for new users
  - gratifying for software development (in terms of e.g. algorithm and ranking capabilities)
  - ...but largely irrelevant from the idea generation point of view

- New transannular C-C bond formation across decalin precursors involved in:
  - ...most of the new suggested routes
  - ...and in the shortest one reported in literature
  - Not surprising because already historically identified as most strategic bond

- Considerations in precursor ranking within ICSYNTH
  - Application of complexity measure to identify which bond(s) lead to the greatest molecular simplification in the retro direction (bond-breaking), coupled with reactions resulting in multiple bond formation
  - Most favorable of 4 possible C-C breakages in twistane case results in decalin skeleton precursor
  - Less attractive bond-making combination are more complex and lead to more complex synthesis
Example: Oxaspiroketone

Med Chem route not suitable for scale-up:
- Yields not reproducible and < 10%
- Some reagents not optimal (TMSCN, LiHMDS)
- Tedious chromatographic separation from mixture

Additional task aspect:
- Correct stereoisomer generation, critical for biological activity of final compound
- At least easy isolation method of desired isomer

Disconnection of bonds a, b, c and d studied:
- a and b disconnections appealing at first sight
- ...but use of non-desired reagents involved

ICSYNTH unconventional suggestion:
- Non intuitive disconnection of c in last step
Example: Oxaspiroketone

Details

Suggestion 17

Represents a group of 3 ensembles with common target bond changes.

Rating: 108477

Suggested Reaction

Precedent Reactions 15 citation(s)


Yield: 85%; Similarity: 86%; Conditions: 1 h, 50 degree, DMF, potassium acetate;
Example: Oxaspiroketone

- Initial conventional retrosynthetic analysis could not identify the illustrated bond connection.
- Although the reported precedent could not be used as such, a new route could be readily developed.

Conclusion: IC SYNTH used as idea generator identifies unconventional and unusual transformations that lead to not expected but in this case highly relevant solutions.
**Outline**

- Setting the scene
  - CASD and InfoChem
  - New motivations

- Retrosynthesis tool: IC\textit{SYNTH}
  - Workflow overview
  - Examples
    - Twistane
    - Oxaspiroketone

- Forward reaction prediction tool: IC\textit{FRP}
  - Workflow overview
  - Application fields
  - Interfaces and examples

- Conclusions
Software Reaction Prediction Tools: IC SYNTH, IC FRP

Target

Starting material

Targeting Reaction Prediction Tools: IC SYNTH, IC FRP
ICFRP: Workflow Overview

1. Pre-processing
   - Automatic transform extraction (ICMAP/CLASSIFY)
   - Remove bond 1-3
   - Make new single bond between atoms 3 and 5
   - Remove bond 1-2
   - Make new single bond between atoms 3 and 4

2. Reaction Prediction
   - Starting material
   - Transform library
   - Lookup stored examples
   - Knowledge-based compound design

Reaction database
   - e.g. SPRESI
   - proprietary databases
   - commercial databases
ICFRP Application Fields

Reactivity mapping
Overview of which positions in the core can be modified, and how

Scaffold modification
Structural changes of the core structure

Library synthesis
Broad range of possible modifications (including, e.g., bioisosteres) of a given position in the molecule

Different outputs obtained by tailoring the query strategy parameters
ICFRP User Interfaces

Knime node

Web-application

Pipeline Pilot component
Example 1: Knowledge-based Drug Design

- Proof of concept on known molecule: Diazepam

80 new suggested molecules (with related synthesis reactions)

2564 known related compounds (with associated physicochemical properties)
Example 1: Knowledge-based Drug Design

- Fingerprints calculated
- Identical compounds filtered off (Tanimoto = 1)
- 13 suggested products with a Tanimoto 0.7-0.9
- Physicochemical properties calculated and compared with reported ones

New “suggested” molecules show similar properties to the known ones
Example 2: Knowledge-based Drug Design

- Proof of concept inspired by “de novo design using reaction vectors”

Penicillin G
Example 2: Knowledge-based Drug Design

- **ICFRP** with *medium* precision

- **ICFRP** with *high* precision
Example 3: Multistep Synthesis Starting from a Building Block
Outline

- Setting the scene
  - CASD and InfoChem
  - New motivations

- Retrosynthesis tool: IC SYNTH
  - Workflow overview
  - Examples
    - Twistane
    - Oxaspiroketone

- Forward reaction prediction tool: IC FRP
  - Workflow overview
  - Application fields
  - Interfaces and examples

- Conclusions
Conclusions

ICSYNTH is used as idea generator for retrosynthetic analysis helps identifying synthesis routes that go beyond known chemist-derived suggestions

- It is a unique predictive tool for route design in Chemical Development at AstraZeneca
- Unwanted/chemically incorrect suggestions generate fastidious noise
- Improvements under development:
  - Stereo-chemistry handling
  - Better chemical selectivity and reaction conflict predictions
  - New algorithm to enable starting material-based strategies

"Younger" ICFRP has proved its potential in assisting research scientists in drug discovery and drug design by generating new valuable candidates in regards of:

- novelty
- chemical diversity
- synthesizability
Conclusions

Retrosynthesis, synthesis planning and forward reaction prediction: When will computers meet the needs of the synthetic chemist?
Acknowledgements

- H. Saller, M. Hutchings (ICSYNTH)
- AZ Process Development Team
- F. Huerta (ICFRP)
- The InfoChem Team