

Boehringer Ingelheim Comprehensive Library of Accessible Innovative Molecules (BICLAIM)

Uta Lessel
Computational Chemistry



Contributors

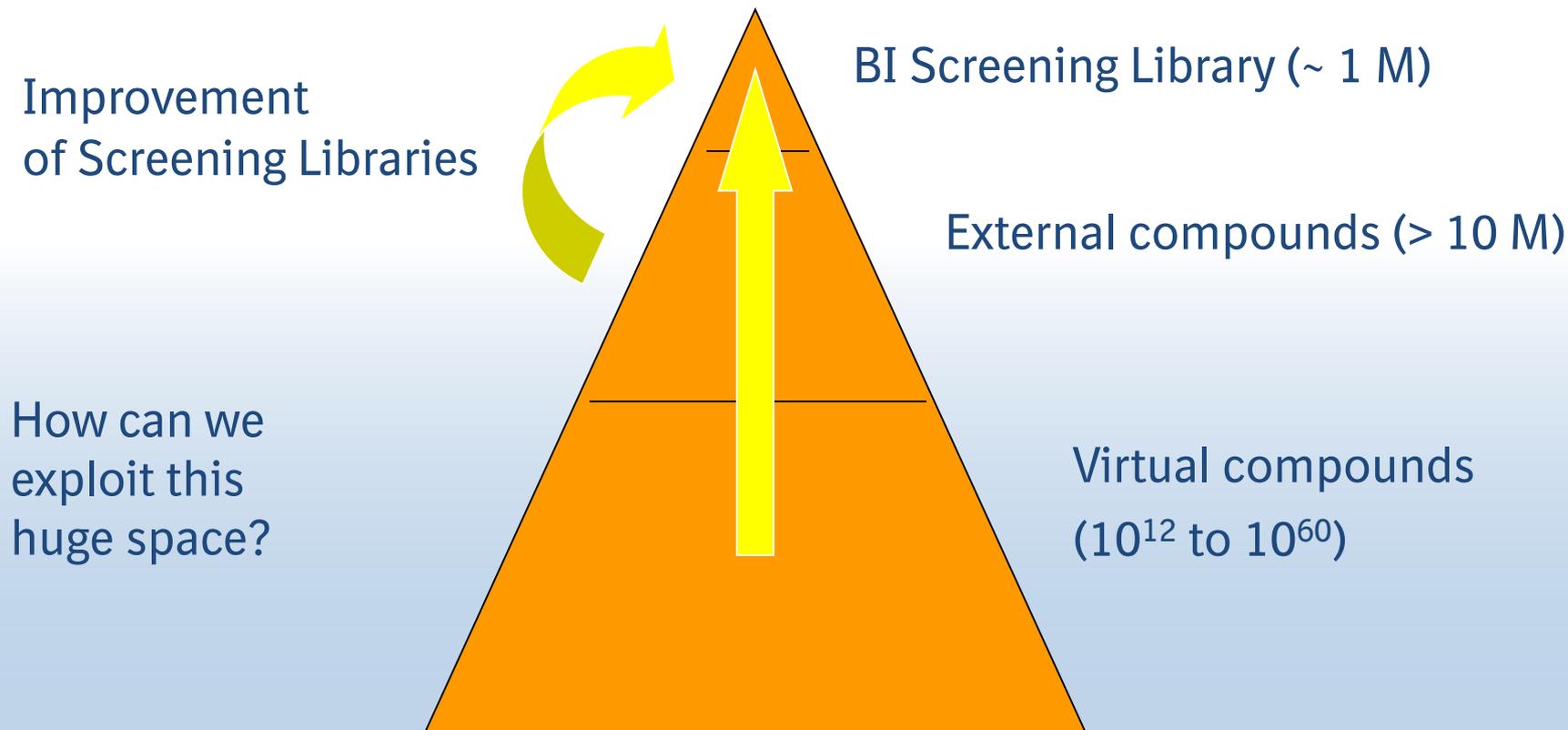
Michael Bieler

Christoph Hoenke

Alexander Weber

CompChem group
and former colleagues

Introduction: The Chemical Universe



Different approaches

- Diverse methods for De novo Design
- Recombination of fragments generated according to RECAP rules
- ...

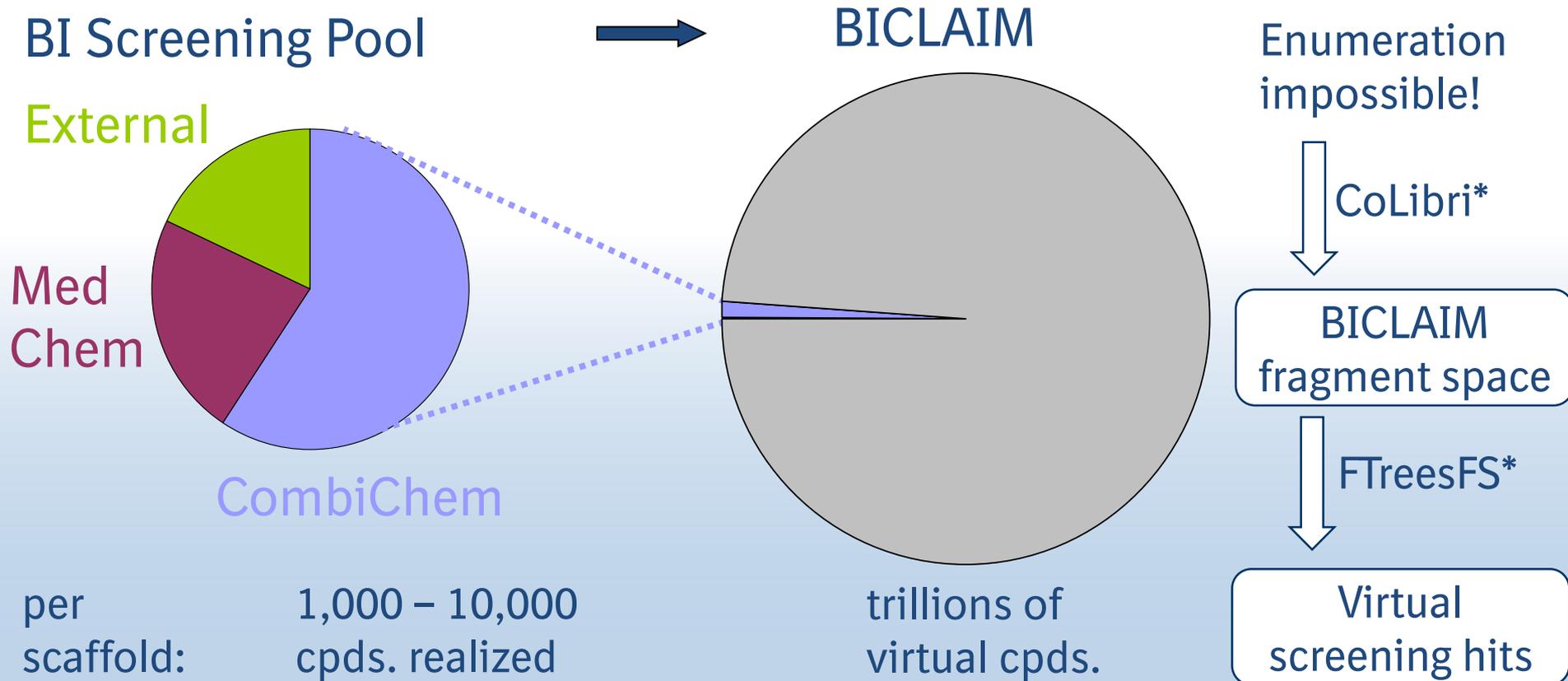
but lacking chemical feasibility of most products is a severe disadvantage

Goal: Get access to compounds with known chemical feasibility by VS

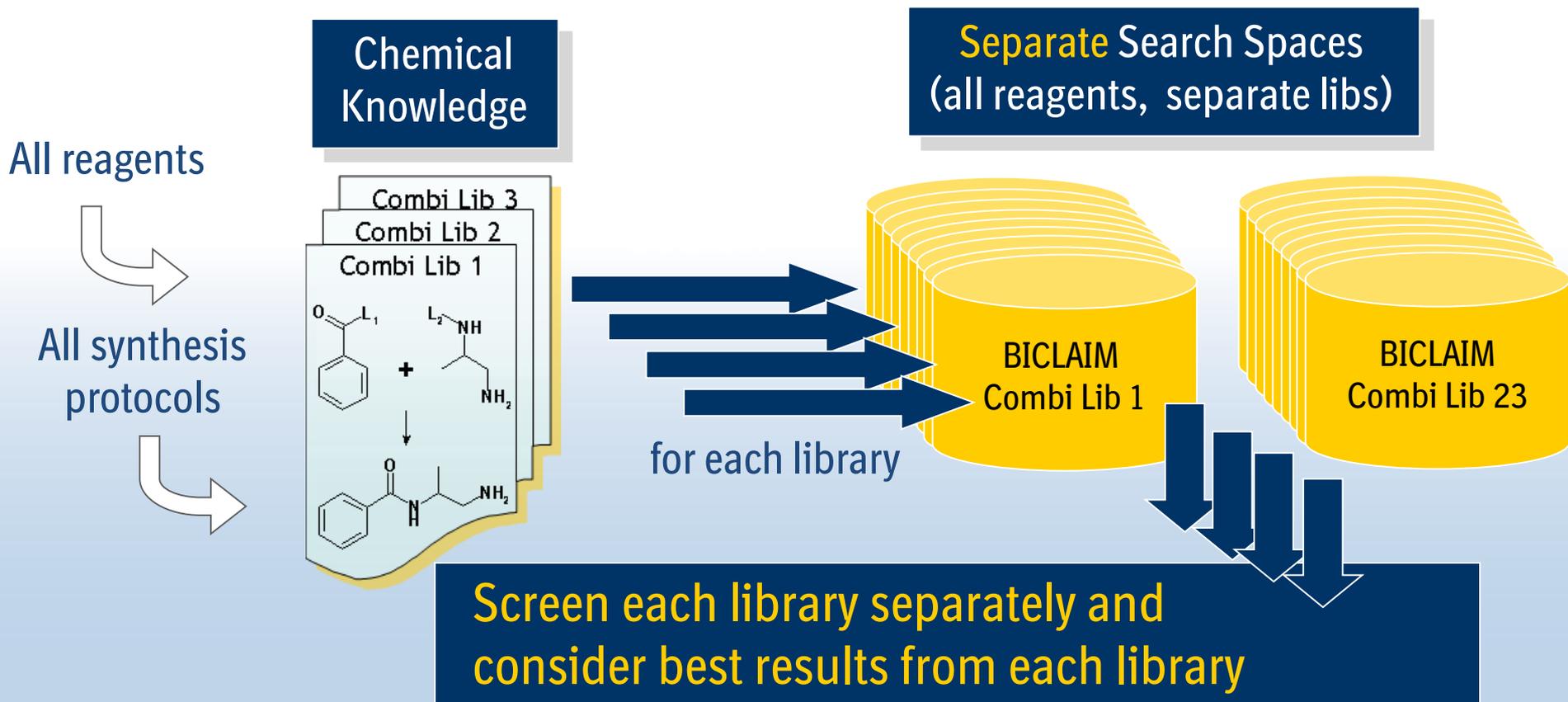
References to published solutions by university research groups and other companies are summarized in the backup.

Basic Idea

*: BioSolveIT, St. Augustin



Reagent-based Generation of Virtual Libraries

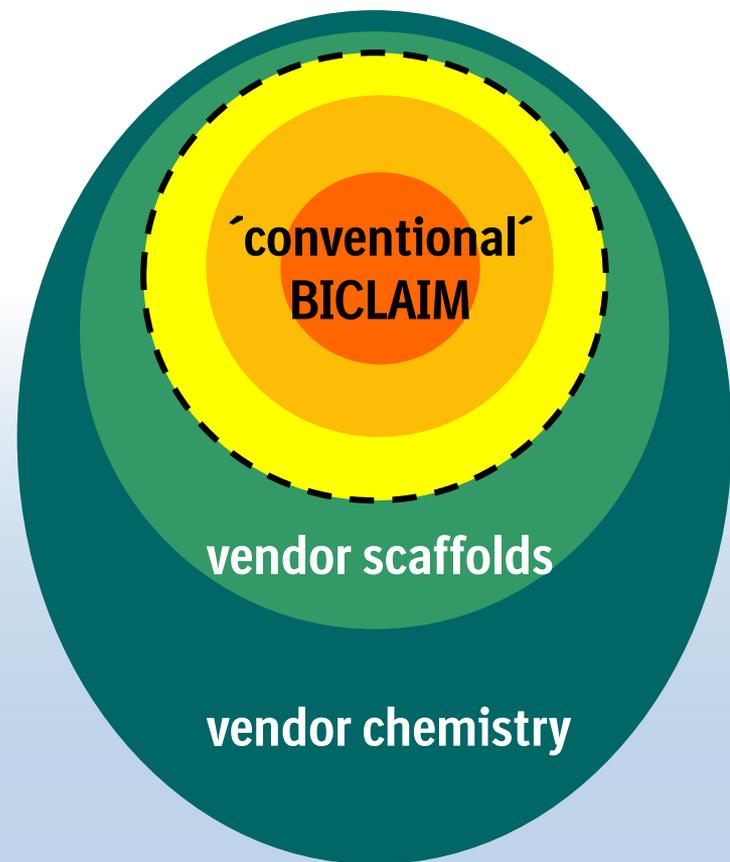
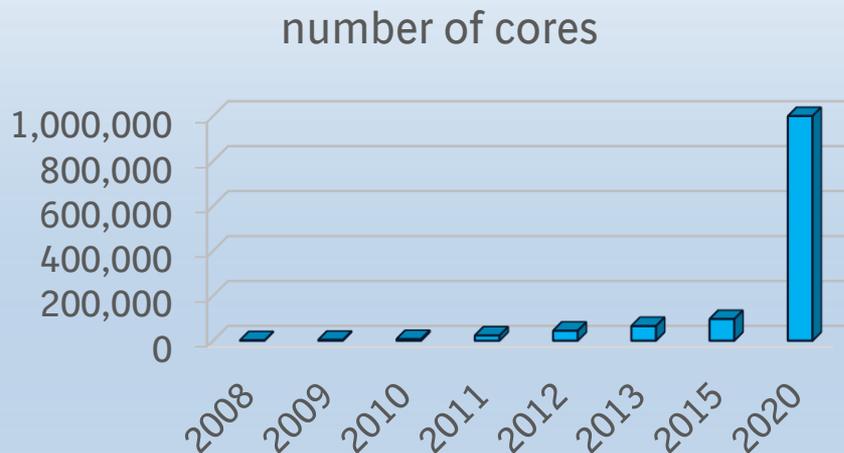


Development of the BICLAIM Space

2008: ~ 1,600 cores from combinatorial libraries
~ 30,000 reagents
encoding ~ 5×10^{11} cpds.

systematic growth by:

- inclusion of external reagents
- expansion to cores from additional internal and external sources



Workflow @ BI

Target Identification
Assay Development

Lead
Identification

Lead
Optimization

Development

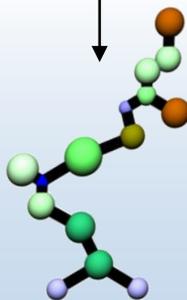
Reference

HTS hit

Lead structure

Fragment
Spaces

Query



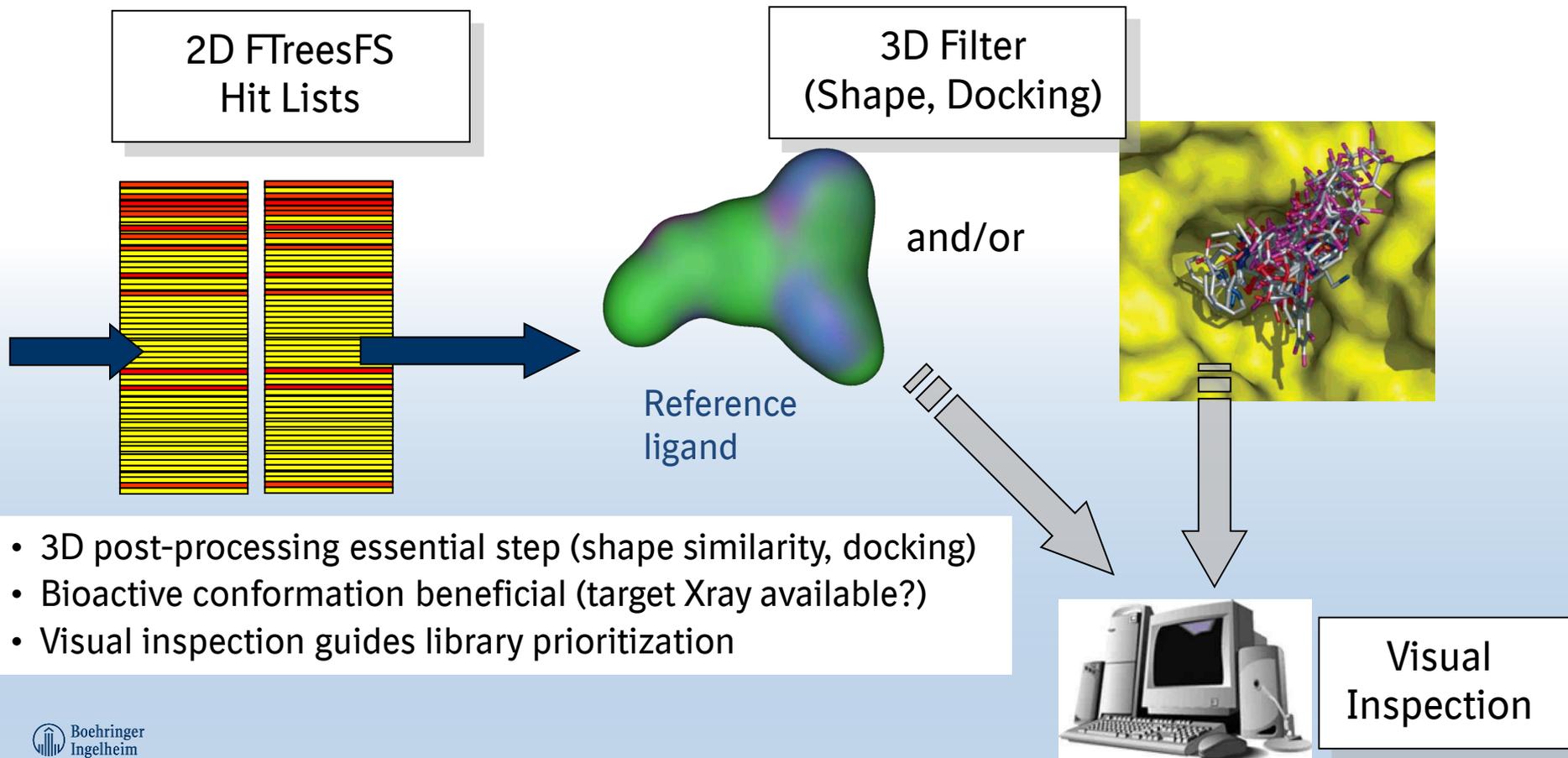
FTreesFS

2D similarity to query molecule

2D FTreesFS
Hit Lists

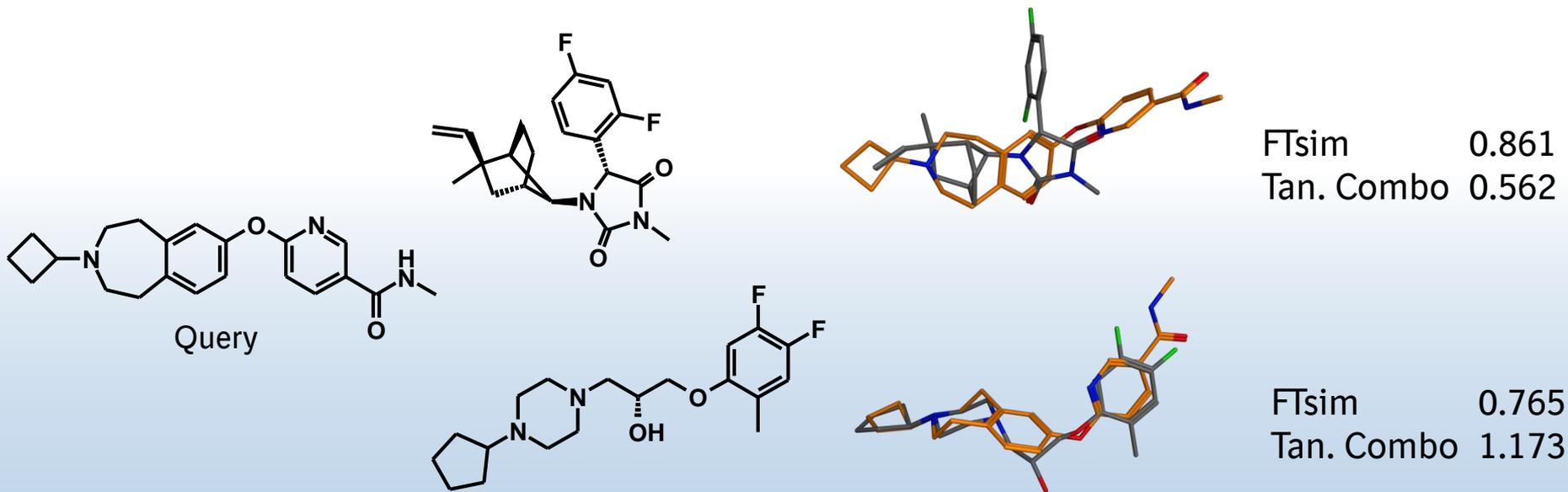
$\sim 10^5$ cpds.

3D Post processing



Why is a post-processing step needed?

FTrees are 2D descriptors, but finally 3D matches are necessary



3D alignments serve to select the best matching scaffolds and help to convince chemists to synthesize focused libraries

Tool Box

Hit
generation,
e.g.:

- FTrees-FS search
- 2D hits: Corina -> Omega -> ROCS alignments
- sort hits by TanimotoCombo
- visual inspection

Hit
characte-
rization

Database linking CoreID with

- Synthesis protocol
- Starting material and its availability
- Existing compounds with the same scaffold

3D alignments and information about chemical feasibility strongly support prioritization of scaffolds in the project teams

General remarks

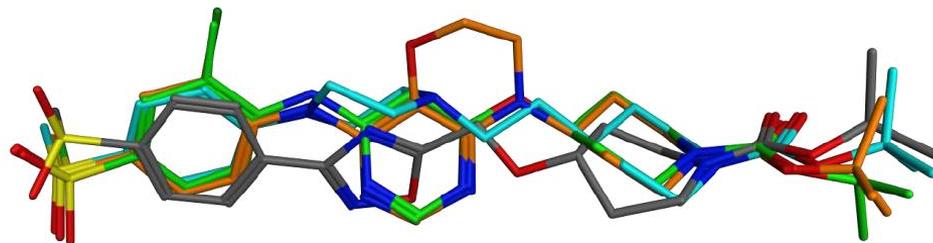
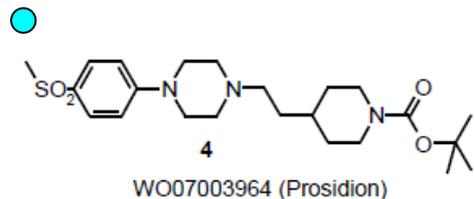
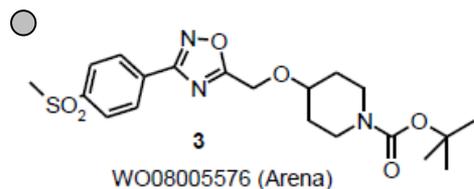
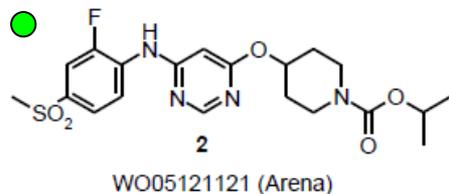
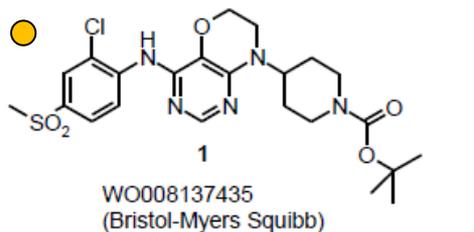
- Powerful procedure for the **detection of new leads**
- **Careful generation of fragment space** huge upfront investment but provides synthesis route for any virtual hit
- Inclusion of library ideas and poorly evaluated chemistry as well as reagents with unknown availability influences **timelines**
- Cores taken up should be **broadly explored**
- Include **project specific knowledge** as far as possible

- BI:
- Key element in lead identification strategy
 - Applied on a routine basis:
before, in parallel, instead HTS or in LO phase depending on project needs (different knowledge which can be considered, timelines, expectations)
 - Leads for many projects successfully provided

Application example (GPR119 agonists)¹³:

Query and Bioactive Conformation Hypothesis

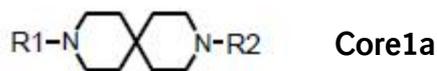
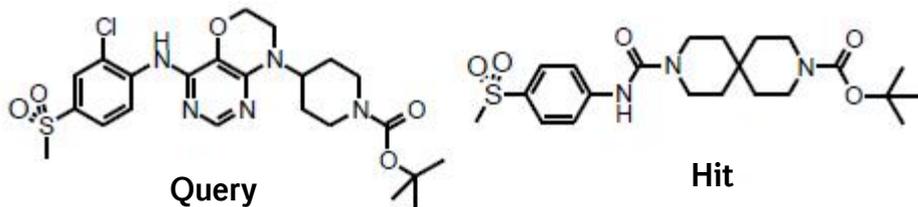
Competitor
GPR119
agonists:



3D alignment suggests
similar binding mode

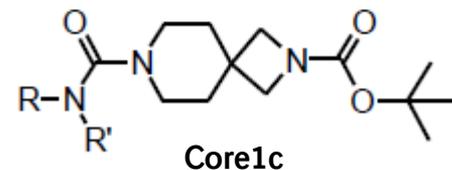
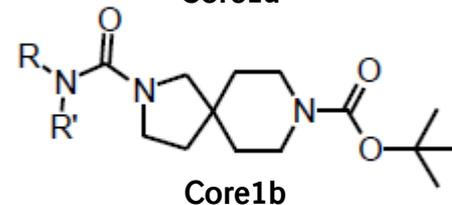
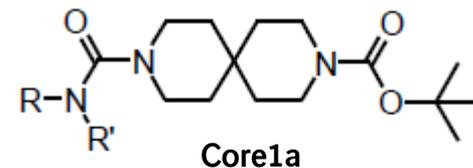
- FTreesFS search for each of the queries
- Merge and unify hit lists
- 3D alignment on most rigid query (1)
- Visual inspection
(focus on activity anchors)

Selected Virtual Hits I



R1: from aldehydes, carboxylic acids, isocyanates, carbamoyl chlorides
R2: from aldehydes, carboxylic acids, isocyanates, carbamoyl chlorides

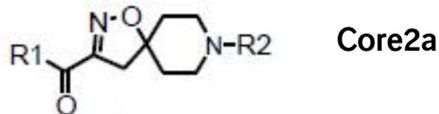
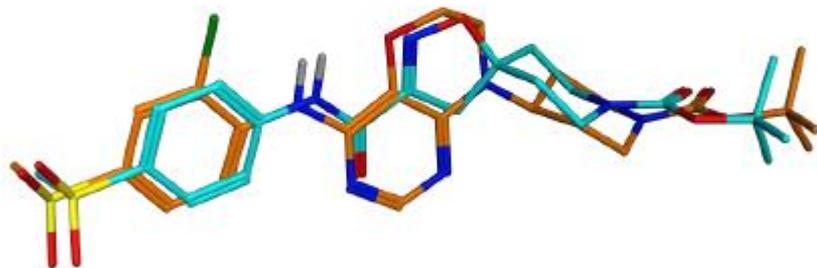
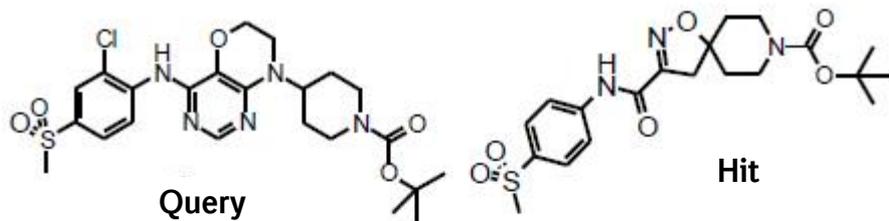
Variants:



Library design:

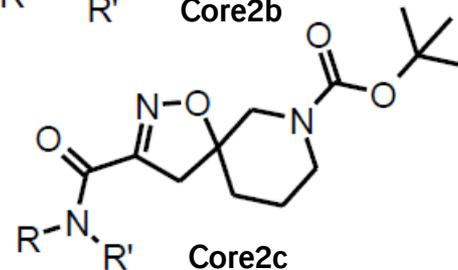
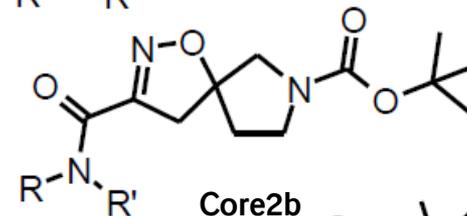
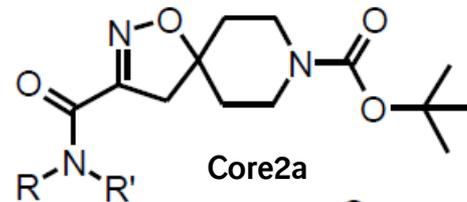
keep carbamoyl moiety and choose reagents similar to the methylsulfonyl-aryl moiety

Selected Virtual Hits II



R1: from primary amines, secondary amines, primary anilines
R2: from aldehydes, carboxylic acids, isocyanates, carbamoyl chlorides

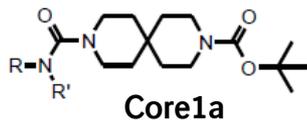
Variants:



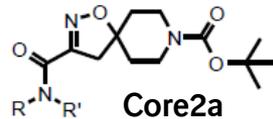
Library design:

keep carbamoyl moiety and choose amines similar to the methylsulfonyl-aryl moiety

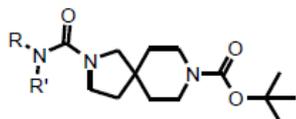
Syntheses



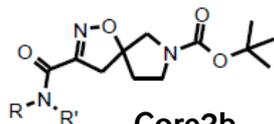
51 cpds.: 5 cpds. < 10 μ M



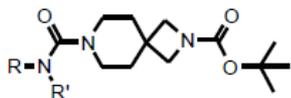
89 cpds.: 38 cpds. < 10 μ M



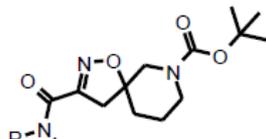
55 cpds.: all inactive



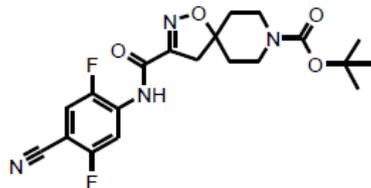
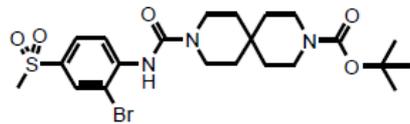
87 cpds.: all inactive



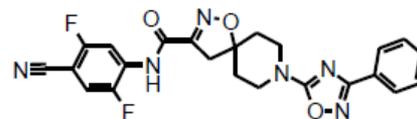
56 cpds.: all inactive



84 cpds.: all inactive

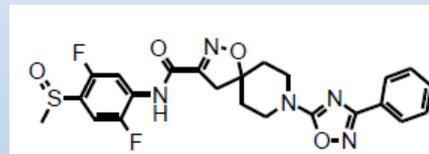


Most promising compound
from first follow up library:



$EC_{50} = 30$ nM
(173% IA)

Most promising compound
from second follow up library:



$EC_{50} = 21$ nM
(160% IA)



Start of LO

Summary and Conclusions

- (FTrees) fragment space searches represent a powerful tool for the detection of new leads
- Success depends on
 - quality of the fragment space
 - combinatorial chemistry know how
 - sufficient synthesis and test capacity
 - matching timelines for virtual screening, synthesis and testing including one to two follow up cycles

and most important

a good team effort including computational, combinatorial, and medicinal chemists

BACKUP

Published solutions

E.g.:

- G. Schneider et al.: De Novo Design based on fragment space from WDI using CATs descriptors (Topas and further developments) [1-3]
- FLEXNovo: FlexX Docking in fragment spaces [4]
- Tripos: synthons + universal reactions (ChemSpace techn., AllChem) [5-7]
- Nikitin et al.: fragment space encoding 10^{13} cpds. from 400 comb. libraries [8]
- Pfizer: FTrees fragment space based on 358 comb. synthesis protocols [9]
extension to parallel synthesis [10]
- AZ : fragment space from an analysis of their Elab [11]
- Eli Lilly: The Proximal Lilly Collection [14]
- Evotec: partnership with BiosolveIT announced to create FTrees fragment space
- Merck KGaA: MASSIV (2018)
- BiosolveIT: KnowledgeSpace, RealSpace, Galaxy
- GSK: GSK space

References

Published solutions:

- [1]: Schneider, G., et al., *Angew. Chem. Int. Ed. Engl.*, **39**, 4130-4133 (2000)
- [2]: Schneider, G., et al., *Angew. Chem. Int. Ed. Engl.*, **38**, 2894-2896 (1999)
- [3]: Hartenfeller, M., et al., *Chem. Biol. Drug Des.*, **72** (1), 16-26 (2008)
- [4]: Degen, J. and Rarey, M., *ChemMedChem*, **1**, 854-868 (2006)
- [5]: Cramer, R.D., et al., *J. Chem. Inf. Comput. Sci.*, **38**, 1010-1023 (1998)
- [6]: Andrews, K.M. and Cramer, R.D., *J. Med. Chem.*, **43**, 1723-1740 (2000)
- [7]: Cramer, R.D., et al., *J. Comput.-Aided Mol. Des.*, **21**, 341-350 (2007)
- [8]: Nikitin, S., et al. *J. Comput.-Aided Mol. Des.*, **19**, 47-63 (2005)
- [9]: Boehm, M., et al., *J. Med. Chem.*, **51**, 2468-2480 (2008)
- [10]: Hu, Q., et al., *ACS Comb. Sci.*, **14** (11), 579-589 (2012)
- [11]: Vainio, M.J., et al., *J. Chem. Inf. Mod.*, **52** (7), 1777-1786 (2012)
- [12]: Lessel, U., et al., *J. Chem. Inf. Model.*, **49**, 270-279 (2009)
- [13]: Wellenzohn, B., et al., *J. Med. Chem.*, **55**, 11031-11041 (2012)
- [14]: Nicolaou, C. A., et al., *J. Chem. Inf. Model.*, **56**, 1253-1266 (2016)

Useful reactions:

- Hartenfeller, M., et al., *J. Chem. Inf. Model.*, **51**, 3093-3098 (2011)
- Chevillard, F. and Kolb, P., *J. Chem. Inf. Model.*, **55**, 1824-1835 (2015)