CCP5 Summer School

July 2023

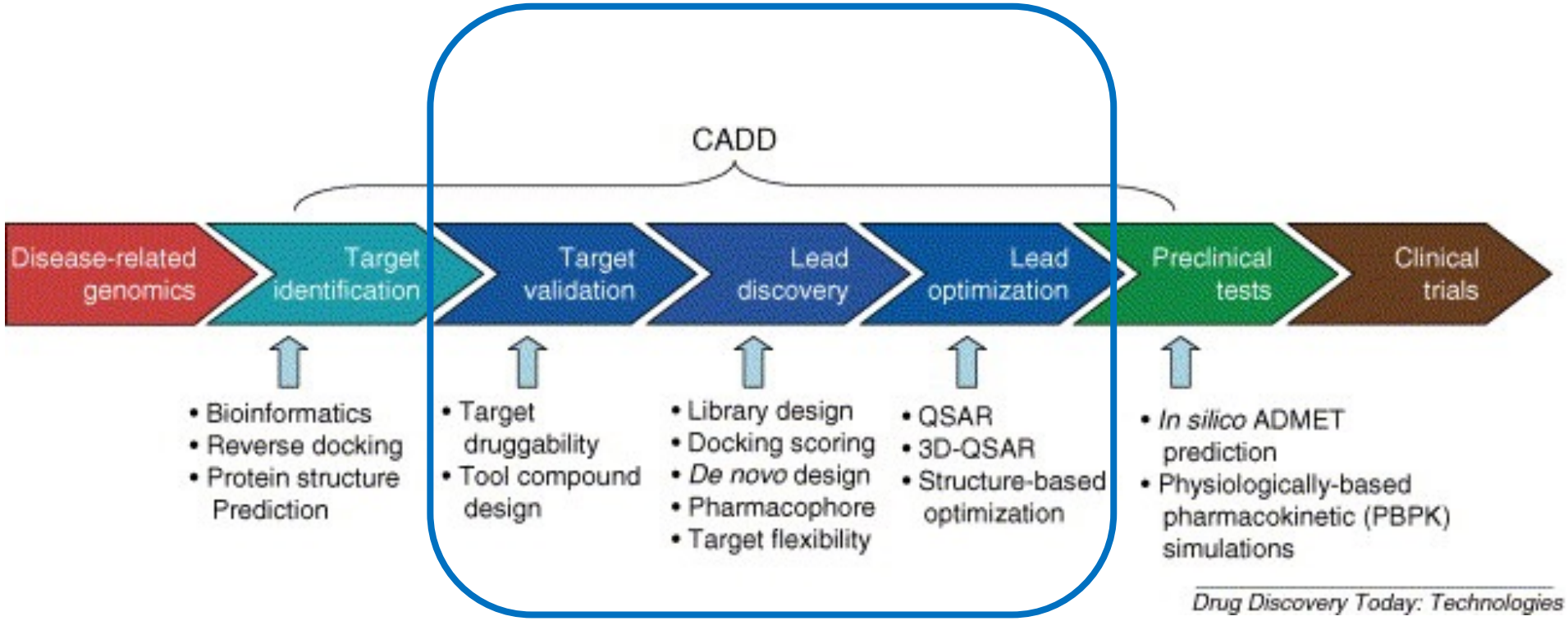
Simulations of biomolecules: design of small molecule complement C5 inhibitors

DR. AGNIESZKA K. BRONOWSKA

SCHOOL OF NATURAL ENVIRONMENTAL SCIENCES

NEWCASTLE UNIVERSITY

CADD pipeline



Molecular docking

The process of predicting the stable 3D geometry
of an interacting pair of molecules - a ***binding mode/pose***
Energy-based docking relies on molecular mechanical force fields (energy minimisation, sampling, scoring function)

Molecular docking: two main tasks

- **Sampling** of ligand conformational space and pose generation (geometry)

- **Scoring** protein-ligand complexes (energetics)

Molecular docking: to-do list

Problem: a pair of molecules represented by their 3D coordinates

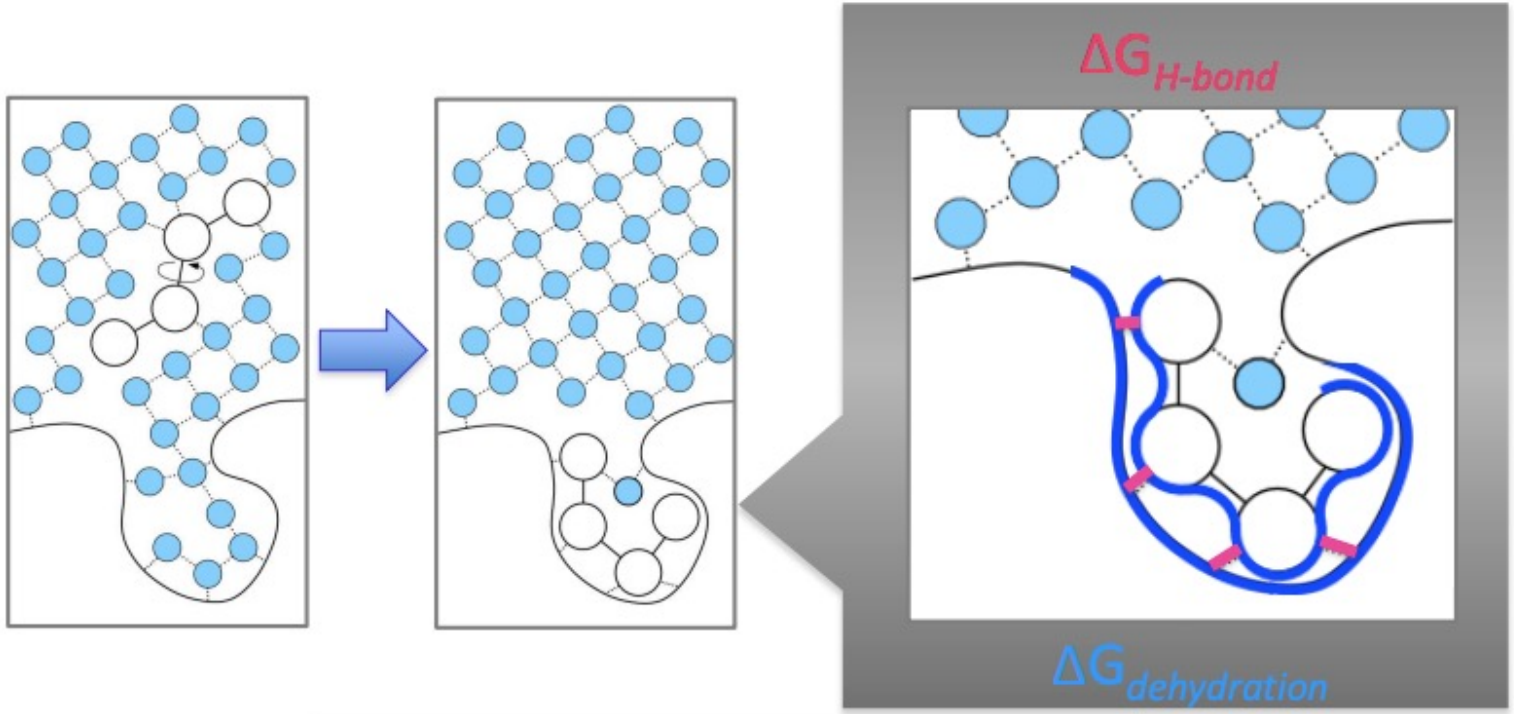
- Decide whether the molecules will form a complex;
- Determine the binding affinity (free binding energy);
- Predict the 3D structure of the complex (binding mode);
- Deduce function (agonist/antagonist);

SeeSAR

SeeSAR

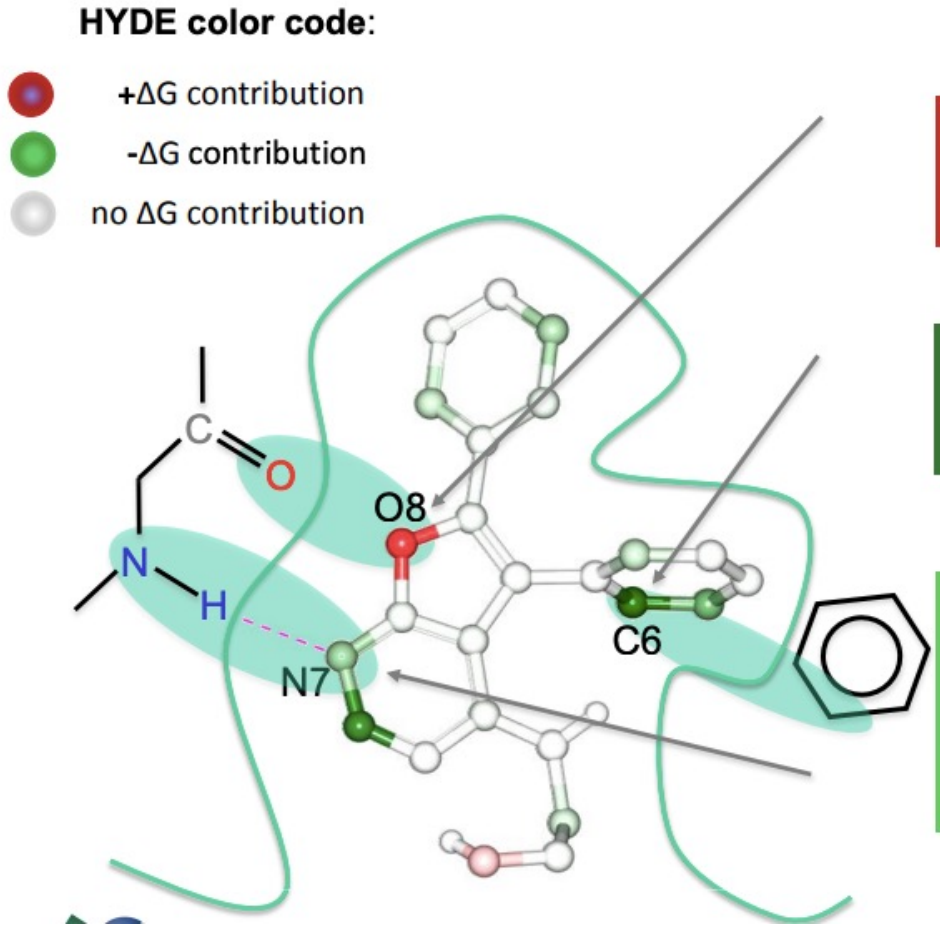
- Platform-independent
- Easy to use (for basic functionalities) – developed for bench chemists
- Decent scoring function (interactions + solvation term used)
- Minimum prep and requirement for e.g. format parsing
- Rapid core expansion, core replacement and analogue search enabled
- Handles multiple file formats
- Covalent docking enabled
- Integrated with other tools, e.g. InfiniSee (chemical space search), RDKit, KNIME

HYDE scoring function: the concept



$$\Delta G^{i}_{\mathrm{HYDE}} = \sum_{atom\ i} \Delta G^{i}_{dehydration} + \Delta G^{i}_{H-bond}$$

HYDE – visual affinities



receptor carbonyl oxygen	8.2 kJ/mol
ligand aromatic oxygen	2.4 kJ/mol
total desolvation cost	**10.6 kJ/mol**

receptor aromatic carbons	-5.2 kJ/mol
ligand aromatic carbon	-2.0 kJ/mol
total desolvation gain	**-7.2 kJ/mol**

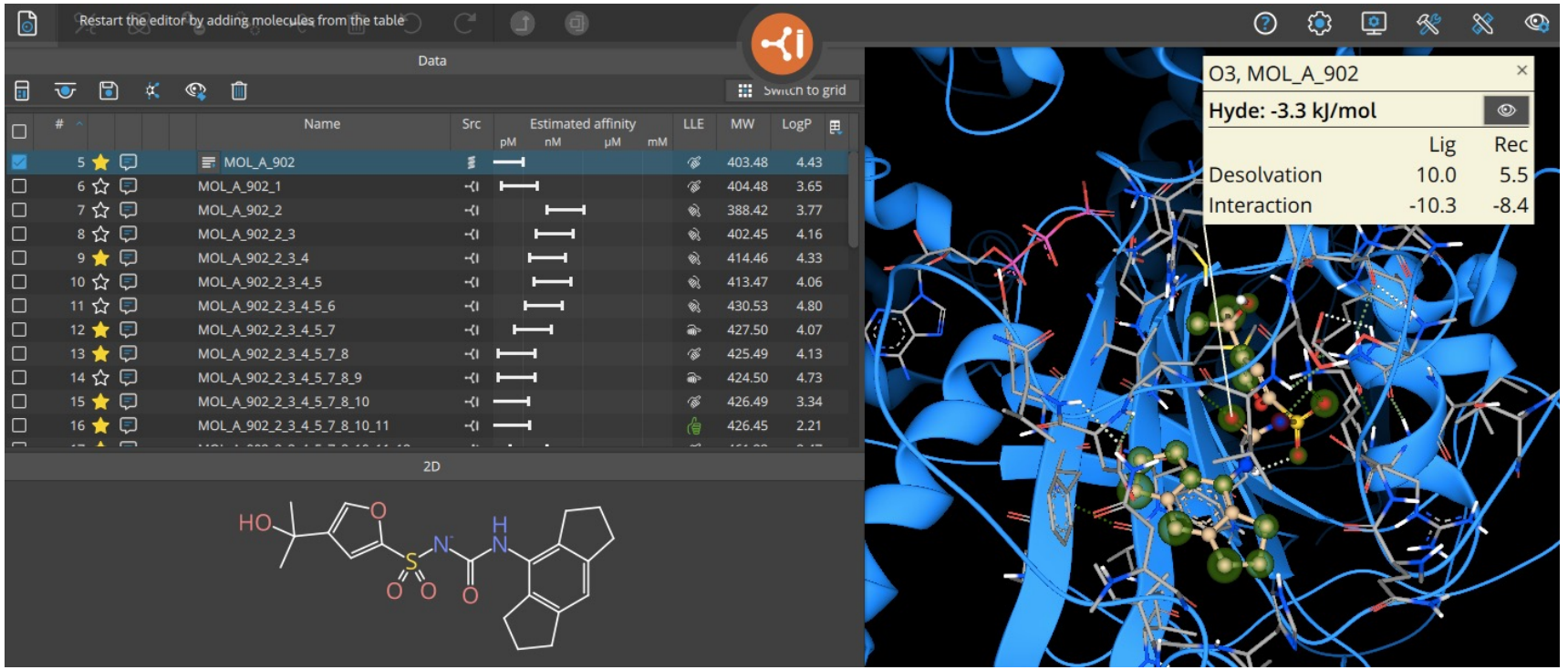
receptor amide N dehydrat	6.3 kJ/mol
interaction energy	-7.4 kJ/mol
ligand aromatic N dehydrat	6.4 kJ/mol
interaction energy	-7.5 kJ/mol
total H-bond energy	**-2.2 kJ/mol**

SeeSAR: a working example

Ligand moves (optimisation after every modification), protein fixed: single conformation

Exhaustive conformational search on ligands prior to docking

Scoring function (HYDE) gives contributions from intrinsic interactions and desolvation



MCC950 bound to NLRP3 inflammasome

Predicted affinity range: High pM to low nM

Ki (SPR and MST): Low nM (8 – 24 nM)

Limitations and some known issues

- Protein is considered rigid: ideally, you should follow your calculations by running short MD simulations on the complexes and recalculating the binding affinities
- Results are very sensitive to even small changes in the conformation of the protein
- Binding affinities for certain groups are not reproduced well: hydrophobic effect tends to be overestimated, while highly polar groups are underestimated
- Every now and then, weird protonation states suggested (you can always manually adjust) and med-chem nonsense molecules suggested in core expansion in Inspirator
- Workflows are limited to small molecules
- Med-chem properties and/or synthetic feasibility of suggested analogues may be problematic
- Technical: 50,000 compounds/rows in SeeSAR GUI (you need to use KNIME to “downsize” very large data sets, or use non-GUI version)

Virtual Screening

Virtual screening using SeeSAR and InfiniSee

- Docking of a large virtual libraries (1,000+) of compounds
- Libraries: collections of small molecules for virtual screening
- Sources: open-access (e.g. ZINC, ChEMBL) or commercial (e.g. Enamine) virtual libraries
- Types of libraries commonly used: fragments, diversity sets, target-focused, custom

Fragment libraries: Enamine

Fragment Collection

MiniFrag Library

sp^3 Rich Fragments

Single Pharmacophore

Carboxylic Acid Fragments

Fluorinated Fragments

Covalent Fragments

Warhead Subsets

Cysteine focused Covalent Fragments

Serine focused Covalent Fragments

Lysine focused Covalent Fragments

Acrylamides

Enamine Essential Fragment Library

- 320 fragments
- Universal tool for initial screen of novel targets
- Designed in collaboration with research group at University of Cambridge
- All fragments have been tested for water solubility and chemical stability in buffer solution
- Increased hit probability: the structures are based on frequently reported fragment hits and scaffolds derived from experimentally determined structures of protein-ligand complexes

Diversity libraries

DIVERSITY LIBRARIES

10 240 compounds

50 240 compounds

Hit Locator Library

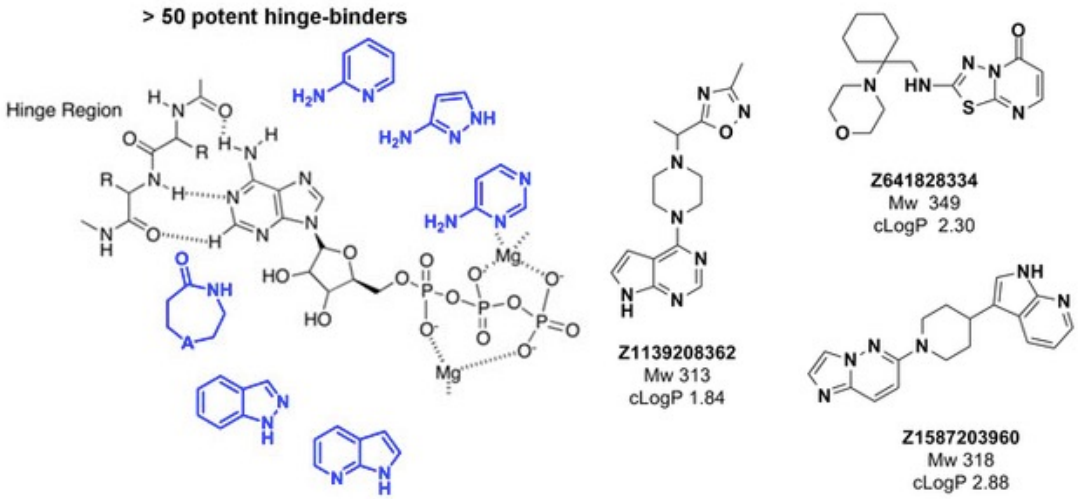
Phenotypic Screening Library

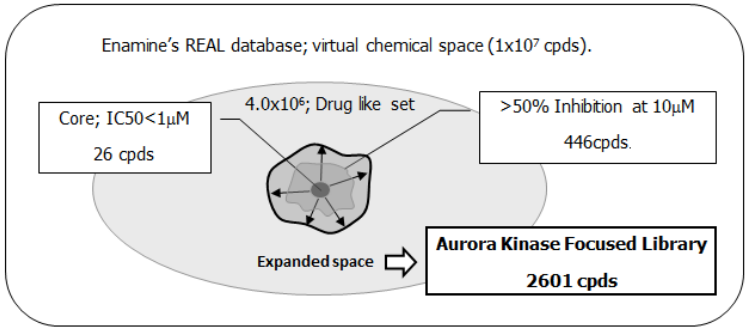
Covalent Screening Library

Target-focused virtual libraries

Kinase libraries:

- Hinge binders (18,000 molecules)
- Aurora A targeted set (2,600 molecules)
- Synthetically feasible
- Comply with drug-like rules
- Evaluated *in vitro*





Building your own set: InfiniSee

- Developed from REAL Space Navigator;
- REAL: 21 bln make-on-demand compounds;
- The current version (Artemis) has a choice of several massive libraries (REAL, CHEMriya, GalaXi, eXplore, KnowledgeSpace)

Size matters: screening ultra-large libraries

Ultra-large library docking for discovering new chemotypes

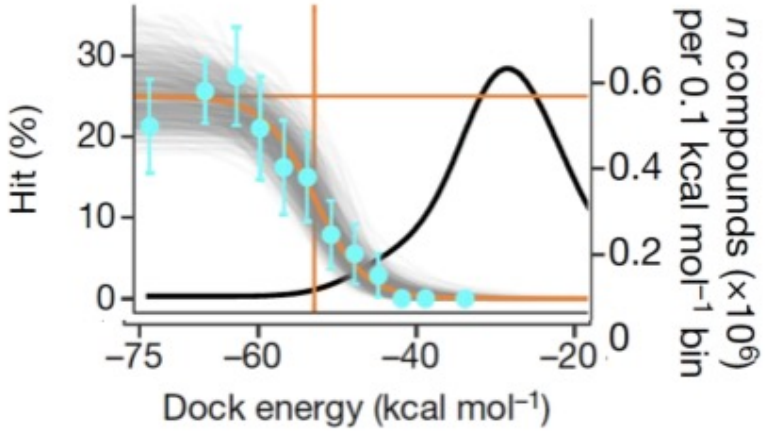
Jiankun Lyu[1,2,10], Sheng Wang[3,4,10], Trent E. Balius[1,10], Isha Singh[1,10], Anat Levit[5], Yurii S. Moroz[5,6], Matthew J. O'Meara[1], Tao Che[4], Enkhjargal Algaa[1], Kateryna Tolmachova[7], Andrey A. Tolmachev[7], Brian K. Shoichet[1]*, Bryan L. Roth[4,8,9]* & John J. Irwin[1]*

Nature. 2019; 566(7743): 224–229.

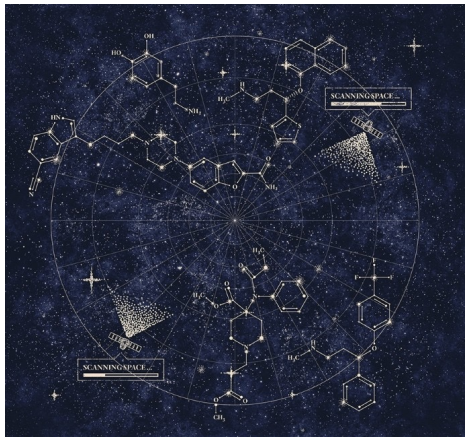
RESEARCH HIGHLIGHT | 05 January 2022

Screening ultra-large virtual libraries

Synthesis-on-demand compounds were chosen from 12 different docking score bins (cyan)

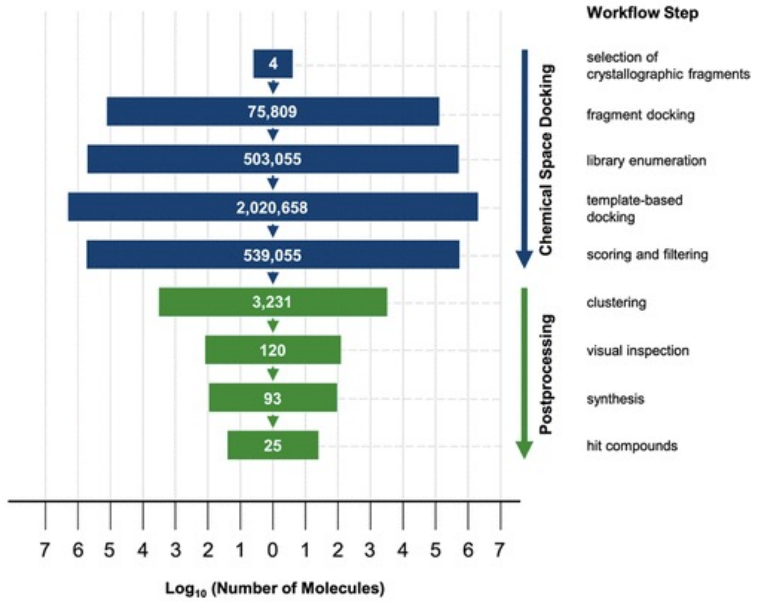


A group of new (no precedent) AmpC inhibitors has been identified

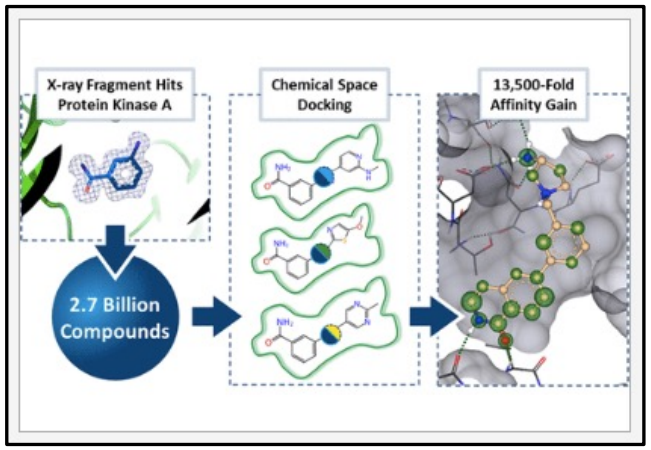


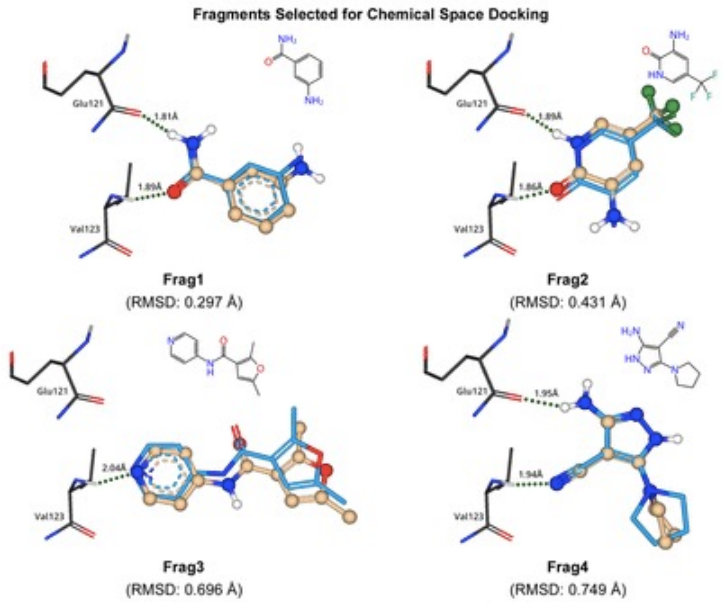
Message: we should aim to explore as much of the chemical space as possible

Screening of large chemical spaces: “crystals first”



Start: 4 PKA-fragment complexes (X-ray)



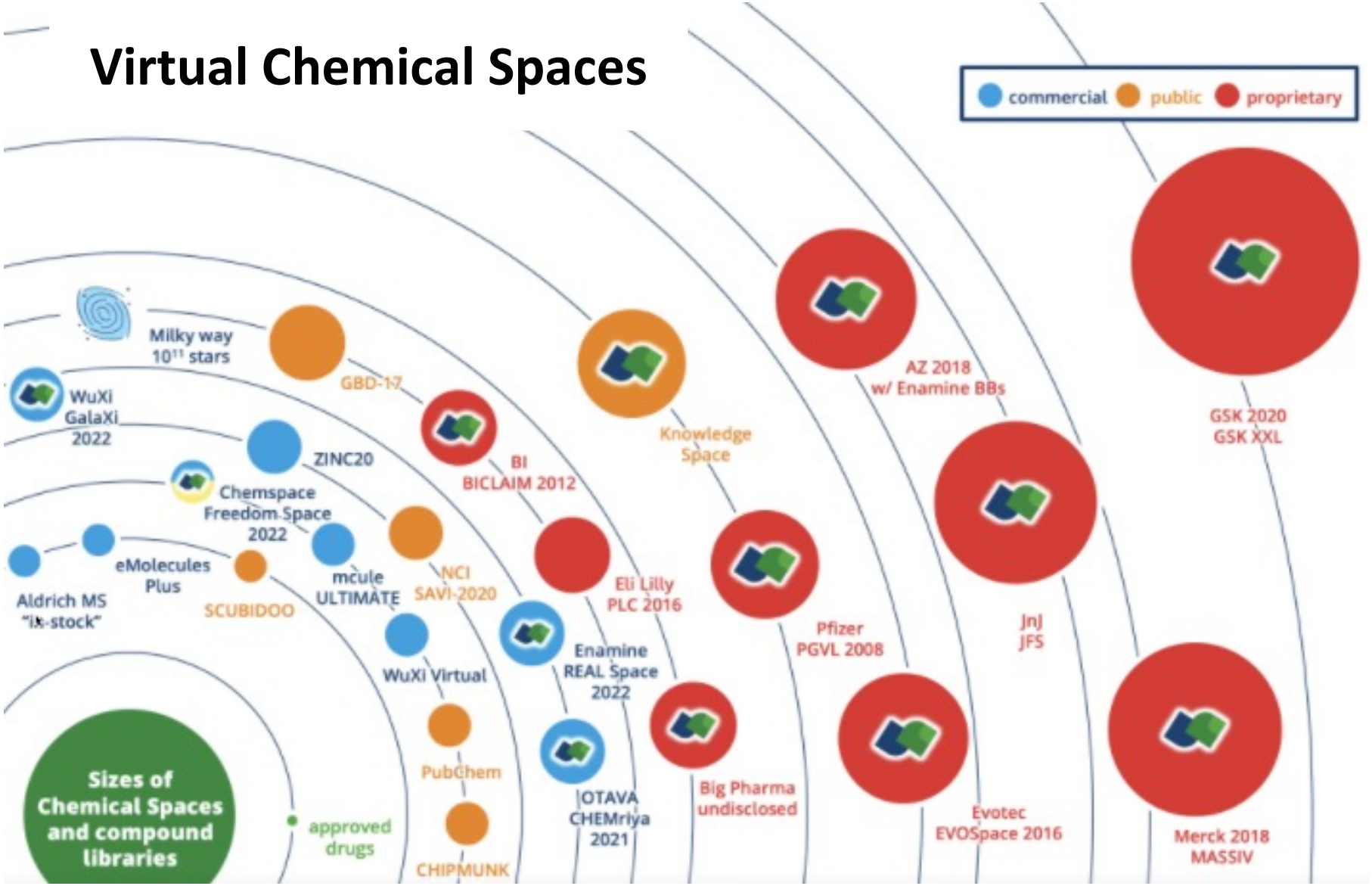


- Template-based docking using Enamine REAL space (20 bn)
- 93 molecules out of 106 selected compounds synthesized
- 40 compounds were active in at least one validation assay
- Most active follow-up having a 13,500-fold gain in affinity

Crystal structures for 6 of the most promising binders were obtained, verifying the predicted binding modes

Müller *et al.*, JCIM 2022

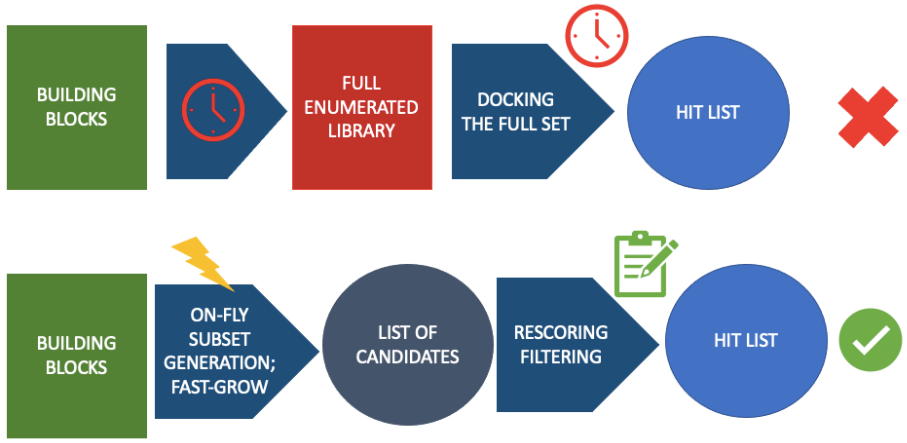
Virtual Chemical Spaces



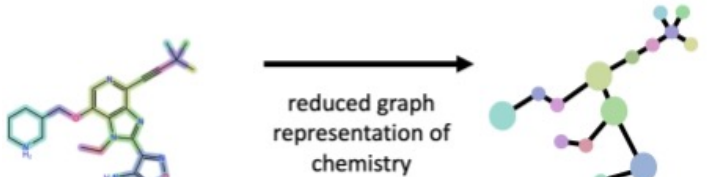
The ”combinatorial challenge”

- Enumerating the GSK space would take 400 Yottabytes (400,000000000000 GB)
- Estimated to take ~11 trillion years to download using high-power broadband
- Storage of all products of a library of such size is unfeasible
- **Way out:** storing building blocks and reaction rules allowing on-fly generation of products

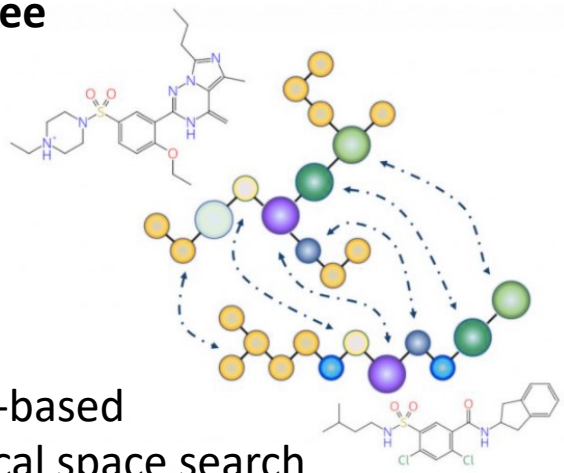
Virtual library	Size
GalaXi	10^9
Enamine REAL	10^{10}
eXplore	10^{12}
BiosolveIT KnowledgeSpace	10^{14}
GSK XXL	10^{26}



Three-stage workflow for ulvHTS



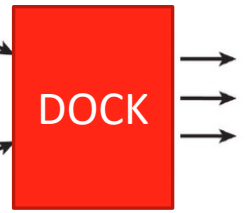
InfiniSee

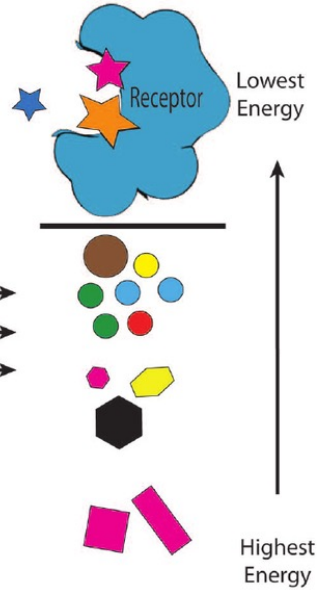


Ligand-based
Chemical space search
Finding “hidden” similarities

Primary

RockDock



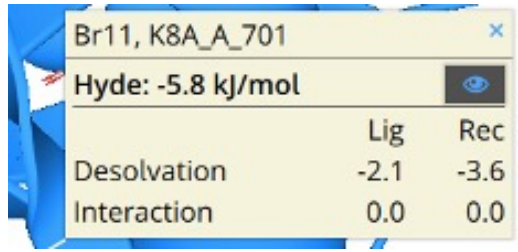


slurm
workload manager

Structure-based
Fast
Hits to be scored by more accurate methods

Secondary

SeeSAR

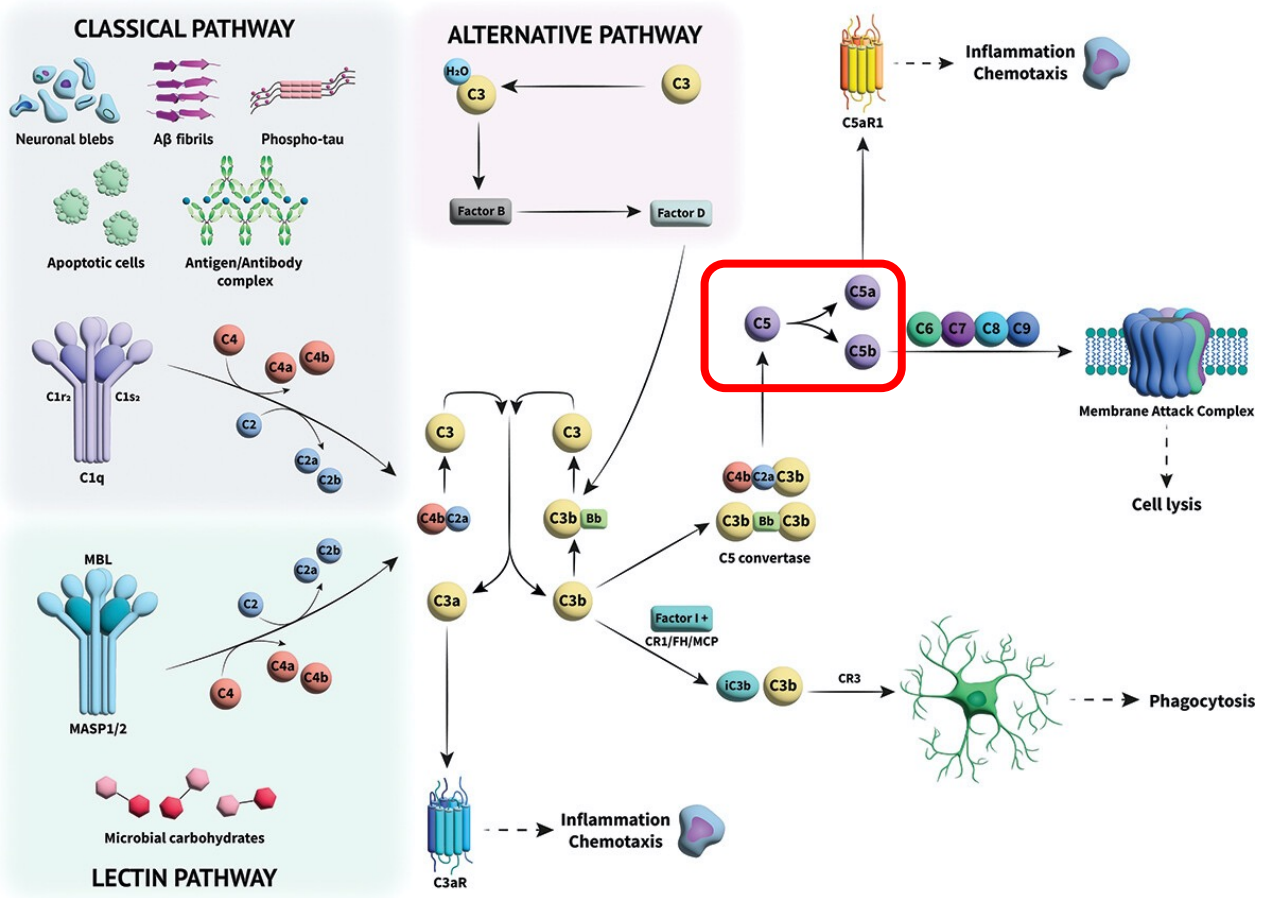


Structure-based
HYDE scoring function

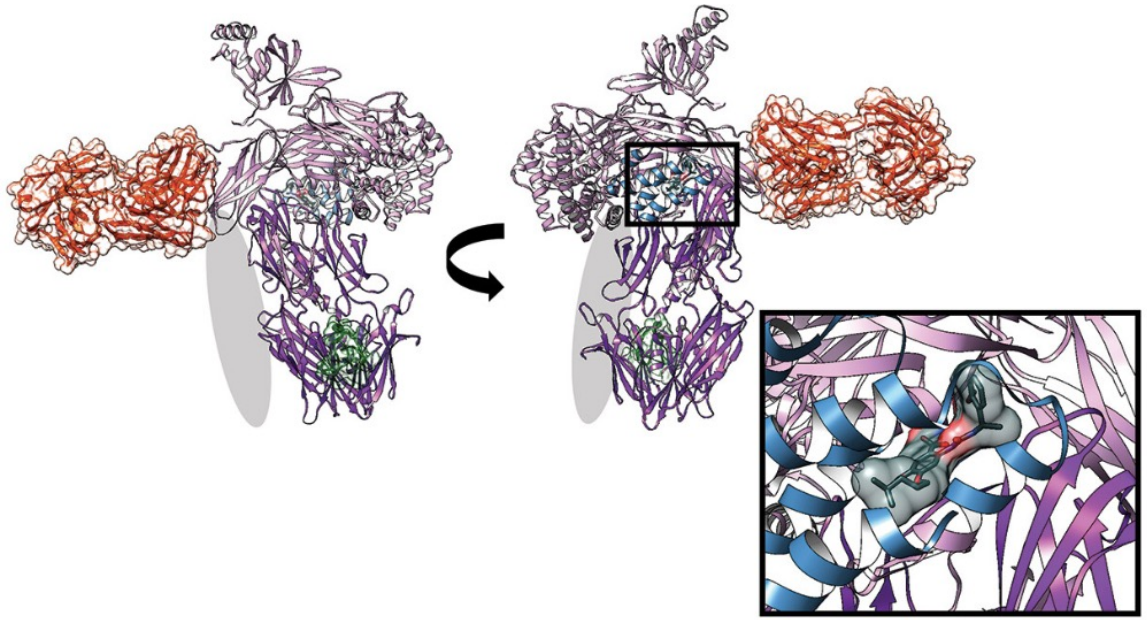
Case Study

COMPLEMENT C5

Complement pathway and C5

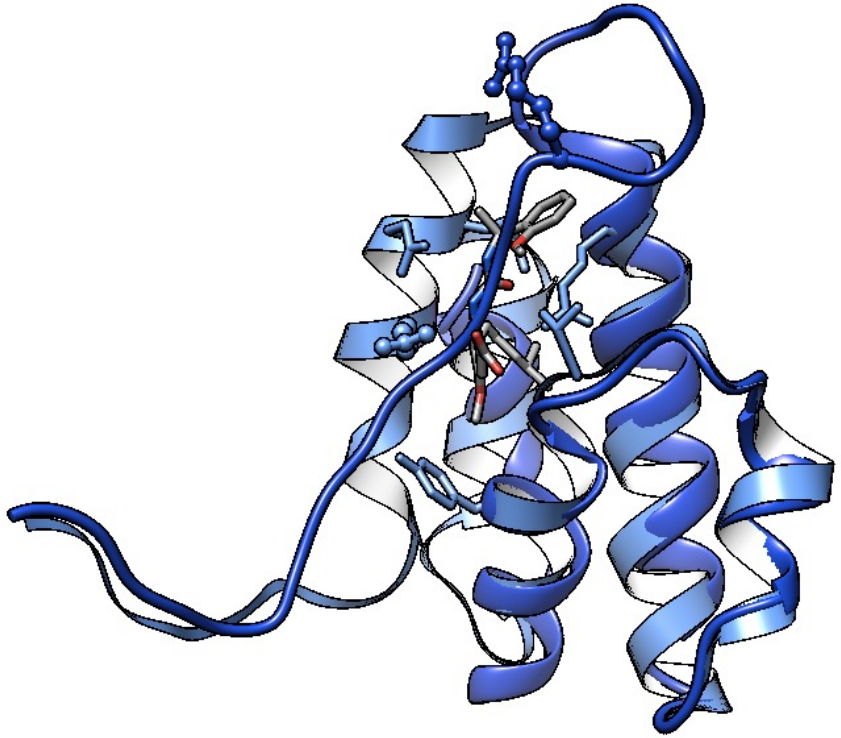


PDB: 8AYH



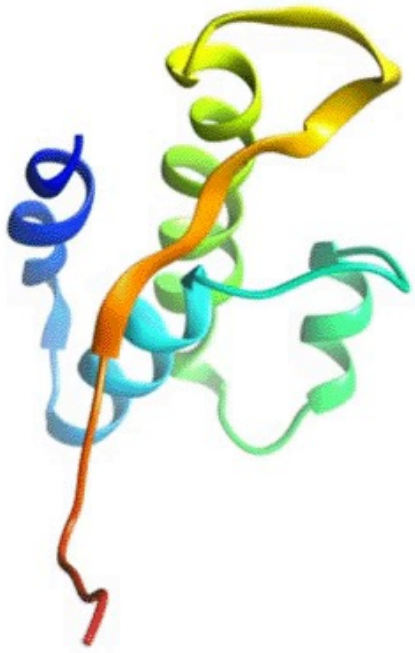
Stennett *et al*., 2023

Complement C5 inhibitors: MoA



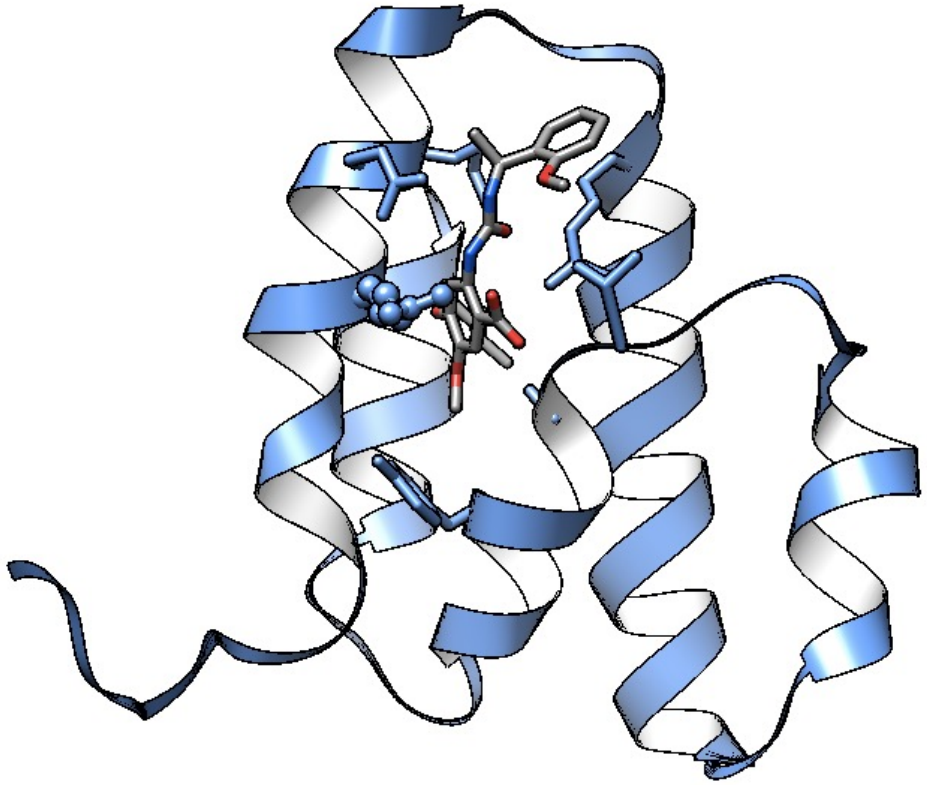
PDB: 3CU7

Ensemble (apo)



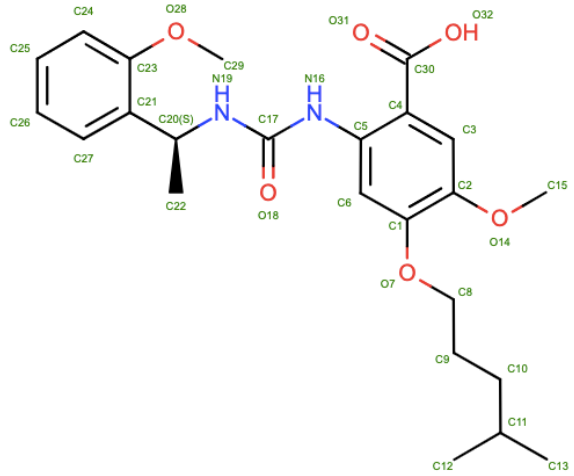
PDB: 8AYH

Complement C5 inhibitors



CryoEM structure (3.35 Å resolution) of human complement C5 in complex with small molecule inhibitor (H1H) and CVF

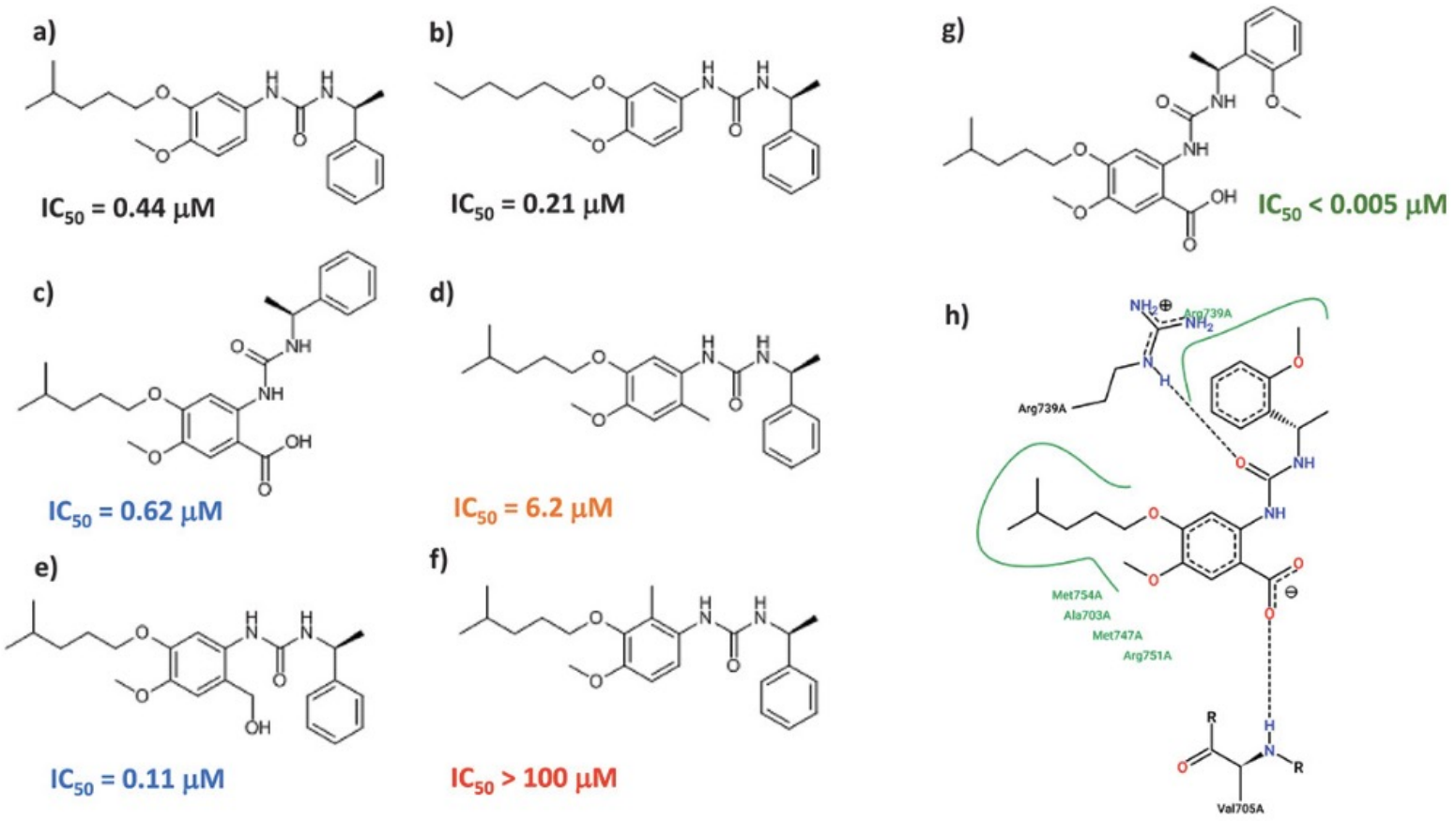
IC_{50}: 0.1 - 5 nM, from 3 assays



H1H

PDB: 8AYH

Complement C5 inhibitors



Stennett *et al*., 2023

Contents of practical sessions

- Introduction to SeeSAR
- Application 1: lead optimisation
- Application 2: core expansion
- Application 3: core replacement/scaffold-hopping
- Application 4: fast generation and evaluation of analogues
- Application 5: Virtual screening of massive virtual libraries with InfiniSee (Scaffold Hopper) and SeeSAR

Credits

Shangze Xu

Amelia Stennett

Junya Zhang

Ayaz Ahmad

Dominic Alderson

Kate Madden

Adam Wollman

Wyatt Yue

Ehmke Pohl (Durham)

João de Souza (RxCelerate)

Jon Heal (RxCelerate)

Christian Lemmen (BiosolveIT)

EPSRC
Engineering and Physical Sciences Research Council

BBSRC
bioscience for the future

RxCelerate

BioSolveIT