CCP5 Summer School

July 2023

Practical sessions: Structure-based and ligand-based design of C5 inhibitors with BioSolveIT suite

DR. AGNIESZKA K. BRONOWSKA

SCHOOL OF NATURAL ENVIRONMENTAL SCIENCES

NEWCASTLE UNIVERSITY

Contents

- Molecular docking and virtual screening
- 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

Molecular docking: two main tasks

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

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

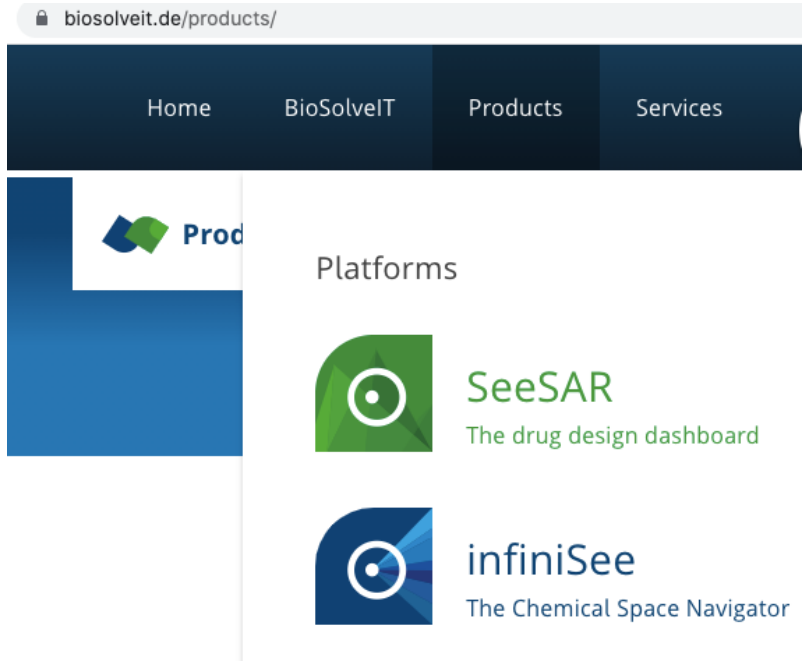
Molecular docking: to-do list

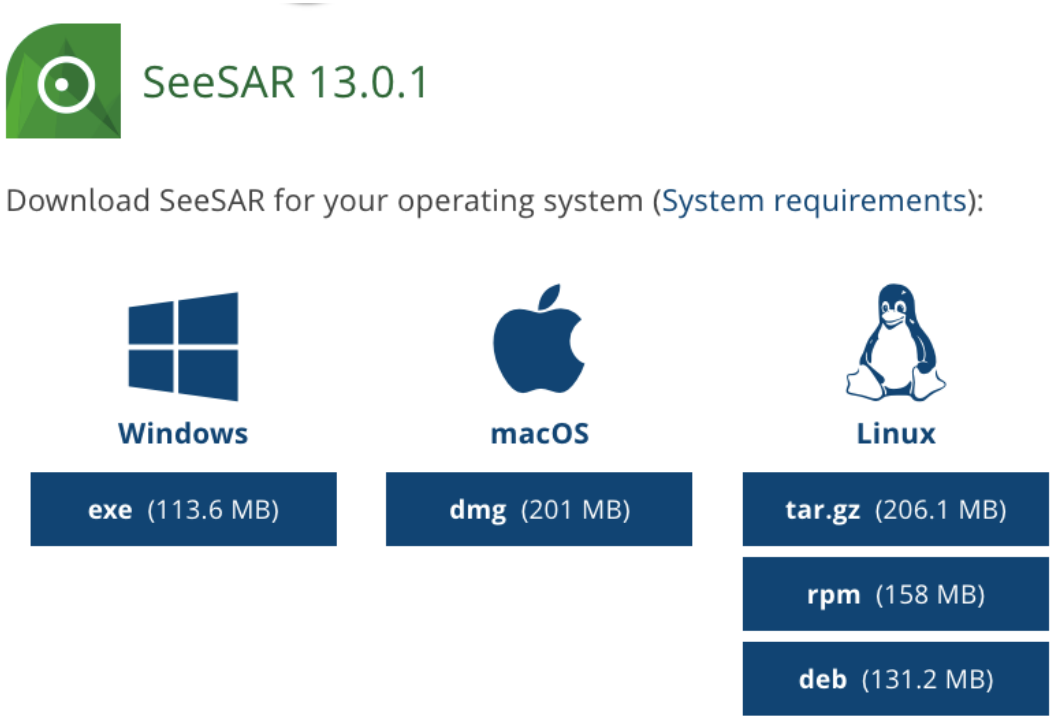
Problem: a pair of molecules represented by their 3D coordinates

- 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)*;

Installing SeeSAR and InfiniSee

- Platform-independent
- Quick to install
- Temporary license key (works for both) provided





Introduction to SeeSAR

Tasks 1 and 2

- Import your target structure to SeeSAR

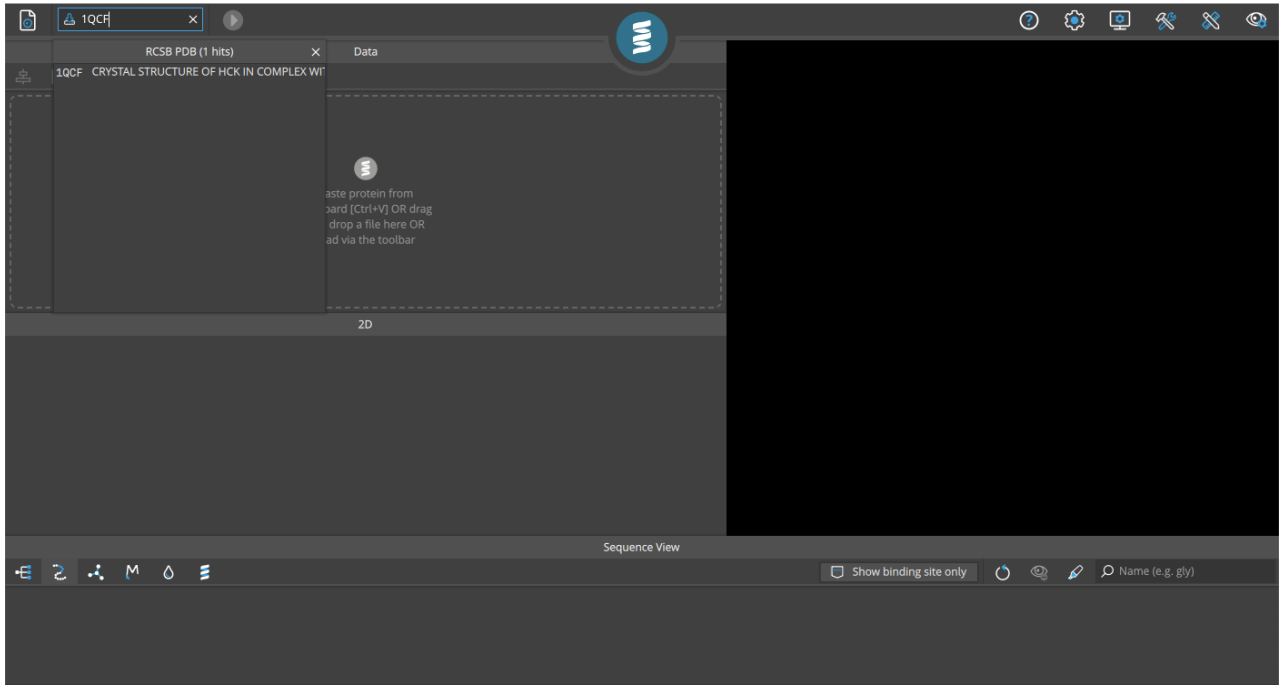
You can either import directly from RCSB PDB Data Bank, or read your local file

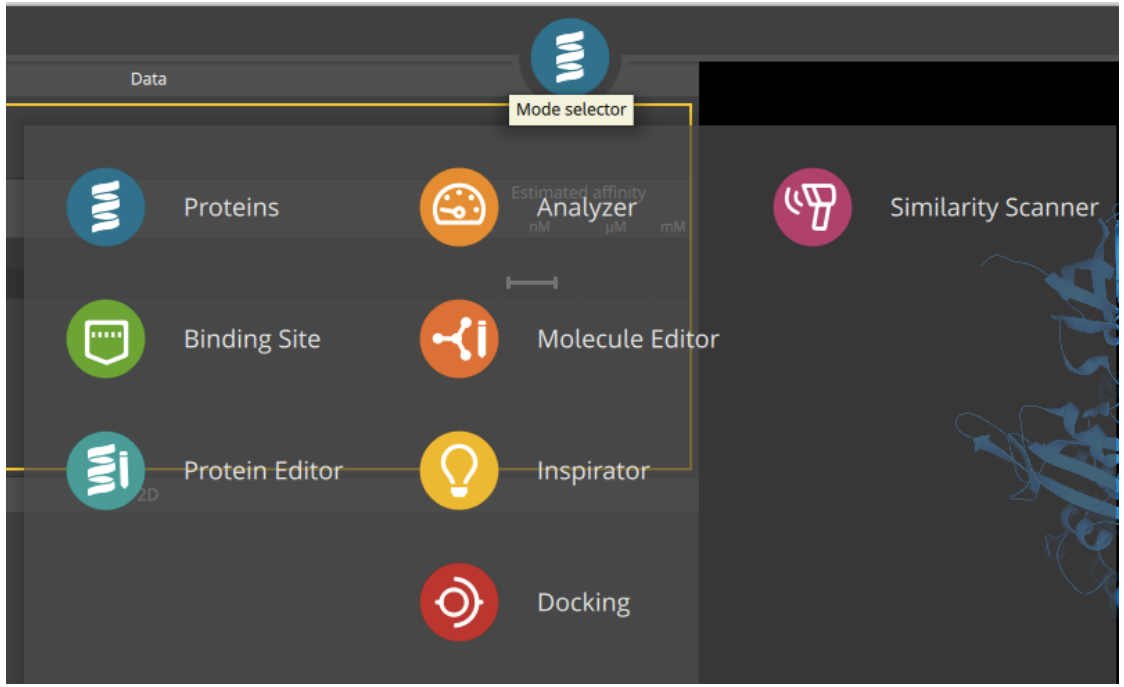
- Assign the binding site

The process is automated: user defines the site either by existing ligand (docked or experimental), or by unoccupied pockets

For the covalent docking, you need to prepare your site (target residue) – or to start with the ligand covalently bound

Getting started



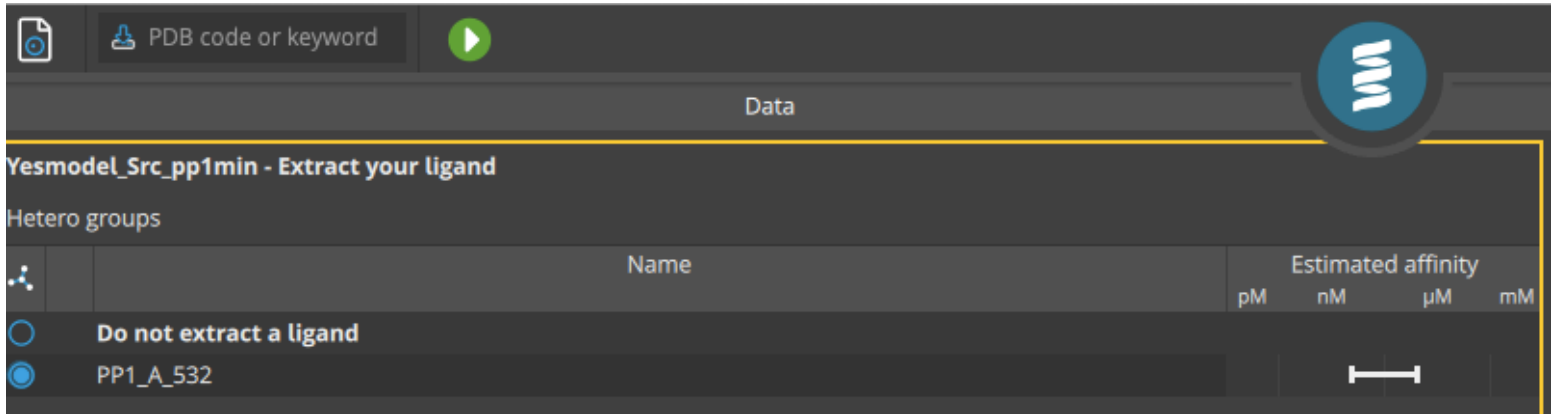


The binding site definition in SeeSAR (1)

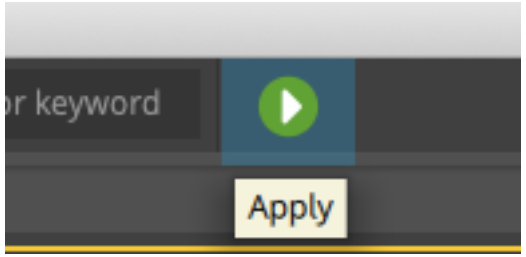
PDB code: 1QCF

1.

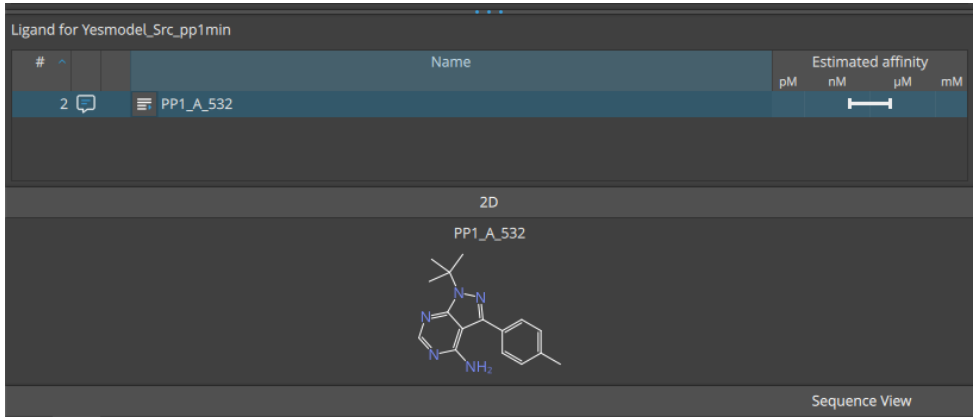
You must tick here



2.

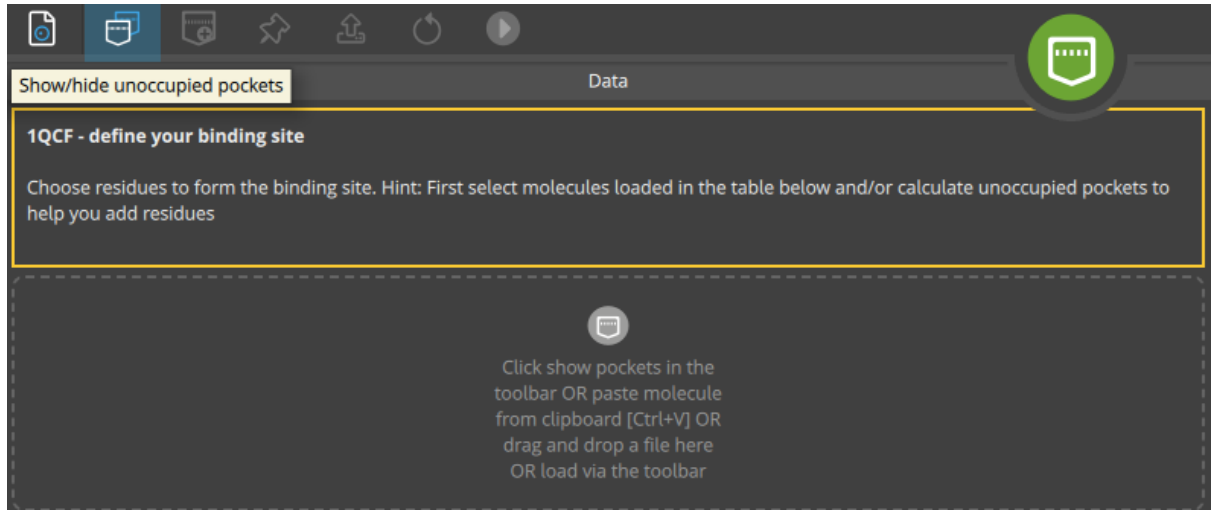


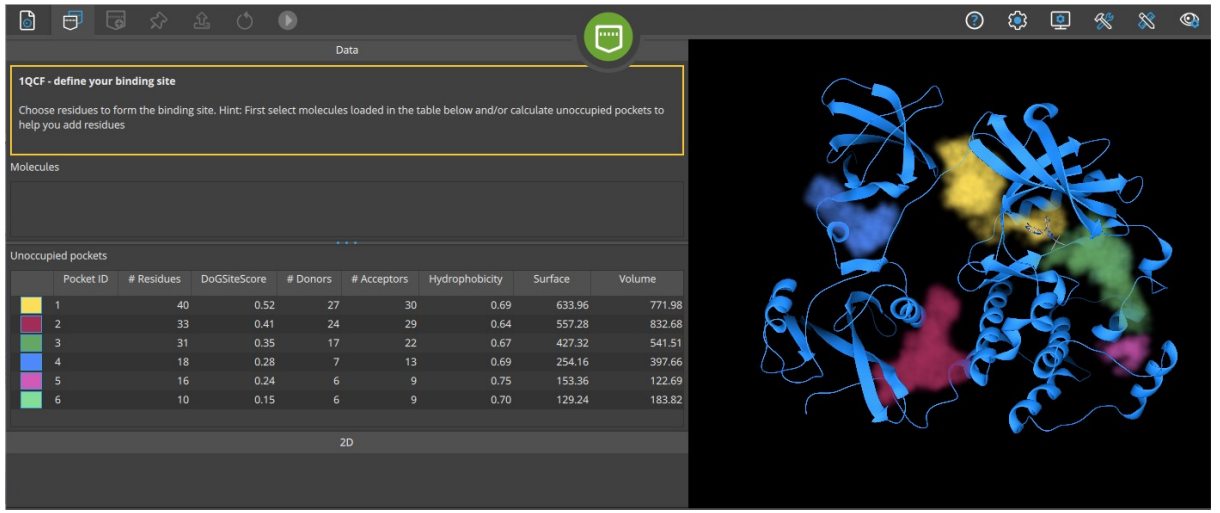
3. The outcome



The binding site definition in SeeSAR (2)

- Using unoccupied pockets





Lead optimisation

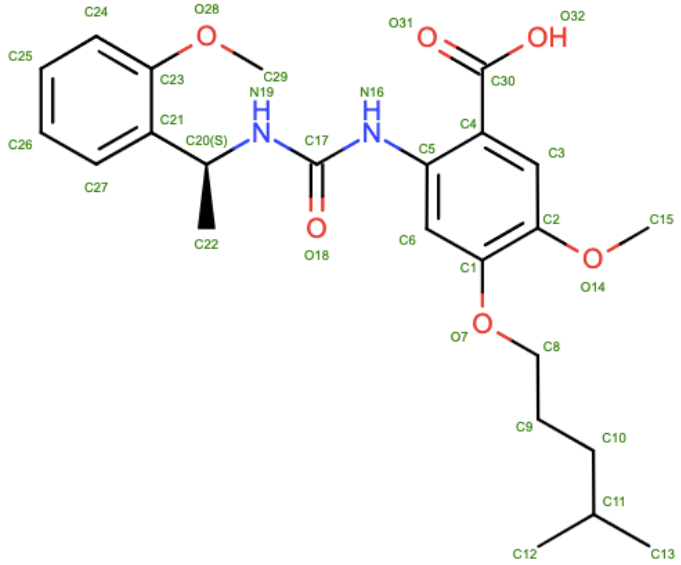
Complement C5 inhibitor

PDB code (source): 8AYH

Use PDB file provided: C5_H1H_CCP5.pdb

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



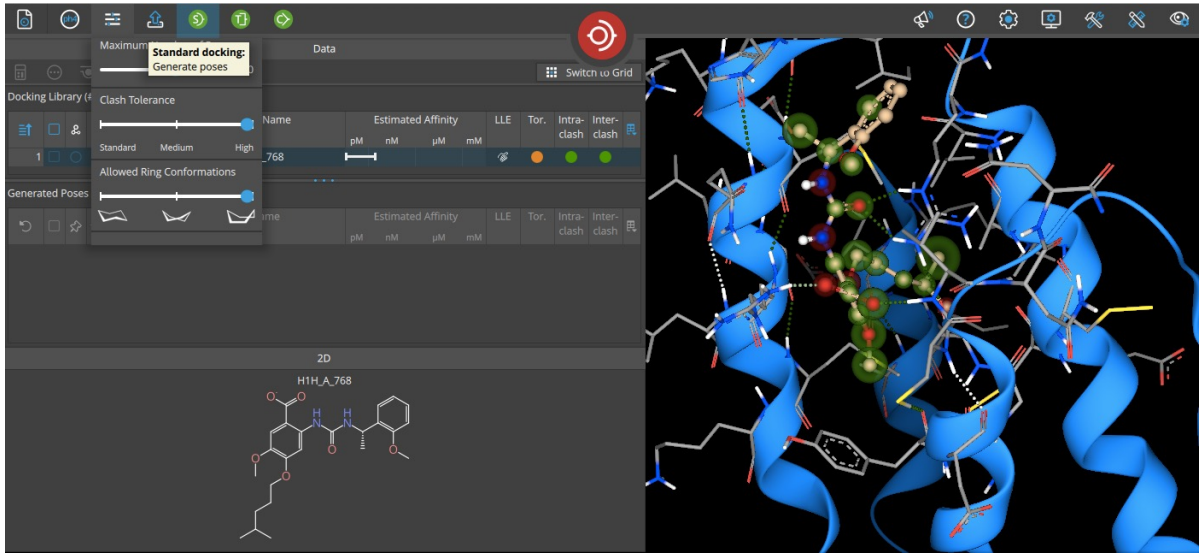
Complement C5 inhibitor

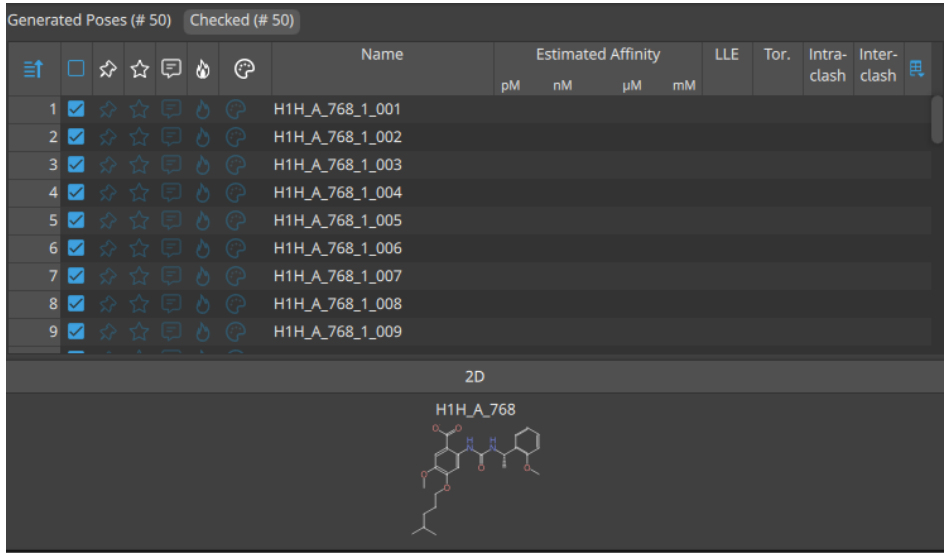
- **Task 3:** open the target-inhibitor complex in SeeSAR, indicate the binding site by the inhibitor (H1H), and bring it to the Molecule Editor mode
- Once you are in the Editor mode, you can manually modify/grow your compound
- **Before you start modifying:** what is the affinity range calculated? How does this compare to the experimental binding affinity?
- What you may try (a good practice): **molecular docking**
- Bring your compound to the **Docking mode**

Complement C5 inhibitor: docking

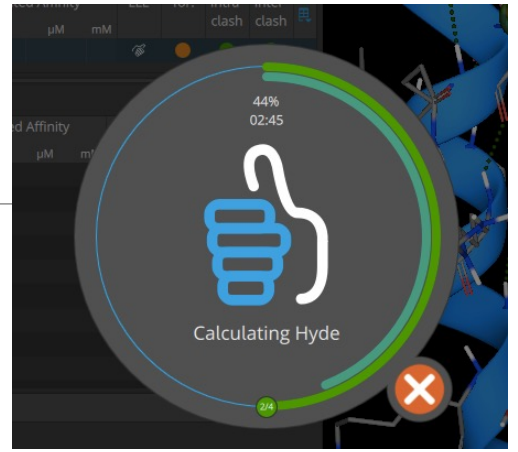
Task 4: Use standard (noncovalent) docking, 50 poses, high clash tolerance and maximum allowed ring conformations

First you generate the poses, then you select them (select all) and rank them by calculating their binding affinities





Complement C5 inhibitor



Task 5: Sort the poses from the best to the worst and bring the best one back to the Editor. Compare it with the starting structure.
What do you see?

Complement C5 inhibitor

Task 6: Try 5-10 modifications to improve the binding affinity (5-10 minutes).

Do not forget to save the project file every now and then.

Complement C5 inhibitor

Task 7: Bring an “external” molecule to SeeSAR (Editor mode) and dock it to C5.

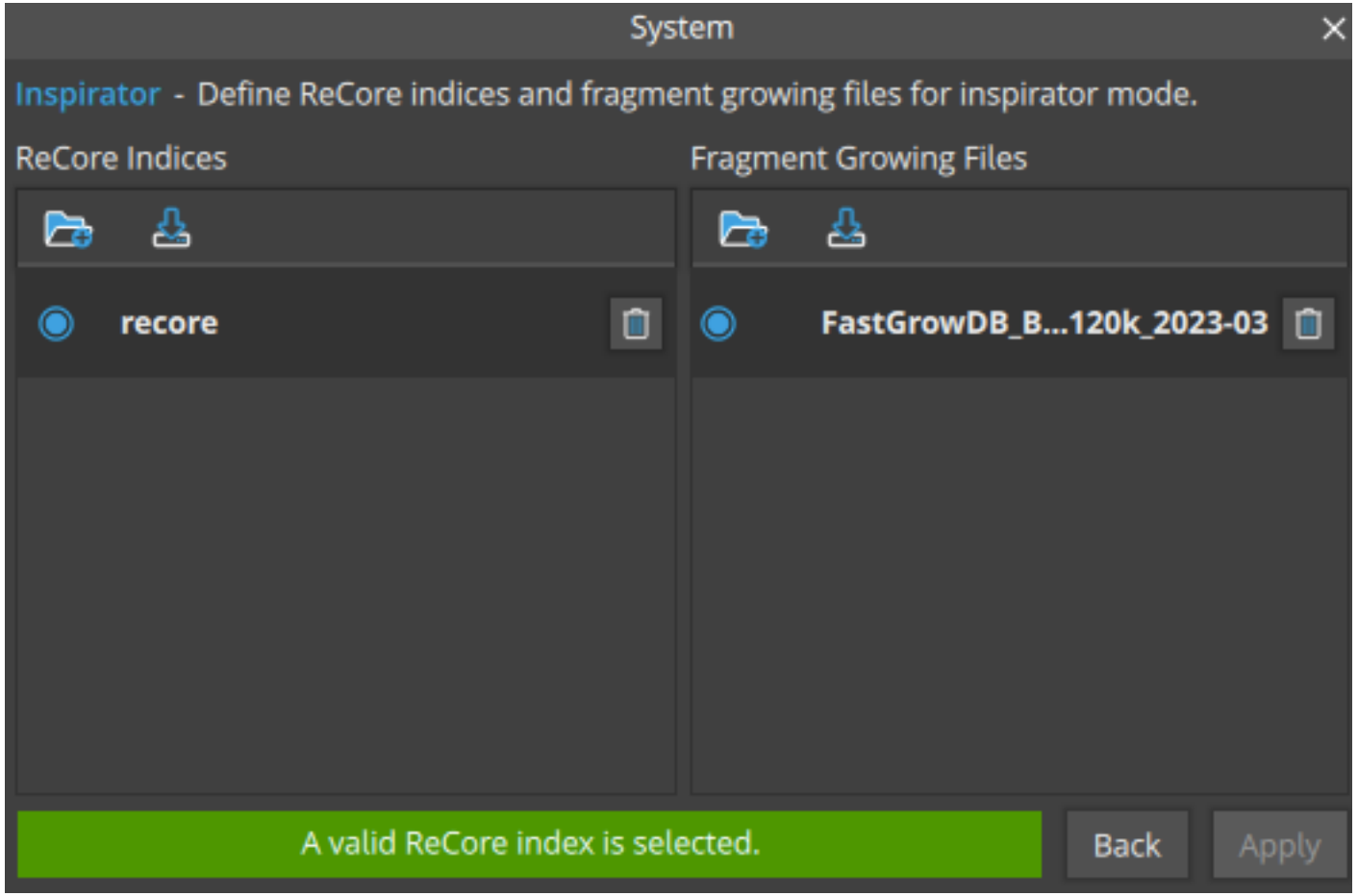
(I will let you to figure out how)

Remember: SeeSAR can deal with a lot of file formats – including SMILES strings.

Good practice: after scoring and ranking, delete the worst/clashed poses to avoid massive files

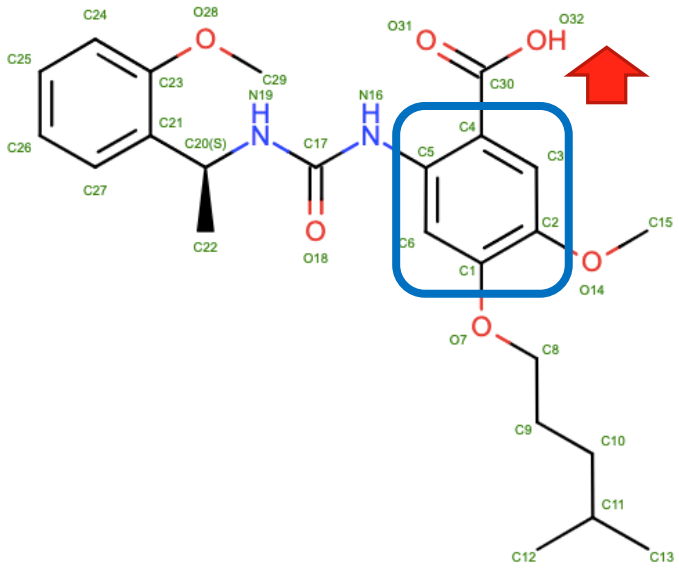
Core expansion in SeeSAR

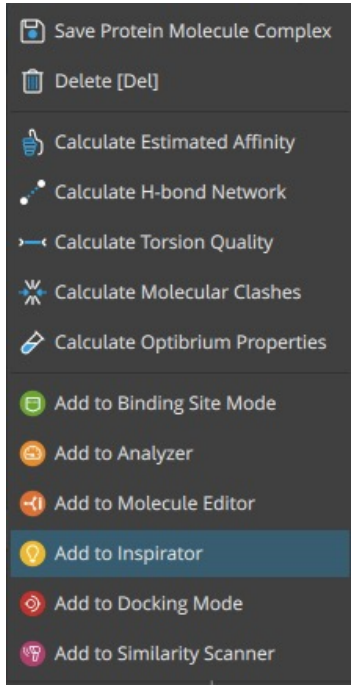
Configuring Inspirator mode



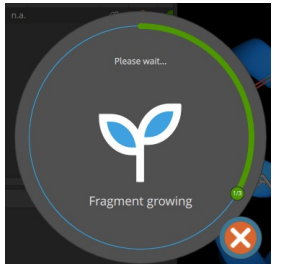
Task 8: Core expansion in SeeSAR

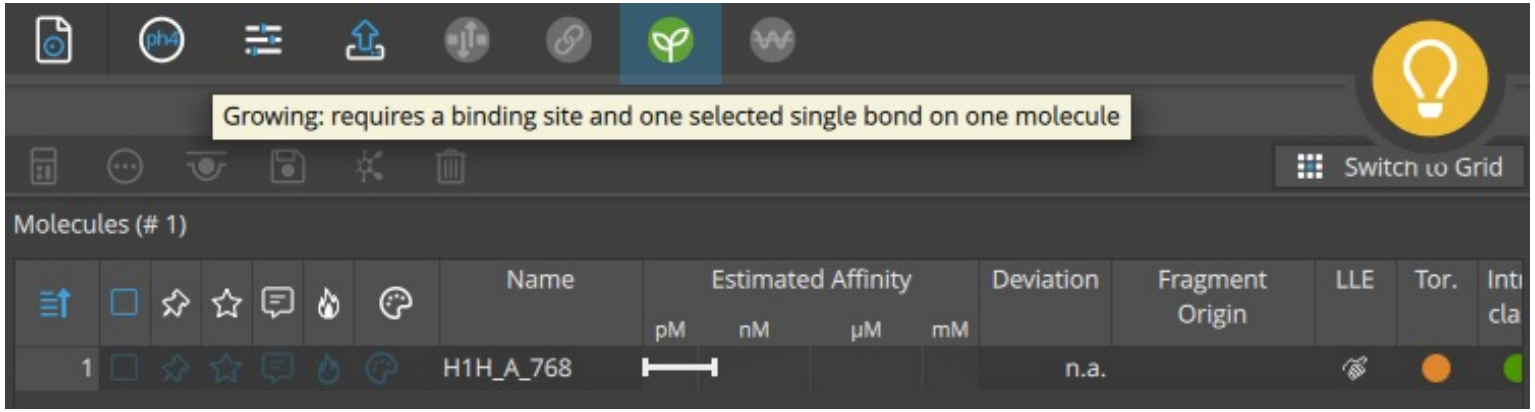
- We will select a central core of H1H for the expansion
- Bring H1H compound from the Editor to the Inspirator mode

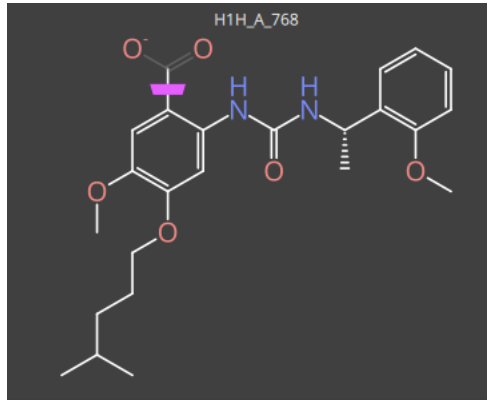


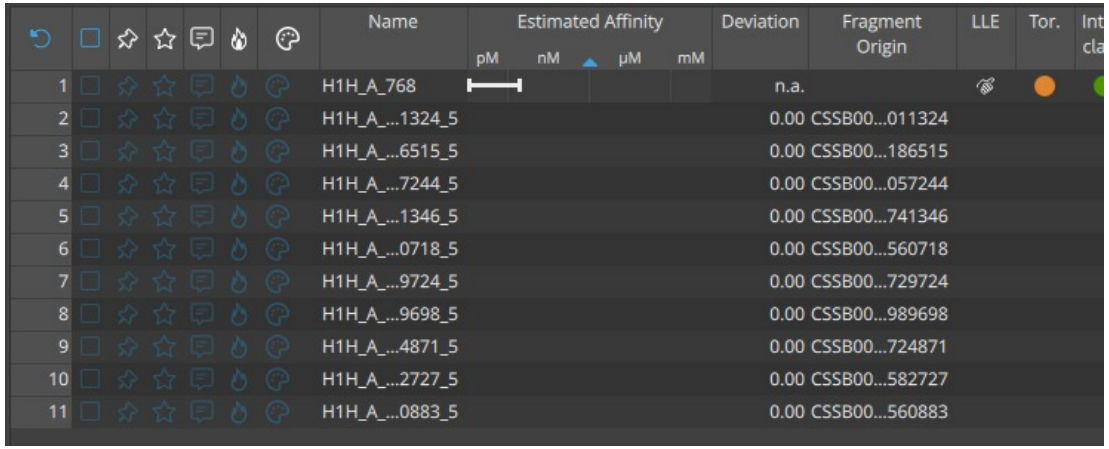


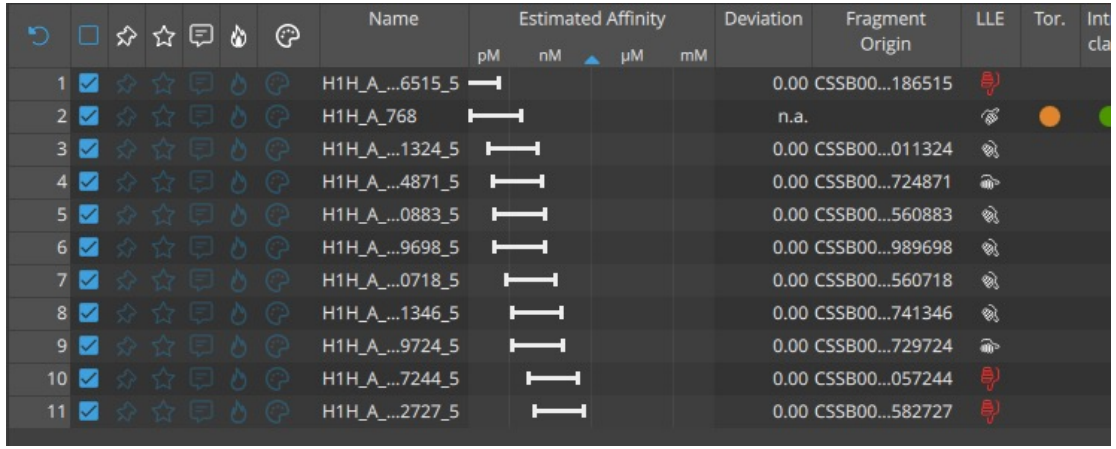
Inspirator


Please wait...
Fragment growing

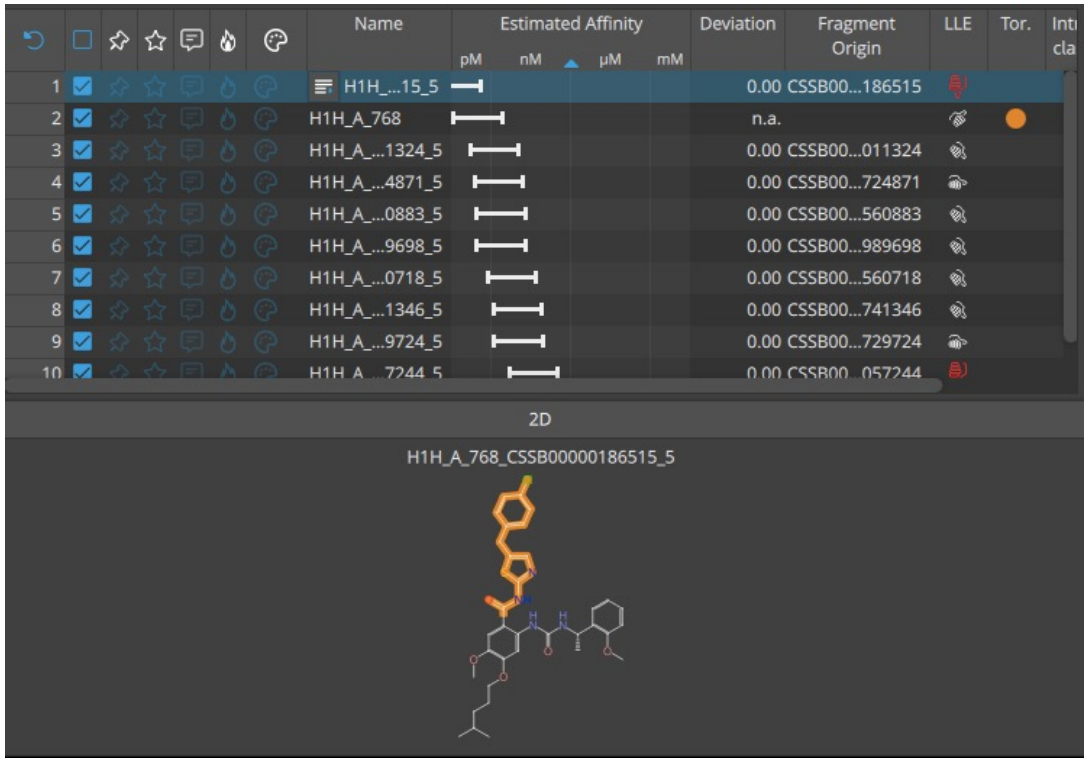

Growing: requires a binding site and one selected single bond on one molecule
Switch to Grid
Molecules (# 1)
Name
Estimated Affinity
pM
nM
μM
mM
Deviation
Fragment Origin
LLE
Tor.
1
H1H_A_768
n.a.

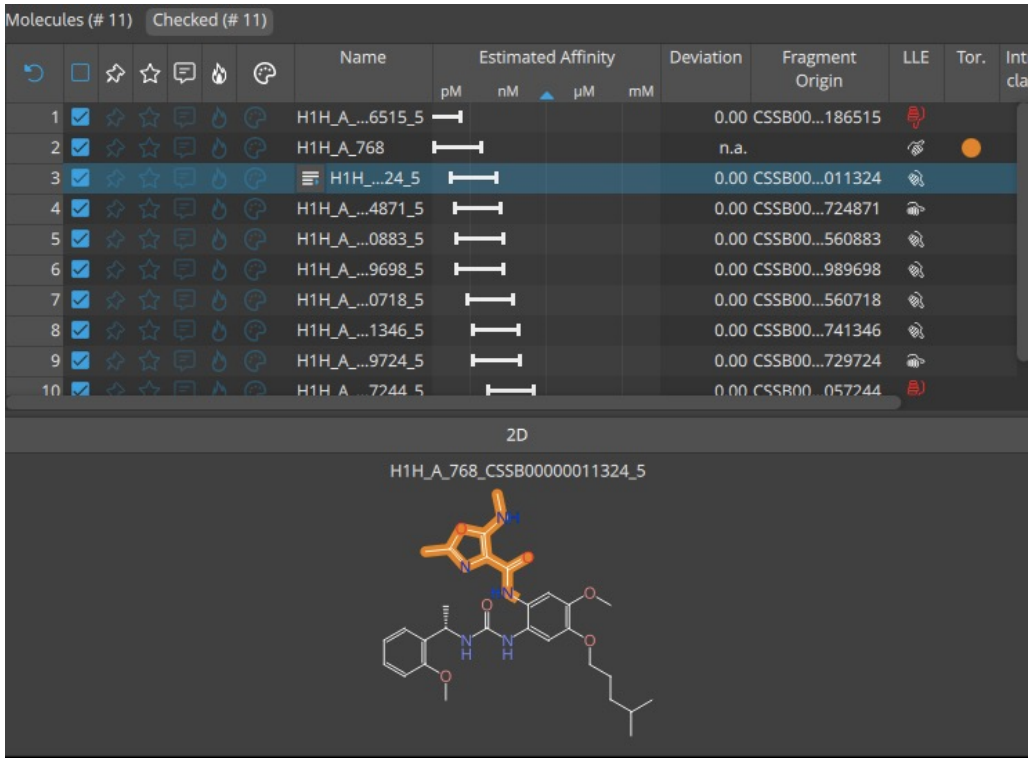

H1H_A_768


Name
Estimated Affinity
pM
nM
μM
mM
Deviation
Fragment Origin
LLE
Tor.
1 H1H_A_768 n.a.
2 H1H_A_...1324_5 0.00 CSSB00...011324
3 H1H_A_...6515_5 0.00 CSSB00...186515
4 H1H_A_...7244_5 0.00 CSSB00...057244
5 H1H_A_...1346_5 0.00 CSSB00...741346
6 H1H_A_...0718_5 0.00 CSSB00...560718
7 H1H_A_...9724_5 0.00 CSSB00...729724
8 H1H_A_...9698_5 0.00 CSSB00...989698
9 H1H_A_...4871_5 0.00 CSSB00...724871
10 H1H_A_...2727_5 0.00 CSSB00...582727
11 H1H_A_...0883_5 0.00 CSSB00...560883


Name
Estimated Affinity
pM
nM
μM
mM
Deviation
Fragment Origin
LLE
Tor.
1 H1H_A_...6515_5 0.00 CSSB00...186515
2 H1H_A_768 n.a.
3 H1H_A_...1324_5 0.00 CSSB00...011324
4 H1H_A_...4871_5 0.00 CSSB00...724871
5 H1H_A_...0883_5 0.00 CSSB00...560883
6 H1H_A_...9698_5 0.00 CSSB00...989698
7 H1H_A_...0718_5 0.00 CSSB00...560718
8 H1H_A_...1346_5 0.00 CSSB00...741346
9 H1H_A_...9724_5 0.00 CSSB00...729724
10 H1H_A_...7244_5 0.00 CSSB00...057244
11 H1H_A_...2727_5 0.00 CSSB00...582727

Inspirator: example results

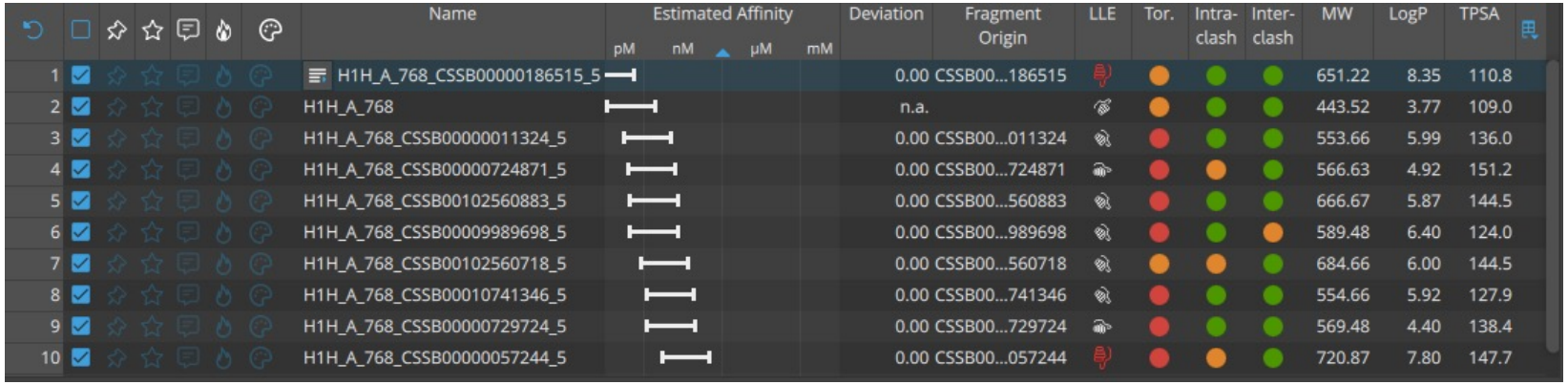




Inspirator

It may be useful to output the following:

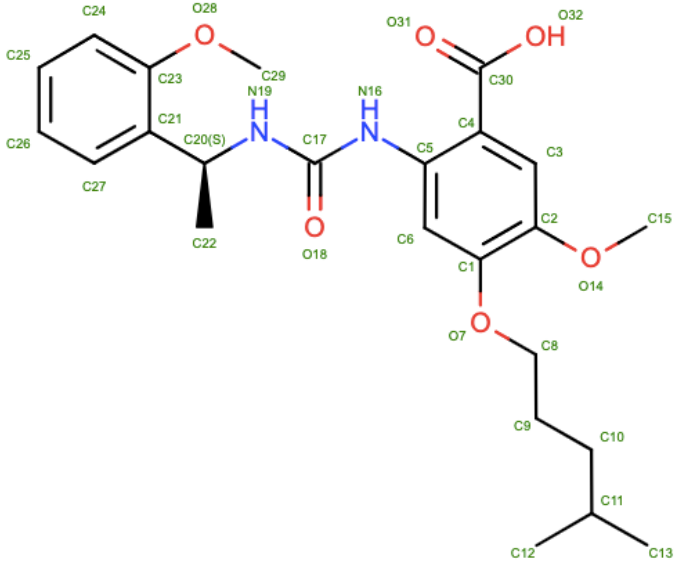
- MW
- logP
- TPSA
- Torsional quality estimates
- Clashes



	Name	Estimated Affinity (pM nM µM mM)	Deviation	Fragment Origin	LLE	Tor.	Intra-clash	Inter-clash	MW	LogP	TPSA
1	H1H_A_768_CSSB00000186515_5		0.00	CSSB00...186515					651.22	8.35	110.8
2	H1H_A_768		n.a.						443.52	3.77	109.0
3	H1H_A_768_CSSB00000011324_5		0.00	CSSB00...011324					553.66	5.99	136.0
4	H1H_A_768_CSSB00000724871_5		0.00	CSSB00...724871					566.63	4.92	151.2
5	H1H_A_768_CSSB00102560883_5		0.00	CSSB00...560883					666.67	5.87	144.5
6	H1H_A_768_CSSB00009989698_5		0.00	CSSB00...989698					589.48	6.40	124.0
7	H1H_A_768_CSSB00102560718_5		0.00	CSSB00...560718					684.66	6.00	144.5
8	H1H_A_768_CSSB00010741346_5		0.00	CSSB00...741346					554.66	5.92	127.9
9	H1H_A_768_CSSB00000729724_5		0.00	CSSB00...729724					569.48	4.40	138.4
10	H1H_A_768_CSSB00000057244_5		0.00	CSSB00...057244					720.87	7.80	147.7

Task 9: Core expansion in SeeSAR

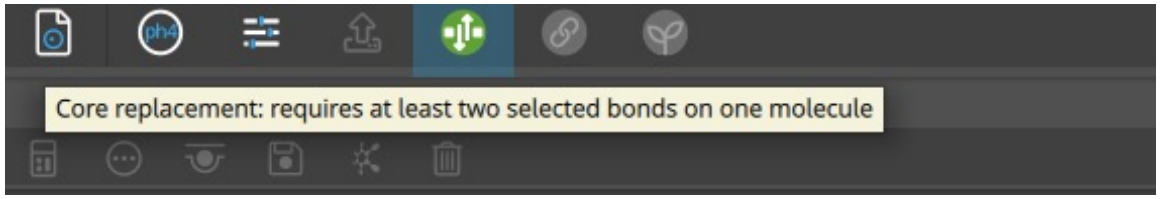
- Repeat the expansion using different direction of the growth
- You may start from H1H, of your “best” analogue from the previous growth iteration

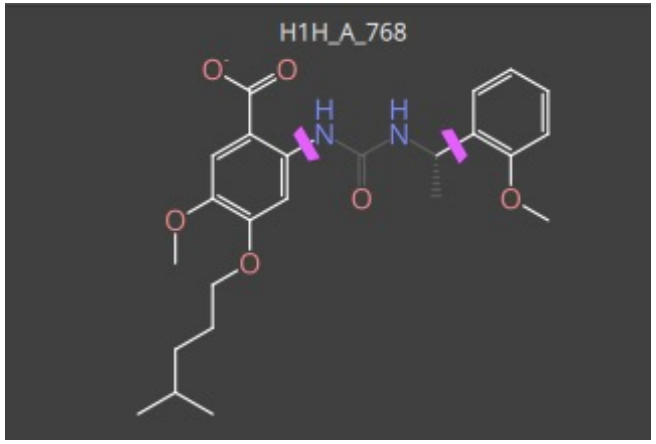


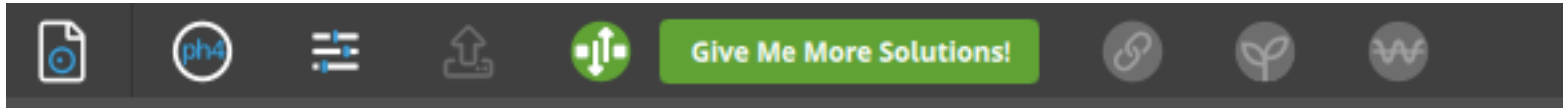
Core replacement

Task 10: Core replacement in Inspirator

- Requires 2 bonds selected
- Does not cut through the ring
- May be done iteratively with core expansion/growth





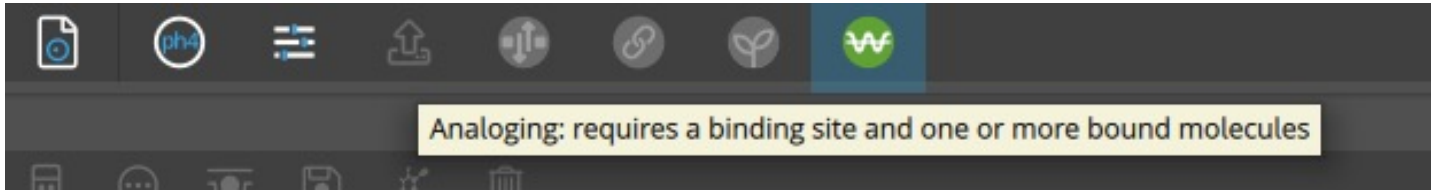


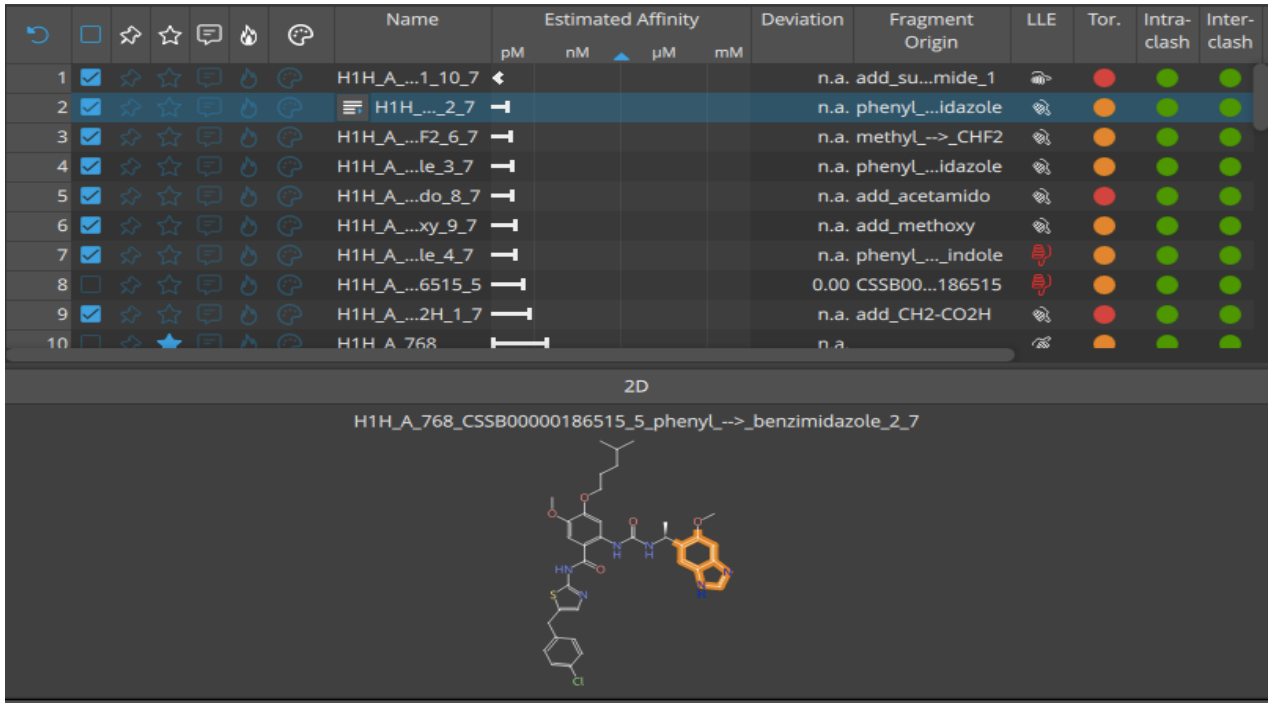
Generation of analogues

SEESAR VERSION 13+ ONLY

Task 11: Analogues in Inspirator

- Requires a binding site and at least one bound molecule





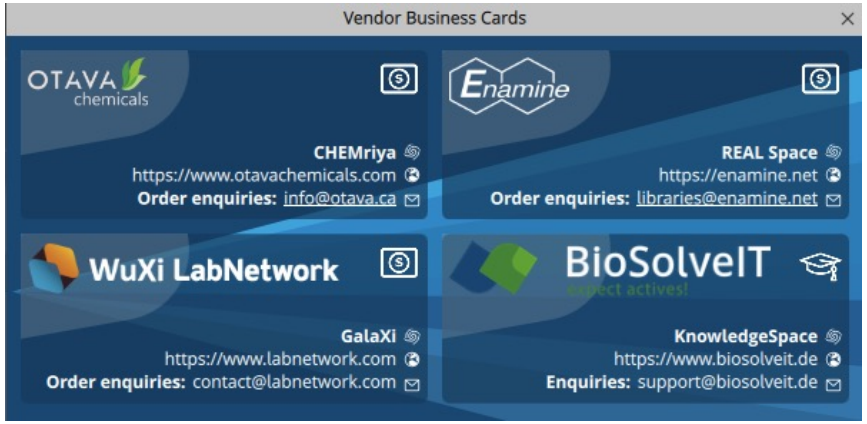
Virtual Screening

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: diversity sets, target-focused, custom

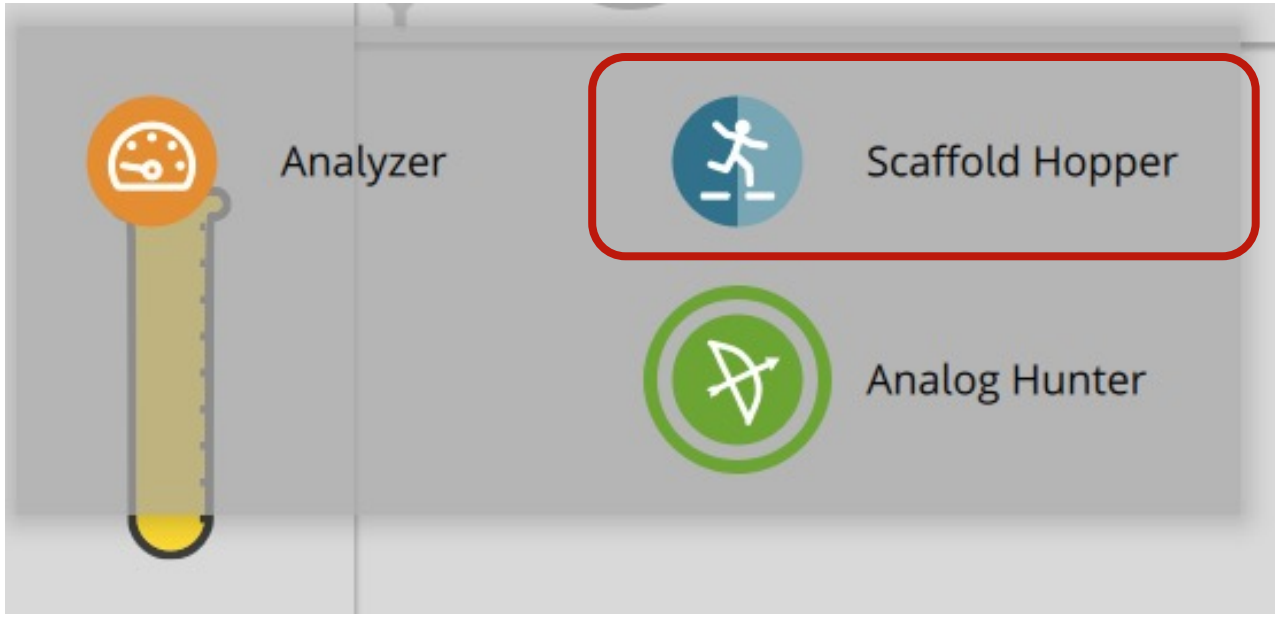
Building your own C5 focused library in InfiniSee

- You may use H1H as a query (“bait”)
- Load it into InfiniSee (Scaffold Hopper mode)
- Choose the chemical space

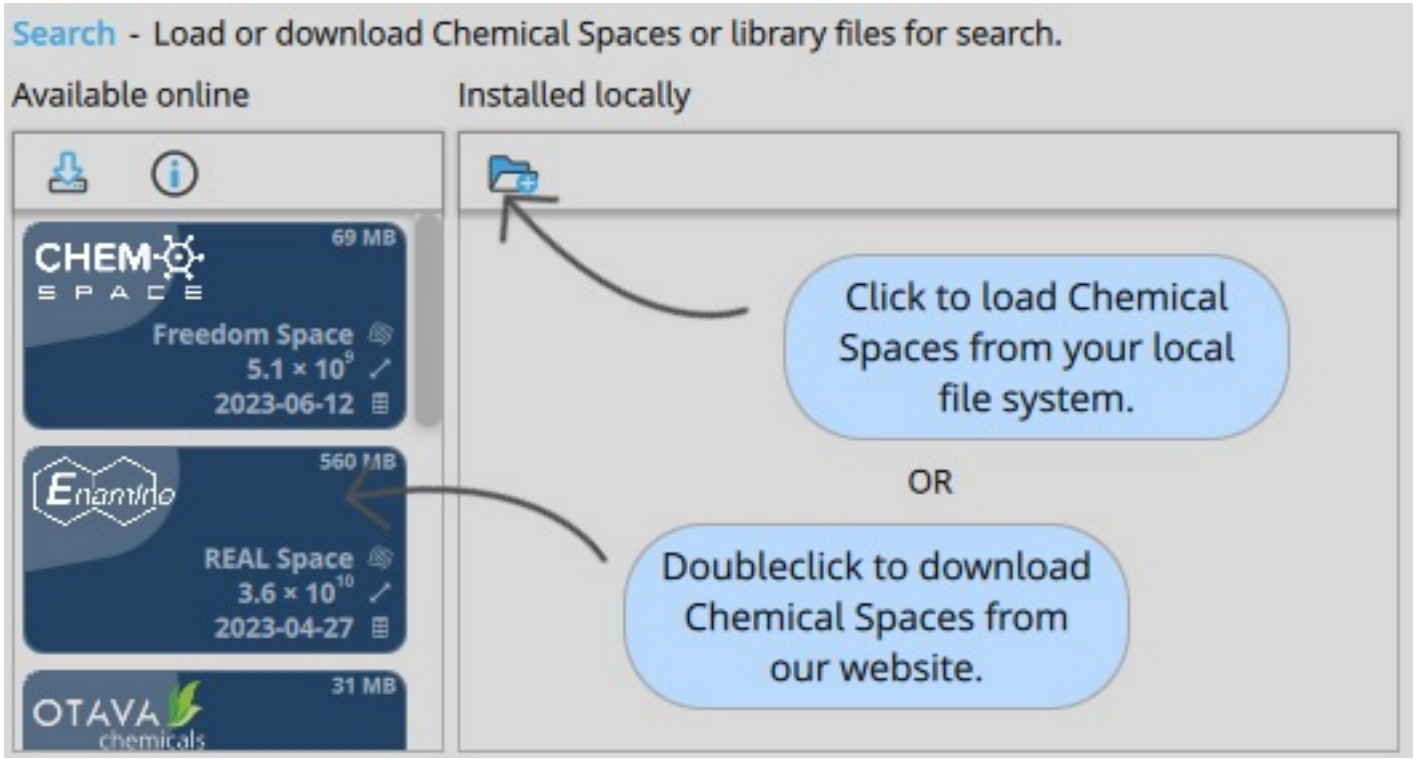


- You can pull thousands of synthetically-feasible analogues in minutes to create your own bespoke focused libraries (SDF/SMILES formats)
- Next, you can virtually screen these analogues

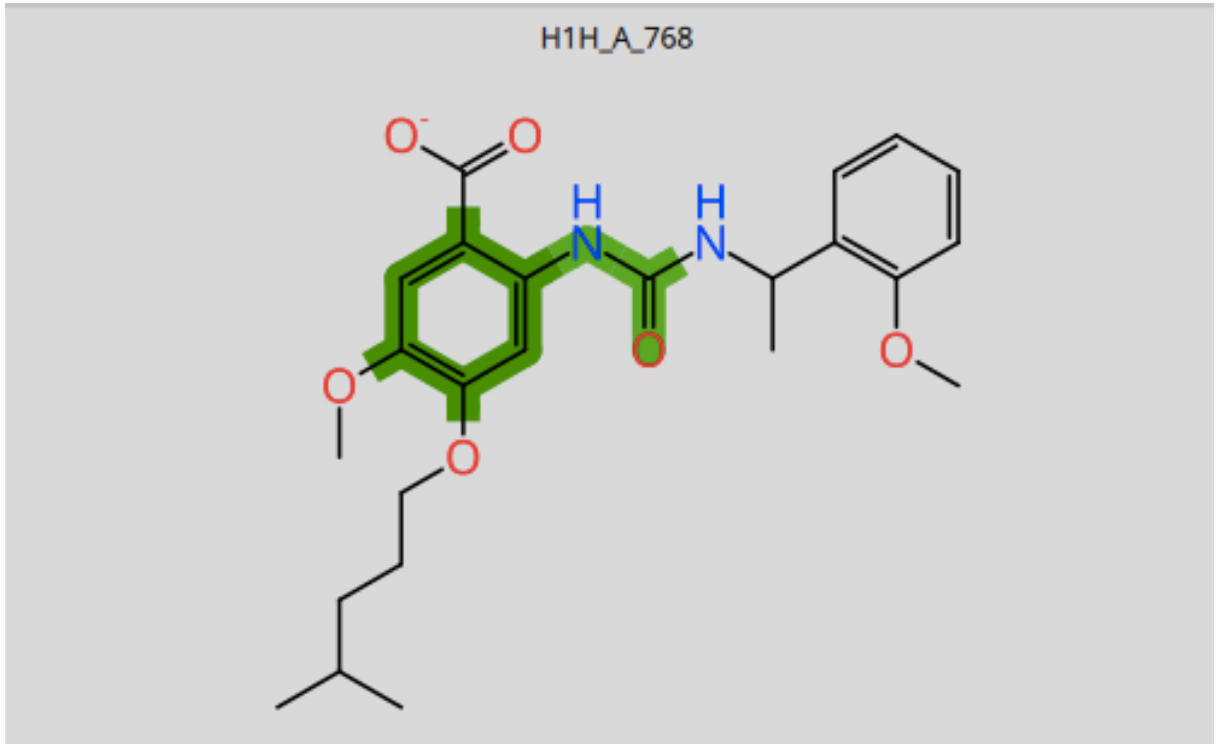
Scaffold Hopper in InfiniSee

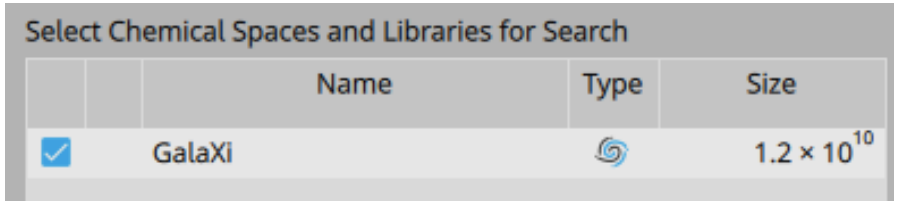


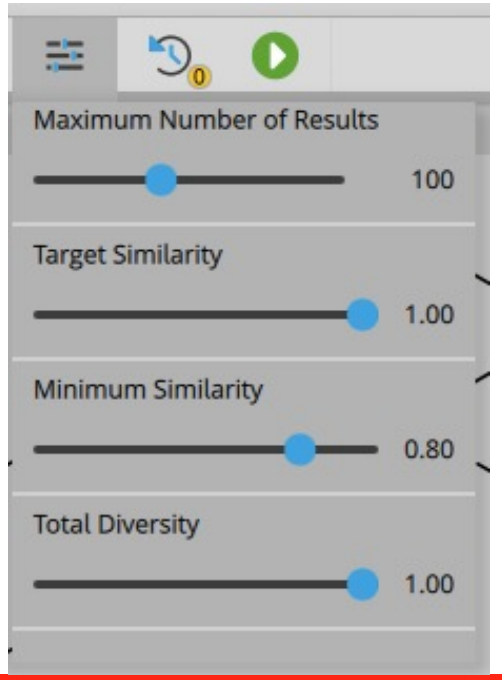
Configuring InfiniSee



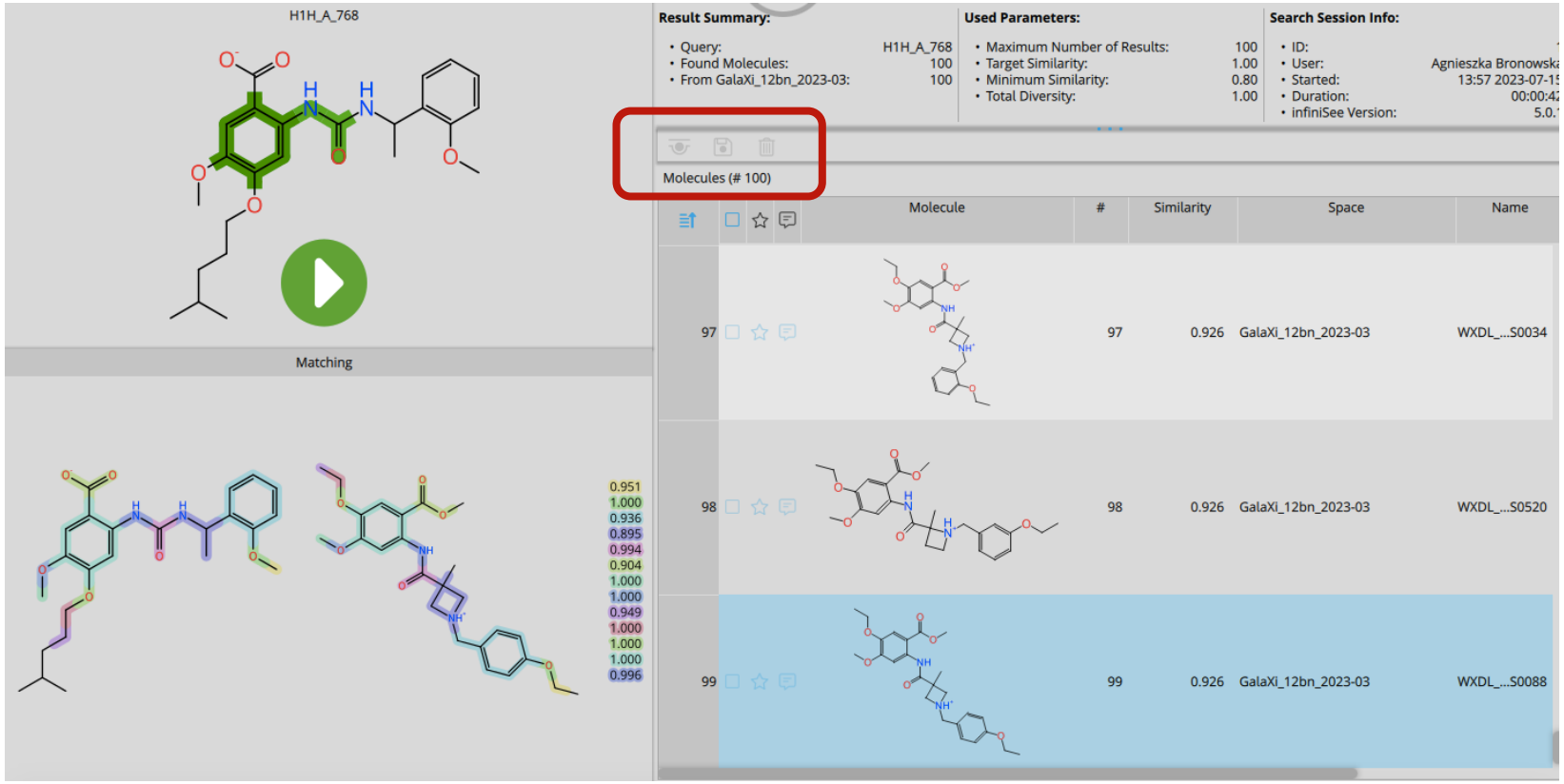
Customise similarity features and results

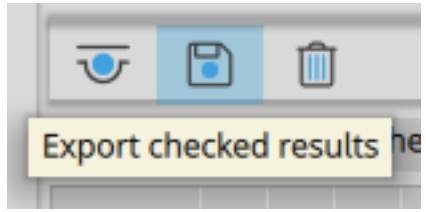






InfiniSee/Scaffold Hopper: results





Chemical space search and virtual screening

Task 12: Using you “best” C5 inhibitor as a “bait” (query), find 100 compounds in GalaXi space.

Out of these 100, pick 10, save them as either SDF or SMILES, upload to SeeSAR and virtually screen them against C5.

Of all compounds virtually tested for C5 binding: select your favourite, export the complex as the PDB file, and you may follow it up using all-atom MD simulations.

Recap: limitations and 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)