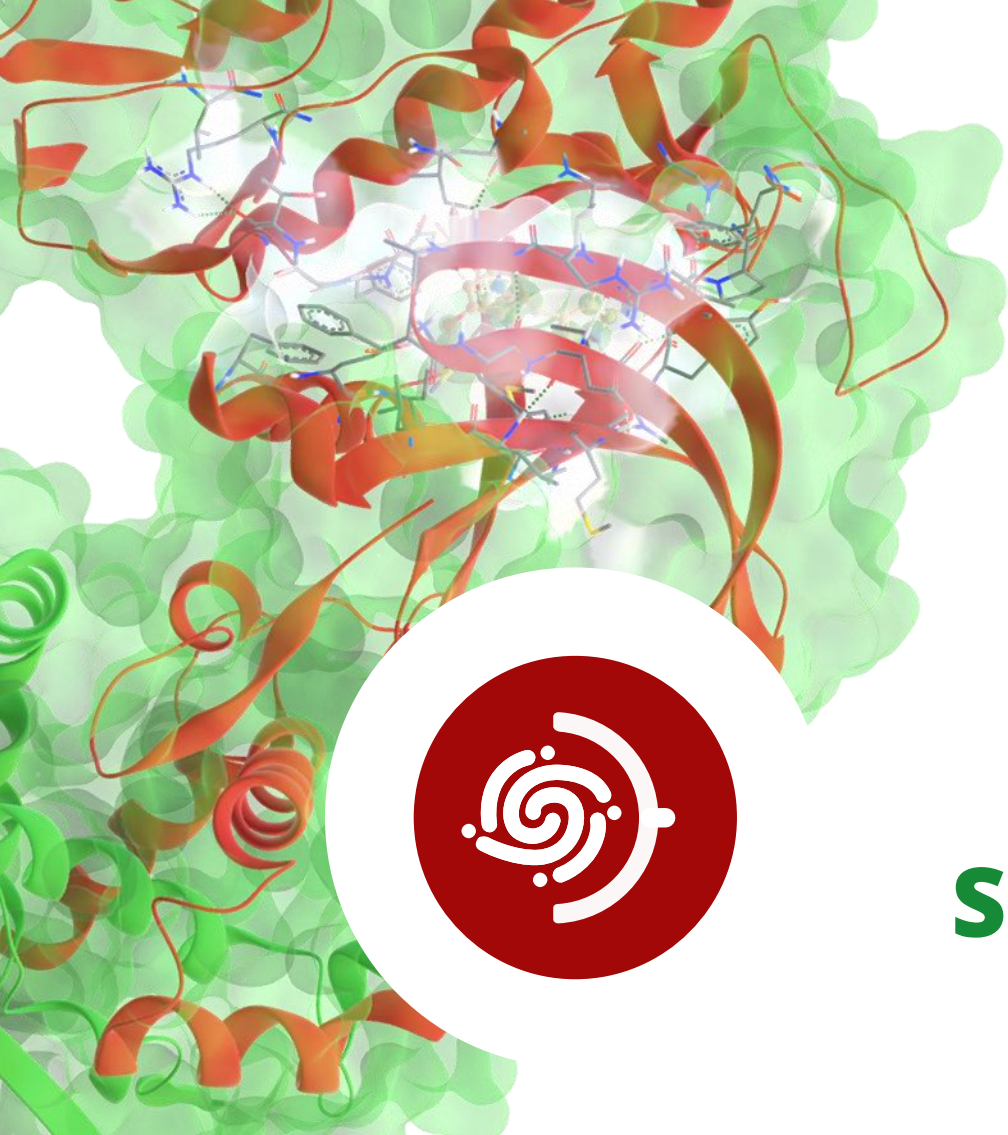


BioSolveIT
expect actives!

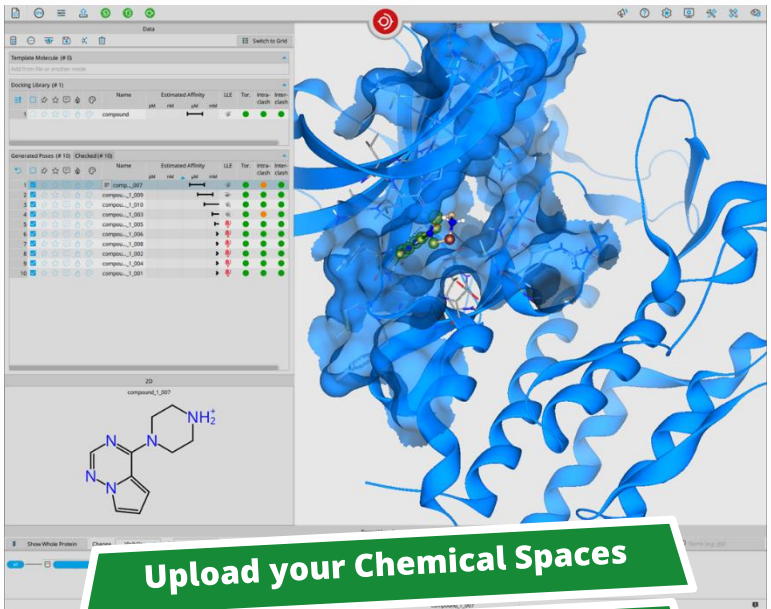
Good Practice for Chemical Space Docking™

Drug discovery, even with the most sophisticated tools, remains a complex endeavor. Therefore, we have compiled some Good Practice recommendations to help users of SeeSAR and Chemical Space Docking™ (C-S-D) to achieve the best and most satisfying results.

Binding Site Definition

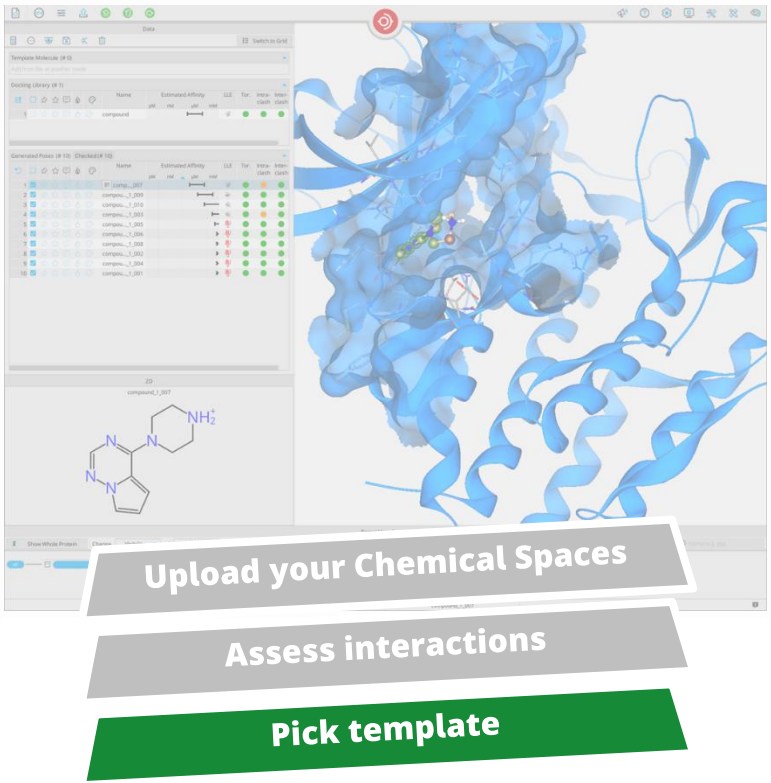
- If you are working with a PDB that contains a ligand, you can use the ligand to define the binding site. All components (residues, water molecule, metals, other co-crystallized molecules) around 6.5 Å of the ligand will become part of the binding site.
- Especially for smaller molecules (e.g., fragments) the binding site definition may end up too small to cover enough space for larger compounds that are generated during the C-S-D workflow. Therefore, it is recommended to extend the binding site in the **Binding Site Mode** to provide enough space for the extension of the anchors. It is also possible to manually add individual residues to the ones proposed by the 'Show/hide unoccupied pockets' feature.
- You can define water molecules in the **Binding Site Mode** to be conserved during the docking. Pose generation will take them into account for interactions and availability of space.

Before Starting the C-S-D Workflow



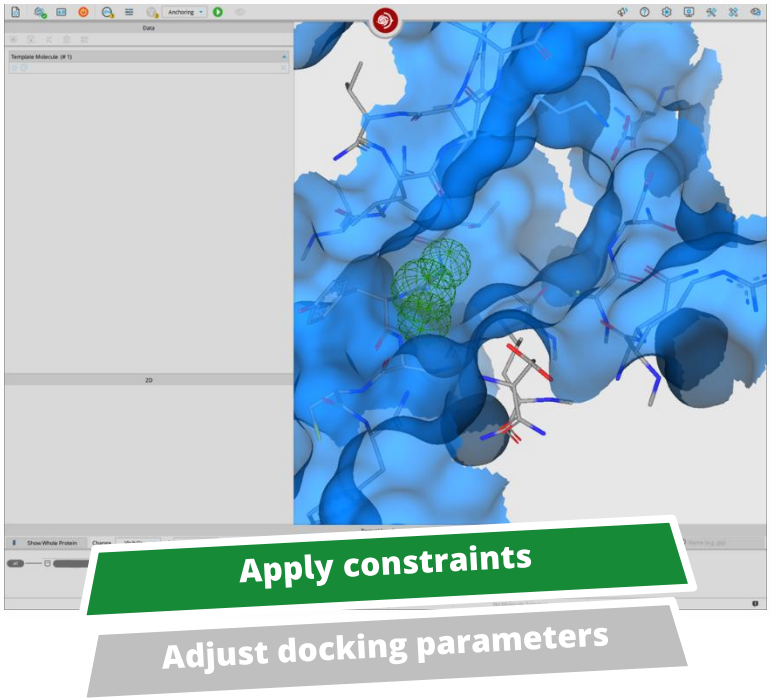
- Upload your Chemical Space. To start a C-S-D worklow you need to upload the Chemical Space to HPSee first. Chemical Spaces can be downloaded here. The upload may take some time to be available in SeeSAR. This is due to the fact that the 3D coordinates of the synthons need to be generated first.
- Validate your docking setup by re-docking your co-crystallized compound in the normal Docking Mode.
- Adjust the clash tolerance in the docking settings if you cannot reproduce the native bound state due to clashes in the tight binding site.
- You can also dock some other known binders to get a feeling for potential interaction hot spots that you can later utilize for pharmacophore constraints.

Before Starting the C-S-D Workflow (2)



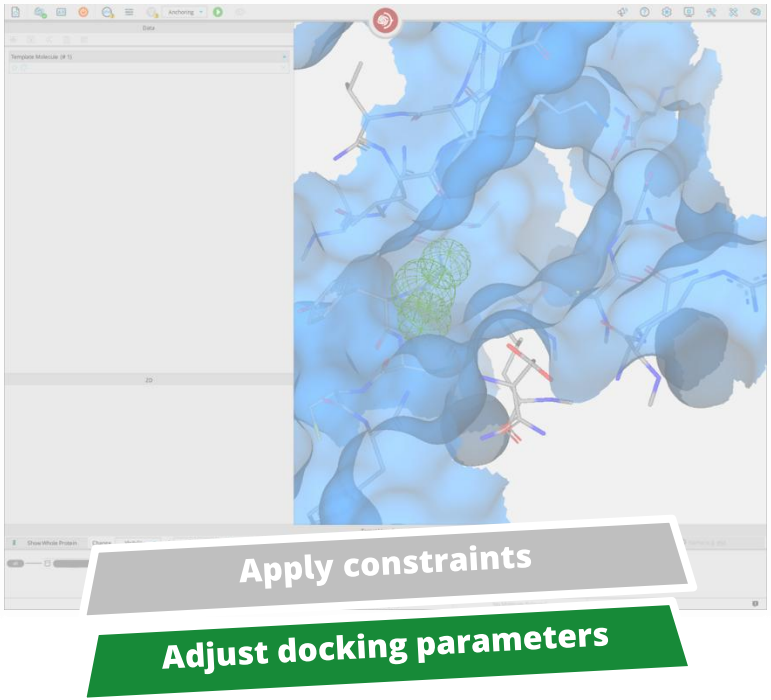
- You can load related PDB structures into the Protein Mode to align them and check for differences. The 'Hide components with similar conformation to binding site protein' feature of the Target View Control helps you to highlight binding site differences and to spot highly flexible rotamers (e.g., lysine, arginine, ...). Keep those in mind, when evaluation poses that form interactions with those side chains.
- A template molecule can be used for C-S-D. The template molecule can be a co-complexed ligand, the substructure of a known binder, or a docking prediction. A minumum common substructure of five atoms is required for matching. Transfer the molecule pose as template to the Space Docking Mode to perform a template-based C-S-D. Using a template accelerates the Anchoring Step of the C-S-D workflow severalfold, as an initial fragment is used as a seed. Keep in mind that not all anchors (molecules containing a linker) will match on the template and no poses will be generated for those.

Inside the Space Docking Mode: Before Starting the Workflow



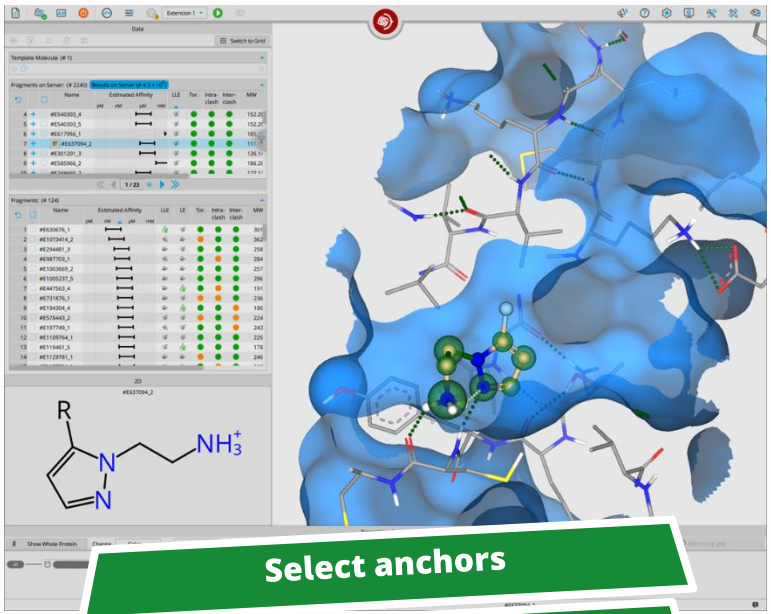
- Select desired docking parameters and apply **pharmacophore constraints** if desired.
  - Docking poses will be generated to comply with the constraints. This is not a post-filtering to remove poses that do not comply with the constraints.
  - You can apply constraints to define a starting sport for your anchors. If you would like to explore two different distant starting points for anchors, we recommend to run two different setups if the binding cavities are too far away, because the workflow will try it's best to generate a pose that satisfies the constraints which extends the runtime if both constraints are mandatory.
  - You can apply a pharmacophore constraint for the linker atom to define the extension direction.
  - You can apply SMARTS definition ('SMARTS (user defined)') to request particular motifs in the placed constraints.
  - You can exclude undesired features with constraints to avoid the growing into a particular direction like the solvent exposed site.

Inside the Space Docking Mode: Before Starting the Workflow (2)



- **No postprocessing of results with pharmacophore constraints is possible.** Generated poses cannot be filtered for pharmacophore features on the server. Define the features you want to see in your results before starting the docking run.
- Per default, five poses are generated during the Anchoring Step for each molecule. It is also possible to increase the number to 10 in the docking settings.
- Save the project once you are content with your setup and start the Anchoring Step. You can access your C-S-D project at any time by loading it into SeeSAR.

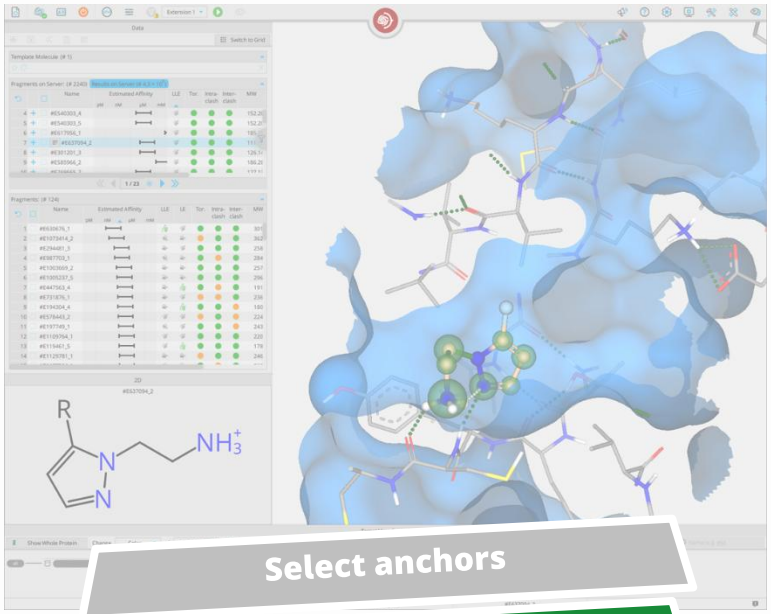
Anchoring Step: Placing the Starting Points



- After anchoring has finished, results are automatically filtered for better results and the best poses are presented. The scoring algorithm HYDE is unreliable for unfavorable torsions and clashes.
- You see all poses, removing duplicates is not possible (yet).
- Select chemically diverse anchors for the Extension Step. By enriching chemical diversity in each step, you cover a broader area of the chemical space and increase the likelihood of finding more active novel chemotypes.
  - The scoring favors larger anchors to display better predicted affinity because they they possess additional heavy atoms, which can form further interactions within the binding pocket. Yet, smallest fragments (e.g., a single pyridine with a linker) can end up as full-fledged drug candidates at the end of the workflow.
  - To check for smaller, valid anchors that form great interactions within the binding site, you can apply following filter or select parameters of it:

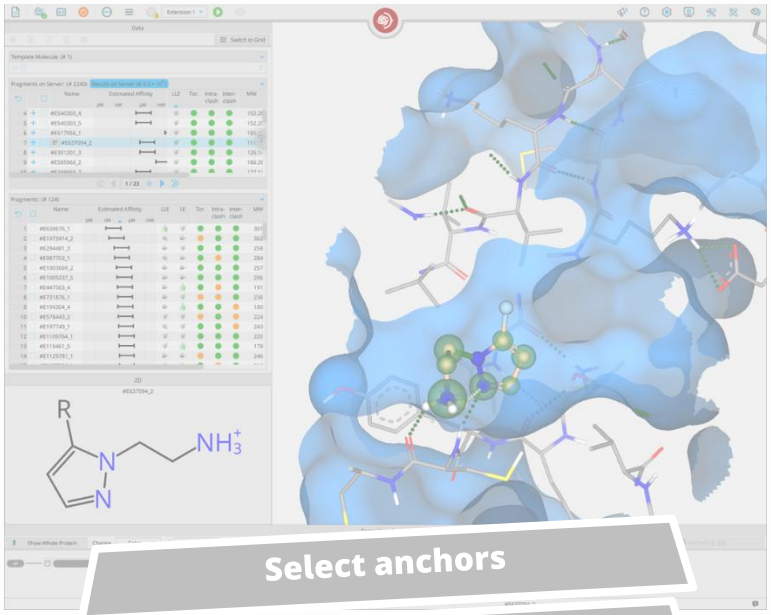
    [good LE (ligand efficiency) + molecular weight < 200 + a desired number of H-bond + min. 1 ring]

Anchoring Step:
Placing the Starting Points (2)



- We recommend to select 50 to 150 anchors for the Extension Step, and a maximum of 300.
- Tables can be sorted for different parameters. Filtering of poses happens on the hardware running HPSee.
- Based on your binding site, you may end up with very flexible compounds. It is recommended to reduce the number of rotatable bonds to max. 3 with a filter.
- Some anchors are already finished compounds. Finished compounds are appearing the in the 'Result Window' instead. This applies to the Extension Step as well, so please check there for finished result compounds.

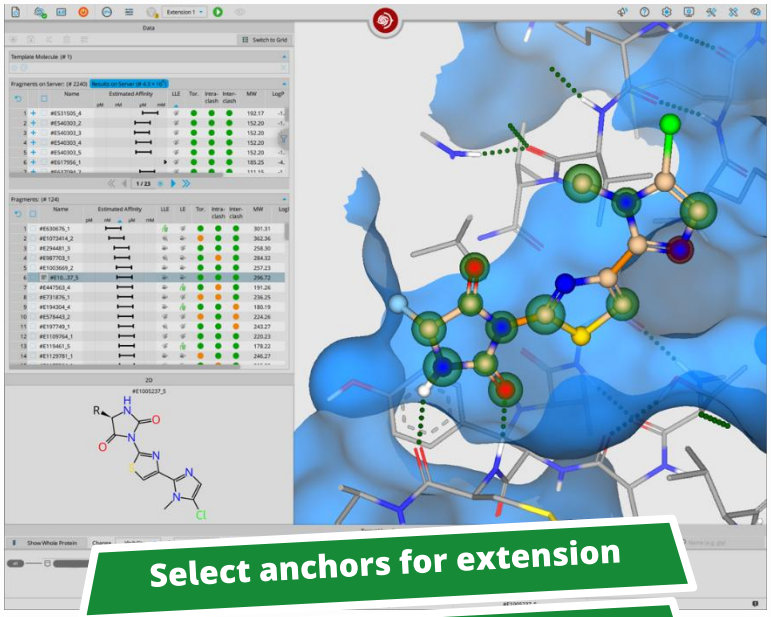
Anchoring Step: Placing the Starting Points (3)



After selecting the anchors for the Extension Step:

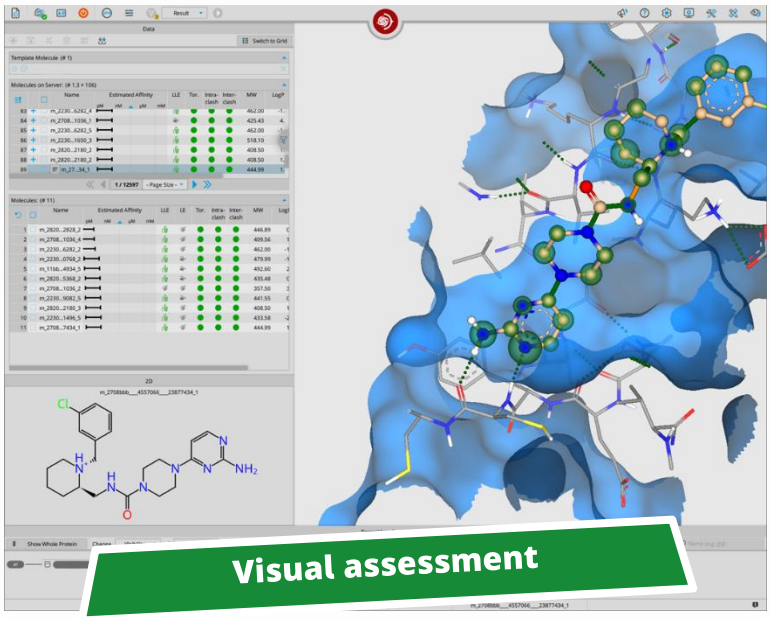
- Setting new pharmacophore constraints is possible for the Extension Step.
- You can apply a new constraint for a linker, to define where the last extension should happen.
- Sublibraries for all selected anchors will be enumerated after starting the Extension Step.
- An anchor may lead to millions of extended structures that will be enumerated and docked afterwards. Therefore, if you are unsure about a fragment, do not pick it.
- **Structurally very similar anchors can lead to different compounds.** Be generous when selecting anchors.
- Per default, five poses are generated during the Extension Step for each extended molecule. It is also possible to increase the number up to 10 in the docking settings.

Extension Step 1: Expanding the Anchors



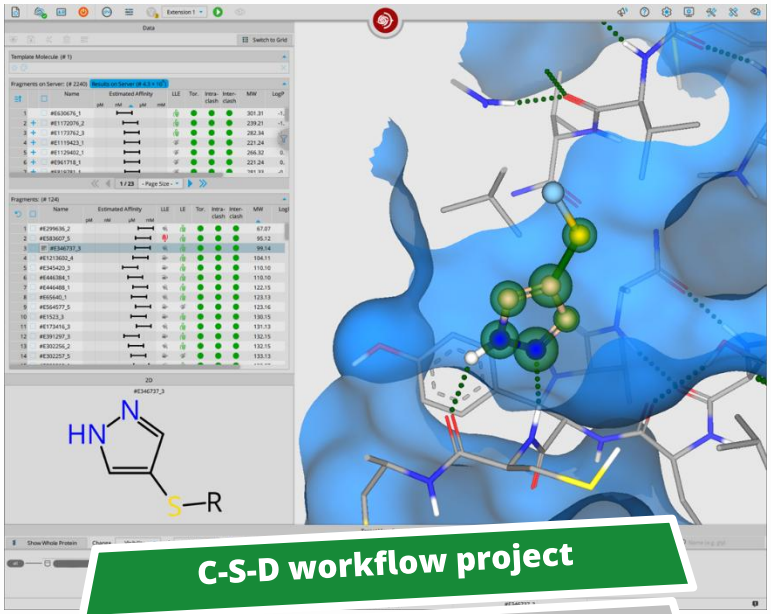
- **Stereochemistry:** For all generated finished compounds and extended anchors all stereoisomers (E/Z, R/S) can emerge during docking.
- All sublibraries are docked based on the initial template anchor, which significantly optimizes computation time while maintaining the integrity of the original binding mode.
- Select extended anchors for the last Extension Step. Again, pay attention to select chemically diverse molecules. Feel free to select as many as you like and feel fitting.
- Sometimes it is possible that no second extension is possible which depends on the selected fragments. Check the results window.
- For the second Extension Step, you can set new pharmacophore and linker constraints. If a molecule does not longer contain a linker, the linker constraint will be ignored.

Extension Step 2:
Finishing the C-S-D Workflow



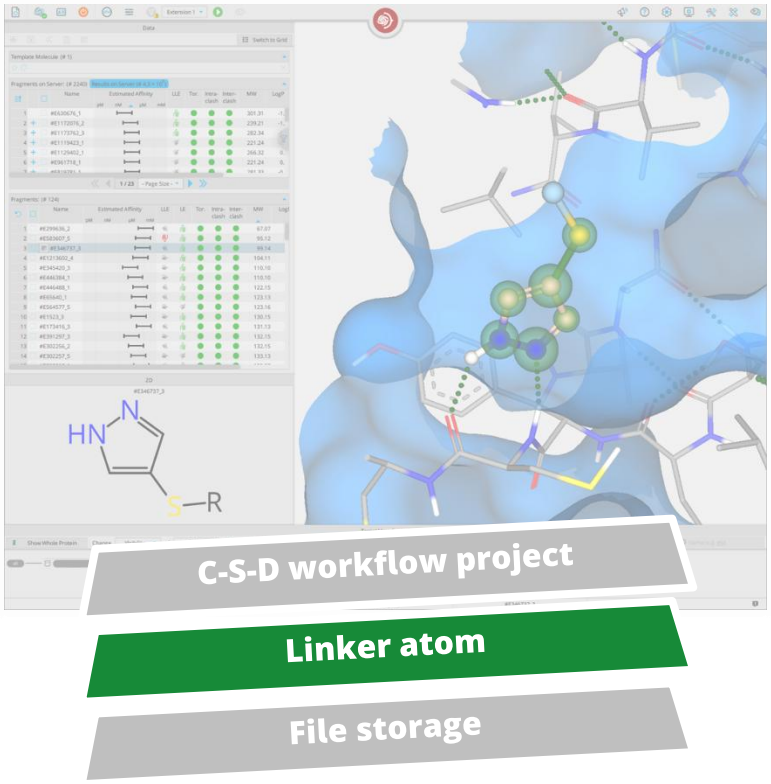
- Bulk download of max. 50k results is possible. Apply filters to download only the relevant ones.
- You can load the finished, selected compounds into the **Analyzer Mode** or load the bulk download into it for convenient visual assessment on your local hardware.
- **Caveat:** Once a molecule was loaded from the Space Docking Mode it cannot be returned back. The Space Docking Mode is a one-way street.
- Again, pay attention to chemical diversity of your selection.

Nice To Know



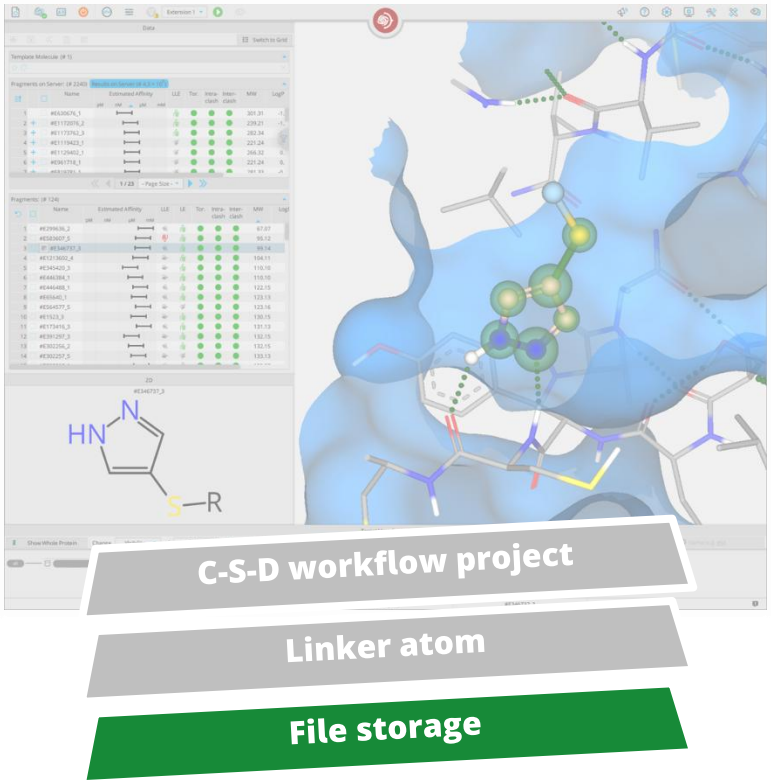
- The Chemical Spaces contain compounds that can be synthesized within one or two steps.
- It is not possible to filter for interactions after the docking in the Space Docking Mode. Transfer results to Analyzer Mode to do so. Pay attention that results cannot be pushed back into the Space Docking Mode after that.
- You can access the project by loading the project file and browsing through the steps of C-S-D workflow.
- Your project file is the only access to the file. Copying the project and loading it removes the results from the original file.

Nice To Know (2)



- The linker atom has a volume of a carbon atom. Furthermore, it samples for potential clashes with the target surface with another carbon-sized phantom atom to avoid placement in a dead-end. Imagine the linker atom as an ethyl group to understand the surface sampling. This behavior of the linker atom cannot be changed.
- The linker conformation is not fixed and is flexible during extension.
- The linker atom is chemically inert and cannot form interactions with the target structure. Compared to other synthon-based approaches, it is not necessary to dock the smallest product, which avoids introducing a scoring bias towards the anchor.
- The α-β-linker-atom torsion is colored at the moment. This is without any meaning but is still filtered out with the default filter if it is a red torsion, affecting the poses you can see. This will be changed in the SeeSAR 14.1 release to grey, no longer affecting the filtering process.

Nice To Know (3)



- There is no possibility to manually connect a SeeSAR project to a C-S-D workflow.
- For each step, a sd file is created on HPSee that can manually be downloaded via API.

How can I order compounds from on-demand Chemical Spaces?

After performing a search with infiniSee your results will be presented in a table. The column "Source" tells you the origin of the Chemical Space that contains your solution; the ID of the respective result molecule is shown in the "Name" column.
Compounds can be ordered by sending a quote request to the compound vendor with the following information:
Requested structures in SMILES or SD format, Compound ID (concatenated), and amount requested.

- For compounds from Ambinter's AMBrosia Space, send your request to ambrosia@greenpharma.com.
- For compounds from eMolecule's eXplore Space, send your request to explore@emolecules.com.
- For compounds from Enamine's REAL Space, send your request to libraries@enamine.net.
- For compounds from WuXi's GalaXi Space please send your request to contact@labnetwork.com
- For compounds from OTAVA's CHEMriya Space please send your request to info@otava.ca.
- For compounds from Chemspace's Freedom Space please send your request to sales@chem-space.com.

Have fun and enjoy your interactive drug discovery journey with SeeSAR!

If you have any problems, please reach out to us:
support@biosolveit.de