REC

First steps in protein-ligand docking with GOLD

CCDC Virtual Workshop

October 2024

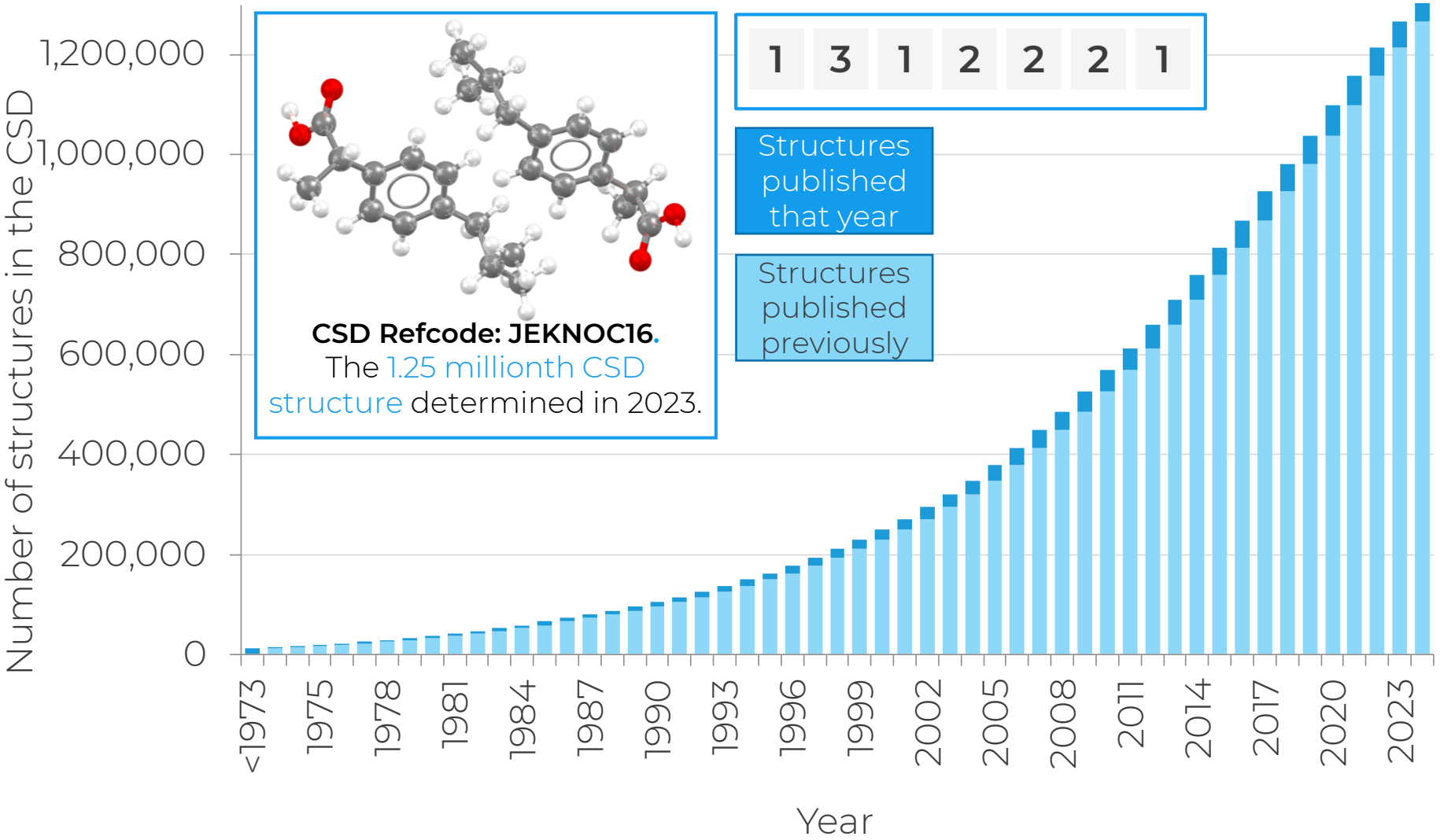
CCDC
advancing structural science

Learning outcomes for today

- The basics of GOLD and the Hermes interface.
- Step-by-step preparation of protein, and ligand(s) in GOLD.
- How to set-up and run a standard protein ligand dock.
  - What configuration options.
  - How to run GOLD in the background or interactive mode.
- How to identify the correct binding modes reliably and with confidence.
- The basics of how GOLD can be used in virtual screening and lead optimisation.
  - Through example case studies.

CCDC

The Cambridge Structural Database



- Every published structure
  - Inc. ASAP & early view
  - *CSD Communications*
  - Patents
  - University repositories
  - Thesis
- Every entry enriched and annotated by experts
- Discoverability of data and knowledge
- Sustainable for over 59 years
- A trusted CoreTrustSeal repository

Certified as Trustworthy by CoreTrustSeal

Inside the Cambridge Structural Database

The CSD is a database of all the published organic and metal-organic experimental crystal structures

Organic
45%

Metal-Organic
55%

At least one transition metal, lanthanide, actinide or any of Al, Ga, In, Tl, Ge, Sn, Pb, Sb, Bi, Po

Not Polymeric
89%

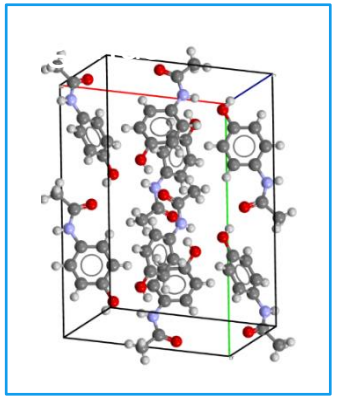
Polymeric: 11%

Single Component
58%

Multi Component
42%

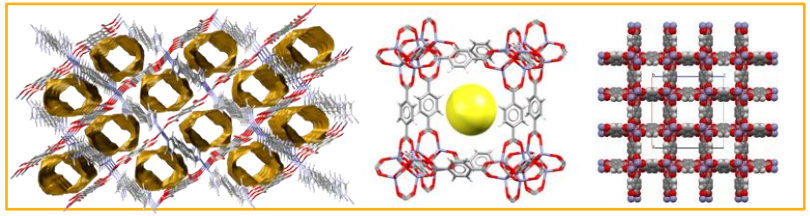
Organic

- Drugs
- Agrochemicals
- Pigments
- Explosives
- Protein ligands



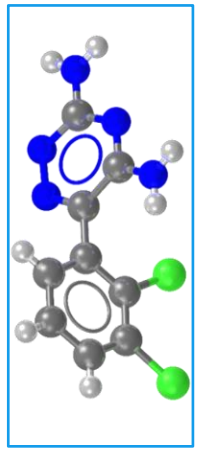
Metal-Organic

- Metal Organic Frameworks
- Models for new catalysts
- Porous frameworks for gas storage
- Fundamental chemical bonding



Additional data

- 13,478 polymorph families
- 174,987 melting points
- 1,075,904 crystal colours
- 951,746 crystal shapes
- 30,275 bioactivity details
- 13,641 natural source data
- > 350,000 oxidation states



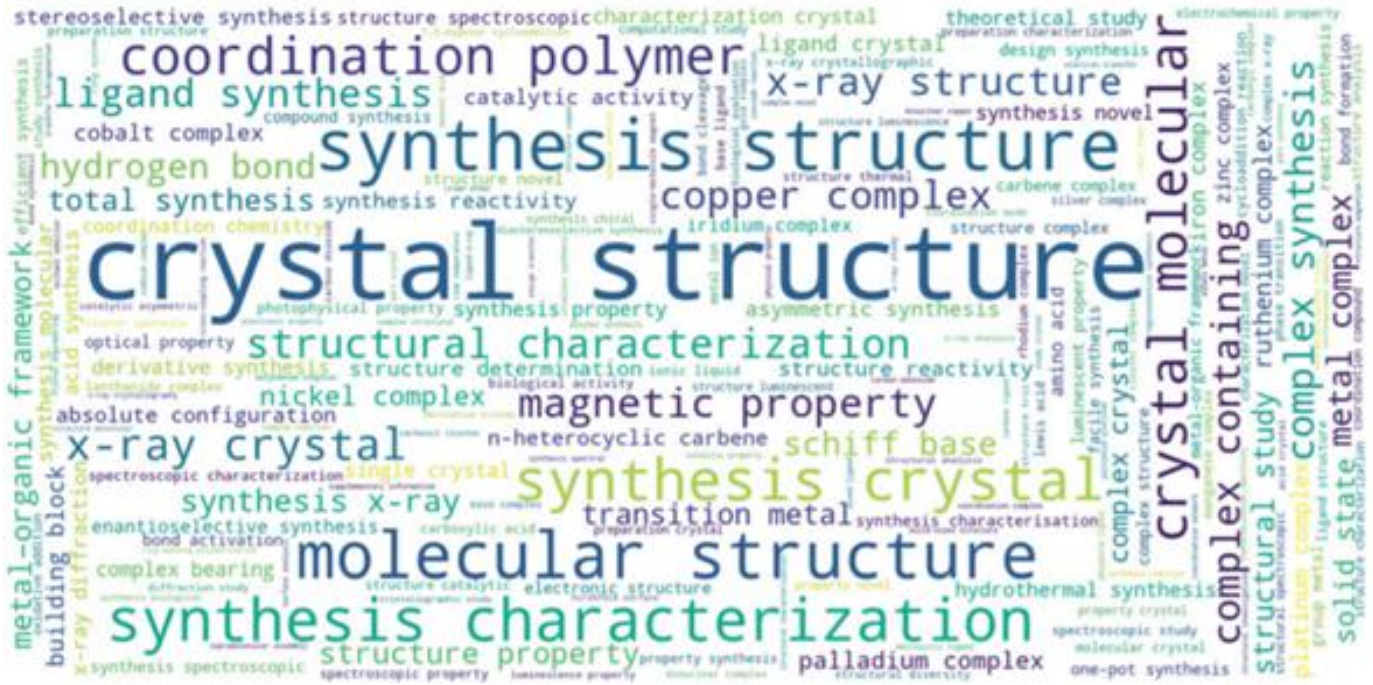
Links and subsets

- DrugBank
- Druglike
- MOFs
- PDB ligands
- PubChem
- ChemSpider
- Pesticide PDB

The CSD in research

Subject category	Citations
Crystallography	3030
Chemistry Multidisciplinary	2746
Chemistry Inorganic Nuclear	1096
Chemistry Physical	1002
Materials Science Multidisciplinary	919
Biochemistry Molecular Biology	413
Physics Atomic Molecular Chemical	310
Chemistry Organic	253
Chemistry Medicinal	185
Nanoscience Nanotechnology	162
Pharmacology Pharmacy	134

The *Web of Science* subject categories that cite the 2016 standard CSD reference* most frequently.



A word cloud of common bigrams in the titles of publications containing CSD-compliant crystal structures.
Based on analysis in P. Willett *et al, CrystEngComm*, 2020,22, 7233-7241
DOI: 10.1039/D0CE00045K

C. R. Groom. et al, Acta. Cryst. B, 2016, 72, 171-179-7241, DOI10.1107/S2052520616003954

CCDC

Small molecules, big impact

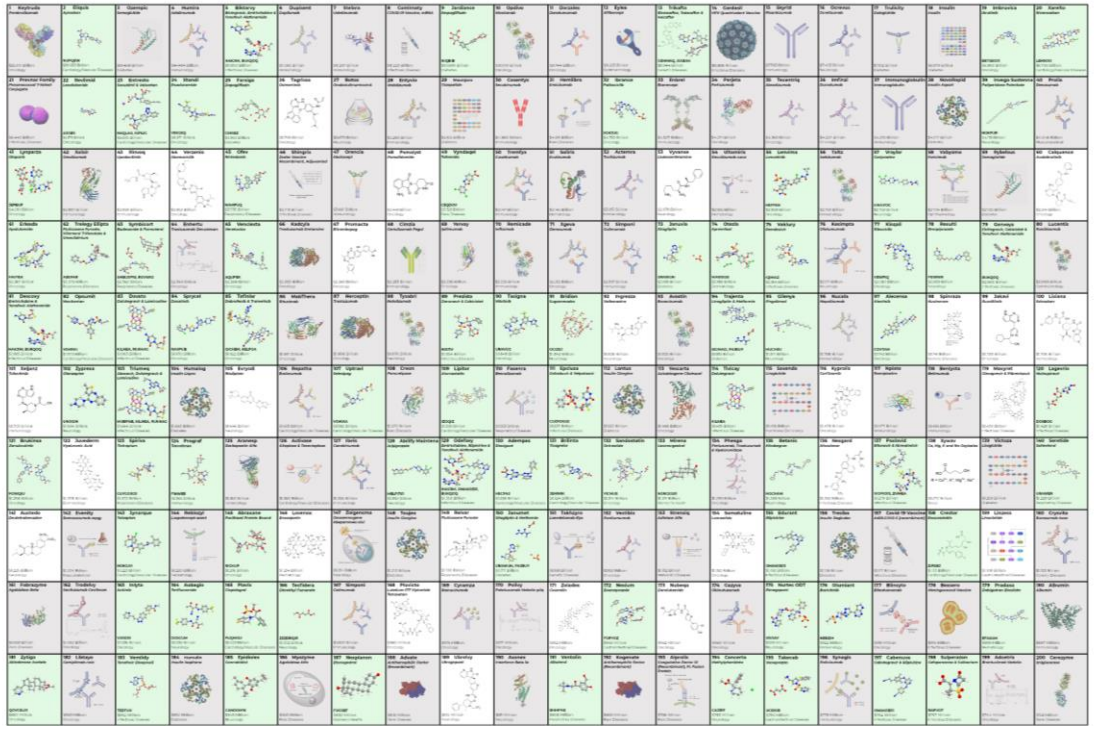
200 top drugs by retail sales in 2023

Small molecule structure in the CSD

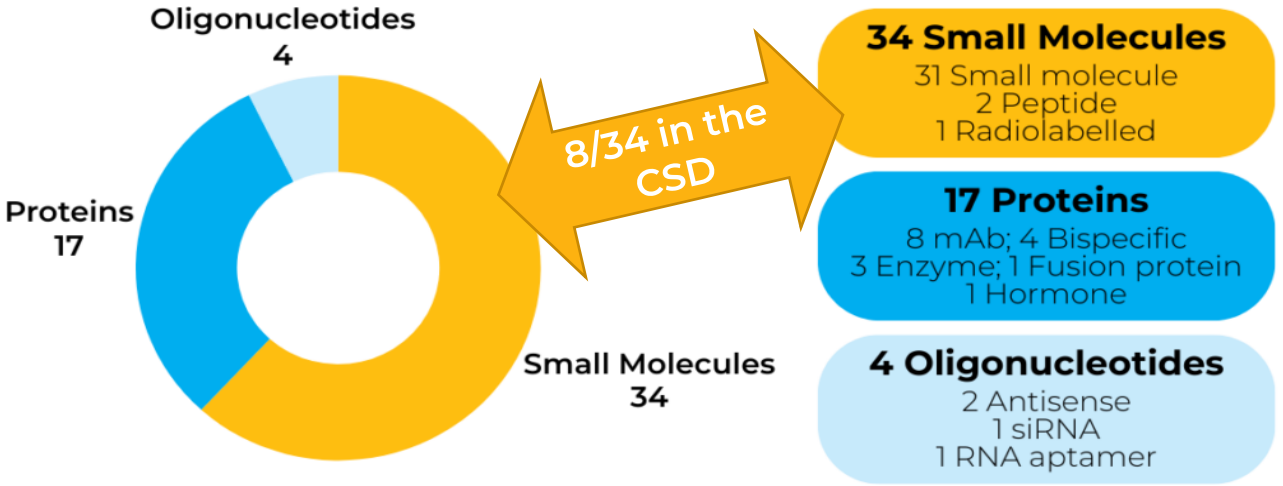
- Open-up new biological and therapeutic opportunities
- Compete against emergent modalities for rare diseases
- Are used to target RNA*

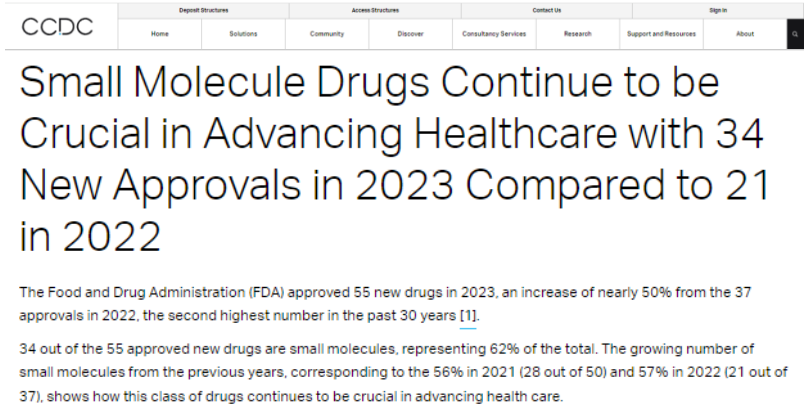
Small molecule structure CSD entry pending

Biological structure in the PDB



FDA novel drug approvals 2023





CCDC

Small Molecule Drugs Continue to be Crucial in Advancing Healthcare with 34 New Approvals in 2023 Compared to 21 in 2022

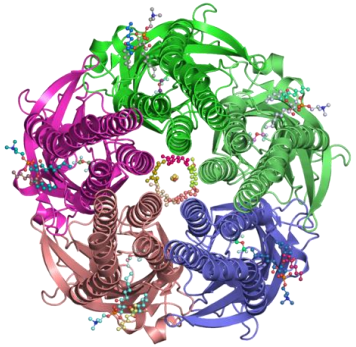
The Food and Drug Administration (FDA) approved 55 new drugs in 2023, an increase of nearly 50% from the 37 approvals in 2022, the second highest number in the past 30 years [1].

34 out of the 55 approved new drugs are small molecules, representing 62% of the total. The growing number of small molecules from the previous years, corresponding to the 56% in 2021 (28 out of 50) and 57% in 2022 (21 out of 37), shows how this class of drugs continues to be crucial in advancing health care.

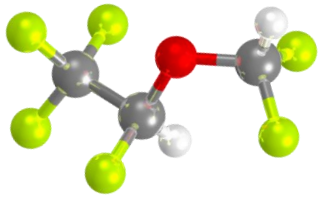
Adapted from poster compiled and produced by the Njardarson Group (The University of Arizona)
Nature Reviews Drug Discovery* **20, 85-90 (2021)

CCDC

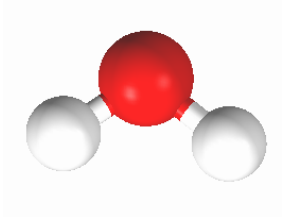
Proteins and ligands



Proteins are large biomolecules and macromolecules that comprise of one or more long chains of amino acids.



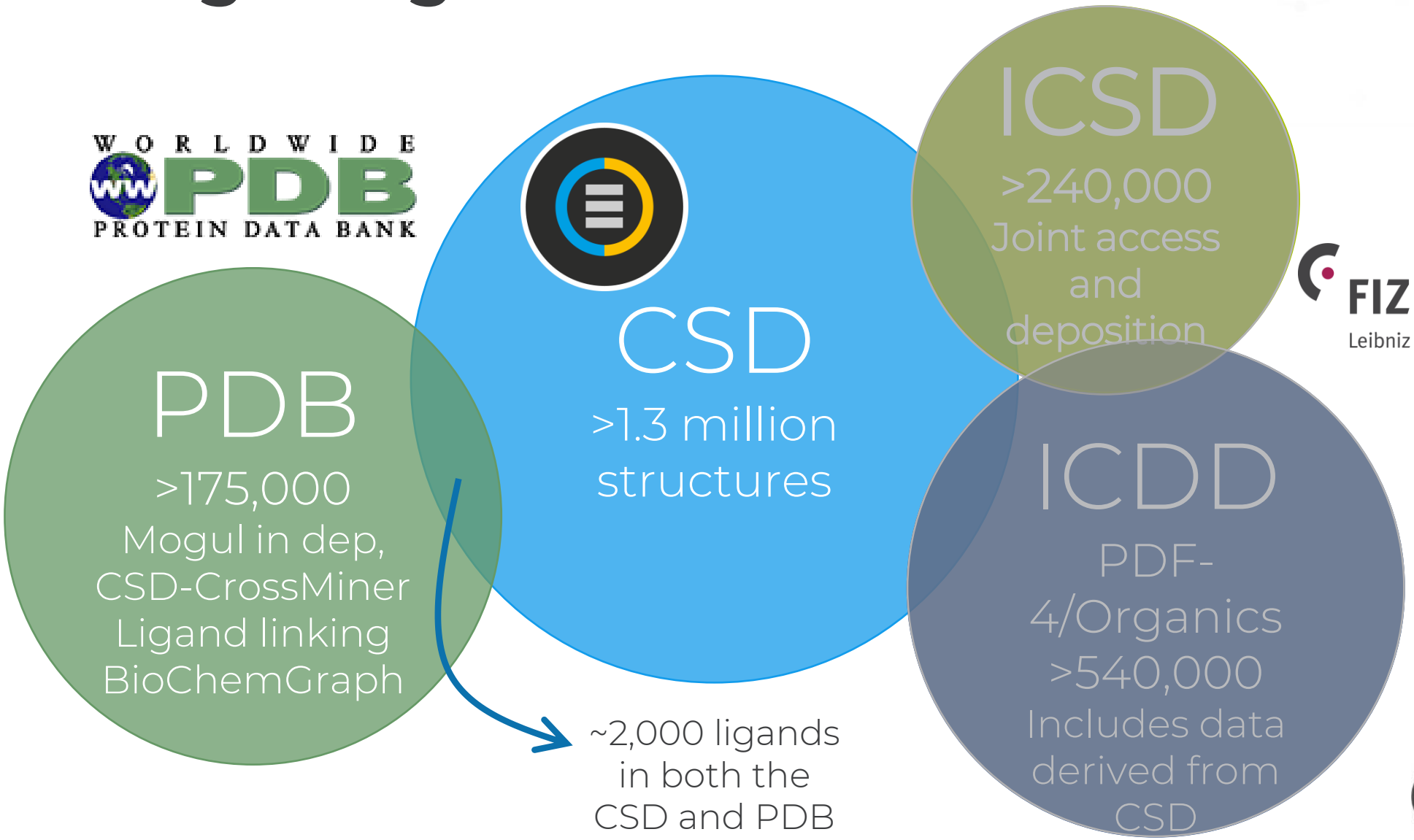
Ligands are small molecules that bind to the protein and can change the protein function.



Functional waters are found in the binding site of a protein and mediate the interactions between the ligand and the protein.

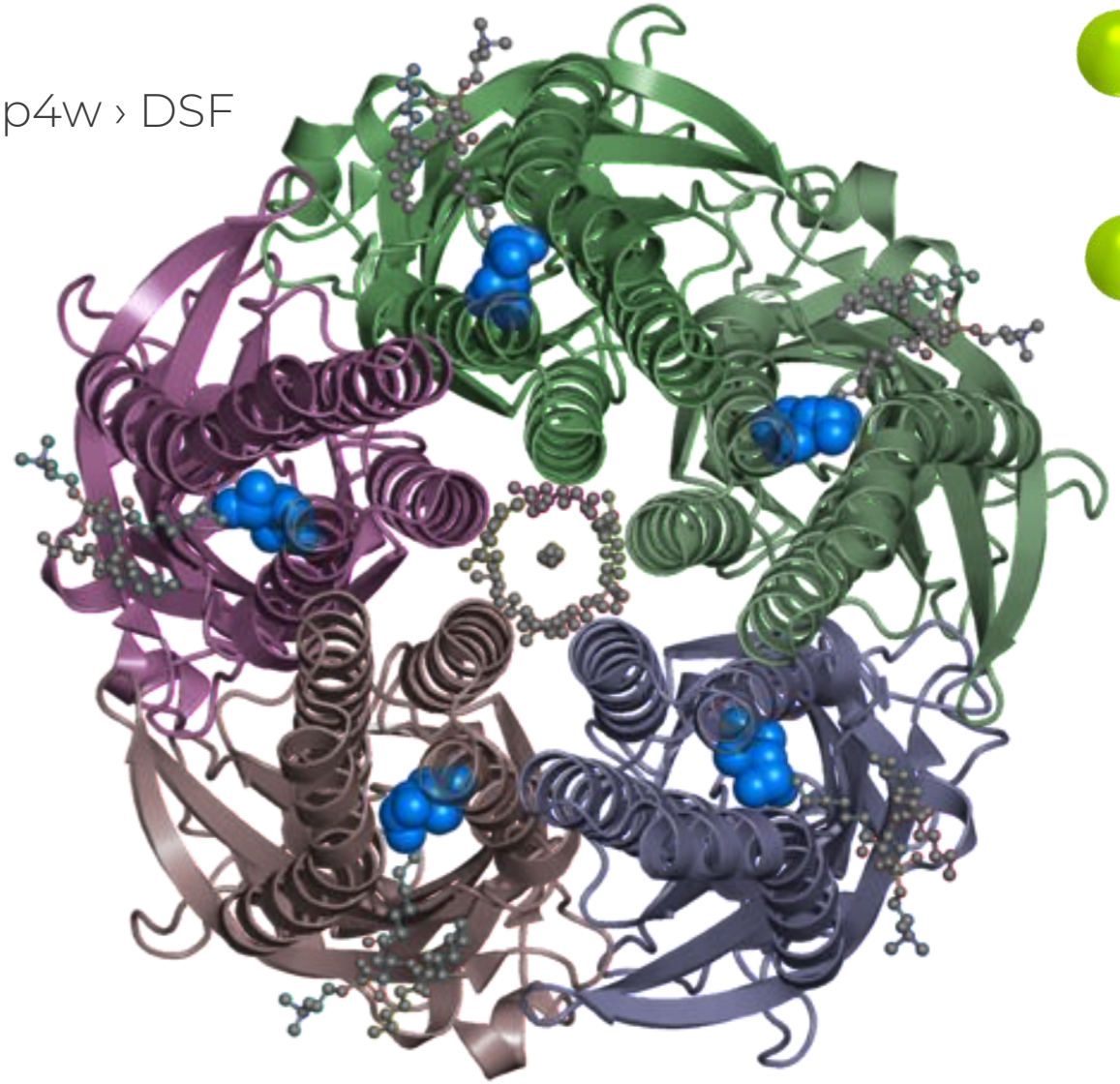
CCDC

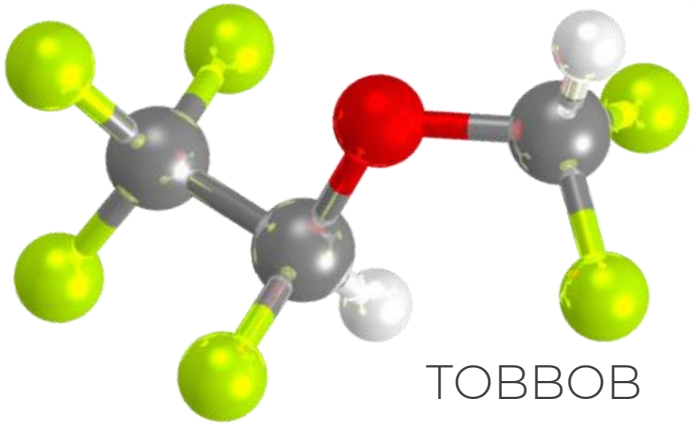
Using integrated structural databases



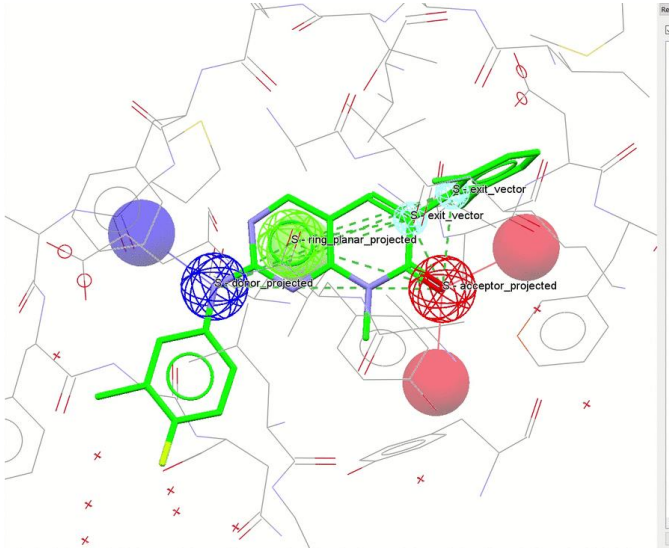
Connecting chemistry and biology

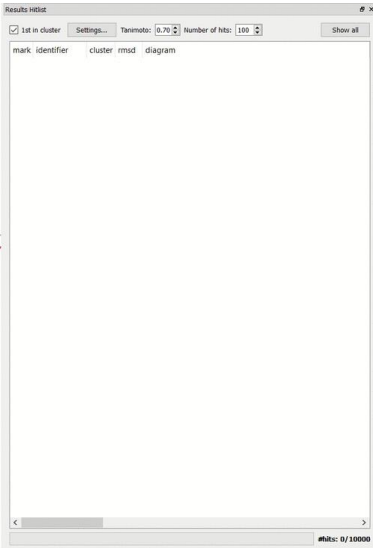
3p4w › DSF





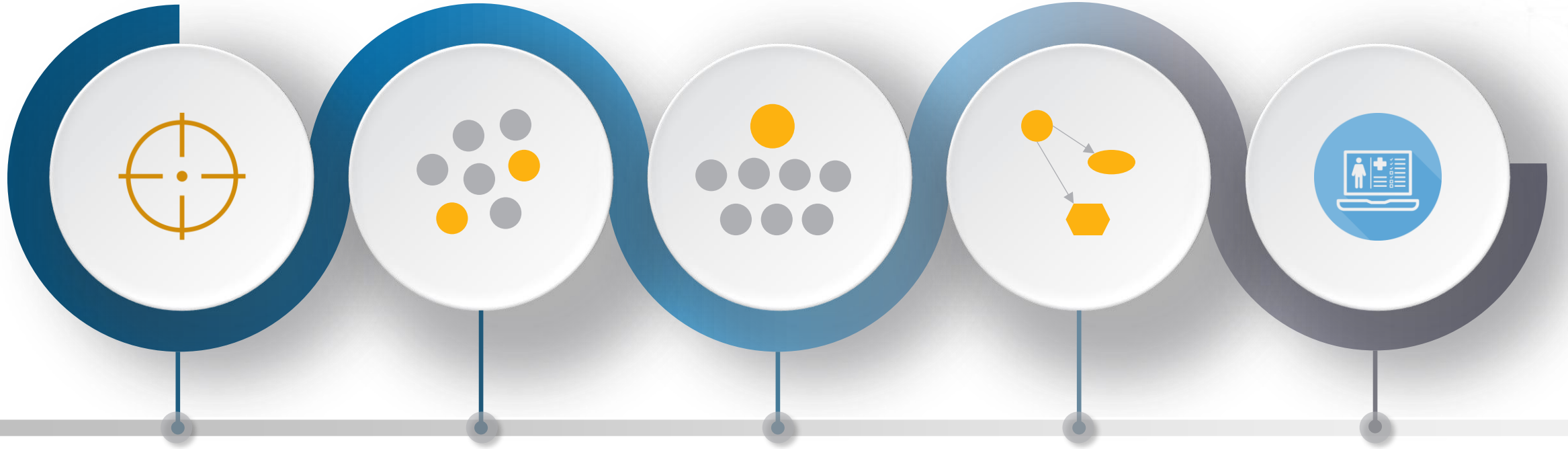
TOBBOB





CCDC

Drug Discovery Pipeline



Target selection

Hit-identification

Structure- based virtual screening.

Hit to lead

Assess how changes affect binding.
Optimize compound geometry.
Predict binding of small molecules to active pockets in proteins.

Lead optimisation

Check the impact of changes with docking pose prediction.
Understand how changes affect conformations.

Drug development

CCDC

The CSD Portfolio

CSDEnterprise.

CSDCore.

Hermes

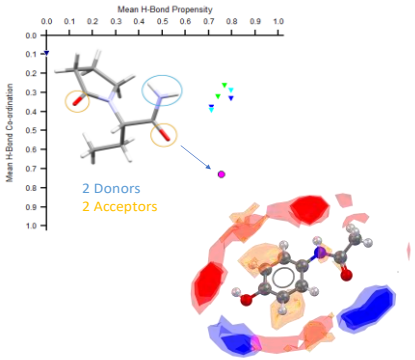
CSDDiscovery.

Design of new molecules

GOLD

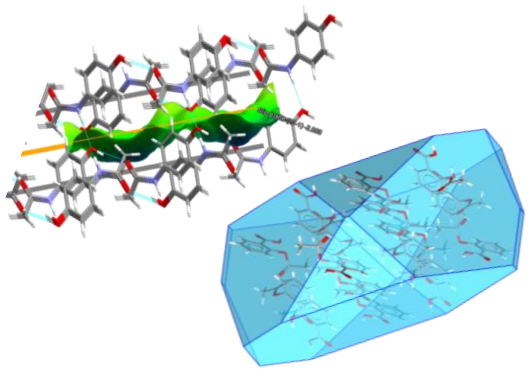
CSDMaterials.

Assessment of solid form stability and properties



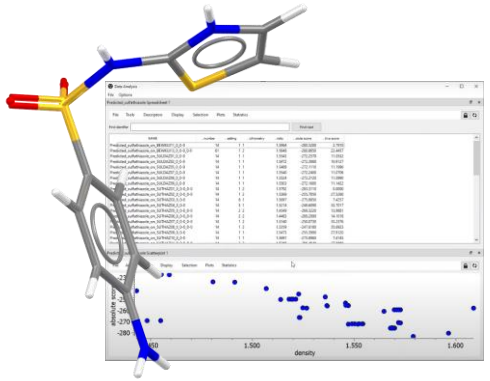
CSDParticle.

Anticipate particle properties and behaviour



CSDTheory.

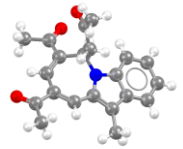
Generate solid form landscapes

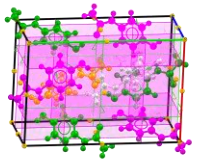


CSDCommunity.

Deposit, publish, access and visualise structural data

Free functionality to share and learn from structures



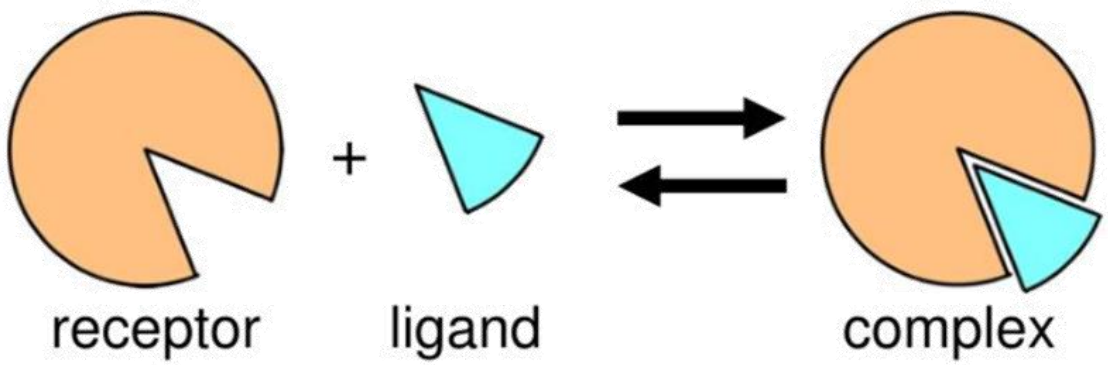


CSDU

Medicinal & Computational Chemists ♦ Crystallographers & Structural Biologists ♦ Solid Form & Crystallisation Scientists ♦ Functional Materials Scientists ♦ Educators ♦ Industry and Academia

CCDC

Docking



Docking studies are computational techniques for the exploitation of the possible binding modes of a substrate to a given receptor, enzyme or other binding site.

Glossary of Terms Used In Computational Drug Design, Pure&Appl. Chem.,Vol. 69, No. 5, pp. 1137-1152, 1997

CCDC

GOLD: Protein-Ligand Docking Software

- **GOLD (Genetic Optimisation for Ligand Docking)** is a genetic algorithm for docking flexible ligands into protein binding sites.

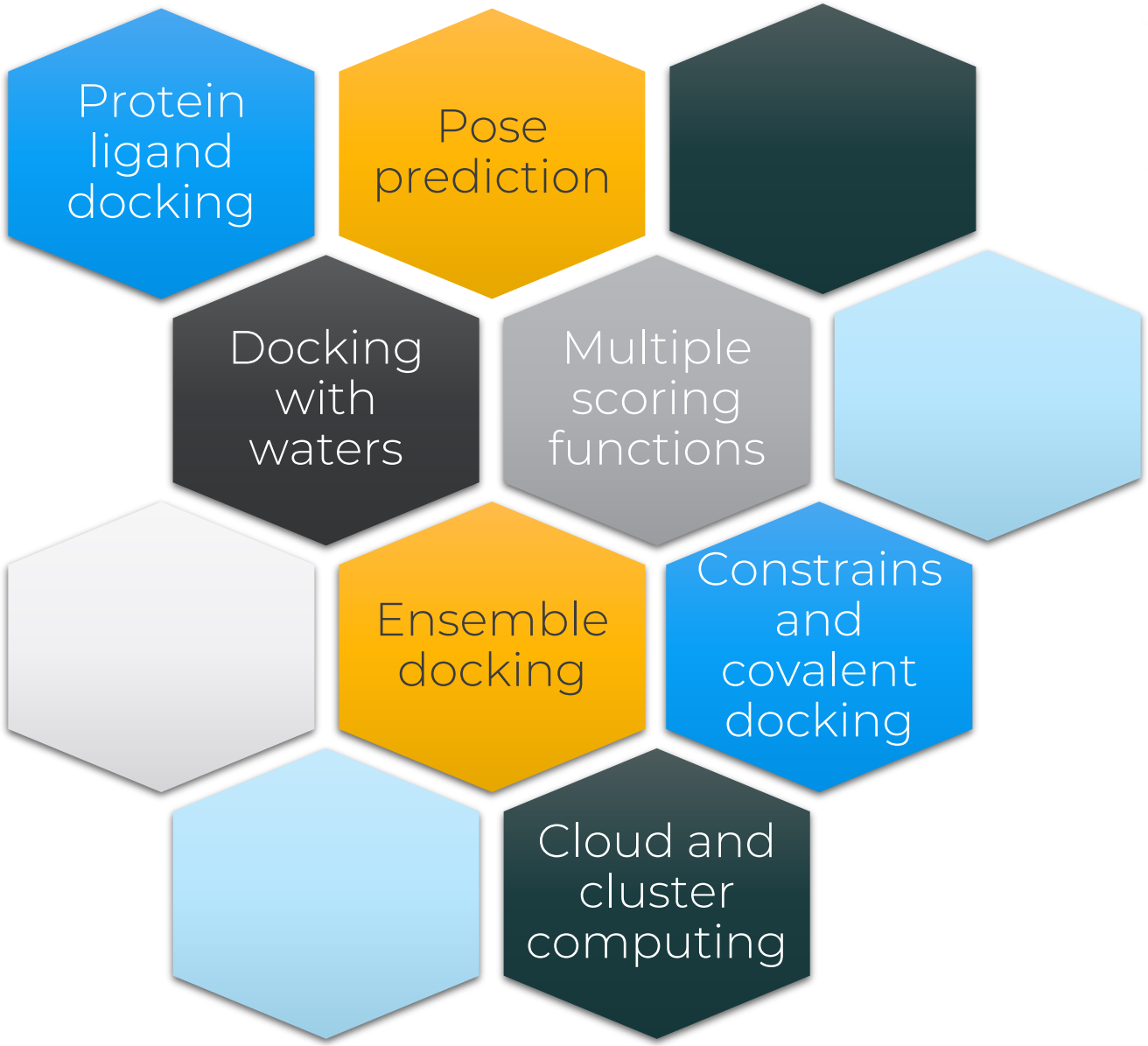
- **GOLD** has proven success in virtual screening, lead optimisation, and identifying the correct binding mode of active molecules.

- Relied on by researchers in academia and industry worldwide.

- Reliable
- Flexible
- Configurable

CCDC

GOLD: All in one molecular docking package



CCDC

What are we going to learn today?

- Protein-ligand docking of a Kinase inhibitor
  - We would learn how to…
    - Import a protein co-crystal from the Protein Data Bank (PDB).
    - Prepare the protein crystal structure for docking.
    - Perform molecular docking experiment in GOLD.
    - Analyse the results obtained from docking experiment.

Molecular Recognition of Receptor Sites Using a Genetic Algorithm with a Description of Desolvation, J. Mol. Biol., Vol. 245, pp. 43-53, 1995.

CCDC

Steps in molecular docking

Target selection

- Co-crystal structure of protein-ligand complex
- from PDB

Protein Preparation

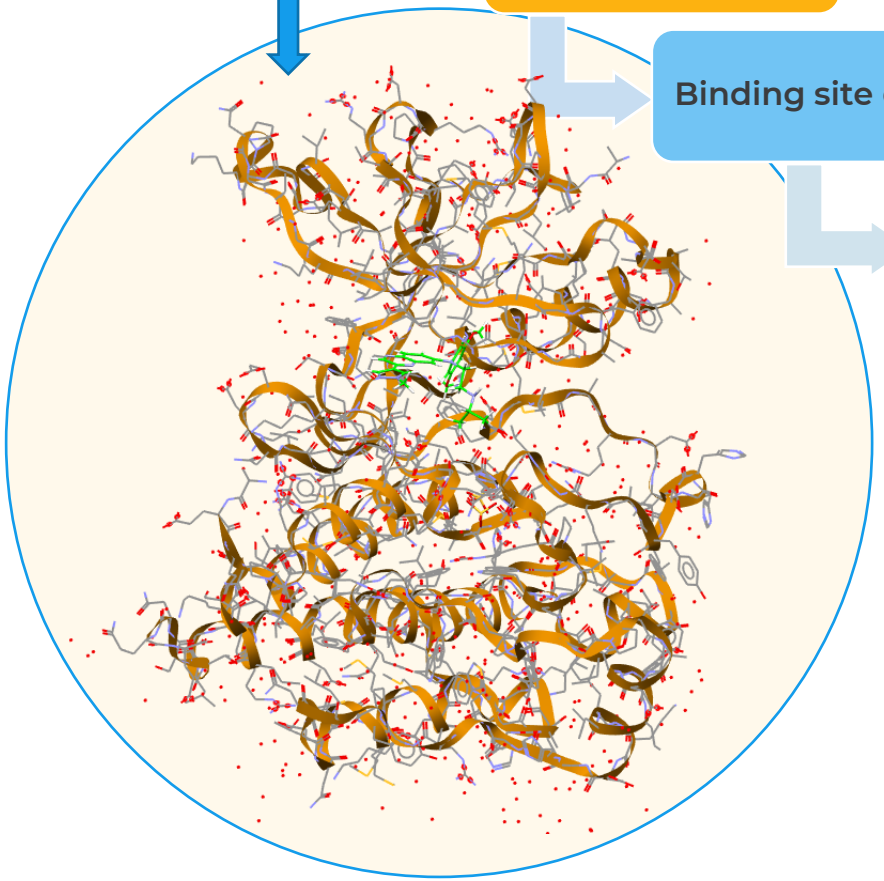
Binding site definition

Ligand(s) selection for docking

Scoring function selection

Setup and running the docking calculation

Analysis of the results and visualisation



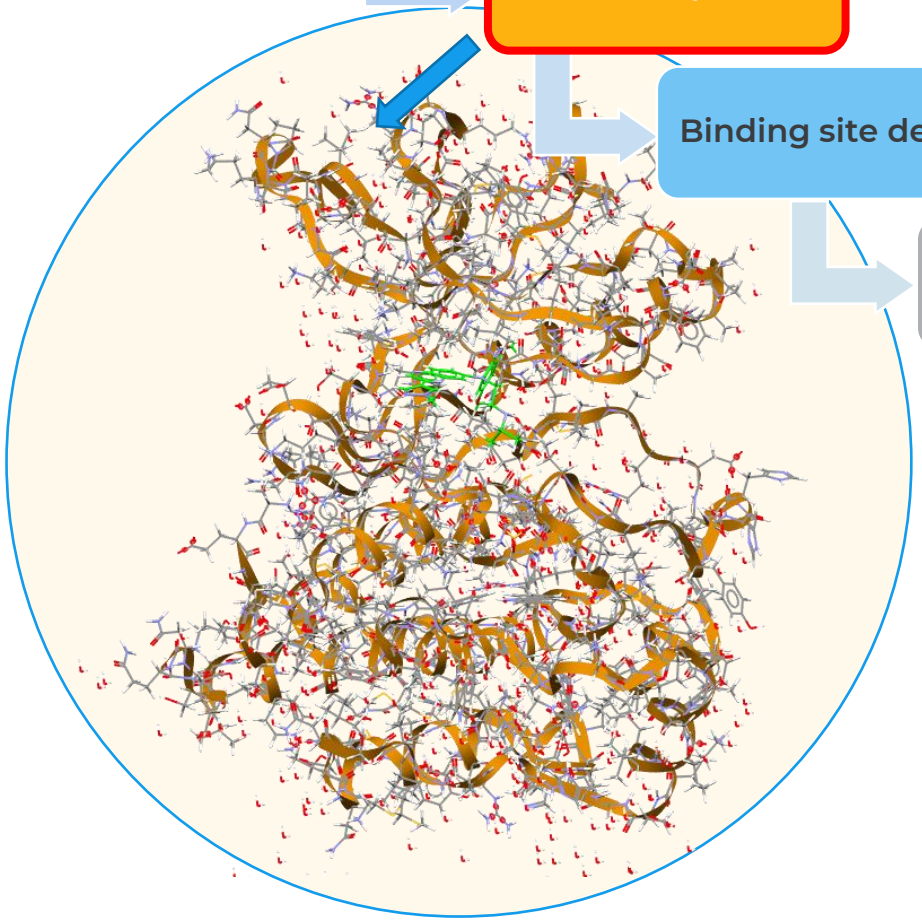
CCDC

Steps in molecular docking

Target selection

Protein Preparation

- Add hydrogens



Binding site definition

Ligand(s) selection for docking

Scoring function selection

Setup and running the docking calculation

Analysis of the results and visualisation

CCDC

Steps in molecular docking

Target selection

Protein Preparation

- Add hydrogens
- Check protonation of His, Gln, Asn

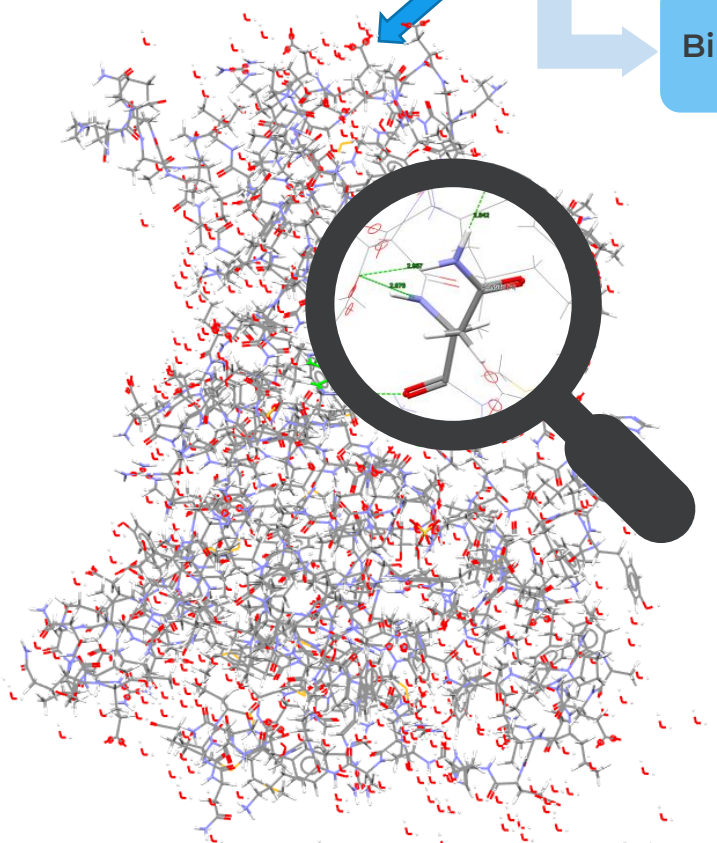
Binding site definition

Ligand(s) selection for docking

Scoring function selection

Setup and running the docking calculation

Analysis of the results and visualisation



CCDC

Steps in molecular docking

Target selection

Protein Preparation

- Add hydrogens
- Check protonation of His, Gln, Asn.
- Remove waters or retain required ones

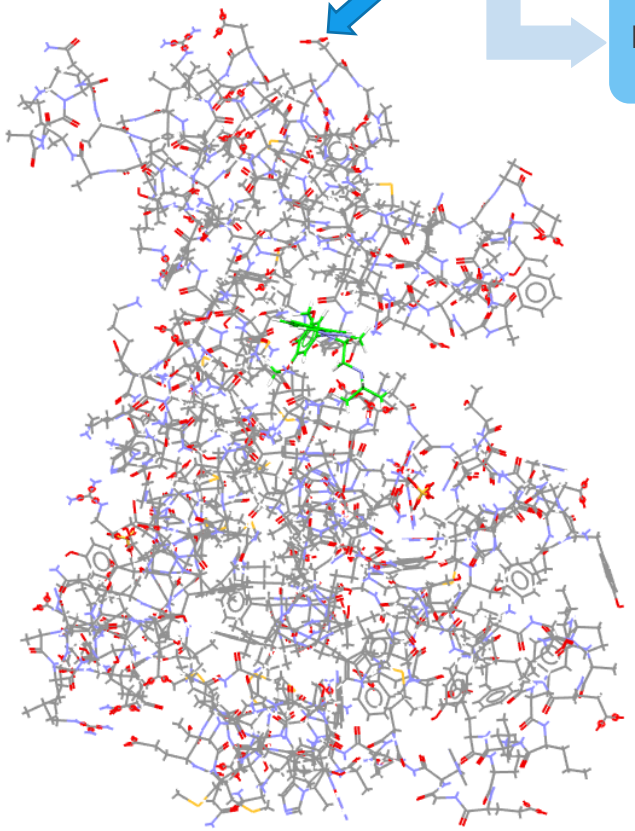
Binding site definition

Ligand(s) selection for docking

Scoring function selection

Setup and running the docking calculation

Analysis of the results and visualisation



CCDC

Steps in molecular docking

Target selection

Protein Preparation

- Add hydrogens
- Check protonation of His, Gln, Asn.
- Remove waters or retain required ones
- Extract ligand and save

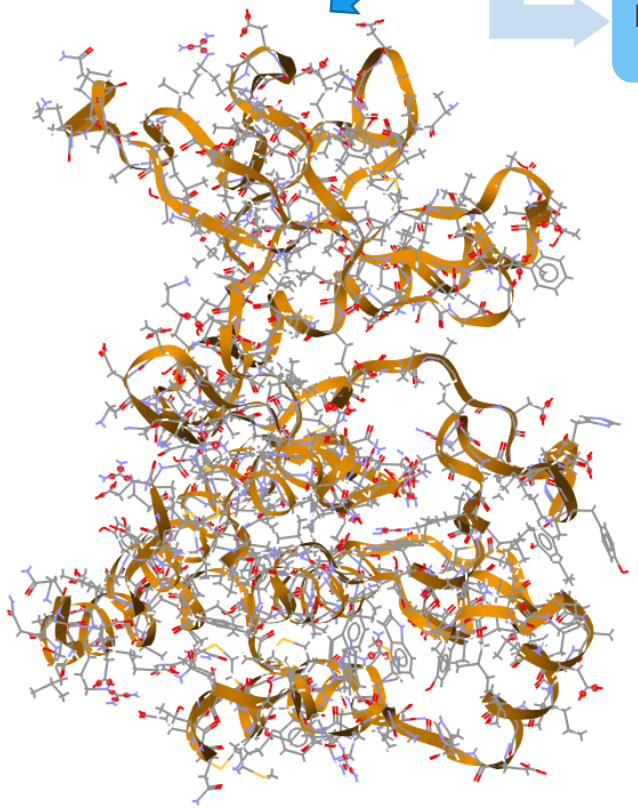
Binding site definition

Ligand(s) selection for docking

Scoring function selection

Setup and running the docking calculation

Analysis of the results and visualisation



CCDC

Steps in molecular docking

Target selection

Protein Preparation

Binding site definition

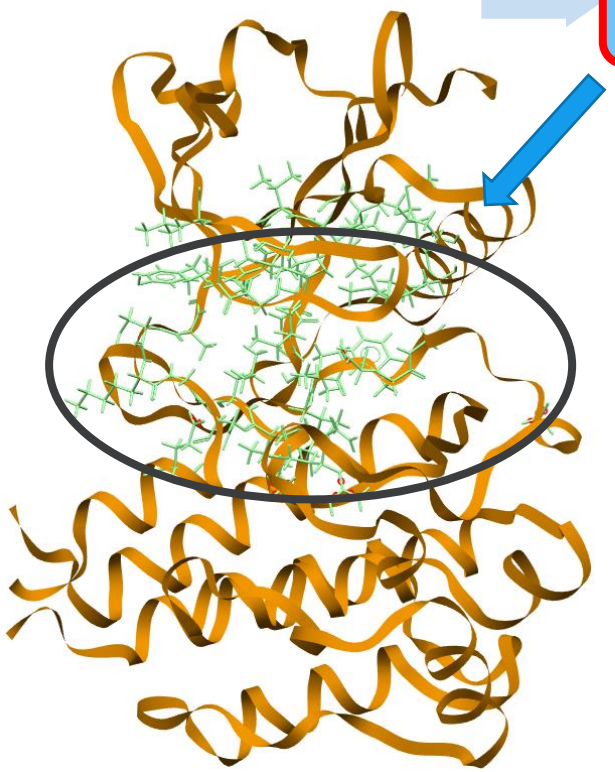
- Select the binding site for docking

Ligand(s) selection for docking

Scoring function selection

Setup and running the docking calculation

Analysis of the results and visualisation



Binding site: A specific region (or atom) in a molecular entity that is capable of entering into a stabilizing interaction with another molecular entity.

Read more in the Glossary in the handout .

CCDC

Steps in molecular docking

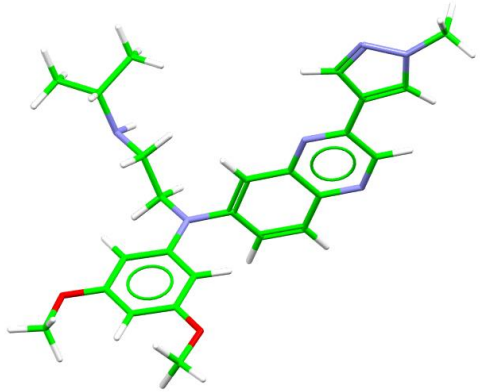
Target selection

Protein Preparation

Binding site definition

Ligand(s) selection for docking

- SDF, MOL2



Scoring function selection

Setup and running the docking calculation

Analysis of the results and visualisation

CCDC

Steps in molecular docking

Target selection

Protein Preparation

Binding site definition

Ligand(s) selection for docking

Scoring function selection

- Optionally rescoring

Setup and running the docking calculation

Analysis of the results and visualisation

Scoring functions in GOLD

- GoldScore
- ChemScore
- ASP (the Astex Statistical Potential)
- ChemPLP

Learn more in the FAQ

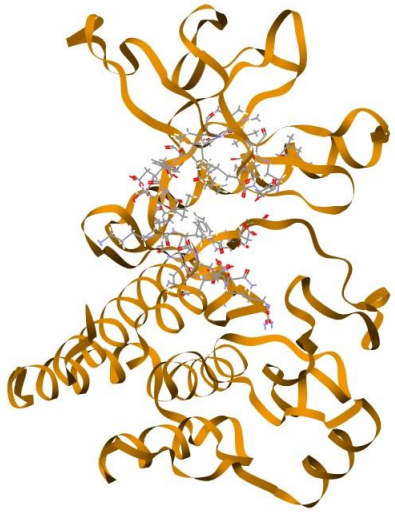
Scoring functions: mathematical functions used to approximately predict the binding affinity between two molecules after they have been docked.

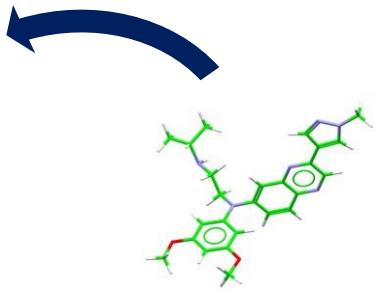
Read more in the Glossary in the handout .

CCDC

Steps in molecular docking

- Target selection
- Protein Preparation
- Binding site definition
- Ligand(s) selection for docking
- Scoring function selection
- Setup and running the docking calculation
- Analysis of the results and visualisation

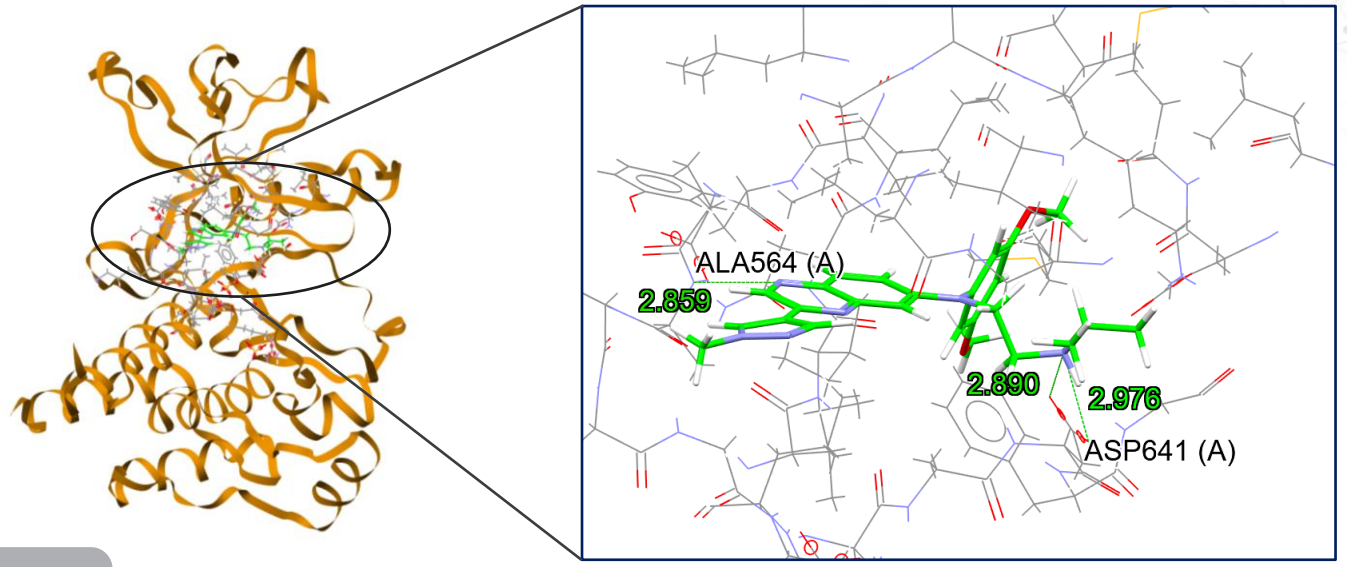




CCDC

Steps in molecular docking

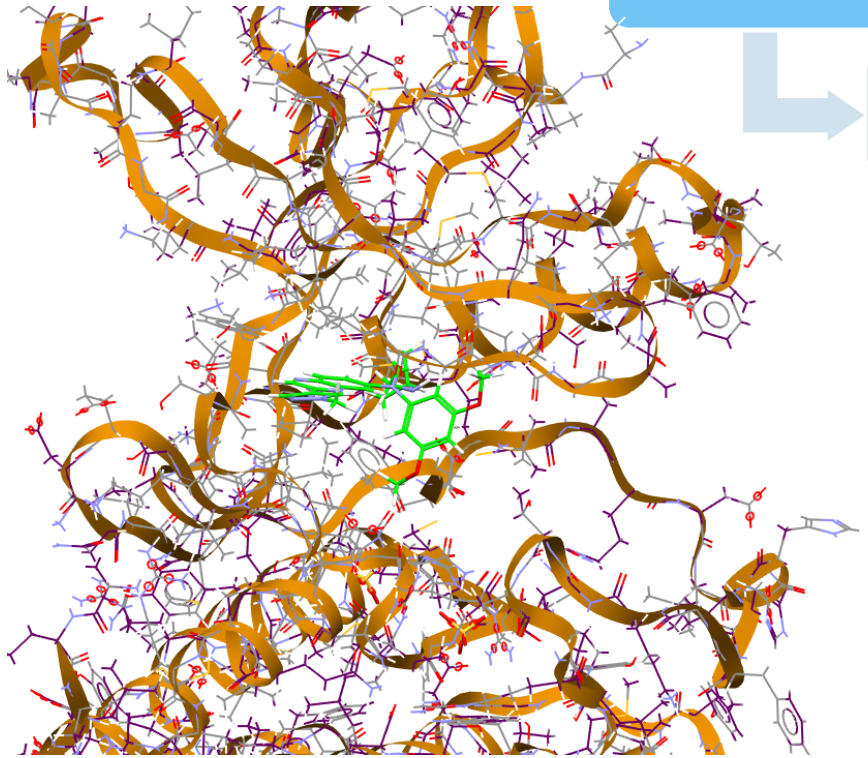
- Target selection
- Protein Preparation
- Binding site definition
- Ligand(s) selection for docking
- Scoring function selection
- Setup and running the docking calculation
- Analysis of the results and visualisation

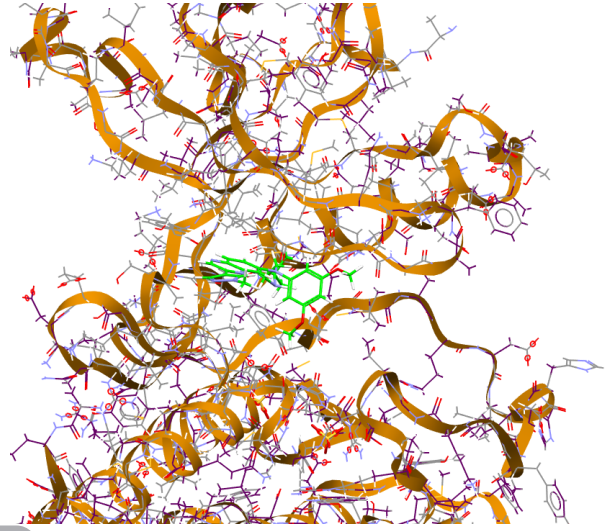


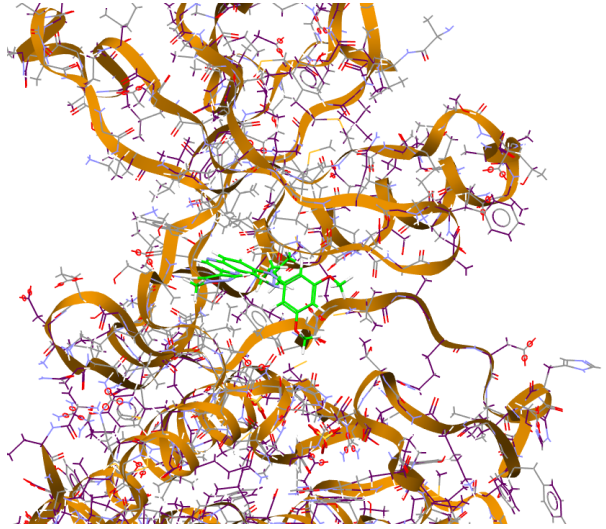
CCDC

Steps in molecular docking

- Target selection
- Protein Preparation
- Binding site definition
- Ligand(s) selection for docking
- Scoring function selection
- Setup and running the docking calculation
- Analysis of the results and visualisation



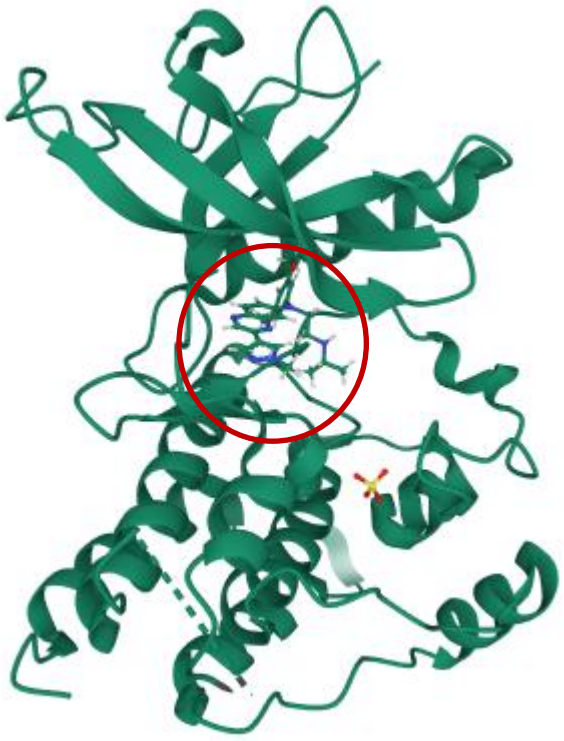


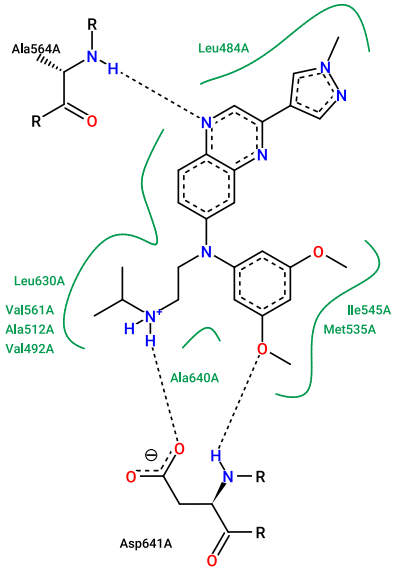


CCDC

Docking with GOLD: Case Study

PDB: 5EW8



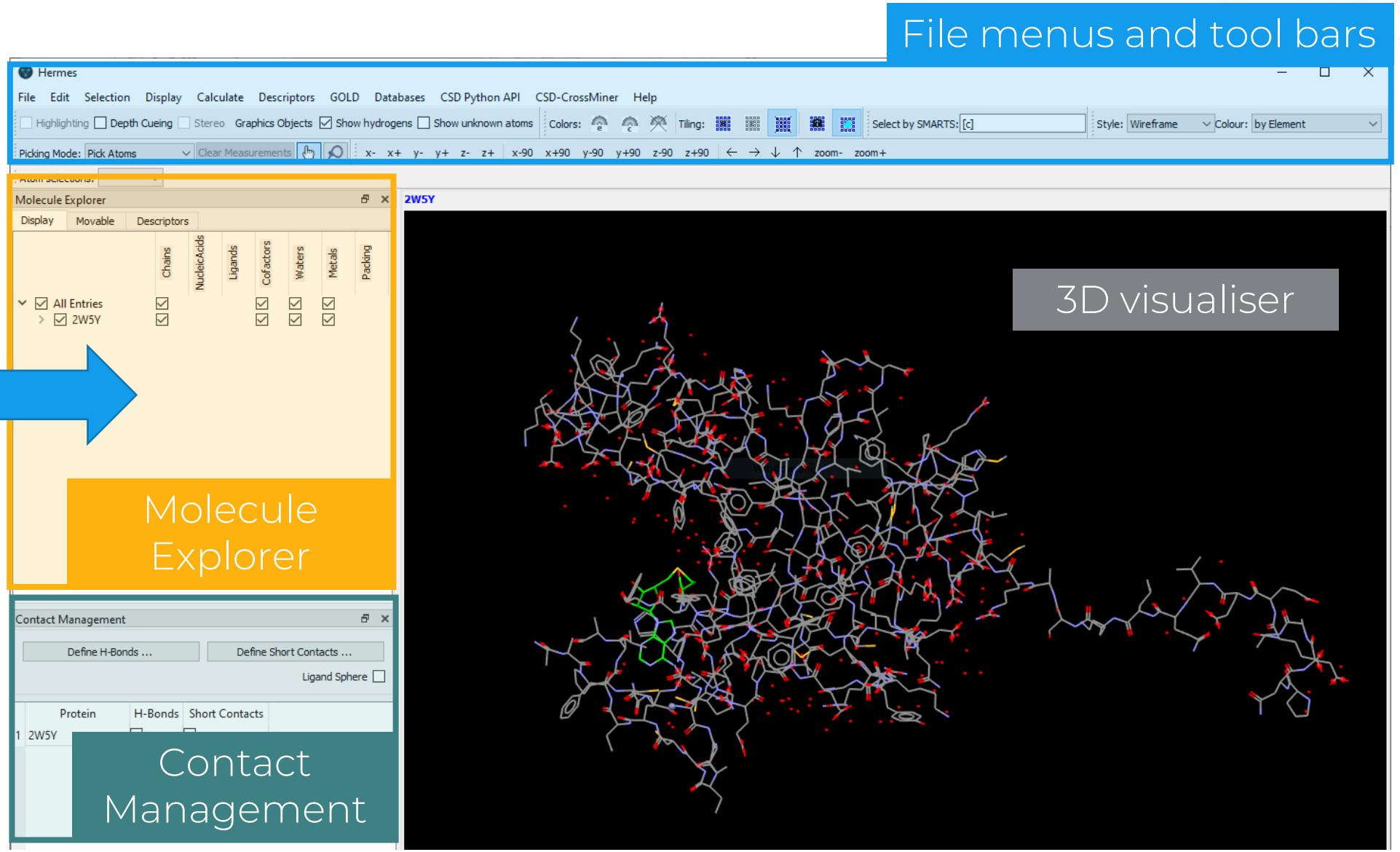


- Fibroblast growth factor receptors (FGFRs) are a family of receptor **tyrosine kinases** expressed on the cell membrane that play crucial roles in both developmental and adult cells.
- Our **ligand** is (3,5-dimethoxyphenyl)-~{N}'-[3-(1-methylpyrazol-4-yl)quinoxalin-6-yl]-~{N}-propan-2-yl-ethane-1,2-diamine), aka Erdafitinib.
- It is the first-ever fibroblast growth factor receptor (FGFR) kinase inhibitor indicated for patients with locally advanced or metastatic urothelial carcinoma.

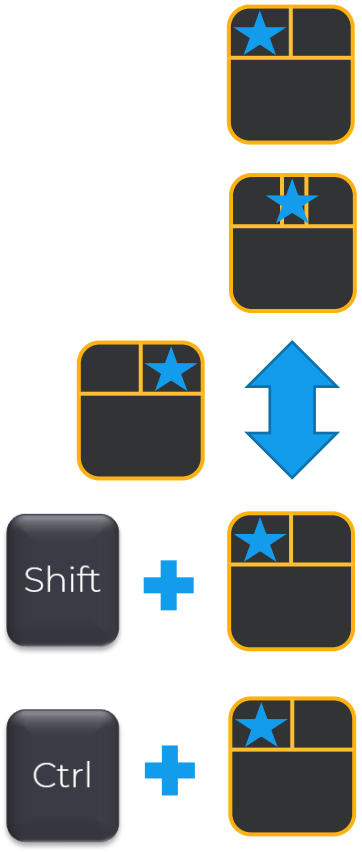
Patani H., et al., Landscape of activating cancer mutations in FGFR kinases and their differential responses to inhibitors in clinical use. Oncotarget. 2016; 7: 24252-24268.

CCDC

Using GOLD through the Hermes Interface



The 3D window basics

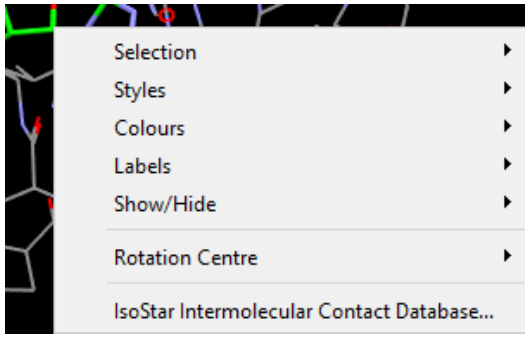


- Left mouse button and move – rotate molecules
- Middle Mouse wheel – move molecules up and down
- Right mouse button and move up and down – zoom in and out of molecules
- Shift + Left mouse button and move - rotate in the plane molecules
- Ctrl + Left mouse button and move - translate molecules

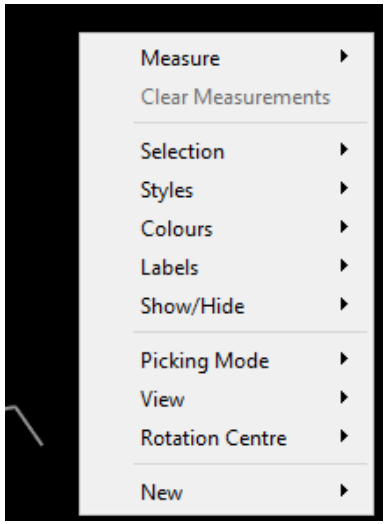
This is for windows and a right-handed mouse – your set up may differ

Opening menus

On a feature:



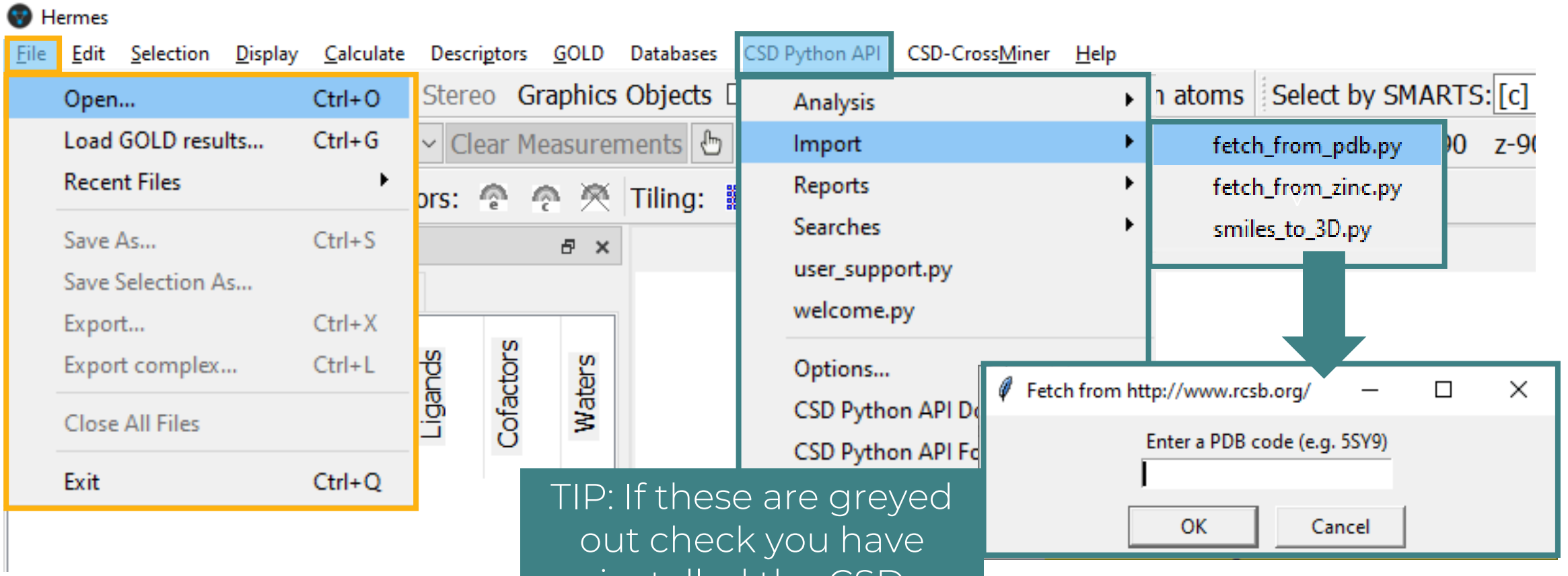
Away from a feature:



Loading structures into Hermes

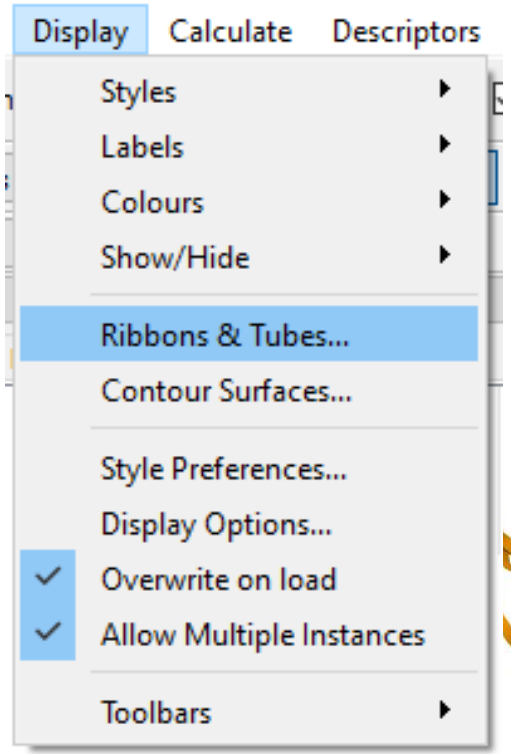
From a file: File > Open

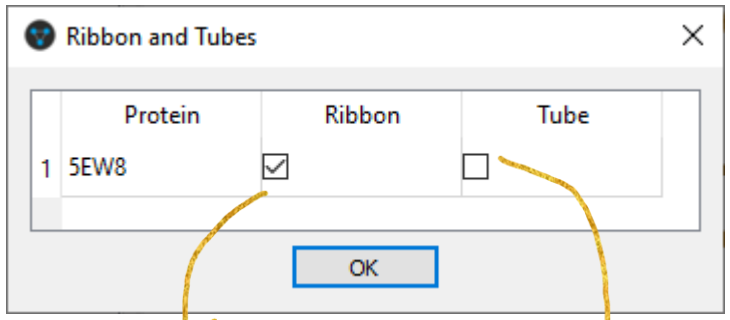
From a DB or SMILES: CSD Python API > Import

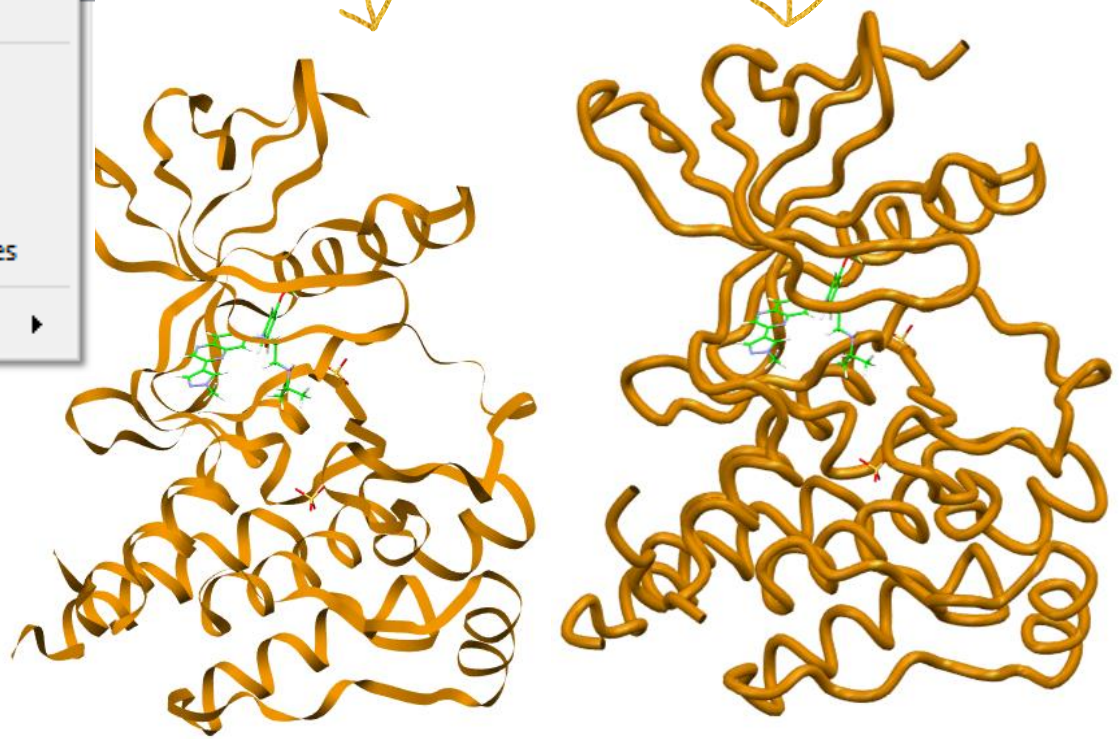


TIP: If these are greyed out check you have installed the CSD Python API

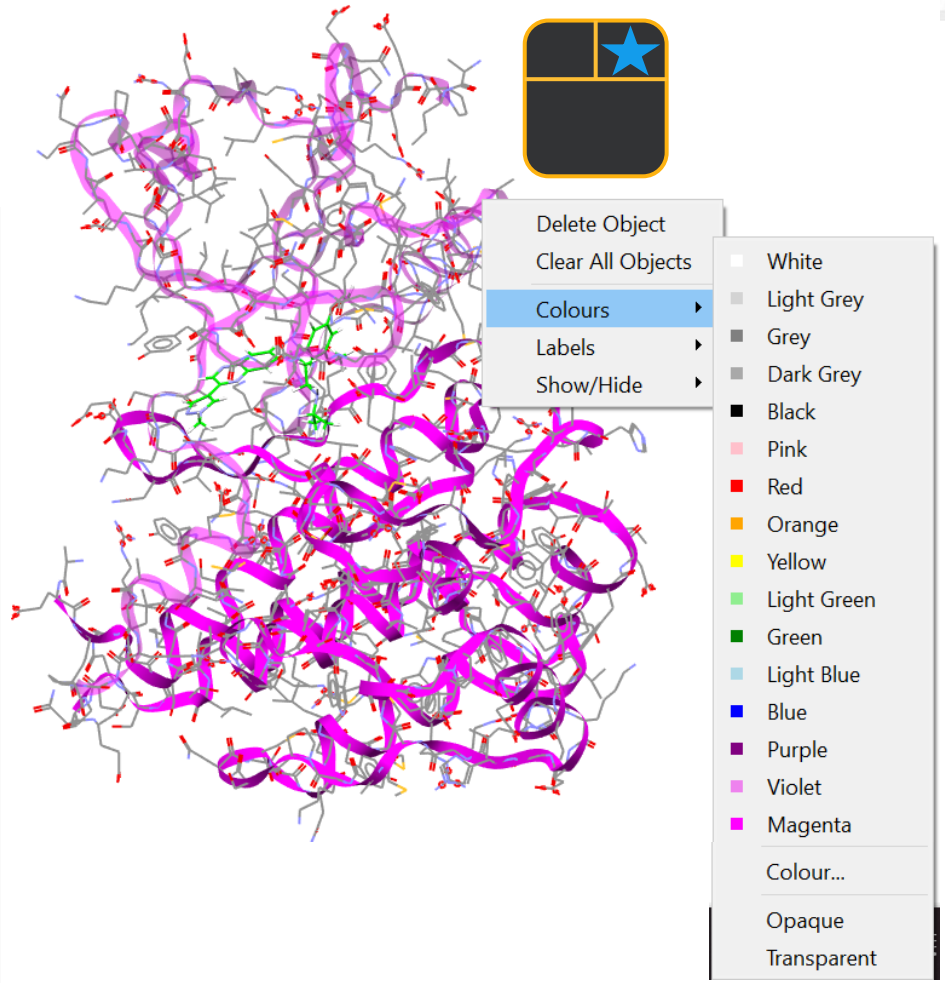
Ribbon & Tubes options



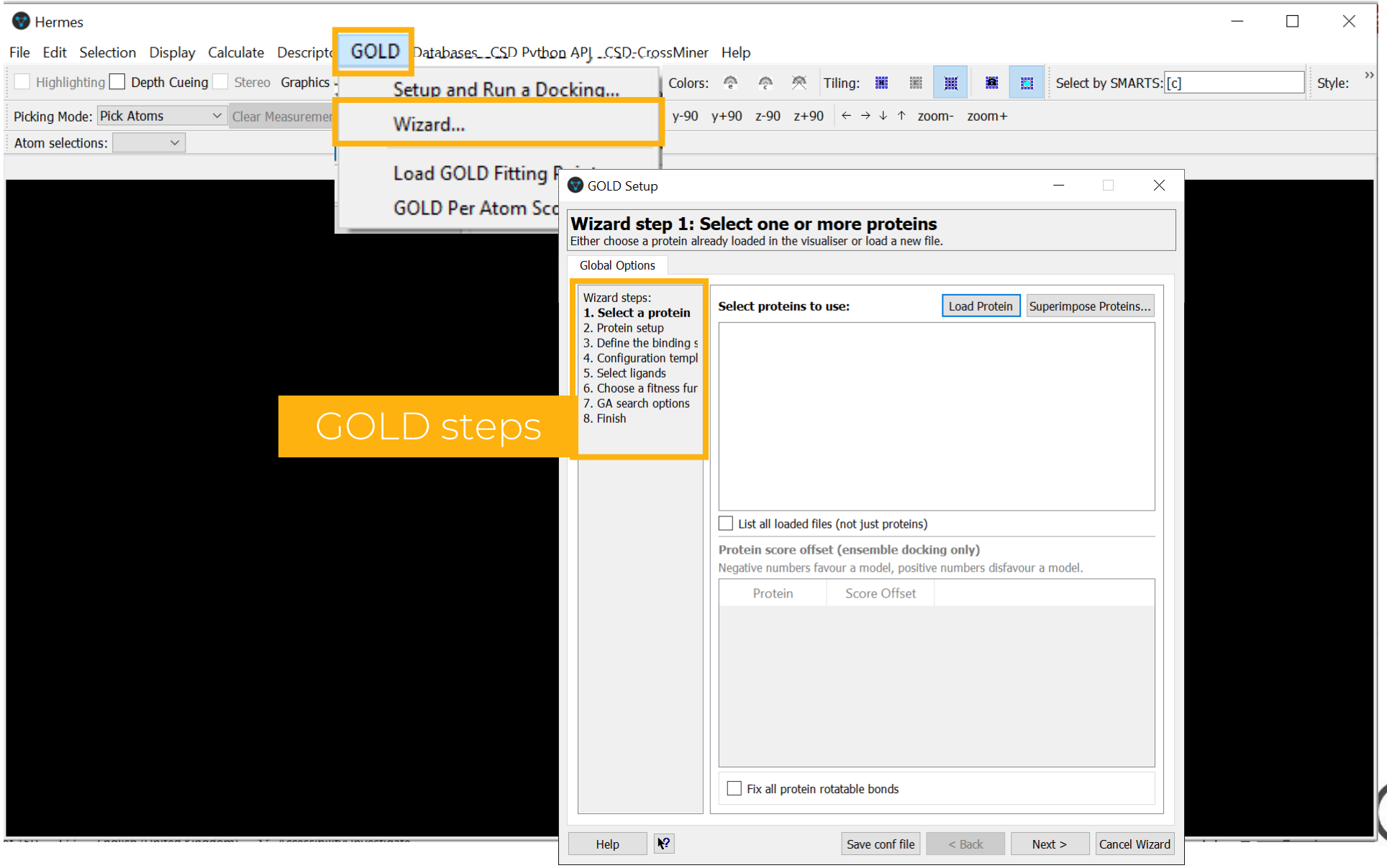




Colours and transparency



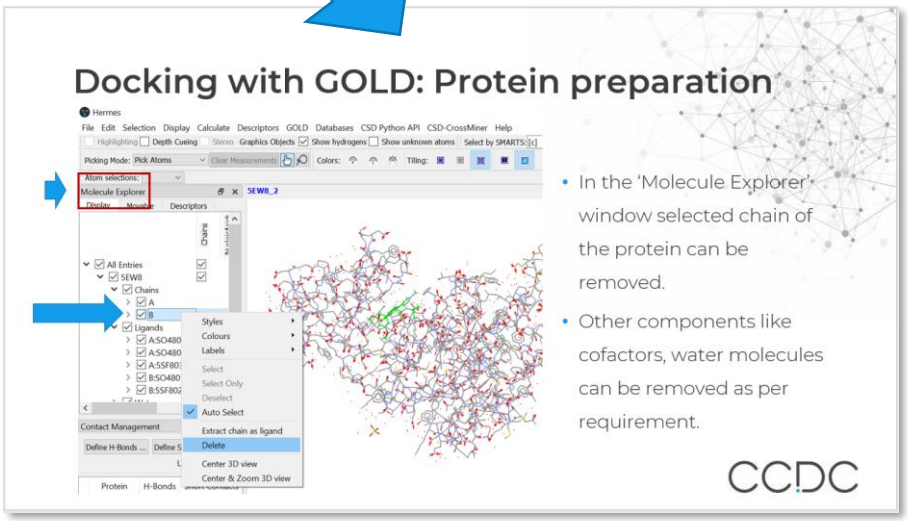
Running GOLD



Show One: Demo of GOLD

We will make the recording available to you in the next few days.

- Get ready to follow the demo and see a GOLD run from fetching the protein to analysing the results.
- In the Notes available on the webpage you will find all the steps and you can refer to these while following along.



CCDC

Try One: hands-on exercise

We will make the recording available to you in the next few days.

It's your turn!

- Try the case study from the handout.
- Your tutors are on hand to help you!
- To ask questions during this time type a message in the chat box.
- If you finish early, ask us for more challenging examples.

https://info.ccdc.cam.ac.uk/2024-autumn-virtual-workshop

Introduction to Protein-ligand docking with GOLD

Developed using 2024.1 CSD Release

Table of Contents

CCDC advancing structural science

Explore More: what can you do with GOLD?

- There is not enough time to explore more advanced functionality today, but we will briefly introduce some extra tips and examples.

CCDC

GOLD: Complete molecular docking suite

Pose prediction

Validate your ligand docking results and optimise hits to leads.

Highly configurable constraints

Use your existing knowledge of the system to bias results and focus on known features and behaviours.

Multiple scoring functions

Score and rescore to build a full picture of your system or perform consensus scoring.

Flexible docking

Perform ensemble docking or handle flexible side-chains with soft potentials.

Water handling

Assess how structural waters affect binding, see if the ligand displaces waters or mediates the interaction during docking.

CCDC

GOLD: Complete molecular docking suite

Virtual screening

Unlimited potential with virtual screening powered by cloud or cluster (HPC).

Python API access

Run dockings programmatically - for parameter optimisation and workflow incorporation.

KNIME component

Perform protein-ligand docking in the KNIME interface to easily build into pipelines of work.

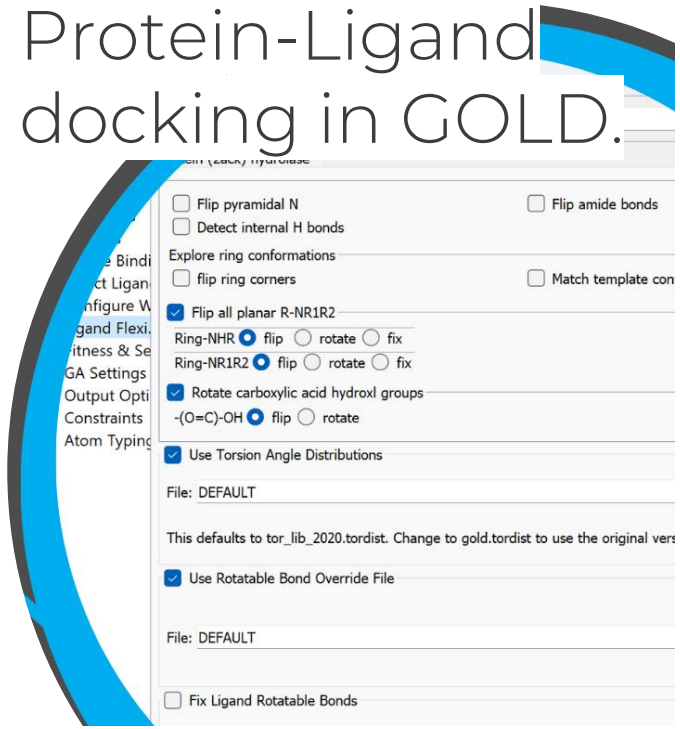
Covalent docking

Understand irreversible binding with covalent docking to explore cancer, immunology and infectious disease targets.

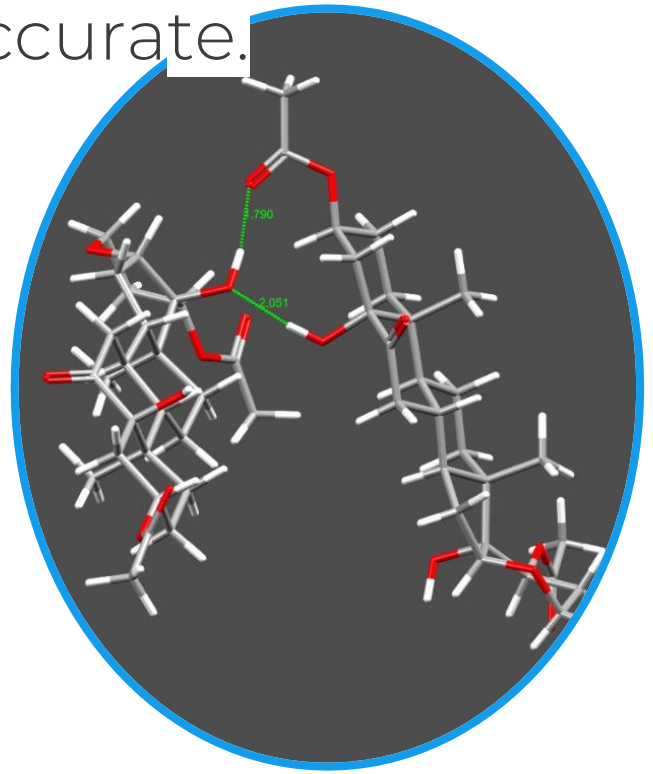
CCDC

What's new in GOLD and Hermes

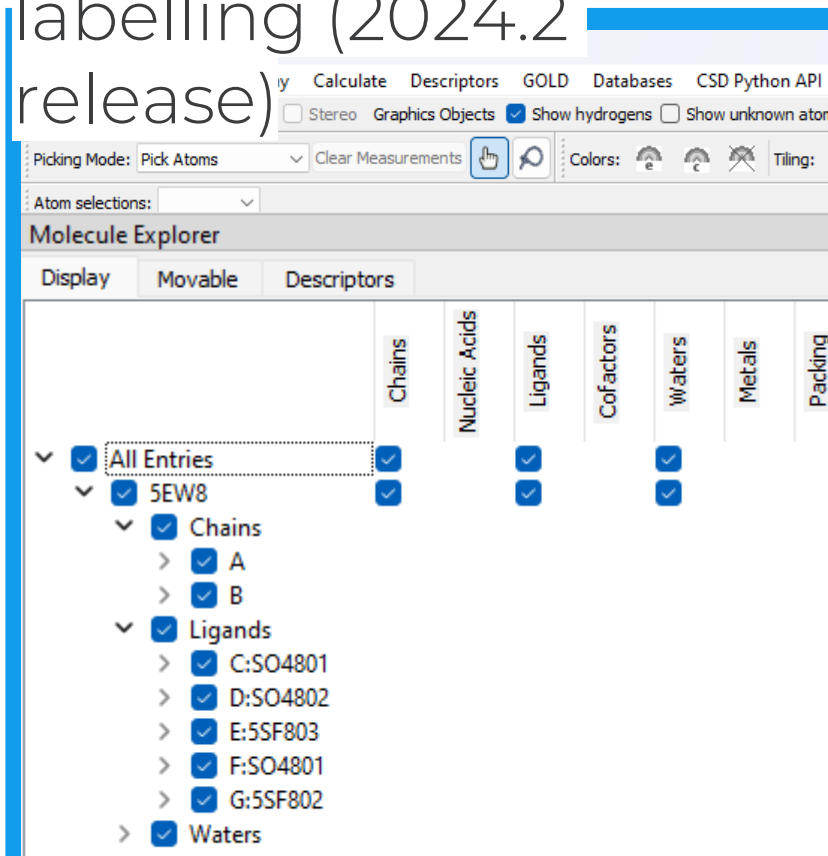
- **Rotamer library** – Improved torsion distribution for Protein-Ligand docking in GOLD.



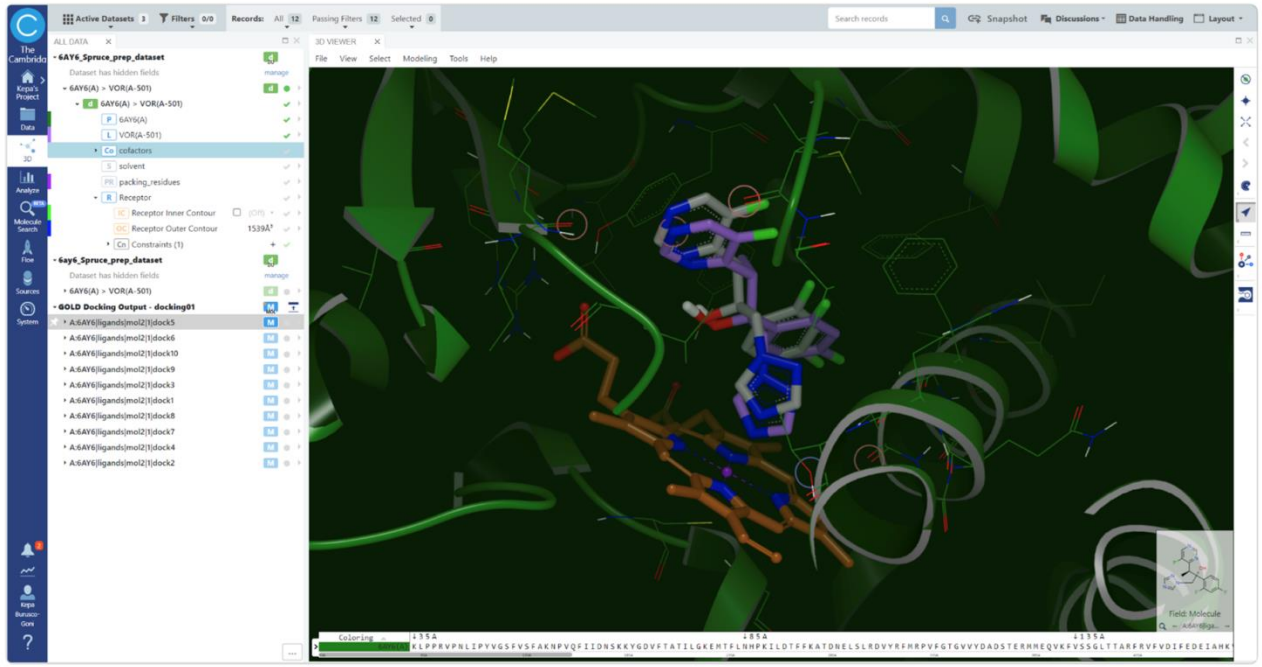
- **Waters** – better placed, not unidirectional, more accurate.



- **Chain labelling** – added progressive labelling (2024.2 release)



GOLD is Now Integrated with Cadence's Orion® Platform



With this integration, computational and medicinal chemists on Orion can:

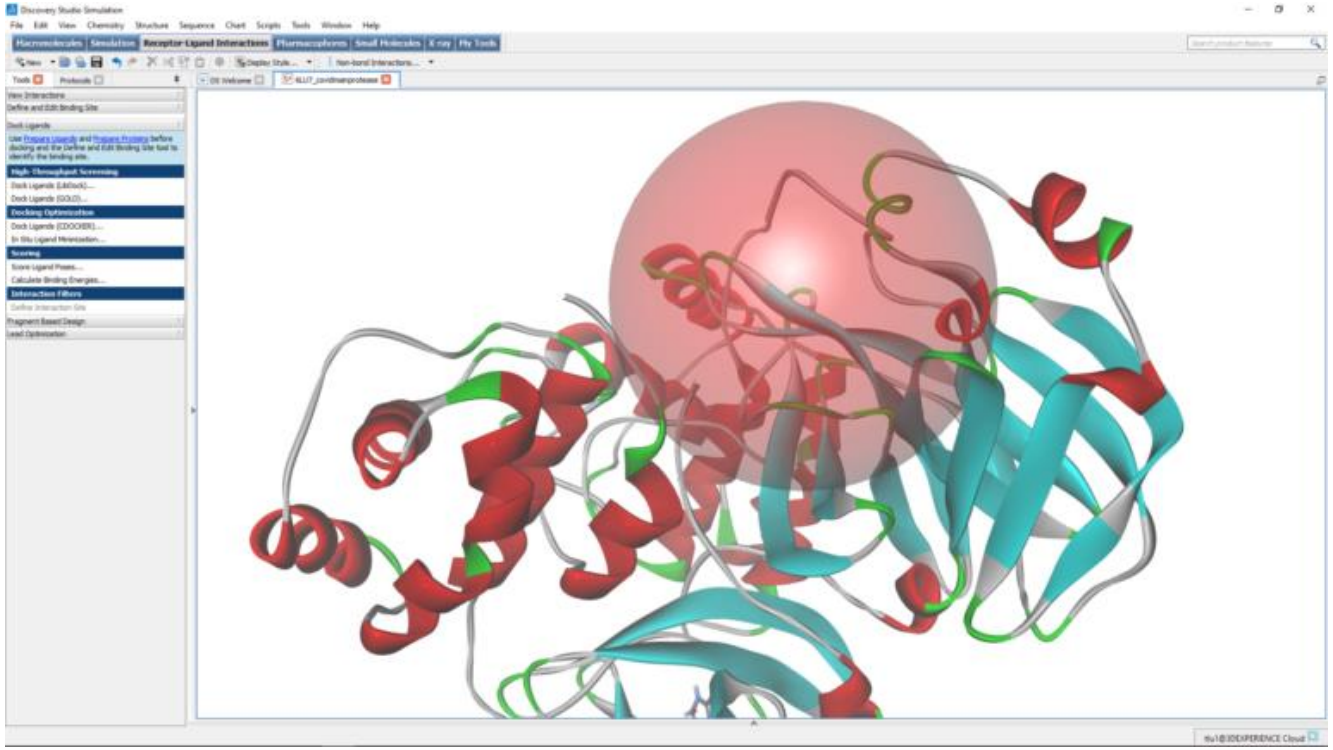
- Run GOLD Docking Workflows Within Orion
- Automate and Scale
- Collaborate and Share
- Access Integrated Data Management

https://www.ccdc.cam.ac.uk/discover/blog/ccdcs-gold-protein-ligand-docking-software-now-integrated-with-cadences-orion-platform/

CCDC

BIOVIA - Docking in the Cloud with GOLD

GOLD algorithm is integrated with BIOVIA cloud-based 3DEXPERIENCE® platform since 2022, combining scientific excellence with cloud flexibility and scalability.



https://www.ccdc.cam.ac.uk/csd-integratable-software/biovia/

CCDC

GOLD in the CSD Python API

Functionality includes:

- Binding site definition: from a reference ligand, protein atom, residue(s) or point
- Docking constraints: protein H-bond, distance, substructure, H-bond, scaffold, similarity, region
- Four fitness functions for score & rescore, GA settings (autoscale)
- Early termination, diverse solutions, reference ligand
- Export protein-ligand complexes with rotated protein atom positions
- Protein & ligand preparation tasks
- Ensemble docking
- Can also supply a gold.conf file for any additional parameters/settings
- Docking post-processing: easy to calculate enrichment metrics (AUC, EF, BEDROC…)

CSD Python API 3.0.14 documentation » Descriptive documentation » Docking and scoring

CCDC

Table of Contents

Docking and scoring
- Introduction
- Setting up a docking run
  - Essential steps
  - Inspecting results
  - Binding Site Definition
    - Defining a binding site from a reference ligand
    - Defining a binding site from a protein atom, residue(s) or point
  - Docking Constraints
    - Setting up protein HBond constraints
    - Setting up distance or HBond constraints
    - Setting up substructure-based distance constraints
    - Setting up scaffold match constraints
    - Setting up similarity constraints
    - Setting up region constraints
    - Setting up pharmacophore constraints
  - Flexible side chains
  - Additional settings
- Interactive docking
- Rescoring
- Protein and Ligand Preparation
- Output File Write Options
- Setting up ensemble docking with waters

Previous topic
Working with cavities

Next topic
Pharmacophore searching

Quick search

Docking and scoring

Note: The `ccdc.docking` module is available only to CSD-Discovery, CSDS+GOLD and CSD-Enterp…

Introduction

Molecular docking is a widely-used computational tool for understanding molecular recognition, which aims to predict the interactions established in a complex formed by two or more constituent molecules with known structures. A key type of molecular docking is protein-ligand docking because of its therapeutic applications in modern structure-based drug design. In this context, the docking process involves the prediction of the binding conformation of small-molecule ligands to the targeted protein binding site. The `ccdc.docking` module provides an API to protein-ligand docking.

The module contains a single class `ccdc.docking.Docker` which, like other classes of the CSD Python API, contains a nested `ccdc.docking.Docker.Settings` class which will be used to specify the desired docking. Once the nested settings class is appropriately configured, `ccdc.docking.Docker.dock()` can be called to perform the docking.

Note: For more information on the docking algorithm please see: "Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation", G. Jones, P. Willett and R. C. Glen, J. Mol. Biol., 245, 43-53, 1995, DOI: 10.1016/S0022-2836(95)80037-9.

See also: API documentation for the docking module

Setting up a docking run

Essential steps

Let us import the appropriate `ccdc.docking` module and …
interest.

```
>>> from ccdc.docking import Docker
>>> from ccdc.io import MoleculeReader, EntryReader
```

A docking requires one or more protein files, one or …
many other optional settings which can be passed to …
them up:

```
>>> docker = Docker()
>>> settings = docker.settings
```

Now get the protein:

```
>>> MLL1_protein_file = '2w5y_protein_prepared.mol2'
```

and load it into the settings:

```
>>> settings.add_protein_file(MLL1_protein_file)
```

Programmatic access with CSD Python API

Free online training module

How are others using GOLD?

GOLD in Action: A Binding Mode Study

Here we highlight a paper by Nunzia Cristiano from the Université Paris Cité, and co-workers from BIOVIA Dassault Systèmes, Novartis Biomedical Research, the University of Montpellier and the University of Regensburg.

In this work, the CCDC's protein-ligand docking software GOLD was used to perform docking studies on selective ligands for a receptor involved in glutamate regulation.

Why?

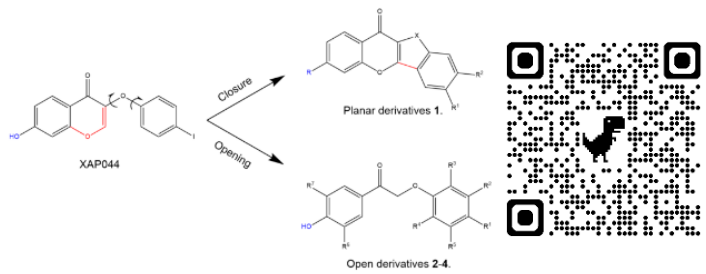
Glutamate is the most abundant excitatory neurotransmitter in the central nervous system (CNS). Its dysregulation has been associated with ischemia, epilepsy, Parkinson's disease, and some psychiatric and mood disorders. Metabotropic glutamate (mGlu) receptors play a key role in modulating the synaptic transmission and are hence optimal drug targets for various CNS disorders.

Despite the hard work of researchers, to date no drugs that target mGlu receptors are on the market. Gee et al. have identified XAP044 as a selective ligand for the mGlu7 presynaptic receptor, and showed that the ligand unexpectedly binds to the extracellular domain of mGlu7 (*J. Biol. Chem.* 2014, *289* (16), 10975–10987). In the work herein presented, the team identified the mode of action of XAP044 by combining various experiments that include synthesizing derivatives, molecular modeling and docking.

How?

The team started by designing and synthesizing several derivatives of XAP044.

As can be seen in **Figure 1**, the ligand XAP044 presents two rotatable bonds that can orient the two aromatic groups either in a planar or in an out-of-plane conformation. To probe the two different orientations, both constrained and open-ring derivatives were prepared. The structure of some of these are shown in **Figure 1**, top right (constrained derivatives, **1**) and in **Figure 1**, bottom right (open-ring derivatives, **2–4**).

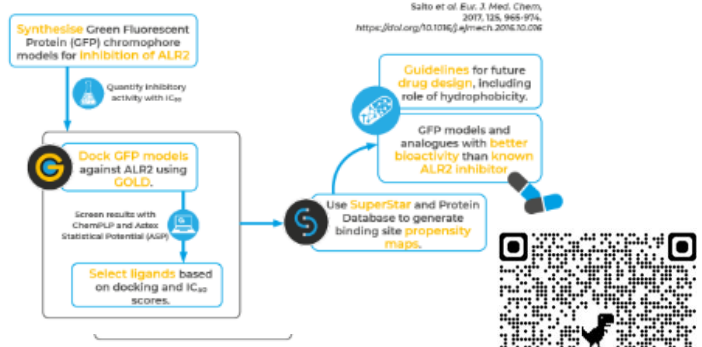


Nunzia Cristiano *et al. J. Med. Chem*. 2024, 67, 14, 11662–11687.

Identifying Aldose Reductase Inhibitors from Green Fluorescent Protein to Treat Diabetes Complications

In this work, authors from Toho University and Kowa Company, Ltd used CSD-Discovery to evaluate the aldose-reductase-inhibiting properties of synthesized Green Fluorescent Protein (GFP) model compounds and their analogues.

Aldose reductase inhibitors are used to treat complications associated with diabetes. The authors confirmed the bioactivity of GFP chromophore models by measuring their IC_{50} values. One of the GFP chromophore models and its analogues exhibited very strong bioactivity, with efficacies higher than that of Sorbinil – a known, highly potent aldose reductase inhibitor. Along with their bioactivity study, the authors performed docking simulations in GOLD and then confirmed docking poses using SuperStar to define the interaction mode of the newly synthesized inhibitors toward the target protein and to identify the molecular features required to gain high inhibitory activity.



Why

Aldose reductase (ALR2) is an NADPH-dependent enzyme that catalyzes D-glucose reduction to D-sorbitol using NADPH as a reductant in the polyol pathway. Most glucose is metabolized via the tricarboxylic acid (TCA) cycle in the glycolytic system. But in patients with hyperglycemia, glucose actively fluxes into the glycolytic system and into the polyol pathway – where ALR2 is activated to consume the flooded glucose. As a result, high levels of D-sorbitol are formed intracellularly. The oxidation of D-sorbitol to D-fructose is very slow, which causes the intracellular concentration of D-sorbitol to increase and a large amount of NADPH is consumed. This results in an osmotic pressure imbalance that can cause diabetes complications, like peripheral neuropathy, cataracts and damage to the kidneys and small blood vessels. The inhibition of ALR2 activity may alleviate or prevent such complications.

Ryota Saito *et al. Eur. J. Med. Chem*, 2017, 125, 965-974.

Investigation of Interactions in Organometallic Compounds Using GOLD

Antimitotic agents are important drugs in anticancer therapy. Owing to their strong side effects and the development of resistance, scientists are searching for more selective antimitotic drugs exhibiting a lower systemic toxicity.

Kinesins are proteins involved in mitosis that are particularly important targets for therapy. New kinesin spindle protein (KSP) inhibitors have been explored and discovered, and ispinesib is one of those. As the most promising results in clinical trials were obtained by combining ispinesib with capecitabine and carboplatin, the scientists started investigating other metal-based drug candidates.

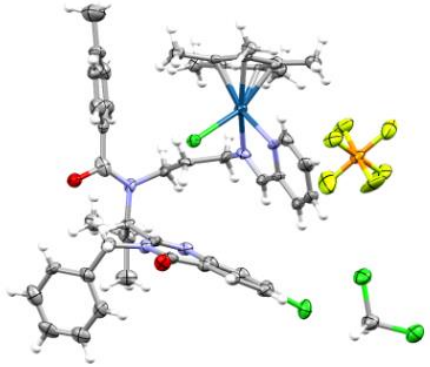
Introducing an organometallic moiety, such as sandwich compounds of Fe and Ru, or half-sandwich compounds of Ru, Os, Rh and Ir, into the structure of an organic pharmacophore often enhances its biological activity. Important advantages can derive from the use of organometallic compounds: they achieve structures that can't be obtained by using organic scaffolds only; the formation of reactive oxygen species (ROS) detrimental to the cell is more likely to occur when in the presence of an organometallic moiety.

Proceeding with their work on Rh and Ir half-sandwich complexes that exhibited high KSP inhibitory activity, the group investigated the effect of combining ispinesib and half-sandwich complexes into the same molecule on the biological activity.

Results and Discussion

The synthesis of the organometallic complexes was performed starting from the (*R*)- and (*S*)-enantiomers of ispinesib. Those precursors were reacted with 2-pyridinecarboxaldehyde in ethanol to form the (*R*)- or (*S*)-imines, and then with the metal dimers to form the half-sandwich conjugates of ispinesib. The metal dimers used were $[(cym)MCl_2]_2$ (M = Ru, Os; cym = η^6–*p*-cymene), $[(1,3,5\text{-}^iPr_3C_6H_3)RuCl_2]_2$ or $[(Cp^*)MCl_2]_2$ (M = Rh, Ir; Cp^* = η^5-pentamethylcyclopentadienyl).

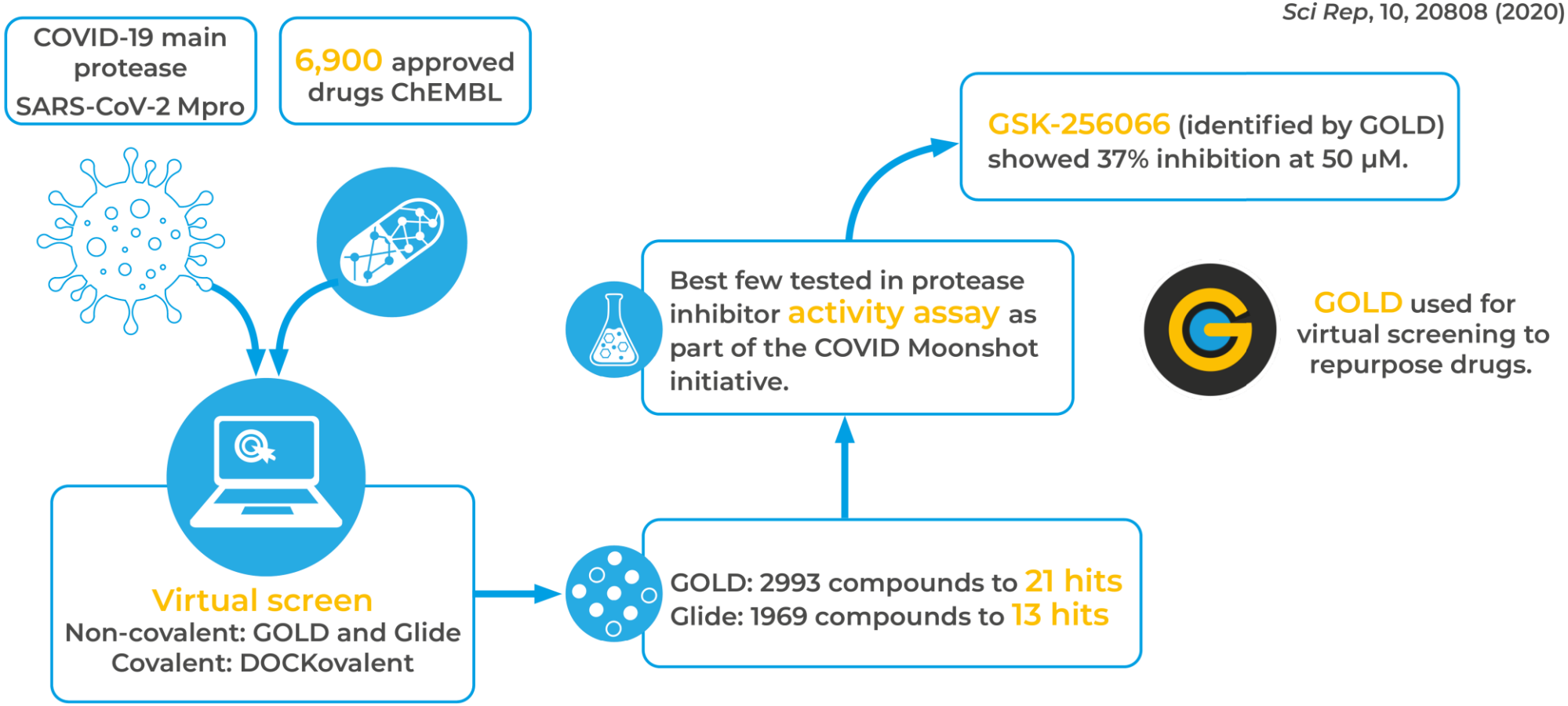
The characterization of the products was performed with NMR spectroscopy, elemental analysis, and electrospray ionization-mass spectrometry (ESI-MS). For the Ir-based complex, the crystal structure was also obtained via single crystal X-ray diffraction (Figure 1).



Damian Plażuk *et al. Dalton Trans.*, 2023,52, 11859-11874.

Virtual screening to repurpose drugs for COVID-19

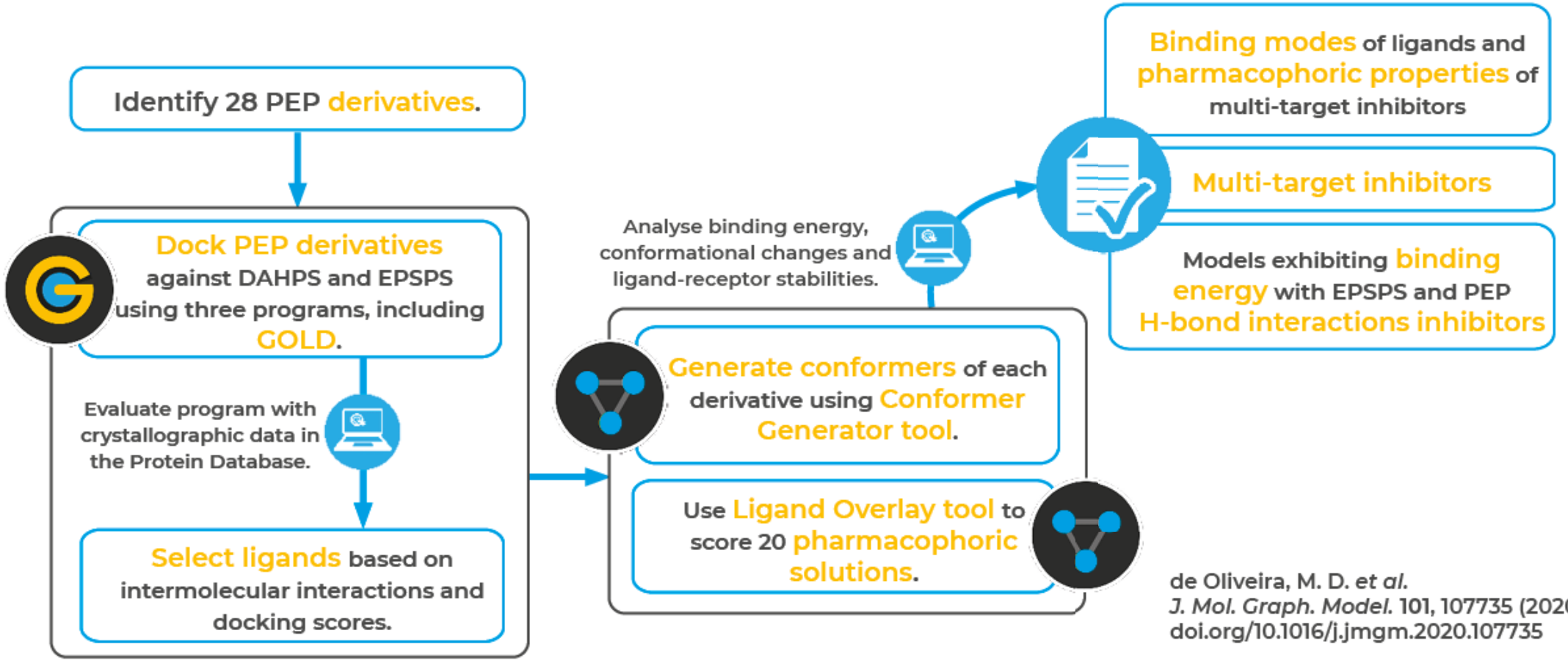
Alina Shitrit *et al.*
Sci Rep, 10, 20808 (2020)



https://www.ccdc.cam.ac.uk/Community/blog/gold-virtual-screen-covid-drug-repurposing/

CCDC

Ligand-based drug design of competitive inhibitors against DAHPS and EPSPS



de Oliveira, M. D. *et al.*
J. Mol. Graph. Model. 101, 107735 (2020).
doi.org/10.1016/j.jmgm.2020.107735

https://www.ccdc.cam.ac.uk/Community/blog/ligand-based-drug-design-of-competitive-inhibitors/

CCDC

What have we learnt?

- The basics of GOLD and the Hermes interface.
- Step-by-step preparation of protein, and ligand(s) in GOLD.
- How to set-up and run a standard protein ligand dock.
  - Configuration options.
  - Run GOLD in the background or interactive mode.
- How to identify the correct binding modes reliably and with confidence.
- The basics of how GOLD can be used in virtual screening and lead optimisation.
  - Example case studies.

CCDC

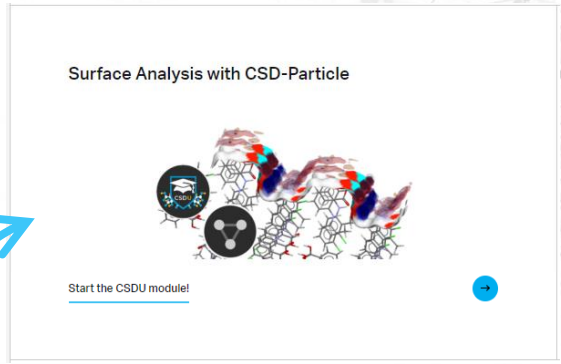
Want to explore more?

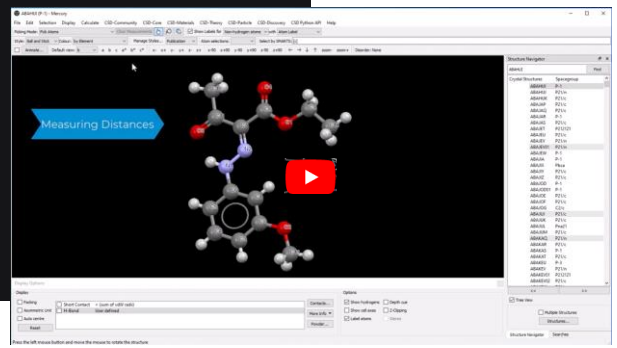
On-demand training resources

Solutions | Community | Discover | Consultancy

- Access/Deposit Structures
- CCDC for the Community
- Education and Outreach
- Events
- Free Products
- Training and Learning

- Free Online Courses
- Self-Guided Workshops
- Training and Support Videos





CCDC

Free online training courses

With completion certificates!

CSDU

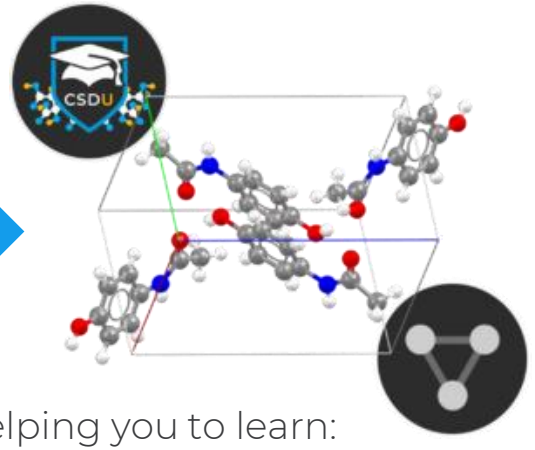
On-demand modules to learn how to use the CSD software at your own pace.

UWatch · UTry · UTest

Mercury · Python API · WebCSD · GOLD Docking

Mogul · FIMs · Particle

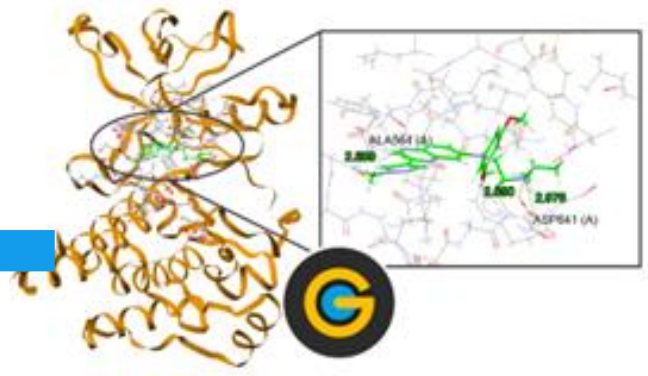
Visualization 101



Helping you to learn:

- The basics of Mercury software.
- How to explore and pack structures.
- How to create high resolution images.

GOLD Docking

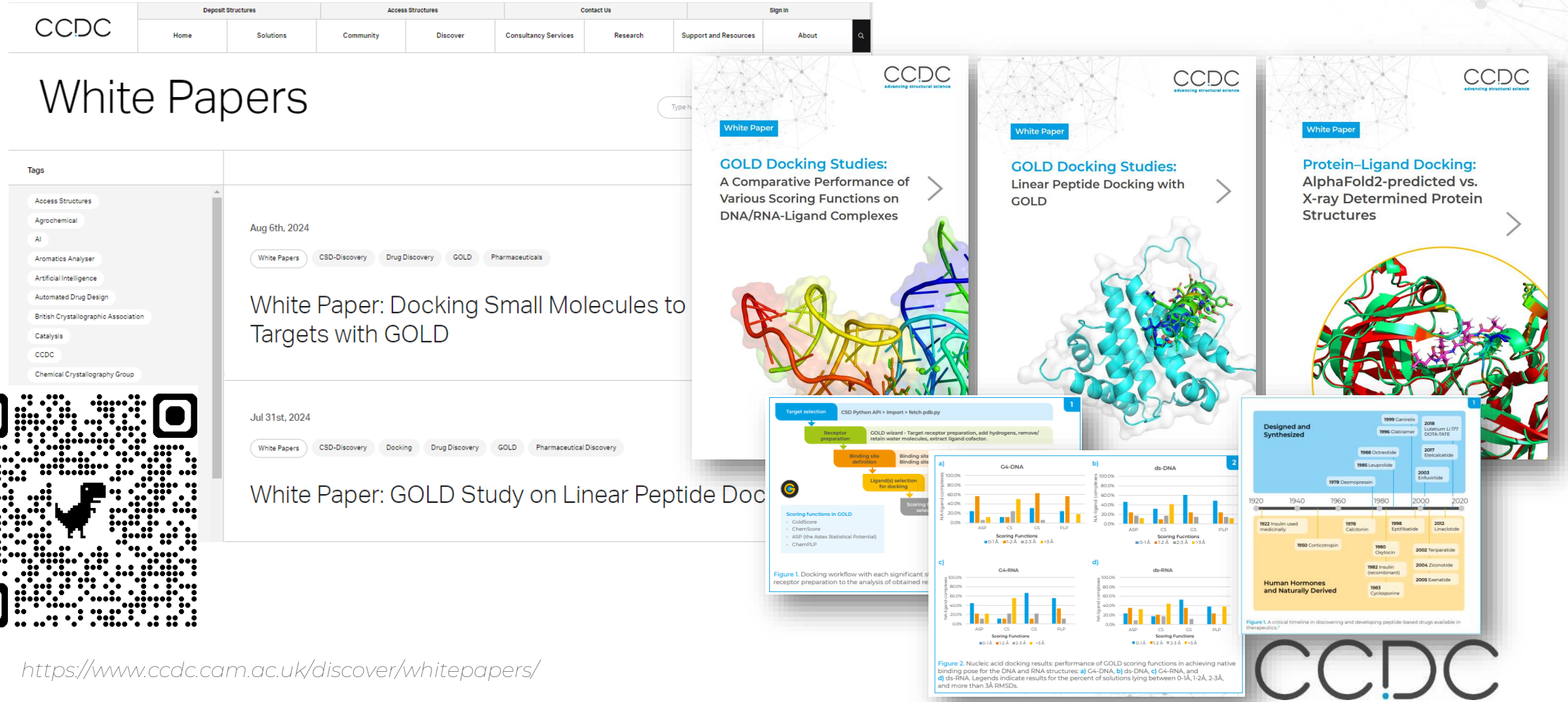


Helping you to learn:

- The basics of Hermes.
- The basics of protein-ligand docking.
- Step-by-step guide on how to use GOLD to perform protein ligand docking.
- Where to get started with your docking simulation.

https://www.ccdc.cam.ac.uk/community/training-and-learning/csdu-modules/

A collection of white papers



https://www.ccdc.cam.ac.uk/discover/whitepapers/

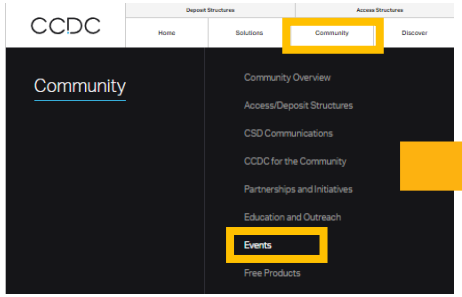
More learning events

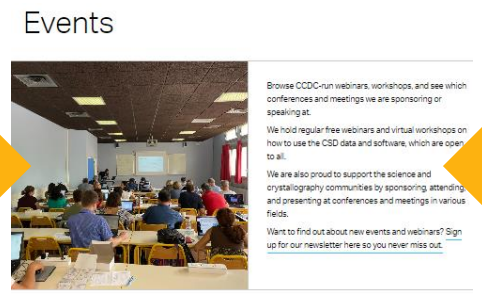
CCDC Virtual Workshops

- **5th Nov CSD-CrossMiner**: Introducing interactive pharmacophore searching across the CSD and the PDB.

CCDC Webinars

- **Coming soon!** The CCDC also regularly host webinars. Check out our website, social media or sign-up to our newsletter to stay up to date.





https://www.ccdc.cam.ac.uk/community/events/

CCDC

FREE VIRTUAL WORKSHOPS

WHAT?

The CCDC Virtual Workshops are a series of hands-on, guided training sessions, where you learn how to use different components of CSD software. These sessions are free and open to both beginners and more experienced users of the CSD Software.

HOW?

The format is 90 minutes and *Show One, Try One, Explore More*:

- *Show One*: A guided demo of the software by the CCDC tutors.
- *Try One*: Hands-on examples for you to try with CCDC tutors on hand to help.
- *Explore More*: Learning outcomes recap, challenges, and quizzes.

OCTOBER-NOVEMBER 2024

8th Oct - 4 pm (BST)

First Steps in Protein-Ligand Docking With GOLD

22nd Oct - 1 pm (BST)

ConQuest to Mercury – From Searching to Data Analysis

5th Nov - 10:30 am (GMT)

Introduction to Pharmacophore Searching Using CSD-CrossMiner

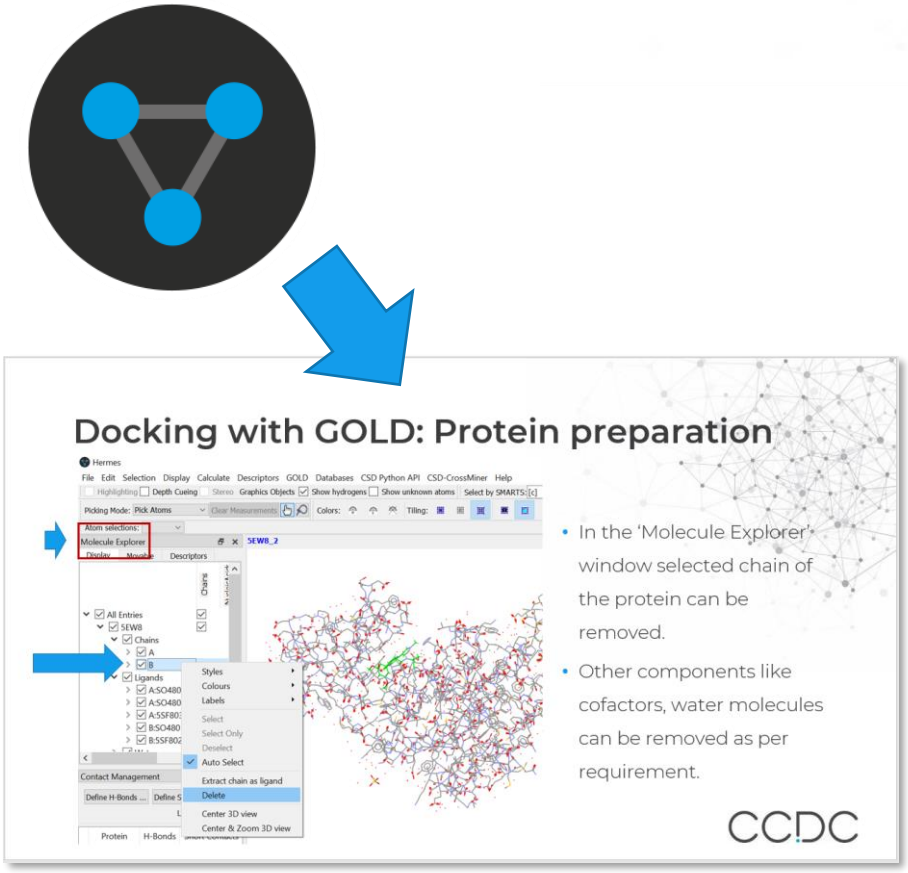
Registration is now open - just scan the QR code

SCAN ME

hello@ccdc.cam.ac.uk

Show One: demo of GOLD

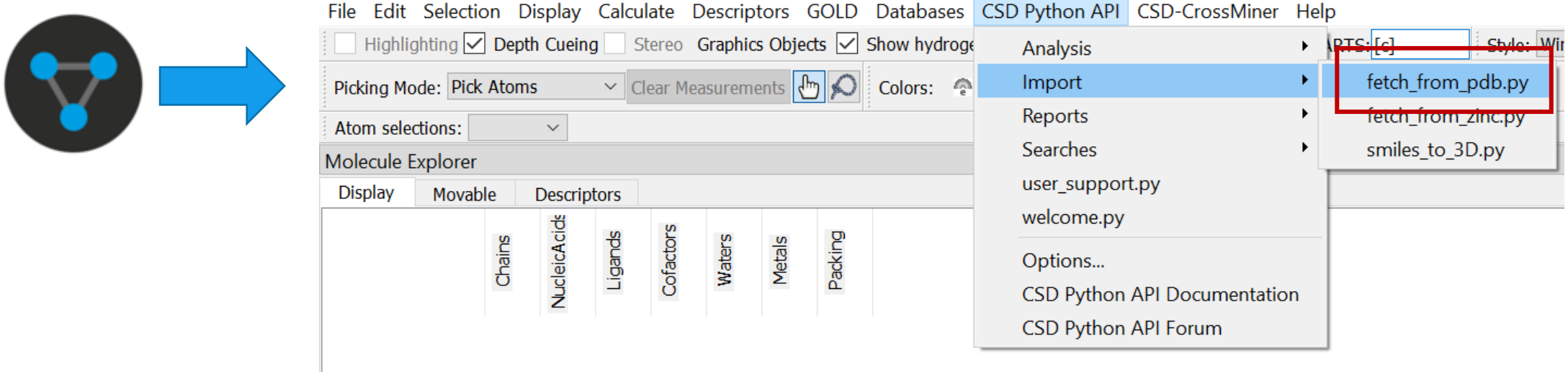
- Get ready to follow the demo and see a GOLD run from fetching the protein to analysing the results.
- In the Notes available on the webpage you will find all the steps and you can refer to these while following along.



CCDC

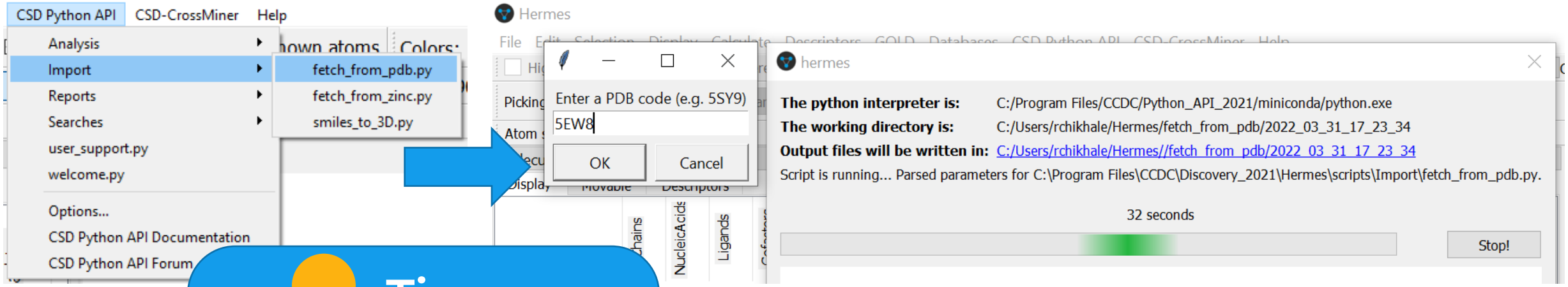
Docking with GOLD: Importing Protein

- Open the Hermes Interface and import the protein crystal structure from the PDB using the ‘fetch_from_pdb.py’ function.



Docking with GOLD: Importing Protein

- Provide with a PDB code in the '**fetch_from_pdb.py**' function search dialogue box.

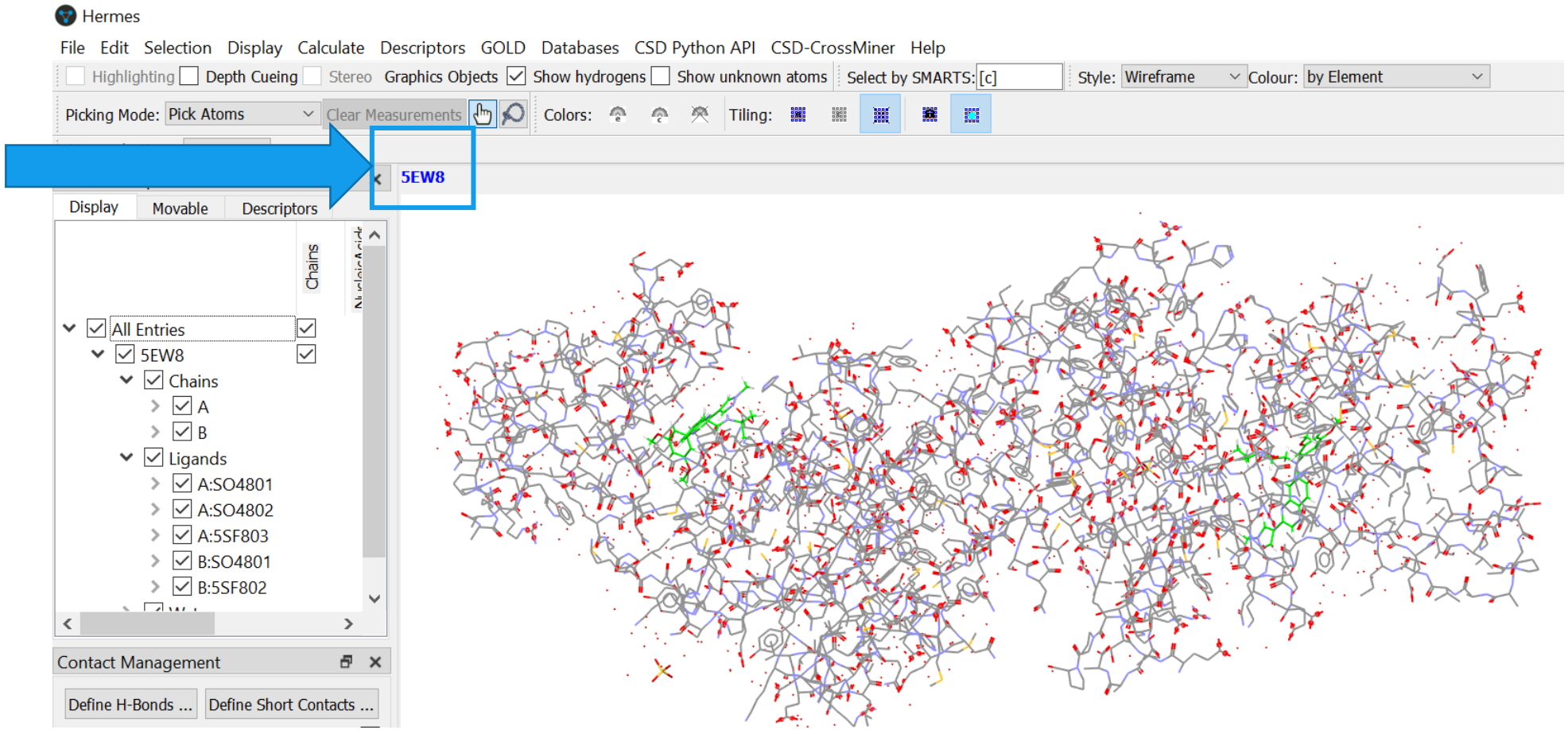


Tip

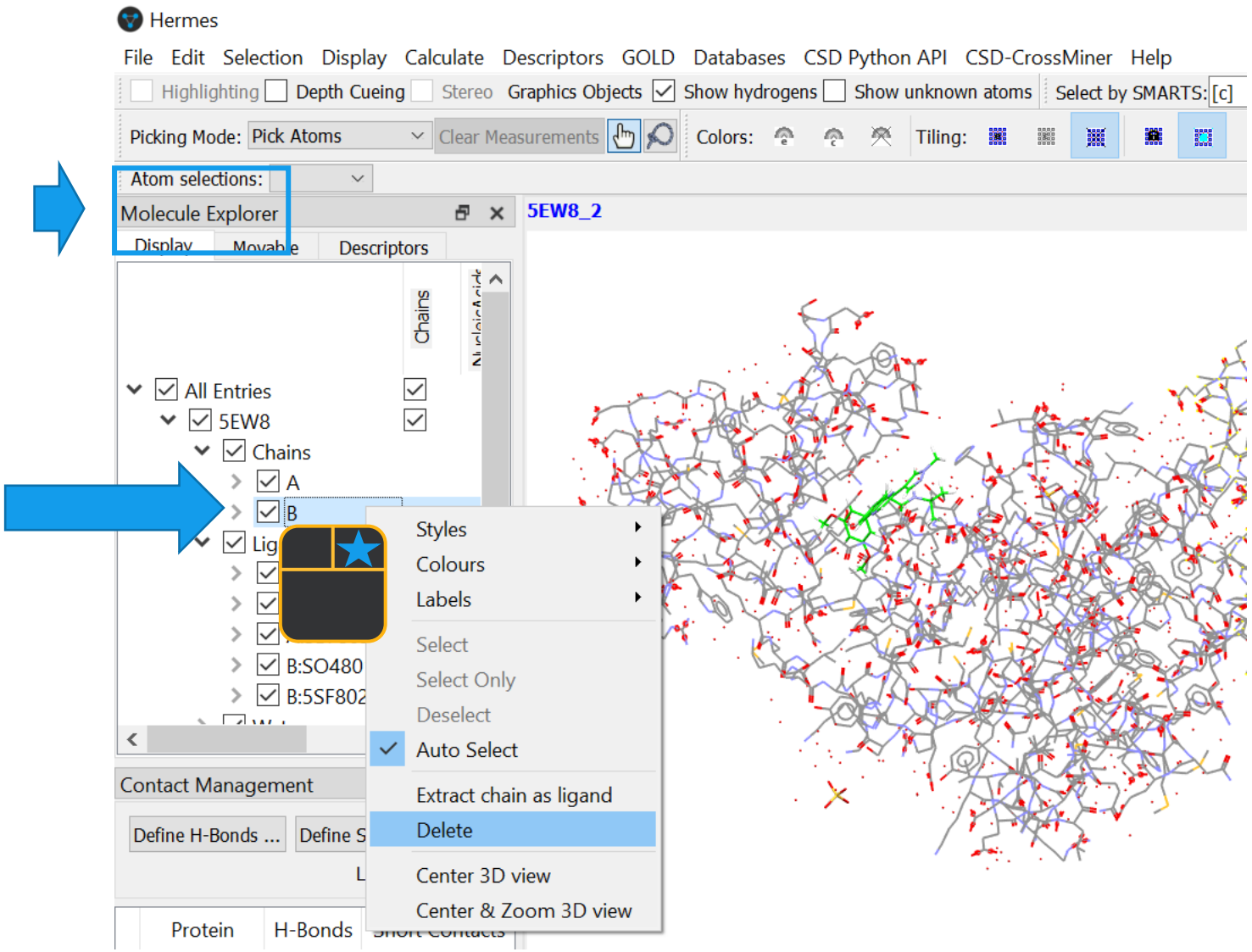
Be aware that the window for entering the PDB code might pop up in a corner of your screen.

Docking with GOLD: Importing Protein

- Once imported, the crystal structure will look like this in Hermes.

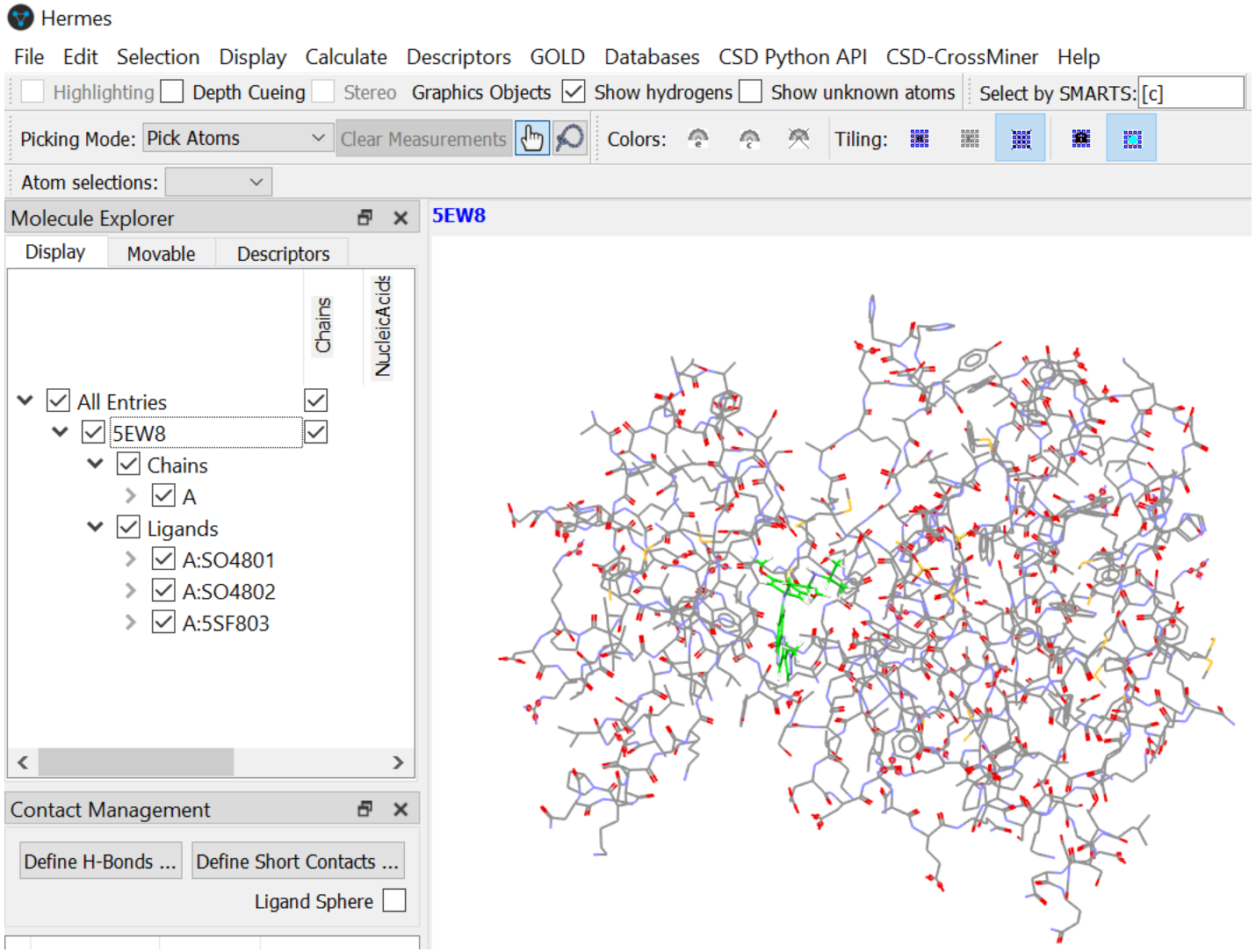


Docking with GOLD: Protein preparation



- In the *Molecule Explorer* window selected chain of the protein can be removed.
- Other components like cofactors, water molecules can be removed as per requirement.

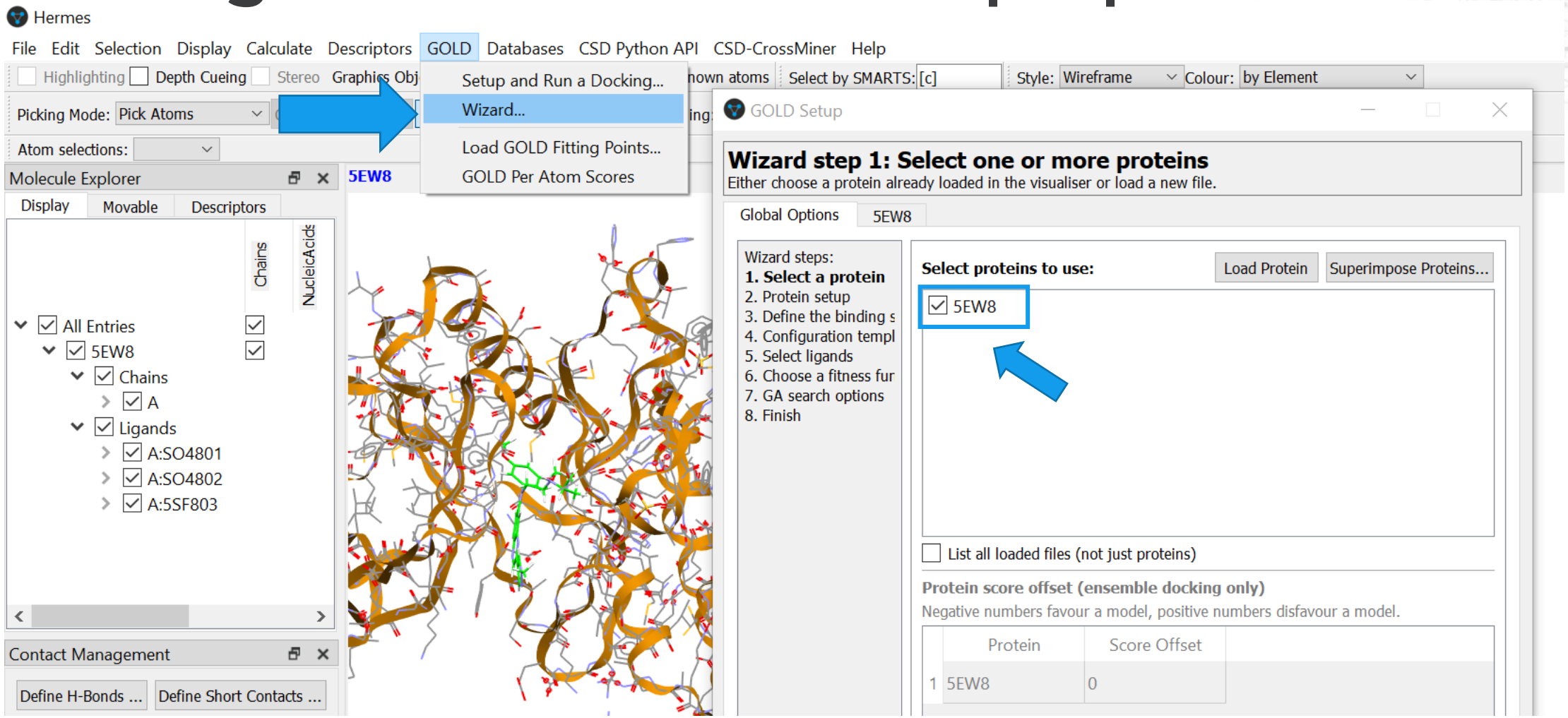
Docking with GOLD: Protein preparation



- Now we are ready for the next stage.

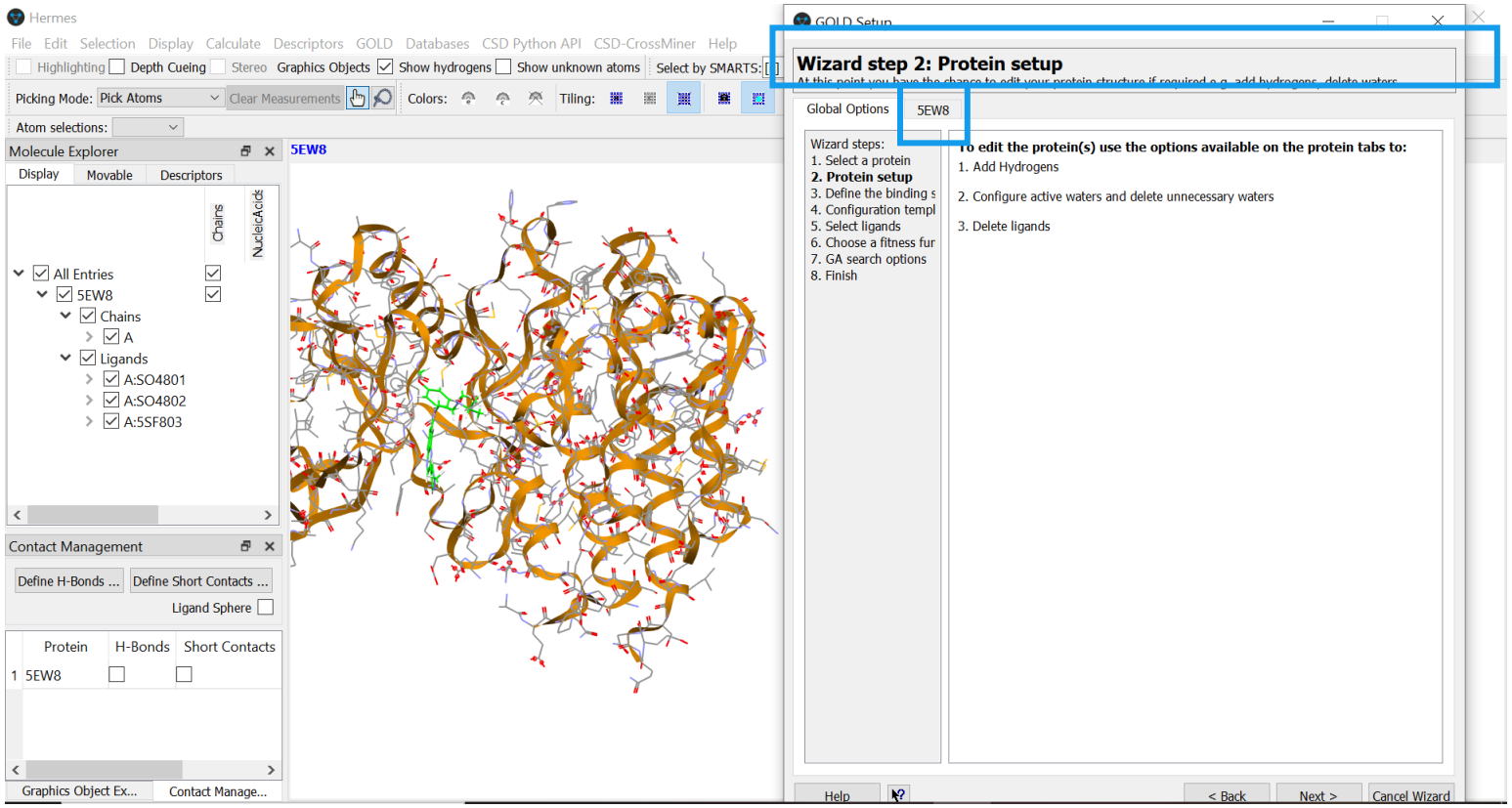
CCDC

Docking with GOLD: Protein preparation

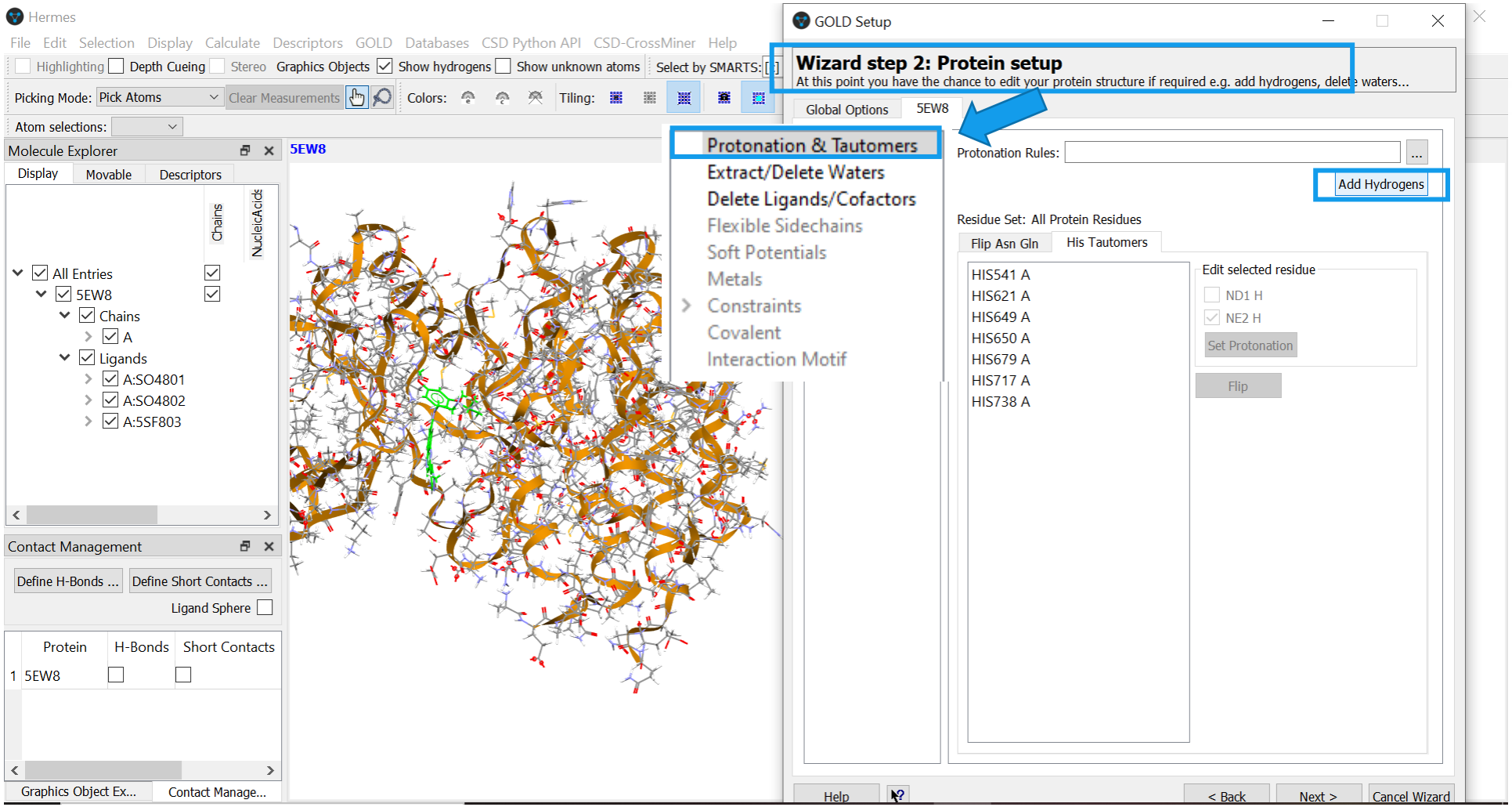


- Launch the GOLD wizard as shown above.
- Then, select the protein to investigate.

Docking with GOLD: Protein preparation



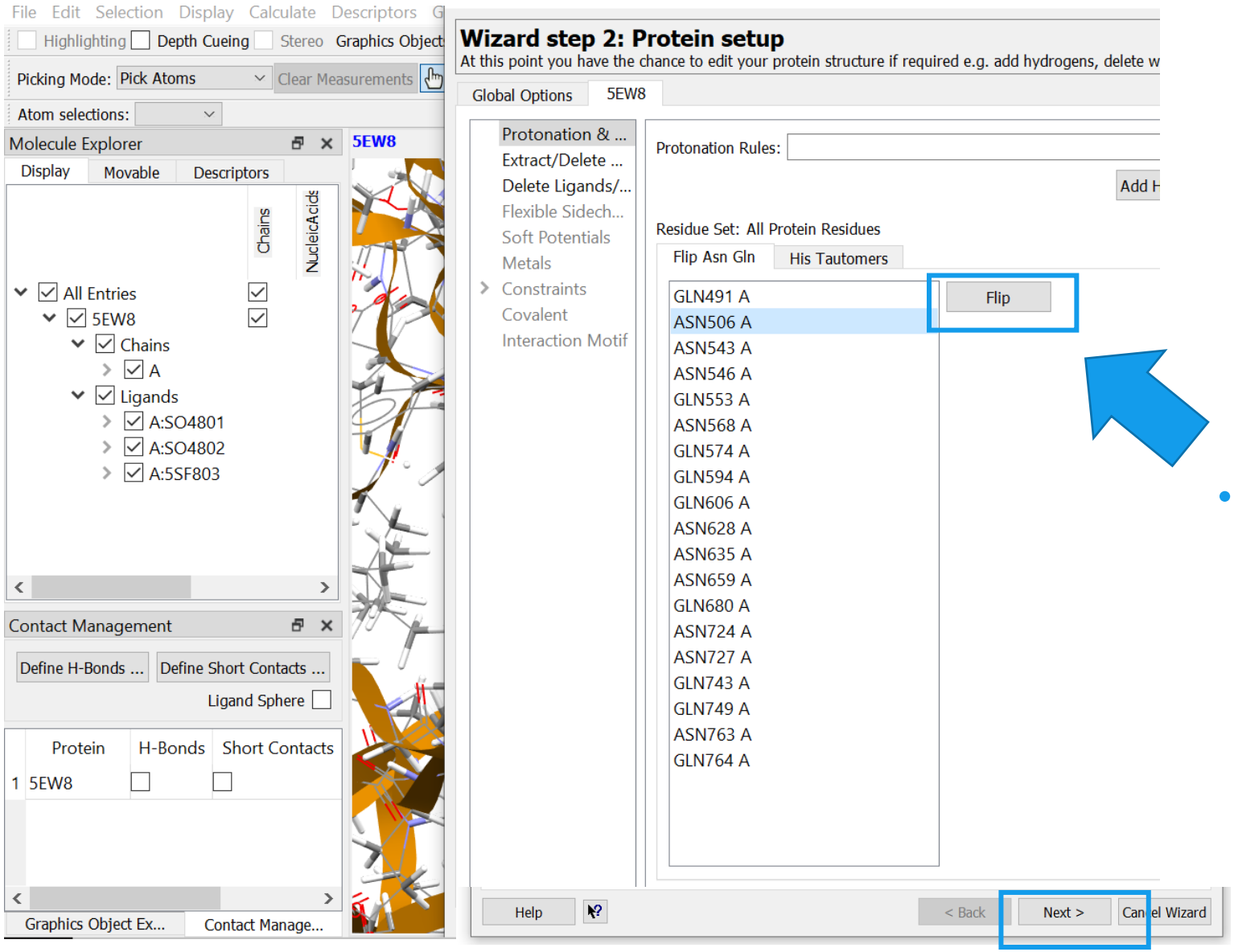
Docking with GOLD: Protein preparation



- Add missing hydrogens.

CCDC

Docking with GOLD: Protein preparation



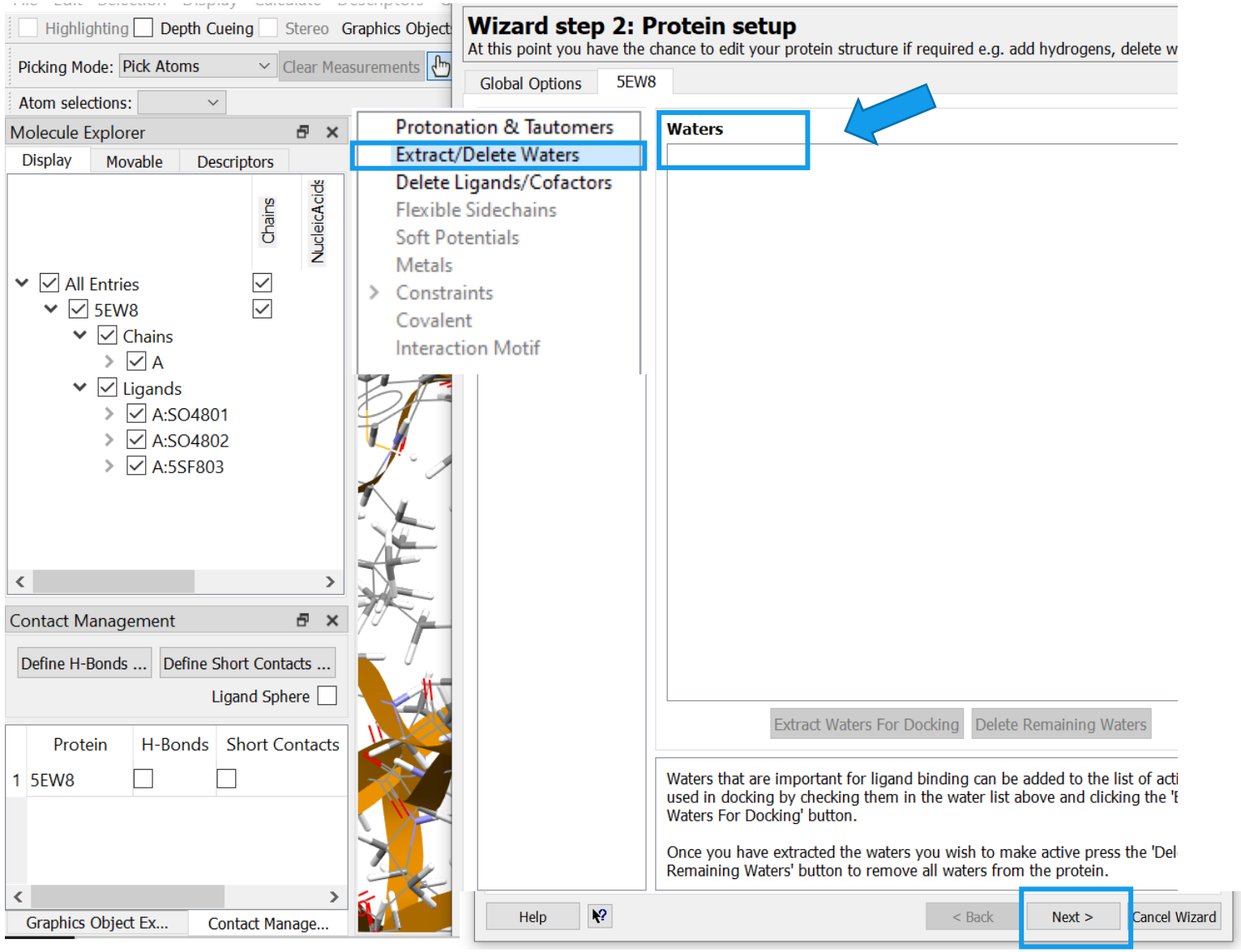
Tip

Always check the protonation around the binding site.

- You can flip the GLN and ASN based on the observation of these residues.

CCDC

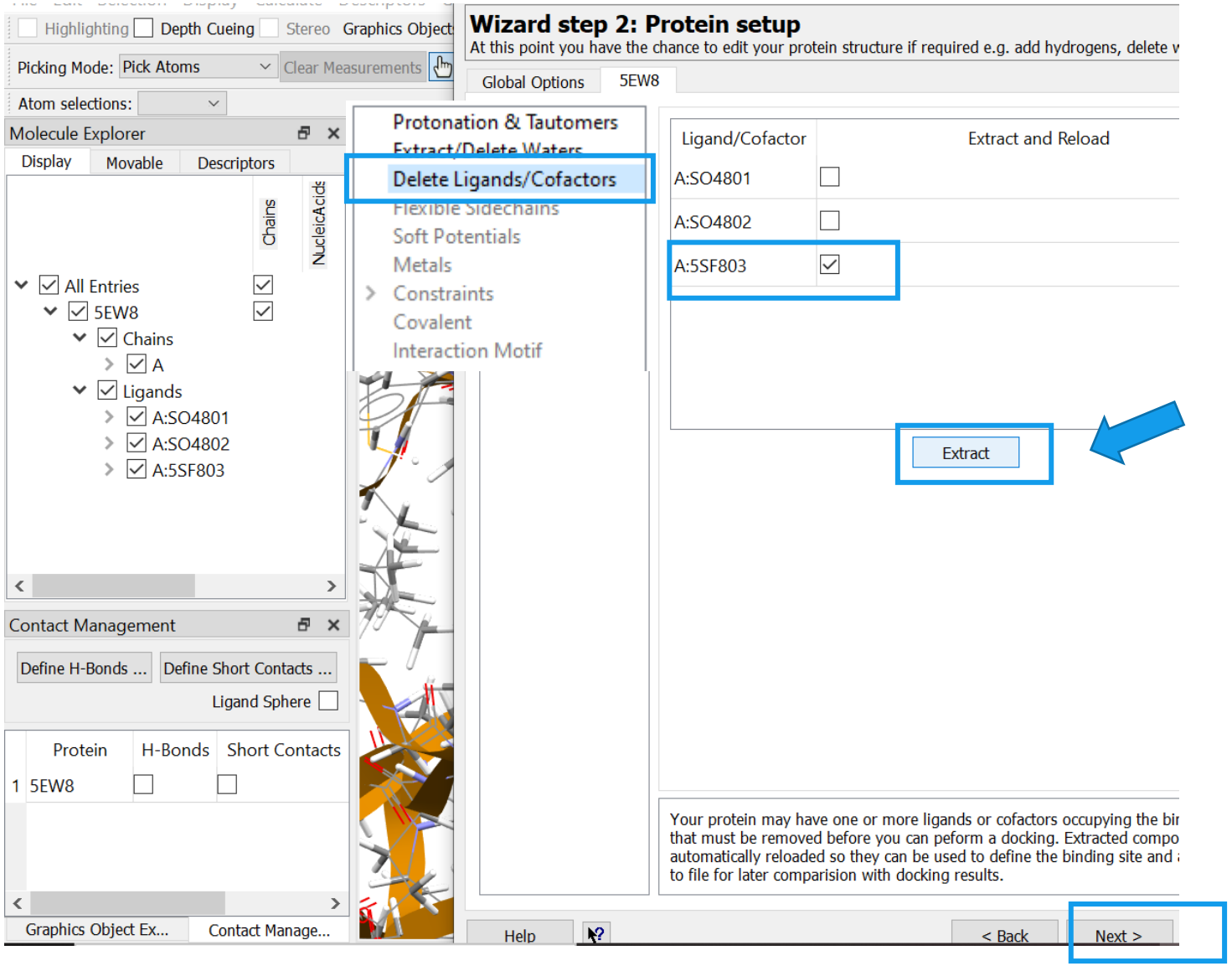
Docking with GOLD: Protein preparation



- You can extract and delete water molecules.
- In the example shown there are no water molecules to remove as we deleted them after loading the protein (before opening the GOLD Wizard).

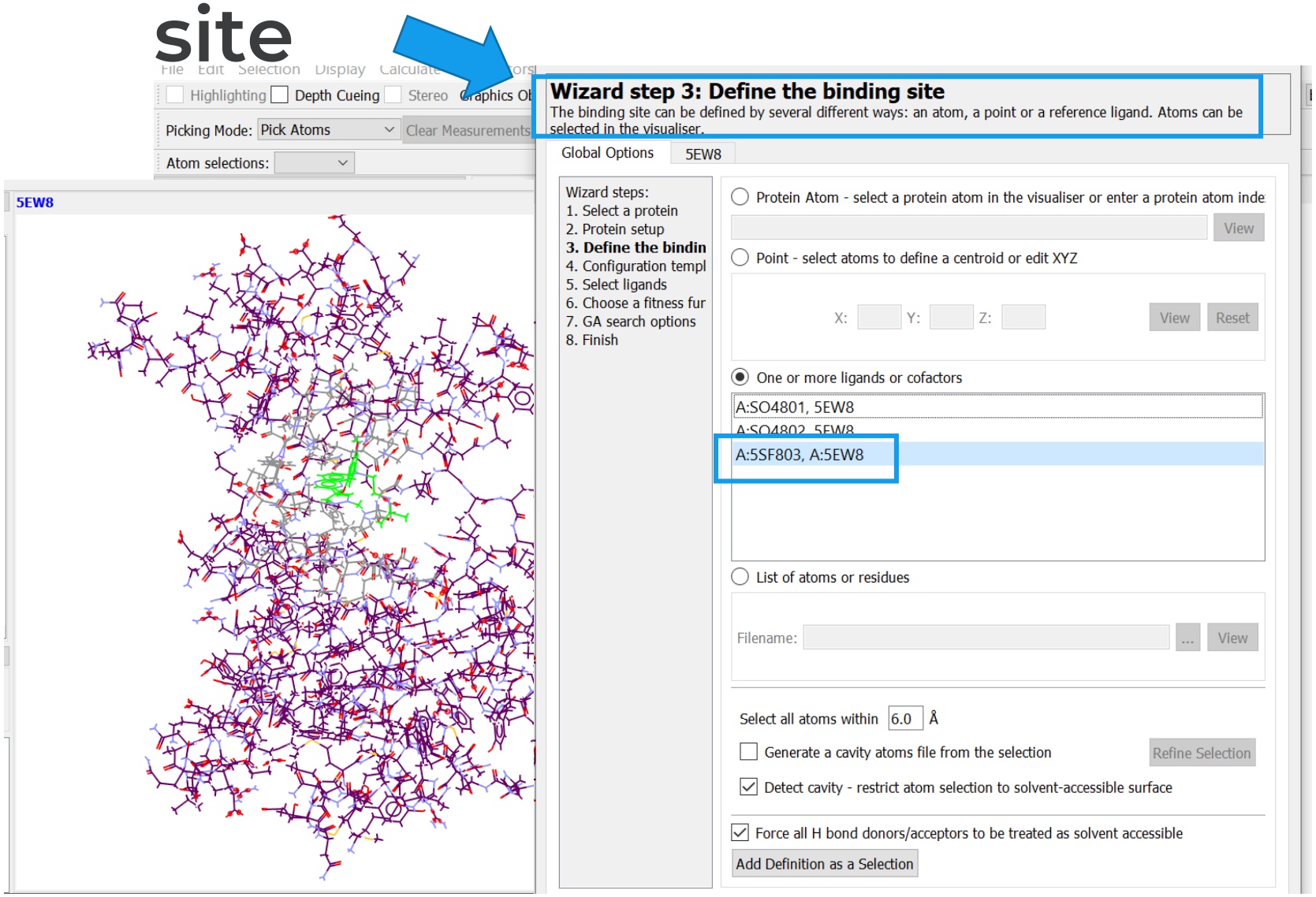
CCDC

Docking with GOLD: Protein preparation



- You can extract and save the bound ligands.

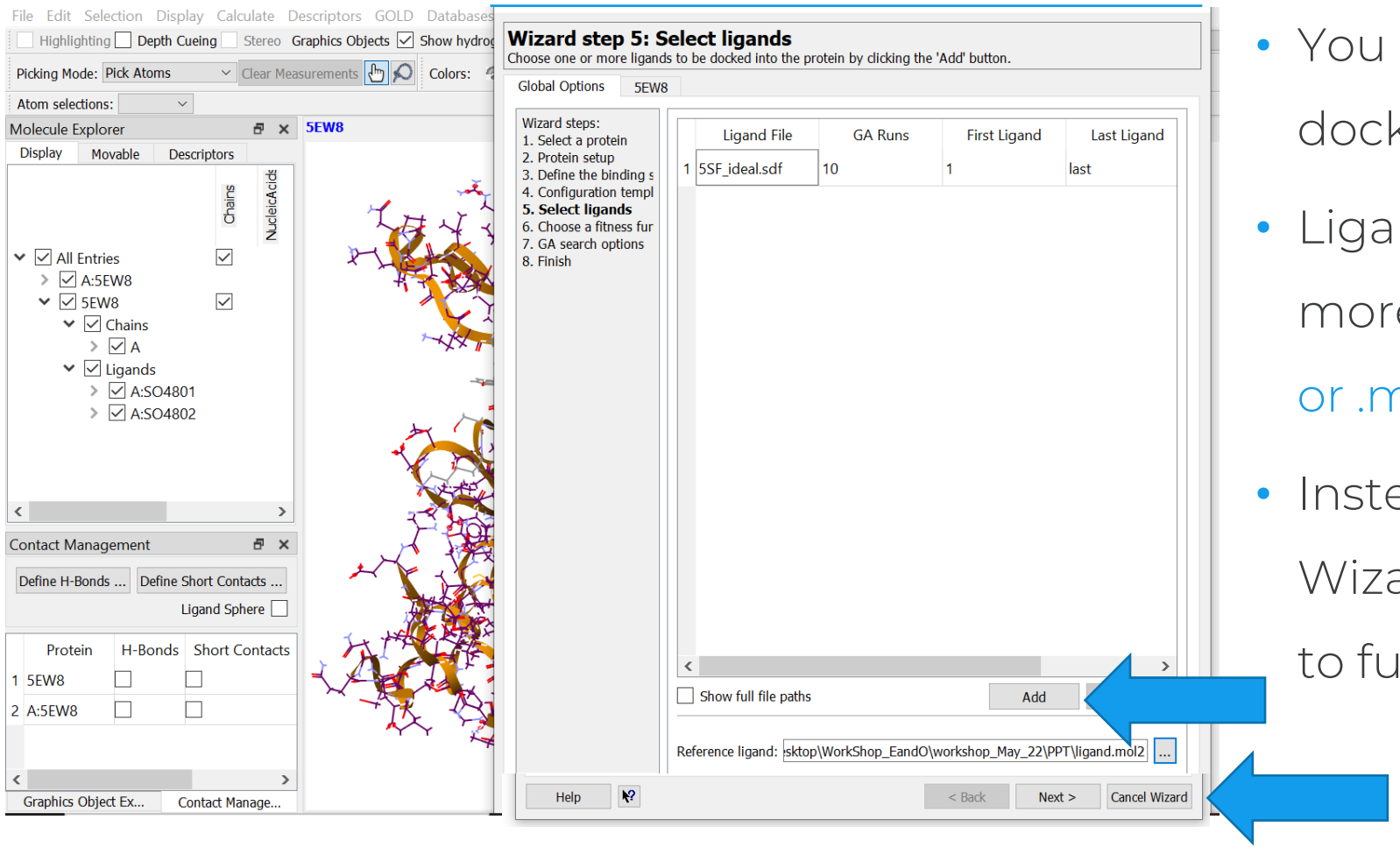
Docking with GOLD: Defining the binding site



- There are various ways in which you can define the binding site.
- Decide and select the one depending on the target protein or specific needs.
- You will see the binding site highlighted in the 3D visualizer.

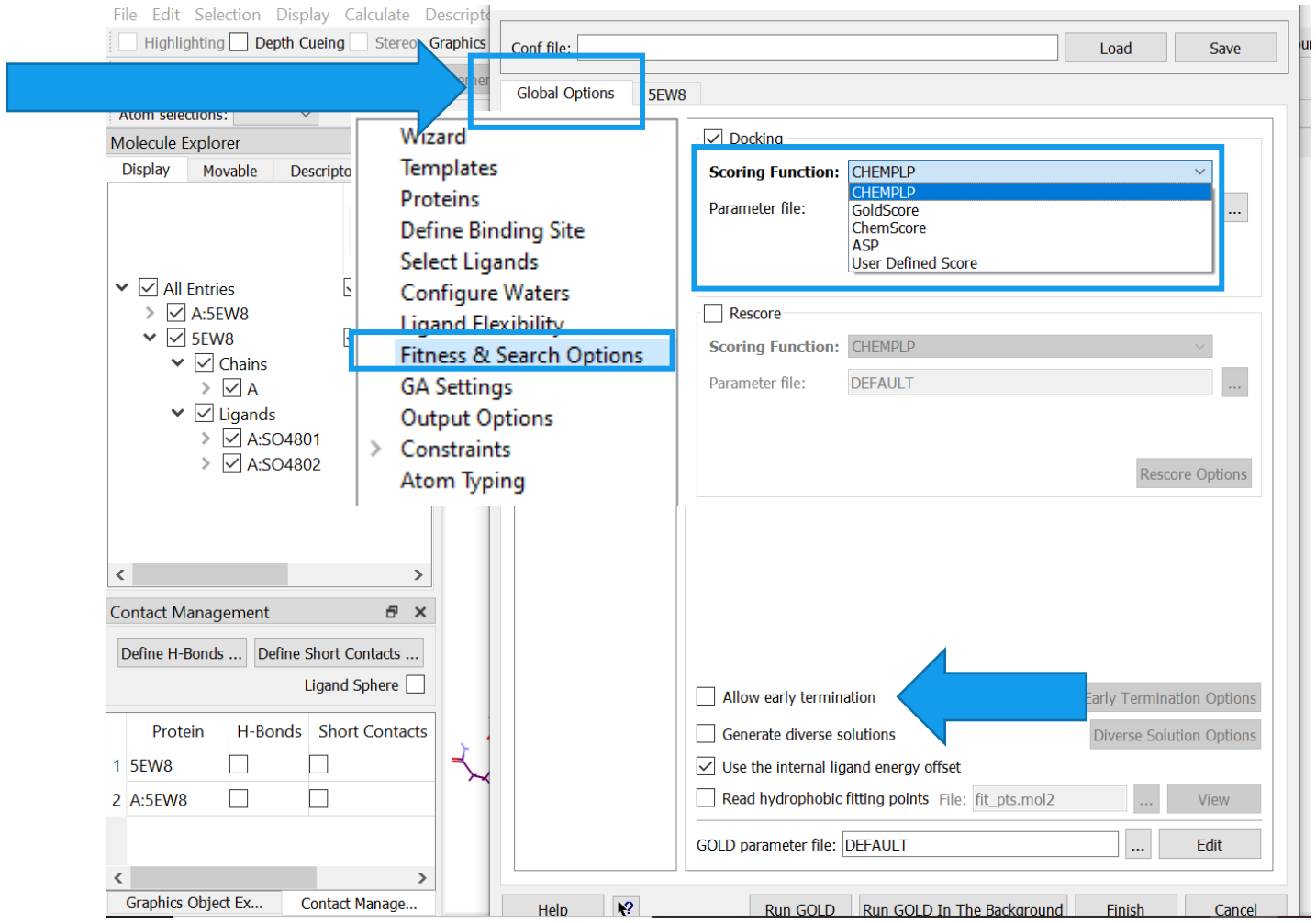
CCDC

Docking with GOLD: Select ligand/s



- You can select the ligands to be docked.
- Ligand file could contain one or more than one molecule in .sdf or .mol format.
- Instead of continuing with the Wizard, cancel the wizard now to further configure.

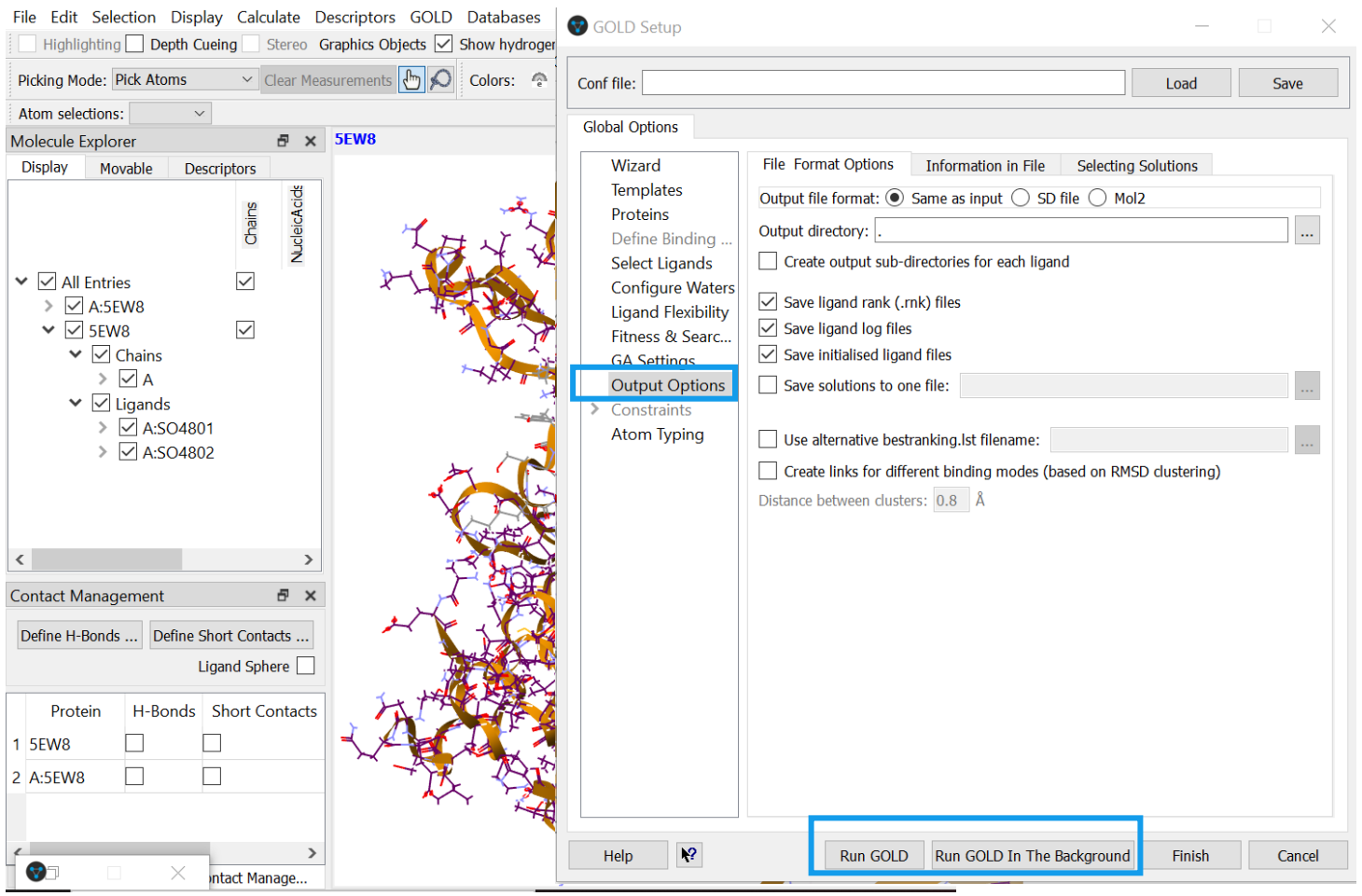
Docking with GOLD: Scoring function



- Under the Global Options, Fitness and Search Options you can select the Scoring Function.
- Here rescoring function is available for generating consensus.

- Untick *Allow early termination*

Docking with GOLD: Run the setup



In Output Options

- Select output destination/directory.
- Format of output you require.

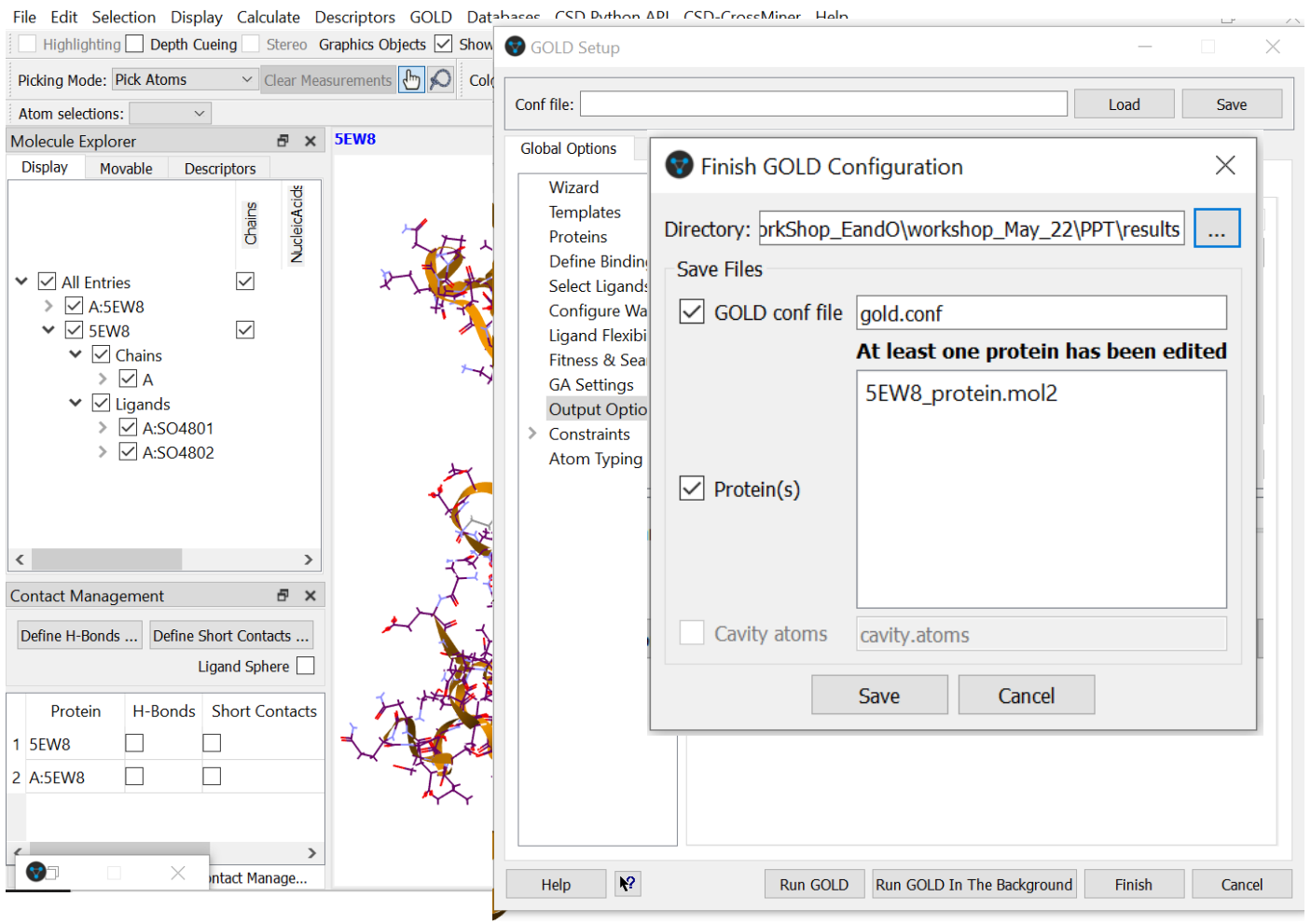
Run the GOLD calculations.

- Two options are available;

1. Run Gold: Interactive mode.

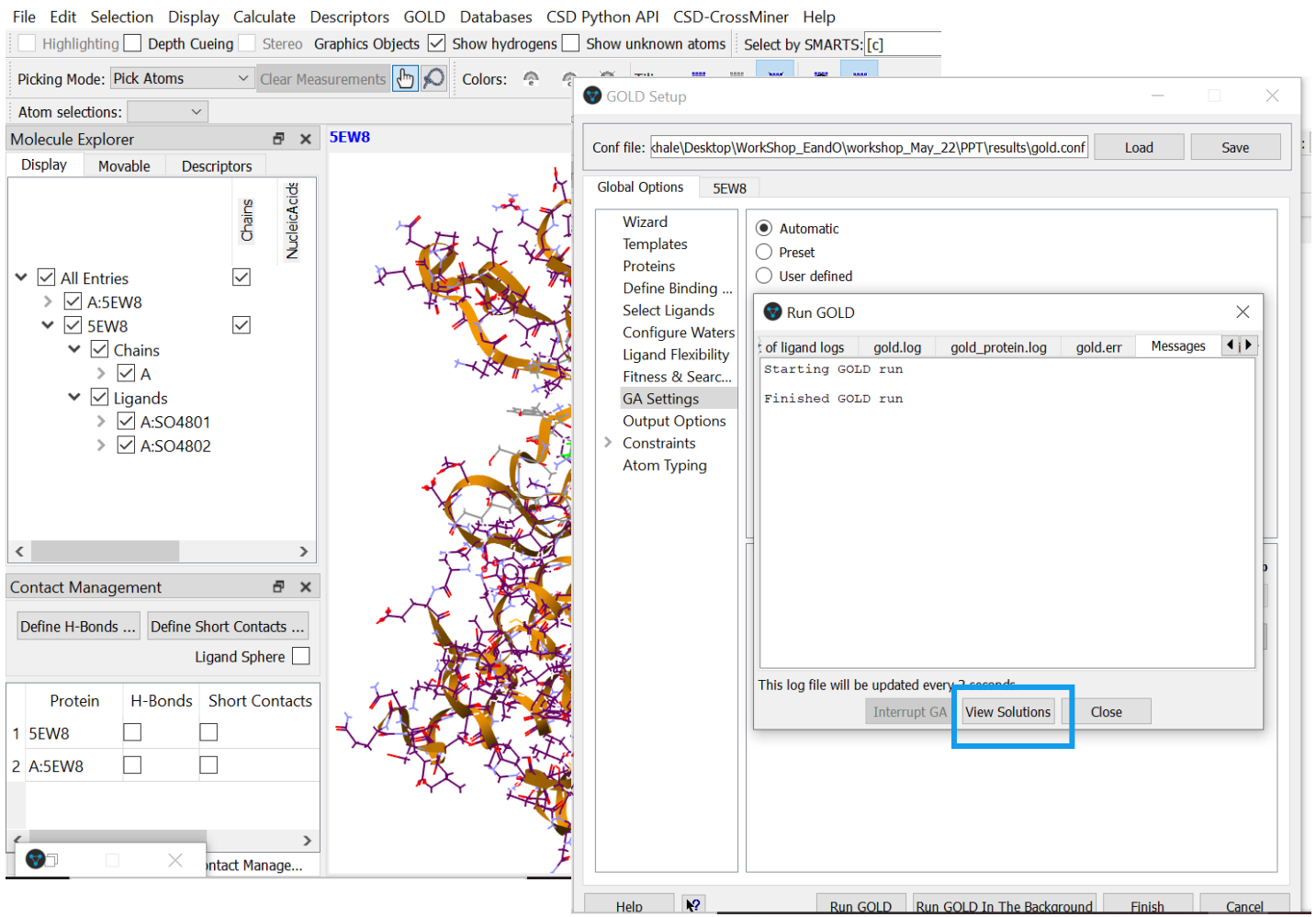
2. Run GOLD in Background: Results are seen in the output folder.

Docking with GOLD: Run the setup



- The run function asks for various options.
- The *gold.conf* file is the one with all the details. It is an editable file.

Docking with GOLD: Docking calculations

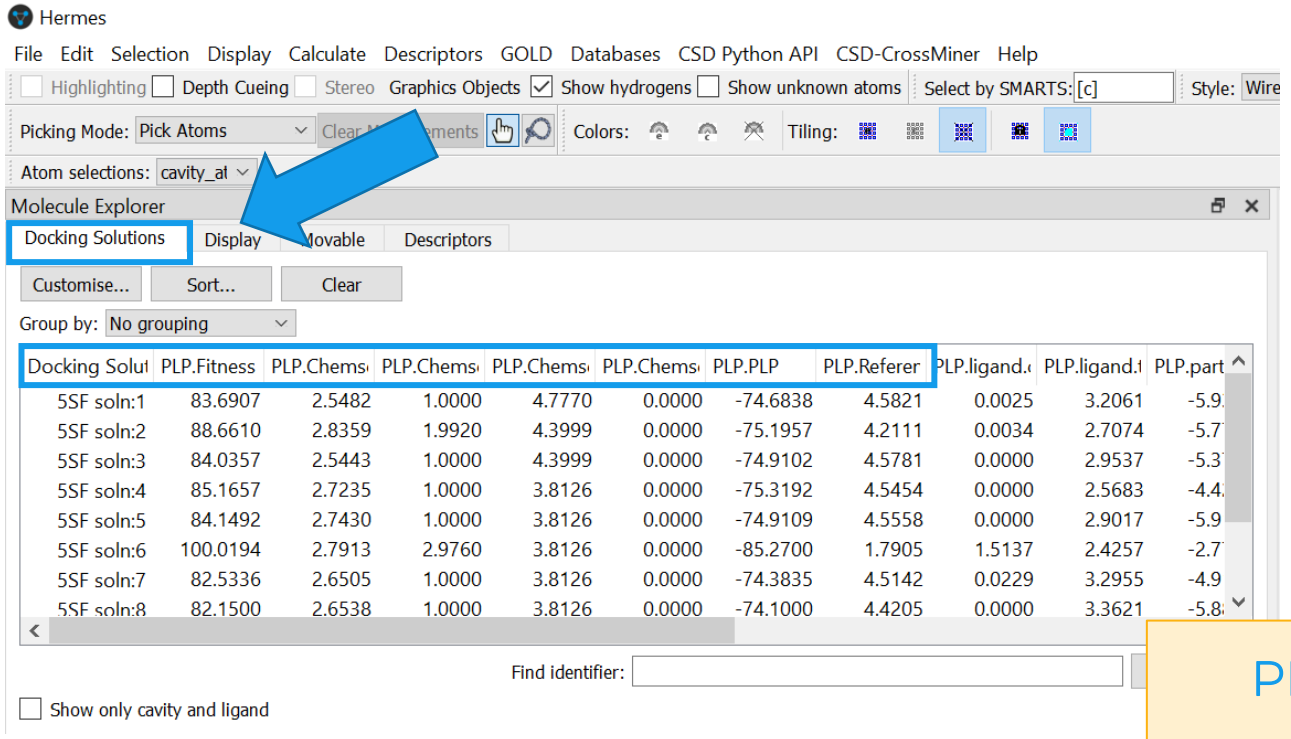


- You can check the progress of the calculation.
- The *Run GOLD* window has various tabs, which provide with on-the-fly status of your calculations.
- Once the run is complete, you can click *View Solutions*.

CCDC

Docking with GOLD: Analysis of results

- Results are displayed in the *Molecular Explorer*.

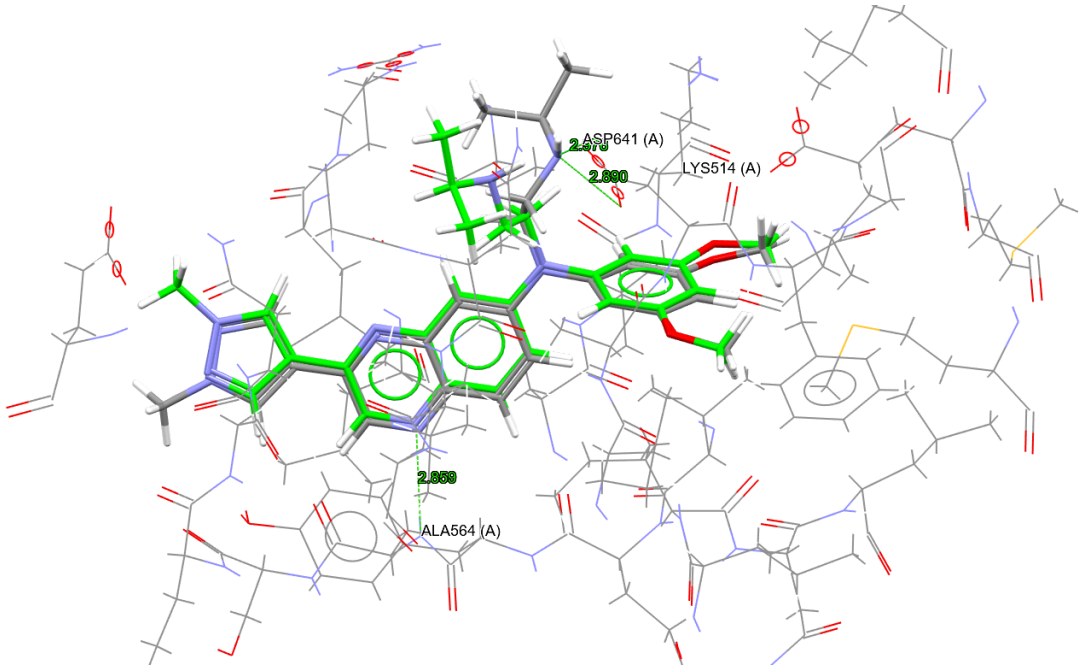


- In most cases we are interested in the PLP fitness score and the RMSD.
- For PLP scoring function, higher scores and lower RMSD means better results.

PLP: Piecewise Linear Potential, it is an empirical fitness functions optimised for pose prediction.

CCDC

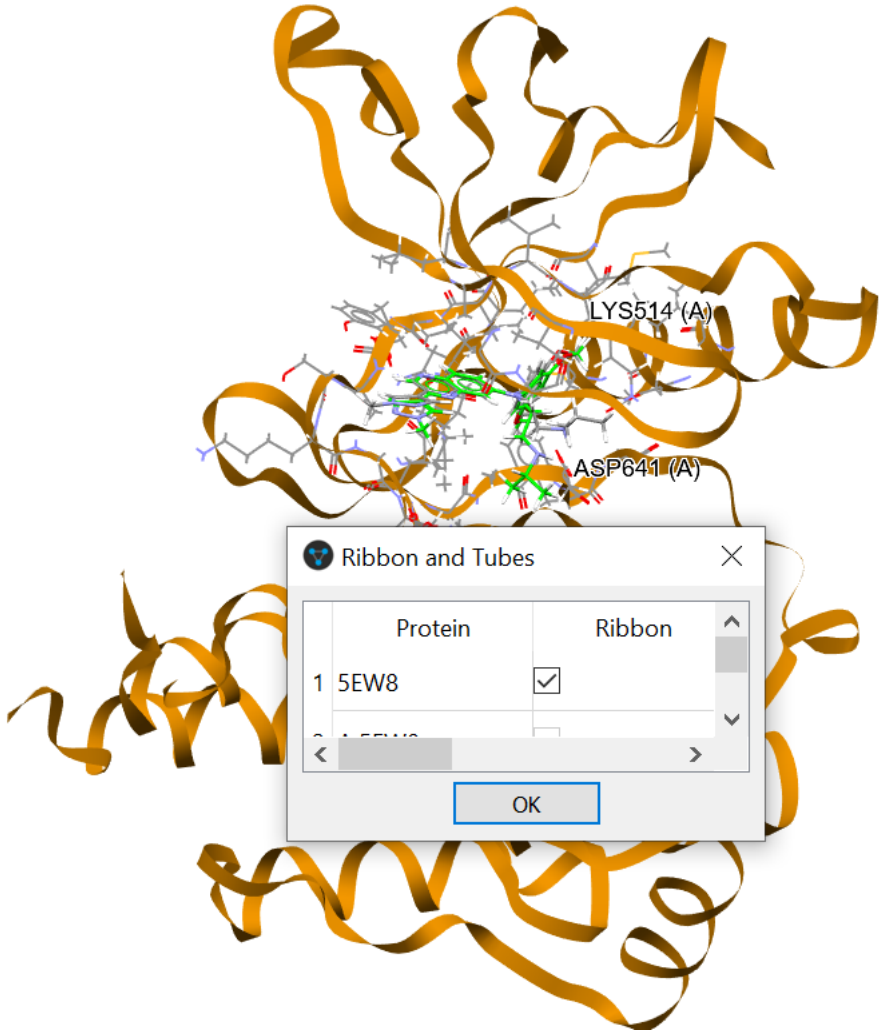
Docking with GOLD: Analysis of results



- Use the *Molecular Explorer* to display docking solutions, component of the system and the molecules.
- Manage the views and study the interactions.

CCDC

Docking with GOLD: Analysis of results



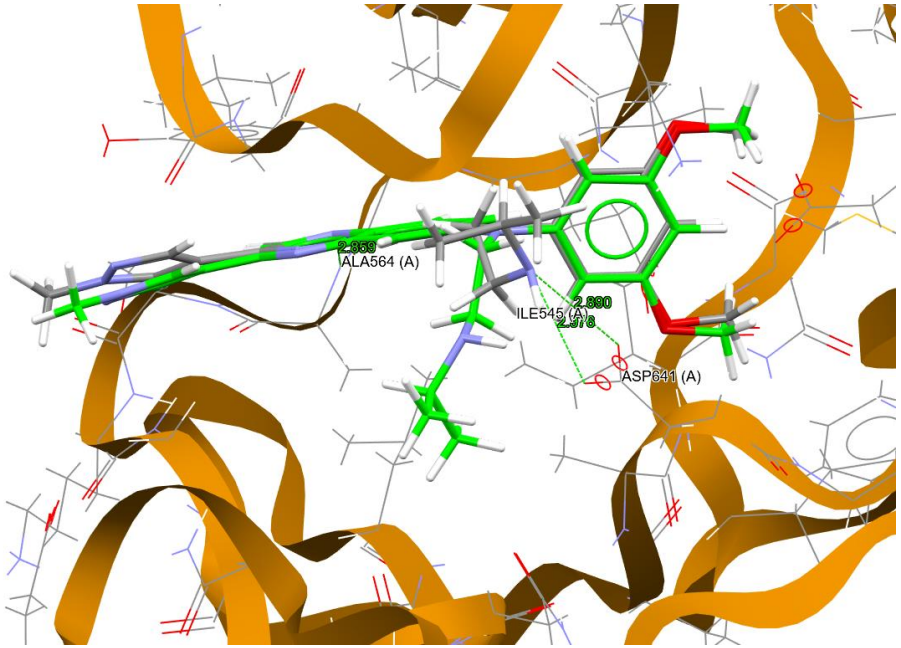
- Select Display

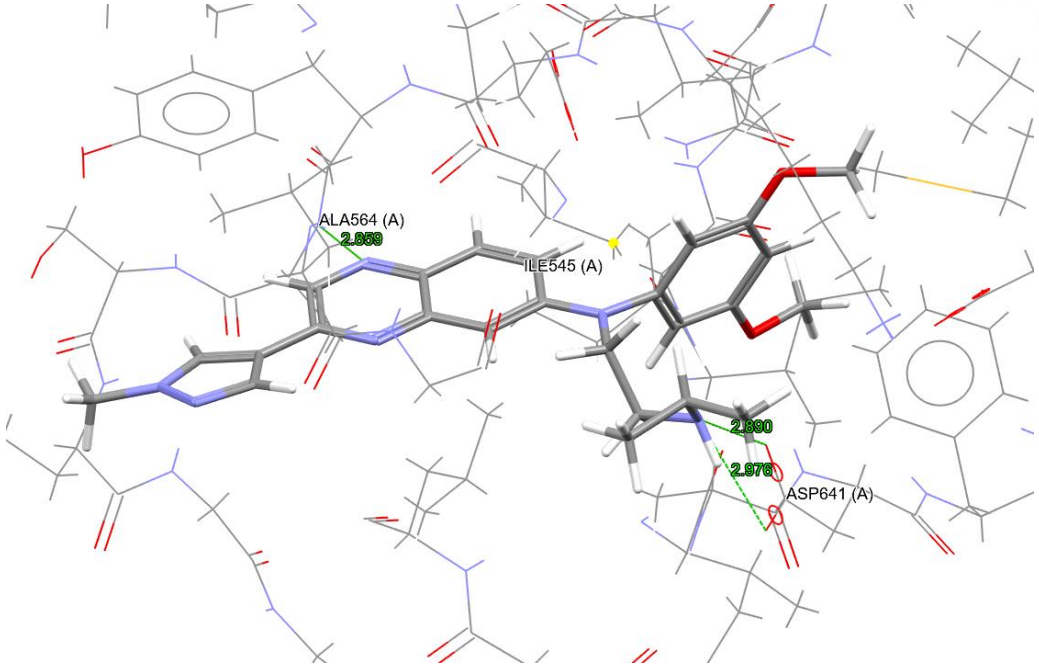
Ribbons and Tubes

Explore various options to create various colour combinations and displays

CCDC

Docking with GOLD: Analysis of results





- Results can be visualised in various ways and representations.
- Can display hydrogen bonds for the docked and reference ligands.

CCDC