

Protein-ligand docking 101 - running a simulation in GOLD

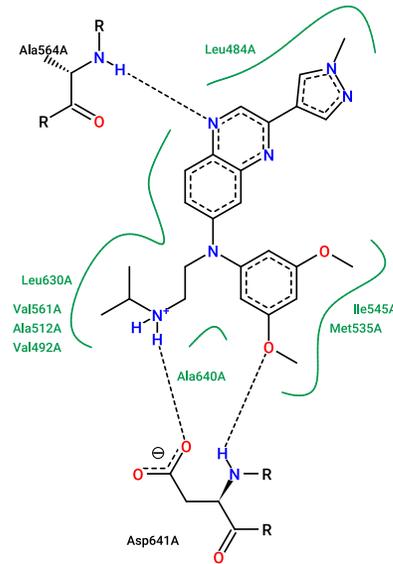
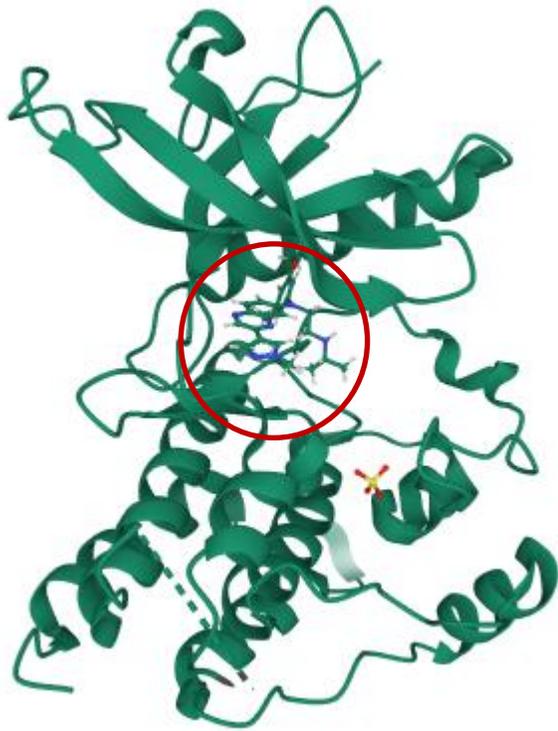
The following slides report all the steps that are shown in the tutorial video for this CSDU module, so you can follow each step in the format you prefer.



CCDC
advancing structural science

Docking with GOLD: Case Study

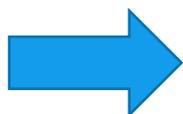
PDB: 5EW8



- **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.

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.



The screenshot shows the Hermes software interface. The 'CSD Python API' menu is open, and the 'Import' option is selected. The 'fetch_from_pdb.py' option is highlighted with a red box. The interface includes a menu bar with 'File', 'Edit', 'Selection', 'Display', 'Calculate', 'Descriptors', 'GOLD', 'Databases', 'CSD Python API', 'CSD-CrossMiner', and 'Help'. Below the menu bar are various toolbars and a 'Molecule Explorer' panel with tabs for 'Display', 'Movable', and 'Descriptors'. The 'Descriptors' tab is active, showing a list of descriptors: 'Chains', 'NucleicAcids', 'Ligands', 'Cofactors', 'Waters', 'Metals', and 'Packing'.

Docking with GOLD: Importing Protein

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

The screenshot shows the CSD Python API interface. The 'Import' menu is open, and 'fetch_from_pdb.py' is selected. A dialog box prompts for a PDB code (e.g. 5SY9) with '5EW8' entered. A terminal window shows the script execution details:

```
hermes
The python interpreter is: C:/Program Files/CCDC/Python_API_2021/miniconda/python.exe
The working directory is: C:/Users/rchikhale/Hermes/fetch_from_pdb/2022_03_31_17_23_34
Output files will be written in: C:/Users/rchikhale/Hermes//fetch_from_pdb/2022_03_31_17_23_34
Script is running... Parsed parameters for C:\Program Files\CCDC\Discovery_2021\Hermes\scripts\Import\fetch_from_pdb.py.
32 seconds
Stop!
```

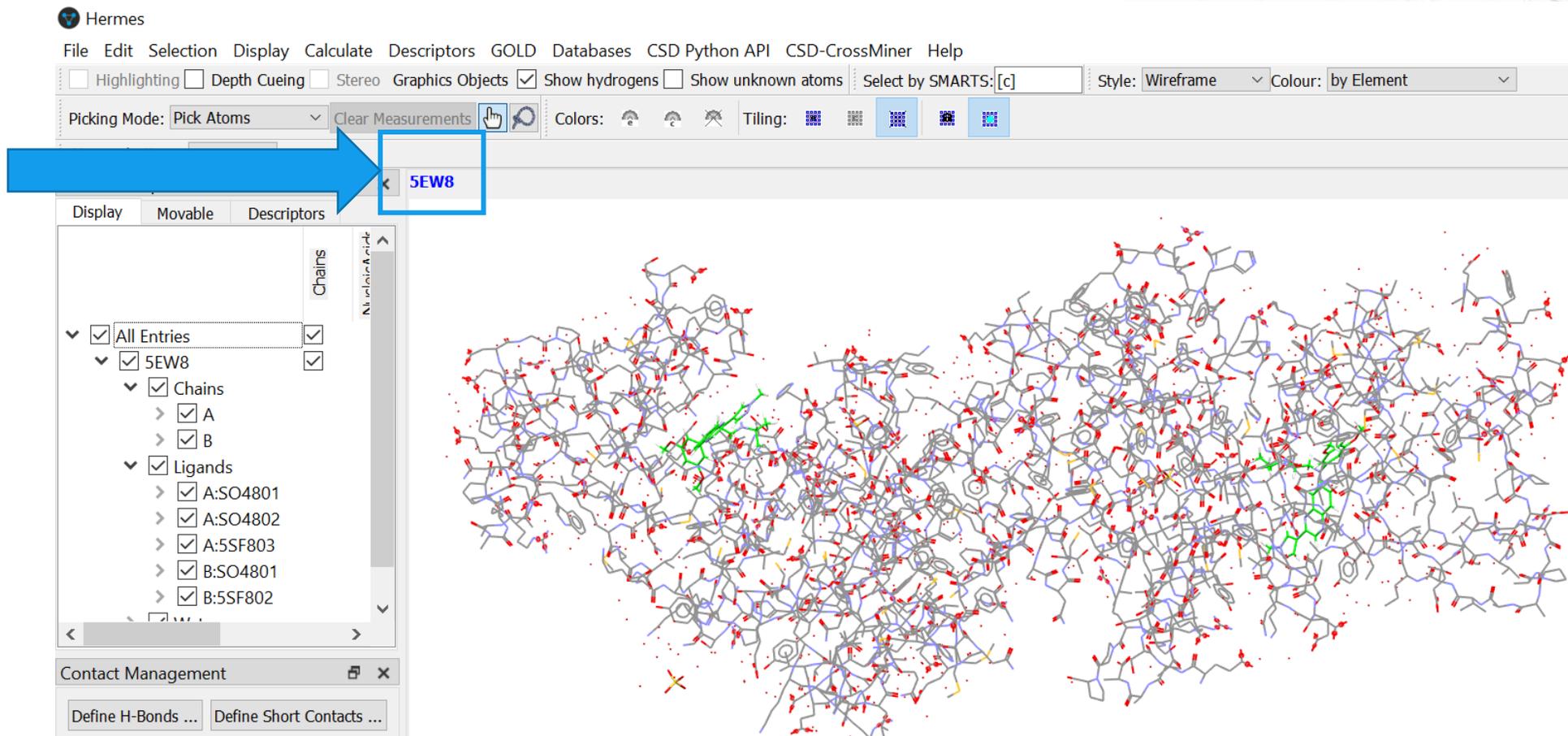


Tip

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.

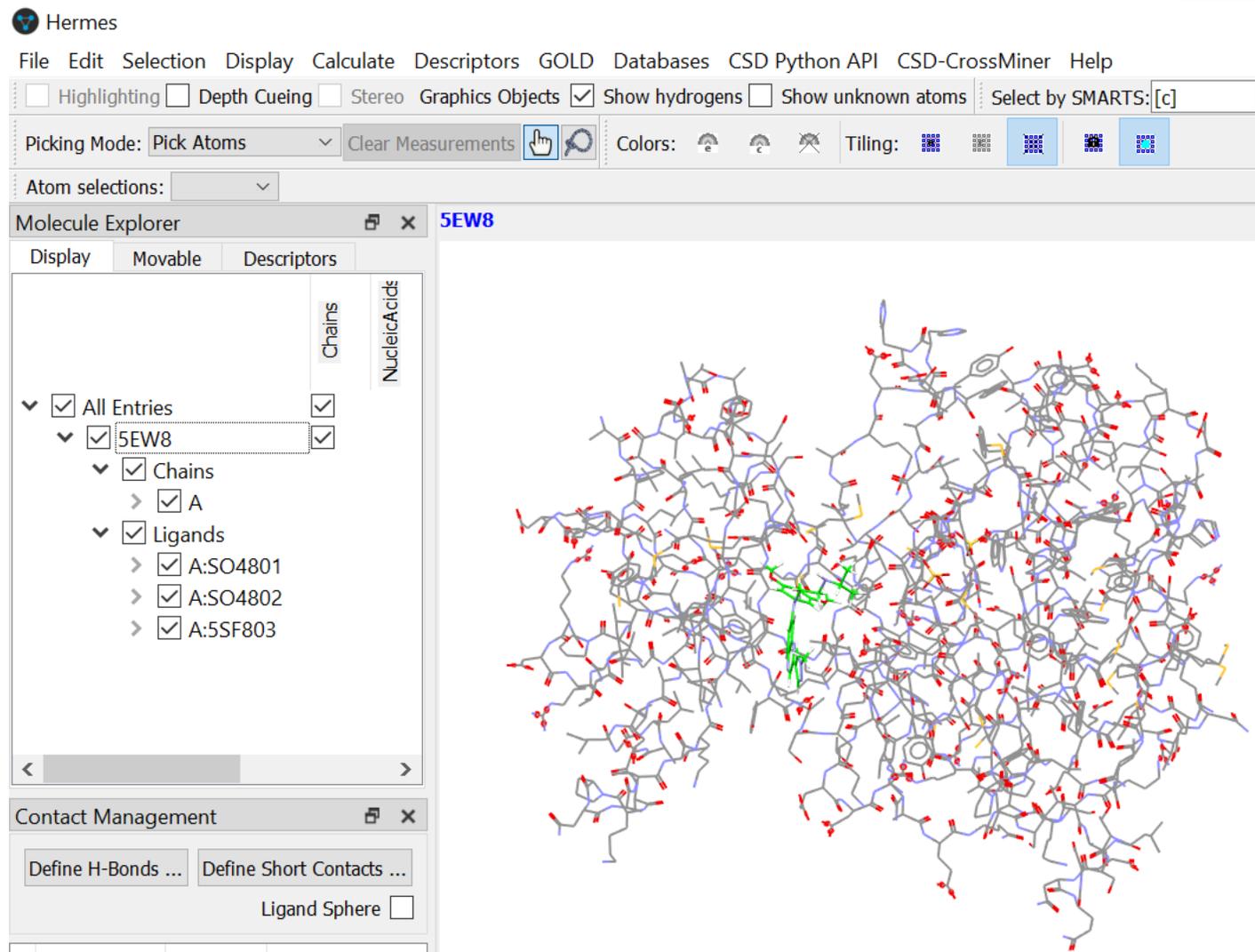


Docking with GOLD: Protein preparation

The screenshot displays the GOLD software interface. At the top, the menu bar includes File, Edit, Selection, Display, Calculate, Descriptors, GOLD, Databases, CSD Python API, CSD-CrossMiner, and Help. Below the menu bar, there are various toolbars and options, including 'Picking Mode: Pick Atoms', 'Clear Measurements', 'Colors', and 'Tiling'. A blue arrow points to the 'Atom selections:' dropdown menu. Below this, the 'Molecule Explorer' window is visible, showing a tree view of the protein structure. A blue arrow points to the 'Chains' section, where 'A' and 'B' are selected. A context menu is open over the 'Chains' section, with 'Delete' highlighted. The main 3D view shows a protein structure with a blue star icon on one of the chains.

- 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.

Docking with GOLD: Protein preparation

The screenshot displays the GOLD software interface. The main window shows a protein structure (5EW8) in a wireframe representation. The 'GOLD' menu is open, and the 'Wizard...' option is highlighted with a blue arrow. The 'GOLD Setup' dialog box is open, showing the 'Wizard step 1: Select one or more proteins' screen. The 'Select proteins to use:' section has a checkbox for '5EW8' checked, with a blue arrow pointing to it. The 'Wizard steps' list includes: 1. Select a protein, 2. Protein setup, 3. Define the binding site, 4. Configuration template, 5. Select ligands, 6. Choose a fitness function, 7. GA search options, and 8. Finish. The 'Protein score offset (ensemble docking only)' section is also visible, with a table showing the score offset for protein 5EW8 as 0.

Protein	Score Offset
1 5EW8	0



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

Docking with GOLD: Protein preparation

The screenshot displays the GOLD software interface. The main window shows a 3D ribbon representation of a protein structure (5EW8) in orange and yellow, with various atoms colored (red, blue, grey). The interface includes a menu bar (File, Edit, Selection, Display, Calculate, Descriptors, GOLD, Databases, CSD Python API, CSD-CrossMiner, Help) and a toolbar with options like Highlighting, Depth Cueing, Stereo, Graphics Objects, Show hydrogens, and Show unknown atoms. The Molecule Explorer on the left shows a tree view for 5EW8, including Chains (A) and Ligands (A:SO4801, A:SO4802, A:SSF803). The Contact Management panel at the bottom left shows a table for H-Bonds and Short Contacts.

Wizard step 2: Protein setup
At this point you have the chance to edit your protein structure if required e.g. add hydrogens, delete waters

Global Options: 5EW8

Wizard steps:

1. Select a protein
- 2. Protein setup**
3. Define the binding site
4. Configuration template
5. Select ligands
6. Choose a fitness function
7. GA search options
8. Finish

To edit the protein(s) use the options available on the protein tabs to:

1. Add Hydrogens
2. Configure active waters and delete unnecessary waters
3. Delete ligands

Buttons: Help, < Back, Next >, Cancel Wizard

Docking with GOLD: Protein preparation

Wizard step 2: Protein setup
At this point you have the chance to edit your protein structure if required e.g. add hydrogens, delete waters...

Global Options SEW8

Protonation & Tautomers

- Extract/Delete Waters
- Delete Ligands/Cofactors
- Flexible Sidechains
- Soft Potentials
- Metals
- Constraints
- Covalent
- Interaction Motif

Protonation Rules: [] ...

Add Hydrogens

Residue Set: All Protein Residues

Flip Asn Gln His Tautomers

HIS541 A
HIS621 A
HIS649 A
HIS650 A
HIS679 A
HIS717 A
HIS738 A

Edit selected residue

ND1 H
 NE2 H

Set Protonation

Flip

Protein H-Bonds Short Contacts

Protein	H-Bonds	Short Contacts
1 SEW8	<input type="checkbox"/>	<input type="checkbox"/>

- Add missing hydrogens.

Docking with GOLD: Protein preparation

Wizard step 2: Protein setup
At this point you have the chance to edit your protein structure if required e.g. add hydrogens, delete w

Global Options 5EW8

Protonation & Tautomers
Extract/Delete Waters
Delete Ligands/Cofactors
Flexible Sidechains
Soft Potentials
Metals
> Constraints
Covalent
Interaction Motif

Waters

Extract Waters For Docking Delete Remaining Waters

Waters that are important for ligand binding can be added to the list of act used in docking by checking them in the water list above and clicking the 't Waters For Docking' button.

Once you have extracted the waters you wish to make active press the 'Del Remaining Waters' button to remove all waters from the protein.

Help ? < Back Next > Cancel Wizard

Molecule Explorer
Display Movable Descriptors
Chains NucleicAcid
NucleicAcid
All Entries
5EW8
Chains
A
Ligands
A:SO4801
A:SO4802
A:5SF803

Contact Management
Define H-Bonds ... Define Short Contacts ...
Ligand Sphere
Protein H-Bonds Short Contacts
1 5EW8

- 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).

Docking with GOLD: Protein preparation

The screenshot displays the GOLD software interface during the 'Wizard step 2: Protein setup'. The main window is titled 'Protonation & Tautomers' and contains a list of actions: 'Extract/Delete Waters', 'Delete Ligands/Cofactors', 'Flexible Sidechains', 'Soft Potentials', 'Metals', 'Constraints', 'Covalent', and 'Interaction Motif'. The 'Delete Ligands/Cofactors' option is highlighted with a blue box. Below this list is a table with the following data:

Ligand/Cofactor	Extract and Reload
A:SO4801	<input type="checkbox"/>
A:SO4802	<input type="checkbox"/>
A:5SF803	<input checked="" type="checkbox"/>

The 'A:5SF803' row is also highlighted with a blue box. Below the table is an 'Extract' button, which is also highlighted with a blue box and pointed to by a blue arrow. At the bottom of the window, there are 'Back' and 'Next >' buttons, with the 'Next >' button highlighted by a blue box. On the left side, the 'Molecule Explorer' panel shows a tree view with 'All Entries', '5EW8', 'Chains', 'A', 'Ligands', 'A:SO4801', 'A:SO4802', and 'A:5SF803'. The 'Contact Management' panel at the bottom left shows a table with the following data:

Protein	H-Bonds	Short Contacts
1 5EW8	<input type="checkbox"/>	<input type="checkbox"/>

The 'Next >' button is highlighted with a blue box.

- You can extract and save the bound ligands.

Docking with GOLD: Defining the binding site

Wizard step 3: Define the binding site
The binding site can be defined by several different ways: an atom, a point or a reference ligand. Atoms can be selected in the visualiser.

Global Options 5EW8

Wizard steps:
1. Select a protein
2. Protein setup
3. **Define the binding site**
4. Configuration template
5. Select ligands
6. Choose a fitness function
7. GA search options
8. Finish

Protein Atom - select a protein atom in the visualiser or enter a protein atom index
 Point - select atoms to define a centroid or edit XYZ
X: Y: Z:

One or more ligands or cofactors

A:SO4801, 5EW8
A:SO4802, 5EW8
A:5SF803, A:5EW8

List of atoms or residues
Filename:

Select all atoms within Å

Generate a cavity atoms file from the selection

Detect cavity - restrict atom selection to solvent-accessible surface

Force all H bond donors/acceptors to be treated as solvent accessible

- 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.

Docking with GOLD: Select ligand/s

Wizard step 5: Select ligands
Choose one or more ligands to be docked into the protein by clicking the 'Add' button.

Global Options 5EW8

Wizard steps:
1. Select a protein
2. Protein setup
3. Define the binding site
4. Configuration template
5. **Select ligands**
6. Choose a fitness function
7. GA search options
8. Finish

Ligand File	GA Runs	First Ligand	Last Ligand
1 5SF_ideal.sdf	10	1	last

Show full file paths

Reference ligand: ...

Buttons: Help, < Back, Next >, Cancel Wizard

- 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

The screenshot shows the GOLD software interface. The 'Global Options' dialog box is open, and the 'Fitness & Search Options' tab is selected. The 'Scoring Function' dropdown is set to 'CHEMPLP'. A blue arrow points to the 'Global Options' dialog, and another blue arrow points to the 'Allow early termination' checkbox.

Wizard

- Templates
- Proteins
- Define Binding Site
- Select Ligands
- Configure Waters
- Ligand Flexibility
- Fitness & Search Options**
- GA Settings
- Output Options
- Constraints
- Atom Typing

Global Options 5EW8

Docking

Scoring Function: CHEMPLP

Parameter file: CHEMPLP, GoldScore, ChemScore, ASP, User Defined Score

Rescore

Scoring Function: CHEMPLP

Parameter file: DEFAULT

Rescore Options

Allow early termination

Generate diverse solutions

Use the internal ligand energy offset

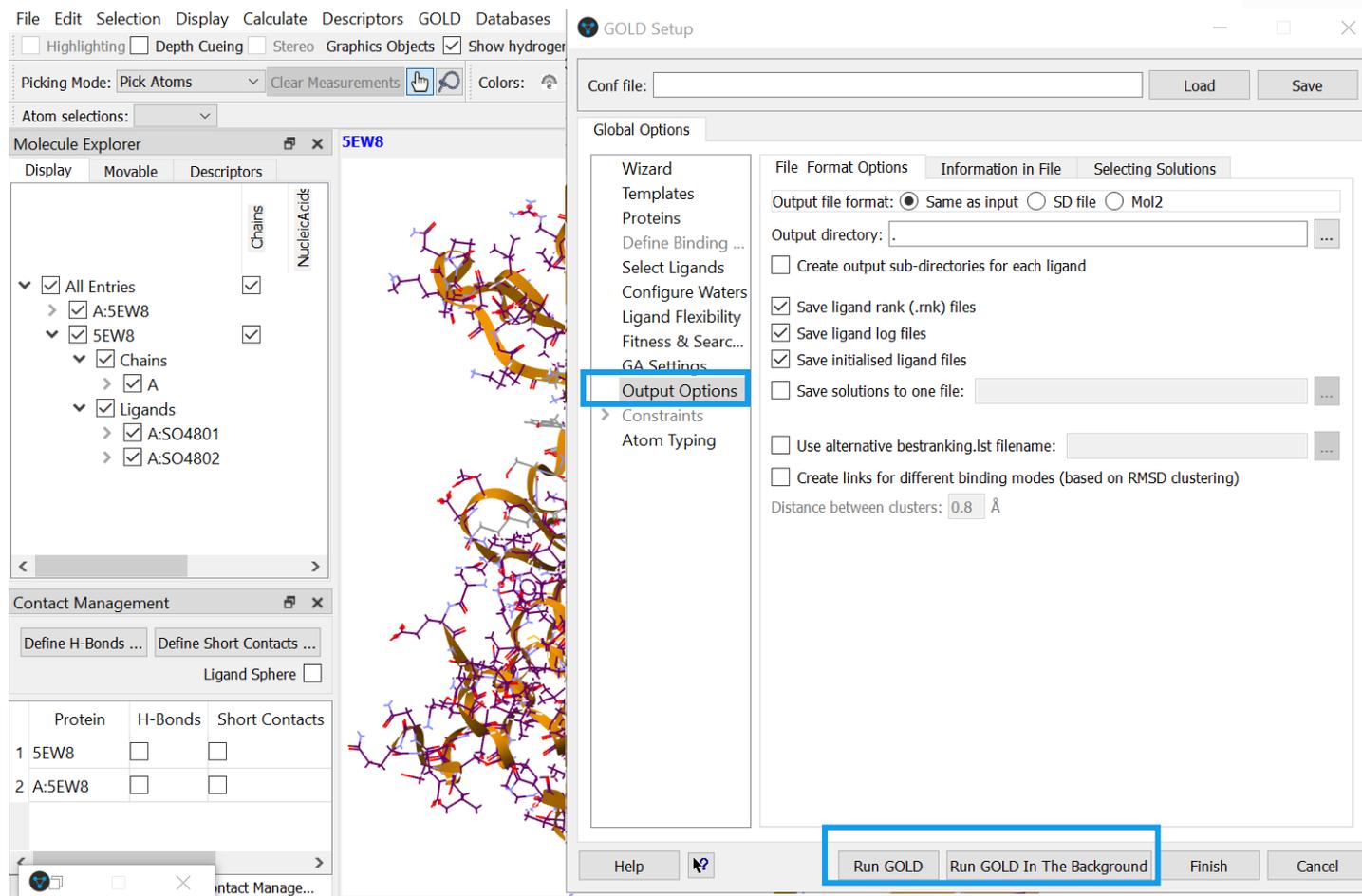
Read hydrophobic fitting points File: fit_pts.mol2 View

GOLD parameter file: DEFAULT Edit

Run GOLD Run GOLD In The Background Finish Cancel

- 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

File Edit Selection Display Calculate Descriptors GOLD Database CSD Python API CSD-CrossMiner Help

Highlighting Depth Cueing Stereo Graphics Objects Show

Picking Mode: Pick Atoms Clear Measurements

Atom selections:

Molecule Explorer SEW8

Display Movable Descriptors

Chains NucleicAcids

✓ All Entries

> ✓ A:5EW8

✓ 5EW8

✓ Chains

> ✓ A

✓ Ligands

> ✓ A:SO4801

> ✓ A:SO4802

Contact Management

Define H-Bonds ... Define Short Contacts ...

Ligand Sphere

Protein	H-Bonds	Short Contacts
1 5EW8	<input type="checkbox"/>	<input type="checkbox"/>
2 A:5EW8	<input type="checkbox"/>	<input type="checkbox"/>

Global Options

Wizard

Templates

Proteins

Define Binding

Select Ligands

Configure Wa

Ligand Flexibi

Fitness & Sea

GA Settings

Output Optio

> Constrains

Atom Typing

Conf file: Load Save

Finish GOLD Configuration

Directory: prkShop_EandO\workshop_May_22\PPT\results

Save Files

GOLD conf file gold.conf

At least one protein has been edited

5EW8_protein.mol2

Protein(s)

Cavity atoms cavity.atoms

Save Cancel

Help ? Run GOLD Run GOLD In The Background Finish Cancel

- 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

The screenshot displays the GOLD software interface. The main window shows a 3D molecular model of a protein-ligand complex. The Molecule Explorer on the left lists the following entries:

- ✓ All Entries
 - ✓ A:5EW8
 - ✓ 5EW8
 - ✓ Chains
 - ✓ A
 - ✓ Ligands
 - ✓ A:SO4801
 - ✓ A:SO4802

The GOLD Setup window is open, showing the Global Options for 5EW8. The Wizard is set to Automatic. The Run GOLD window is also open, showing the progress of the calculation. The Run GOLD window has several tabs: of ligand logs, gold.log, gold_protein.log, gold.err, and Messages. The Messages tab is active, showing the following text:

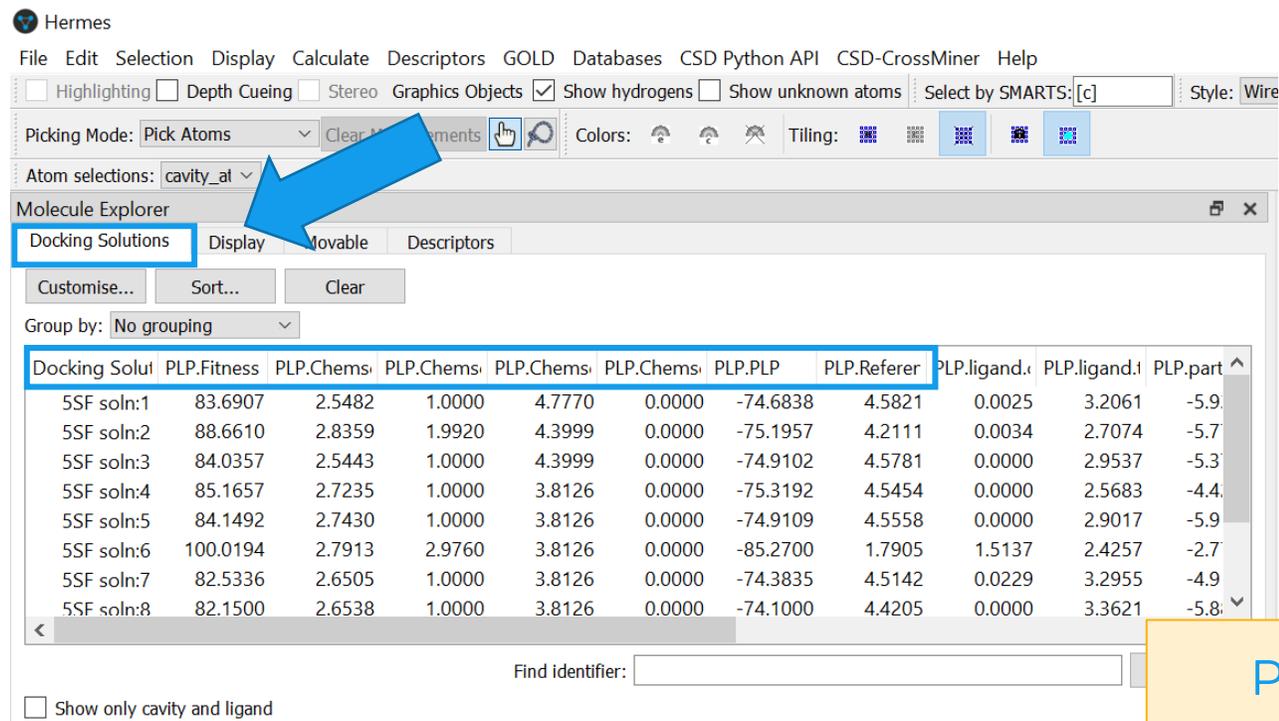
```
Starting GOLD run
Finished GOLD run
```

The Run GOLD window also has a "View Solutions" button highlighted with a blue box. The bottom of the Run GOLD window shows a progress bar and a "View Solutions" button.

- 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**.

Docking with GOLD: Analysis of results

- Results are displayed in the *Molecular Explorer*.



Atom selections: cavity_at

Molecule Explorer

Docking Solutions | Display | Movable | Descriptors

Group by: No grouping

Docking Solu	PLP.Fitness	PLP.Chems	PLP.Chems	PLP.Chems	PLP.Chems	PLP.PLP	PLP.Referer	PLP.ligand.i	PLP.ligand.l	PLP.part
5SF soln:1	83.6907	2.5482	1.0000	4.7770	0.0000	-74.6838	4.5821	0.0025	3.2061	-5.9
5SF soln:2	88.6610	2.8359	1.9920	4.3999	0.0000	-75.1957	4.2111	0.0034	2.7074	-5.7
5SF soln:3	84.0357	2.5443	1.0000	4.3999	0.0000	-74.9102	4.5781	0.0000	2.9537	-5.3
5SF soln:4	85.1657	2.7235	1.0000	3.8126	0.0000	-75.3192	4.5454	0.0000	2.5683	-4.4
5SF soln:5	84.1492	2.7430	1.0000	3.8126	0.0000	-74.9109	4.5558	0.0000	2.9017	-5.9
5SF soln:6	100.0194	2.7913	2.9760	3.8126	0.0000	-85.2700	1.7905	1.5137	2.4257	-2.7
5SF soln:7	82.5336	2.6505	1.0000	3.8126	0.0000	-74.3835	4.5142	0.0229	3.2955	-4.9
5SF soln:8	82.1500	2.6538	1.0000	3.8126	0.0000	-74.1000	4.4205	0.0000	3.3621	-5.8

Find identifier:

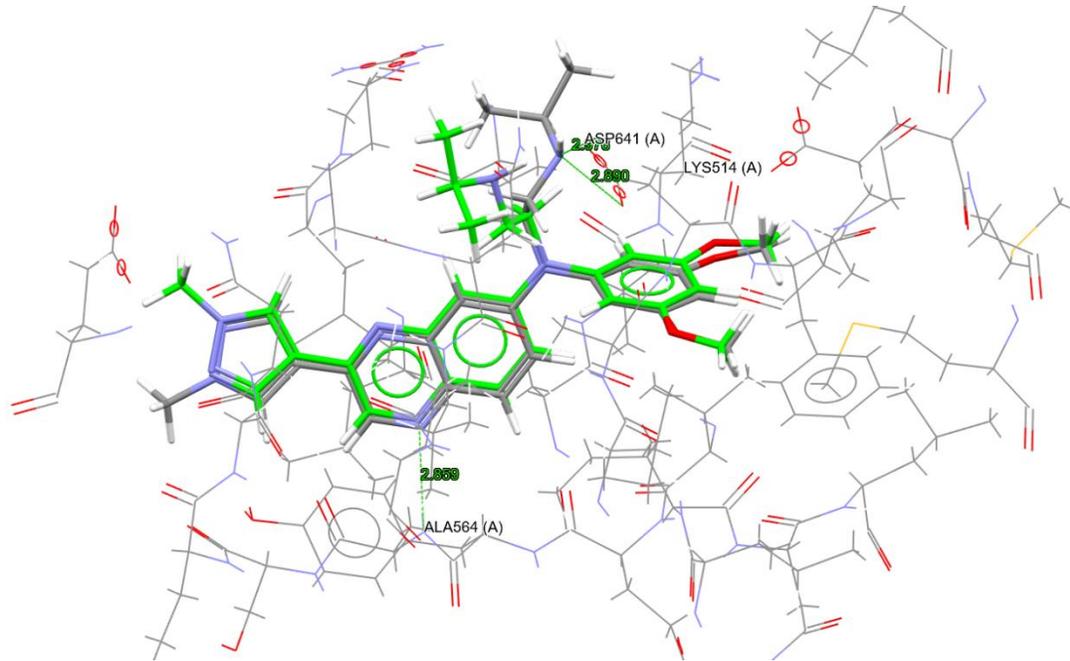
Show only cavity and ligand

- 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.

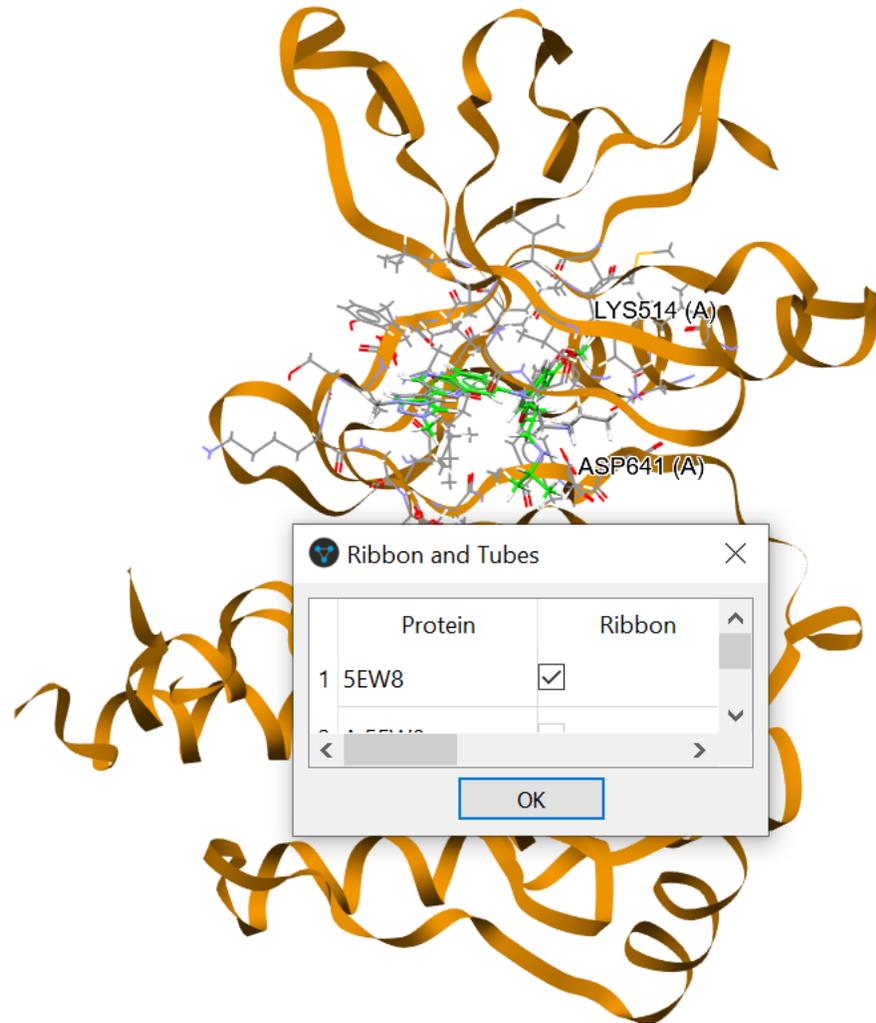
PLP: Piecewise Linear Potential, it is an empirical fitness functions optimised for pose prediction.

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.

Docking with GOLD: Analysis of results



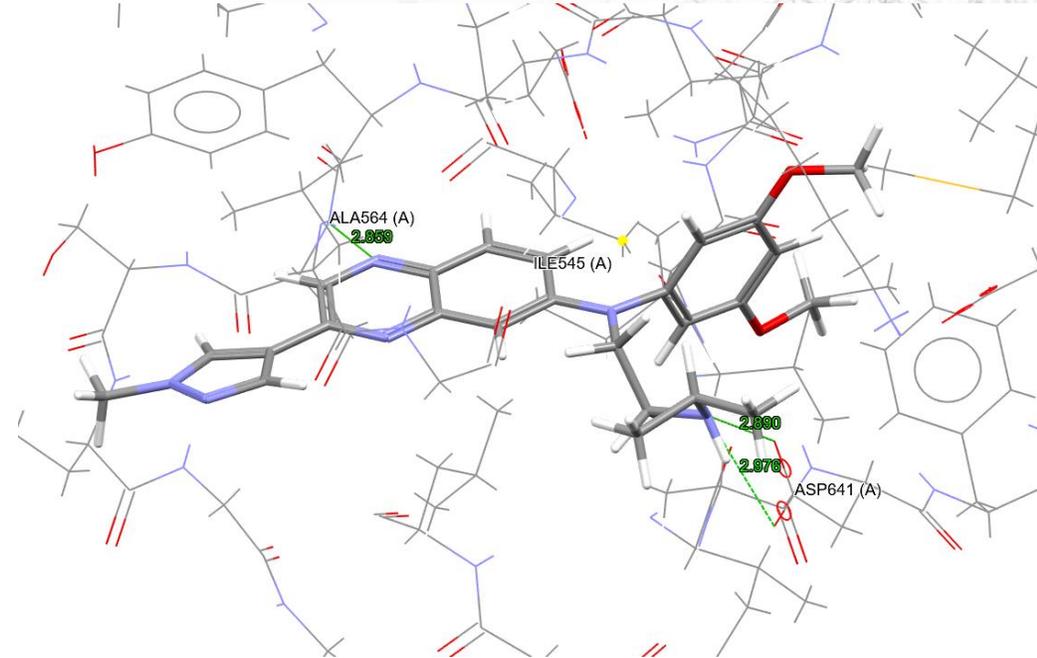
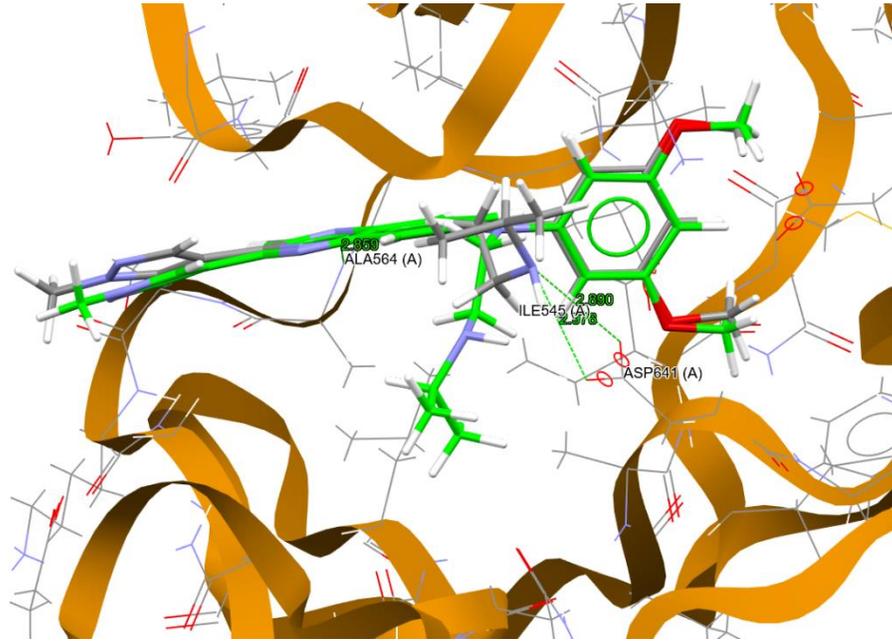
- Select **Display**



Ribbons and Tubes

Explore various options to create various colour combinations and displays

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.