Virtual Screening of a Library of KRAS Inhibitors using GOLD

Developed using
2025.1 CSD Release

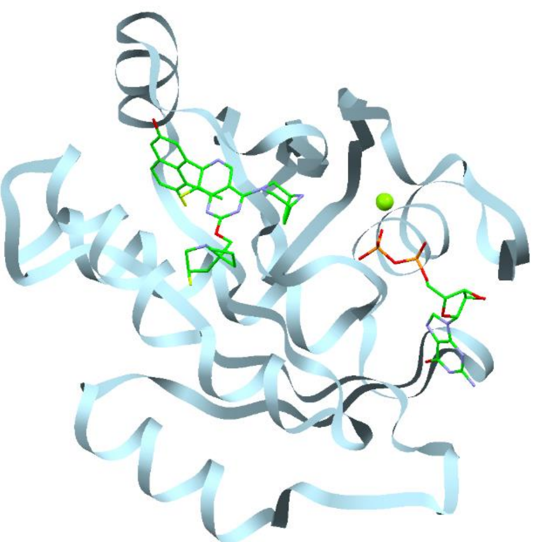


Table of Contents

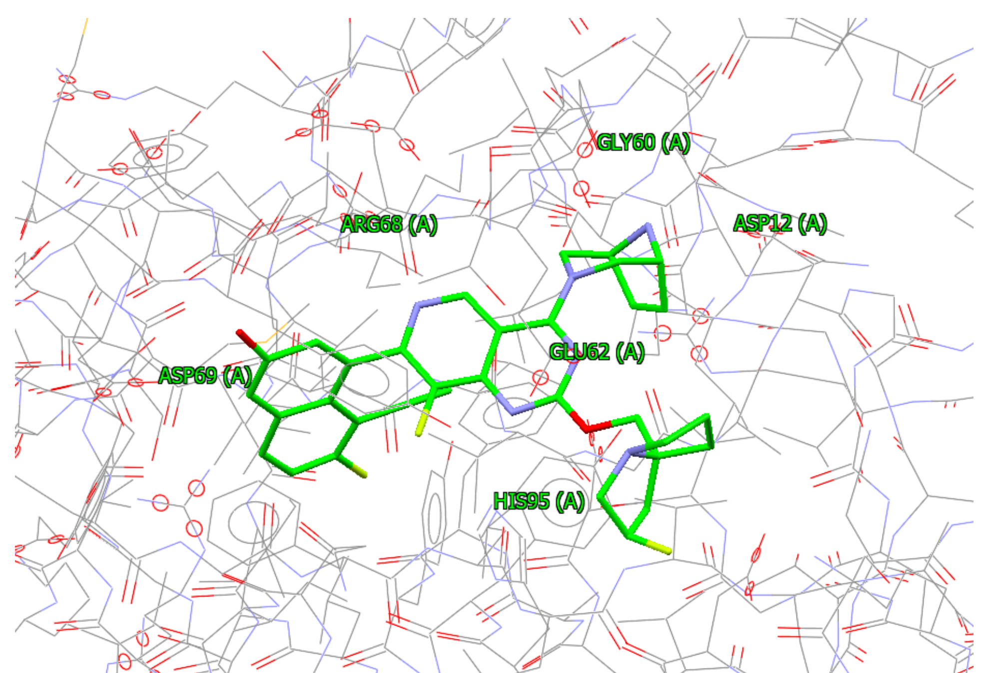
Learning Outcomes 3
Pre-required Skills 3
Materials........ 3
Example: Screening a Library of KRAS inhibitors within the Binding Pocket of MRTX-1133 4
Introduction........ 4
Challenges 5
Additional Files 6
Importing the Protein-Ligand Complex from the PDB 7
Editing the Structure 8
Exploring the Binding Site 10
Set up the GOLD Virtual Screening........ 12
Visualization of and Analysis of the Results 20
Conclusions........ 23
Summary........ 23
Next Steps........ 23
Analysing Interactions using SuperStar........ 24
Challenge 31
Glossary 32
Basics of Hermes Visualisation........ 35

Virtual Screening of a library of KRAS inhibitors using GOLD

Molecular docking is the most common computational structure-based drug design (SBDD) method and has been widely used ever since the early 1980s. It is the tool of choice when the three-dimensional (3D) structure of the protein target is available. Molecular docking popularity has been facilitated by the dramatic growth in availability and power of computers, and the increasing number of and ease of access to small molecule and protein structures. The main goal of molecular docking is to understand and predict molecular recognition, both structurally (i.e., finding possible binding modes) and energetically (i.e., predicting binding affinity). Molecular docking was originally designed to be performed between a small molecule (ligand) and a target macromolecule (protein).

Molecular docking applications in drug discovery are varied, including structure-activity studies, lead optimisation, finding potential leads by virtual screening, providing binding hypotheses to facilitate predictions for mutagenesis studies and, also in assisting X-ray crystallography and cryogenic electron microscopy (cryo-EM) in the fitting of substrates and inhibitors to electron density.[1]

GOLD (Genetic Optimisation for Ligand Docking) is a genetic algorithm for docking flexible ligands into protein binding sites. GOLD has been extensively tested and has shown excellent performance for pose prediction and good results for virtual screening.



[1] Francesca Stanzione, Ilenia Giangreco, Jason C. Cole, Chapter Four - Use of molecular docking computational tools in drug discovery, Editor(s): David R. Witty, Brian Cox, Progress in Medicinal Chemistry, Elsevier, Volume 60, 2021, Pages 273-343.

Learning Outcomes

In this workshop, you will carry out a virtual screening of a library of compounds. We will only use a small number of compounds for demonstration purposes. Your library would likely be much larger in a real virtual screening. After this workshop, you will:

- Know what a virtual screening is and understand the different processes involved.
- Be able to navigate Hermes, a 3D visualizer for proteins and the interface to GOLD.
- Be able to edit a protein-ligand complex from the PDB in Hermes to isolate the binding site.
- Know how to set up a virtual screening simulation in GOLD using the wizard in Hermes.
- Run a virtual screening and analyse virtual screening results.
- Be aware of other tools that can be deployed in virtual screening processes and the benefits of a data-driven approach.

The *Extension Exercise* at the end will show you how to calculate SuperStar interaction maps and evaluate virtual screening results using these.

This workshop will take approximately ***20-25 minutes*** to complete. The words in *Blue Italic* in the text are reported in the Glossary at the end of this handout.

Pre-required Skills

Familiarity with the Hermes interface and basic knowledge of docking with GOLD would be advantageous. For further guidance on protein-ligand docking see the "Introduction to Protein-Ligand Docking with GOLD" self-guided workshop available on this webpage or try the Protein-ligand docking 101 - running a simulation in GOLD CSDU module. The basics visualisation options for Hermes are at the end of this handout.

Materials

For this workshop we will use the files that you can download from here.

Example: Screening a Library of KRAS inhibitors within the Binding Pocket of MRTX-1133

Introduction

KRAS^{G12D} is the most frequent KRAS oncogenic mutation in solid tumors and a key therapeutic target.[2,3] Unlike **KRAS^{G12C}**, which can be inhibited covalently via a reactive cysteine, KRAS^{G12D} lacks such a residue, making selective inhibition more difficult.

Historically, KRAS has been a challenging target due to its high affinity for GDP/GTP and lack of accessible binding pockets. Previous attempts to identify hits with shallow pocket binders and cyclic peptides showed limited cellular activity due to poor physicochemical properties. MRTX1133[4] (Figure 1A) overcomes these limitations with picomolar binding affinity and nanomolar cellular potency.

The discovery process of this inhibitor began with analogues of a pyrido[4,3-d]pyrimidine scaffold. Structural analysis revealed key interactions with **Asp12**, **Arg68**, and **Glu62**, and highlighted the importance of the scaffold's substituents. The protonated piperazinyl group formed a salt bridge with **Asp12**, enhancing selectivity. Additional interactions with **Gly60**, **Gly10**, **Thr58**, and a hydrophobic pocket contributed to potency.

These insights guided the development of MRTX1133, demonstrating that noncovalent inhibition of KRAS^{G12D} is achievable and therapeutically promising.

Your protein: KRAS [PDB 7RPZ] (Figure 1B)

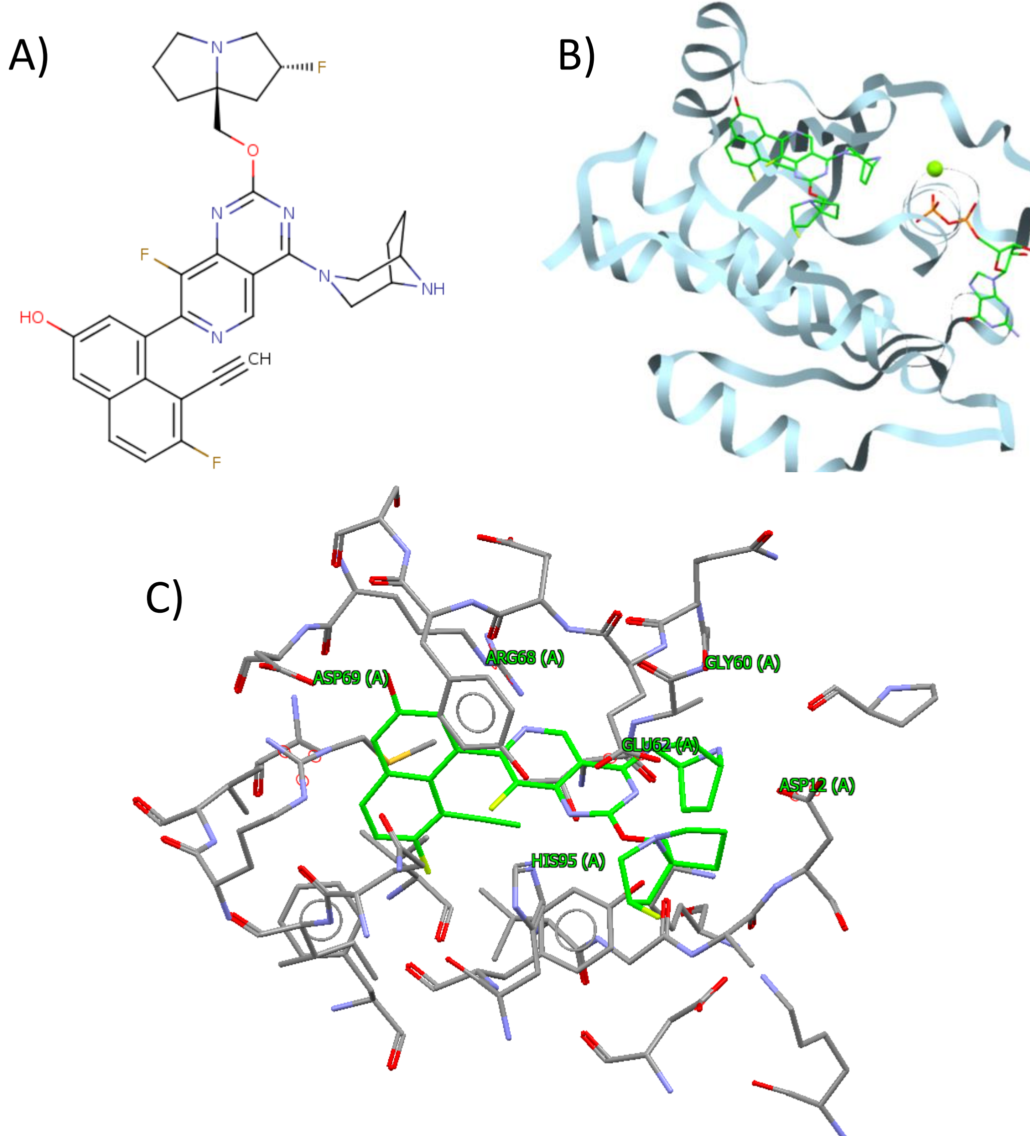


Figure 1. A) Structure of MRTX-1133. B) KRAS protein co-crystallized with inhibitor MRTX-1133. A magnesium ion, and cofactor Guanosine-5'-diphosphate are also shown. C) Inhibitor-pocket interactions with key residues.

---

[2] Identification of MRTX1133, a Noncovalent, Potent, and Selective KRASG12D Inhibitor; J. Med. Chem. 2022 65 (4), 3123-3133; https://pubs.acs.org/doi/10.1021/acs.jmedchem.1c01688

[3] RCSB PDB - 7RPZ: KRAS G12D in complex with MRTX-1133

[4] 4-(4-(3,8-Diazabicyclo[3.2.1]octan-3-yl)-8-fluoro-2-(((2R,7aS)-2-fluorohexahydro-1H-pyrrolizin-7a-yl)methoxy)pyrido[4,3-d]pyrimidin-7-yl)-5-ethynyl-6-fluoronaphthalen-2-ol | C33H31F3N6O2 | CID 156124857 - PubChem

Your ligand: 4-(4-[(1R,5S)-3,8-diazabicyclo[3.2.1]octan-3-yl]-8-fluoro-2-{[(2R,4R,7aS)-2-fluorotetrahydro-1H-pyrrolizin-7a(5H)-yl]methoxy}pyrido[4,3-d]pyrimidin-7-yl)-5-ethynyl-6-fluoronaphthalen-2-ol, an experimental inhibitor that interacts with pocket residues Asp12, Arg68, and Glu62 of KRAS protein (PDB 7RPZ) as shown in the cocrystal structure (Figure 1B and 1C).

The discovery of novel drug-like hit candidates remains one of the most significant challenges in modern drug development, despite the rapid advancement of sophisticated computational methodologies. Over the past decade, a common strategy during the early phases of drug discovery has involved the virtual screening of ultra-large compound libraries using *high-throughput screening* techniques. This approach aims to identify chemical entities capable of modulating the biological target of interest.

In addition to high-throughput screening, several *de novo drug design* strategies have gained prominence. These include fragment-based drug design (FBDD), which encompasses techniques such as fragment growth, fragment merging, and scaffold replacement. Furthermore, artificial intelligence (AI)-driven methods—particularly those employing generative algorithms—have emerged as powerful tools for designing ligands within target binding pockets. These algorithms incorporate structural and physicochemical biases to optimize protein-ligand interactions, thereby enhancing the likelihood of identifying viable therapeutic candidates.

Despite the significant advancements in artificial intelligence technologies in the last years, recent comparative studies have demonstrated that classical physics-based approaches continue to outperform AI-driven methods in certain key areas of structure-based drug design. Notably, traditional virtual screening techniques—such as molecular docking using tools like GOLD—still exhibit superior performance in pose prediction accuracy when compared to current AI-based methodologies.[5]

[5] PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. arXiv 2023. https://doi.org/10.48550/arXiv.2308.05777

Challenges

In this workshop, as an example, we will screen a small library of KRAS inhibitors within the binding pocket of **MRTX-1133** (Figure 1) by using GOLD docking.

After running the virtual screening campaign, we will visualize the docking solutions to identify which KRAS inhibitor/s could retain the same interactions as the **MRTX-1133** ligand and evaluate potential scaffold replacement.

Furthermore, we will validate docking solutions using SuperStar to assess which novel interactions within the binding pocket could be favourable. Superstar calculates binding site interactions as propensity maps by mining the PDB or the CSD databases and can be used to validate docking poses when kinase inhibitors fit the propensity maps.

Provided Input Files

Please download and unzip the input files from this link. In the folder you will find:

- **7RPZ.pdb** & **7RPZ.cif** files: the crystal structure of the KRAS protein co-crystallized with inhibitor MRTX-1133: 4-(4-[(1R,5S)-3,8-diazabicyclo[3.2.1]octan-3-yl]-8-fluoro-2-{[(2R,4R,7aS)-2-fluorotetrahydro-1H-pyrrolizin-7a(5H)-yl]methoxy}pyrido[4,3-d]pyrimidin-7-yl)-5-ethynyl-6-fluoronaphthalen-2-ol. You can use this protein to walk through the key steps required to prepare a protein for virtual screening docking or you can follow the instructions illustrated in the first step of this workshop. Note: The legacy pdb file format is now deprecated, and it has been replaced by the new cif format. Hermes and GOLD can read both.

- The ligands file (**KRAS_inhibitors_lite.mol2**). The docking library constituted by a small set of KRAS inhibitors selected for this workshop together with the ligand MRTX-1133 co-crystalised with target 7RPZ. Note: this has been set up in accordance with the guidelines for the preparation of input files that you can find in the GOLD user guide (Setting Up the Protein(s) and Setting Up Ligands chapters). **Important:** use this file in **today's hands-on session**.

- The ligands file (**KRAS_inhibitors.mol2**). The docking library constituted by a set of KRAS inhibitors downloaded from the ChEMBL 35 website[6] and the ligand MRTX-1133 co-crystalised with target 7RPZ. Note: this has been set up in accordance with the guidelines for the preparation of input files that you can find in the GOLD user guide (Setting Up the Protein(s) and Setting Up Ligands chapters). **Important:** use this file for the **extension exercise** if you wish.

Additional Files

- **binding_site.mol2** – the extracted binding site.

- **ligand.mol2** – the inhibitor MRTX-1133 extracted from 7RPZ X-ray structure.

- **Demo_lite** – folder containing input and output files from the virtual screening run and SuperStar analysis **for today's session.**

- **gold_KRAS_lite.conf** – GOLD configuration file (also located in the Demo_lite folder)

- **Demo** – folder containing input and output files from the virtual screening run and SuperStar analysis for the extension execise

Note that if you skip the GOLD docking run, to load the configuration file and view the results, you will need to edit the path to the files so that Hermes can load the necessary files for visualization. Open the configuration file in a text editor then edit the following paths and filenames:

5. In the FLOOD FILL section, update the cavity_file path to point to the folder where you have the files saved.

6. In the DATA FILES section, update the ligand_data_file to the folder where you have the files saved.

7. In the DATA FILES section, update the directory to the folder where you have the files saved.

8. In the PROTEIN DATA section, update the protein_datafile to point to the folder where you have the files saved.

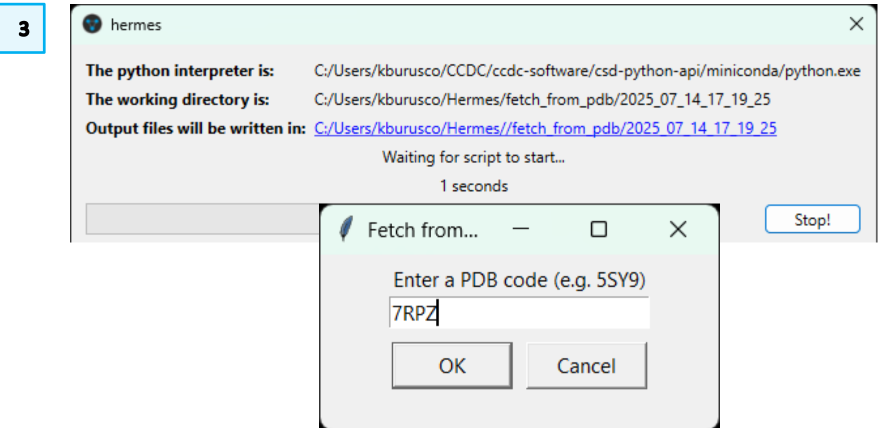
[6] https://www.ebi.ac.uk/chembl/

Importing the Protein-Ligand Complex from the PDB

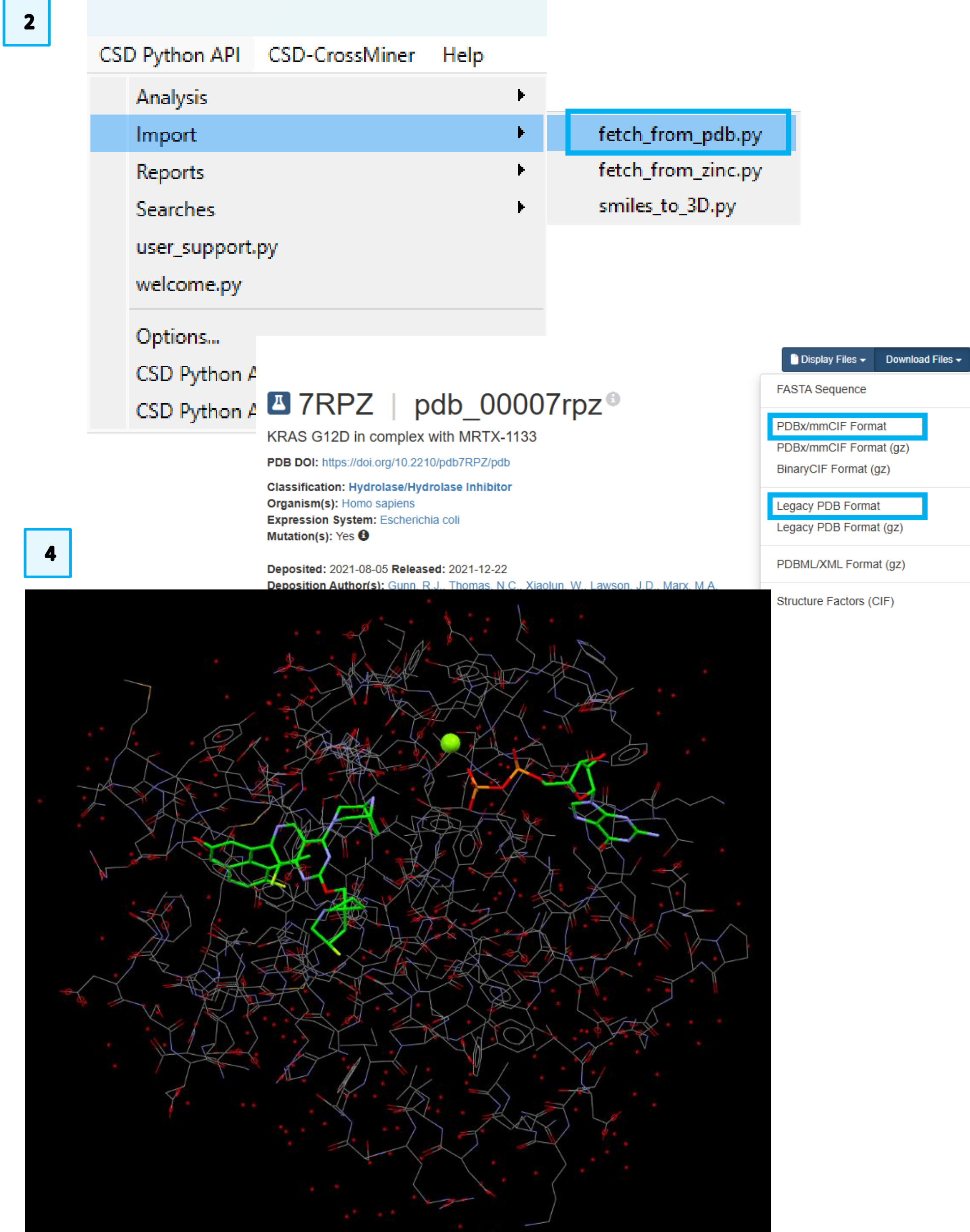
1. Launch Hermes by clicking its icon

2. We are going to import the file of our protein and ligand directly from the PDB using the CSD Python API functionality. From the menus at the top, click on *CSD Python API*, then go to *Import*, and choose *fetch_from_pdb.py*. (If the CSD Python API is not installed, the cif/pdb file is also included in this **zipped folder** or can be downloaded directly from the pdb website. Load the file using *File > Open*.)

3. This will bring up 2 windows: one displaying a loading bar, which show the script is running, and a smaller one, where we will insert the PDB code. For this case study we will type: **7RPZ**. Then click **OK**. It will take a few seconds to load the selected complex in Hermes. Once the structure is loaded, close the script window.



4. To better visualise the structure, you can zoom out using the **zoom-** button above the visualizer or right-clicking while moving the mouse down. The ligands are shown with green carbon atoms by default, the proteins with grey carbons, while the red dots around the structure are oxygen atoms of water molecules.



Editing the Structure

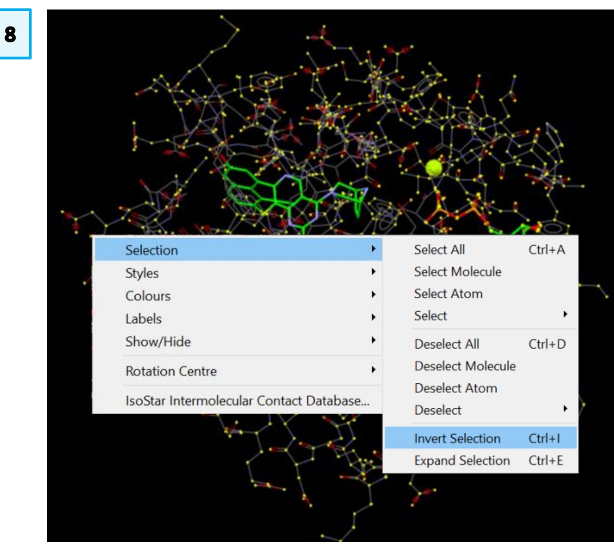
In this virtual screening exercise, we wish to work only with the ligand binding site so we will remove unnecessary components of the structure and create a new structure file just containing the ligand and its binding site.

5. In the *Molecule Explorer* click on the arrow next to **7RPZ**, then next to Ligands to expand and access all the components. If you do not have the Molecule Explorer displayed, see the instructions at the end of this document. You can try ticking/unticking each component to see that they will appear/disappear in the visualizer.

6. For this example, there are no structural waters that need to be included, so we can delete them all. You do not need to expand the entry for *Waters* and delete the molecules on-by-one, but you can **right-click** directly on **Waters** and select **Delete**.

7. **Right click** on **C:6IC202** and choose *Select Ligand Binding Site*. This selects all the atoms in the ligand and in the portion of the protein to which it binds. The selected atoms are highlighted in yellow

8. **Right click** in the Display window and from the menu pick *Selection > Invert selection*. Alternatively, you can try **Ctrl + I**.

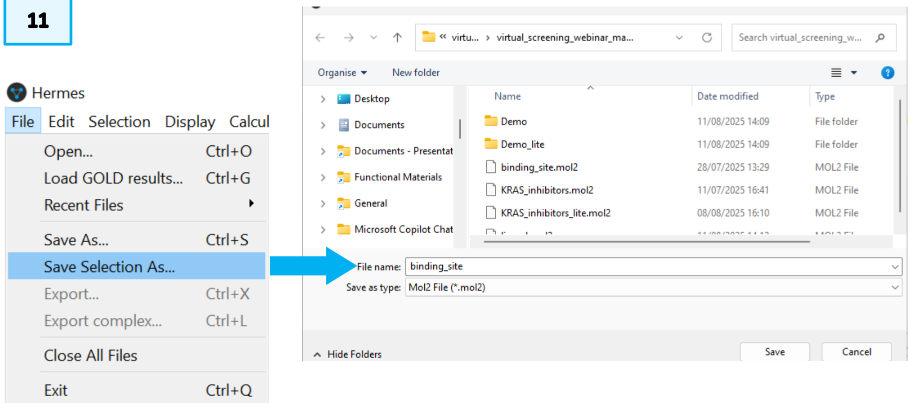


9. **Right click** and select *Show/Hide* > Hide atoms.

10. Using the **Lasso** tool select all of the atoms shown in the **Display** window.

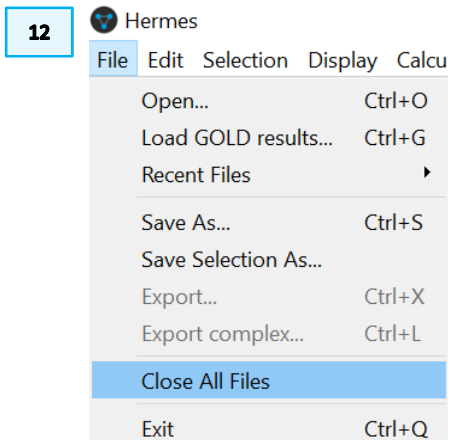
11. From the top-level menus select *File* > *Save Selection As* to save the file in an appropriate location that will be your working directory for this workshop. Type "binding_site" as File name, and ensure the Save as type is set to **Mol2 File (*.mol2)**.

11

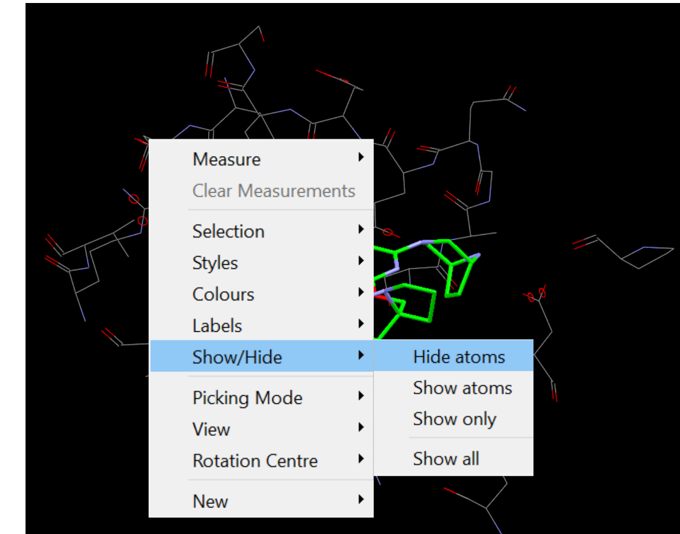


12. Close all files by selecting *File* > *Close All Files* from the top-level menus.

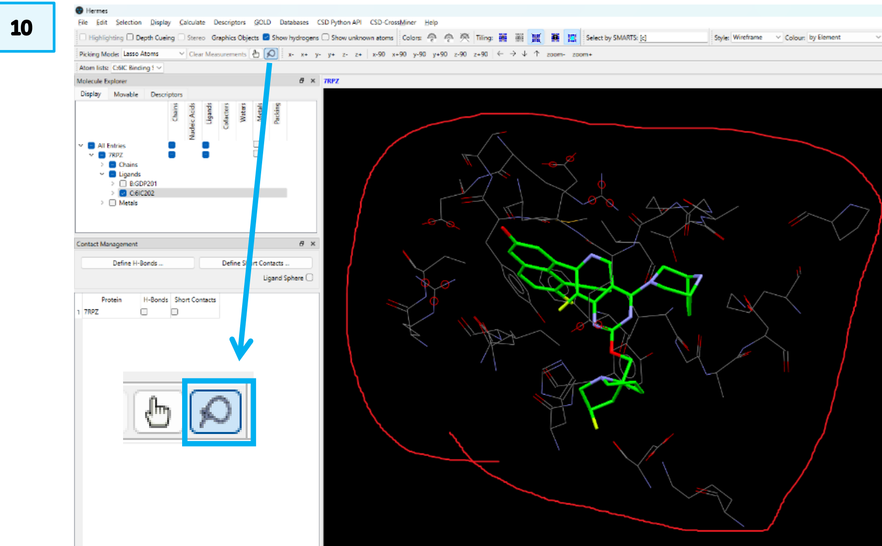
12



9



10



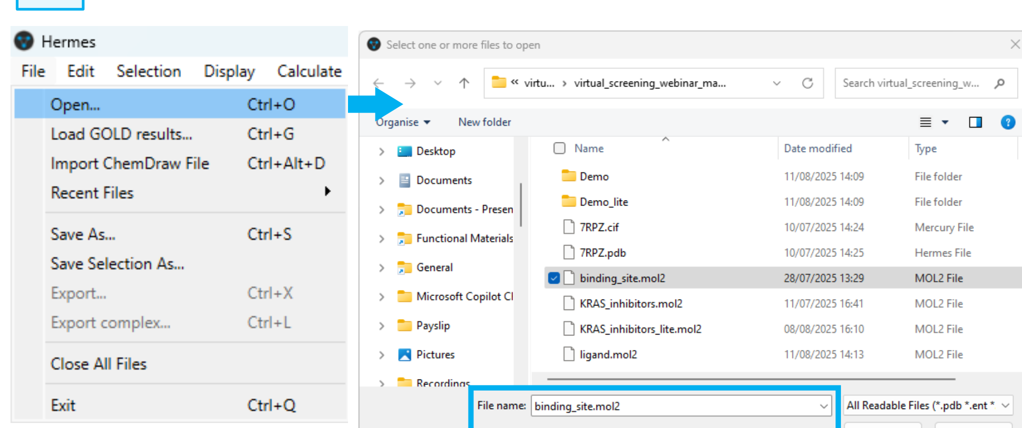
Exploring the Binding Site

13. Reload the binding site by selecting *File > Open* and navigating to the location of “binding_site.mol2” created in **Step 11**. Select this file and open it. You may also select this file from the bundle of files supplied with this workshop, if you prefer.

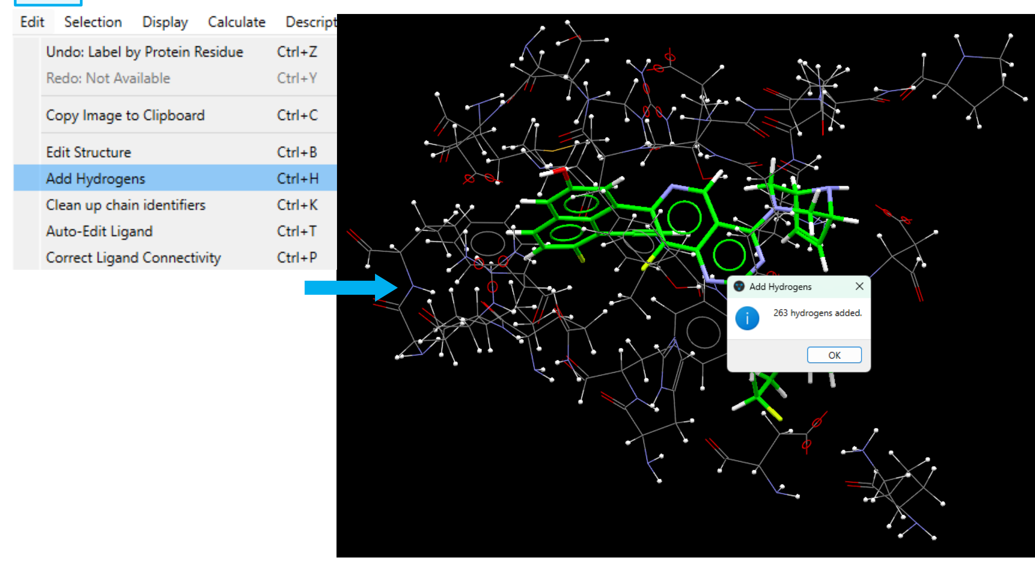
14. From the top-level menus select *Edit > Add Hydrogens*. This ensures that the fragment of the protein used for screening has the appropriate number of hydrogen atoms. The *Add Hydrogens* dialogue box should report “263 hydrogens added”. Click **OK** to continue.

15. You can customise the style of the various components to make them easier to visualize. From the top-level menus, select *Display > Style Preferences*. From the Style Preferences dialogue select **Stick** from the drop-down menu next to Ligand. The ligand will now be displayed in capped-stick style. ***Tip:*** *you can adjust the thickness of the sticks by Right Clicking then selecting Styles > Capped Stick Settings… and adjusting the Capped Stick Radius.*

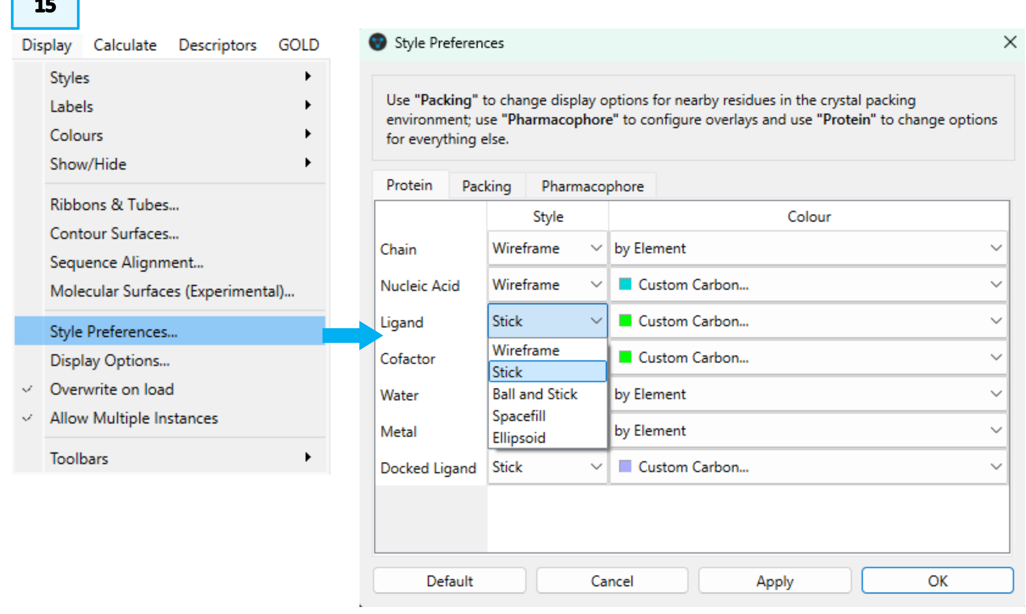
13



14



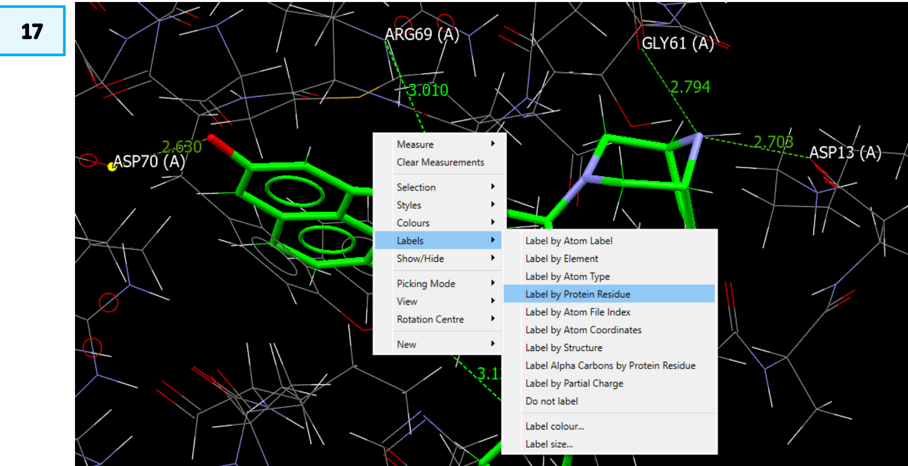
15



16. To visualize the hydrogen bonds made between the ligand and protein residues, check the **H-bonds** tick box in the *Contact Management window*.

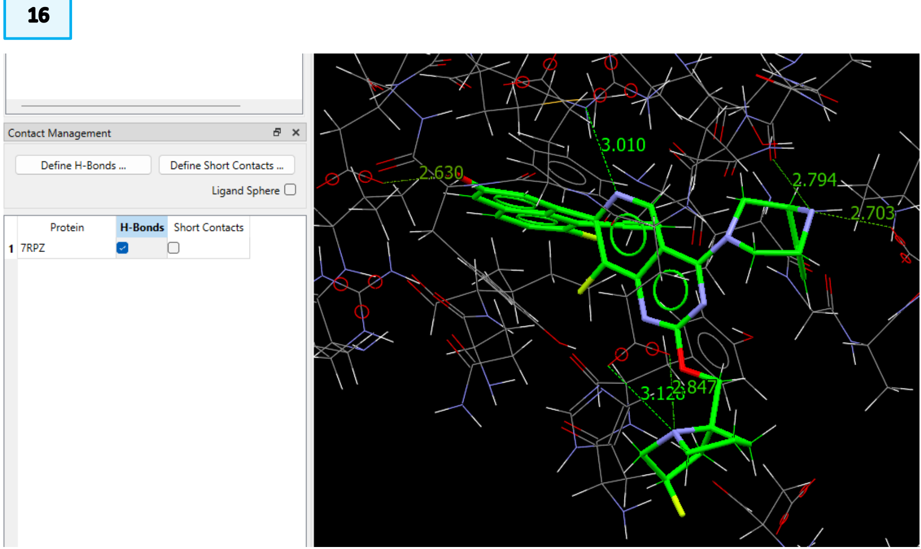
17. In order to highlight individual interactions, pick the donor/acceptor atoms on the protein residues which are indicated as forming a hydrogen bond by the green dashed line. Then, **right click** and *select Labels > Label by Protein Residue*. Labels **"ASP70", "ARG69", "GLY61**", **"GLU63",** and **"ASP13"** should now appear next to the corresponding residue. ***Tip:*** *you can change the size of the labels by right clicking, then selecting Labels > Label Size… and entering the desired font size in the dialogue box.*

17

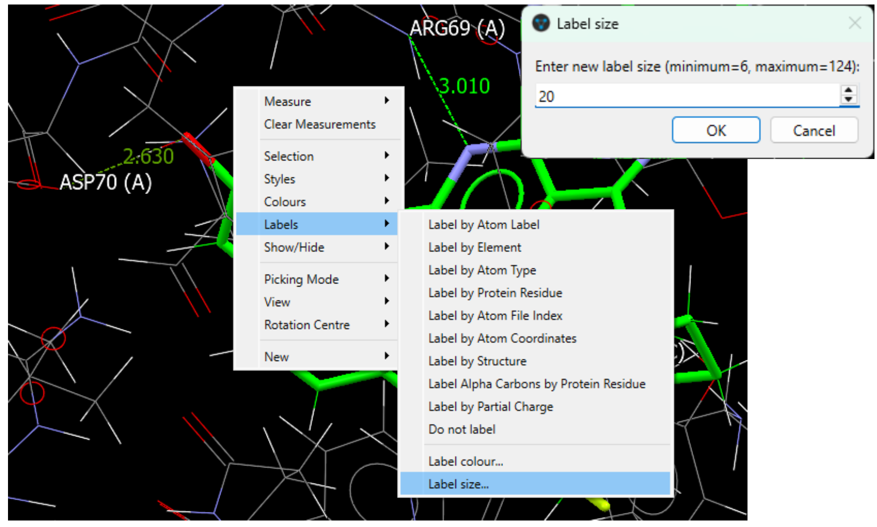


18. Visually inspect the hydrogen bond interactions. In the [extension exercise](), you will be able to explore this further by calculating **SuperStar** interaction maps. From the *File* menu **Save** the curated pocket (e.g. as binding_pocket_curated.mol2) and **Close All files**, as in Steps 11-12.

16



Tip:

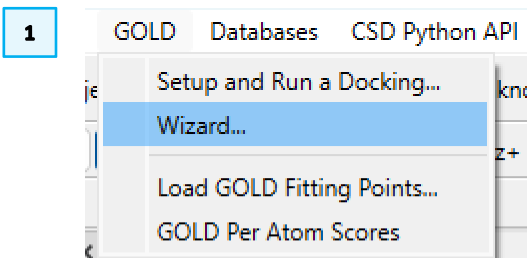


Set up the GOLD Virtual Screening

First, we are going to import again the file of our protein and ligand. For this we can either repeat the instructions in **Step 1** in the previous section or simply open the file **7RPZ.cif** (or **7RPZ.pdb**) provided with the bundle of files supplied with this workshop.

1. Ensure that all files from the previous section are closed. To start our docking simulation, click GOLD from the top menus, then pick Wizard...

1

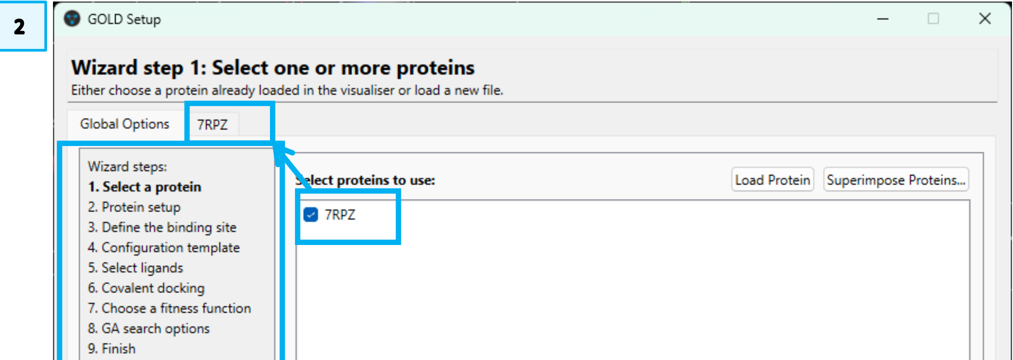


from the resultant pull-down menu.

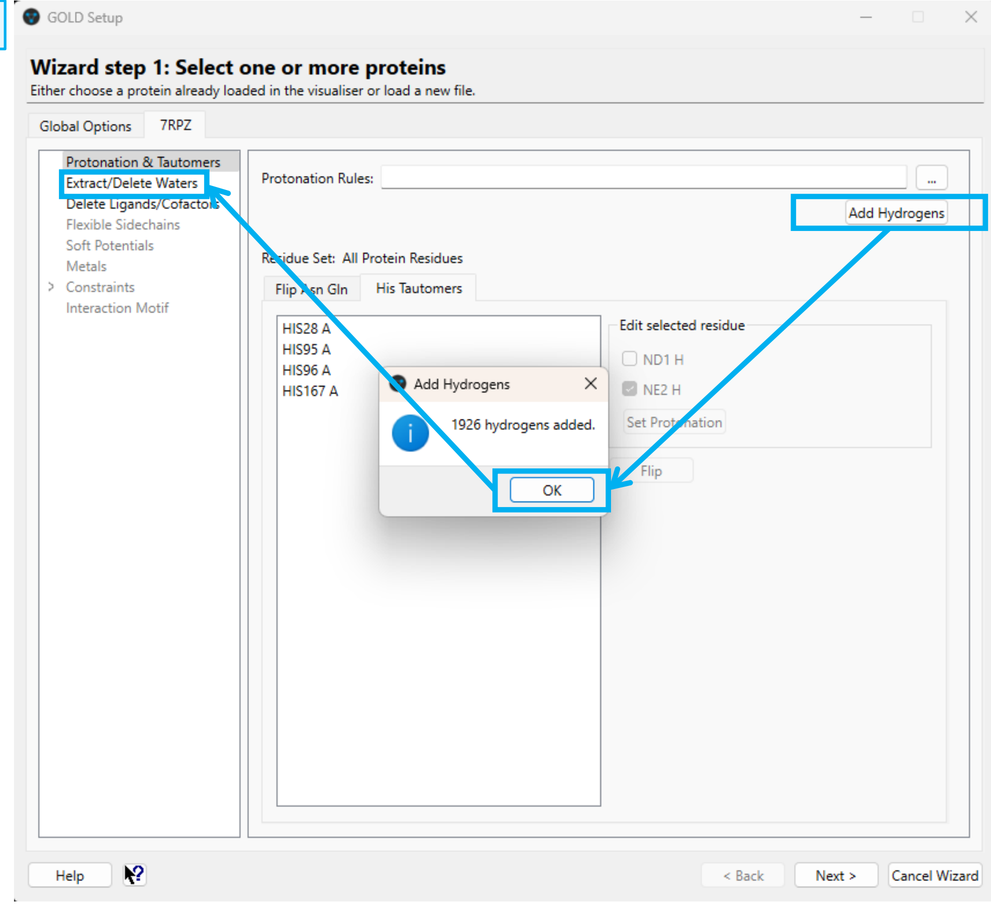
2. In the *GOLD Setup window*, we see that we are at the **Wizard step 1: Select one or more proteins**. To select our protein, in the **Global Options** tab tick on the box next to the protein code, **7RPZ**. This will bring up a new tab labelled with the name of the protein file, **7RPZ**.

3. Click on the **7RPZ** tab. This panel contains different options to clean and prepare the protein for the docking calculation. From the top-right click *Add Hydrogens*. This ensures that the protein used for screening has the appropriate number of hydrogen atoms. The Add Hydrogens dialogue box should report "1926 hydrogens added". Click OK to continue and press *Extract / Delete Waters* on the top of the left-hand panel to move to the next section.

2

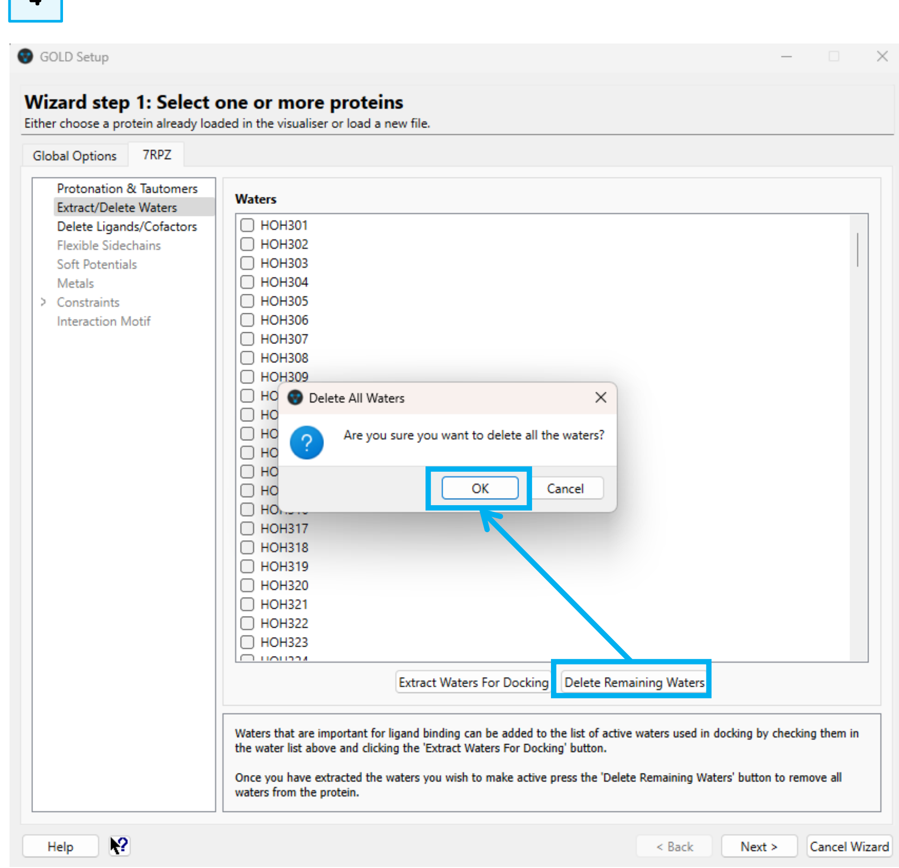


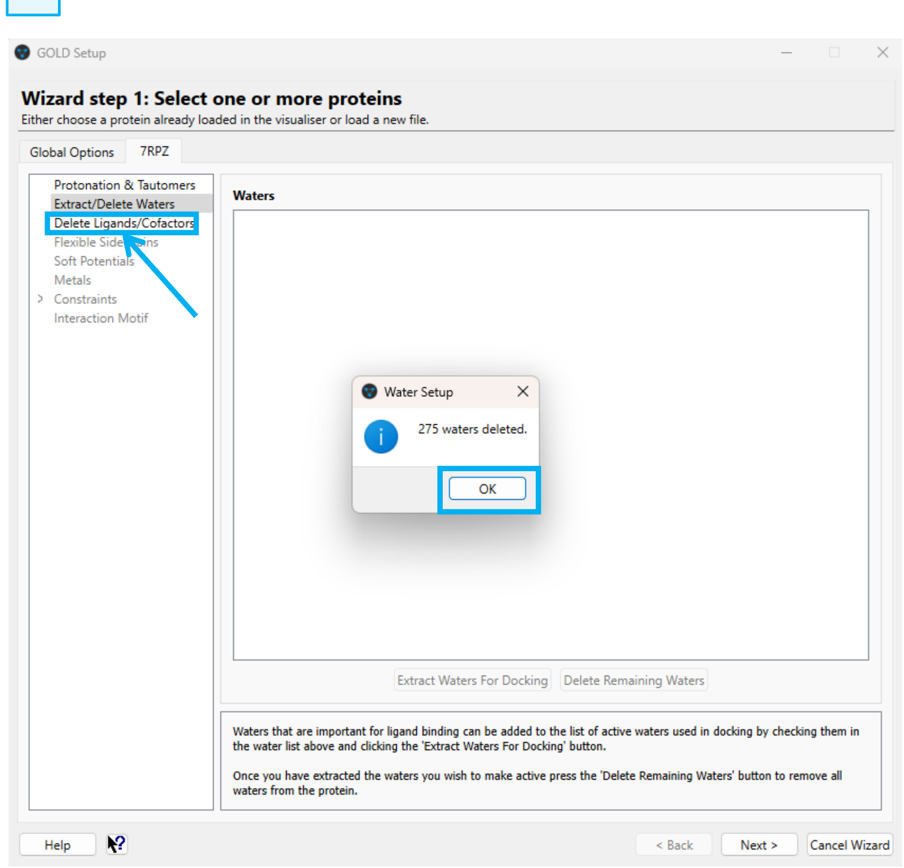
3



4. As we mentioned before, for this example there are no structural waters that need to be included, so we can delete them all. Click on **Delete Remaining Waters**, and then click **OK** in the pop-up window asking for confirmation.

5. Another pop-up window will display that "275 water molecules have been deleted". Press **OK** and click on *Delete Ligands/Cofactors* from the menu on the left-hand side to move to the next section.



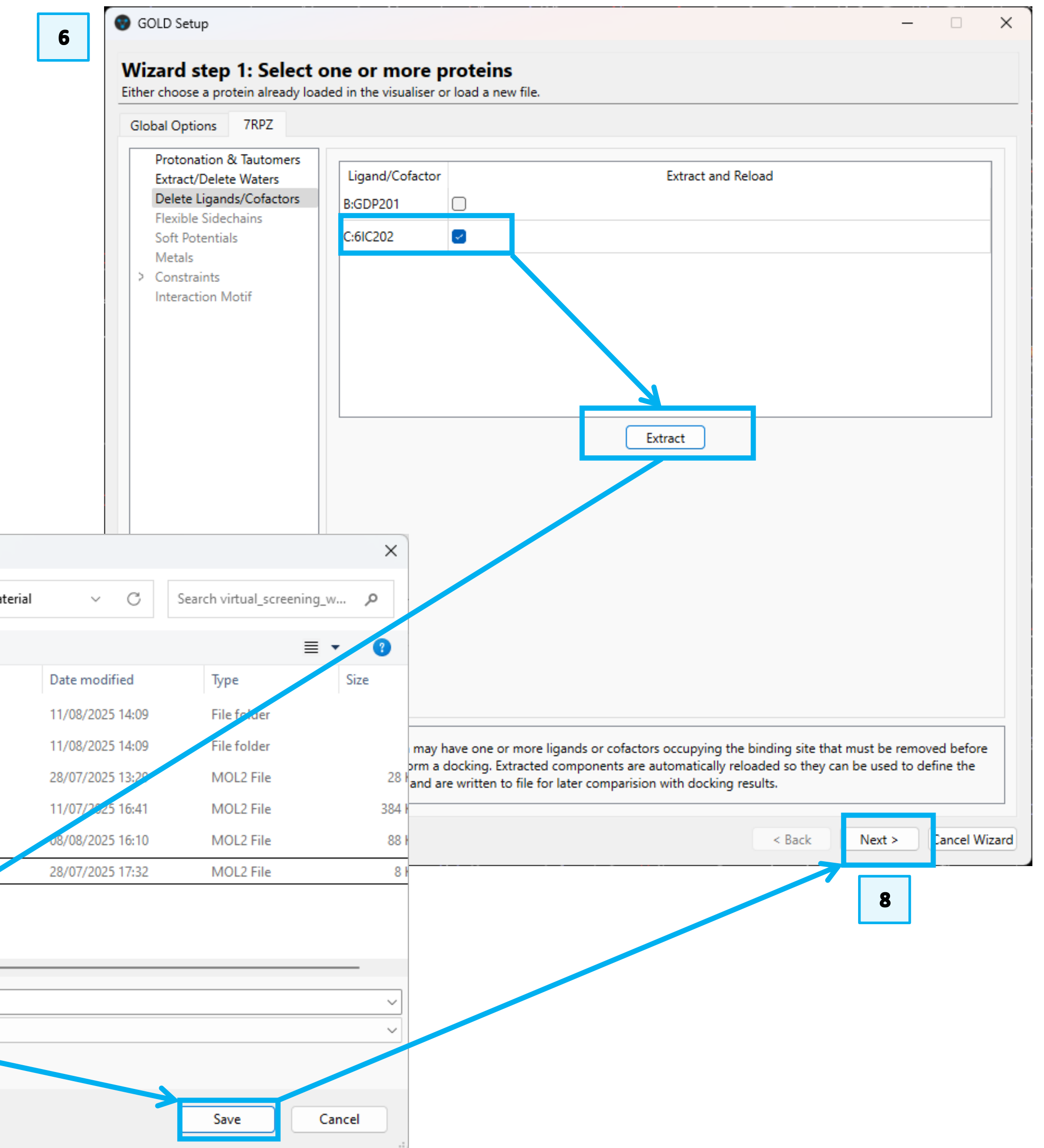


The protein **7RPZ** has been co-crystallised with a **GDP** cofactor (B:GDP201), a **Mg** metal ion, and the inhibitor **MRTX-1133** (C:6IC202). Both cofactor and metal ion are located in a separated area from the binding pocket so they will not interfere in the docking calculation, therefore, we will extract the inhibitor from the binding site to prepare the protein and leave the cofactor and ion untouched.

6. Tick the box next to **C:6IC202** and click **Extract**.

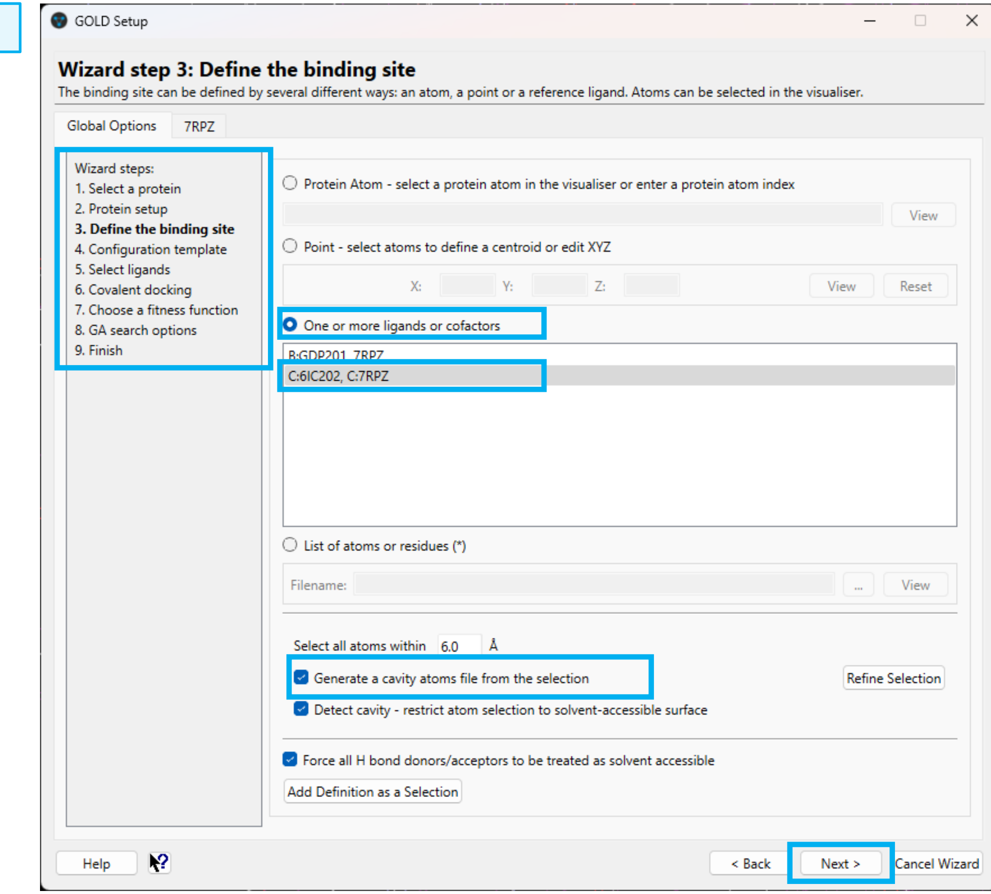
7. When prompted, save the ligand file as **ligand.mol2**, click **Save**.

8. Click **Next** twice to move on.

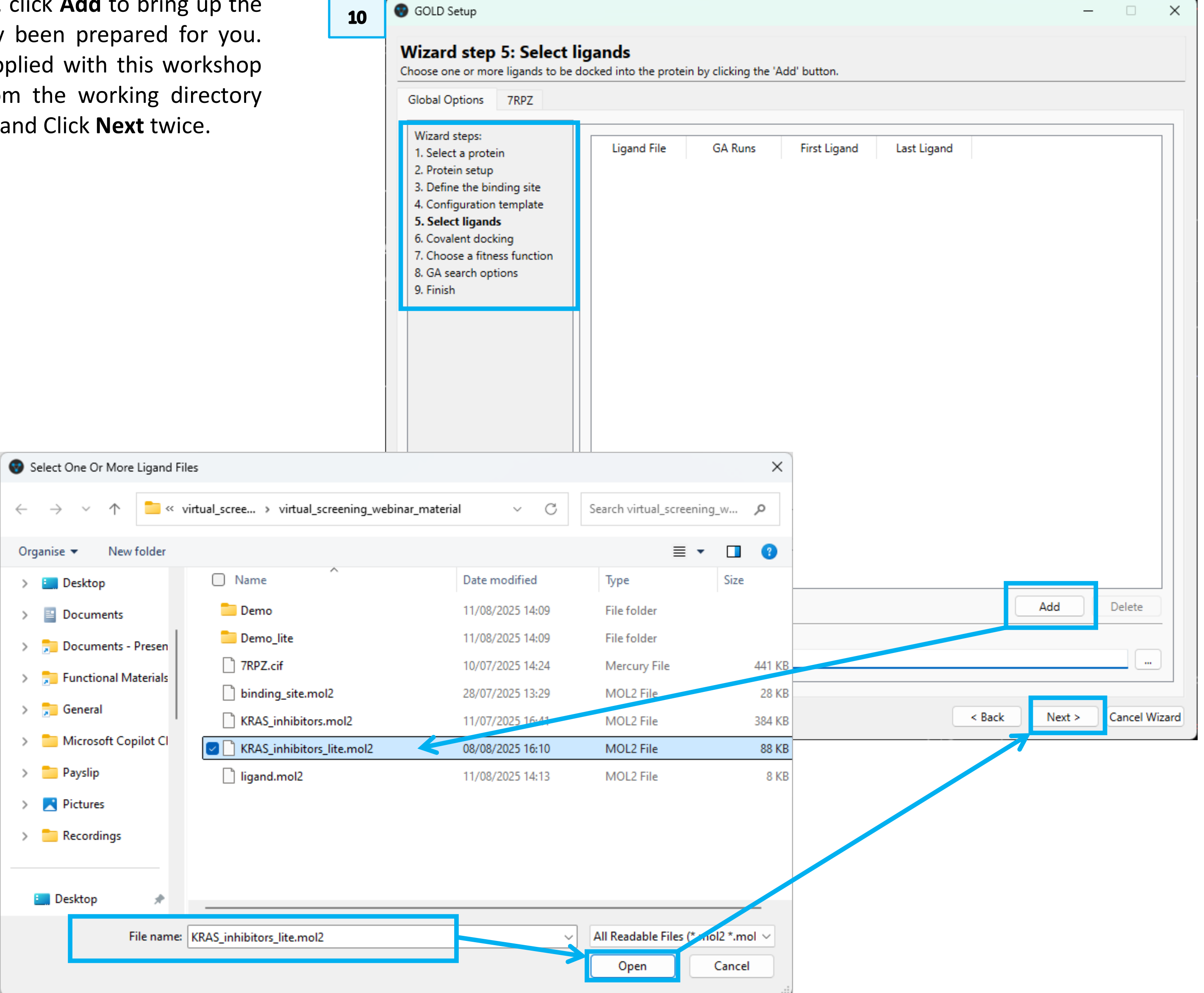


9. Return to the Global Options tab to **Define the binding site** (**Wizard step 3**). Click the radio button next to One or more ligands or cofactors. Make sure the ligand option **“C:6IC202: C:7RPZ”** is selected in the box. Tick the box **Generate a cavity atoms file from the selection**. Click **Next** twice to move to **Select ligands** (**Wizard Step 5**).

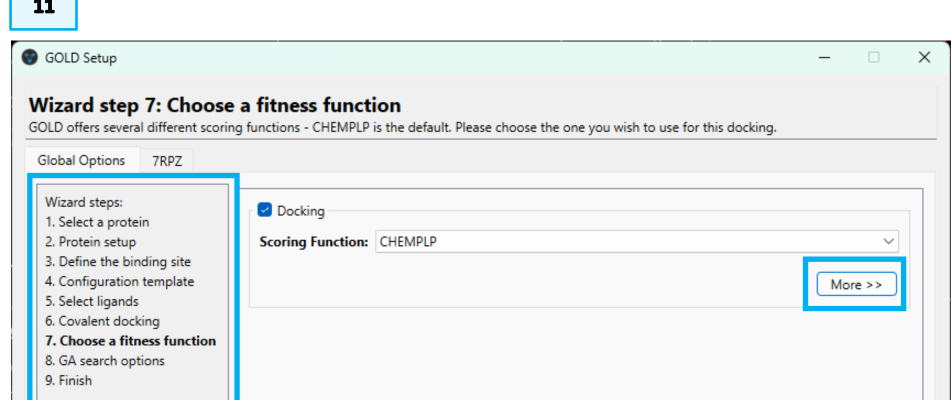
9

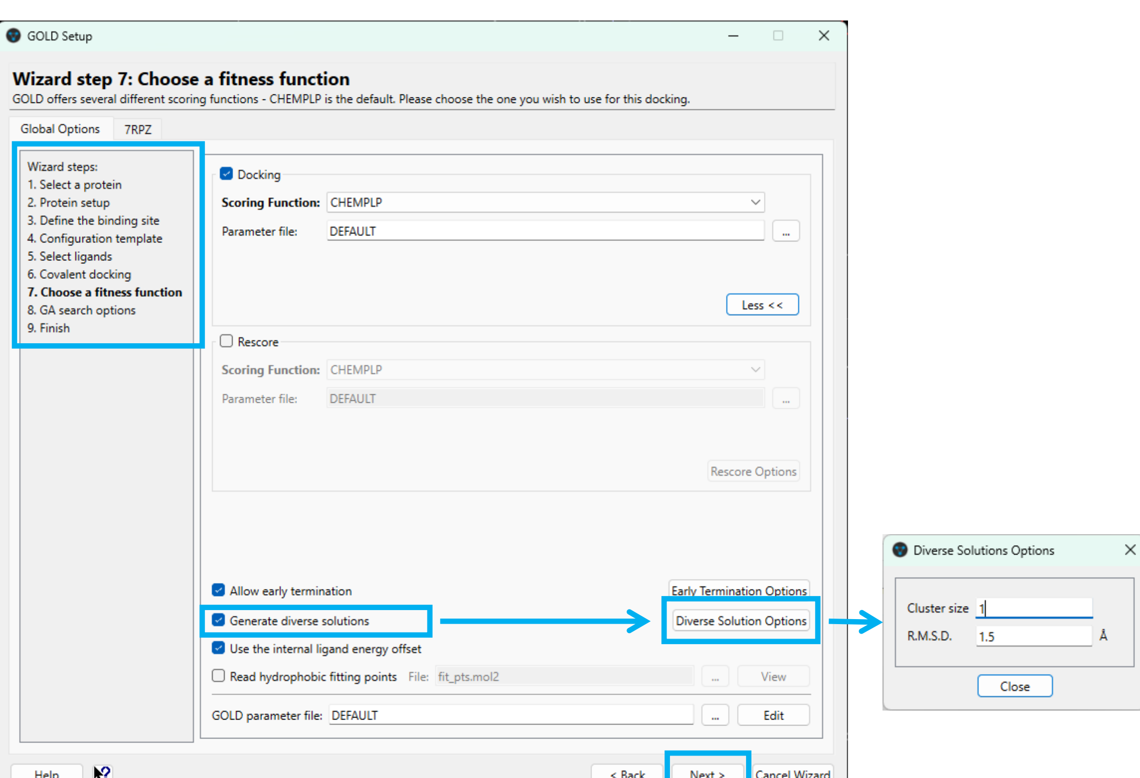


10. In **Select ligands**, near the bottom of the window, click **Add** to bring up the File Explorer. The library of ligands have already been prepared for you. Navigate to the directory containing the files supplied with this workshop (Input) and select **KRAS_inhibitors_lite.mol2** from the working directory where you unzipped the files supplied. Click **Open** and Click **Next** twice.

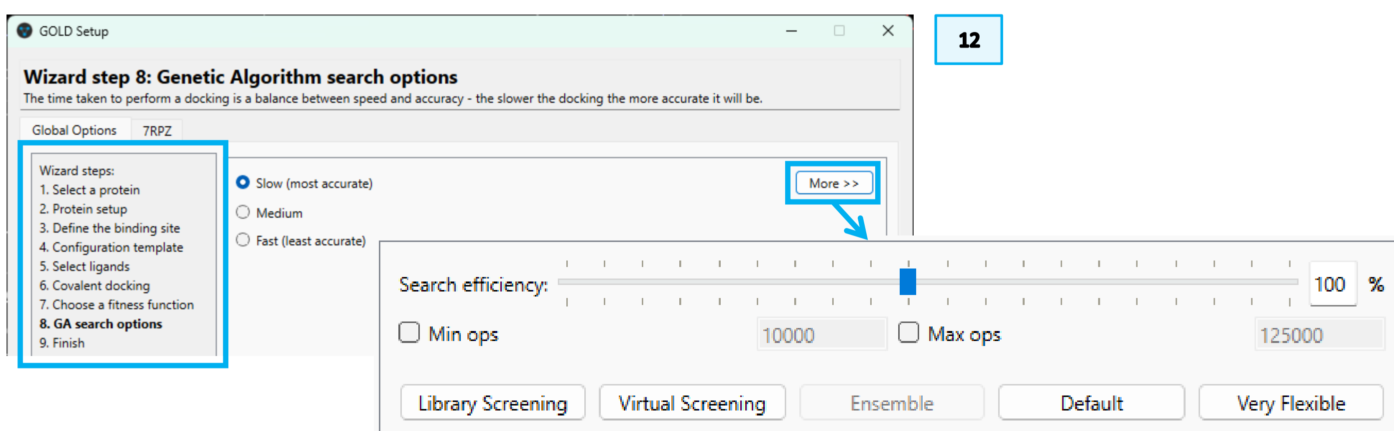


11. **Fitness function (Wizard step 7).** Leave the default scoring function as "CHEMPLP" and click *More >>.* Tick the box next to **Generate diverse solutions**, click on **Diverse Solution Options**. You can adjust these values as you wish but we will use the default settings. Click **Close**.





12. Click **Next** to proceed to the **Genetic Algorithm search options** **(Wizard step 8)** and leave the default option (Slow – most accurate) selected. You can customise this by click on *More >>. **Tip:** for larger ligands libraries, you could consider different settings, such as the Virtual Screening preset, which is faster, with less search efficiency.*

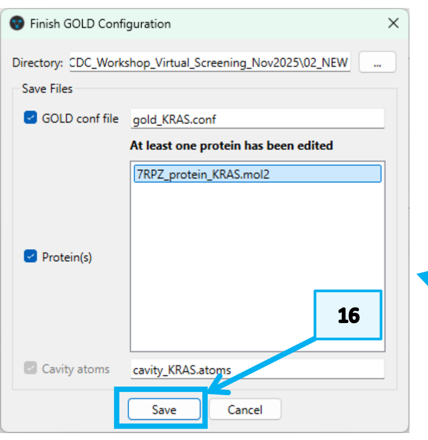


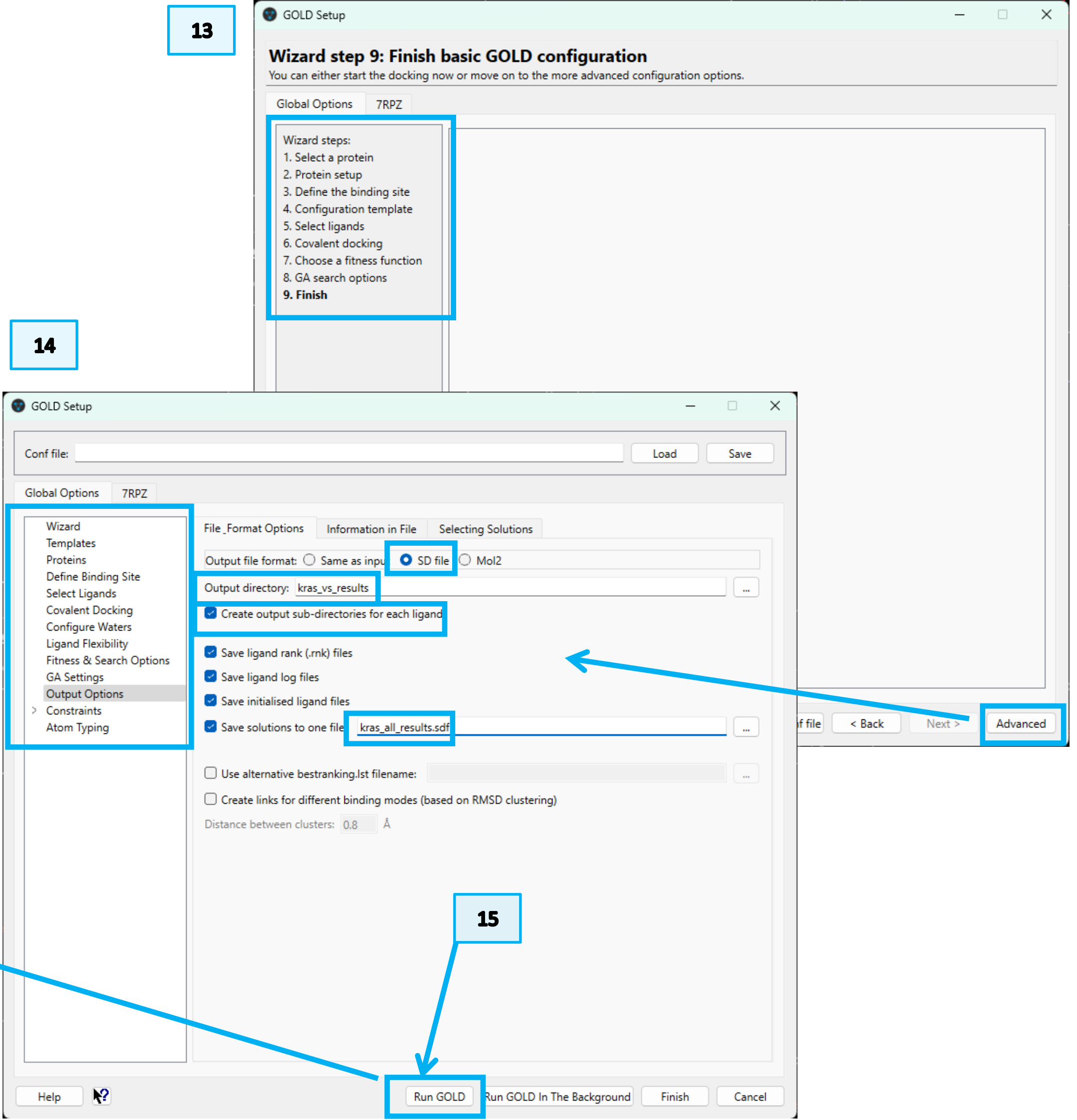
13. Click **Next** to move to **Finish (Wizard step 9)** but **DO NOT click on Run Gold yet**. Select **Advanced** at the bottom right-hand corner of the Wizard window.

14. Select the *Output Options* tab on the left panel and click the radio button **SD file**. In the field *Output Directory*, write the name of a folder to store all the output files (i.e. **results**). Tick the box *Create output subdirectories for each ligand*, to have the output files per ligand organised in individual folders. Tick the box next to **Save solutions to one file** and give it a name (i.e. **kras_all_results.sdf,** including the sdf extension) to collect all the results in a single file.

15. At the bottom of the **GOLD Setup window** click the **Run GOLD** button to start the GOLD run.

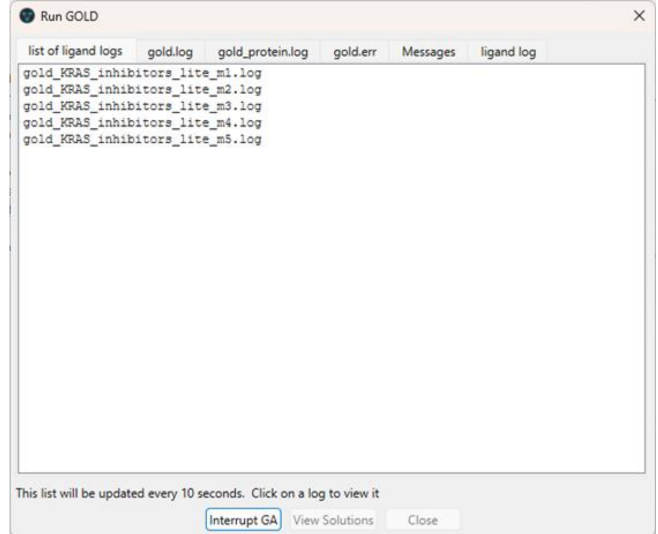
16. A pop-un window will appear, named Finish GOLD Configuration. You may choose to create a new folder to store the files by clicking on the three dots next to *Directory* or rename the files suggested. When you are satisfied with your selection, click **Save** to start the virtual screening calculation. It should take 3-4 minutes to complete.

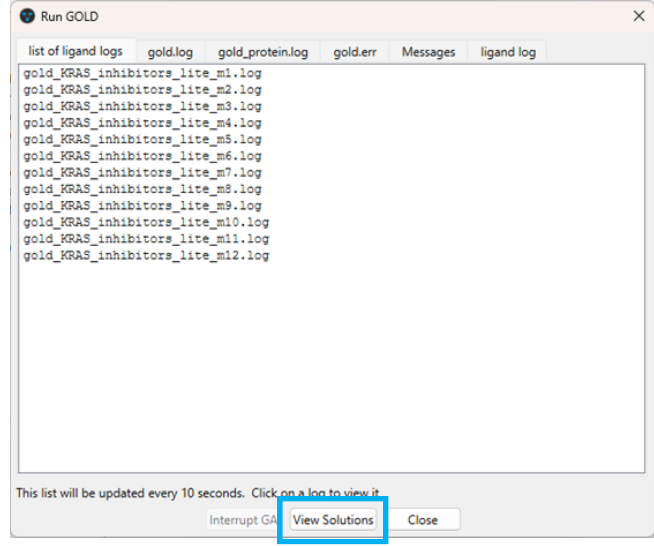


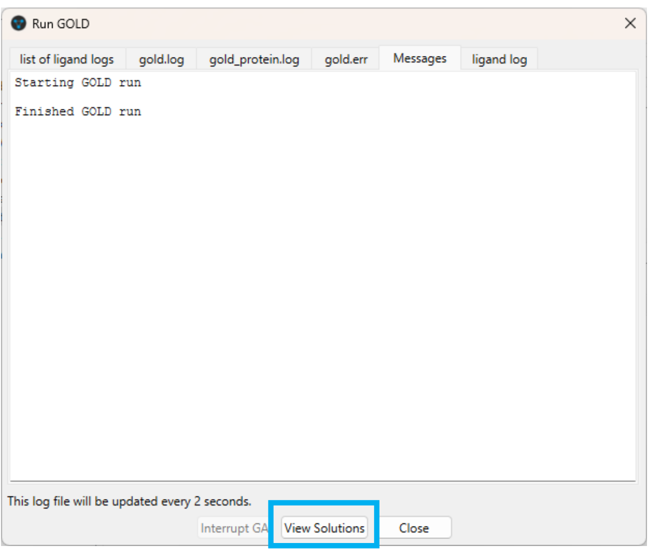


17. This will bring up a new window, Run GOLD, where you can follow the run, by going to the Messages tab. Once the run is completed, click **View Solutions** to visualize the results in Hermes. We can then click **Close** in the **Run GOLD** window and click **Cancel** in the **GOLD Setup window**. .

    Note: If you do not want to complete the GOLD run you can click Interrupt GA, close all files (*File > Close All Files*) and load the gold_KRAS.conf file (*File > Load GOLD results…)* supplied with this workshop in the **results** folder instead.[7] You may also need to re-open the reference ligand file.



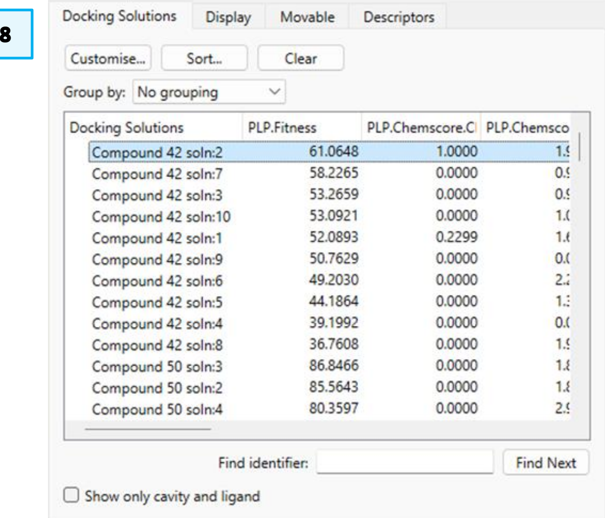




[7] In order to open the .conf file, you will need to edit the file with a text editor according to the instructions in the section Additional Files.

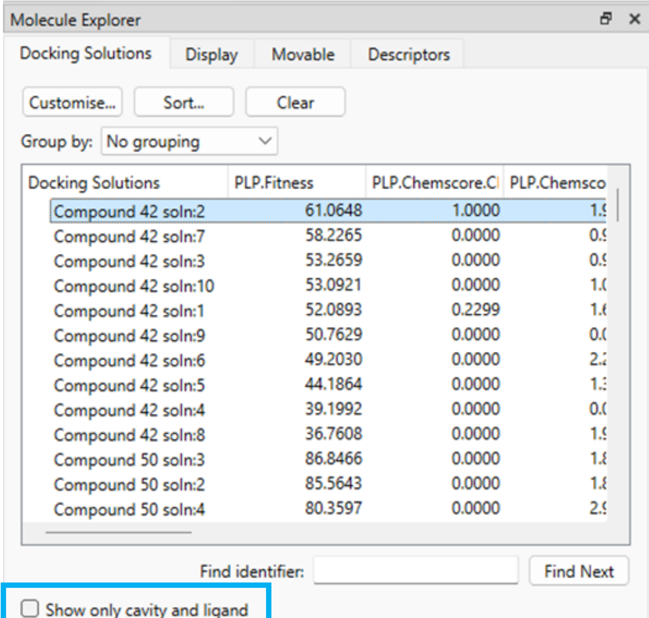
Visualization of and Analysis of the Results

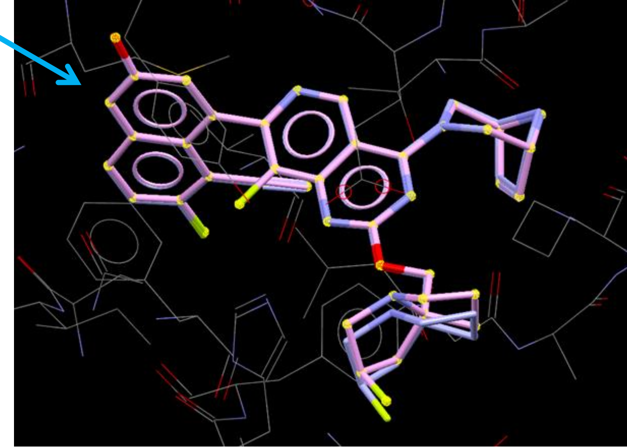
18. The solutions are listed on the left-hand side in the *Docking Solutions* tab of the *Molecule Explorer*. This tab becomes available only after the GOLD run has completed. **Do not click on any Docking Solution yet**



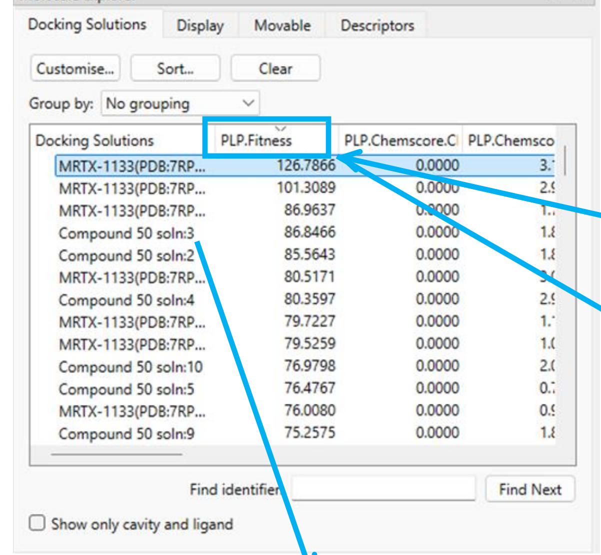
19. To make it easier to distinguish the docked ligands from the reference ligand you can change the colour of the reference ligand. We recommend changing the carbon atoms to a pink colour. To do this select the ligand (hold **Shift** while clicking on an atom of the ligand), **right click,** then select *Styles > Capped-sticks.* Again, **right-click**, select *Colour > Custom carbon* and choose the shade of pink (or whichever colour) which you prefer. On the menu bar, top-left, untick the box next to *Show Hydrogens*, to improve the pocket view.

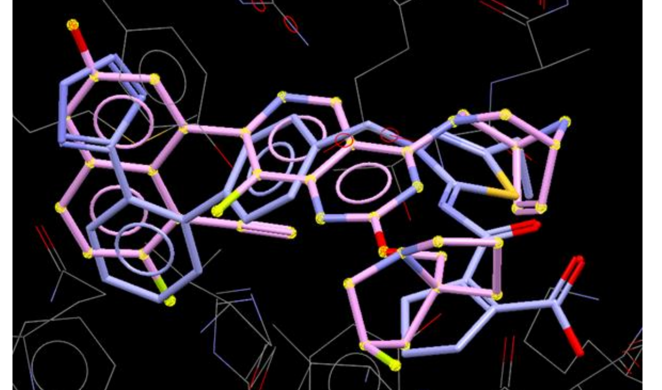
20. If you click on each entry in the *Molecule Explorer*, you will see it displayed in the visualizer in a blue color. The docked ligand will be overlaid with the (pink) original ligand. Note: GOLD is stochastic, so your results are likely to differ from those displayed in the pictures in this example. You can tick the **Show only cavity and ligand box** if you wish to simplify the display.





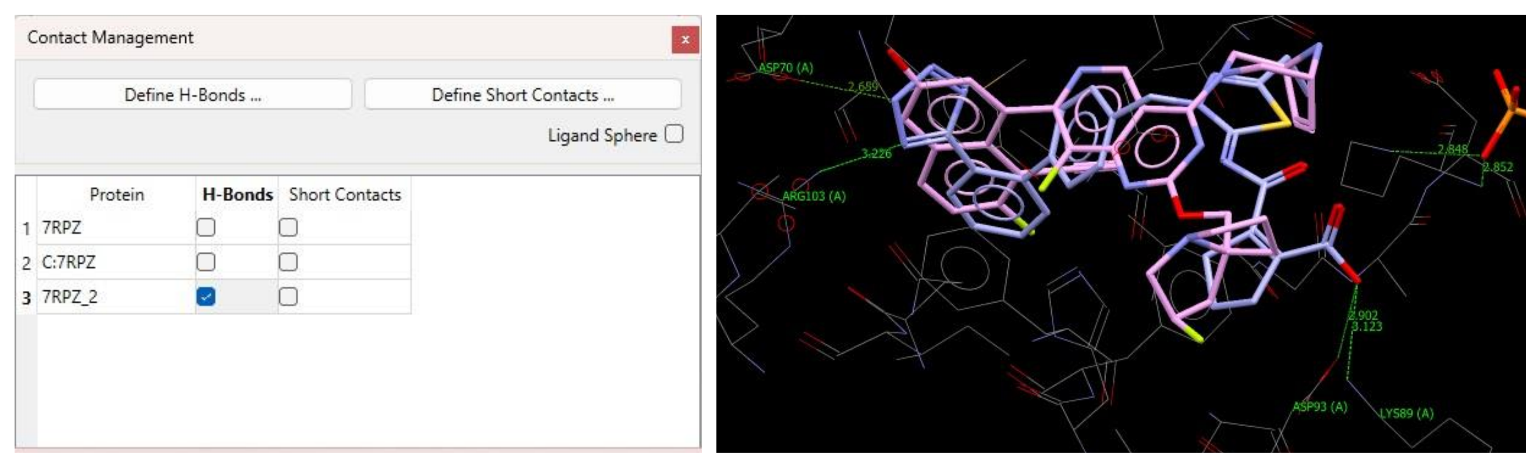
21. You can sort the solutions based on their PLP fitness score by clicking on **PLP.Fitness** in the Docking Solutions tab of the *Molecule Explorer* (click twice to sort by high to small). You can now observe and compare the solutions. You can also sort them also based on the other scores.





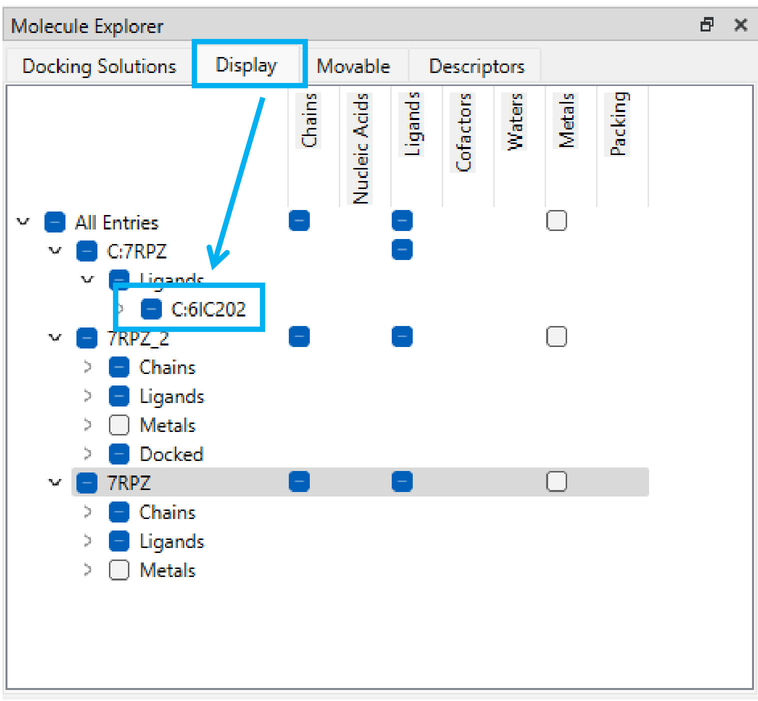
22. You may also explore the hydrogen bonds formed. To do so, in the *Contact Management* panel on the left-hand side, tick the box in the **H-bonds** column for the component of interest. The docked solutions correspond to entry **7RPZ_2**. If you don't see this tab, follow the instructions at the end of this handout to display it.

22



23. Observe:
    - How do the different ligands overlay with the co-crystallised ligand?
    - How do the hydrogen bonds change in the different solutions?

Hint: you may wish to hide the original ligand to view the interactions of the docked ligands more clearly. You can do this by deselecting **C:6IC202** in the Display tab of the Molecule Explorer window. You may also find it helpful to label the interacting protein residues as it **Step 17** of Exploring the Binding Site.



Conclusions

- We performed an in silico virtual screening of a library of inhibitors towards a KRAS protein co-crystallized with MRTX-1133: 4-(4-[(1R,5S)-3,8-diazabicyclo[3.2.1]octan-3-yl]-8-fluoro-2-{[(2R,4R,7aS)-2-fluorotetrahydro-1H-pyrrolizin-7a(5H)-yl]methoxy}pyrido[4,3-d]pyrimidin-7-yl)-5-ethynyl-6-fluoronaphthalen-2-ol (PDB 7RPZ).
- As expected, some in silico hits replicate interactions observed with the ligand A5G in the crystal structure.
- Interestingly, other hits that contain different scaffolds show interactions with the Asp12, Arg68, Glu62 in the binding region and with Gly60, Gly10, and Thr58 elsewhere within the pocket, suggesting that binding interactions or binding affinities can be enhanced while identifying potential new chemotypes that could represent new chemical entities to claim.

Summary

After this workshop, you should be able to:

- Import a file in Hermes directly from the PDB using the CSD Python API functionality.
- Set up protein, ligands, and parameters for a GOLD run, and run it.
- Visualize and analyse the GOLD virtual screening results with Hermes.

Next Steps

After the Explore More part of today's workshop, you can try the docking again using the full library of ligands from ChEMBL 35 website contained in the KRAS_inhibitors.mol2 file (see Additional files). Alternatively, try the section on Analysing Interactions Using SuperStar on the following page for an additional analysis of the results.

After this workshop, you can continue learning about GOLD with more exercises available in the self-guided workshops available in the CSD-Discovery workshops area on our website: https://www.ccdc.cam.ac.uk/community/training-and-learning/workshop-materials/csd-discovery-workshops/

Analysing Interactions using SuperStar

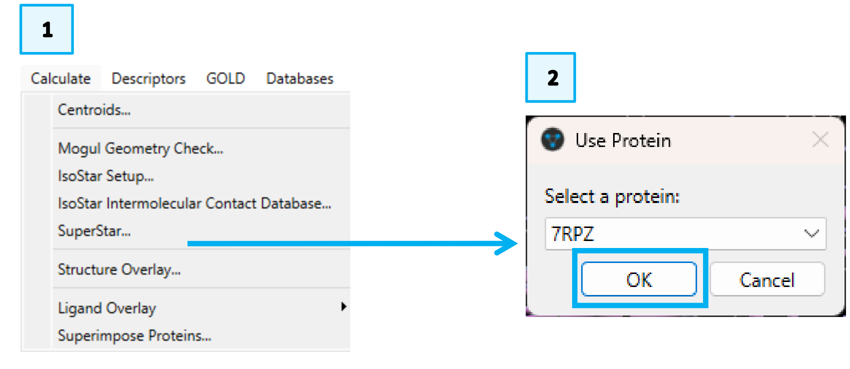
A more sophisticated, knowledge-based approach to exploring the protein-ligand interactions uses [SuperStar interaction maps](). SuperStar uses crystallographic information about non-bonded interactions to generate interaction maps within protein binding sites or around small molecules, i.e., it predicts 'hot-spots' where a chosen interaction-type is exceptionally favourable.

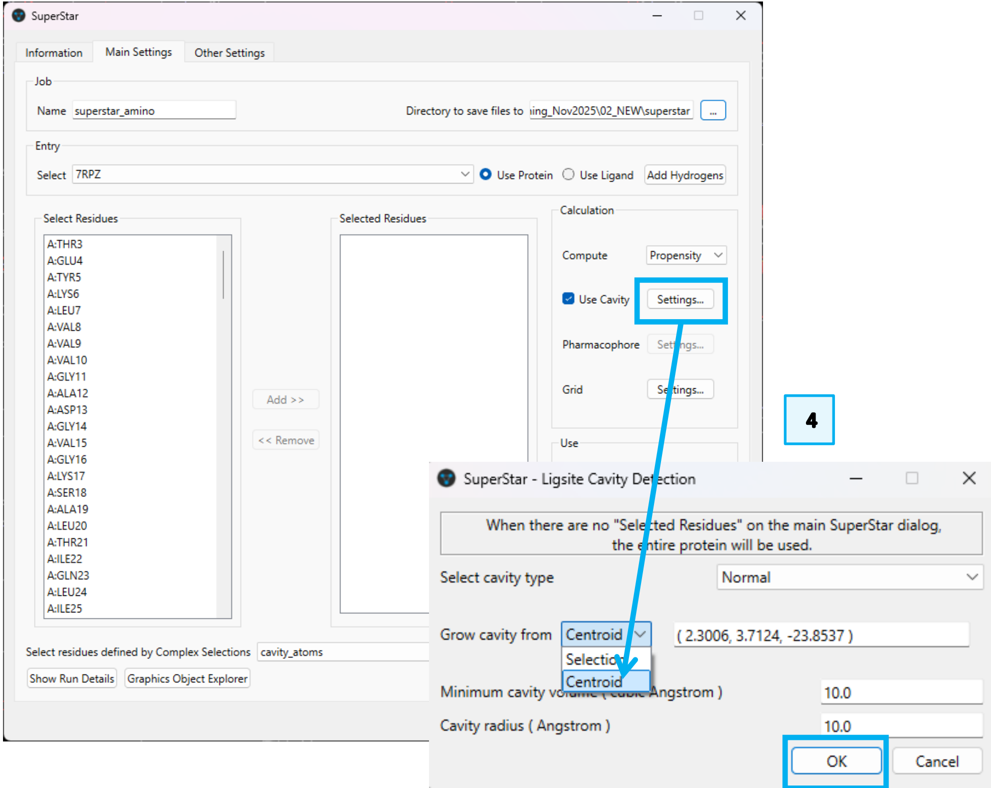
1. From the top-level menus, select *Calculate > SuperStar* to bring up the **SuperStar** window.

2. The *Use Protein* dialogue box will appear, from the drop-down menu, select **7RPZ**; this will launch the **SuperStar** window.

3. Go to the Calculation panel on the right and tick the box next to *Use Cavity*. Click on the *Settings* button to bring the **SuperStar – Ligsite Cavity Detection** window.

4. Select *Grow cavity from > Centroid*. Click on any atom in the inhibitor, sidechain, or any point in space in the binding site to select the cavity. Leave the rest parameters with their by-default values. Click on **OK**.



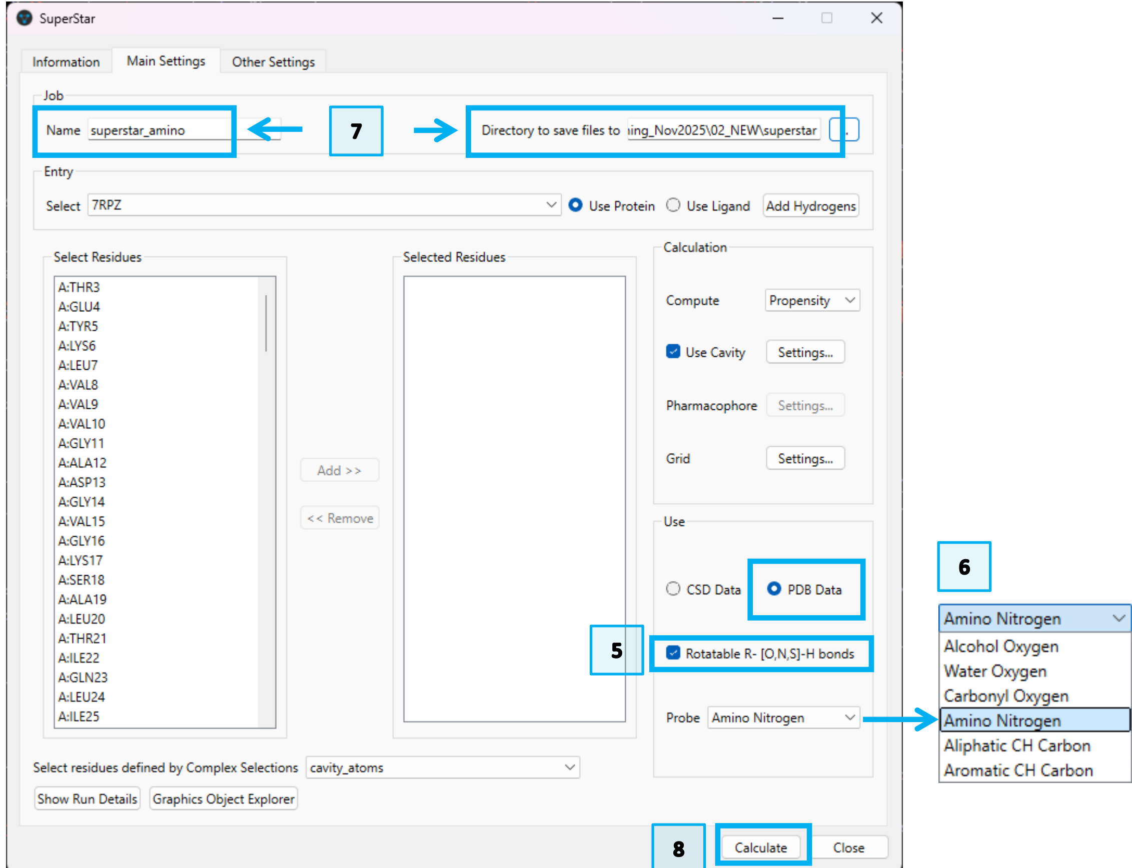


5. In the *Use* box on the right-hand side, select **PDB Data** and tick **Rotatable R-[O,N,S]-H bonds** to use rotatable bonds.

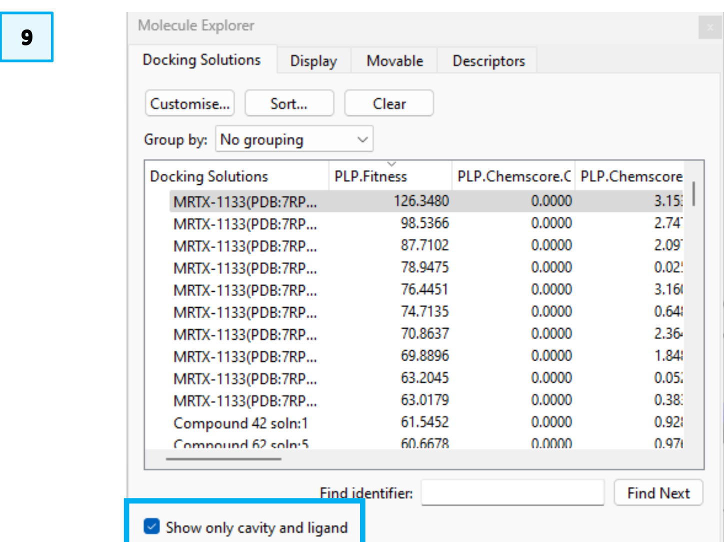
6. From the *Probe* drop-down menu, select **Amino Nitrogen**. A probe is a small functional group which is used to generate individual interaction maps.

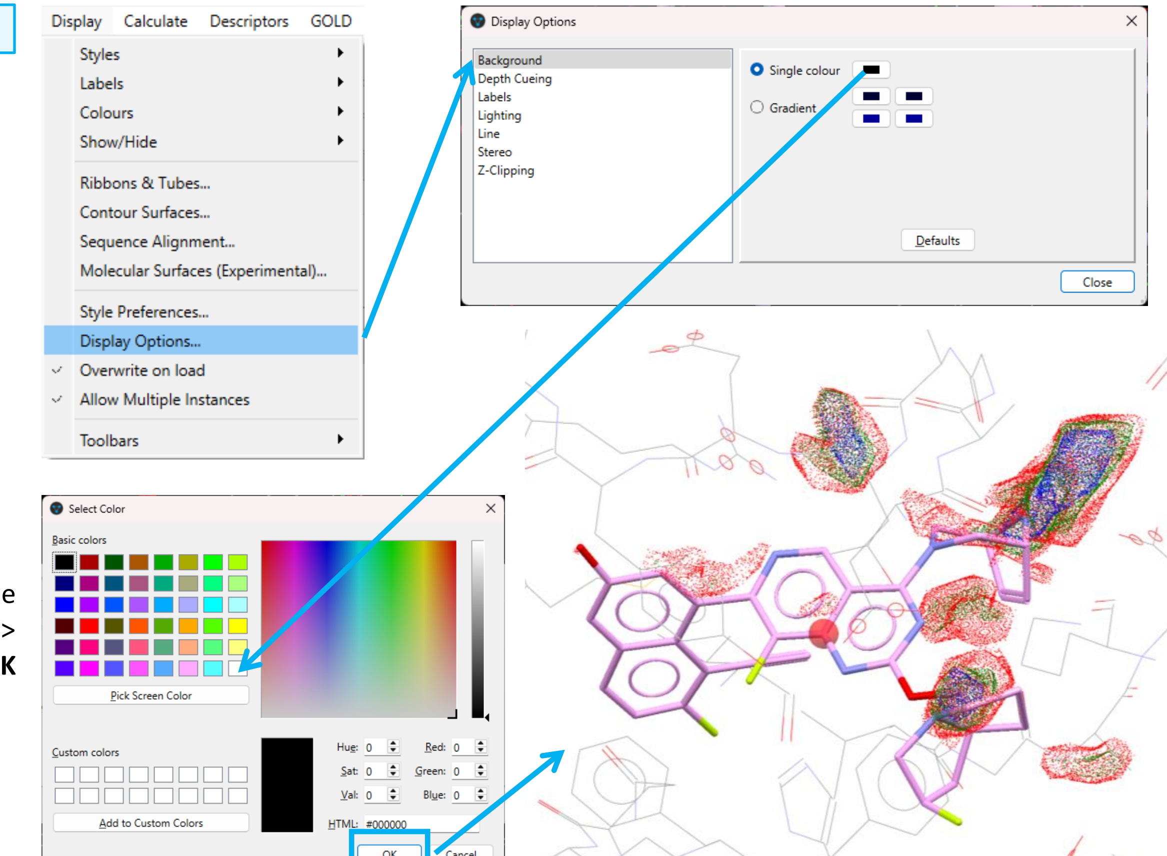
7. In the **Name** box, give the map an appropriate name such as “superstar_amino” and select a suitable directory to save to using the ... button.

8. Click **Calculate** to generate the interaction map. DO NOT close the window. We will use it again.



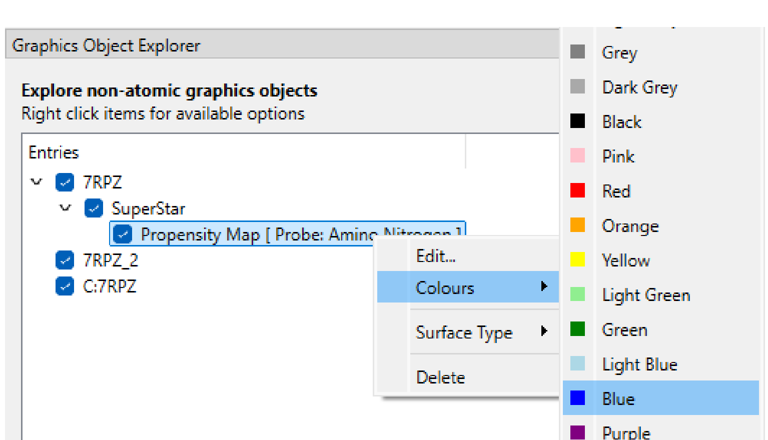
9. In the **Molecule Explorer** panel, on the left, tick the box next to *Show only cavity and ligand*. This will show only the residues and ligands in the cavity for improved visualization and analysis.

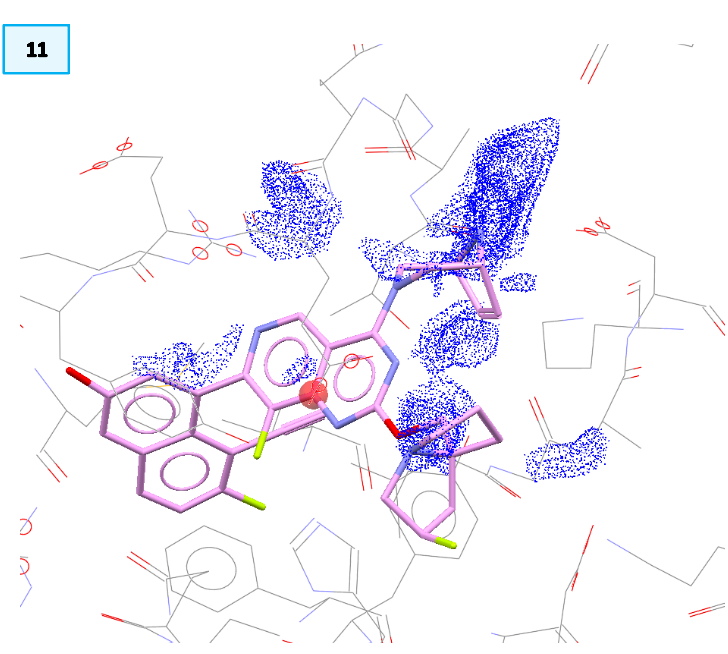




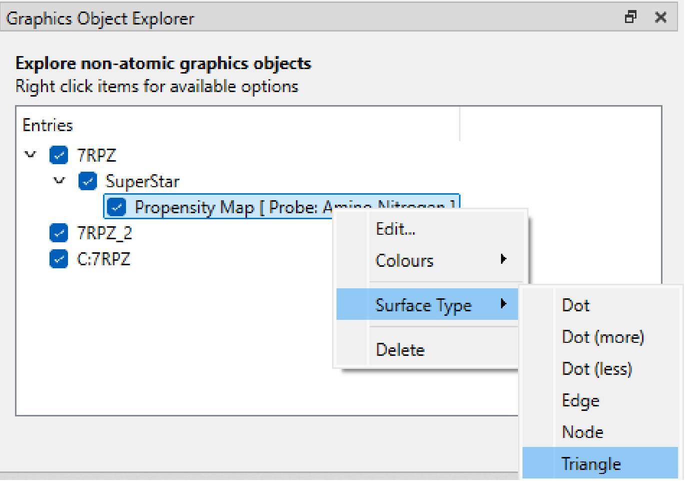
10. The interaction map will now be shown in the **Display window**. The background colour can be selected via Display > Display Options > Background. Click the box next to the radio button “Single Colour”. Click **OK** and **Close the Display Options**.

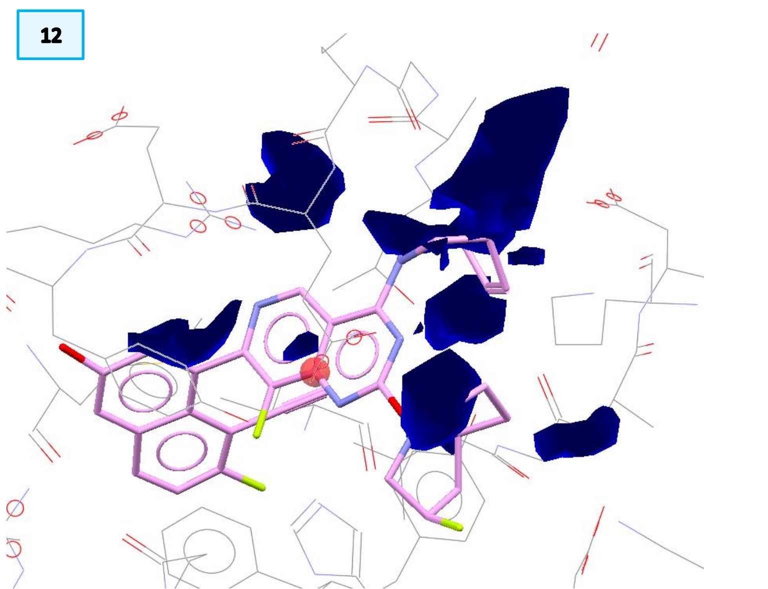
11. In the **Graphical Object Explorer** window **right click** on **Propensity Map [Probe:Amino Nitrogen]** and from the drop-down menu select *Colours > Blue*.



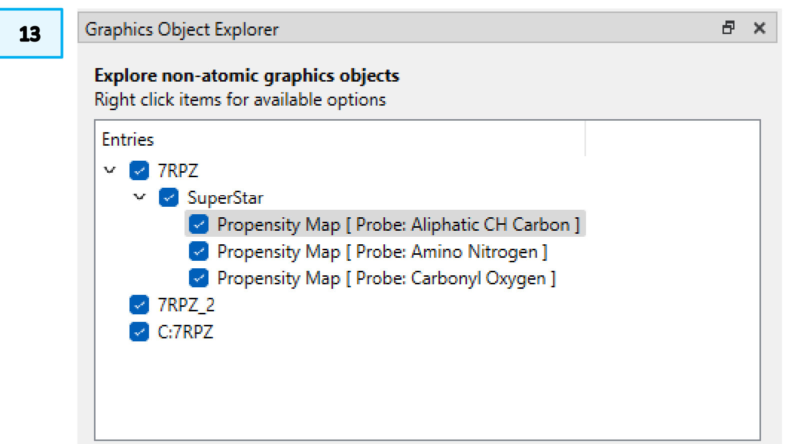


12. **Right click** again and select *Surface Type > Triangle*.

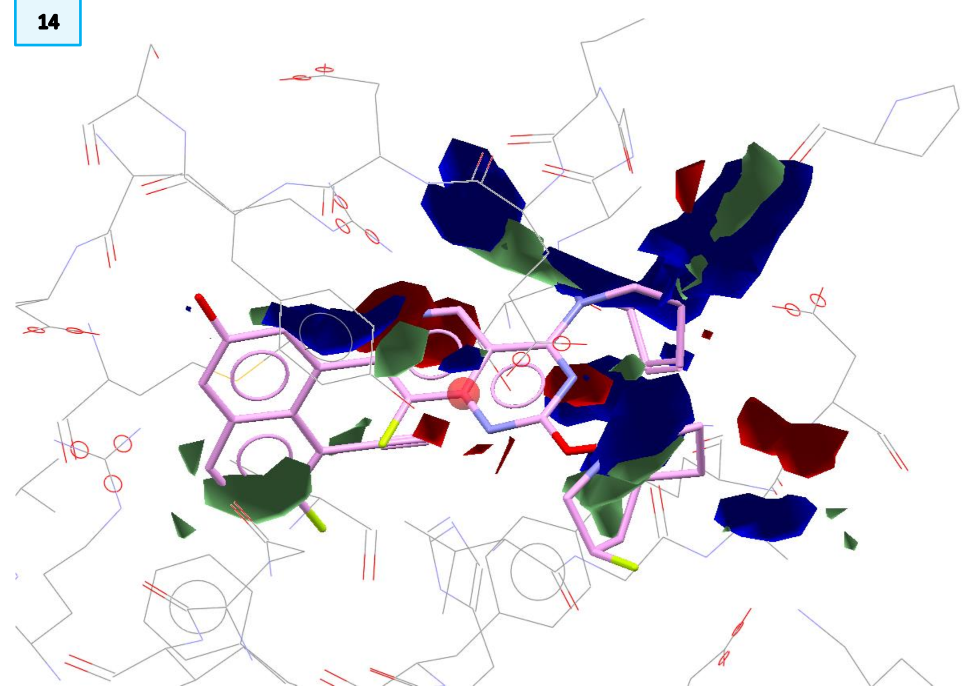




13. **Repeat Steps 6-12** to calculate the interaction map **for the Carbonyl Oxygen probe**. Customise the **colour to red** for this map.
14.

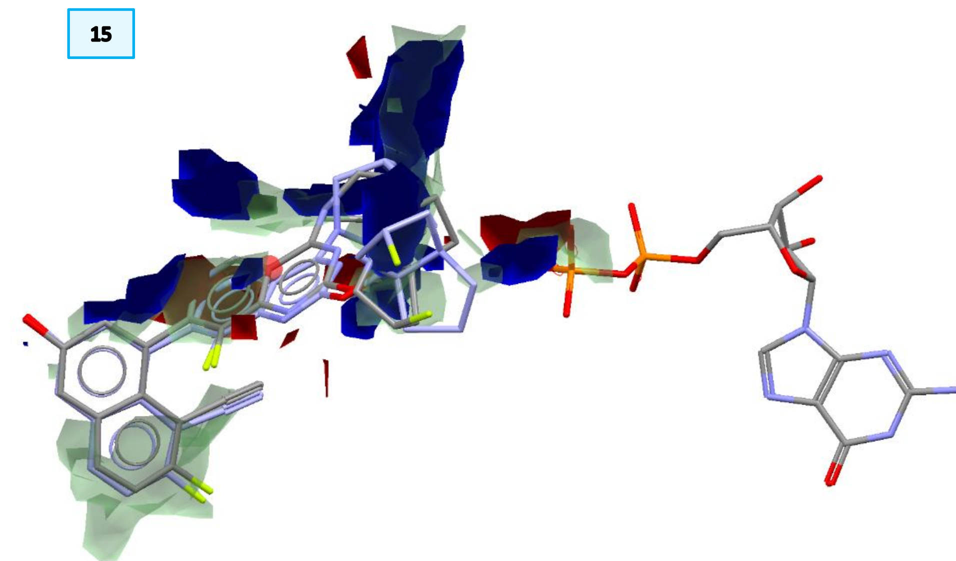


15. **Repeat Steps 6-12** to calculate the interaction map **for the Aliphatic CH probe**. Customise the **colour to green** for this map.

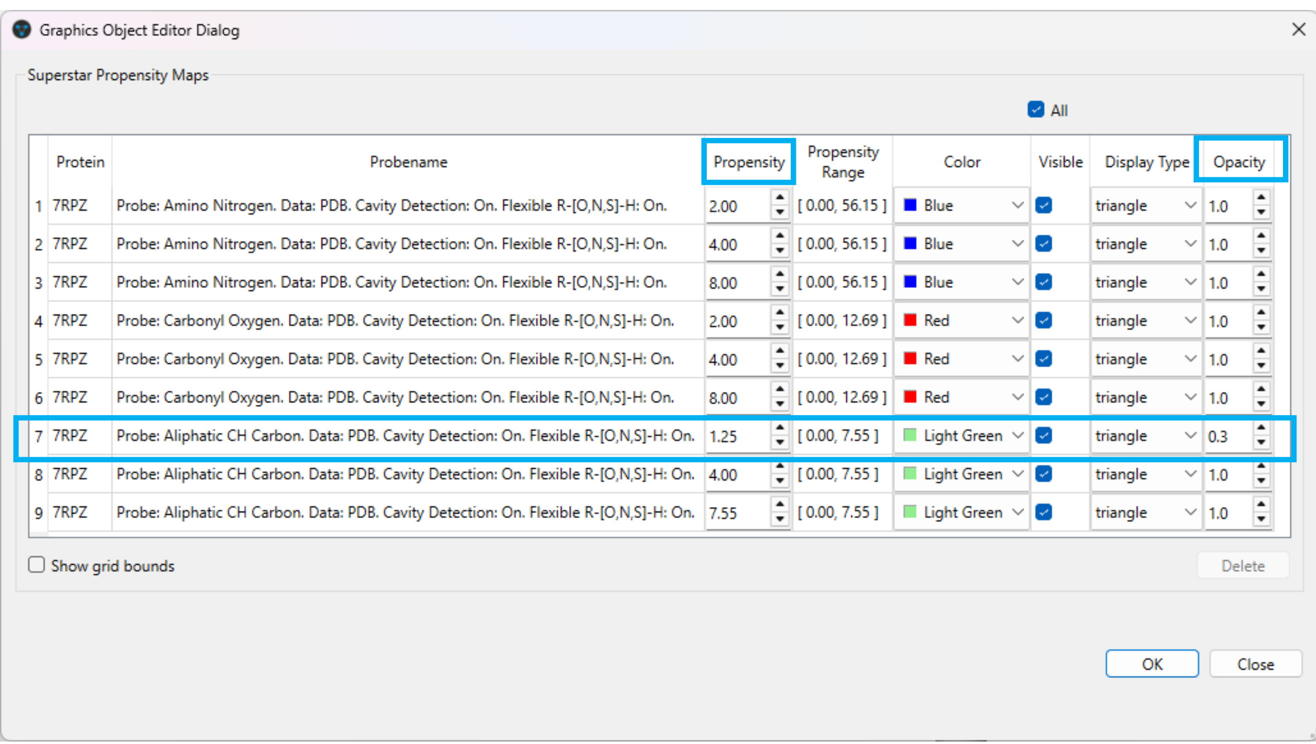


16. At times, certain interaction maps may not be immediately clear, and adjusting the visualization parameters can help reveal important details. To explore these options, try modifying the propensity levels in the visualizer.

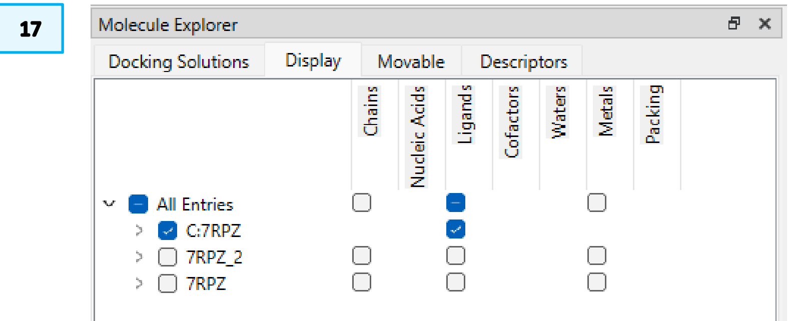
17. **Right-click** on **Propensity Map [Probe: Aromatic CH carbon]** and select **Edit**. Then, update the propensity value of the first aliphatic map to 1.25 and set the opacity to 0.3. Click **OK** to apply the changes.

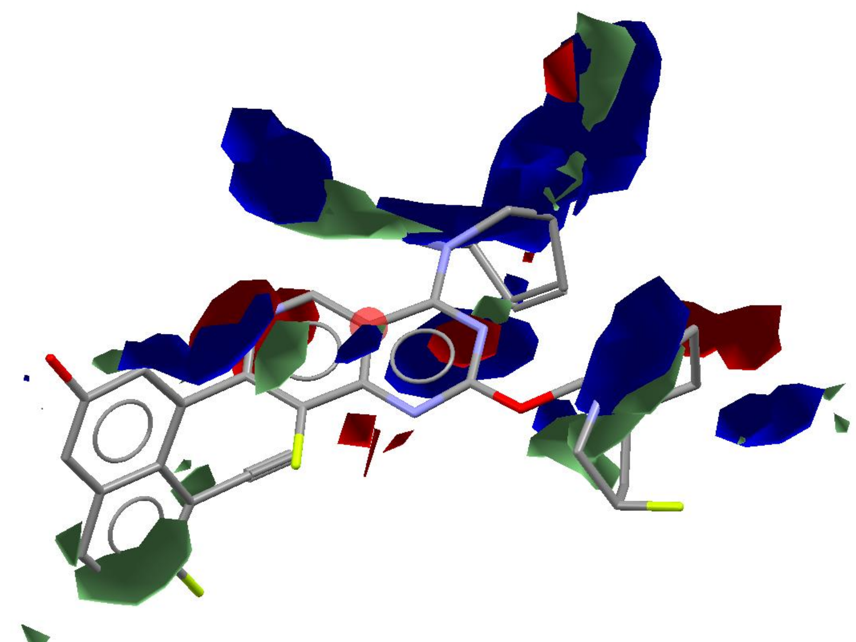


16

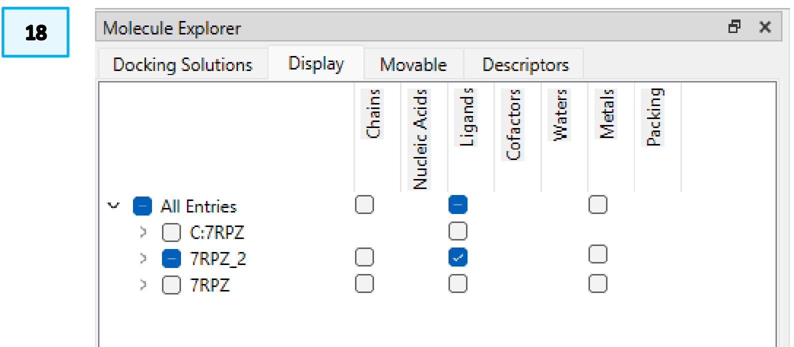


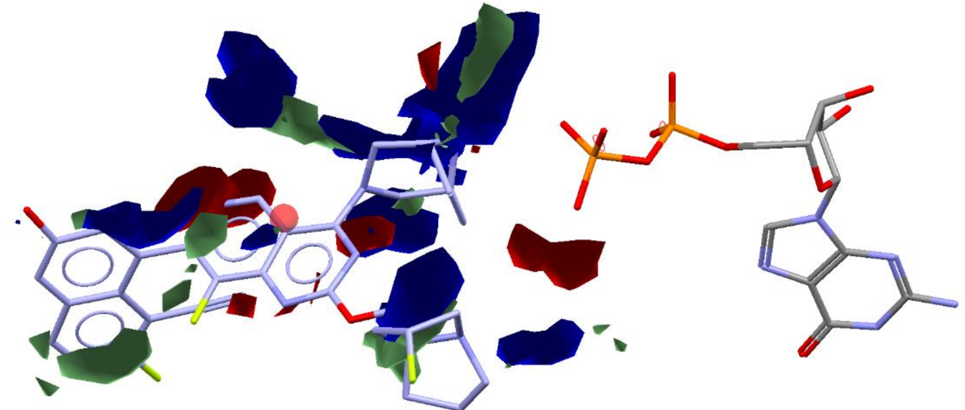
18. To focus on the interaction map it may be helpful to hide the side chain and docked ligands in the display, to do so untick **Chains** adjacent to **All Entries** in the **Display** tab of the **Molecule Explorer**. Also, untick **7RPZ_2** in the **Ligands** column.



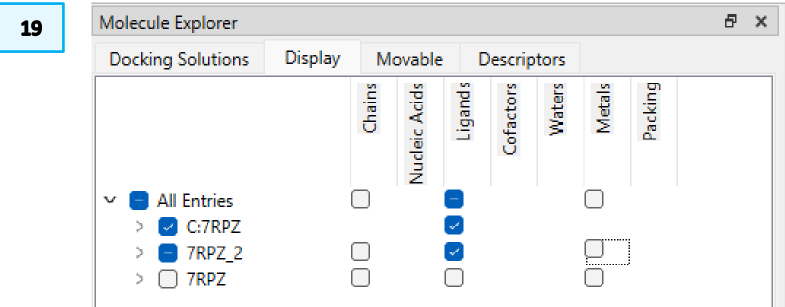


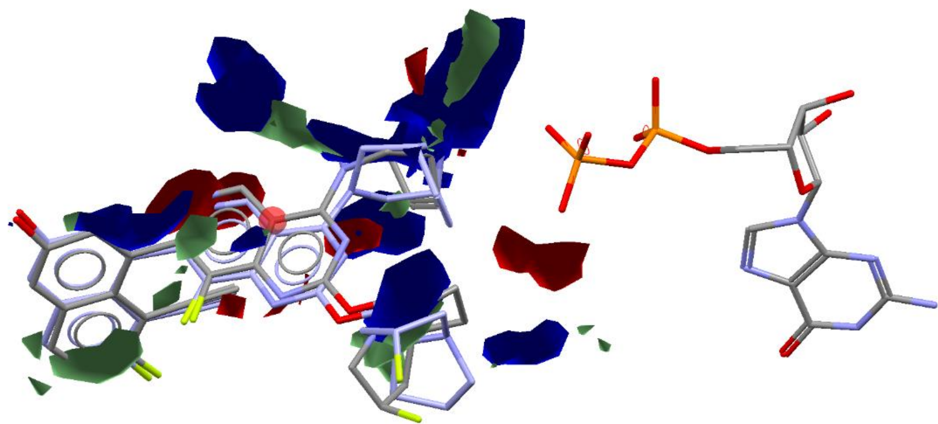
19. Untick **C:7RPZ** in the **Ligands** column of the molecule explorer and tick **7RPZ_2** in the same column. Scroll through the docking solutions to see how well each docked ligand satisfies the SuperStar interaction map.





20. Keep exploring different options and combinations in the **Display** tab of the **Molecule Explorer** for more custom-made visualizations of protein-ligand interactions.

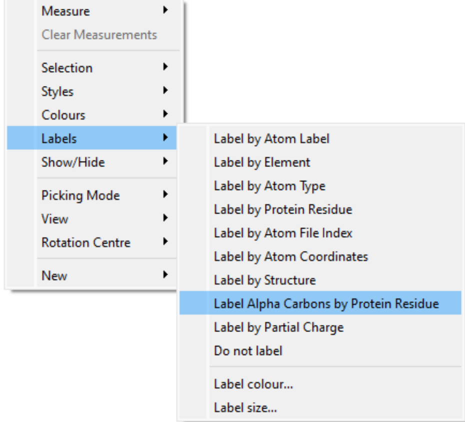




Challenge

It was said in the introduction that structural analysis revealed key interactions with **Asp12**, **Arg68**, and **Glu62**. Additional interactions with **Gly60**, **Gly10**, **Thr58**, and a hydrophobic pocket contributed to potency. *Can you identify any of these key residues near the hotspots calculated with SuperStar?*

Tip: you can label protein residues most clearly by right clicking in the visualiser area then select *Labels > Label Alpha Carbons by Protein Residue*.



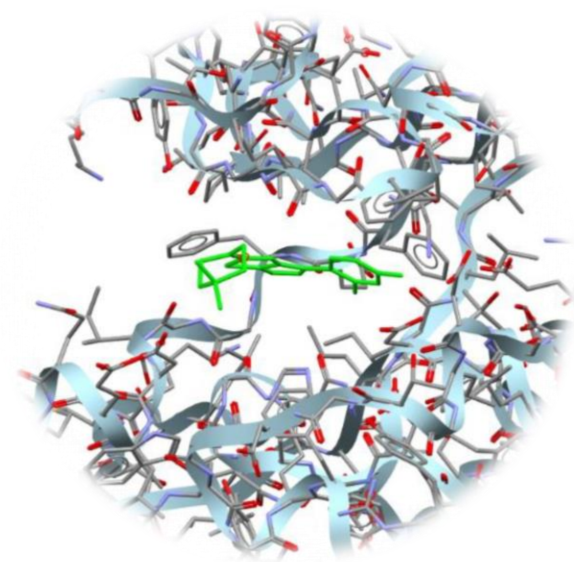
Glossary

Binding site

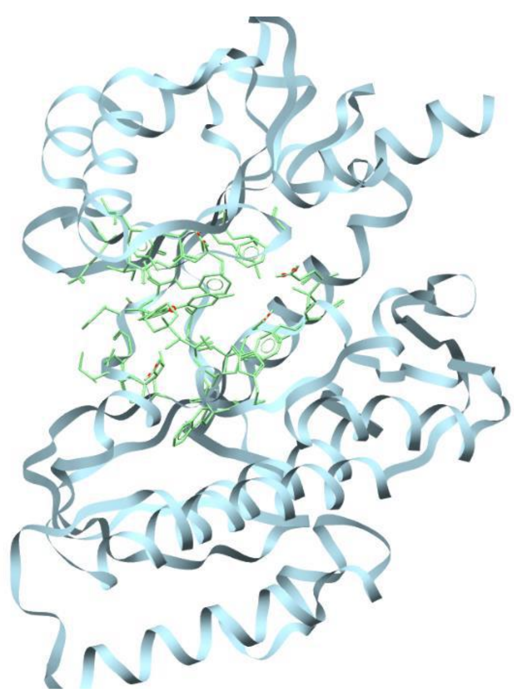
A specific region (or atom) in a molecular entity that is capable of entering into a stabilizing interaction with another molecular entity. An example of such an interaction is that of an active site in an enzyme with its substrate. Typical forms of interaction are by hydrogen bonding, coordination and ion pair formation. Two binding sites in different molecular entities are said to be complementary if their interaction is stabilizing. Source: PAC, 1994, 66, 1077. (Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)) on page 1089.

Docking

Docking studies are computational techniques for the expl oration of the possible binding modes of a substrate to a given receptor, enzyme or other binding site. Source: PAC, 1997, 69, 1137. (Glossary of terms used in computational drug design (IUPAC Recommendations 1997)) on page 1142.



A protein displayed as a ribbon with its binding site represented in capped sticks style in green colour.



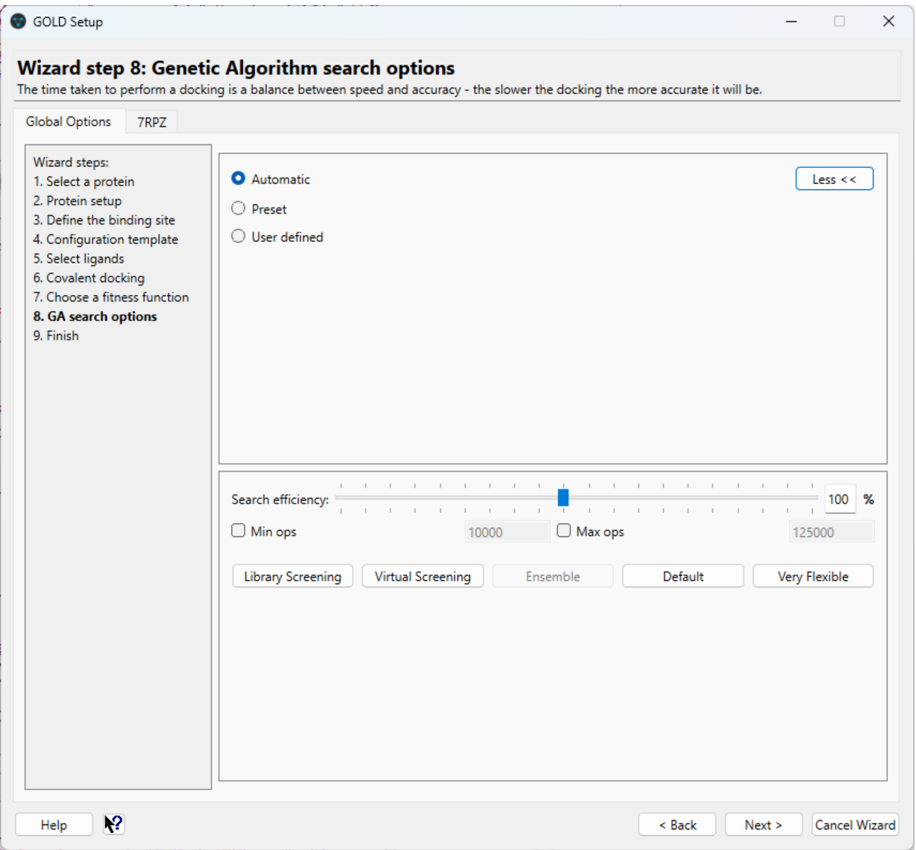
GA setting

A genetic algorithm is an optimization algorithm based on the mechanisms of Darwinian evolution which uses random mutation, crossover and selection procedures to breed better models or solutions from an originally random starting population or sample. Source: PAC, 1997, 69, 1137. (Glossary of terms used in computational drug design (IUPAC Recommendations 1997)) on page 1144.

In GOLD, number of genetic operations performed (crossover, migration, mutation) is the key parameter in determining how long a GOLD run will take (i.e. this parameter controls the coverage of the search space). GOLD can automatically calculate an optimal number of operations for a given ligand, thereby making the most efficient use of search time, e.g. small ligands containing only one or two rotatable bonds will generally require fewer genetic operations than larger, highly flexible ligands. The criteria used by GOLD to determine the optimal GA parameter settings for a given ligand include:

- the number of rotatable bonds in the ligand,
- ligand flexibility, i.e. number of flexible ring corners, flippable nitrogens, etc. (see Ligand Flexibility),
- the volume of the protein binding site, and
- the number of water molecules considered during docking (see Water Molecules).

The exact number of GA operations contributed, e.g. for each rotatable bond in the ligand, are defined in the gold.params file (see Altering GOLD Parameters: the gold.params File). For further information please refer the GOLD user manual.



GOLD

GOLD (Genetic Optimisation for Ligand Docking) is a genetic algorithm for docking flexible ligands into protein binding sites. GOLD has been extensively tested and has shown excellent performance for pose prediction and good results for virtual screening.

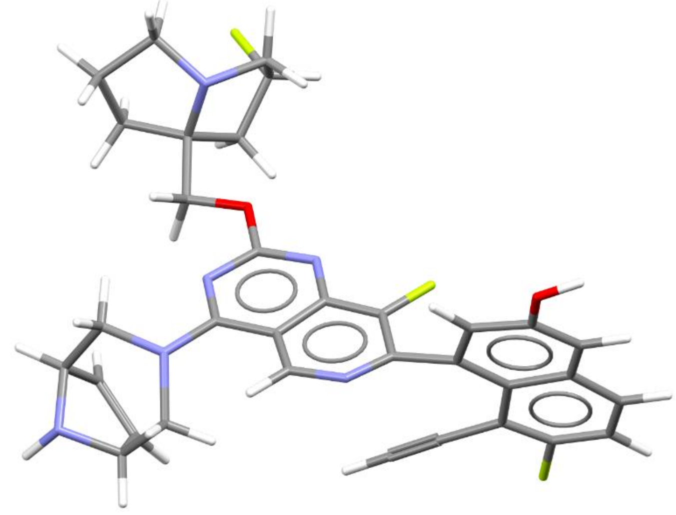
Ligand
In this context, these are the molecules under investigation. These could be organic, inorganic and macromolecules.

Scoring function
In the fields of computational chemistry and molecular modelling, scoring functions are mathematical functions used to approximately predict the binding affinity between two molecules after they have been docked. Most commonly one of the molecules is a small organic compound such as a drug and the second is the drug's biological target such as a protein receptor.

Stochastic sampling
Random sampling; random sample The sample so selected that any portion of the population has an equal (or known) chance of being chosen. Haphazard or arbitrary choice of units is generally insufficient to guarantee randomness. Source: PAC, 1990, 62, 1193 (Nomenclature for sampling in analytical chemistry (Recommendations 1990)) on page 1202.

SuperStar interaction map
SuperStar is a program for identifying regions within a protein binding site or around a small molecule where particular functional groups (probes) are likely to be found. The propensity of finding probes in different regions of space is mapped out using experimental data from crystal structures in the CSD and PDB, to give a propensity map which indicates regions of favourable interaction.

Virtual screening
Virtual screening is a computational technique used in drug discovery which automatically evaluates a library of small-molecule compounds (which may be very large in number) for their ability to bind to a drug target such as a protein.



4-(4-[(1R,5S)-3,8-diazabicyclo[3.2.1]octan-3-yl]-8-fluoro-2-{[(2R,4R,7aS)-2-fluorotetrahydro-1H-pyrrolizin-7a(5H)-yl]methoxy}pyrido[4,3-d]pyrimidin-7-yl)-5-ethynyl-6-fluoronaphthalen-2-ol co-crystallised with KRAS^{G12D} protein studied in this workshop.

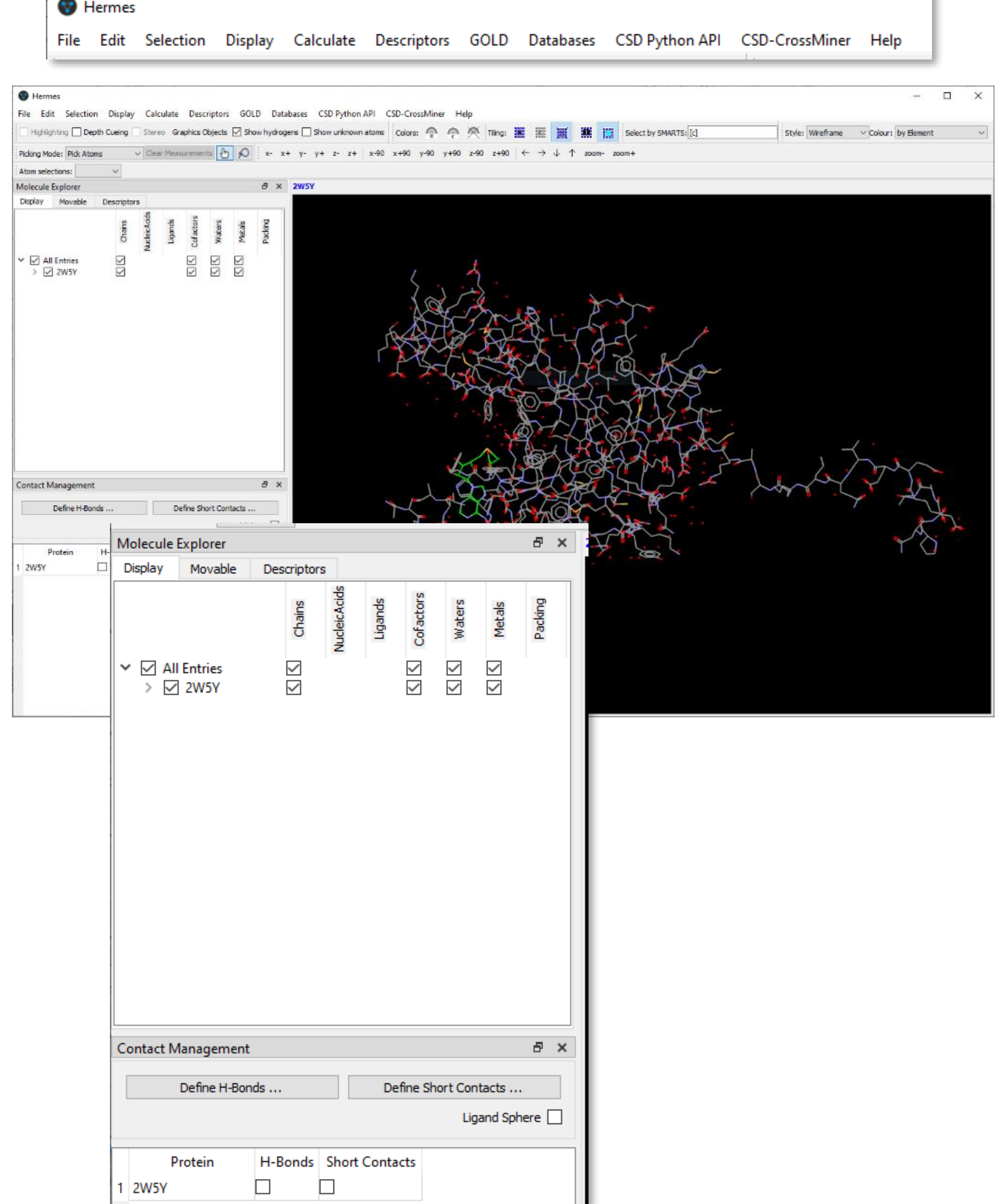
Basics of Hermes Visualisation

Hermes is the CCDC's visualization software to study protein-ligands interaction, and it is the interface for protein docking with GOLD.

In the following we will see some of the basics of navigation and visualization in Hermes that you will find helpful to support your analysis.

In the **Hermes interface** we find:

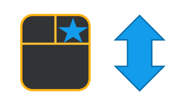
- **At the top**: list of menus from which we can access visualization and analysis options, and other CSD-Discovery components.
- **On the left-hand side**: the **Molecule Explorer**, where the proteins, ligands, water molecules, and other molecules present in the structures are listed in the *Display* tab. You can expand each entry to see the components, and tick or untick the corresponding boxes to display / not display them. The solutions from a GOLD run will also be displayed in the Molecule Explorer.
- **On the left-hand side**, below the Molecule Explorer: the **Contact Management** box, which allows to display H-bonds and short contacts for selected molecules by ticking the corresponding box.



Using the **mouse to enhance visualization**:

- 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

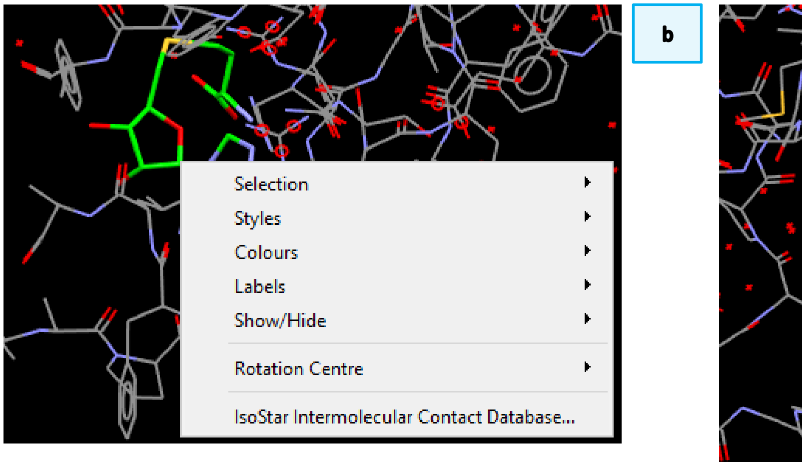
- Shift + Left mouse button and move - rotate in the plane molecules

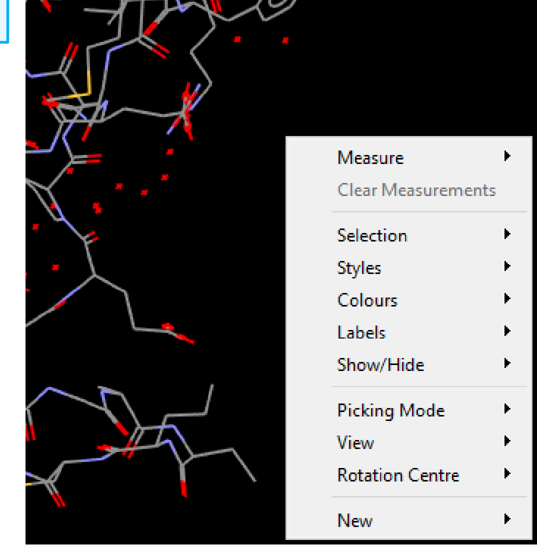
- Ctrl + Left mouse button and move - translate molecules

Right click:

a) On a feature

b) Away from a feature



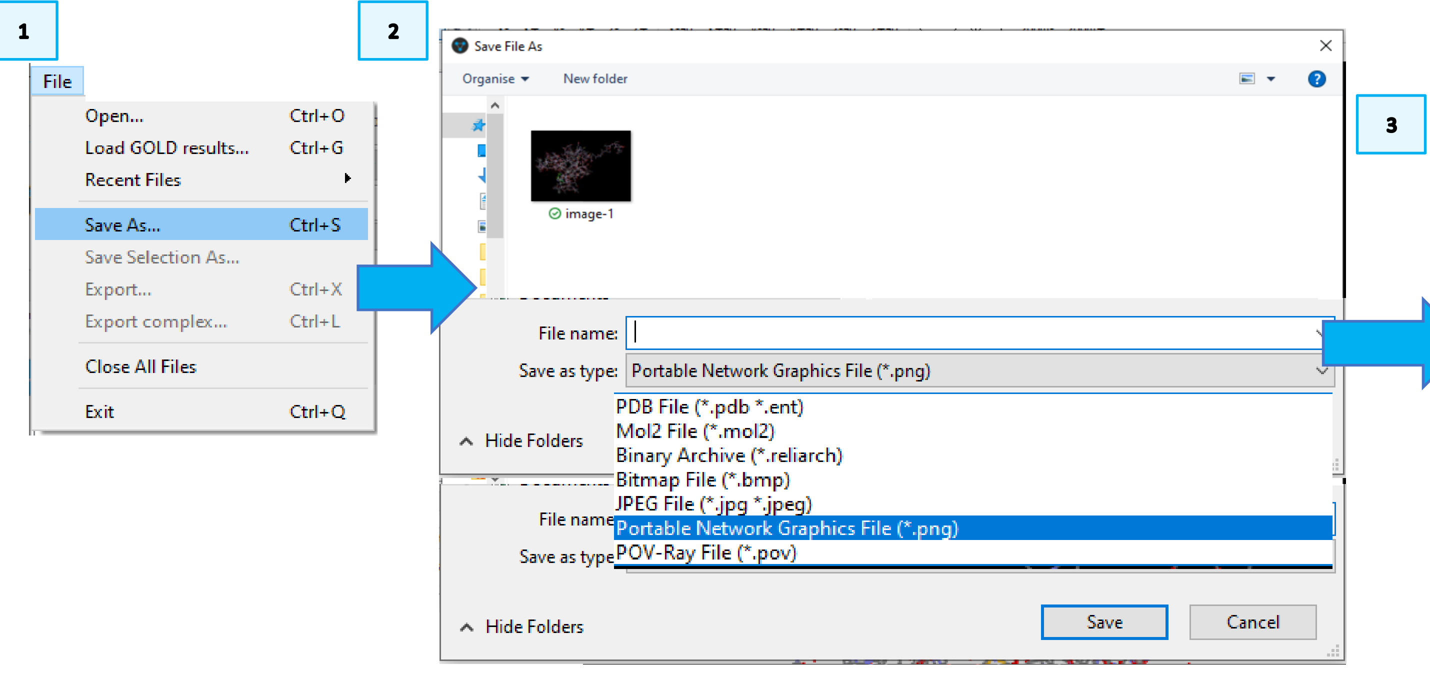


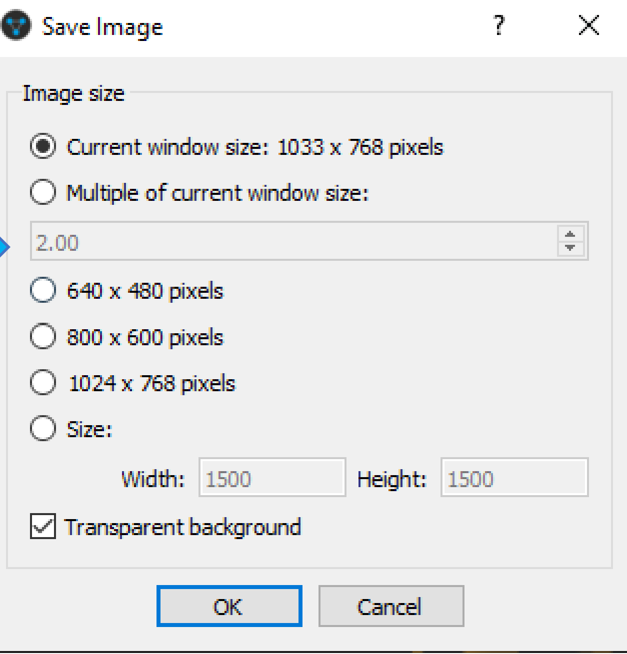
Creating high-resolution images

1. Go to the top menus and click on **File**, then **Save A**s. Alternatively you can use the keyboard and do **Ctrl + S**.

2. In the *Save File As* window, select destination folder, file type and choose a name. Then click **Save**.

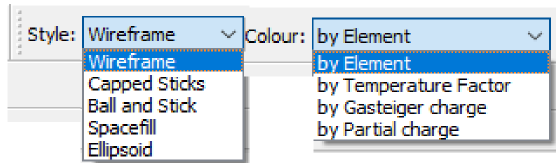
3. This will bring up the *Save Image* menu. Here you can select the resolution and tick whether you want a transparent background. Click **Ok**.

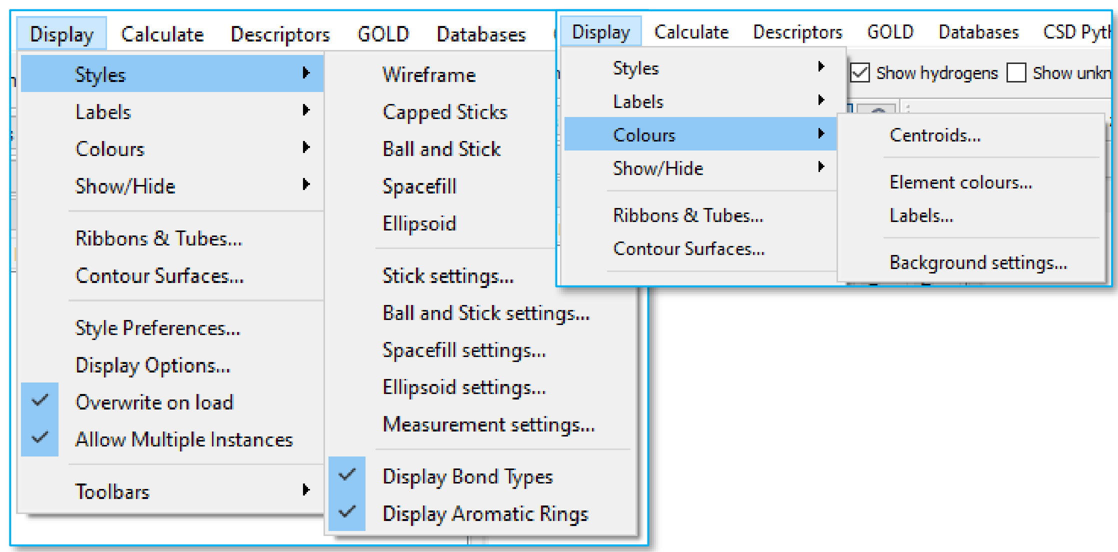




You can access visualization options from the **Display** menu.

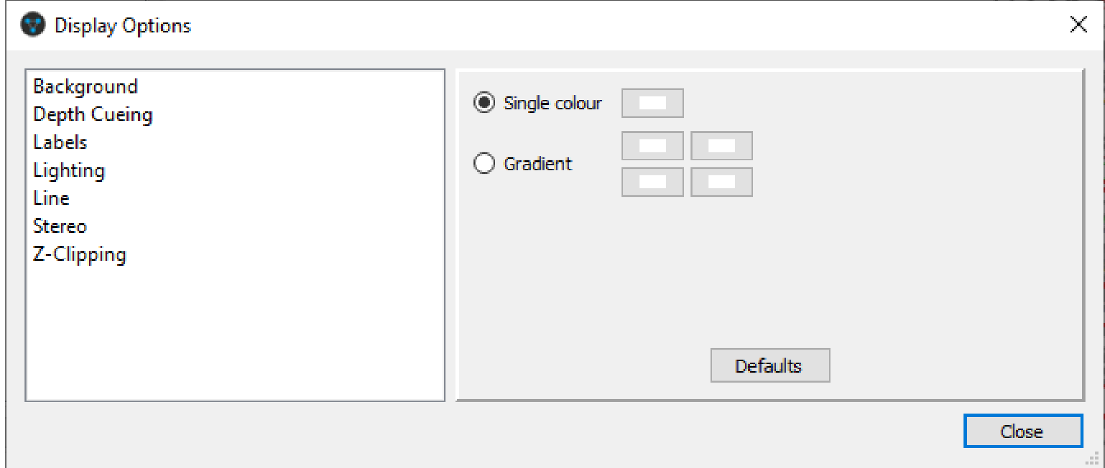
- From **Display > Styles**, you can pick the representation style of atoms and bonds and change settings. **Note**: you can also access this feature from the dropdown menu at the top right of the 3D visualizer window.





- From **Display > Colours**, you can edit the colour for elements and object. **Note**: you can change the colouring style for the structure from the dropdown menu at the top right of the 3D visualizer window.

- Clicking on Display > Display Options brings up the Display Options menu, where you can edit settings for the 3D visualizer.



- If any of the toolbars at the top of the Hermes interface is not displayed and you wish to have it shown, or vice versa, you can do so from **Display > Toolbars**, where you can ensure to tick the ones you wish to have on.