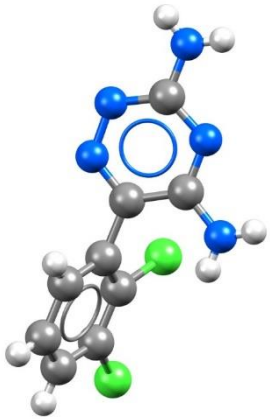
Validating Docking Poses with SuperStar and Full Interaction Maps (FIMs)

Version 1.0 - February 2019
GOLD v5.7.0, Hermes v1.10.0, SuperStar v2.2.0, Mercury v4.0.0



The Cambridge Crystallographic Data Centre

Table of Contents

Validating Docking Poses with SuperStar and Full Interaction Maps (FIMs)

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 interact favourably.

SuperStar uses real experimental information about intermolecular interactions, derived from either the Cambridge Structural Database (CSD) or the Protein Data Bank (PDB).

It applies a fully knowledge-based approach to identify protein-ligand interaction hot spots.

On the other hand, Full Interaction Maps (FIMs) provide a small molecule's interaction preferences, highlighting the regions around the molecule (maps) where chemical probe groups are likely to be found, based on the IsoStar interaction data extracted from the CSD. FIMs can be used to understand likely interaction patterns between a ligand and protein from the ligand's perspective.

In this workshop, you will use both SuperStar and FIMs to validate docking poses obtained with GOLD. You will learn how to visually identify what are the most likely interactions within the protein's pocket and the preferred interactions of the docked molecule.

The files to perform this workshop are provided in the `workshop` folder here.

This example assumes you are already largely familiar with GOLD and with how to setup proteins and ligands for docking calculations. If not, please refer to the following sections of the GOLD User Guide:

- Setting Up the Protein
- Essential Steps

Case Study

Introduction

Indole-3-glycerol phosphate synthase (IGPS) catalyses the fifth reaction in the pathway for the biosynthesis of tryptophan, in which the substrate 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate (CdRP, Figure 1a) undergoes a ring closure reaction to the product indole-3-glycerol phosphate (IGP, Figure 1d). Hence, this enzyme has one substrate, CdRP, but three products: 1-C-(indol-3-yl)-glycerol 3-phosphate, CO_2, and H_2O.

IGPS from the hyperthermophile *Sulfolobus solfataricus* (sIGPS) has been determined at 2 Å resolution (1igs.pdb).

Your task:

The object of this workshop is to perform a docking calculation to dock the IGP product into the sIGPS receptor, and then use SuperStar and FIMs to determine the important structural features needed for the IGP ligand to interact with the IGPS receptor, helping to improve the understanding of the catalytic mechanism of IGPS. This will allow the validation of the best-ranked docking pose. Note that a co-crystal structure of the sIGPG-IGP resolved at 2 Å resolution (1a53.pdb) is available for comparison.

Provided input files in the `workshop` *folder:*

- *1IGS_protein.mol2*: sIGPS protein coordinates derived from the crystallographic structure *1igs.pdb*. The protein structure has already been set up in accordance with the guidelines for the preparation of protein input files in the GOLD User Guide.
- *igp_ligand.mol2*: We will use the idealised version of igp_ideal.sdf downloaded from the PDB, which has no resemblance to the original binding mode. The *igp_ligand.mol2* has been set up in accordance with the guidelines for the preparation of input files in the GOLD User Guide.
- *cavity.txt*: A text file containing the list of atoms included in the binding site. This file was generated using the cavity detection available through SuperStar in Hermes.

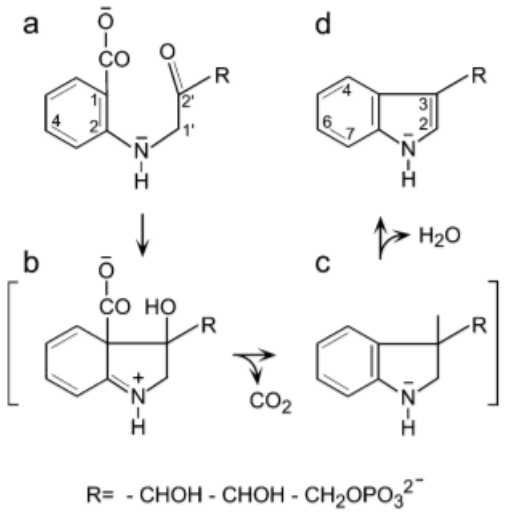


Figure 1. *The mechanism of the indole-3-glycerol phosphate synthase reaction. CdRP, 1-(o-carboxyphenylamino)-1-deoxyribulose 5-phosphate, undergoes a ring closure reaction to the product indole-3-glycerol phosphate.*

- *gold.conf*: The GOLD configuration file has been provided for this workshop and can be used to perform the docking calculation.

Challenges:

The example used here mimics the situation where a researcher has a crystal structure of a protein and wants to explore the binding site of this protein to determine the key features required for a ligand to powerfully and selectively bind the protein and for a protein to satisfy the ligand's interaction preferences. This workshop example is suitable for experienced users of CCDC tools.

Docking the IGP ligand in IGPS protein

1. Open Hermes. Select **GOLD** and then **Setup and Run a Docking…** from the resultant pull-down menu.

2. Select **Load Existing** from the *GOLD* pop-up window and navigate to the `workshop` folder to load the provided *gold.conf* configuration file.

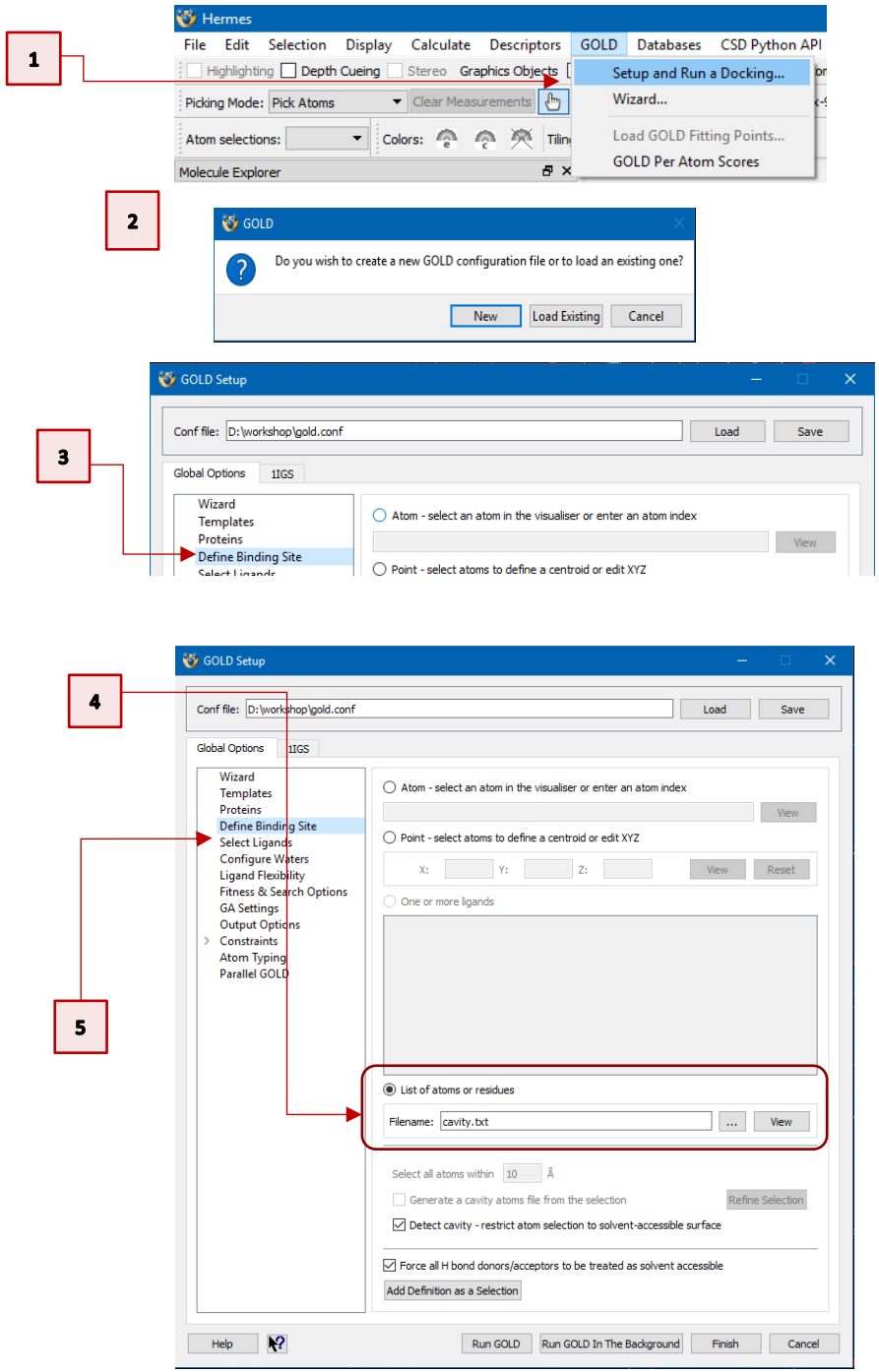
This will automatically load the settings and parameter values for this workshop into the *GOLD Setup* window in addition to the specified protein file (1IGS). We will take a look at the settings defined for this docking job. The GOLD interface contains two tabbed views; the default is *Global Options* which allows you to specify particulars of the docking in general; the other displays the protein name, in this case *1IGS*, and allows you to edit the protein and set up parameters specific to the protein such as flexible sidechains.

3. Click on the **Define Binding Site** option from the *Global Options* tab of the *GOLD Setup* window.

4. In this example, the binding site is defined as a list of atoms, which is provided in the file *cavity.txt.* Carbon atoms outside of the binding site will be displayed in purple. Leave all other parameters to their defaults. For visualisation purposes, it can help here to switch off the display H-atoms using the **Show hydrogens** tick box in the top-level menu of Hermes.

The cavity was determined using the cavity detection functionality of SuperStar run on the entire protein, and the atom indexes composing the cavity were saved into a text file (*cavity.txt)*. Cavities are identified using the LIGSITE algorithm (M. Hendlich *et al.*, *J. Mol. Graph. Model.*, **15**, 359-363, 1997) by scoring all grid points according to how deeply they are buried. The higher the value, the more deeply buried the point is within the cavity. If you wish to generate this file by yourself, you can use the SuperStar User Guide to guide you.

5. Click on **Select Ligands**. The ligand (*igp_ligand.mol2*) is listed under the *Ligand File* column. As with the protein file, the ligand has been supplied with all hydrogen atoms added.



6. Click on **Ligand Flexibility** and make sure that **Flip pyramidal N**, **Flip amide bonds** and **flip ring corners** tick-boxes are all ticked.

    a. **Flip pyramidal N** check-box: allows pyramidal (i.e. non-planar sp^3) nitrogen atoms of the ligand to invert during docking.

    b. **Flip amide bonds** check-box: allows amides, thioamides, ureas, and thioureas in the ligand to flip between *cis* and *trans* conformations during docking.

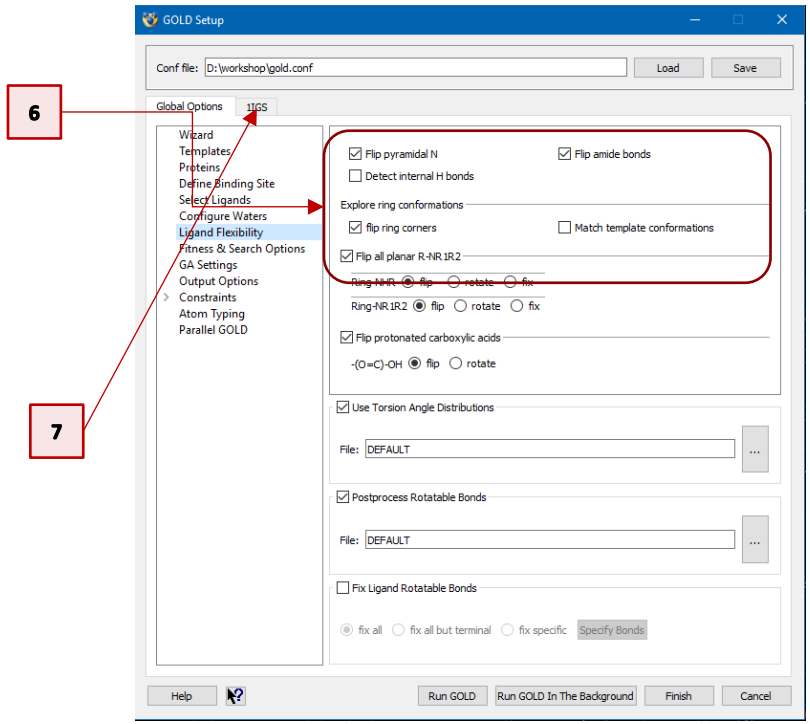
    c. **flip ring corners** check-box: allows GOLD to perform a limited conformational search of cyclic systems by allowing free corners of the rings in the ligand to flip above and below the plane of their neighbouring atoms.

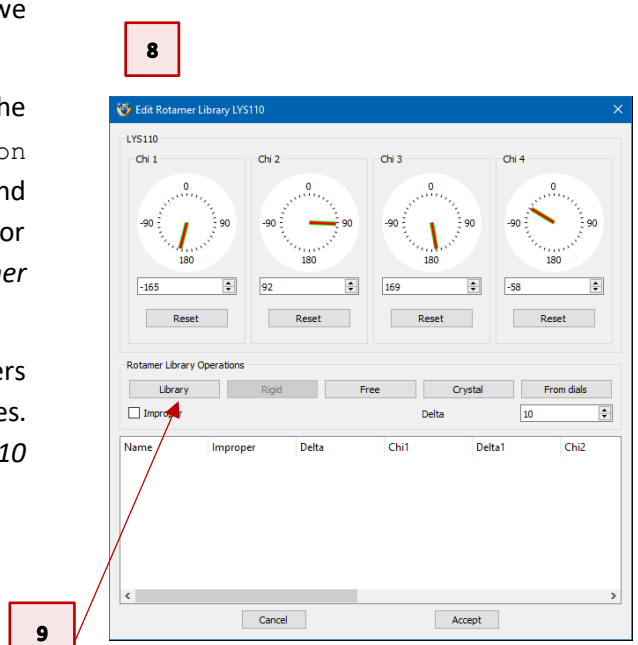
7. Now, click on the *1IGS* tab of the *GOLD Setup* window and then click on **Flexible Sidechains** from the resulting *1IGS* tab option. Here we will specify sidechains that will be allowed to rotate during docking.

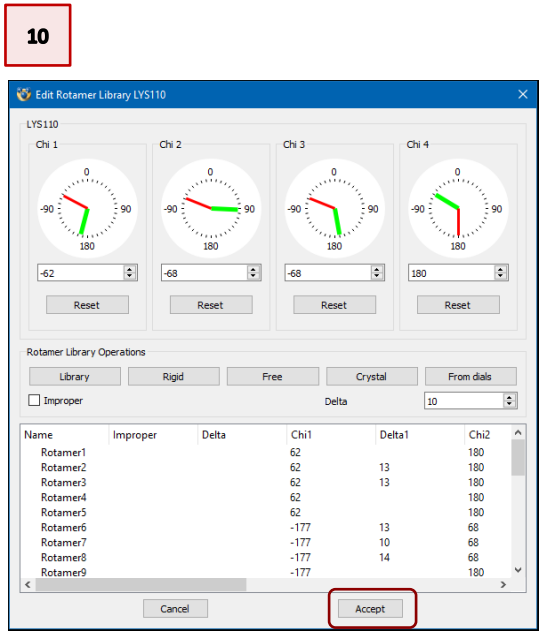
8. Scroll down the list of residues until you find *LYS110.* Click on *LYS110* and then click on the **Edit** button. This will open a dialogue window where we can define the rotamers for Lys110.

9. Click on the **Library** button. The dialogue has been updated to reflect the rotamers defined from a rotamer library, available as `<Installation folder>/Discovery_2019/GOLD/gold/rotamer_library.txt`, and which compiles the most commonly observed side chain conformations for the naturally occurring amino acids (derived from *The Penultimate Rotamer Library*, S. C. Lovell *et al.*, *Proteins*, **40**, 389-408, 2000).

10. If you scroll down the list, you will notice that there are 27 allowed rotamers listed in the *Edit Rotamer Library LYS110* for the 4 rotatable torsion angles. Click **Accept** to use this list and close the *Edit Rotamer Library LYS110* window.

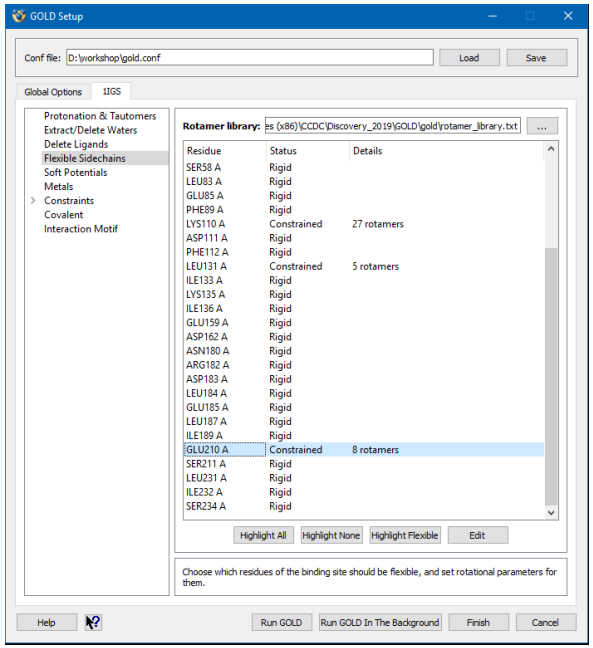






11. Repeat the same procedure for *LEU131* and *GLU210*. You should have a rotamer list as that showed on the left.

11



Setting the docking parameters

Selecting a fitness function

Return to the *Global Options* tab of the *GOLD Setup* window to proceed with the rest of the docking setup.

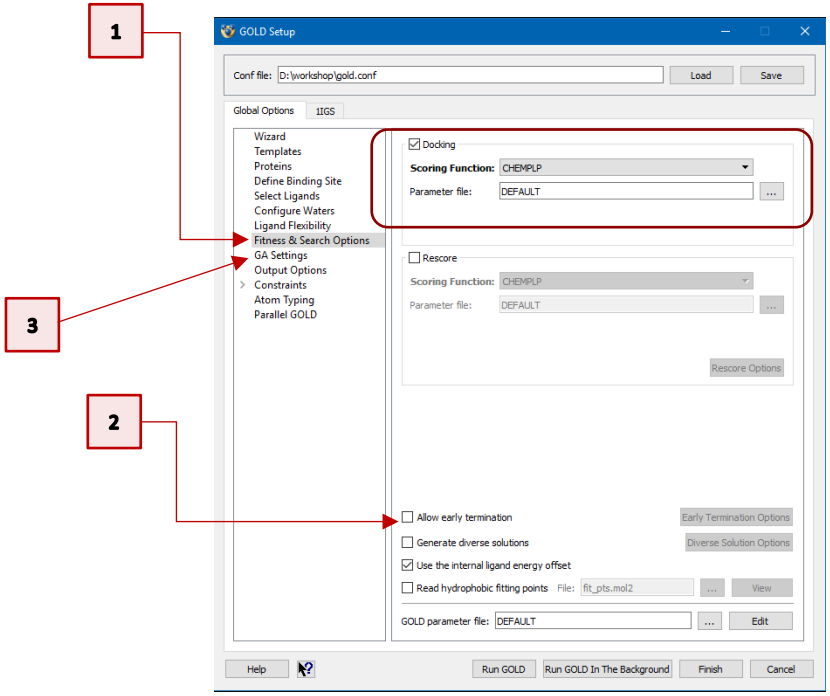
1. Click on the **Fitness & Search Options**. Ensure that the default *CHEMPLP* scoring function is selected under *Docking*.
2. By default, the **Allow early termination** check box is ticked. Switch it off by removing the tick next to *Allow early termination*. This will ensure that as many solutions as possible are explored.
3. Click on the **GA Settings** option.

Specifying the GA settings

GOLD optimises the fitness score using a genetic algorithm (GA). A number of parameters control the precise operation of this genetic algorithm. The 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.

4. Enable automatic (i.e. ligand-independent) GA settings by choosing the **Automatic** radio button and ensure the *Search efficiency* is set to 100%.



Running the docking

5. Still on the *Global Options* tab of the *GOLD Setup* window, select **Output Options**. Click on the **…** button next to *Output directory* and specify a directory to which you have write permissions; this is where the GOLD output files will be written (`results` in this case).

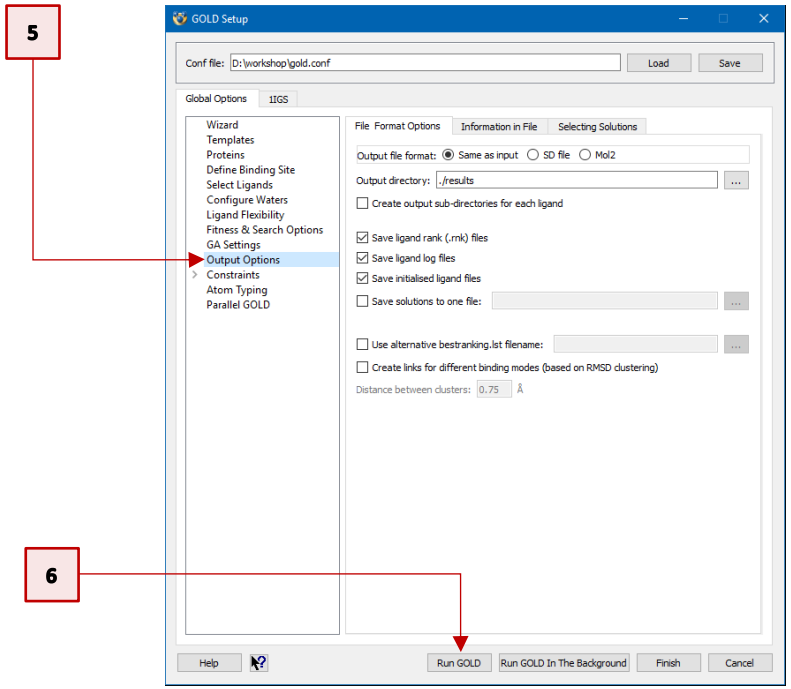
6. We have now finished setting up our docking, so click on the **Run GOLD** button at the bottom of the GOLD interface. You will be presented with a *Finish GOLD Configuration* window containing *Save Files* options:

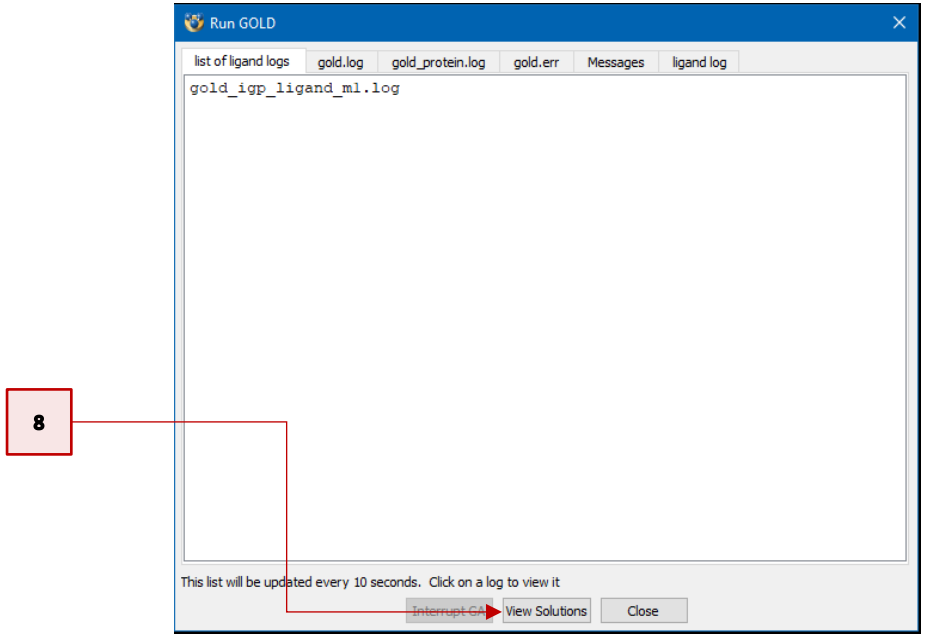
    a. Ensure the *GOLD conf file* tick-box is activated and rename this file as *gold_flexible.conf.*

    b. We did not edit the protein; therefore, you can leave the Protein(s) tick-box unticked. Click **Save** to start the docking.

7. As the job progresses output will be displayed in several tabs in the Run GOLD window.

8. Once the job is complete, load the docking results into Hermes by clicking on the **View Solutions** button in the *Run GOLD* window. We have finished with the *Run GOLD* window now; therefore, close the window by clicking on the **Close** button. In the *GOLD Setup* window, click on the **Cancel** button to close this window as it is also no longer needed.





Analysis of the results

The docking solutions are given in their docked order with their corresponding fitness score under the column headed *PLP.Fitness*. We have obtained 10 docking solutions as this is the default number of how many times (i.e. GA runs) our ligand was docked in the IGPS protein.

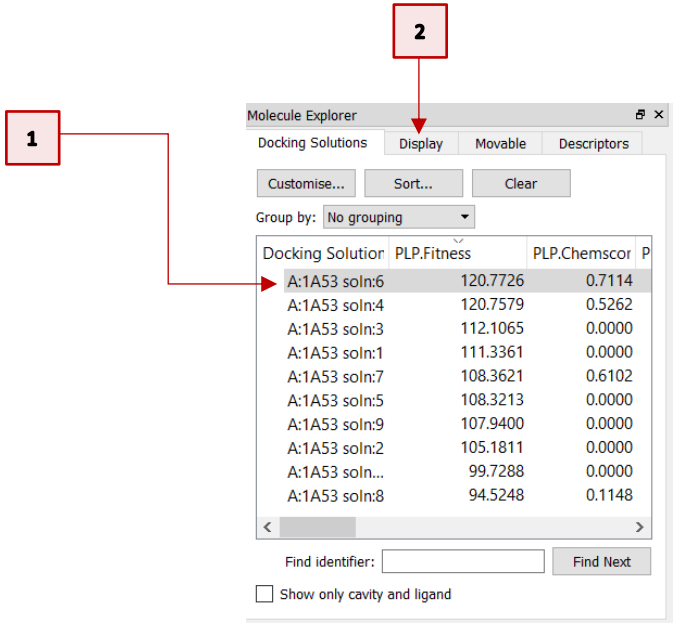
1. Sort the solutions based on their overall fitness score by clicking on the *PLP.Fitness* column header in the *Docking Solutions* tab of the *Molecule Explorer* window. Then click on the docking solution with the highest score. This will load the solution in the 3D view.

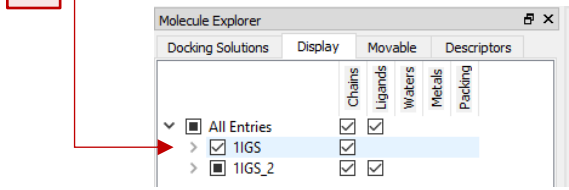
Please note: Due to the non-deterministic nature of GOLD your results may vary from those described in this workshop. In addition, the precise values of scoring function terms in your docking results will differ from the ones illustrated here.

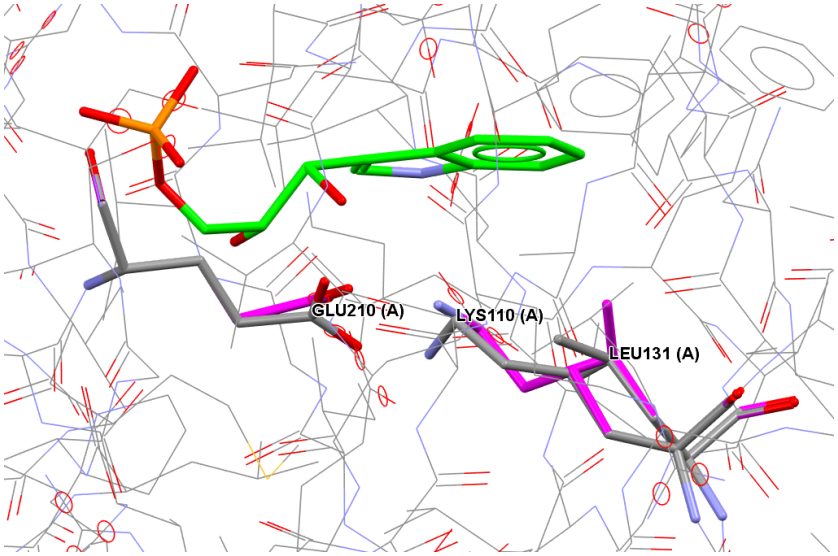
2. Click on the *Display* tab of the *Molecule Explorer* window to see all the entries available in the 3D view.

3. Show the original protein structure in the 3D view by ticking the box next to the *1IGS* entry, listed under *All Entries*.

4. Change the style and colour of carbon atoms of the of the side chain atoms of Lys110, Leu131 and Glu210 of both the original protein structure and of the best ranked docking pose. It will be clear that these moved in order to optimise the conformation of the ligand and its interactions with the binding site.







Generating SuperStar Maps using Hermes

Generate a propensity map using a carbonyl oxygen probe

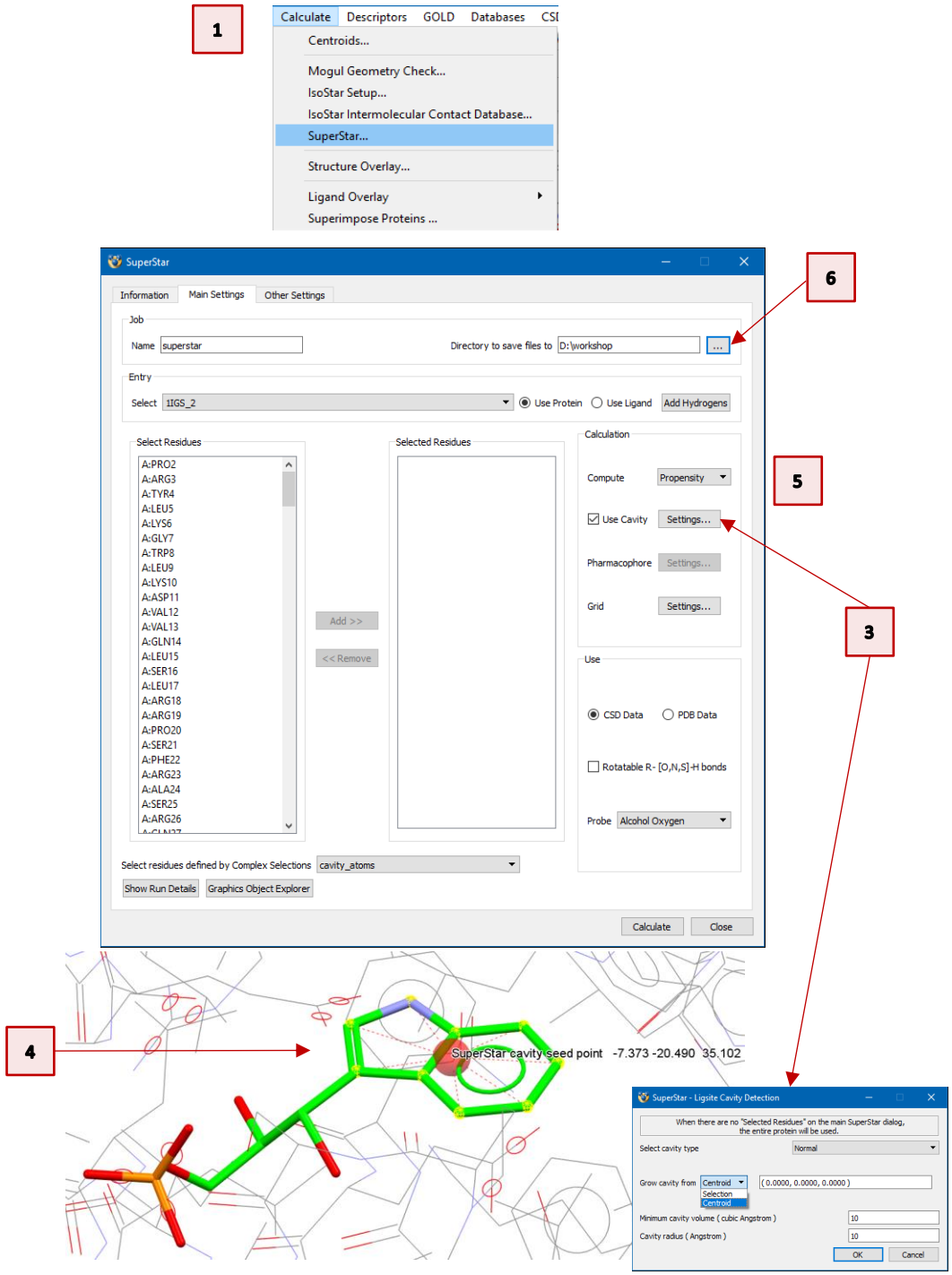
We will now use SuperStar to visually identify what are the most likely interaction hotspots within a protein's binding pocket in order to validate the best ranked docking pose.

1. Click on **Calculate** in the Hermes top-level menu and select **SuperStar** from the drop-down menu
2. Choose *1IGS_2* from the list of available options in *Use Protein* pop-up window and click **OK.** This will open the *SuperStar* window.
3. In the *SuperStar* window, click on **Settings** next to *Use Cavity*. In the *Ligsite Cavity Detection* dialogue, choose *Centroid* from the *Grow cavity from* drop-down menu.
4. Select the atoms of the indole ring of the IGP ligand in the visualiser. The centroid will be displayed in the visualiser and its XYZ coordinates will be automatically entered to the right of the *Grow cavity from* option. Leave all other settings as default. All residues included in the cavity to be detected from this centroid will then be used for the map calculation.
5. Make sure that **Propensity** is selected next to *Compute* in the *Calculation* section of the *SuperStar* window.

Propensities are measured relative to the expected (random) chance of finding a group at a certain position, e.g. a propensity of 4 indicates that the chance of finding the probe group at that point is 4 times as high as random. As a general rule, propensities of 2 and higher indicate favourable interaction sites (although, for hydrophobic probes, values between 1 and 2 can also be meaningful).

See the SuperStar User Guide for more details.

6. Select an output directory where you have write permissions, e.g. `workshop`.



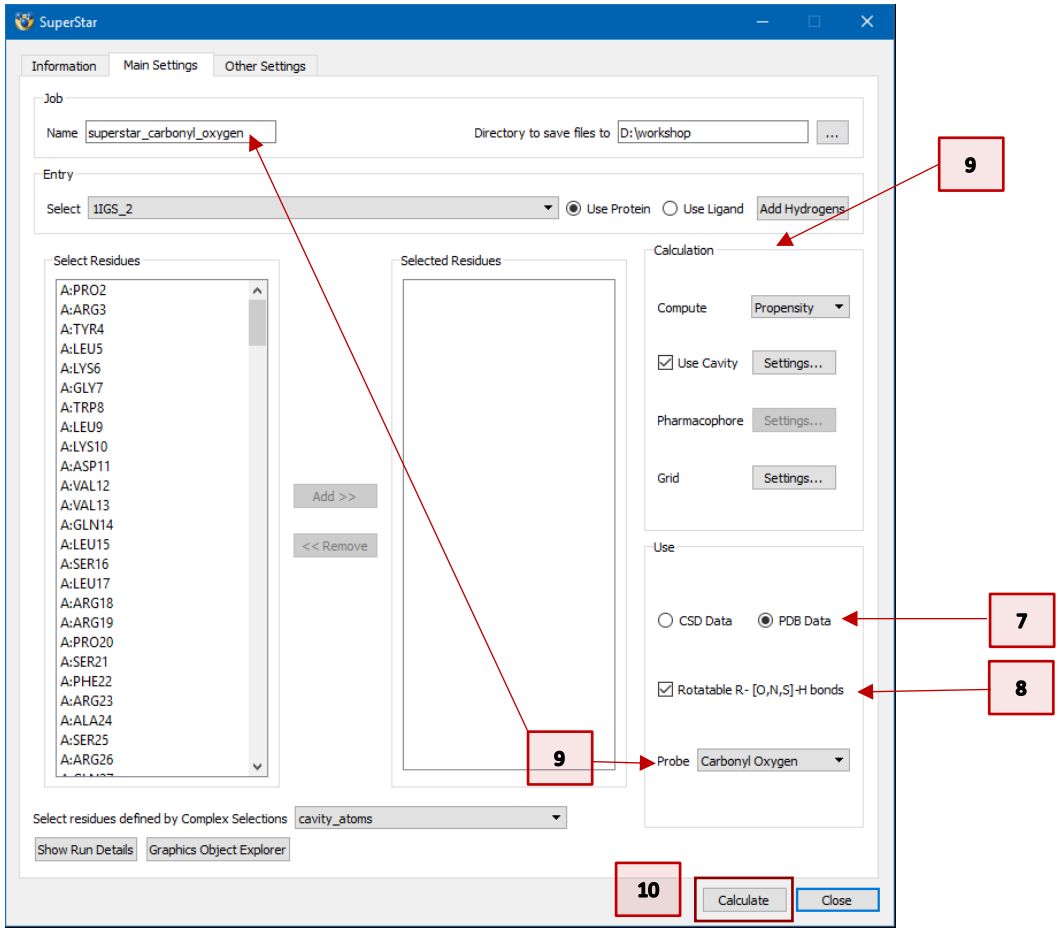
7. SuperStar allows data from either the Cambridge Structural Database (CSD) or the Protein Data Bank (PDB) to be used for propensity map calculations. The *Data source* to be used for the map calculation can be specified within the *Use* area of *SuperStar* window. Enable the **PDB Data** button to retrieve PDB-based scatterplot data for the compilation of the SuperStar map.

8. For protein residues, SuperStar can take into account the fact that OH, NH and SH groups may rotate. To enable this feature tick the **Rotatable R-[O,N,S]-H bonds** tick-box in the *Use* area of the *SuperStar* window. This feature applies to serine, threonine, tyrosine, lysine, and cysteine residues. If this check-box is not enabled, the -XH bond for the above residues will be treated as being fixed in space.

Before a SuperStar map can be calculated, a probe needs to be selected. Contouring will subsequently take place for one of the atoms of the chosen probe. The list of available probes for CSD and PDB data are listed in the **Probe** drop-down menu within the *Use* area of *SuperStar* window.

9. Select **Carbonyl Oxygen** as probe atom form the **Probe** drop-down menu. The carbonyl oxygen can accept hydrogen bonds, so favourable spots for acceptor groups in the ligand will be detected. Note that you may wish to change the default name of the job from `superstar` to e.g. `superstar_carbonyl_oxygen`.

10. Click **Calculate.** The resulting PDB-based map will be displayed as a 3D contoured map in the Hermes 3D view.



SuperStar propensity maps

Once the SuperStar maps have been created, their display can be controlled by the Hermes *Graphics Object Explorer*. This will be opened the first time a map is calculated. It can also be opened by clicking on the **Graphics Object Explorer** button at the base of the *SuperStar* window. In addition, it can be opened from the **Display** drop-down menu in the top-level Hermes menu bar.

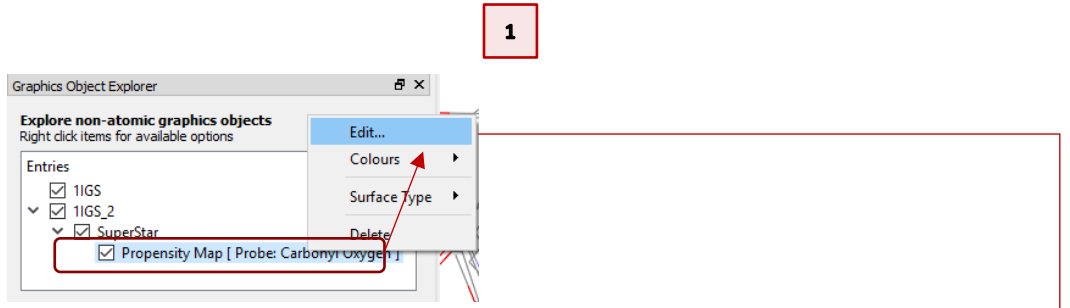
Customising SuperStar maps

1. Right-click on **Propensity Map [Probe: Carbonyl Oxygen]** in the *Graphics Objects Explorer* window and then choose **Edit.** This will open the *Graphics Object Editor Dialog* window. For each contour surface listed, it is possible to edit the propensity level, colour, display type, opacity and visibility. Here, three propensity maps are listed, corresponding to the maps contoured at levels 2.0 (red) and 4.0 (green) and 8.0 (blue). These SuperStar contour maps depict the spatial distribution of propensities.

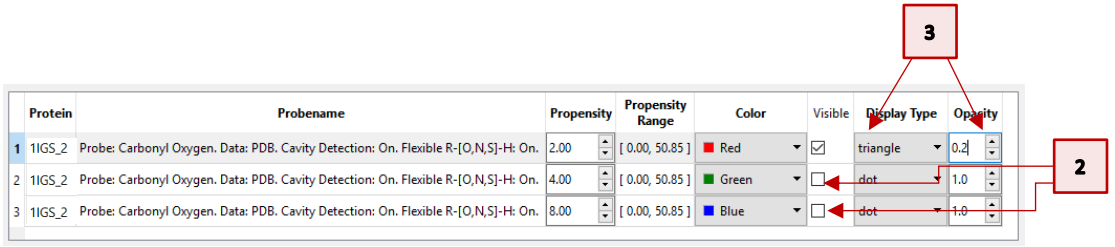
2. From the *Graphics Object Editor Dialog* window, turn-off the second and third propensity map (corresponding to **Propensity** of *4.0* and *8.0*) by disabling their tick-boxes in the **Visible** column.

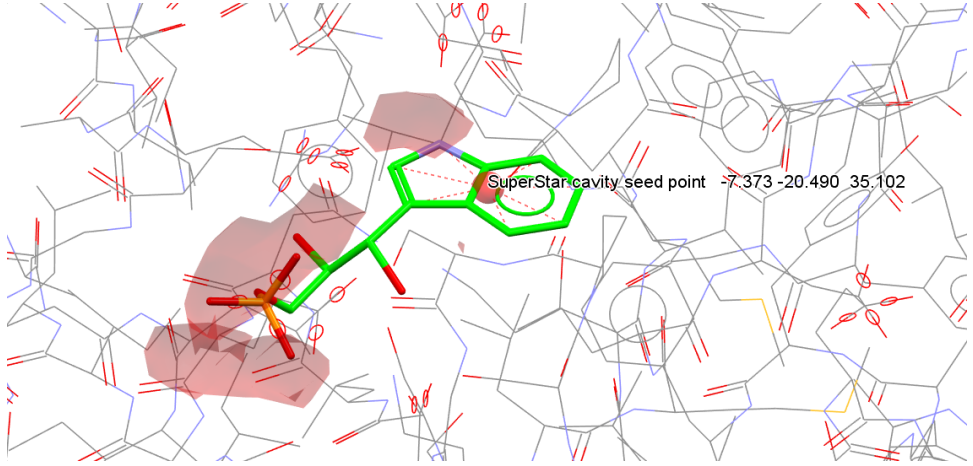
3. Customise the first propensity map, corresponding to the map with **Propensity** of *2.0*, by changing its **Display Type** from *dot* to *triangle* using the drop-down menu, and then changing its **Opacity** from *1.0* to *0.2* using the spin-box.

This will result in the Hermes 3D view showing only the map of propensity 2 for carbonyl oxygen probe, indicating that the chance of finding this probe atom at that point is twice as high as random.



	Protein	Probename	Propensity	Propensity Range	Color	Visible	Display Type	Opacity
1	1IGS_2	Probe: Carbonyl Oxygen. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: On.	2.00	[0.00, 50.85]	Red	☑	dot	1.0
2	1IGS_2	Probe: Carbonyl Oxygen. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: On.	4.00	[0.00, 50.85]	Green	☑	dot	1.0
3	1IGS_2	Probe: Carbonyl Oxygen. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: On.	8.00	[0.00, 50.85]	Blue	☑	dot	1.0

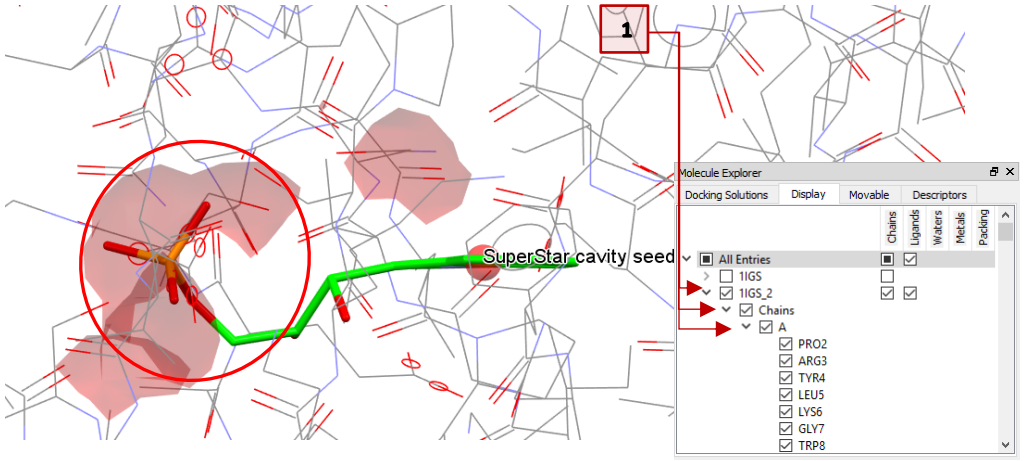


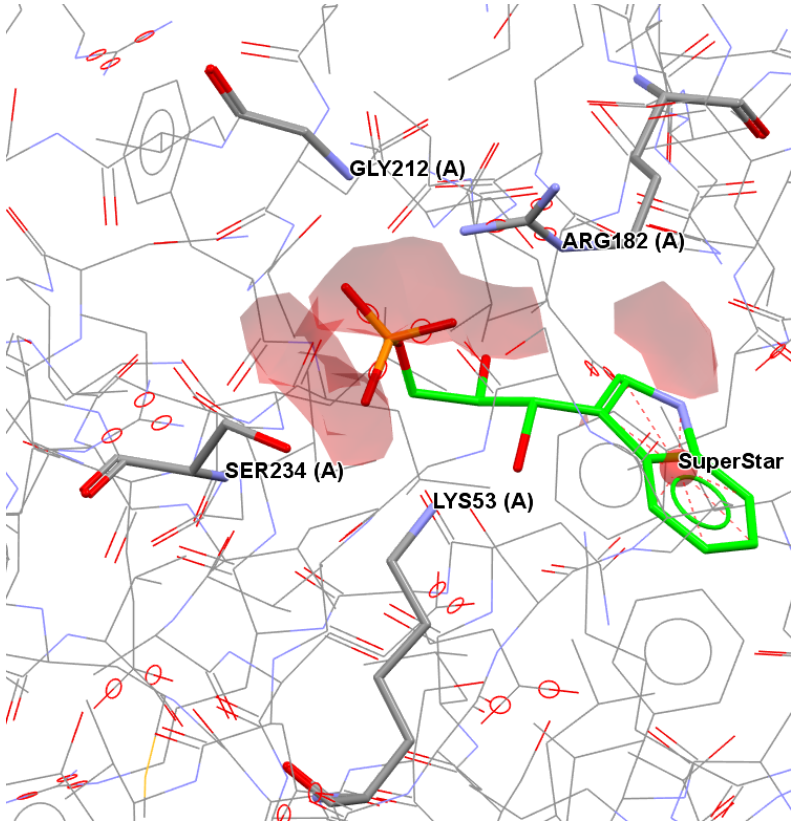


Analysing SuperStar propensity map derived from a carbonyl oxygen probe

The propensity map derived by using the carbonyl oxygen as probe reveals that the triosephoshate moiety of the IGP ligand matches very well in the H-bond acceptor areas predicted by SuperStar (see red circle in the figure on the right).

1. In Hermes, in the *Molecule Explorer* window click on the "*>*" adjacent to 1IGS_2 and underneath *All Entries*. Then click on "*>*" adjacent to Chains and the "*>*" adjacent to *A.* This will show the list of all the residues of the protein.
2. Scroll through the list in *Molecule Explorer* and select LYS53, ARG182, GLY212 and SER234. Multi-selection is possible using **CTRL + left-click**.
3. Right-click in the *Molecule Explorer* area and select **Styles** > **Capped Sticks** from the pull-down menu. This will show the selected residues as capped sticks in the Hermes 3D view.
4. The predicted docked pose of compound IGP in the protein binding site shows that there are several key hydrogen bond interactions between the triosephoshate moiety of IGP (as H-bond acceptors) and the backbone NH group of Gly212 & the terminal OH/NH groups of the Lys53, Arg182 and Ser234 side chains (as H-bond donors) in the protein binding site (see figure on the right, H atoms not displayed for clarity).

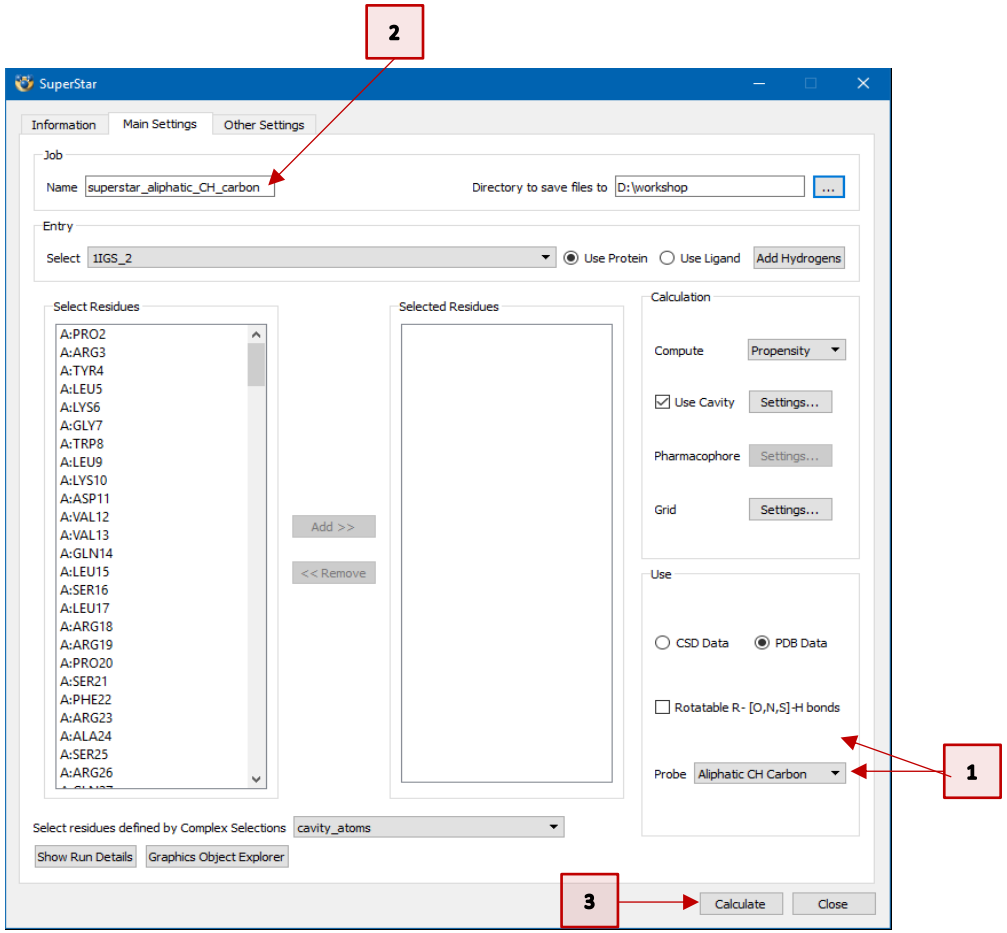


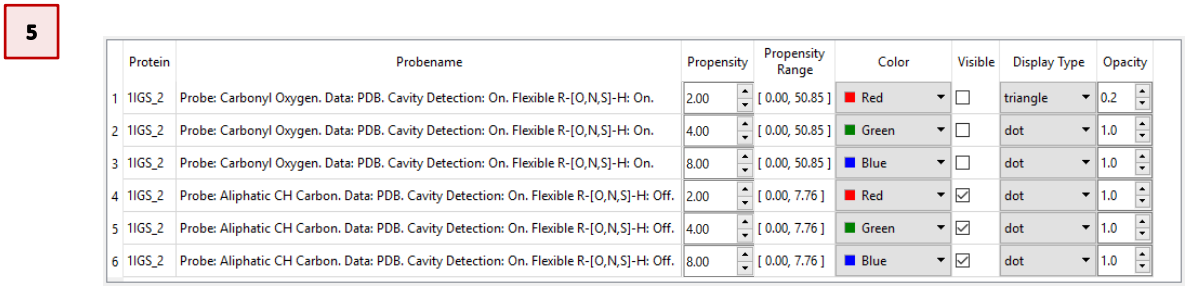


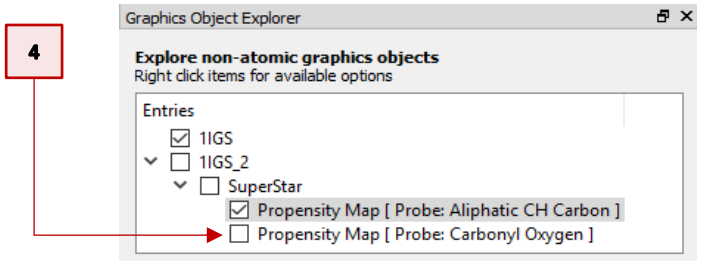
Generate a propensity map using an aliphatic CH carbon probe

We are now going to explore the hydrophobic interactions in the protein binding site. To do so, return to the *SuperStar* window. If you have closed it, click on **Calculate** > **SuperStar** in the Hermes top-level menu.

1. Select **Aliphatic CH Carbon** from the **Probe** drop-down menu, and untick the **Rotatable R-[O,N,S]-H bonds** option.
2. From the *Job* area of the *SuperStar* window, change the name of the job to e.g. *superstar_aliphatic_CH_carbon*; this will ensure that the previous maps are not overwritten.
3. Click **Calculate.** The resulting PDB-based map will be displayed as a 3D contoured map in the Hermes 3D view.
4. Hide the propensity map previously obtained using the carbonyl oxygen probe by unticking the box for **Propensity Map [Probe: Carbonyl Oxygen]** in the *Graphics Object Explorer* window.
5. Right-click on **Propensity Map [Probe: Aliphatic CH Carbon]** in the *Graphics Object Explorer* window and then choose **Edit.** This will open the *Graphics Object Editor Dialog* window. This time, six propensity maps are listed. Note that the maps generated with the carbonyl oxygen probe have the **Visible** tick-box disabled.





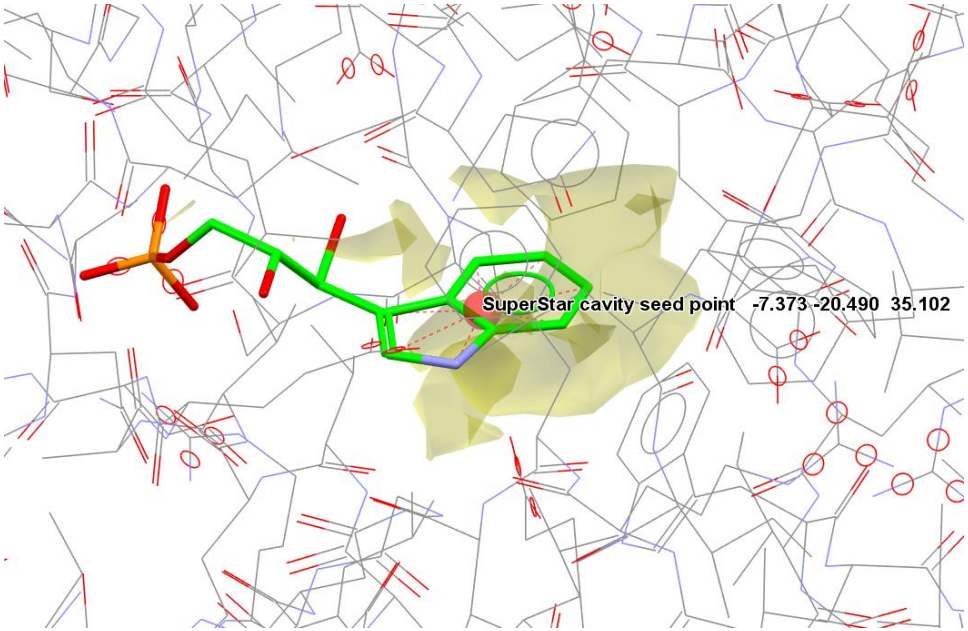


6. As we did for the carbonyl oxygen probe, from the *Graphics Object Editor Dialog* window, turn-off the propensity maps for the aliphatic CH carbon probe corresponding to **Propensity** of *4.0* and *8.0*) (rows 5 and 6 in the figure above).

7. Customise the map for the aliphatic CH carbon probe with a **Propensity** of *2.0* in this manner:

    - Change its **Display Type** by clicking on *dot* and select *triangle* from the drop-down menu.
    - Then change the **Opacity** from *1.0* to *0.2* using the **Opacity** spin-box.
    - Then change the colour of the contour surface by selecting **Yellow** from the **Color** drop-down menu.

This will result in the Hermes 3D view showing only the map of propensity 2 for the aliphatic CH carbon probe, indicating that the chance of finding this probe atom at that point is twice as high as random.



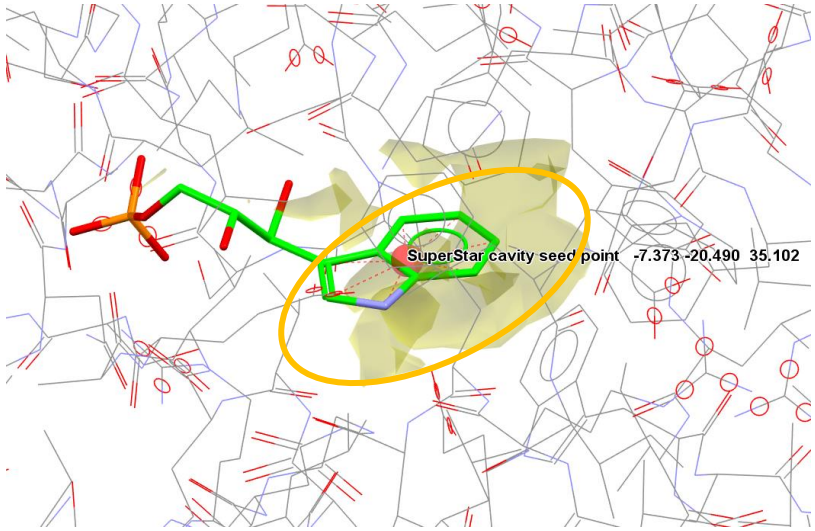
Analysing the SuperStar propensity map derived from an aliphatic CH carbon probe

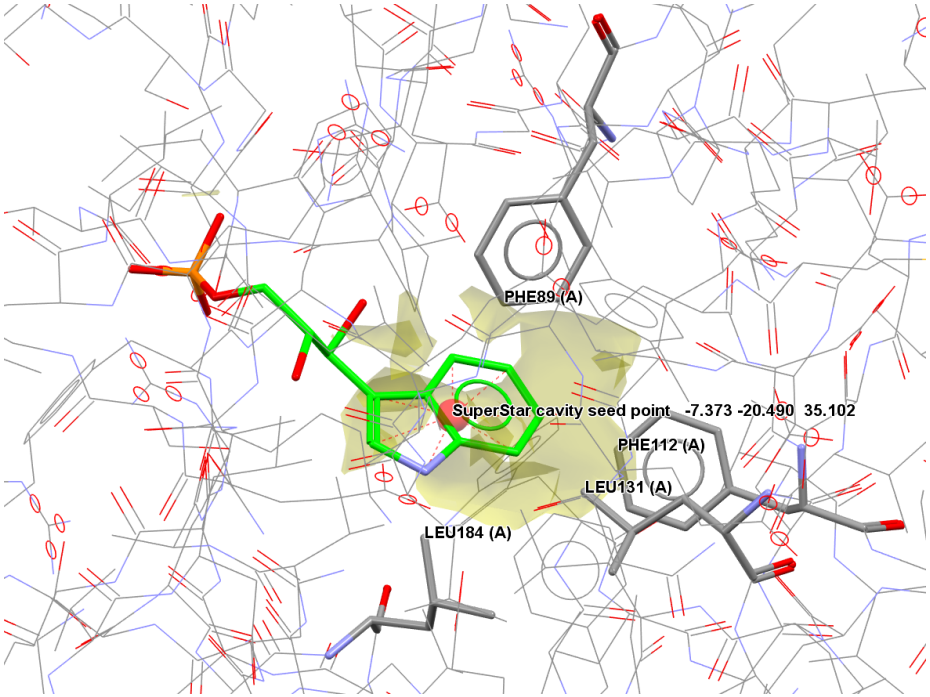
The propensity map derived by using the aliphatic CH carbon probe reveals that the docked pose of compound IGP has the benzene ring of the indole moiety that matches very well in the hydrophobic areas predicted by SuperStar (see orange circle in the figure on the right).

8. From the *Molecule Explorer* window, select PHE89, PHE112, LEU131 and LEU184. Multi-selection is possible using **CTRL + left-click**.

9. Right-click in the *Molecule Explorer* area and select **Styles** > **Capped Sticks** from the pull-down menu. This will show the selected residues as capped sticks in the Hermes 3D view.

The benzene ring of the indole moiety fills a hydrophobic region of the IGPS binding site (see figure on the right).

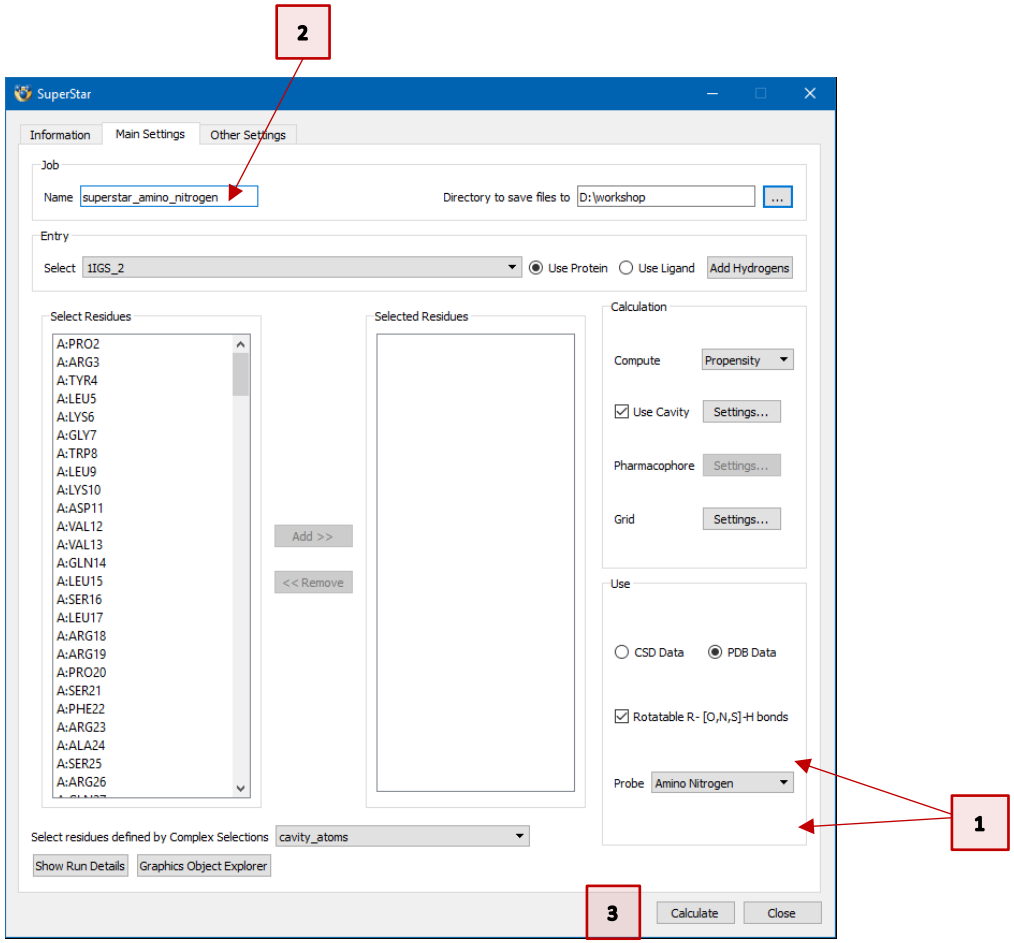




Generate a propensity map using an amino nitrogen probe

We are now going to use an amino nitrogen probe to point out potential H-bond donor areas of the binding site of the IGPS protein. To do so, return to the *SuperStar* window; if you have closed it, click on **Calculate** > **SuperStar** in the Hermes top-level menu.

1. Select **Amino Nitrogen** from the **Probe** drop-down menu and tick the option for **Rotatable R-[O,N,S]-H bonds**.
2. From the *Job* area of the *SuperStar* window, change the name of the job to e.g. *superstar_amino_nitrogen*; this will ensure that the previous maps are not overwritten.
3. Click **Calculate.** The resulting PDB-based map will be displayed as a 3D contoured map in the Hermes 3D view.
4. As in the previous section of the workshop, hide the propensity map obtained using the aliphatic CH carbon probe by disabling the relevant propensity maps in the *Graphics Object Explorer* window, and then edit the propensity maps obtained with the amino nitrogen probe in this manner:
    - Turn off the propensity maps of the amino nitrogen probe corresponding to **Propensity** of *4.0* and *8.0*) (rows 8 and 9 in the figure on the right).
    - Customise the map corresponding to **Propensity Map [Probe: Amino Nitrogen]** with **Propensity** of *2.0***.**
    - Change its **Display Type** by clicking on *dot* and selecting *triangle* from the drop-down menu.
    - Change its **Opacity** from *1.0* to *0.2* using the **Opacity** spin- box.
    - Finally, change the colour of the propensity map by selecting **Cyan** from the **Color** drop-down menu.



	Protein	Probename	Propensity	Propensity Range	Color	Visible	Display Type	Opacity
1	1IGS_2	Probe: Carbonyl Oxygen. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: On.	2.00	[0.00, 50.85]	Red	☐	triangle	0.2
2	1IGS_2	Probe: Carbonyl Oxygen. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: On.	4.00	[0.00, 50.85]	Green	☐	dot	1.0
3	1IGS_2	Probe: Carbonyl Oxygen. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: On.	8.00	[0.00, 50.85]	Blue	☐	dot	1.0
4	1IGS_2	Probe: Aliphatic CH Carbon. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: Off.	2.00	[0.00, 9.54]	Yellow	☐	triangle	0.2
5	1IGS_2	Probe: Aliphatic CH Carbon. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: Off.	4.00	[0.00, 9.54]	Green	☐	dot	1.0
6	1IGS_2	Probe: Aliphatic CH Carbon. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: Off.	8.00	[0.00, 9.54]	Blue	☐	dot	1.0
7	1IGS_2	Probe: Amino Nitrogen. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: On.	2.00	[0.00, 33.22]	Cyan	☑	triangle	0.2
8	1IGS_2	Probe: Amino Nitrogen. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: On.	4.00	[0.00, 33.22]	Green	☐	dot	1.0
9	1IGS_2	Probe: Amino Nitrogen. Data: PDB. Cavity Detection: On. Flexible R-[O,N,S]-H: On.	8.00	[0.00, 33.22]	Blue	☐	dot	1.0

Analysing SuperStar propensity map derived from an amino nitrogen probe

The propensity map derived by using an amino nitrogen probe reveals that the docked pose of compound IGP has the nitrogen atom of the indole moiety and the hydroxy group closest to the indole matching the H-bond donor areas predicted by SuperStar (see blue circles in the figure on the right).

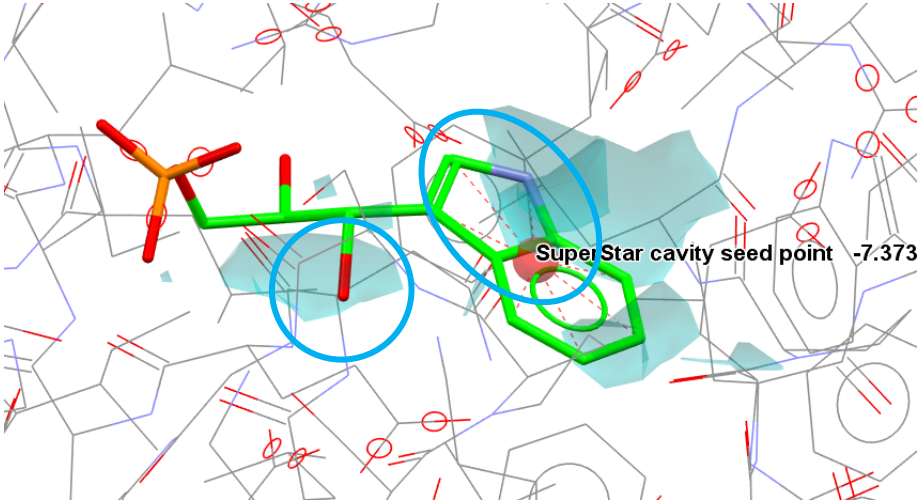
5. From the *Molecule Explorer* window, select GLU51, LYS53, LYS110, GLU159 and GLU210. Multi-selection is possible using **CTRL + left-click**.

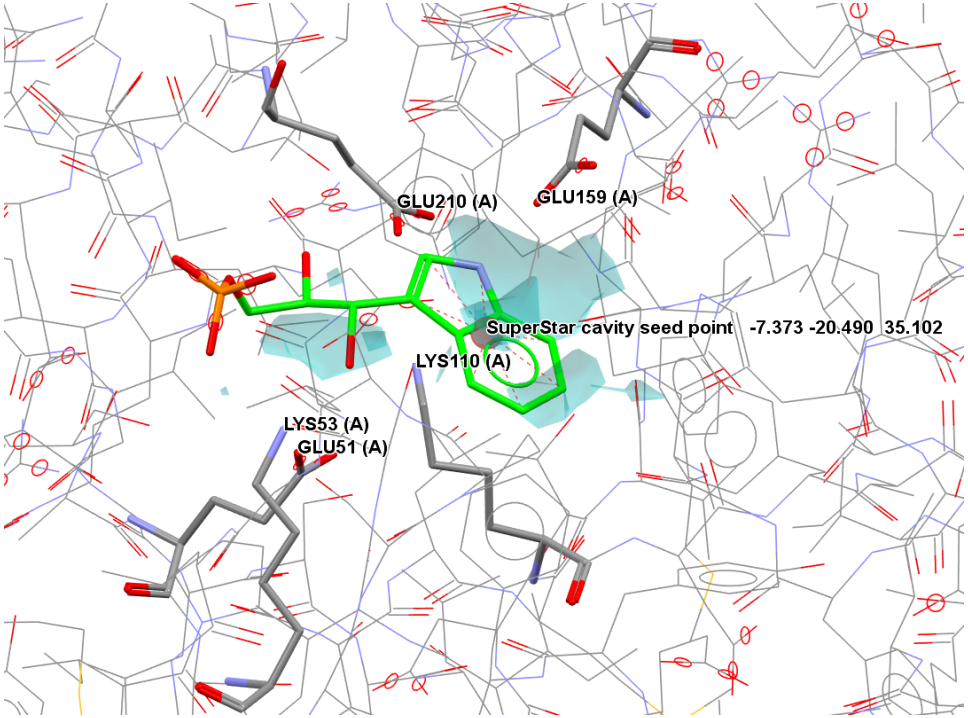
6. Right-click in the *Molecule Explorer* area and select **Styles** > **Capped Sticks** from the pull-down menu. This will show the selected residues as capped sticks in the Hermes 3D view.

A possible triple salt-bridge cluster is possible between the hydroxyl group and the Glu51, Lys53 and Lys110. In addition, the NH group of the indole moiety is favourably oriented to hydrogen bond to Glu159 and Glu210.

Note that by setting the Lys110 and Glu210 side chains as flexible, we allowed GOLD to find a docking solution that allows the formation of these H-bonds (see the figure on the right).

<u>Do not close your Hermes session.</u>



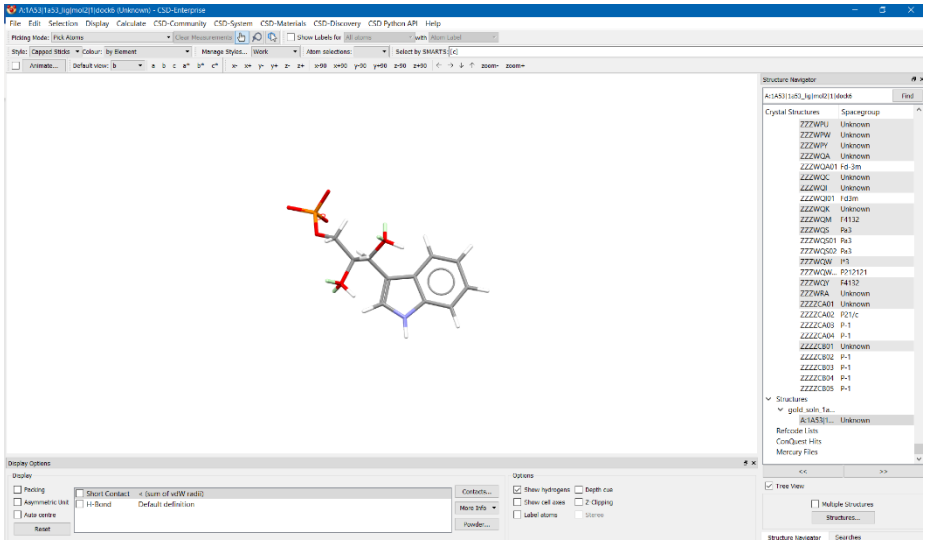


Generating Full Interaction Maps (FIMs) using Mercury

We will now use FIMs to visually identify what are the most likely interaction preferences of the best ranked docking pose obtained from the GOLD calculation. FIMs allows you to assess how the predicted protein binding site can satisfy the interaction preferences of the functional groups in the IGP compound.

Do not close your Hermes session.

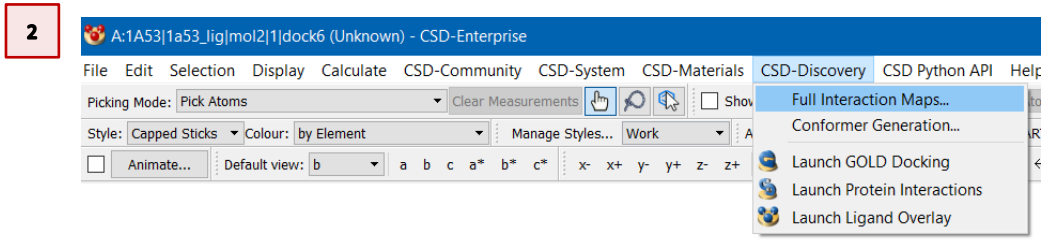
1. To access Full Interaction Maps, launch Mercury by clicking on its icon. Click on **File** and select **Open…** from the pull-down menu to load the best ranked docking pose obtained in your GOLD calculation. You will find the gold_soln_igp_ligand_m1_x.mol2 files for the docking poses in the folder where you have saved your docking solutions (i.e. `results`).



In the example showed here, the best ranked docking pose corresponds to pose number 6 (i.e. `gold_soln_igp_ligand_m1_6.mol2`).

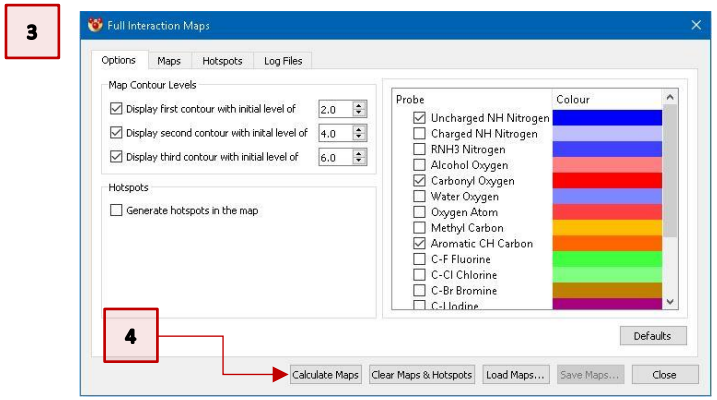
Note that your GOLD results may vary from those described in this workshop document and your best ranking pose could correspond to a different number.

2. Click on the **CSD-Discovery** in the Mercury top-level menu and select **Full Interaction Maps…**



Note that Full Interaction Maps is also available under the **CSD-Materials** menu of Mercury.

3. In the *Full Interaction Maps* window, you will see several options. On the left you will find options to change the display contour levels. On the right, you will see a list of functional groups to be used as probes. For the purposes of this workshop, we will keep the default options.



The default options typically work well for most situations, but if you wish to look for the interaction preferences of the molecule with a specific functional group, or if you want to change the look of the map, you will want to edit these settings.

4. Click on **Calculate Maps** in the *Full Interaction Maps* window.

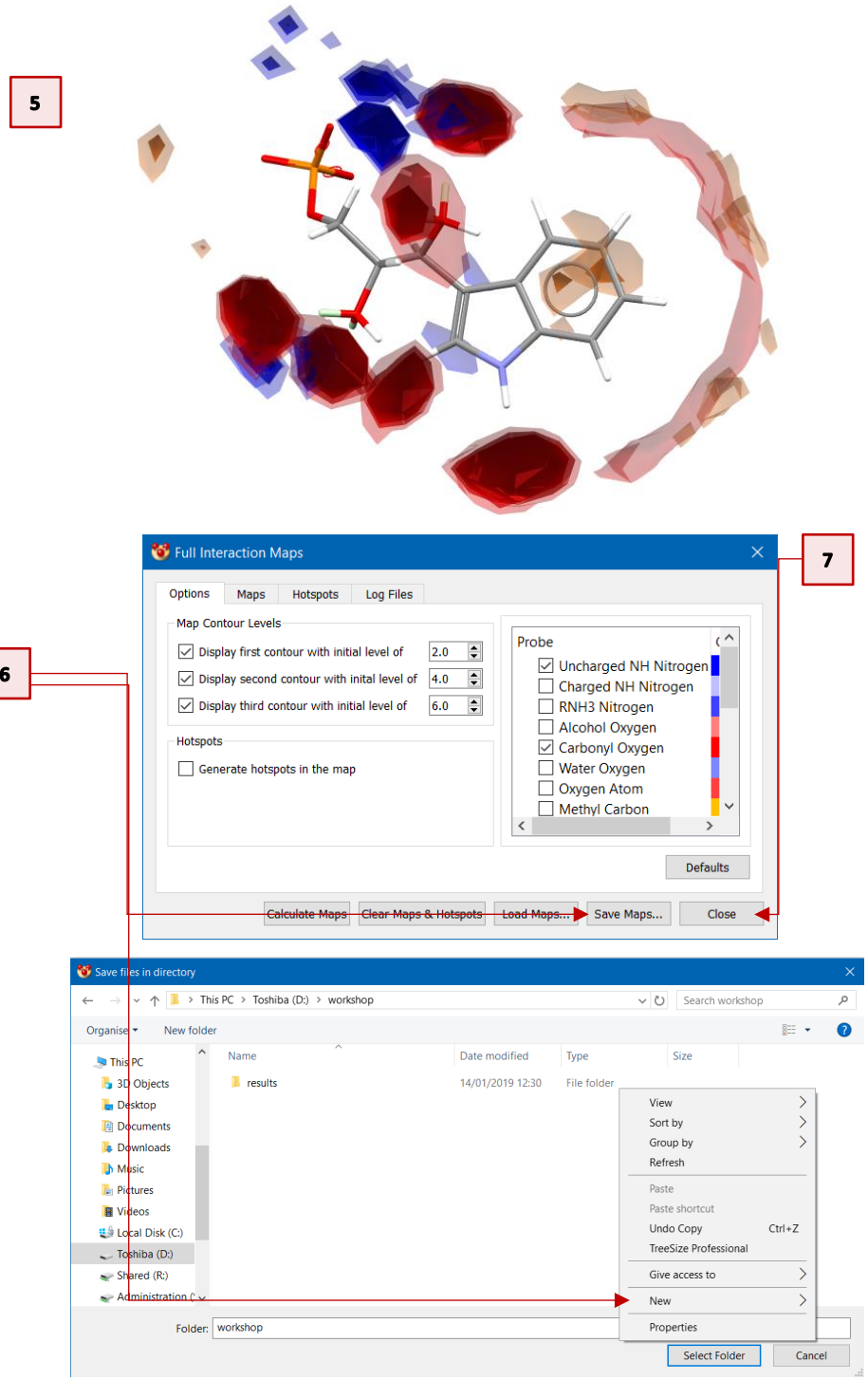
5. The generated maps will now be displayed in the main Mercury window. Notice that:

    - Red regions of the maps denote areas in which there is a high probability of locating a hydrogen bond acceptor.
    - Blue regions denote the preferred locations of an H-bond donor.
    - Brown regions indicate hydrophobic preferences.

We are now going to save these generated maps in order to load them into Hermes. This will allow us to see these maps in the context of the protein binding site.

6. Click on **Save Maps…**in the *Full Interaction Maps* window. Note that the **Save Maps…** option requires the selection (or creation) of an empty directory to save into. Navigate to your previously created `workshop` folder and use the right-click menu to create a new folder (e.g. `FIMs`). Then click on **Select Folder.** The generated maps will then be saved in your defined folder.

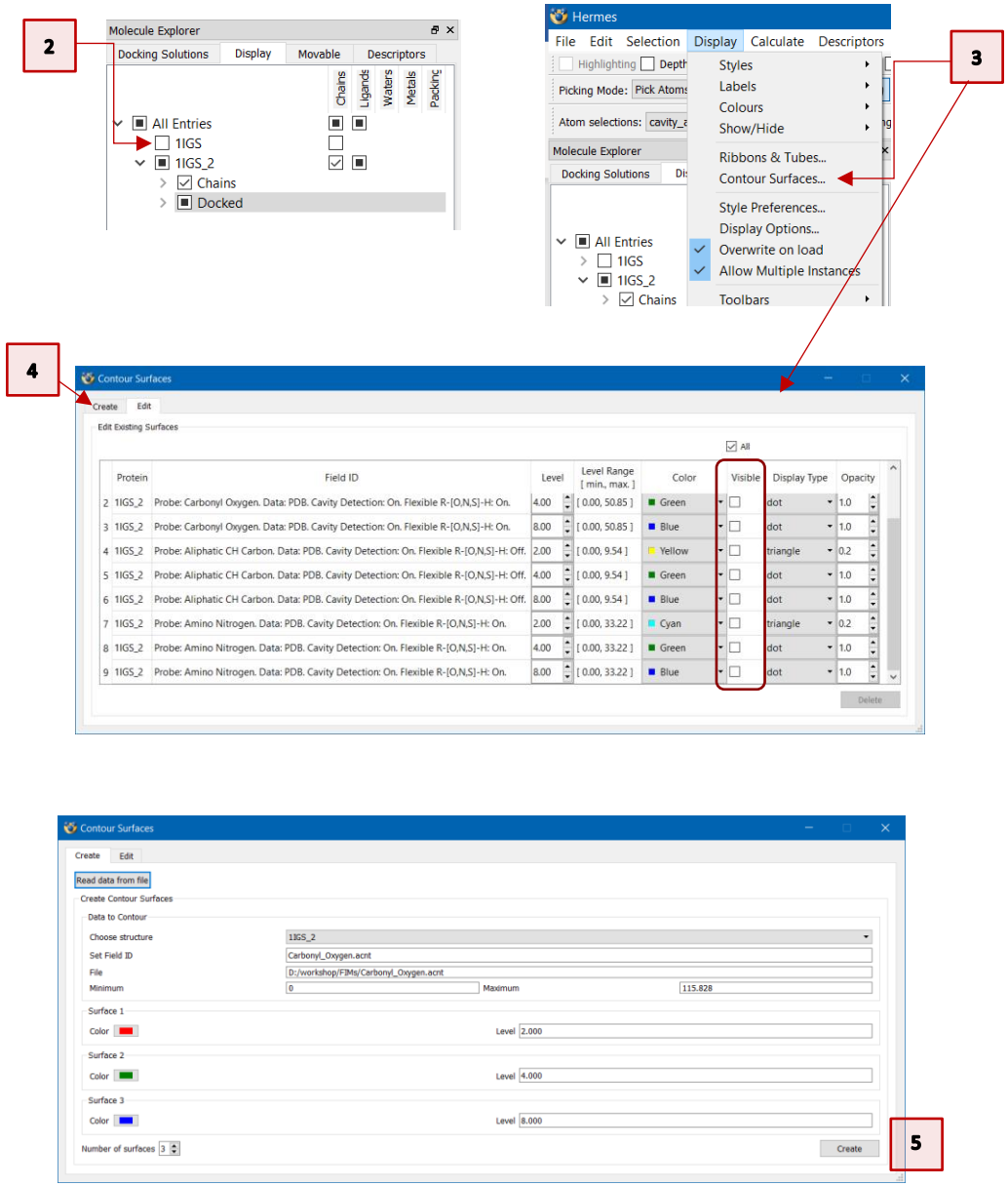
7. Click on **Close** in the *Full Interaction Maps* window and you can also close Mercury if you wish by clicking on **File** and then select **Exit** from the pull-down menu.



Analysing the Full Interaction Maps in Hermes

To analyse the FIMs of the best ranking docking pose we will load the saved maps into the Hermes session.

1. Return to your Hermes session.
2. Hide the original 1IGS protein by unticking the 1IGS tick-box in the *Molecule Explorer* window.
3. Click on **Display** and select **Contour Surfaces** from the pull-down menu. In the resulting *Contour Surfaces* window you will see all the propensity maps that we have generated with SuperStar. Ensure that all propensity maps are hidden by disabling all tick-boxes in the **Visible** column.
4. Click on the **Create** tab of the *Contour Surfaces* window and then click on **Read data from file.** Navigate to the folder where you saved your FIMs and select the *Carbonyl_Oxygen.acnt* file to open.
5. Click on **Create** in the *Contour Surfaces* window. This will load the full interaction map of the best scored docking pose obtained using a carbonyl oxygen as probe.
    a. Turn off the FIMs of the carbonyl oxygen probe corresponding to **Propensity** of *2.0* and *8.0* (rows 10 and 12 in the figure on the right).
    b. Customise the map corresponding to **Carbonyl_Oxygen.acnt** with **Propensity** of *4.0***.**
    c. Change its **Display Type** by clicking on *dot* and selecting *triangle* from the drop-down menu.
    d. Change its **Opacity** from *1.0* to *0.2* using the **Opacity** spin-box and choose **Red** in the **Color** drop-down menu.



6. You should have the *Contour Surfaces* window set as on the figure shown on the right.

7. Repeat steps 4 and 5 for the other two maps saved in the `FIMs` folder (i.e. *Aromatic_CH_Carbon.acnt* and *Uncharged_NH_Nitrogen.acnt*). For the **Aromatic_CH_Carbon.acnt** with **Propensity** of *2.0* set **Yellow** from the **Color** drop-down menu and for **Uncharged_NH_Nitrogen.acnt** with **Propensity** of *4.0* set the colour **Cyan.**

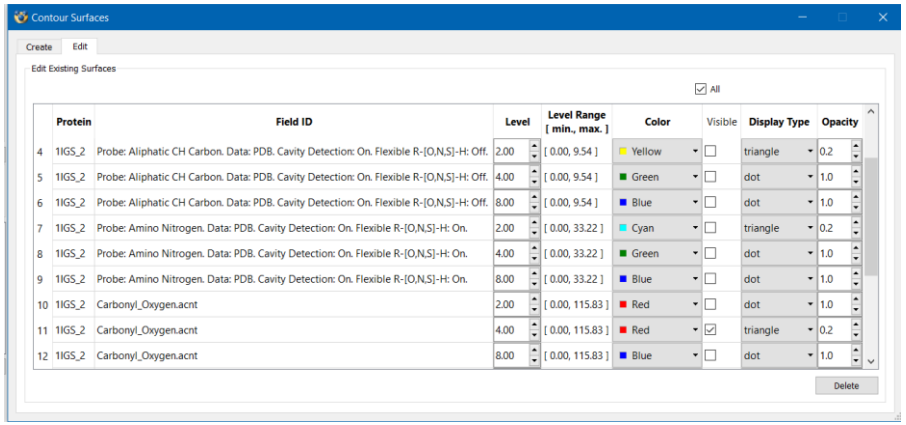
8. Your *Contour Surfaces* window should look like the one shown here on the right.

We will now analyse the interaction maps obtained using the different probes.

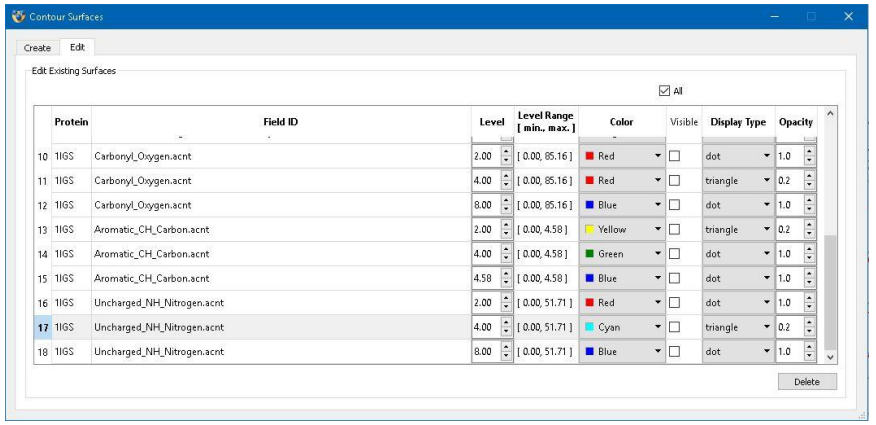
Analysing the full interaction map derived from a carbonyl oxygen probe

9. Hide the FIM of the aromatic CH carbon and that of the uncharged NH nitrogen by disabling the **Acnt grid [Aromatic_CH_Carbon.acnt]** and **Acnt grid [Uncharged_NH_Nitrogen.acnt]** tick-boxes in the *Graphics Object Explorer* window. You will now have only the full interaction map derived by using the carbonyl oxygen probe displayed in the Hermes 3D view.

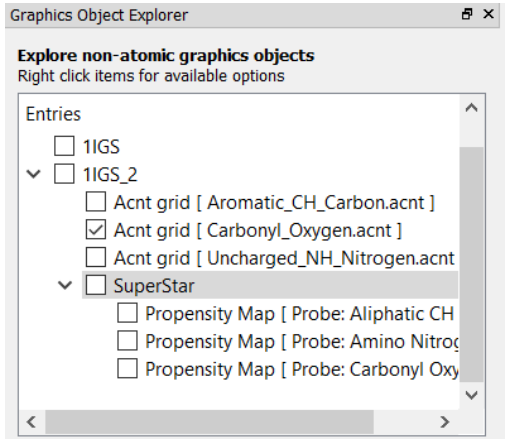
6



8



9



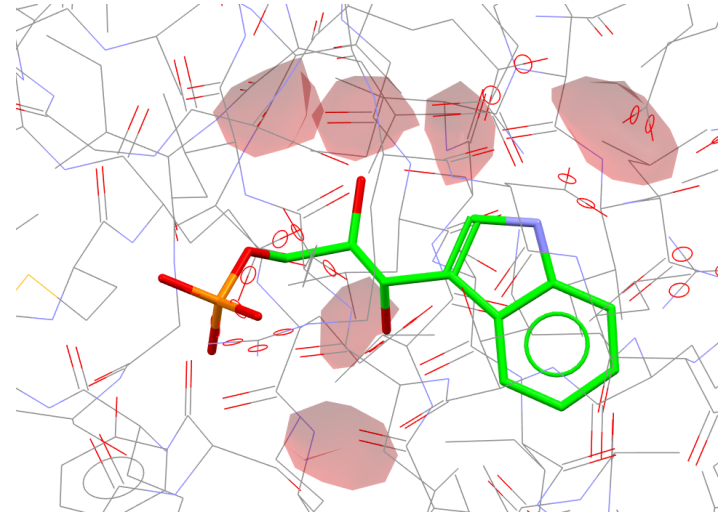
The interaction map derived by using the carbonyl oxygen probe reveals that the binding site of the IGPS protein has several residues that match very well in the H-bond acceptor areas around the ligand predicted by the Full Interaction Maps calculation.

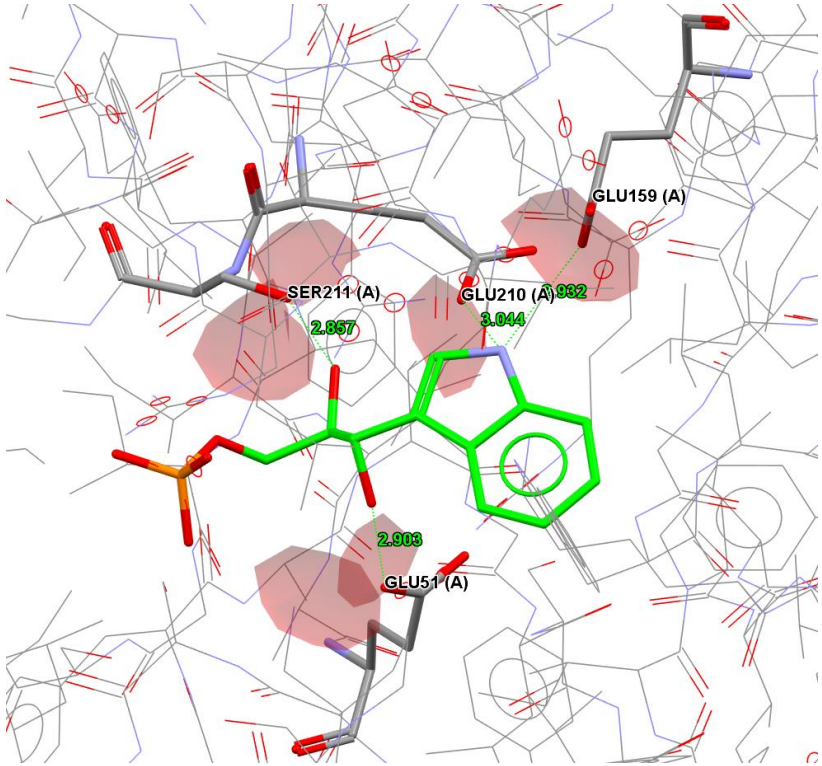
10. From the *Molecule Explorer* window, select GLU51, GLU159, GLU210 and SER211. Multi-selection is possible using **CTRL + left-click**.

11. Right-click in the *Molecule Explorer* area and select **Styles** > **Capped Sticks** from the pull-down menu. This will show the selected residues as capped sticks in the Hermes 3D view.

Most of the predicted H-bond acceptor interaction sites of the best scored docking pose of compound IGP are satisfied by the conformation of the ligand in the protein binding site. Several key hydrogen bond interactions can be formed:

- The indole moiety of the IGP ligand can act as a H-bond donor and form H-bonds with the sidechains of Glu210 and Glu159.
- The hydroxyl groups of the IGP ligand can act as a H-bond donor and form H-bonds with the sidechains of Glu51 and Ser211 (see figure on the right, H atoms not displayed for clarity).

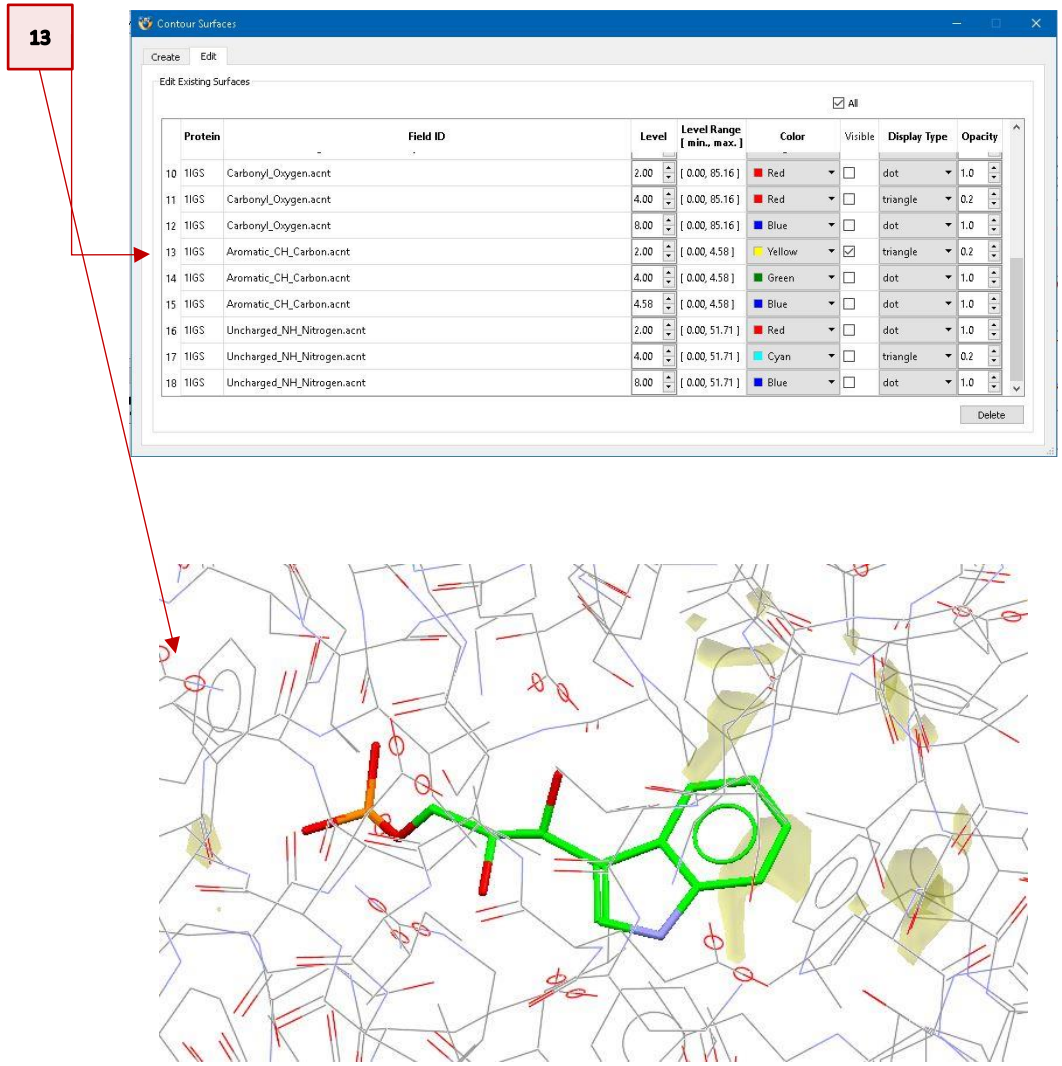




Analysing the full interaction map derived from an aromatic CH carbon probe

12. Hide the grid map of the carbonyl oxygen probe by disabling the **Carbonyl_Oxygen.acnt** tick-box in the **Visible** column of the *Contour Surfaces* window (row 11 in the figure on the right).
13. Now show the grip map of the aromatic CH carbon probe by enabling the tick-box corresponding to **Aromatic_CH_carbon.acnt** with **Propensity** of *2.0* (row 13 in the figure on the right)

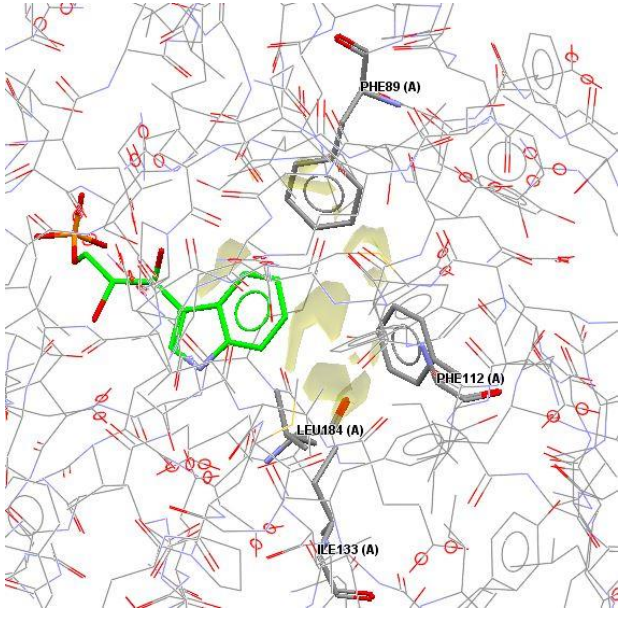
You will now have only the full interaction map derived by using the aromatic CH carbon probe displayed in the Hermes 3D view.



14. From the *Molecule Explorer* window, select PHE89, PHE112, ILE133 and LEU184. Multi-selection is possible using **CTRL + left-click**.

15. Right-click in the *Molecule Explorer* area and select **Styles** > **Capped Sticks** from the pull-down menu. This will show the selected residues as capped sticks in the Hermes 3D view.

The hydrophobic interaction preferences around the IGP molecule show that the benzene ring of the indole moiety fills a hydrophobic region of the IGPS binding site (see figure on the right).

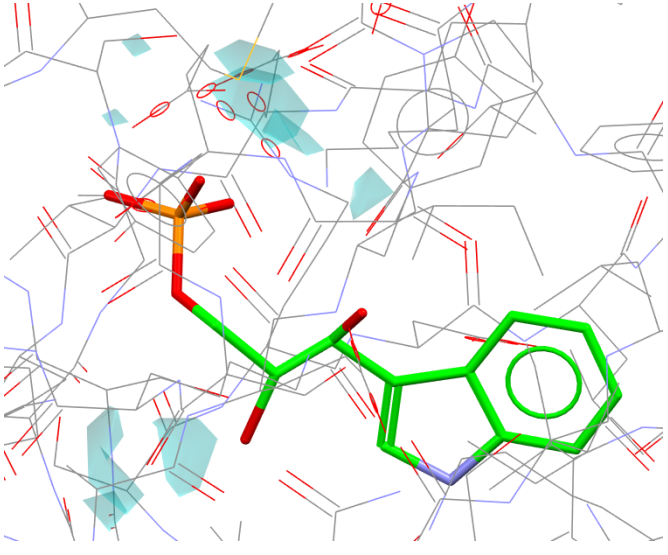


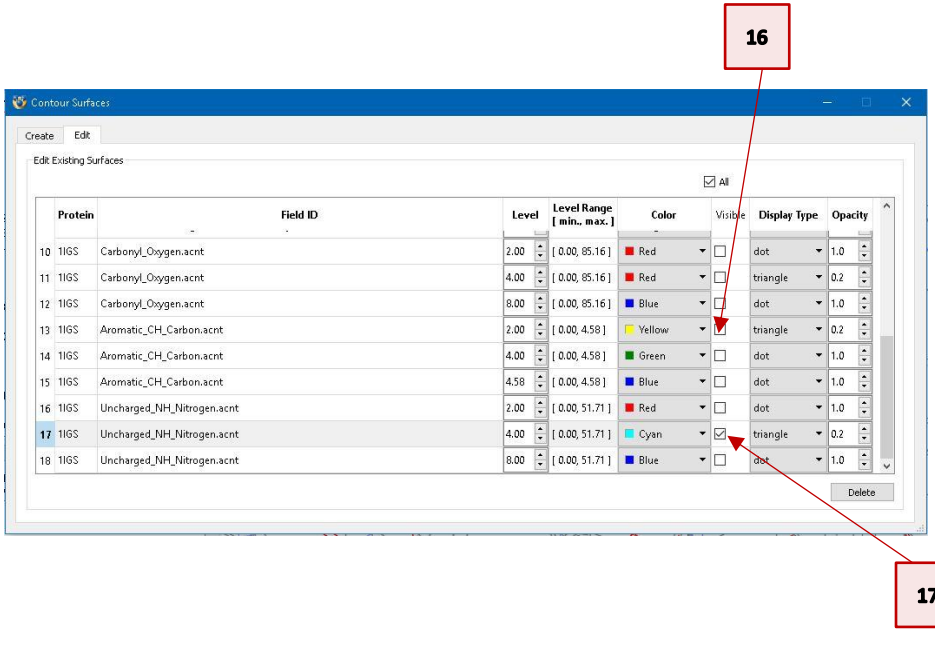
Analysing the full interaction map derived from an uncharged NH nitrogen probe

16. Hide the grid map of the aromatic CH carbon probe by disabling the **Aromatic_CH_Carbon.acnt** tick-boxes in the **Visible** column of the *Contour Surfaces* window (row 13 in the figure on the right).

17. Now show the grip map of the uncharged NH nitrogen probe by enabling the tick-box corresponding to **Uncharged_NH_Nitrogen.acnt** with **Propensity** of *4.0* (row 17 in the figure on the right)

You will now have only the full interaction map derived by using the uncharged NH nitrogen probe displayed in the Hermes 3D view.





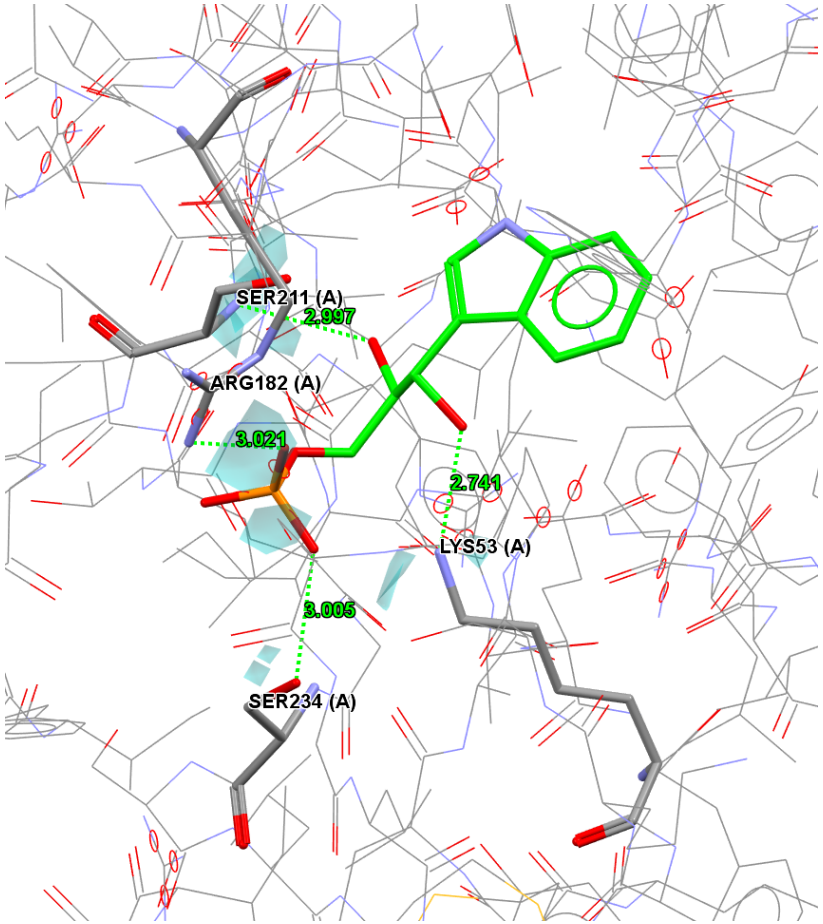
18. The interaction map derived by using the uncharged NH nitrogen as probe reveals that the predicted binding site of IGPS protein has several residues that match in the H-bond donor areas predicted by Full Interaction Maps calculation.

19. From the *Molecule Explorer* window, select LYS53, ARG182, SER211 and SER234. Multi-selection is possible using **CTRL + left-click**.

20. Right-click in the *Molecule Explorer* area and select **Styles** > **Capped Sticks** from the pull-down menu. This will show the selected residues as capped sticks in the Hermes 3D view.

Most of the predicted H-bond donor interaction sites of the best scored docking pose of compound IGP are satisfied by the conformation of the ligand in the protein binding site. Several key hydrogen bond interactions can be formed:

- The triosephosphate moiety of IGP can form H-bonds with the sidechain of Arg182 and with the backbone NH of Ser34.
- The hydroxyl groups of the IGP ligand can act as H-bond acceptors and form H-bonds with the backbone NH of Ser211 and the sidechain of Lys53 (see figure on the right, H atoms not displayed for clarity).



Conclusions

- We used GOLD to dock the IGP molecule in the sIGPS protein. The binding site of the sIGPS protein was identified using the LIGSITE algorithm.
- SuperStar and Full Interaction Maps were used to validate the docking pose of the best scored solution.
    - SuperStar mapped the regions where the IGPS binding site preferred to find H-bond acceptors/donors and hydrophobic groups in the ligand.
    - FIMs mapped the regions where the IGP ligand preferred to find H-bond acceptors/donors and hydrophobic groups in the protein binding site.
- The analyses showed that the binding mode of the best ranked docking pose of the IGP ligand in the sIGPS receptor satisfies both the protein's and ligand's preferred interaction sites predicted with SuperStar and FIMs, respectively. Several H-bonds can be formed, optimising the conformation of the ligand.
- The conformation of the best docking pose of the IGP ligand in the sIGPS protein (with green carbon atoms) is comparable to the native binding mode of the IGP ligand in the 1A53 co-crystal structure of sIGPS-IGP with a 2 Å resolution (with magenta carbon atoms in Figure 2).

Note that the applicability of this approach to other systems depends on the size of the binding site and on the nature of the ligand. Flexible ligands and larger binding sites increase the diversity of docking poses such that the correct binding mode may not be easily identified. In such cases, any previous knowledge for your system could be useful to guide the docking calculation, for example by setting constraints.

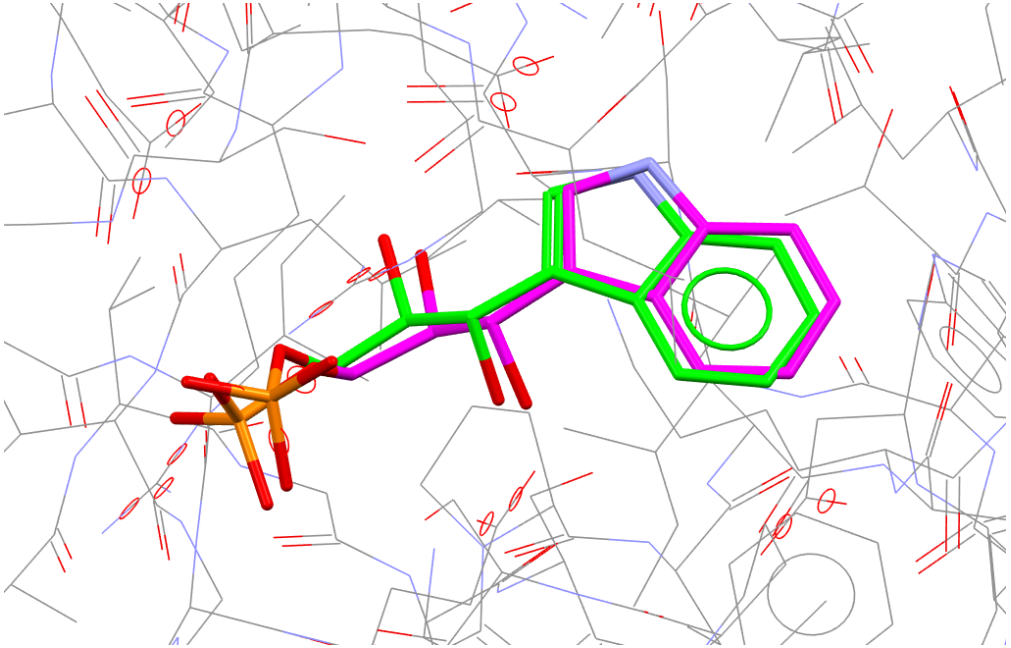


Figure 2. *Comparison of the best scored docking pose of IGP ligand (green carbon atoms) and the crystallographically observed conformation of the ligand in the co-crystal structure of sIGPS-IGP (magenta carbon atoms).*