

# Pose Prediction with GOLD (GLD-003)

Developed using  
2020.0 CSD Release



## Table of Contents

The Purpose of Docking.....	2
GOLD's Evolutionary Algorithm.....	3
A Checklist for Docking.....	3
GOLD and Hermes .....	3
Redocking <i>SAH</i> into <i>MLL1</i> .....	4
Introduction .....	4
Prepare the Coordinates of Your Protein Binding Site and Your Ligand .....	5
Inspect Sidechain Conformations and Protonation States .....	5
Change Selected Protein Sidechains.....	7
Prepare Your Ligand .....	9
Introduction to <i>GOLD</i> Configuration Files.....	10
Set Up a <i>GOLD</i> Run .....	11
Define Your Protein Binding Site .....	11
Enter Your Ligand for Docking .....	12
Set Docking Parameters for Your Ligand .....	12
Output Parameters .....	13
Save Your Setup and Invoke <i>GOLD</i> .....	13
Monitor the Progress of Your Docking.....	14
Visualise Your Results .....	15
High Accuracy Sampling and Scoring .....	16
Specify High Accuracy Settings for Your Docking Run .....	16
Visualise Your Results .....	17
Docking with Protein HBond Constraints.....	18
Introduction.....	18
Set Up Protein HBond Constraints.....	19
Visualise Your Results .....	21

## The Purpose of Docking

When you investigate potential drugs for a new protein target or try to improve existing medicines, it is important to know:

- if and how strongly a proposed molecule binds to the protein (binding affinity prediction)
- **how it binds to the protein (pose prediction)**

Docking 3D coordinates of a ligand into low energy states of a protein binding site (or active site) is commonly used for pose prediction. It approximates the process of a flexible ligand binding to an active site. Docking can be set up to make the most of a scientist's knowledge about:

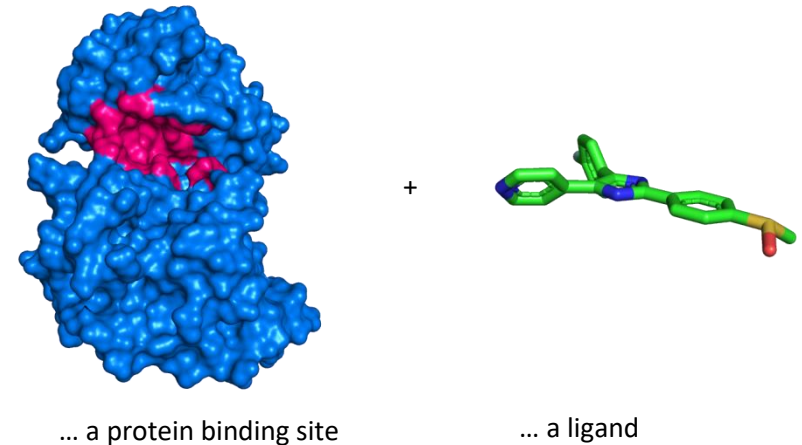
- the flexibility of the target
- structural waters
- metal binding sites
- preferred protein-ligand interactions

Proposed docking poses give chemists insights and testable hypotheses about which parts of the ligand:

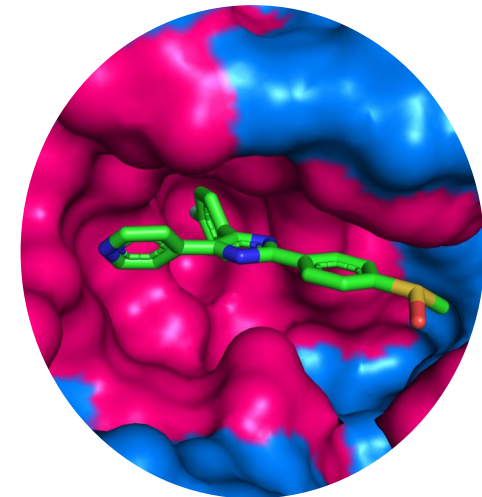
- are responsible for the binding to the protein
- should be optimised to improve the potency of the drug
- would cause clashes with the protein if changed
- can be replaced to improve pharmacokinetic properties and alleviate side effects

The quality of an individual docking result is measured by a relative score. This is used to judge which poses generated for an individual ligand are the most likely ones. Docking scores cannot accurately predict whether a compound binds and what its affinity is. This needs to be confirmed by *experimental* assays.

Given ...



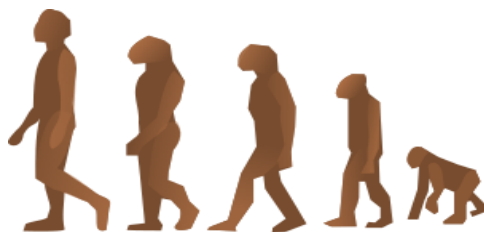
Docking calculates ...



... a pose for the ligand  
... a score for the pose

## GOLD's Evolutionary Algorithm

GOLD stands for “**Genetic Optimisation for Ligand Docking**”. It is one of the most widely used docking programs and available as part of the *CSD-Discovery* and *CSD-Enterprise Suites*.



GOLD's evolutionary algorithm modifies the position, orientation and conformation of a ligand to fit into one or more low energy states of the protein active site. It maps ligand geometry parameters onto populations of chromosomes and then runs evolutionary rounds of **mutation**, **crossover**, **scoring** and **selection** to optimise protein-ligand interactions.

## GOLD and Hermes

GOLD itself is accessible from Windows, Linux and macOS command lines.

In this guide, you will:

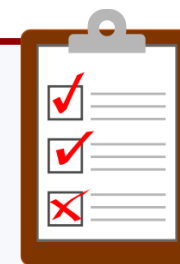
- prepare input for GOLD
- perform dockings
- analyse docking results

To get started, you will use the Hermes visualiser to:

- correctly setup your coordinate data
- choose suitable parameters for GOLD's genetic algorithm
- call GOLD and visualise the result of its dockings

Please note that GOLD can be used without Hermes. For more information, please check the [GOLD Configuration File User Guide](#).

## A Checklist for Docking



Have I got:

- ☐ 3D coordinates for my protein?
- ☐ access to electron density data for my target?
- ☐ knowledge about where the protein active site is?
- ☐ knowledge on whether metal ions are involved in protein-ligand interactions?
  - 3D coordinates for such metal ions?
- ☐ 3D coordinates for my ligand, including the positions of hydrogens?
  - are the bond lengths and angles in my ligand correct?
- ☐ an enumeration of all protonation/tautomeric states for my ligand?
- ☐ any expectations of which interactions are likely to take place within my active site?

Am I aware that:

- ☐ His sidechains can adopt one of three protonation/tautomeric states?
- ☐ Glu and Asp sidechains can interact with the ligand as carboxylates or carboxylic acids?
- ☐ crystallography at medium resolution is not able to provide certainty about the orientation of the amide groups in Asn and Gln sidechains
- ☐ protein sidechains can move?

Have I checked that:

- ☐ there is crystallographic evidence for the position of long and surface exposed sidechains, such as Arg and Lys?
- ☐ there are no unwanted water or buffer molecules in my active site?
- ☐ my assumptions on the positions of protein sidechains are based on electron density data?

## Redocking *SAH* into *MLL1*

### Introduction

#### Your protein target:

*MLL1* is a pathogenic protein which is involved in blood cancer. It alters gene transcription by transferring methyl groups from its cofactor *S*-adenosyl methionine (*SAM*) to histones, the proteins DNA is wrapped around in the cell nucleus. The cofactor is converted to *SAH* *S*-adenosyl-L-homocysteine (*SAH*). Cofactor binding sites can be challenging in drug discovery projects due to the potential for off-target binding to related proteins. Interestingly, *SAM* is very flexible and adopts different conformations in different enzymes. Thus, *SAM* binding sites are considered druggable and currently investigated for the treatments of cancer and neuropsychiatric disorders.<sup>1</sup>

#### Your task:

In this tutorial, you test GOLD's ability to fit a compound, in your case the cofactor product *SAH*, back into its binding site. This is a vital first step to check whether a docking setup has any chance of success in a drug discovery project.

Please download the [example folder of input files](#).

*Is this an easy docking? No! You will meet the following challenges:*

- 11 non-trivial rotatable bonds in the ligand/cofactor
- a flexible sugar ring in the ligand/cofactor
- a partially open shape of the protein binding site

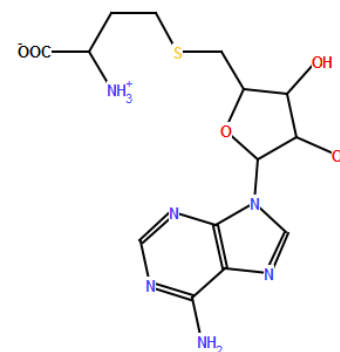
#### You will focus on:

- preparing a correct representation of the protein binding site
- setting an appropriate protonation state for your ligand/cofactor
- considering ring flexibility in the settings for the genetic algorithm

At the end of this tutorial, you will optimise your results by:

- docking with high accuracy settings and
- using Protein HBond constraints.

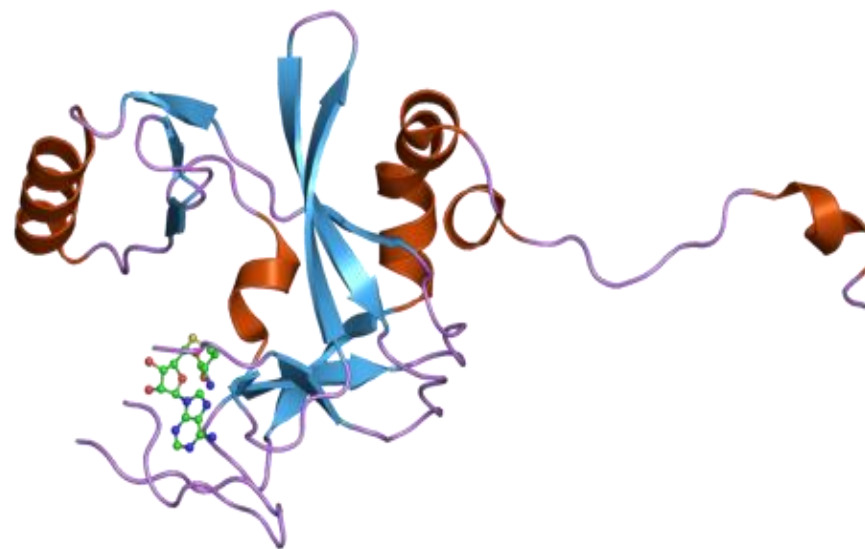
<sup>1</sup> Arrowsmith, Cheryl H., *et al.* "Epigenetic protein families: a new frontier for drug discovery." *Nature Reviews Drug Discovery* (2012). **11**, 384-400.



[Download Input Files](#)



*SAH cofactor, depicted by the Diagram module of the CSD Python API*



*MLL1 complexed to SAH cofactor, PDB-code: 2w5y*

## Prepare the Coordinates of Your Protein Binding Site and Your Ligand

### Inspect Sidechain Conformations and Protonation States

1. Start *Hermes* by double clicking its icon on your desktop. If you are on macOS double click on the *Hermes.app* in the Discovery installation folder.
2. Use **File > Open...** to load *2w5y\_protein.mol2*. This file includes the coordinates deposited as *2w5y* for the heavy atoms. Hydrogens have been added by *Hermes* as demonstrated in the [Introduction to Hermes](#) guide. Hydrogen atoms can also be added by the command line program *gold\_utils -convert* or using the *CSD Python API*.
3. In the same way as 2., load the experimentally solved conformation of the cofactor with added position of hydrogens, i.e. *SAH\_native.mol2*.
4. Find *SAH* in *Molecule Explorer*, right-click on it and select **Center & Zoom 3D view** to focus on the active site.
5. You will now have the cofactor and active site on display. Focus on the selection of sidechains shown in the picture on the right-hand side. You will notice atom pairs at hydrogen bonding distances of less than 3.0Å between the donor and acceptor atoms:

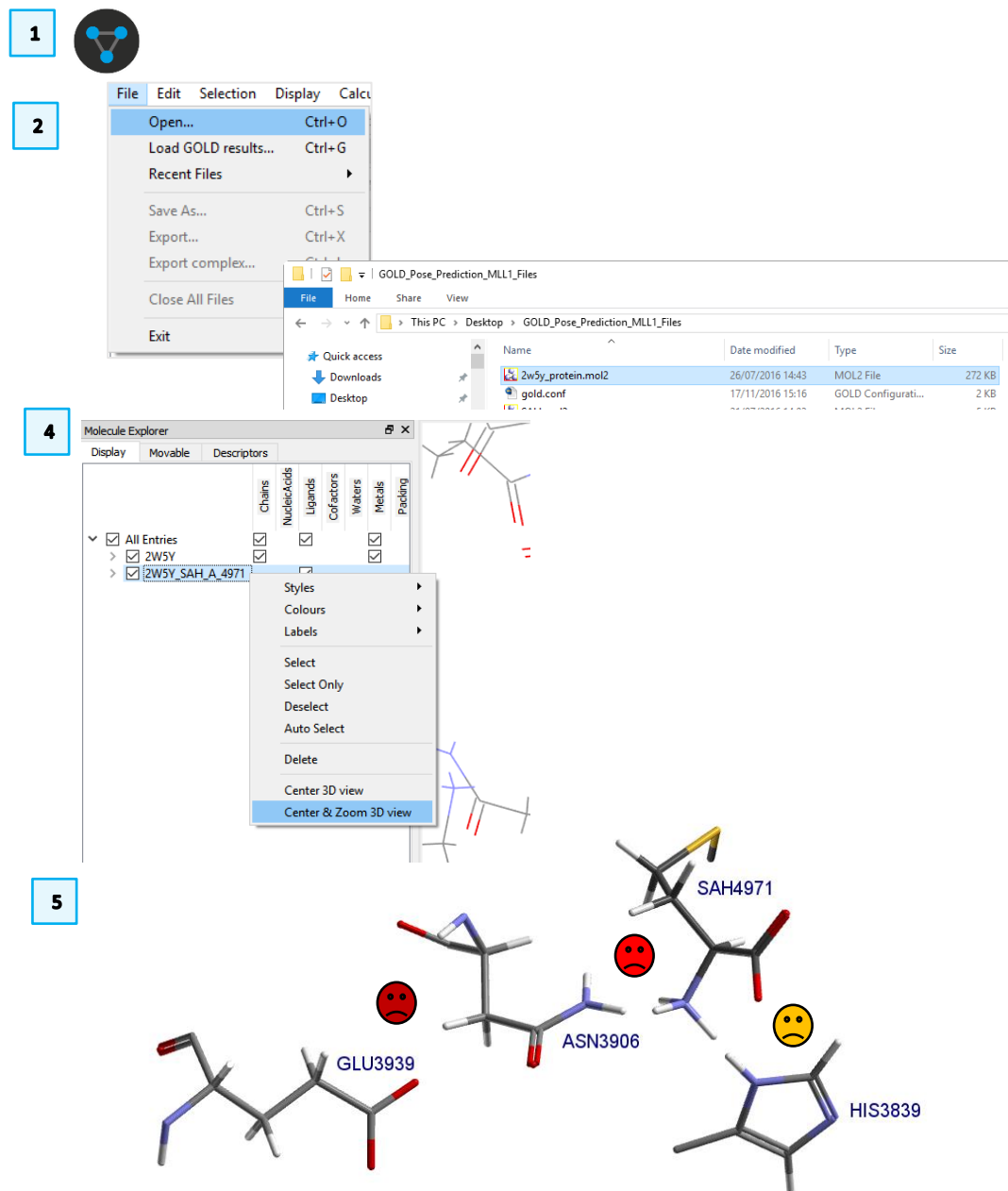
• OD1, Asn3906 and OE1, Glu3939 ☹️

• ND2, Asn3906 and amino N, SAH ☹️

• ND1 of His3839 and amino N, SAH 😞

However, donor atoms point to donors and acceptor atoms to acceptors.<sup>2</sup>

<sup>2</sup> Errors like this one are often curated by PDB REDO. You can download the “Re-refined and rebuilt structure”, *2w5y\_final.pdb*, from [PDB-REDO](#) for comparison.



We can conclude that:

- The orientation of the *Asn3906* amide group in the crystal structure is incorrect.
- The protonation state of *His3839* is unsuitable for the cofactor *SAH*.

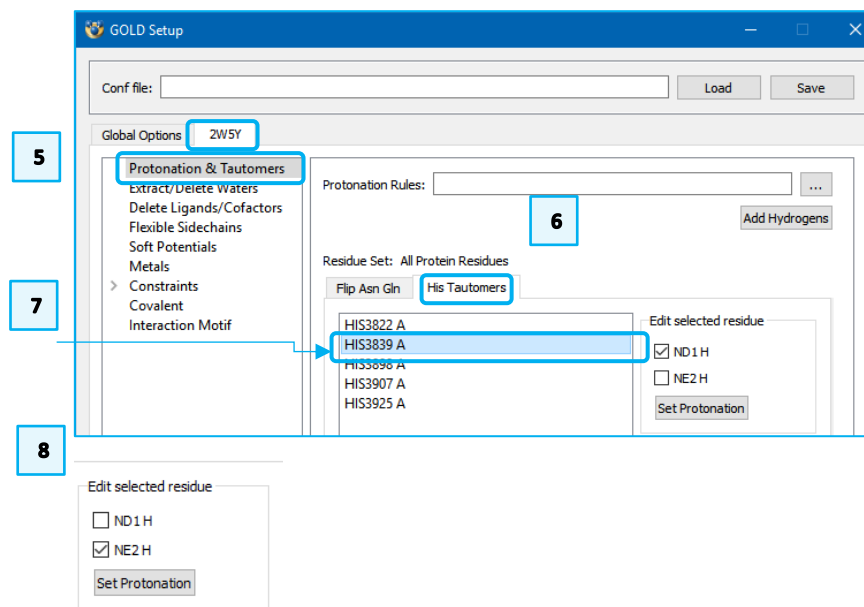
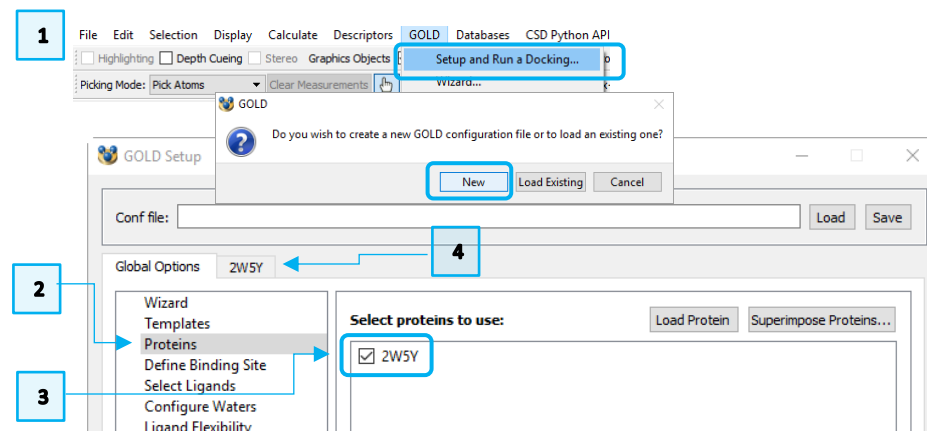
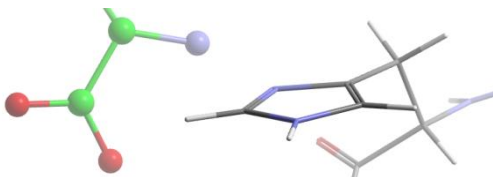
**Note:** Multiple protonation states can be explored using the *ensemble docking* feature of *GOLD*, which can sample several protein protonation states within the same docking run. For more information, check the section on ensemble docking in the [GOLD User Guide](#).

## Change Selected Protein Sidechains

### Modify His3839

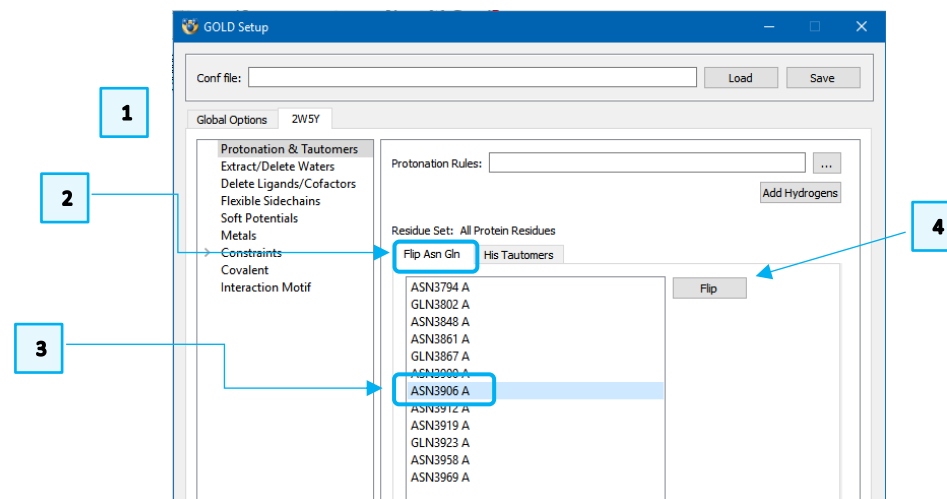
1. Click on **GOLD > Setup and Run a Docking** from Hermes top-level menu and click on **New** button in the *GOLD* pop-up window.
2. In the *Global Options* pane, click *Proteins*.
3. Tick the 2W5Y tick-box as illustrated.
4. A new tab labelling a pane for the protein 2W5Y will appear next to the *Global Options* tab. Click on this tab to enter the protein preparation area for 2W5Y.
5. Within the 2W5Y area, stay on the *Protonation & Tautomers* page.
6. Focus on the *His Tautomers*.
7. Select *HIS3839 A*.
8. In the *Edit selected residue* field
  - uncheck *ND1 H*
  - check *NE2 H*
  - click **Set Protonation**

*His 3839* should now have a deprotonated nitrogen to face the amino group of *SAH*:



### Flip the Amide Group in Asp3906

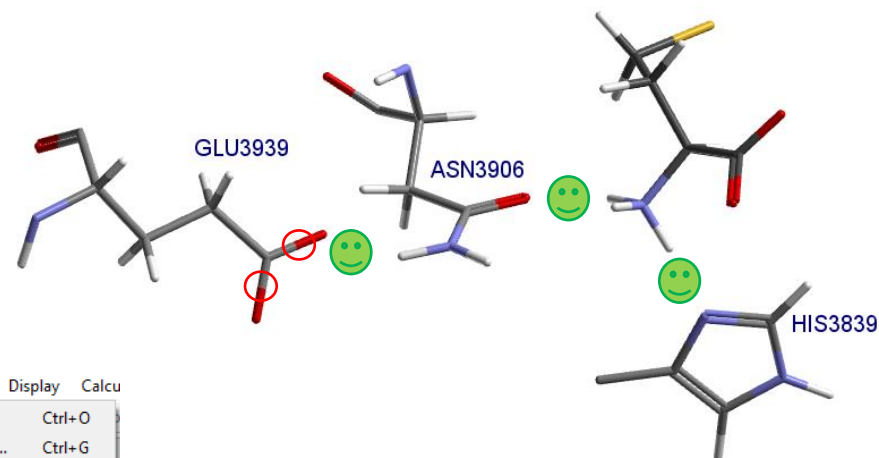
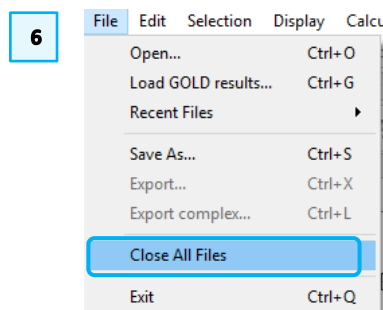
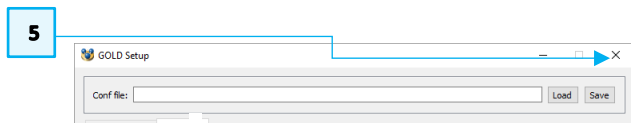
1. Stay in on the 2W5Y page of the *GOLD Setup*.
2. Click on *Flip Asn Gln* tab to enter the pane for flipping *Asp* and *Gln* sidechains.
3. Click *ASN3906 A*.
4. Click **Flip**.
  - *ND2* of the *Asn3906* is now in a position to donate a hydrogen to the carboxylate group of *Glu3939*
  - *OD1* of *Asn3906* and *ND1* of *His3839* can now accept hydrogens from *SAH*.



Use **File > Save As** to preserve the changes made to the protein, as described in the [Introduction to Hermes](#) guide. Use *2w5y\_protein\_prepared.mol2* as the name for your edited protein.

Close:

5. *GOLD Setup* window
6. All files in Hermes, by clicking on **File > Close All Files**





## Prepare Your Ligand

S-adenosyl-L-homocysteine is expected to adopt a zwitterionic state at neutral pH. To test out docking approaches, it is advisable to start with a ligand that does **not** originate from the crystallographic experiment which provides the structure of the protein-ligand binding site. Crystallographic refinement will have aimed to optimise all bond lengths, angles and rings to achieve an optimal fit. To test the quality of a docking setup, you should remove such bias. We will use the idealised version of [SAH\\_ideal.sdf](#) downloaded from the PDB, which has no resemblance to the docked conformation when it comes to e.g. the pucker of the furanose ring.

1. Make sure all previously used files have been closed. Start by loading the structure of SAH in an idealised conformation, *SAH\_ideal.sdf*, using **File > Open** from the top-level menu.
2. Select the ligand/cofactor by right-clicking on SAH in the *Molecule Explorer* window and click on **Select**. The selected atoms are displayed as yellow spheres in the 3D view.
3. Right-click in the 3D view and click on **Label > Label by Atom Label**. This will show the name of each selected atom in the 3D view.
4. Click **Edit > Edit Structure** from the top-level menu to edit the ligand/cofactor.
5. Click on **Hydrogen Atoms** in the *Add* panel of the *Edit Structure* window and subsequently click on the *N1* of the amino acid functional group to add the hydrogen.
6. Click **Atoms & Bonds** in the *Remove* panel of the *Edit Structure* window and select the *H34* atom in the 3D view to remove it and create a carboxylate group.

Now your ligand/cofactor structure is ready for docking. Use the **File > Save As** from the top-level menu to write your changes to a *.mol2* file, named *SAH.mol2*. Now close all files by clicking on **File > Close All Files**.

The figure consists of six numbered screenshots illustrating the steps to prepare the SAH ligand for docking in a molecular modeling software.

- 1**: Screenshot of the top-level menu with **File > Open...** selected.
- 2**: Screenshot of the *Molecule Explorer* window. The **SAH** entry is selected, and the **Select** option is chosen from the context menu.
- 3**: Screenshot of the 3D view with the **Label > Label by Atom Label** option selected from the right-click context menu.
- 4**: Screenshot of the top-level menu with **Edit > Edit Structure** selected.
- 5**: Screenshot of the *Edit Structure* window. In the **Add** panel, **Hydrogen Atoms** is selected. The *N1* atom of the amino acid functional group is highlighted in the 3D view.
- 6**: Screenshot of the *Edit Structure* window. In the **Remove** panel, **Atoms & Bonds** is selected. The *H34* atom is highlighted in the 3D view.

The 3D molecular models show the progression: initial SAH structure, addition of a hydrogen atom to N1, and removal of a hydrogen atom from H34 to form a carboxylate group.

**Note:** GOLD can automatically deduce atom types if bond orders are correctly set up for *GOLD*. Please see Chapter 6 in the [GOLD User Guide](#) for more information.

## Introduction to *GOLD* Configuration Files

*GOLD*'s algorithm is implemented as a command-line tool.

For simple dockings, *GOLD* requires 3 input files:

A *GOLD* configuration file (extension: *.conf*, also see the [GOLD Configuration File User Guide](#) for more information). The *GOLD* configuration file specifies the location of two further files:

- a. the prepared 3D coordinates of your ligand
- b. the prepared 3D coordinates of your protein

An example of a prepared *GOLD* configuration file is included as *MLL1.conf* in your example folder. We are going to prepare another copy of this file in the next sections of this tutorial.

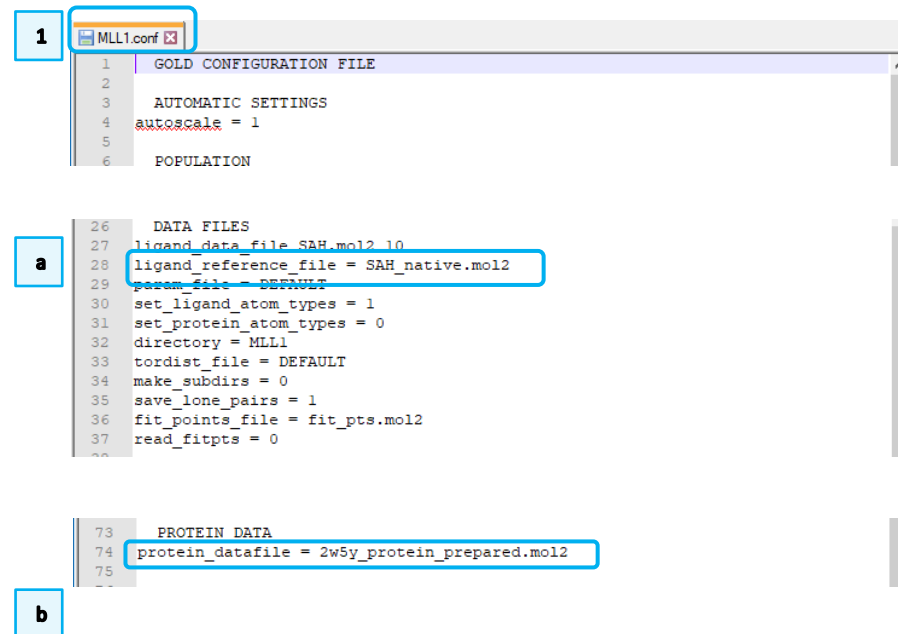
Have a look at the file and find the entries specifying your protein and ligand.

There are several ways to set up and modify *GOLD* configuration files, including:

- with a text editor
- with the *GOLD Setup* menu in *Hermes*
- with the docking module of the *CSD Python API*.

Once set up, *GOLD* can be run from the command line as described by the [GOLD User Guide](#).

In the next sections, you will use the *GOLD Setup* menu in *Hermes* to create a *GOLD* configuration file.



## Set Up a GOLD Run

### Define Your Protein Binding Site

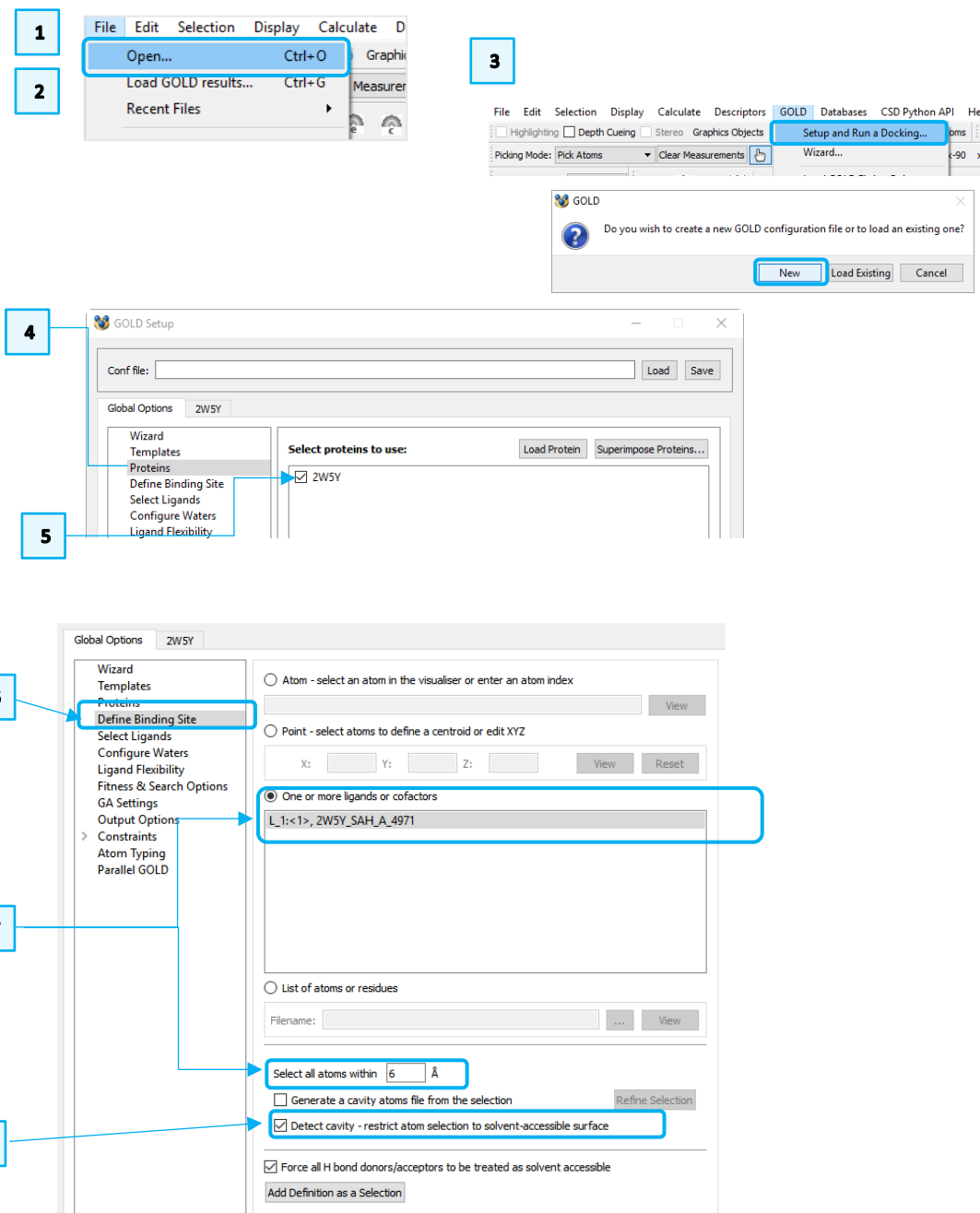
1. In Hermes click **File > Open** from the top-level menu to load your prepared protein, *2w5y\_protein\_prepared.mol2*.
2. Next, click **File > Open** and load *SAH\_native.mol2*. This represents *SAH* in its conformation and orientation in the protein binding site. We will use this file to define the binding site in step 8.
3. Start the *GOLD Setup* by clicking on **GOLD > Setup and Run a Docking** and select **New**.

Step 4-5: Set *2W5Y* as the protein for docking, in the following manner: in the *Global Options* pane, click *Proteins*, and then tick the *2W5Y* tick-box as illustrated on the right.


6. To start the definition of your protein active site, select *Define Binding Site* from the *Global Options* tab.

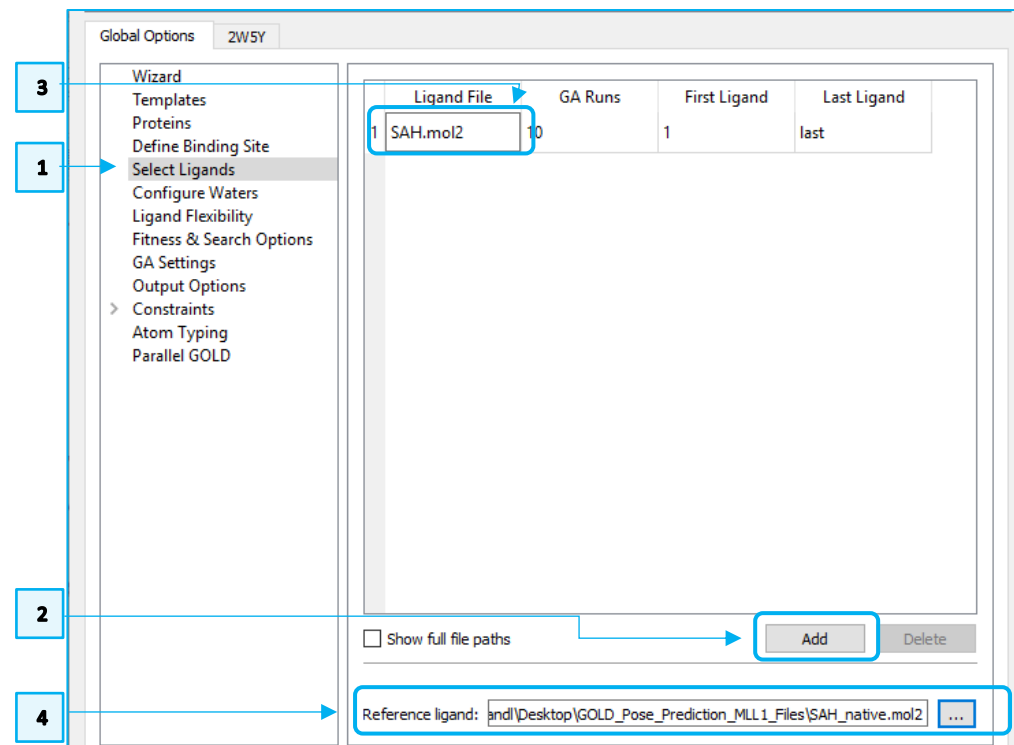
7. In the *Define Binding Site* page, click the **One or more ligands or cofactors** radio button, which will select the already loaded *SAH* ligand/cofactor. Leave the radius for cavity atom selection at the default value of 6Å.

8. Make sure that the **Detect cavity** tick-box is activated.



## Enter Your Ligand for Docking

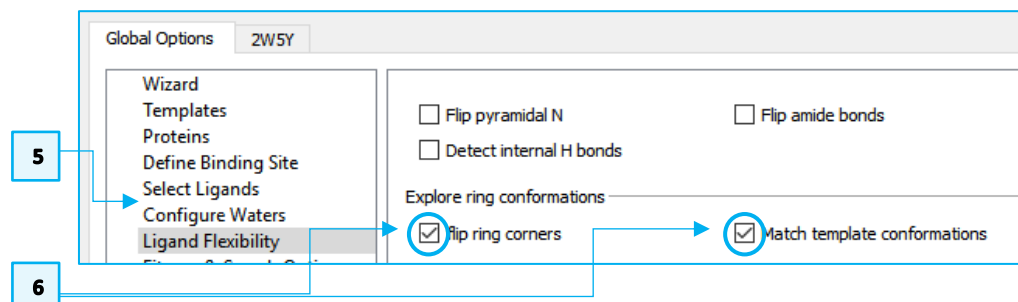
1. To add your ligand/cofactor to the docking setup, click on *Select Ligands* from the *Global Options* tab.
2. Click *Add* to browse for the prepared ligand, *SAH.mol2*.
3. The filename *SAH.mol2* will appear in the *Ligand File* column of the table. Note that, because the position of the undocked ligand can be quite distant from the protein, it is not automatically loaded into Hermes. Optionally, you can load the ligand into Hermes using the **File** menu. You will see that it is quite distant from the active site and has a random position, conformation and orientation.
4. Enable GOLD to judge the success of this protocol by entering the native pose of SAH as a reference ligand. Click on  in the **Reference ligand:** and specify the ligand file, *SAH\_native.mol2*.




## Set Docking Parameters for Your Ligand

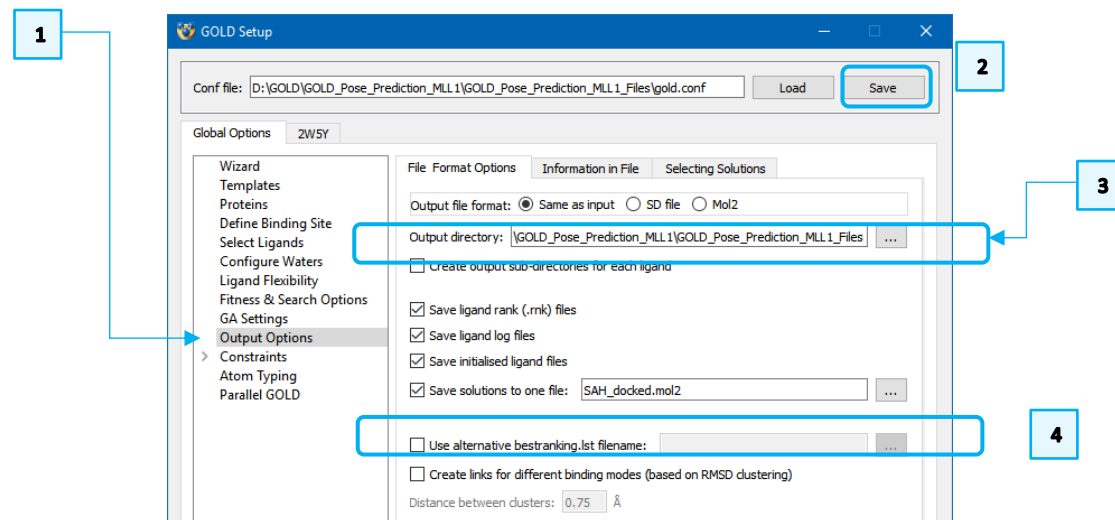
The ribose ring of your input molecule is in one of many possible 3D conformations.

5. To search for alternative puckers, click on *Ligand Flexibility* from the *Global Options* tab.
6. In the *Explore ring conformations* section, activate both **flip ring corners** and **Match template conformations** tick-boxes.



## Output Parameters

1. From the *Global Options* tab click on *Output Options* to explore what is available to you. It is recommended to set an *Output directory* in the *File Format Options* panel with the same base name as your *GOLD* .conf file.
2. Without entering a path in the first step, click **Save** in the **Conf file:** text-box of the interface to trigger a dialogue that allows you to specify a directory and name for your *gold.conf* file. Use *MLL1.conf* as the name.
3. Enter *MLL1* including the full path as the name of your output directory, in the **Output directory**. You can navigate to the desired directory by clicking on .
4. For tidy storage of your solutions, it is advisable to write them to a single SDF or MOL2 file. Tick the **Save solutions to one file** tick-box and enter *SAH\_docked.mol2* as the file name.

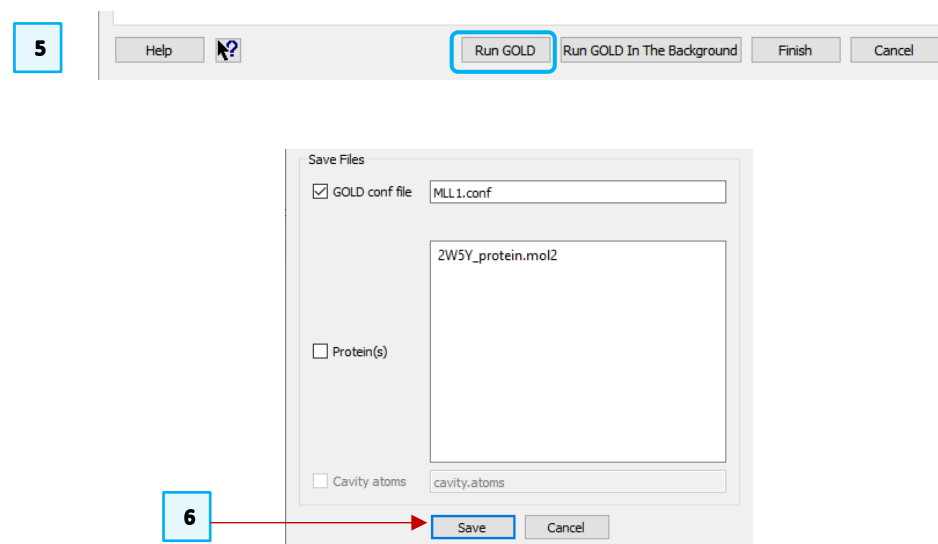


## Save Your Setup and Invoke GOLD

5. Click **Run GOLD** in the *GOLD Setup* window to trigger *GOLD*.
6. When prompted, click **Save**. This will save your .conf file again and use it to run your docking. Open your *GOLD* configuration file in a text editor while *GOLD* runs and look for the settings you have specified.

### Note on the Output Options Section:

By default, *GOLD* specifies the positions of lone electron pairs, e.g. on oxygens and nitrogens, as dummy atoms in the output files of docked ligands. You can exclude dummy atoms by deactivating the **Save lone pairs** tick-box in the *Information in File* panel under *Output Options*. This helps to avoid compatibility issues when post-processing docking poses with third party software. Leave the defaults for this tutorial.



## Monitor the Progress of Your Docking

As the job progresses, output will be displayed in a *Run GOLD* window:

This is a tabbed view that allows inspection of several files:

*list of ligand logs,*

*gold.log,*

*gold\_protein.log,*

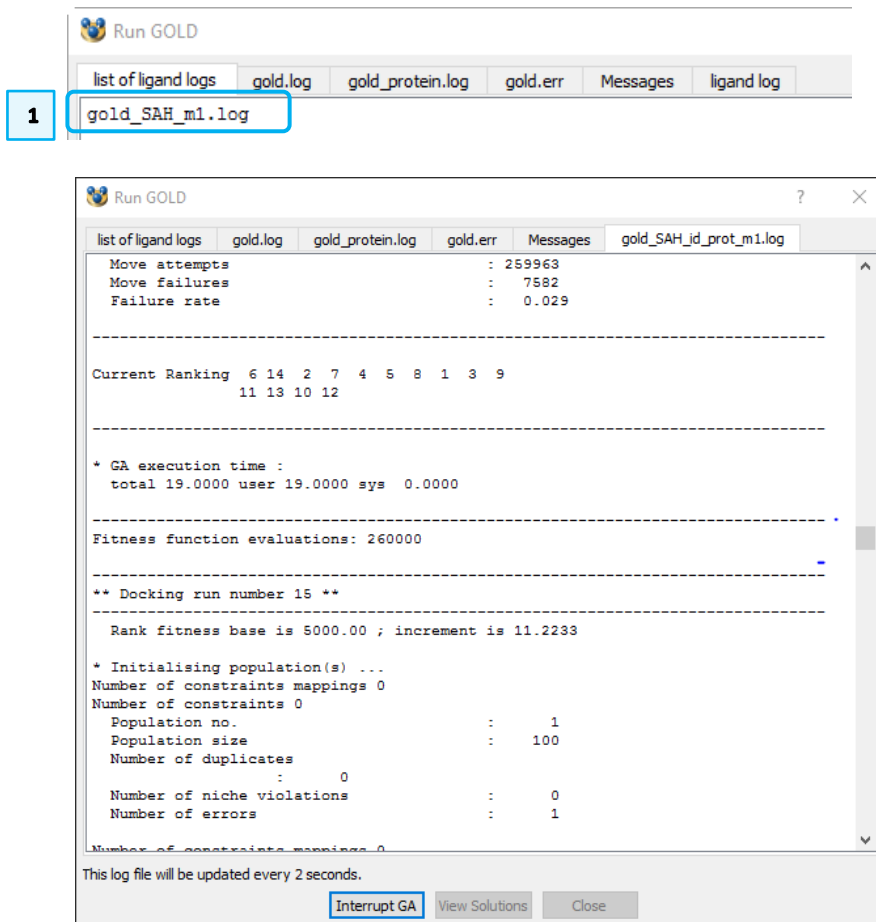
*gold.err,*

*Messages and*

*ligand log.*

During docking, *gold\_protein.log* reflects the setup of *GOLD*. After setup is finished (~10-30s), the ligand is docked and *list of ligand logs* can be used to follow the genetic algorithm sampling through generations of conformations and scores.

1. Click on the ligand log file to view *GOLD*'s progress.



**Note:** Any error or warning messages produced are displayed under the *gold.err* tab. This may contain warning messages relating to the *GOLD* atom type assigner. These messages can be ignored.

## Visualise Your Results

1. Once the job is complete, the docking results can be loaded into Hermes by clicking on the **View Solutions** button in the *Run GOLD* window. Keep the *Run GOLD* window open. Alternatively, use **File > Load GOLD results** and select your *.conf* file.
2. You should see *Molecule Explorer* open on the *Docking Solutions* panel. Use the Up and Down arrows on your keyboard to change between docking solutions. Please use **Customise** to restrict the display of scoring value terms, keeping “*PLP.Fitness*” and “*PLP.Reference.RMSD*” in the display.
3. Click onto the “*PLP.Fitness*” header to sort the scores by their PLP.Fitness value or use the **Sort...** button to that purpose.
4. Load *SAH\_native.mol2* (using **File > Open**) to compare the solutions to the ligand crystal structure, also see below.

Due to the stochastic nature of *GOLD*, the precise values of scoring function terms in your docking will differ from the ones illustrated. But we expect that at least one of the top scoring solutions (green) will be close to the experimental structure (grey).

Congratulations! You have finished your first docking!

*“Hang on, the amino and carboxylate groups of SAH are not forming the hydrogen bonds I expect!”*

True! *SAH* is a long and flexible ligand/cofactor. We will need advanced settings in *GOLD* to consistently achieve better results.

The sections on:

- High Accuracy Sampling and Scoring
- Docking with Protein HBond Constraints

will guide you.

1 This log file will be updated every 2 seconds. Interrupt GA View Solutions Close

2

3

Molecule Explorer

Docking Solutions Display Movable Descriptors

Customise... Sort... Clear

Group by: No grouping

Docking Solutions	PLP.Fitness	PLP.Chemscore.Hbond	PLP.PLIP	PLP.Reference.RMSD
SAH soln:3	64.3772	7.7840	-61.0530	1.8280
SAH soln:4	76.7234	5.8420	-61.6974	0.9364
SAH soln:10	68.4243	4.6530	-54.6411	4.5438
SAH soln:5	67.8070	5.5600	-52.9002	2.4373
SAH soln:2	64.2804	7.7520	-43.4051	3.4305
SAH soln:1	62.7806	7.7680	-41.4384	4.0932
SAH soln:7	60.0197	7.8240	-40.8810	5.8489
SAH soln:9	56.0593	2.9138	-48.4693	7.7528
SAH soln:8	54.5690	2.8560	-46.4323	8.0289
SAH soln:6	52.7087	5.3840	-38.9206	3.3801

4

## High Accuracy Sampling and Scoring

### Specify High Accuracy Settings for Your Docking Run

Please load *MLL1.conf* (**GOLD > Setup and Run a Docking > Load Existing**) as a starting point for this tutorial. We will repeat our *SAH/MLL1* docking with *GOLD*'s premium settings for sampling and scoring.

We increase the number of independent genetic algorithms runs to 30. This means that up to 30 poses will be created.

1. Click on *Select Ligands* in the *Global Options* tab.
2. Double click on *GA Runs* field and edit it to show a value of 30.

We increase the number of operations performed within each individual *GOLD* run to account for the flexibility of the ligand.

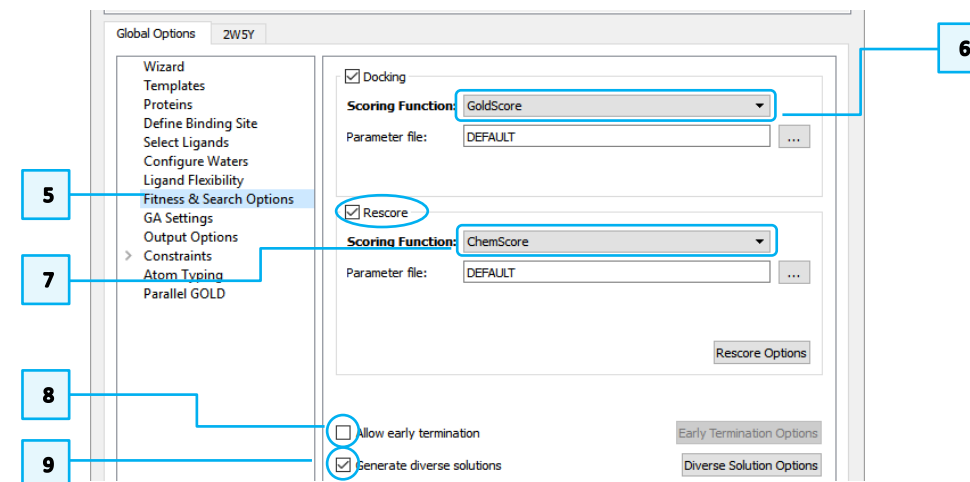
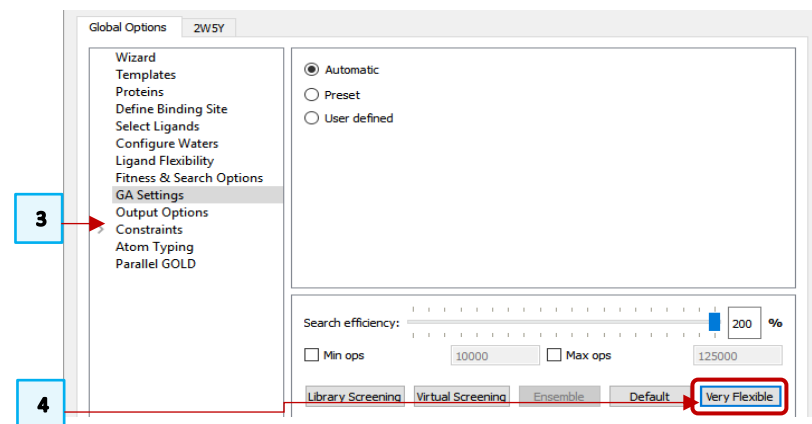
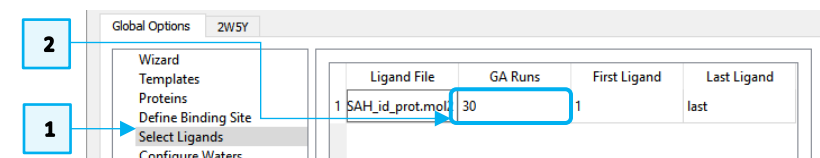
3. Click *GA Settings* in the *Global Options* tab.
4. Click the **Very Flexible** button.

For many targets docking with *GoldScore* and rescoring with *ChemScore* identifies very accurate poses. To give rescoring a set of solutions including poses that do not get the top scores from the scoring function used for docking, we stop *GOLD* from terminating early when it keeps finding the same pose, and instruct it to return ligand poses that are dissimilar from one another.

To achieve these settings:

5. Click on *Fitness & Search Options* of the *Global Options* tab.
6. Set the **Scoring Function:** to *GoldScore*.
7. Click the **Rescore** tick-box and set the **Scoring Function:** to *ChemScore*.
8. Disable the **Allow early termination** option.
9. Enable the **Generate diverse solutions** tick-box.

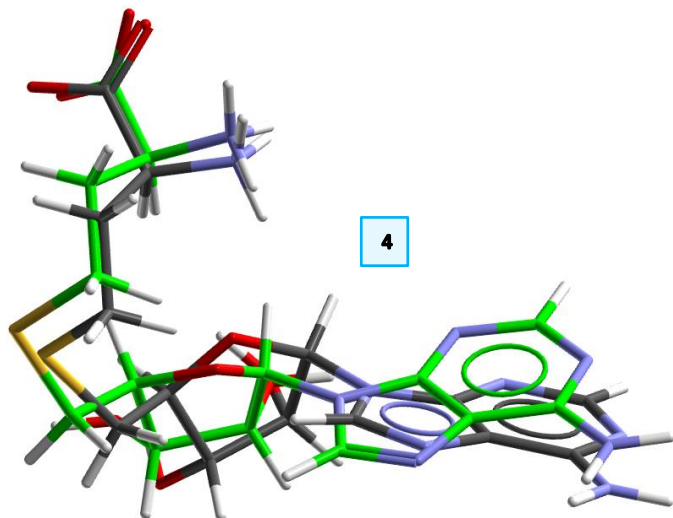
In *Output Options* chose *MLL1\_HA* as the name of your output directory, save your *GOLD* configuration as *MLL1\_HA.conf*. Then, start *GOLD* selecting the **Run GOLD** button. This *GOLD* run will take about 15 minutes.





## Visualise Your Results

1. Once the job is complete, the docking results can be loaded into Hermes by clicking on the **View Solutions** button in the *Run GOLD* window. Keep the *Run GOLD* window open. Alternatively, use **File > Load GOLD results** and select your *MLL1\_HA.conf* file.
2. You should see *Molecule Explorer* open on the *Docking Solutions* panel. Use the Up and Down arrows on your keyboard to change between docking solutions. Click **Customise** to restrict the display of scores to "Goldscore.Fitness", "Chemscore.Fitness", "Chemscore.DEClash.Weighted" and "Goldscore.Reference.RMSD".
3. Use **Sort** to sort by the "Chemscore.Fitness" rescoring function or simply click on the header "Chemscore.Fitness" in the table.
4. Load *SAH\_native.mol2* (**File > Open**) to compare the solutions to the ligand crystal structure. We expect one of your three best scoring solutions (green) to be within an RMSD of less than 2.0 Å to the ligand in the crystal structure (grey).



1 This log file will be updated every 2 seconds.

Interrupt GA View Solutions Close

Molecule Explorer

Docking Solutions Display Movable Descriptors

2 Customise... Sort... Clear

3 Group by: No grouping

Docking Solutions	Goldscore.Fitness	Chemscore.Fit	Chemscore.DEC	Goldscore.Reference.RMSD
SAH soln:25	65.5820	-0.8892	0.1978	0.8484
SAH soln:18	62.9400	-1.2270	3.0006	6.5654
SAH soln:12	51.0433	-2.0408	1.2402	8.4783
SAH soln:6	52.5155	-2.4891	1.8317	8.0619
SAH soln:7	50.0746	-6.6111	6.7693	8.0359
SAH soln:13	50.7905	-6.7661	3.9856	8.1876
SAH soln:9	51.5166	-7.0286	5.8313	7.8399
SAH soln:1	53.3123	-7.3926	2.1098	7.9714
SAH soln:5	42.5890	-7.7771	2.8604	8.4118
SAH soln:26	52.6563	-7.9819	5.4476	8.0243
SAH soln:19	44.5196	-8.0093	2.5300	8.4354
SAH soln:21	54.7799	-8.4011	1.3431	2.9168
SAH soln:28	44.1050	-9.3084	2.9771	8.4713
SAH soln:15	51.2528	-10.0015	3.5088	7.3715
SAH soln:2	50.1114	-10.2025	3.1954	8.4302
SAH soln:22	50.3009	-10.2656	1.9905	8.4186
SAH soln:24	50.5928	-10.7797	2.8521	8.2196
SAH soln:30	41.3535	-11.2844	3.6595	7.9668
SAH soln:14	53.4306	-11.5621	3.9294	2.1963
SAH soln:16	50.5079	-12.1179	1.6438	7.8727
SAH soln:23	45.4939	-12.1699	5.1307	2.8649
SAH soln:29	48.3583	-12.3739	5.7276	8.5326
SAH soln:20	46.7968	-13.9439	5.2948	8.1353

Find identifier: Find Next

## Docking with Protein HBond Constraints

### Introduction

You may have some expectations of where and how your ligands are interacting with the protein target.

Setting up constraints based on your knowledge:

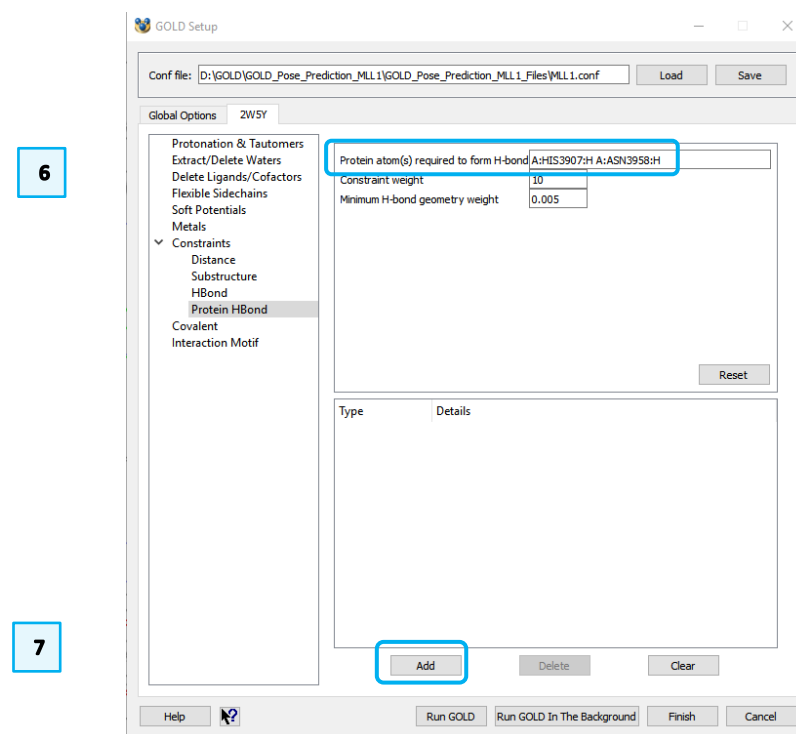
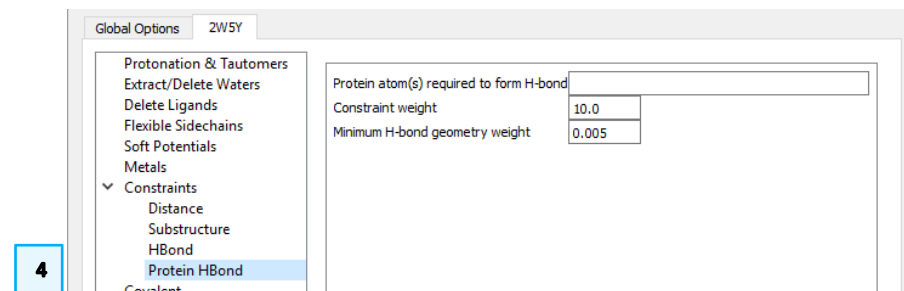
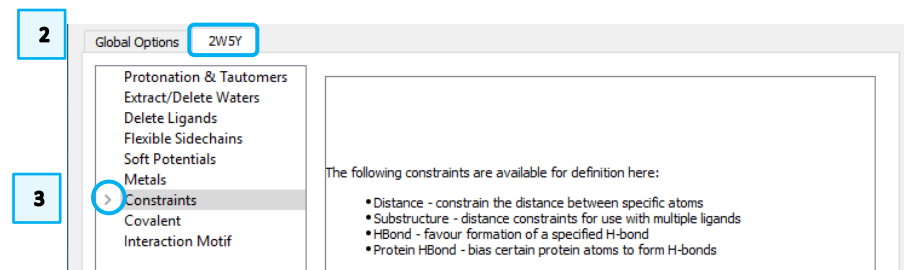
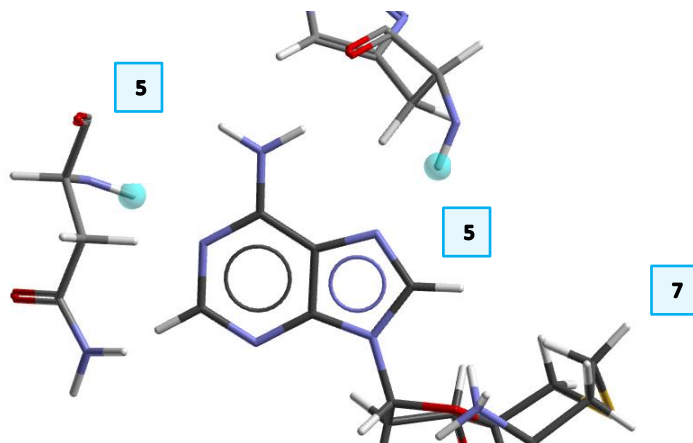
- reduces the number of conformations *GOLD* needs to explore
- speeds up docking
- in the case of Protein HBond Constraints, gently biases *GOLD* towards the interactions you expect, but does not enforce them. It is doing so by introducing penalties for poses that do not show the expected interactions.

## Set Up Protein HBond Constraints

1. Click on **GOLD > Setup and Run a Docking > Load Existing** to load *MLL1.conf*.
2. Click on the *2W5Y* tab.
3. Click on the “>” or “▶” symbol next to “*Constraints*” to open a list of available constraints.
4. Select *Protein HBond* to open the Protein HBond page to set the requirement that at least one of:
  - backbone N-H, *His3907* and
  - backbone N-H, *Asn3958*
 forms an H-bond with *SAH*.

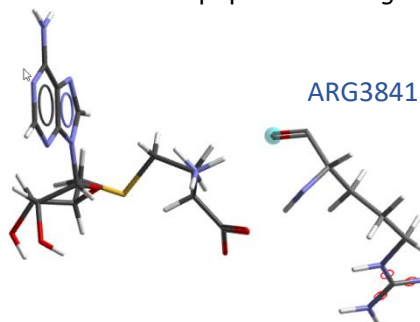
To detect the *His3907* and *Asn3958* residues in the 3D view, use *Molecule Explorer* window. Individual components of the model can be opened by selecting the “>” to the left for Windows. Click on “>” adjacent to *2W5Y* under *All Entries*, then “>” adjacent to *Chains* and then “>” adjacent to *A* to display the list of all the residues. You can then right-click on the desired residue and click on **Center & Zoom 3D view**.

5. Click on the backbone N-H atoms of both *His3907* and *Asn3958* in the 3D view (**H atoms not N**). Blue halos will appear.
6. The selected atoms are listed in the **Protein atom(s) required to form H-bond** field.
7. Click on **Add** to enter the constraint.

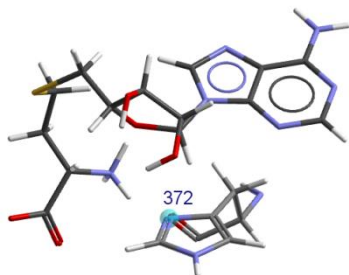


8. Click **Reset** to clear the **Protein atom(s) required to form H-bond** field and set up an individual constraint:

- Click on the backbone peptide *O* of *Arg3841* and then click **Add**.



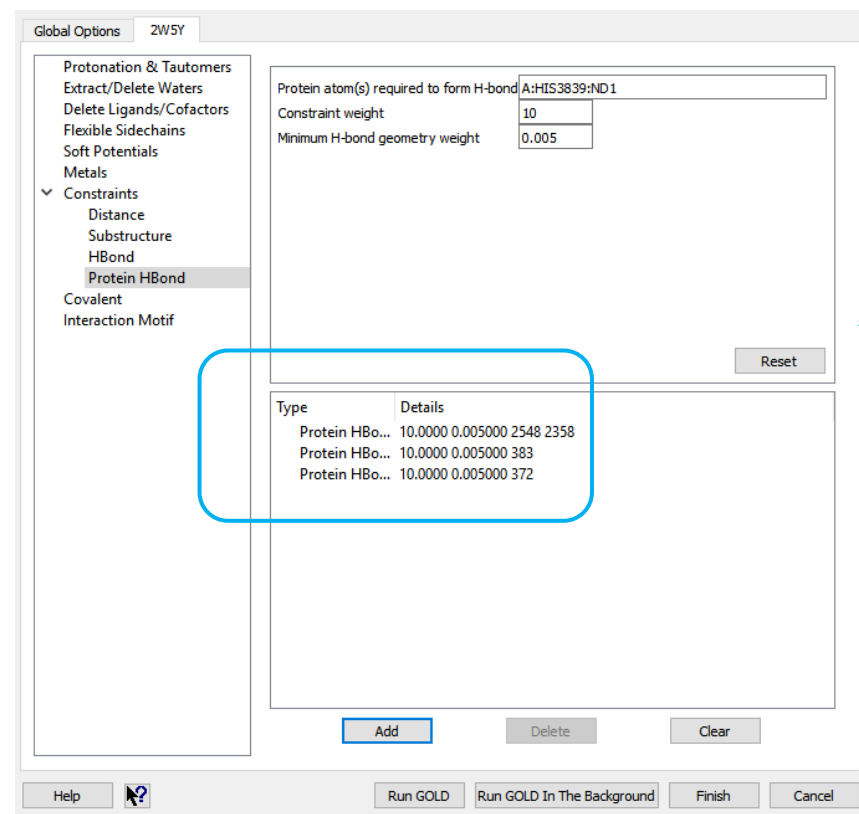
- Repeat the same procedure for *ND1* of *His3839*.



At the end of your setup, you should have:

- one constraint involving two atoms
- two constraints involving one atom each.

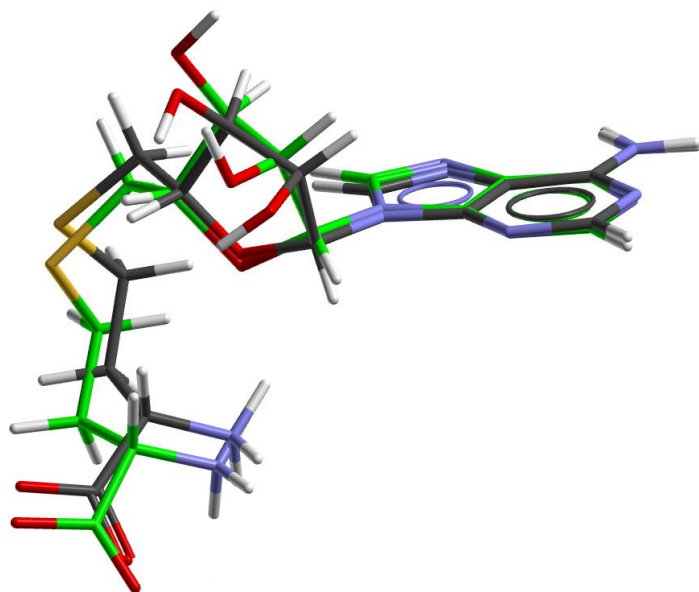
Save your configuration file as *MLL1\_PHB\_Constraints.conf*.



## Visualise Your Results

Enter *MLL1\_PHB\_Constraints* as your output directory (*Global Options > Output Options*) and save your settings as *MLL1\_PHB\_Constraints.conf*.

Please run your docking and view the results. We expect *GOLD* to identify several very good poses with no constraint violations.



Molecule Explorer

Docking Solutions   Display   Movable   Descriptors

Customise...   Sort...   Clear

Group by: No grouping

Docking Solution	PLP.Fitness	PLP.PLP	PLP.Reference.RM	PLP.Constraint
SAH soln:4	85.2265	-63.4952	0.6476	0.0000
SAH soln:3	74.6352	-58.1347	0.9450	0.0000
SAH soln:5	71.7372	-52.7890	1.3902	0.0000
SAH soln:1	62.7020	-57.7021	5.0481	0.0000
SAH soln:2	61.0710	-43.4040	3.3184	0.0000

Find identifier:  Find Next