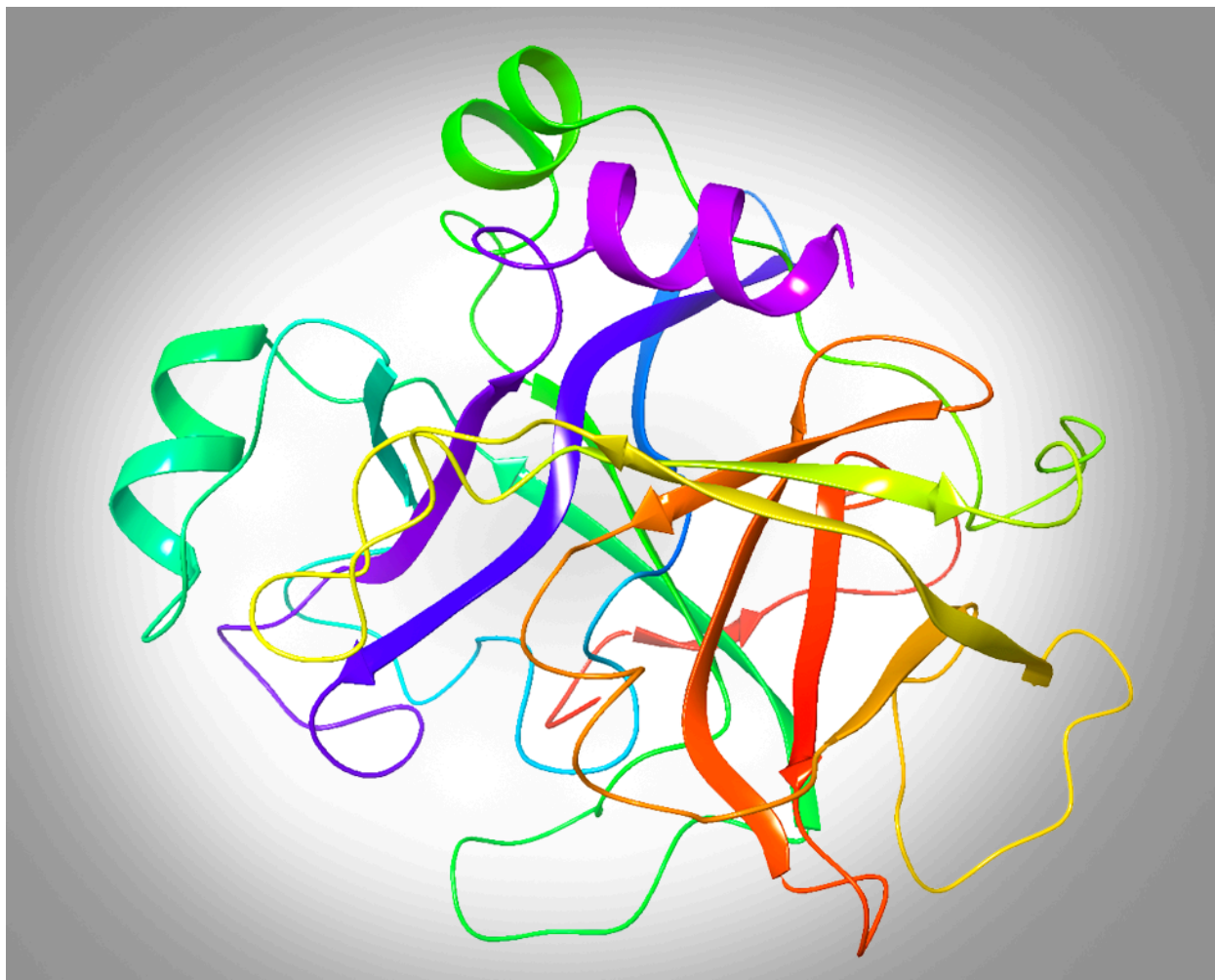


Structure-Based Virtual Screening



Structure-Based Virtual Screening

About this Lesson:

This lesson will focus on an important early stage of drug discovery in which protein structures and molecular modeling are utilized to identify molecules that can be further developed into drugs.

Using Maestro, students will learn how to perform a virtual screen for potential inhibitors of FXa using the ligand docking application Glide. Students will learn how to generate a protein receptor grid, dock a set of ligands into the receptor grid, and analyze the docking results. Students can then sketch their own inhibitor designs and evaluate their value.

Learning Objectives:

- Differentiate between ligand-based and structure-based molecular modeling approaches
- Learn the steps of a molecular docking workflow using Schrödinger's Glide
- Perform a structure-based virtual screen of a small set of ligands
- Design your own inhibitor for FXa and determine its docking score

Lesson Contents:

1. [Setting Up the Maestro Session](#)
2. [Introduction to Structure-Based Virtual Screening](#)
3. [Generating a Receptor Grid](#)
4. [Docking the Cognate Ligand with a Hydrogen-Bond Constraint](#)
5. [Docking the Screening Compounds](#)
6. [Analyzing Results and Binding-Site Characterization](#)
7. [Individual Exercise](#)
8. [Summary, Additional Resources, and References](#)
9. [Glossary of Terms](#)

Standards Alignment:

- *ACS Guidelines*
 - Biological macromolecules ([Section 5.1](#))
- *AAMC MCAT*
 - Structure, function, and reactivity of biologically-relevant molecules ([5D](#))

Assessments for Understanding:

The following types of formative assessments are embedded in this lesson:

- Assessment of student understanding through discussion of warm-up questions and filling in any knowledge gaps about structure-based virtual screen steps
- Visual assessment of student-generated docking scores from their own set of ligands

Associated Documentation Pages: [Glide Documentation](#)

Warm-Up Questions:

Read the article "[Structure-Based Virtual Screening: From Classical to Artificial Intelligence](#)" and answer the following questions.

1. What is a typical drug development timeline from drug target identification to clinical trials?
2. What are some advantages and disadvantages to performing structure-based virtual screens?

Need help? Contact us at teaching@schrodinger.com

1. Setting Up the Maestro Session

At the start of the Maestro session, it is essential to 1) check your mouse actions, 2) change the file path to the Working Directory for this lesson, and 3) save your project file. The working directory indicated in this section contains the files necessary to complete this lesson. If you do not set the appropriate working directory, you will be unable to run any calculations.

1. Launch the Virtual Cluster

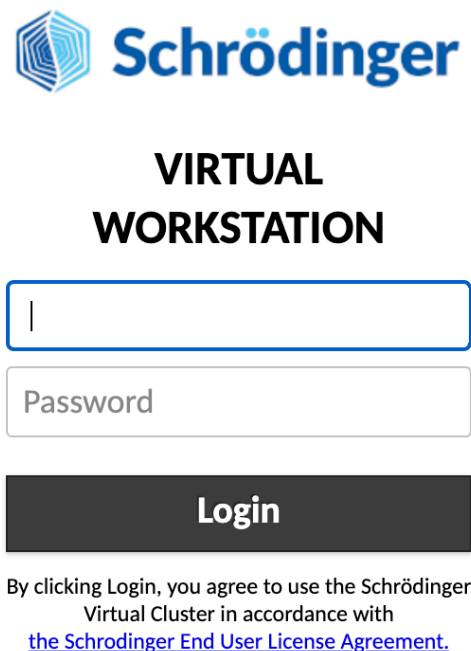
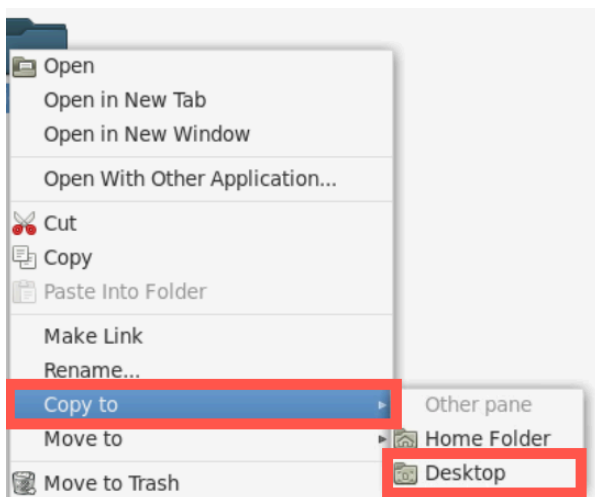


Figure 1-1. Virtual workstation login page.



2. Double-click the **course-data** folder on the desktop

Figure 1-2. Course-data folder on the desktop.



3. Right-click the SBVS folder and select **Copy to > Desktop**

Figure 1-3. Copy the lesson folder to the Desktop.



4. Double-click the Maestro icon on the desktop

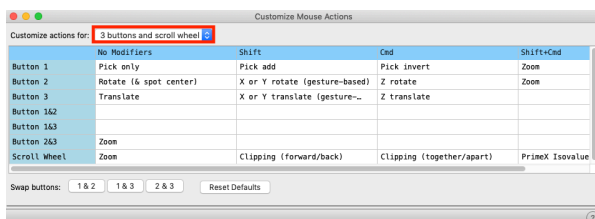
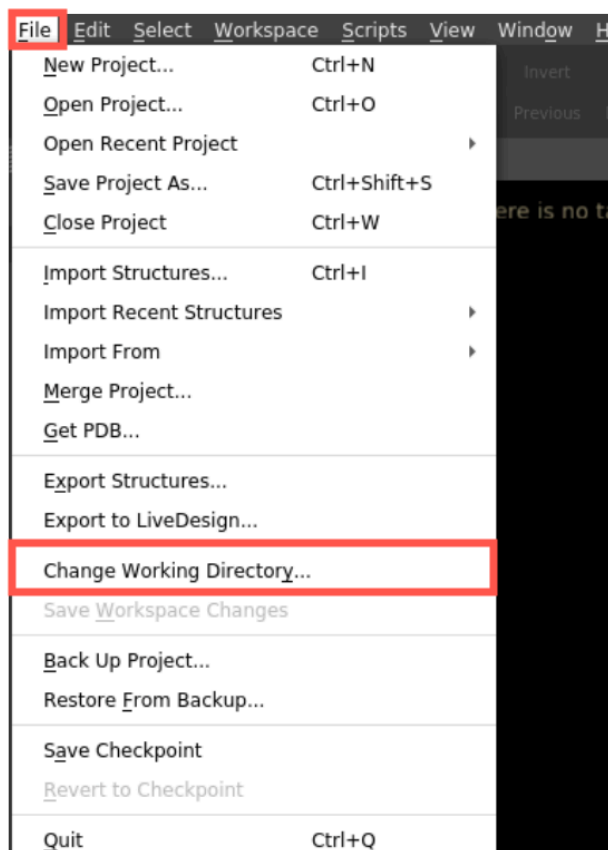


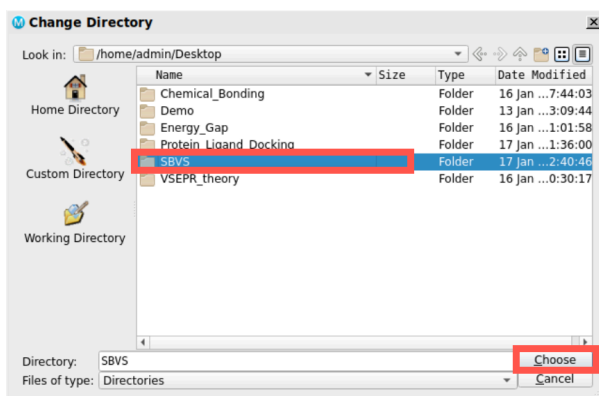
Figure 1-4. Change the mouse actions.

5. Check your mouse actions.
 - o Go to **Workspace > Customize Mouse Actions**
 - o *Note:* This lesson was made with a three-button mouse with a scroll wheel, but a trackpad can still be used
 - o **Trackpad keys:**
 - **Up/Down trackpad** = Zoom In/Out
 - **Option** = Rotate
 - **Control** = Translate



6. Go to File > Change Working Directory

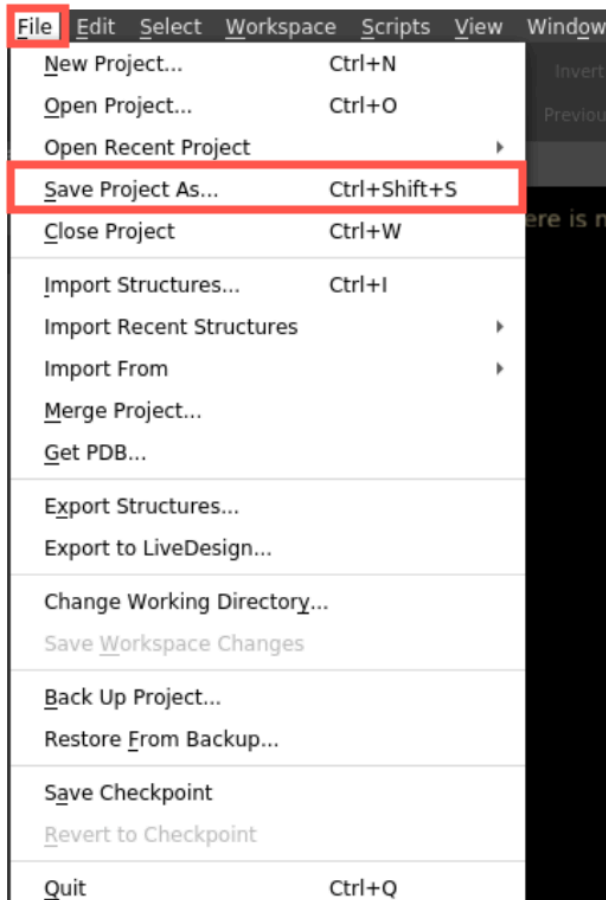
Figure 1-5. Change Working Directory option.



7. Navigate to Desktop > SBVS folder and click **Choose**

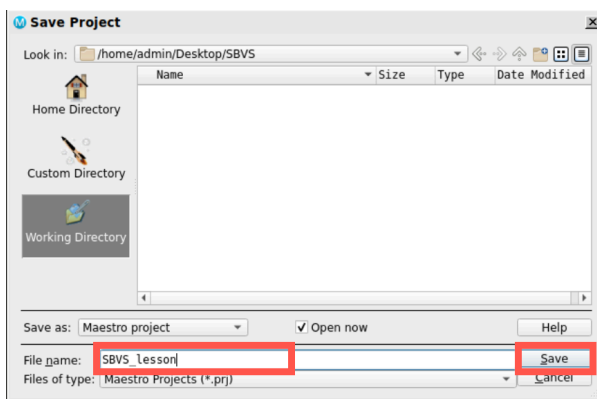
Pre-generated input and results files are included for running jobs or examining output

Figure 1-6. Change Working Directory panel.



8. Go to File > Save Project As

Figure 1-7. Save Project option.



9. Change the *File name* to SBVS_lesson, click Save
- The project is now named SBVS_lesson.prj

Figure 1-8. Save Project panel.

2. Introduction to Structure-Based Virtual Screening

Molecular modeling encompasses a wide array of approaches that impact the early stages of drug discovery: hit ID, hit-to-lead, and lead optimization. Generally, these molecular modeling approaches are broken down into two categories: ligand-based and structure-based approaches. Ligand-based approaches use information from a hit or series of hits to inform on next stages in drug discovery. This information includes molecular fingerprints, shape, charge, and more. Structure-based approaches use information not only from a ligand structure but also from a target structure. A target is another term for a protein or other macromolecule that is being targeted by a drug. Both of these approaches are often used synergistically in all stages of computer-aided drug discovery and in combination with fragment-based approaches as well.

In this lesson, we will be performing a structure-based virtual screen of potential FXa inhibitors. Structure-based approaches allow for the identification of key residues around the ligand and water energetics in the binding site that could modify or enhance the binding of one compound over another. A binding site is a pocket or surface of a target protein where a compound or drug binds to elicit a downstream effect in a disease pathway.

Figure 1 below shows a schematic for the steps involved in a structure-based virtual screen. Structure files obtained from the PDB, vendors, and other sources often lack necessary information for performing modeling-related tasks. Typically, these files are missing hydrogens, partial charges, side chains, and/or whole loop regions. In order to make these structures suitable for modeling tasks, we use the Protein Preparation Workflow to resolve issues. Similarly, ligand files can be sourced from numerous places, such as vendors or databases, often in the form of 1D or 2D structures with

unstandardized chemistry. LigPrep can convert ligand files to 3D structures, with the chemistry properly standardized and extrapolated, ready for use in virtual screening. In this lesson, the protein, cognate ligand, and virtual screening ligands have already been prepared in order to save time. However, these preparation steps are a necessary part of a virtual screen and must be done before docking. Please see the lesson on Protein-Ligand Interactions for guidance on using the Protein Preparation Workflow and LigPrep.

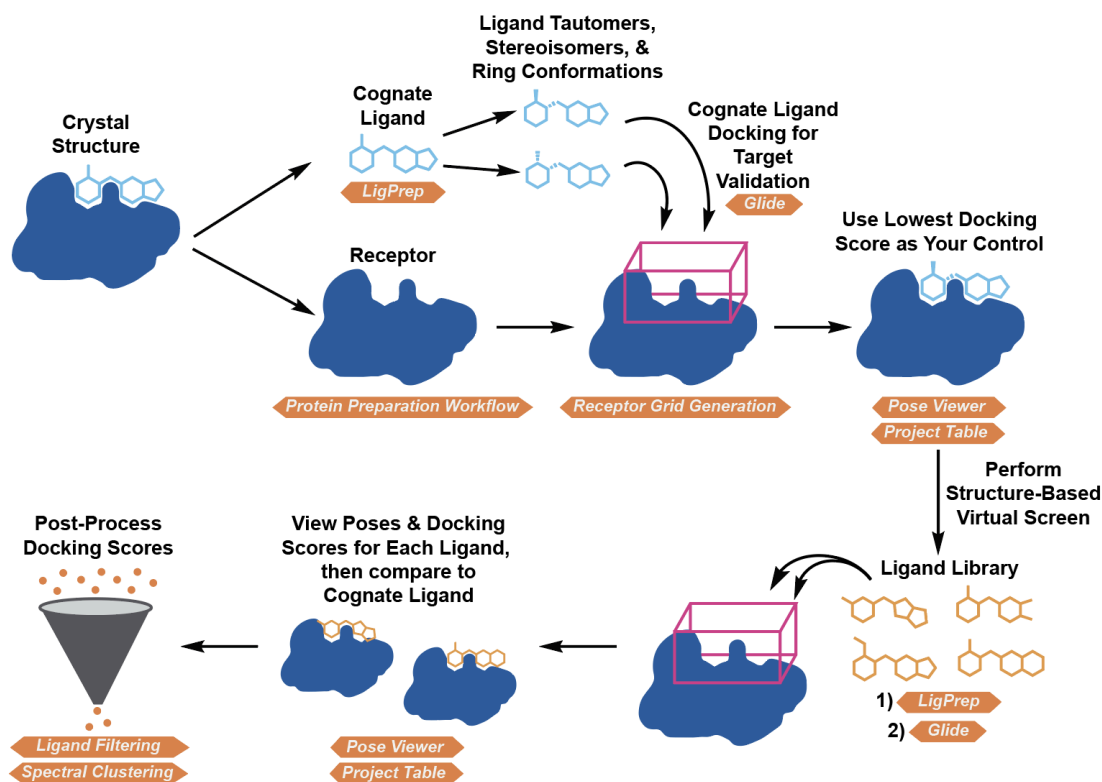


Figure 1. Workflow for a Structure-Based Virtual Screening

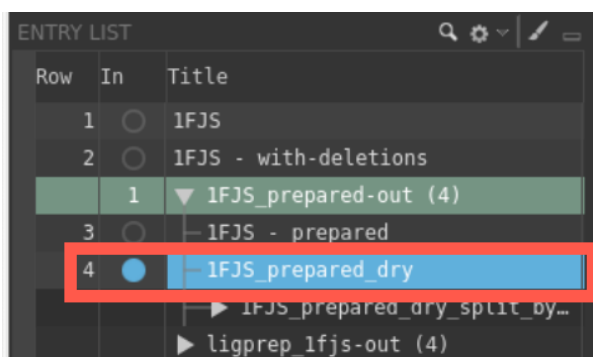
3. Generating a Receptor Grid with Hydrogen-Bond Constraint

Constraint

This section is a continuation of the [Protein Ligand Docking lesson](#), Section 5: Generating a Receptor Grid with Hydrogen-Bond Constraint. If you already completed that lesson, please Skip to Section 5 of the SBVS lesson: Docking the Screening Compounds.

Grid generation must be performed prior to docking with Glide. The shape and properties of the receptor are represented in a grid by fields that become progressively more discriminating during the docking process. To add more information to a receptor grid, different kinds of constraints can be applied during the grid generation stage. For a comprehensive overview of constraint options, see the [Glide User Manual \(Help > Help > User Manuals > Glide User Manual\)](#). In this tutorial, we will set a hydrogen bond constraint in our receptor grid.

3.1. Identify the binding site



1. Click the **In** circle next to **1FJS_prepared_dry** to include it in the Workspace

Figure 3-1. Select and include 1FJS_prepared_dry (Entry #4).

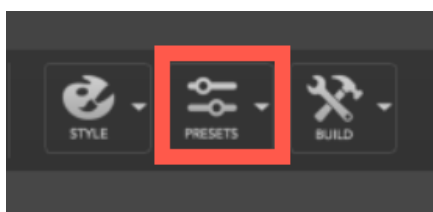


Figure 3-2. Double-clicking Presets to render the protein structure with a Custom Preset.

2. Double-click **Presets**
 - 1FJS_prepared_dry is rendered using the Custom Preset



Figure 3-3. Toggle off the residue labels.

- You may toggle off the residue labels in the bottom right corner if you wish

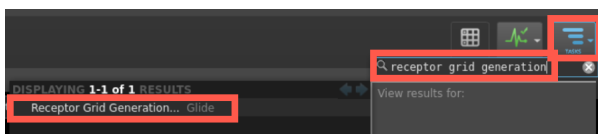


Figure 3-4. Opening up Receptor Grid Generation from the Tasks menu.

- Go to **Tasks** and type **Receptor Grid Generation** in the search bar
- Select **Receptor Grid Generation: Glide**
 - The Receptor Grid Generation panel opens

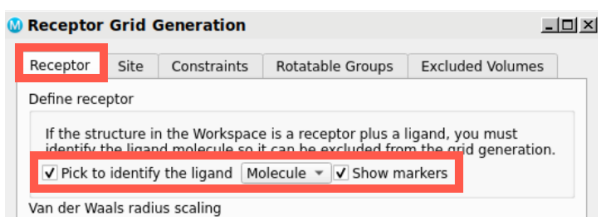


Figure 3-5. The Receptor tab of Receptor Grid Generation.

- Under Define Receptor, check the boxes for **Pick to Identify the ligand (Molecule)** and **Show Markers**
 - A banner in the Workspace will prompt you to click on an atom in the ligand

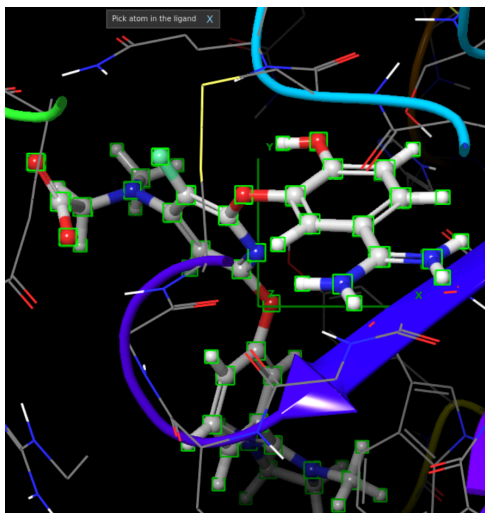


Figure 3-6. Selecting any atom part of the ligand highlights the entire molecule with green cubes around each atom.

- Click any atom on the **ligand**. The ligand is the only molecule in the workspace that is in ball-and-stick representation.

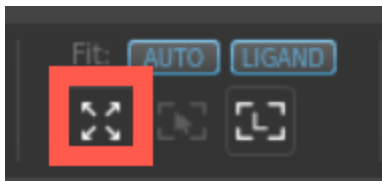


Figure 3-7. Zooming out to visualize the entire structure.

8. You may **zoom out** to visualize the entire protein structure using the **Fit Tools**

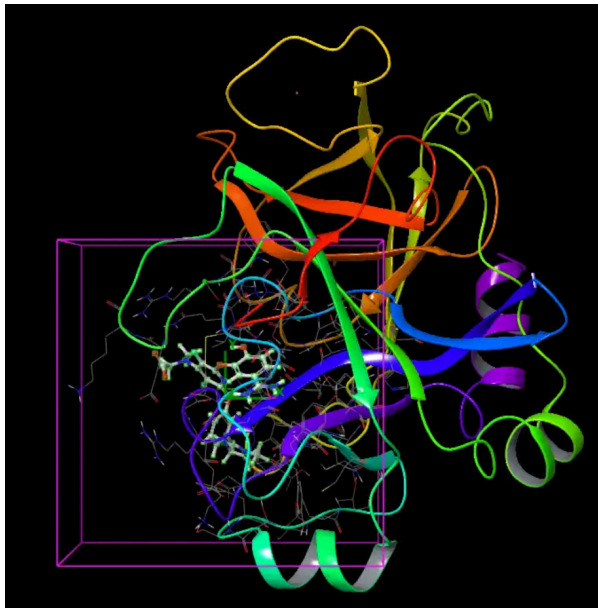


Figure 3-8. The ligand is defined to be excluded from grid generation.

9. The ligand is now highlighted in green with a purple box around it
 - o The purpose of this is to exclude the ligand from the grid generation so it may be reused with a variety of other ligands for docking of a larger ligand library

Note: The purple bounding box defines the region that the docked molecule(s) can occupy to satisfy the initial stages of docking

3.2. Define the bounding box dimensions

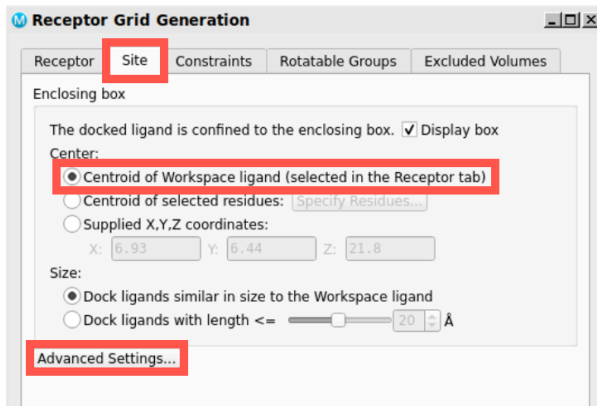


Figure 3-9. The Site tab of Receptor Grid Generation.

1. Click the **Site tab**
2. Select **Centroid of Workspace ligand (selected in the Receptor tab)**
3. Click **Advanced Settings**
 - a. A green inner bounding box appears

Note: The green bounding box defines the region in which the centroid of the docked molecule(s) must occupy to pass the initial stages docking

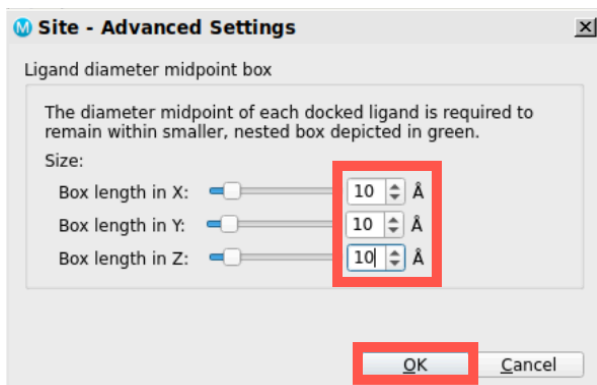


Figure 3-10. Ligand diameter midpoint box panel.

4. Keep the default settings for X, Y, and Z sizes the same at **10, 10, and 10 Å**, respectively.
5. Click **OK**

3.3. Set a hydrogen bonding constraint

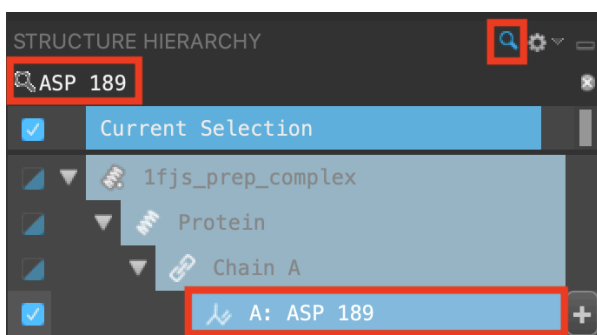


Figure 3-11. Search in the Structure Hierarchy.

1. Minimize the Receptor Grid Generation panel (do not close it)
2. Type **L** to zoom to the ligand
3. In the **Structure Hierarchy** under the Entry List, click the **magnifying glass**
4. In the search field, type **ASP 189**
5. Select **ASP 189**

According to [Adler et al.](#), the salt bridge formed between the inhibitor and Asp189 observed in the crystal structure of the complex contributes to the potency of the ligand. For this tutorial, you will set the constraint for this specific hydrogen bond in the receptor grid. Please see the [Introduction to Structure Preparation and Visualization](#) tutorial for instructions on how to add residue labels and show H-bonds.

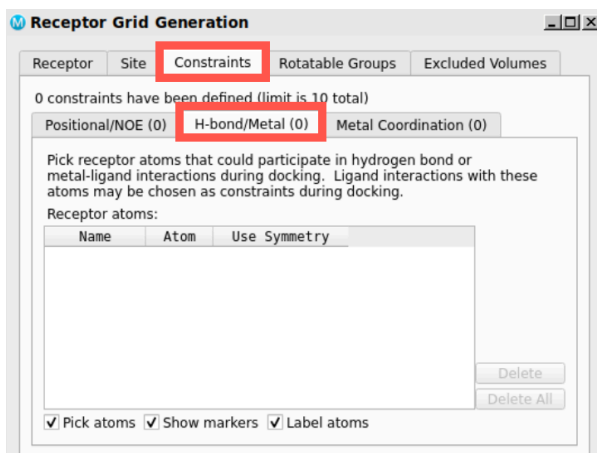


Figure 3-12. The Constraints tab of Receptor Grid Generation.

6. In the Receptor Grid Generation panel, click the **Constraints** tab
7. Click the **H-bond/Metal (0)** tab
 - o A banner appears prompting selection of the receptor atom to be the constraint

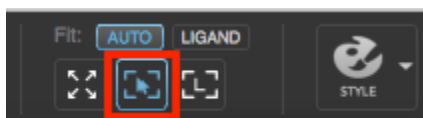


Figure 3-13. Zoom to selected atoms.

8. Under Fit, click **Fit view to selected atoms**

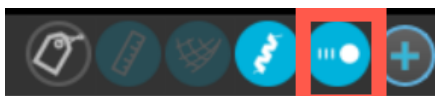


Figure 3-14. Turn on the Interactions toggle in the workspace toggles.

9. Turn the Interactions toggle on to see the non-covalent interactions between the ligand and active site residues

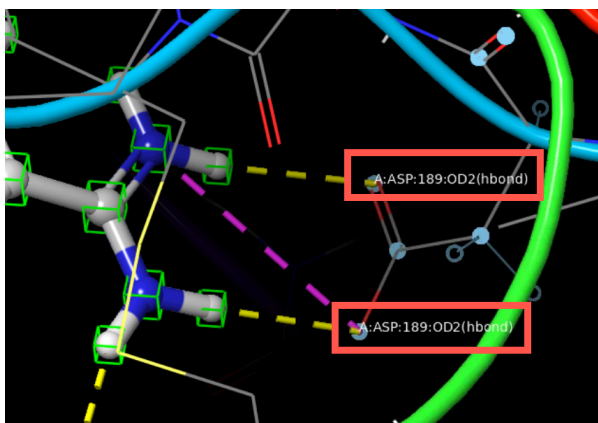


Figure 3-15. Constraint defined on ASP 189.

10. Click an **oxygen atom** of the ASP 189 sidechain
 - Both oxygens are highlighted
 - An H-bond constraint is defined in the Receptor atoms table

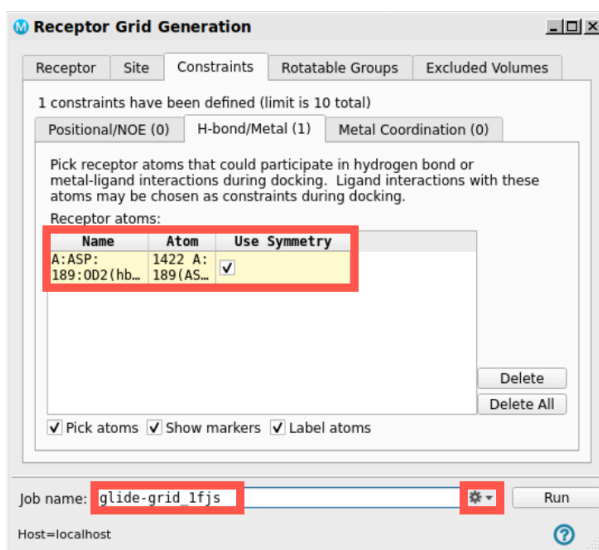


Figure 3-16. Hydrogen-bond constraint is now the receptor grid generation job.

11. Go back to the **Reception Grid Generation** panel and notice that ASP 189 is now listed under the H-bond constraint
12. Change Job name to **glide-grid_1fjs**

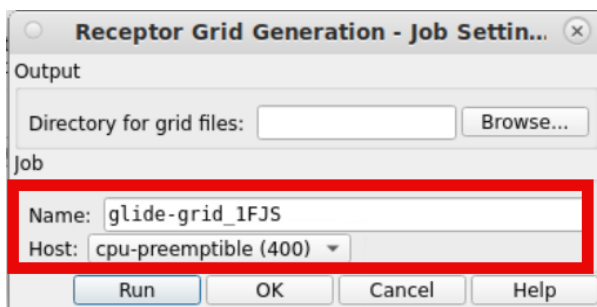


Figure 3-17. Adjusting the job settings

13. Adjust the job settings

- Host: **cpu-preemptible (400)**

14. Click **Run**

- This job requires a CPU host and should complete in under 2 minutes
- A folder named glide-grid_1fjs is written to your Working Directory

Note: All Receptor Grid Generations jobs do not incorporate into the Entry List like other jobs have done in the past. Rather, this creates a folder inside of your Working Directory with a .zip file that we will use for docking with Glide in the next section.



Pause & Think #1: Why is it necessary to generate a receptor grid? What would happen if you proceeded with docking a ligand without a receptor grid?

4. Docking the Cognate Ligand with Hydrogen-Bond Constraint

The minimum requirements for running a Glide virtual screen are a grid file and a ligand file. It is strongly recommended that the grid file be generated from a protein prepared using the Protein Preparation Workflow and the ligand file be prepared using LigPrep. Additionally, you can choose the scoring function, set ligand- and receptor-based constraints, and define the output. Please see the Glide User Manual for more detail. In this section, we will include the hydrogen bonding constraint that was created in the previous step.

In this section, we will dock the cognate ligand to validate the docking protocol which includes the receptor grid and the necessary constraints. The validated protocol can then be used for virtual screening of a ligand library. The information gained from this step can help with evaluating poses and beneficial interactions, which is useful for hit finding. This job will use the receptor grid file that was generated in the previous section.

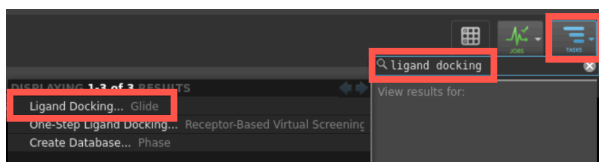


Figure 4-1. Opening Glide Ligand Docking from the Tasks Menu

1. Go to **Tasks** and type **Ligand Docking** in the search bar
2. Select **Ligand Docking: Glide**
 - o The Ligand Docking panel will open

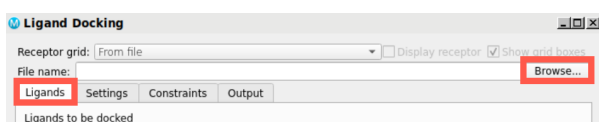


Figure 4-2. The Ligands tab of the Ligand Docking panel.

3. Next to Receptor grid, click **Browse**
4. In the **glide-grid_1fjs** folder, choose **glide-grid_1fjs.zip** and click **Open**

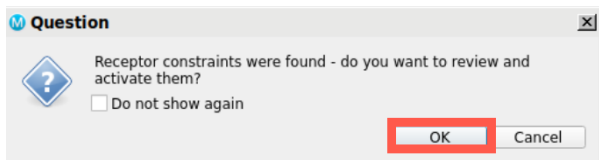


Figure 4-3. A window that verifies receptor constraints.

5. A new window will appear verifying that receptor constraints were found in the receptor grid. Click **OK**.

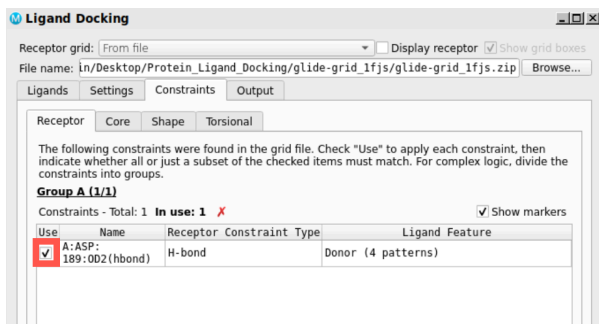


Figure 4-4. The Constraints tab of the Ligand Docking panel.

6. This takes you to the **Constraints** tab. Click on the **Receptor** tab.
7. Under Use, **check** the H-bond constraint for ASP 189

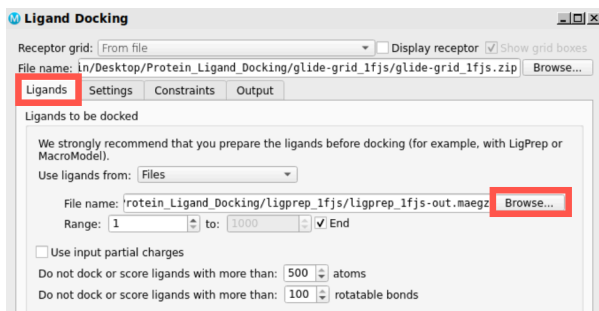


Figure 4-5. The Ligands tab of the Ligand Docking panel.

8. In the Ligands tab, for Use ligands from, choose **Files**
9. Next to File name, click **Browse**
10. Go to the ligprep_1fjs folder and choose **ligprep_1FJS-out.maegz**

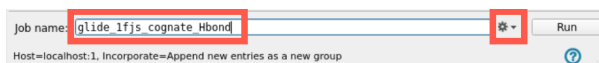


Figure 4-6. Changing the job name.

11. Change Job name to **glide_1FJS_cognate_Hbond**

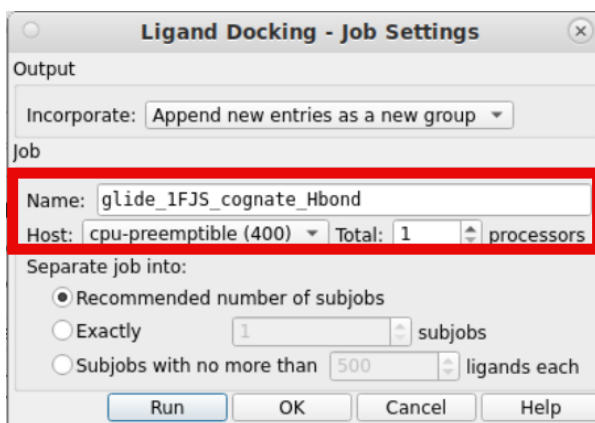


Figure 4-7. Adjusting the Job Settings.

12. Adjust the job settings ()

- Host: **cpu-preemptible (400)**
- Total: **1 processors**

13. Click **Run**

- This job requires a CPU host and should complete in under 2 minutes
- A new group will be added to the Entry List when completed

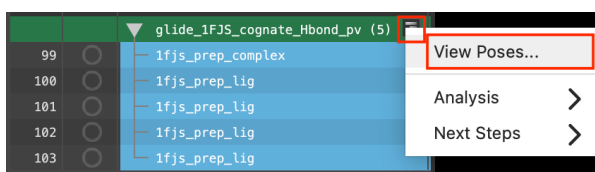


Figure 4-8. View Poses of the cognate ligand. The poses are ranked with the best docking score at the top.

14. Click the **icon** next to the glide-1FJS_cognate_Hbond_pv and select **View Poses**

15. In the Pose Viewer panel, select **Set Up Poses**

16. The 1fjs_prep_complex entry is fixed in the Workspace, the top 1fjs_prep_lig entry is included, and the Pose Viewer panel appears

You may wish to double click the Presets button to have each ligand represented in a unique color.

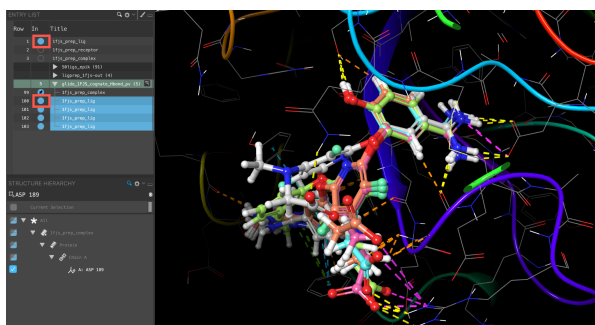


Figure 4-9. Docking result poses shown against the cognate ligand (gray).

17. Include other ligand results

- H-bonds to ASP 189 are conserved

18. Double-click the **In** circle next to **1fjs_prep_complex**

- The entry is no longer fixed in the Workspace

Note: To view all docked ligands in their own color, include all ligands to be evaluated and double click the Presets

button, or include all ligands and choose Binding Mode Comparison in the Presets menu.

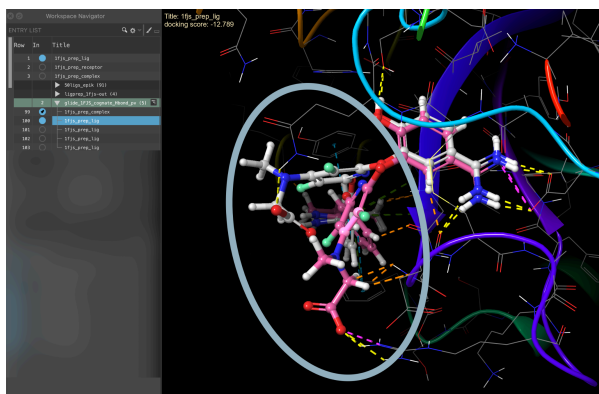


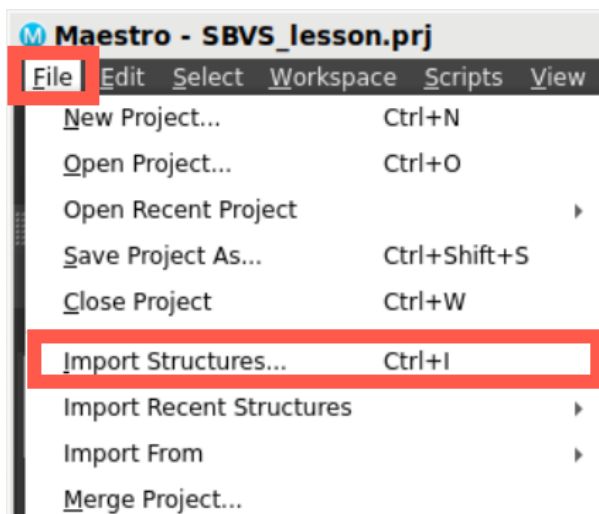
Figure 4-10. An overlay of the best docked pose (pink) with the crystal structure (gray).

19. Include the top docked pose (the pose with the most negative docking score) and **1fjs_prep_lig** to compare the best docked pose to the cognate ligand

As shown in Figure 4-10, the parts of the ligand that form interactions, including the H-Bond at ASP-189, overlap with the cocrystal ligand. However, even in the best docked pose (shown in pink), the solvent-exposed region (circled in gray) does not fully align with that of the cocrystal ligand. It is not surprising given that the binding pocket of 1FJS is solvent exposed and that docking is performed in vacuum. To get a better agreement with the crystal structure pose, there are other constraints that can be considered. Section 7 in the [Protein Ligand Docking lesson](#) focuses on core constraints, which can further improve the alignment between the docked poses and the cocrystal ligand.

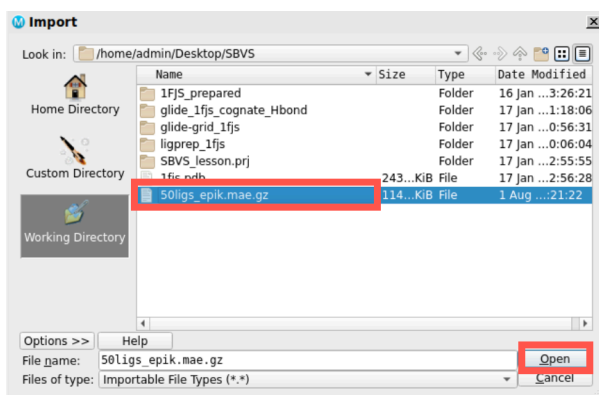
If you would like to go straight to Structure-Based Virtual Screening, continue this lesson with Section 5: Docking the Screening Compounds with Constraints.

5. Docking the Screening Compounds with Constraints



1. Import the prepared ligand library by going to **File > Import Structure**

Figure 5-1. Import structures into Maestro.



2. Select **50ligs_epik.mae.gz** and press **Open**

Figure 5-2. Selecting the pre-prepared ligand library from your Working Directory.

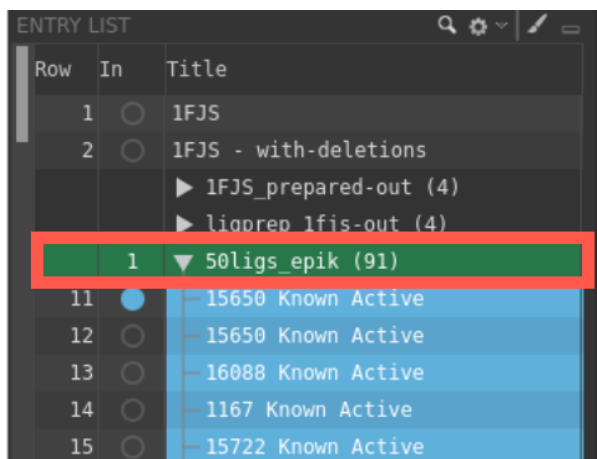


Figure 5-3. Selecting 50ligs_epik in the Entry List.

3. In the Entry List, select the group **50ligs_epik**

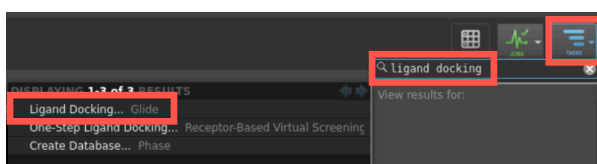


Figure 5-4. Opening Glide Ligand Docking from the Tasks Menu

4. Go to **Tasks** and type **Ligand Docking** in the search bar
5. Select **Ligand Docking: Glide**
 - The Ligand Docking panel will open

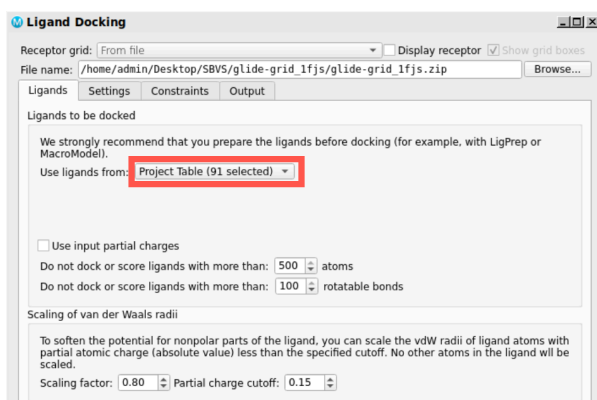


Figure 5-5. Using ligands from selected entries.

6. In the Ligand Docking panel, click the **Ligands** tab
7. For Use ligands from, choose **Project Table (selected entries)**

Note: Keep glide-grid_1fjs.zip as the receptor grid

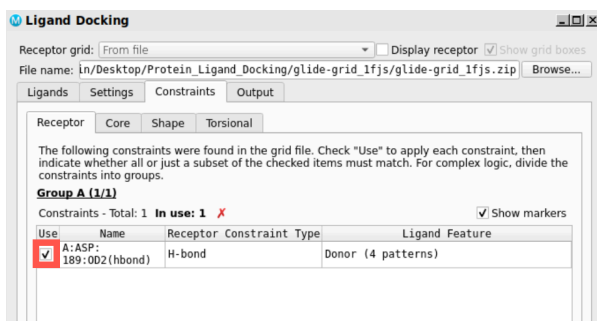


Figure 5-6. The Constraints tab of the Ligand Docking panel.

8. Go to the **Constraints** tab. Click on the **Receptor** sub-tab.
9. Under Use, **check** the H-bond constraint for ASP 189

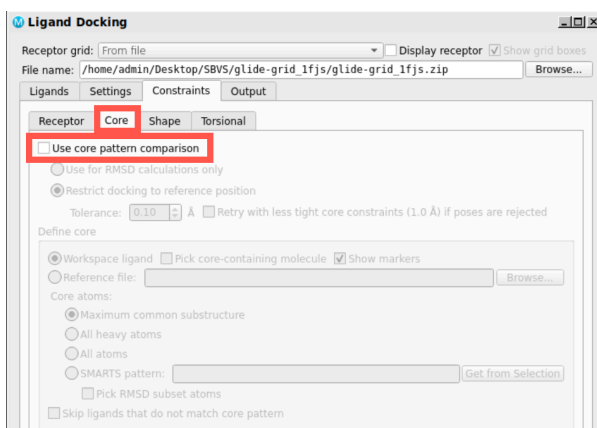


Figure 5-7. Remove the Core Constraint in the virtual screening.

10. Go to the **Core** tab
 11. Ensure that **Use core pattern comparison** is not selected
- Note:** You will dock the screening compounds with only a Hydrogen-Bond constraint

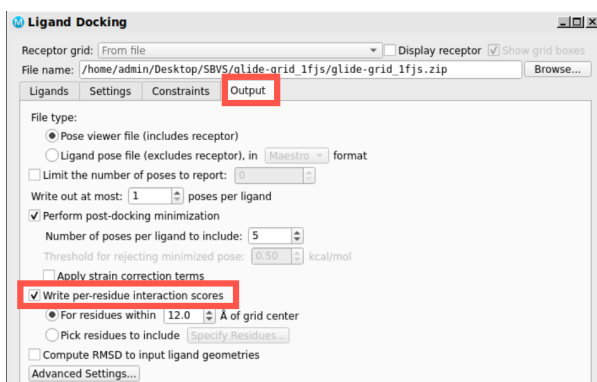


Figure 5-8. The Output tab of the Ligand Docking panel.

12. Click the **Output** tab
13. Check mark **Write per-residue interaction scores**

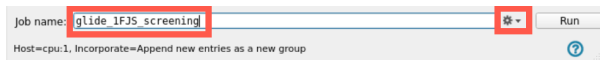


Figure 5-7. Changing the job name.

14. Change Job name to **glide_1FJS_screening**

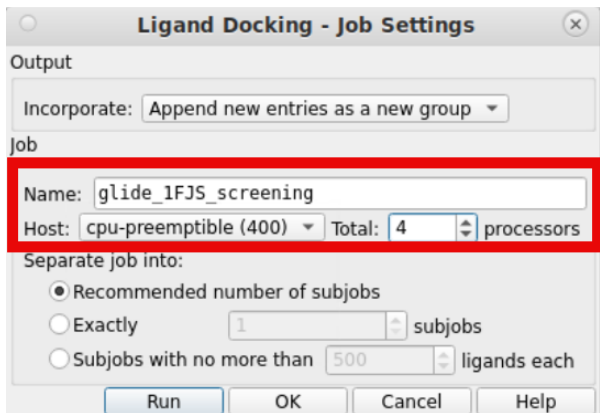


Figure 5-8. Adjusting the Job Settings.

15. Adjust the job settings ()

- Host: **cpu-preemptible (400)**
- Total: **4 processors**

16. Click **Run**

- This job requires a CPU host and should complete in ~10 minutes
- A new group will be added to the Entry List when completed

6. Analyzing Results and Binding-Site Characterization

Multiple Glide docking results can be viewed in the [Entry List](#) and be identified by the job name. Docked results will show the receptor in the first row and the docked ligand(s) in the subsequent row(s), where they are ordered by best to worst docking score, or Glide Gscore if Epik state penalties were not applied in LigPrep. The Glide Gscore is broken down by van der Waals electrostatic components and can be seen in the [Project Table](#), using the Property Tree. You can read more about how docking scores/poses are generated [here](#) and [here](#) and what dependencies they have [here](#) and [here](#).

6.1. Visualize the results using Pose Viewer

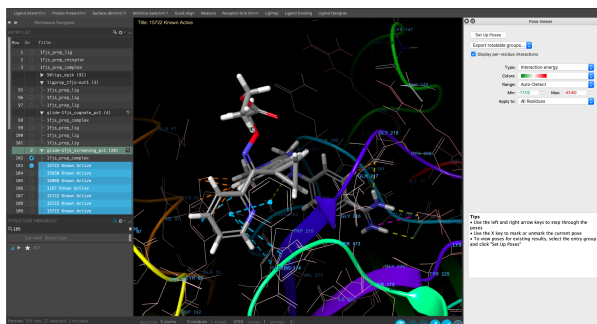


Figure 6-1. Pose Viewer panel.

1. Swipe through the results using the right and left arrow keys
 - Ligand poses are displayed in the Workspace
 - Residues are colored according to their interaction energies, ranging from green (favorable) to red (unfavorable)

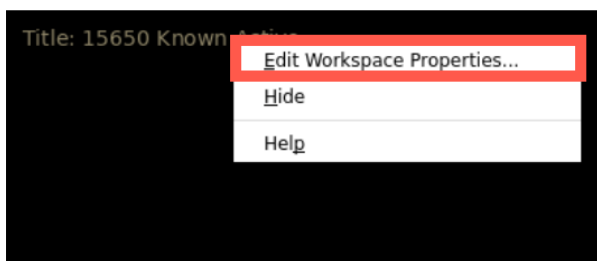
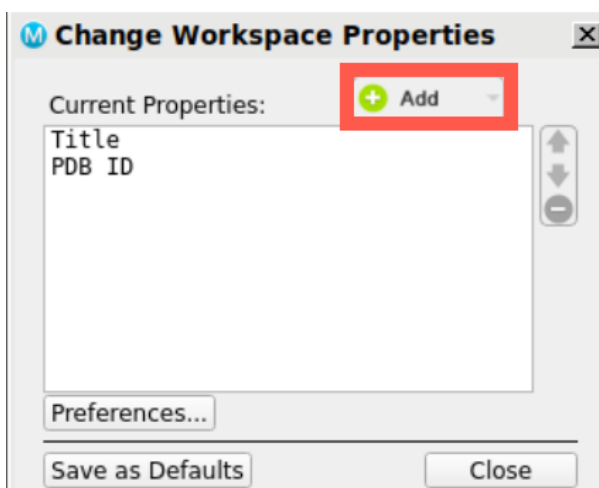


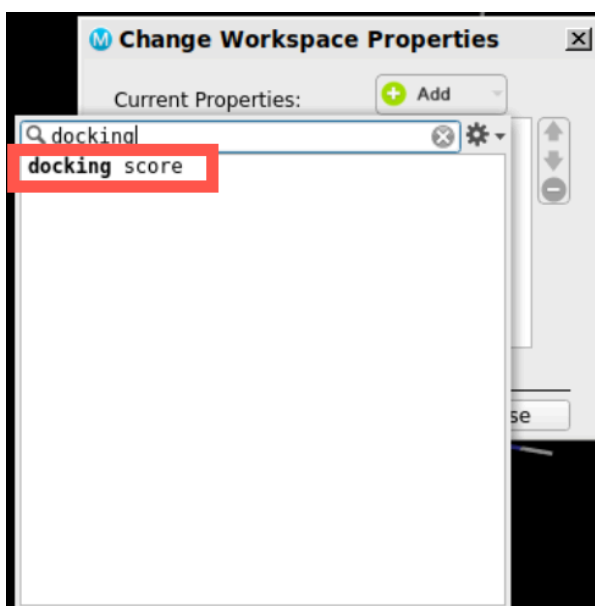
Figure 6-2. Editing Workspace Properties.

2. Right-click the yellow property title in the upper lefthand corner of the workspace
3. Select **Edit Workspace Properties**



4. The **Change Workspace Properties** window opens
5. Select **Add**

Figure 6-3. Adding more properties.



6. Type **docking** in the search bar
7. Add **docking score** to the **Workspace Properties**

Figure 6-4. Adding docking scores.

You will only have to do this once – now every time you perform protein-ligand docking with Glide, the docking score will automatically appear in the upper lefthand corner when looking at a pose in PoseViewer.

Please see [Knowledge Base Article 1027](#) for more information on the difference between docking score, Glide gscore, and glide emodel score.

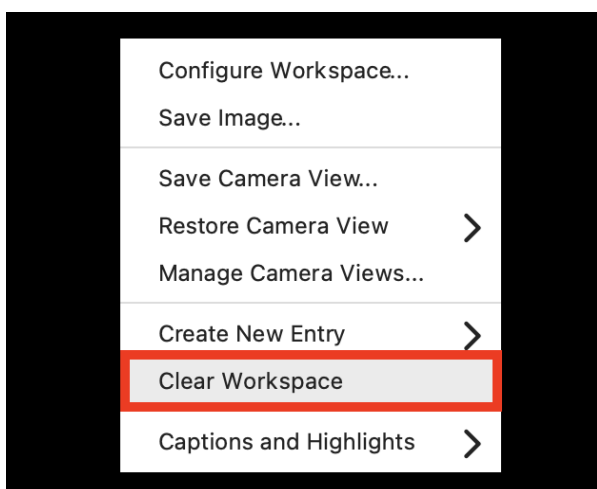


Pause & Think #2: In the table below, select your top 5 ligands with the best docking scores. Remember that the lower the value, the better the docking score. For each ligand, i) write the ligand name and take a screenshot of the pose, ii) list the docking score that can be found in the Project Table, and iii) identify at least 2 types of protein-ligand interactions between that particular ligand and residues within the active site.

Ligand Name & a Screenshot of the Pose	Docking Score (found in Project Table)	Protein-Ligand Interactions

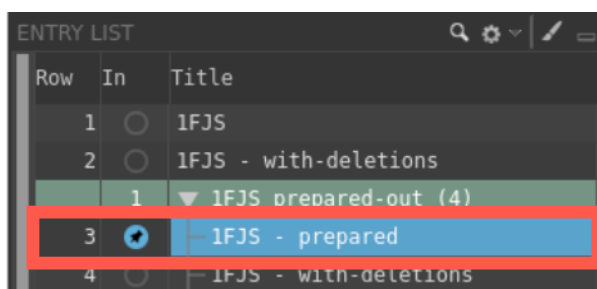
6.1. Visualize the results using Pose Viewer

We will analyze the binding site using SiteMap. SiteMap characterizes hydrophilic, hydrophobic, acceptor, and donor regions of a receptor. This is useful for learning more about an active site, predicting a binding site in an apo structure, or identifying possible allosteric sites. SiteMap ranks the potential binding sites with a druggability score, which can be viewed in the [Project Table](#). The output from a Glide virtual screen can be overlaid with SiteMap information to examine how well the docked ligands explore the various regions in the binding cavity. Sites identified by SiteMap can also be used to create receptor grids for virtual screening experiments. This can be useful for exploring sites without a known active compound.



1. Right-click an empty area in the [Workspace](#) choose **Clear Workspace**

Figure 6-5. Clear Workspace.



2. Double-click the In circle to fix **1FJS - prepared** (in row 3) in the Entry List
3. Go to Tasks and search SiteMap
 - o The SiteMap panel opens

Figure 6-6. Fix 1FJS - prepared in the Workspace.

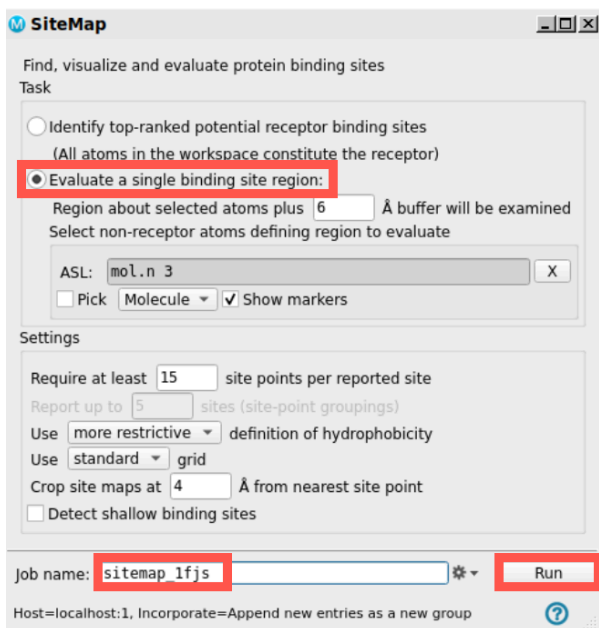


Figure 6-7. SiteMap panel.

4. Under Task, select **Evaluate a Single binding site region**
5. Click on the **ligand** in the Workspace
 - o The ligand is highlighted
 - o SiteMap removes the ligand from the calculation
6. Change the Job name to **sitemap_1fjs**
7. Click **Run**
 - o A banner appears when the job has incorporated
 - o A new group is added to the Entry List

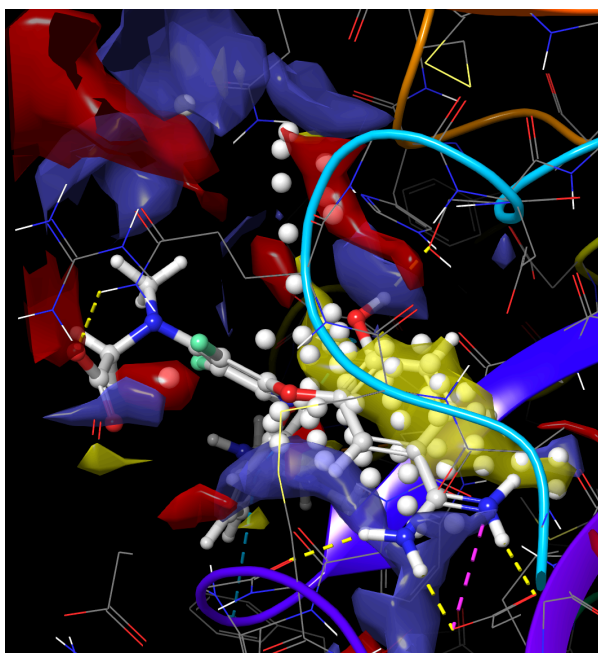


Figure 6-8. SiteMap results in the Workspace.

8. Include **sitemap_1fjs_site_1**, **sitemap_1fjs_ligand**, and **1fjs_protein**
9. Type **L**
 - o Various surfaces are shown representing different regions of hydrophilic property; hydrophobic (yellow), acceptor (red), donor (blue)
 - o The white site-point spheres each represent $\sim 1 \text{ \AA}^3$
10. In the Entry List, click the **S** next to **sitemap_1fjs_site_1** to toggle the surfaces associated with the SiteMap

Note: To find all possible binding sites using SiteMap, under Task select Identify top-ranked potential receptor binding sites.

If you want to detect all pockets, you will need to exclude the cognate ligand from the Workspace.



Pause & Think #3: SiteMap visualization uses a grid of points to identify potential hydrophobic and hydrophilic regions; the hydrophilic regions are further classified into hydrogen-bond donor, hydrogen-bond acceptor, and metal-binding regions, and the surface of the protein is contoured. Take a screenshot of your SiteMap results. Identify which regions of your receptor are hydrophilic and hydrophobic.

7. Individual Exercise

Using the information that you gained from docking the screening ligands, design a new inhibitor that may have a better docking score. Perform LigPrep on your molecule and use Glide to obtain its docking score. Take a screenshot of its pose and paste it below. Then list its docking score. Provide analysis as to why you chose to design this particular inhibitor.

8. Summary, Additional Resources, and References

In this lesson, we imported and prepared a protein and ligand file, then visualized and analyzed the protein-ligand complex. A raw PDB file was made suitable for modeling purposes using the Protein Preparation Workflow, and the cognate ligand was extrapolated using LigPrep in the same fashion that would be used for a multi-ligand file. Then the prepared ligand was docked into the prepared protein using Glide. The Workspace Configuration toolbar allowed for toggling various components in the Workspace and the 2D view in the Ligand Interaction Diagram gave another way to analyze information.

For further learning:

- [Target Analysis with SiteMap and WaterMap](#)
- [Structure-Based Virtual Screening using Glide](#)
- [Ligand-Based Virtual Screening Using Phase](#)
- [Introduction to Molecular Modeling in Drug Discovery Online Course](#)
- [Target Enablement, Preparation, & Validation Online Course](#)

9. Glossary of Terms

Entry List - a simplified view of the Project Table that allows you to perform basic operations such as selection and inclusion

Included - the entry is represented in the Workspace, the circle in the In column is blue

Project Table - displays the contents of a project and is also an interface for performing operations on selected entries, viewing properties, and organizing structures and data

Recent actions - This is a list of your recent actions, which you can use to reopen a panel, displayed below the Browse row. (Right-click to delete.)

Scratch Project - a temporary project in which work is not saved. Closing a scratch project removes all current work and begins a new scratch project

Selected - (1) the atoms are chosen in the Workspace. These atoms are referred to as "the selection" or "the atom selection". Workspace operations are performed on the selected

atoms. (2) The entry is chosen in the Entry List (and Project Table) and the row for the entry is highlighted. Project operations are performed on all selected entries

Working Directory - the location that files are saved

Workspace - the 3D display area in the center of the main window, where molecular structures are displayed