Docking ligands to a receptor and computing scores for the docked poses

Purpose: Introduces setting up and running a docking protocol, docking a ligand, and analyzing the scores.

Modules: Discovery Studio

Time:

Prerequisites: Docking, Consensus Score

Introduction

This tutorial focuses on the computational methods that a computational chemist performs. This tutorial covers:

- Loading the protein receptor
- Defining the receptor and searching it for binding sites
- Running the Docking protocol
- Analyzing the docking results
- Minimizing the docked poses
- Rescoring the minimized ligand poses
- Analyzing the scored minimized poses
- Running a Consensus Score protocol
- Analyzing the Consensus Score results

1. Loading the protein receptor

Start Discovery Studio.

In the Files Explorer, right-click a directory to which you would prefer to save protocol data. Click **Set Default** from the context menu.

This sets the default working directory. When a protocol is run, a subfolder is created in this directory containing the protocol's input and output files.

Choose **File** | **Open...** from the menu bar.

This displays the Open dialog.

Navigate to the tutorials directory and open the file **pdb1kim_protH.msv**.

Note. For information about locating data files, refer to the Tutorials - Receptor-Ligand Interaction Help topic.

This opens the structure in a 3D Structure View. This protein has been already prepared for this lesson. Ligands and all crystallographic waters have been removed, and hydrogens have been added.

2. Defining the receptor and searching it for binding sites

Check to see if the Binding Site tools appear in the Tools Explorer. If the Binding Site tools are not visible, select **View | Tool Panels | Binding Site** from the menu bar. If necessary, double-click the panel header to display the commands within the panel.

This displays the Binding Site tools.

In the 3D Structure View, select any atom of the protein receptor by clicking it.

The selected atom is highlighted with a yellow square.

In the Binding Site tools, click **Define Selected Molecule as Receptor.**

This defines the protein molecule as the receptor. If a Hierarchy View is open, you see **SBD_Receptor** appear as a new group icon. If the Hierarchy View is not open and you wish to see it, select **View | Hierarchy** from the menu bar.

Now, click **Find Sites from Receptor Cavities.**

This employs a cavity detection algorithm that identifies binding site cavities inside the protein receptor. In this case, nine sites are identified. The sites are sorted by size, and the largest site is displayed.

Note. The Binding Site tools may be used to browse the list of sites using the **Next Site** and **Previous Site** commands under **Display**. Each site may be modified with the **Contract Binding Site** and **Expand Binding Site** commands under **Site Editing**.

3. Running the Docking protocol

In the Protocols Explorer, under the **Receptor-Ligand Interactions** group, open the **Docking** protocol by double-clicking it.

This opens the Docking protocol in the Parameters Explorer.

In the Parameters Explorer, the **Input Target Receptor** should be set to **pdb1kim_protH:1kim_proth** and **Input Binding Site** should show **Site 1 ...** as the selected binding site along with the number of points, volume in Angstroms cubed, and partition level of 1.

Note. The **Input Target Receptor** parameter lists all molecules in all 3D Windows using the syntax of "Window name:Molecule name" as potential target receptors.

Click the **Parameter Value** cell next to the **Input Ligands** parameter, and click the  button at the far right of the cell.

This opens the **Specify Ligands** dialog.

Click the **All ligands from a file** toggle and click the  button.

This opens the **Choose a ligands file** dialog.

Navigate to the tutorials directory and open the file **TK_xray_ligs.sd**.

Note. For information about locating data files, refer to the Tutorials - Receptor-Ligand Interaction Help topic.

This selects the file and populates the name into the **Specify Ligands** dialog.

Click **OK** to close the **Specify Ligands** dialog.

The specified SD file is now entered into the **Parameter Value** cell for the **Input Ligands** parameter of the **Docking** protocol.

Click the **Parameter Value** cell next to the **Energy Grid Forcefield** parameter and select **PLP1**.

This specifies **PLP1** as the energy function for docking.

Note. The simple and short-range nature of this function allows for faster docking runs compared with using **Dreiding** or **CFF**, so it has been chosen for use in this tutorial.

Click the **Parameter Value** cell next to the **Conformation search Number of Monte Carlo Trials** parameter. Clear the text in the cell and enter the value **5000**.

This specifies a fixed value of **5000** Monte Carlo trials to sample for each ligand for the conformational searching step.

Note. The default entry for this parameter specifies the number of steps as a function of the number of ligand torsions resulting in longer searches for more flexible ligands. The choice of **5000** enables the tutorial to run somewhat faster without significantly degrading the quality of the results for the purpose of this tutorial.

Click the **Parameter Value** cell next to the **Pose saving Maximum Poses retained** parameter. Enter the value **5**.

This specifies a value of up to **5** poses to save for each ligand. Only poses that are distinct based on RMS and energy criteria are saved.

Click the **Parameter Value** cell next to the **Scoring Scores** parameter. In the popup list that appears, check the scores for **LigScore2, PLP1, Jain, PMF,** and **Ludi Energy Estimate 3**.

The specified scoring functions will be calculated for each docked ligand pose when the protocol is run.

From the Parameters Explorer, click the  **Run** button.

This runs the Docking protocol. This job typically takes several minutes to complete. The status of the job can be monitored in the Jobs Explorer.

4. Analyzing the docking results

After the job is complete, click the **Start Date** column in the Jobs Explorer, and ensure that the column is sorted in descending order such that the most recent job appears at the top. If the arrow is pointing upward, click again so the arrow is pointing down. Double-click any cell in the row corresponding to the **Docking** run. If you have not run any other jobs after submitting the docking run, this should be the first row in the Jobs Explorer.

This opens the Files Explorer to the Output folder of the Docking job.

Double-click the **BrowseMolecules.pl** file in the Output folder.

This simultaneously opens the resulting docked ligand poses of the Docking job into a Table Browser and opens an associated 3D View containing the first docked ligand pose and the protein receptor used for the calculation.

Right-click any cell of the Table Browser, and select the **Group By...** command.

This opens the Group By dialog.

In the **Select column to group by** popup, select **MOL_NUMBER** and click **OK**.

The Table Browser is now split into two views. The left view contains a list of all ligands.

Double-click the first ligand in the right view.

The right view now contains a list of all poses for the first ligand.

Click any cell in the right Table Browser. Now use the up and down arrow keys to scroll through the list of poses.

Observe that the docked pose changes in the 3D View.

Note. You may find it easier to view the poses if the binding site points are not also displayed. To hide the display of the binding site, double-click a binding site point in the 3D View to select the entire site. Then, right-click and select **Hide**. You can alternatively hide the receptor by double-clicking any of its atoms and then double-clicking the selected amino acid to select the whole protein. Then, right-click and select **Hide**.

Select another row in the left Table Browser

Observe that the 3D View now displays the first pose of the selected ligand, and the right Table Browser now contains a list of poses for the selected ligand.

Right-click in the left Table Browser, and use the **Group By...** command to return the Group By selection to **<None>**.

The split view of the Table Browser disappears and all poses of all ligands are now shown in the single Table Browser.

Click the column heading labeled **-PLP1**. Now, with the **CTRL** key depressed, also click the column labeled **LigScore2_Dreiding**.

Both columns are highlighted in the Table Browser.

Choose **Chart | Point Plot** from the menu bar.

A point plot is generated with the **-PLP1** score plotted on the X-axis and the **LigScore2_Dreiding** score on the Y-axis for all the ligands and poses. A number of the LigScore2 values are negative. This is because the PLP1 docking function has a very soft repulsive core, whereas LigScore2 tends to strongly penalize poses forming bumps with the receptor.

In the plot, use the mouse to select the lowest point on the plot by dragging the cursor around the point.

The point is highlighted in yellow.

Return to the **docked - Table Browser** window and scroll the table down to row 19.

This row is now highlighted corresponding to the point picked in the plot. The value of the **LigScore2_Dreiding** cell in this row is approximately -13.5. The ligand pose is now displayed in the 3D View.

In the 3D View, right-click empty space and select **Receptor-Ligand Bumps** from the context menu.

Short contacts between the ligand pose and the protein receptor are displayed in magenta.

5. Minimizing the docked poses

In the Protocols Explorer, double-click **In-Situ Ligand Minimization** under the **Receptor-Ligand Interactions group**.

The **In-Situ Ligand Minimization** protocol opens in the Parameters Explorer.

In the cell next to **Input Receptor**, select **pdb1kim_protH:1kim_proth**.

Note. If you have previously closed the pdb1kim_protH window, please re-open this molecule into a new 3D Window as described above in Section 1 under **Load the protein receptor**.

This selects the protein as the receptor molecule for the **In-Situ Ligand Minimization** protocol.

In the cell next to **Input Ligand File**, click the  button at the far right of the cell.

This opens the **Specify Ligands** dialog.

Click the **All ligands from a table browser** toggle. The item **docked** should already be selected as the name of the Table Browser. Click **OK** to close the dialog.

The name **docked** appears in the cell next to **Input Ligand File**.

Accept all other defaults. Click the  **Run** button to run the protocol.

The protocol runs, and a new row appears in the **Jobs** Explorer table. This protocol can take about 20 minutes to complete.

After the run completes, the value **Success** appears in the **Status** column of the job in the **Jobs** Explorer. In the **Jobs** Explorer, double-click any cell of this job.

This opens up the **Output** folder for this job in the Files Explorer. Note where this folder resides in the folder hierarchy.

6. Rescoring the minimized ligand poses

In the Protocols Explorer, double-click the **Scoring** protocol.

The **Scoring** protocol opens in the Parameters Explorer.

In the cell next to **Input Target Receptor**, select **pdb1kim_protH:1kim_proth**.

pdb1kim_protH:1kim_proth is now displayed in the cell. The **Parameter Value** for **Input Binding Site** should

automatically update and show **Site 1** as the selected binding site along with the number of points, volume in Angstroms cubed, and partition level of 1.

Click the **Parameter Value** cell next to the **Input Ligands** parameter, and click the  button at the far right of the cell.

This opens the **Specify Ligands** dialog.

Click the **All ligands from a file** toggle and click the  button.

This opens the **Choose a ligands file** dialog.

Browse to the **Output** folder of the **In-Situ Ligand Minimization** job, and double-click the file **minimized.sd**.

This selects the file and populates the name into the **Specify Ligands** dialog.

Click **OK** to close the **Specify Ligands** dialog.

The specified SD file is now entered into the **Parameter Value** cell for the **Input Ligands** parameter of the **Scoring** protocol.

Click the **Parameter Value** cell next to the **Scoring Scores** parameter. In the popup list that appears, check on the scores for **LigScore2, PLP1, PMF, Jain,** and **Ludi Energy Estimate 3**.

The specified scoring functions will be calculated for each docked ligand pose when the protocol is run.

From the Parameters Explorer, click the  **Run** button.

This runs the Scoring protocol. This job typically takes about a minute to complete. The status of the job can be monitored in the Jobs Explorer.

7. Analyzing the scored minimized poses

After the **Status** of the **Scoring** job is shown as **Finished** in the Jobs Explorer, double-click in the row corresponding to that job.

This opens the Files Explorer to the Output folder of the Scoring job.

Double-click the **BrowseMolecules.pl** file in the Output folder.

This simultaneously opens the scored ligand poses of the job into a Table Browser, and opens an associated 3D View containing the first ligand pose and the protein receptor used for the calculation.

Scroll down the Table Browser to row 19, which should correspond to **MOL_NUMBER** 4 and **POSE_NUMBER** 4. Use the column scroll bar to scroll over to the column with the **LigScore2_Dreiding** heading.

The **LigScore2_Dreiding** value for this pose is now approximately **5**. Using the unminimized pose, the LigScore2_Dreiding value was approximately -13.5, so the minimization has significantly improved the **LigScore2_Dreiding** estimate for this pose.

Click the column heading labeled **-PLP1**. Now, with the **CTRL** key depressed, also click the column labeled **LigScore2_Dreiding**.

Both columns are highlighted in the Table Browser.

Choose **Chart | Point Plot** from the menu bar.

A point plot is generated with the **-PLP1** score plotted on the X-axis and the **LigScore2_Dreiding** score on the Y-axis for all the ligands and poses. There is a significantly improved correlation between the **PLP1** and **LigScore2_Dreiding** values for the minimized poses compared with the plot previously generated using the unminimized poses from the Docking protocol.

8. Running a Consensus Score protocol

In the Protocols Explorer, double-click the **Consensus Score** protocol.

The **Consensus Score** protocol opens in the Parameters Explorer.

Click the **Parameter Value** cell next to the **Source** parameter, and click the  button at the far right of the cell.

The **Specify Ligands** dialog opens.

Click the **All ligands from a table browser** toggle, and select **scored** as the name of the Table Browser. Click **OK** to close the dialog.

The name **scored** appears in the cell next to **Source**.

Click the **Parameter Value** cell next to the **Scores for Consensus** parameter.

The **Scores for Consensus** dialog opens with 2 listboxes displayed. The list of all available fields from the Table Browser available to choose for the Consensus Score calculation is shown in the listbox on the left. The listbox on the right is empty.

With the CTRL key depressed, select the following entries from the list: **-PLP1, -PMF, Jain, LigScore2_Dreiding,** and **Ludi_3**.

The selected entries are highlighted.

Click the **>>** button to transfer the selection to the right listbox. Click **OK** to close the dialog.

The selected scores are displayed in the cell next to the **Scores for Consensus** parameter in the Parameters Explorer.

From the Parameters Explorer, click the  **Run** button.

This runs the Consensus Score protocol. This job typically takes less than a minute to complete. The status of the job can be monitored in the Jobs Explorer.

9. Analyzing the Consensus Score results

After the **Status** of the **Consensus Score** job is shown as **Finished** in the Jobs Explorer, double-click in the row corresponding to that job.

This opens the Files Explorer to the Output folder of the Consensus Score job.

Double-click the **consensus.sd** file in the Output folder.

This should open the file into a Table Browser labeled consensus.

Note. You may need to ensure that the File Types preferences for sd files is set to open in a Table Browser. This can be done using the **Edit | Preferences...** command from the menu bar to open the **Preferences** dialog. Select the **File Types** preference in the hierarchy list located under the **Files Explorer** group. Ensure that the **Window Type** for **sd** files is set to **Table Browser**.

A new column has been added labeled **Consensus**.

Double-click the **Consensus** column header twice to sort the table by Consensus score in descending order (i.e. higher scores are at the top).

Observe that the first row has a **Consensus** score of **5** and corresponds to the **1e2k_lig** molecule. Scrolling the table to the right shows that this entry corresponds to **POSE_NUMBER 5** for this molecule. Examining the values of the scores for the 5 scoring functions used in the calculation shows they all generated a high score for this pose resulting in its Consensus score value of **5**.

This is the end of the tutorial.