

Superimposing protein structures

Required functionality and modules: Discovery Studio Client.

Required data files: kinases.dsv and kinases.pir.

Time: 10 minutes.

Introduction

The first step in comparing different protein structures is to superimpose the structures based on either their C-alpha or backbone atoms. Since different proteins have different numbers and types of residues, the residues to be used must be known to map the coordinates from different protein structures.

Discovery Studio provides several methods for superimposing protein structures. One is to manually specify matching residues between two proteins and superimpose them based on those matching residues. A more general approach is to superimpose two or more protein structures based on their sequence alignment, and then the matching residues for superimposition are automatically identified based on the aligned residues in the sequence alignment. The second approach can also be used to superimpose the protein structures based on a subset of aligned residues, such as ligand binding site residues.

In this tutorial you will learn how to superimpose a set of proteins based on their sequence alignment

- [Open and view a set of protein molecules and their sequence alignment](#)
- [Superimpose protein structures based on their sequence alignment](#)
- [Select the ligand binding site residues](#)
- [Superimpose the protein structures using the binding site residues](#)

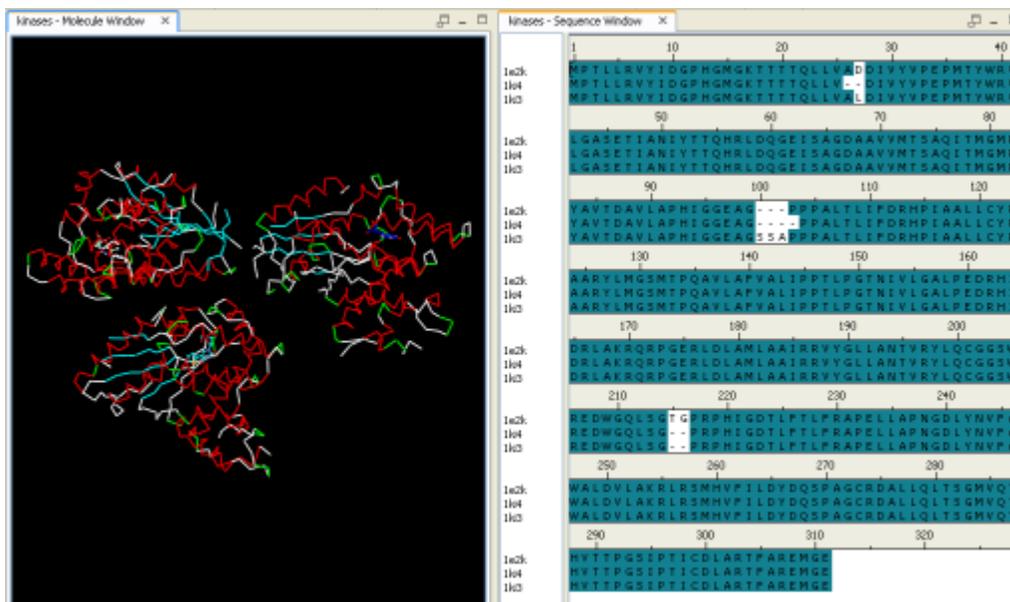
Open and view a set of protein molecules and their sequence alignment

From the Files Explorer, navigate to the **Samples | Tutorials | Quick Start Tutorials** folder and open the protein structure file `kinases.msv` and sequence alignment file `kinases.pir`.

The files are opened two new windows, a Molecule Window with three protein structures, 1e2k, 1ki4, and 1ki3 and a Sequence Window with the corresponding aligned sequences. Close the Hierarchy and Data Table Views if they are open by pressing the CTRL+H and CTRL+T keys.

Drag the **Sequence Window** tab to the lower part of the Molecule Window so that they can be viewed side by side.

The sequences in the Sequence Window are linked to the corresponding proteins in the Molecule Window when the two files are opened. The order in which files are opened does not matter. Any residue selection in the Molecule Window are reflected in the Sequence Window and vice versa. Click on individual or multiple residues in the Sequence Window to demonstrate this.



Superimpose protein structures based on their sequence alignment

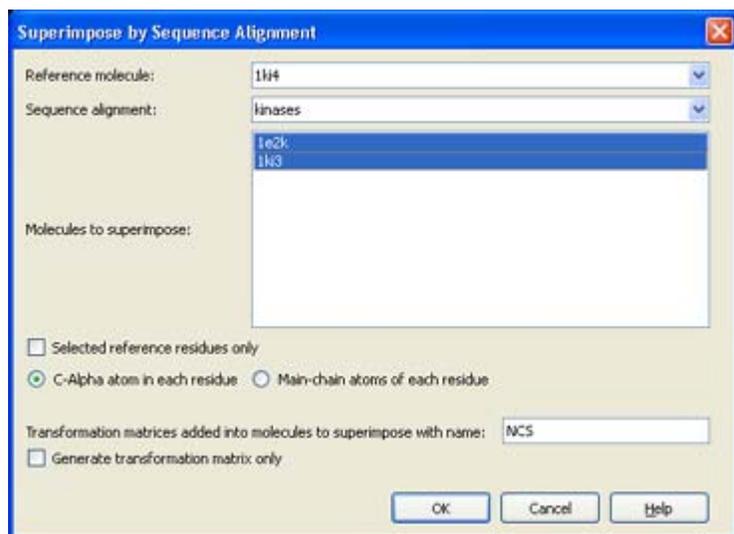
Click the **kinases - Molecule Window** tab to make it active.

From the menu bar, choose *Structure / Superimpose / By Sequence Alignment...* to display the Superimpose by Sequence Alignment dialog.

From the dialog, select 1ki3 as the **Reference molecule** from the list.

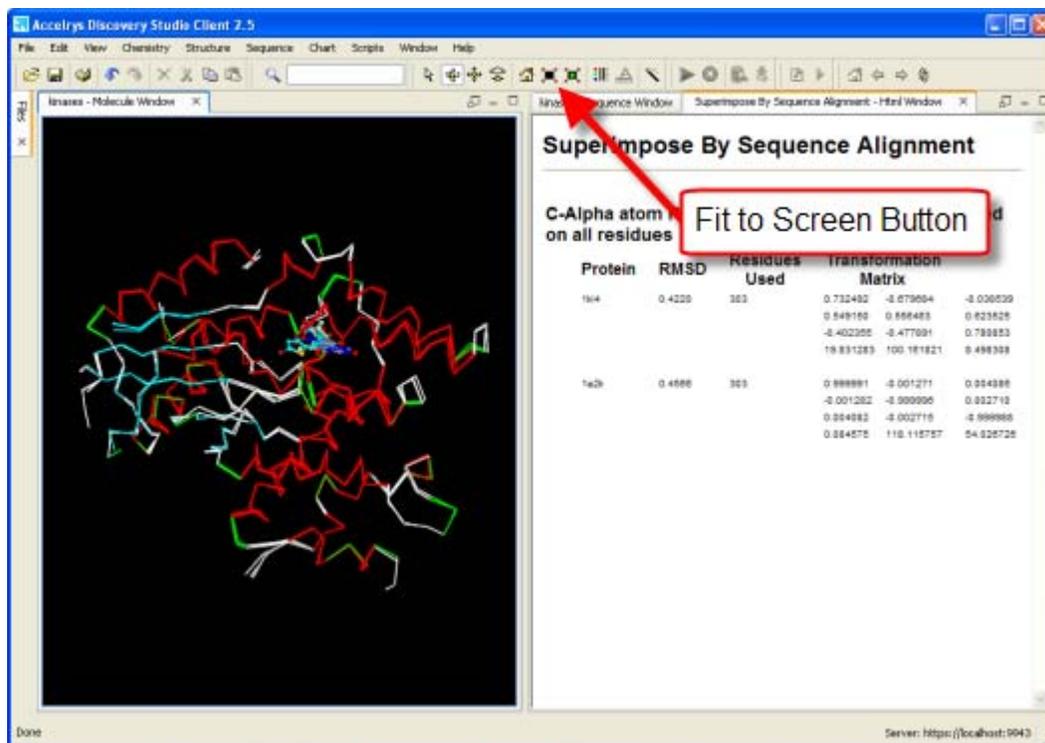
Select 1ki4 and 1e2k as **Molecules to superimpose** by clicking and dragging the over the molecule names.

Select the **Main-chain atoms of each residue** radio button.



This superimposes 1ki4 and 1e2k over the 1ki3 structure based on the aligned residues defined in the Sequence Window. The three structures in the Molecule Window are now on top of each other.

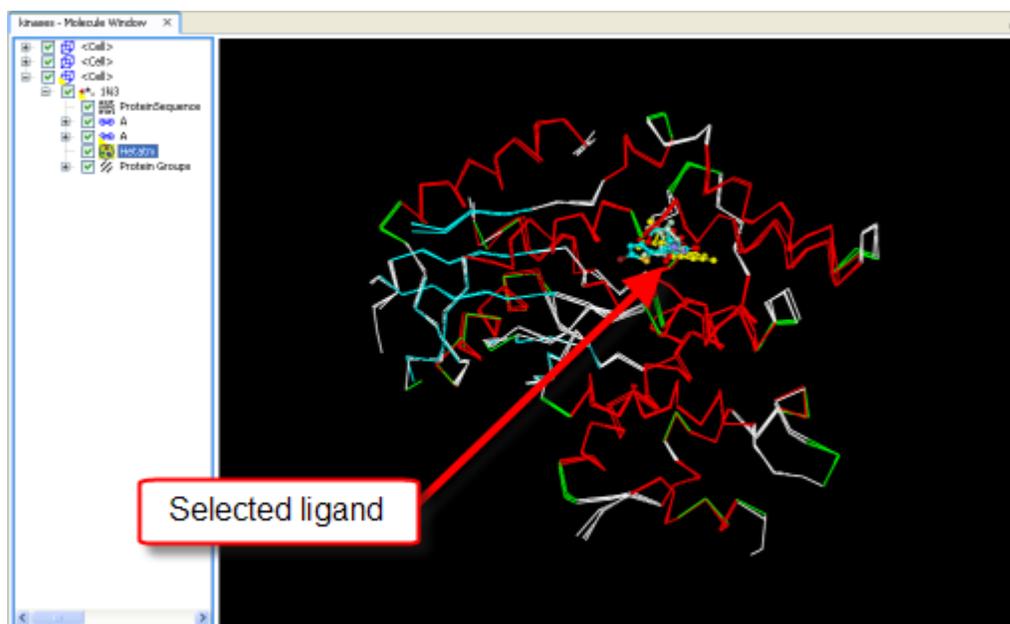
The command also brings up a report that summarizes the RMSD of the superimposition and the number of residues used to superimpose the proteins. Close this window after inspecting the results so that the Molecule Window and Sequence Window are both available to work with for the next section.



Select the ligand binding site residues

Open the Hierarchy View by clicking the Ctrl+H keys. From the **Hierarchy View**, click **Hetatm** under the 1k13 molecule to select its ligand.

This highlights the ligand of 1k13 in the Graphics View.



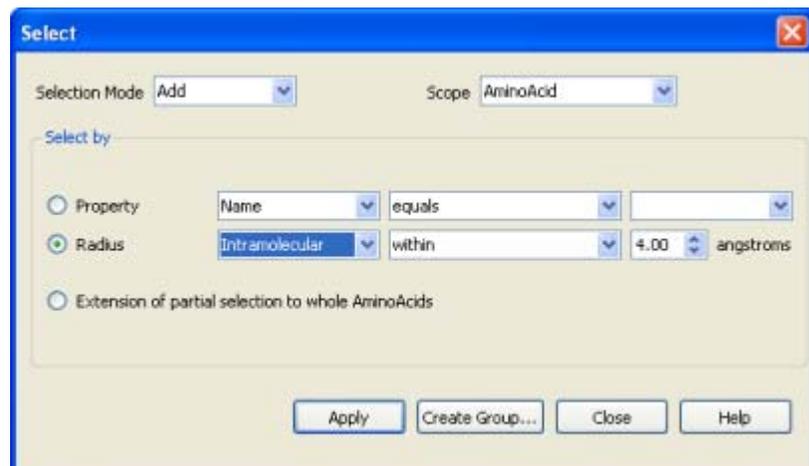
From the menu bar, choose *Edit | Select...* to display the Select dialog.

From the dialog, click the *Scope* checkbox and select *AminoAcid* from the list.

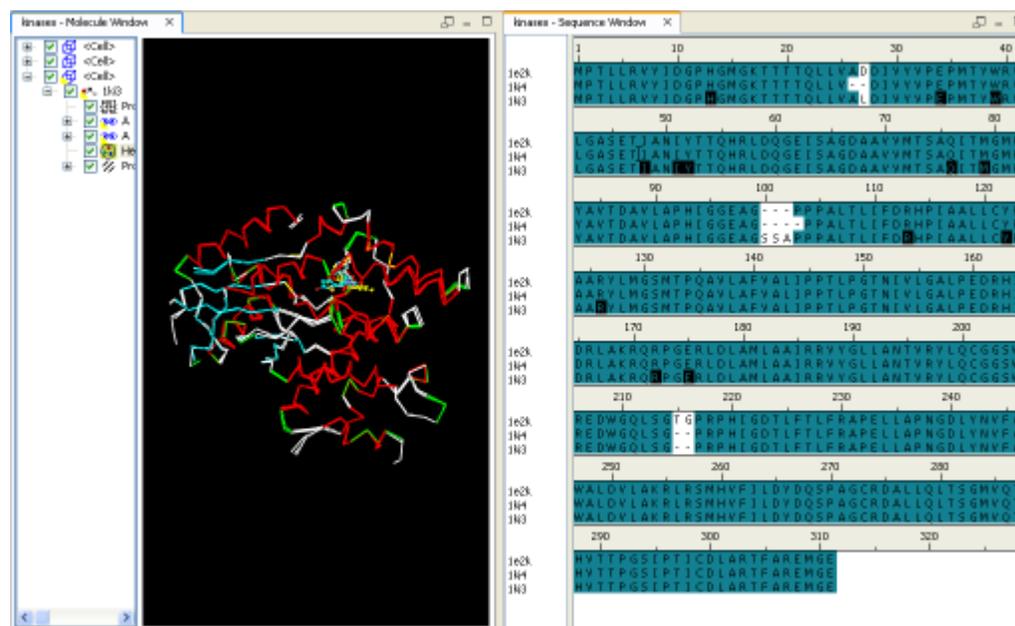
Select the *Radius* radio button and select *Intramolecular* from the first list.

Enter 4 into the distance box.

Click **Apply**, and then **Close**.



This will select any residues in 1ki3 that have an atom within four angstroms of any of the selected ligand atoms. Those selected residues are highlighted in both the Molecule Window and the Sequence Window.



Superimpose the protein structures using the binding site residues

Click the **kinases - Molecule Window** tab to make it active.

From the menu bar, choose *Structure | Superimpose | By Sequence Alignment...* to display the Superimpose by Sequence Alignment dialog.

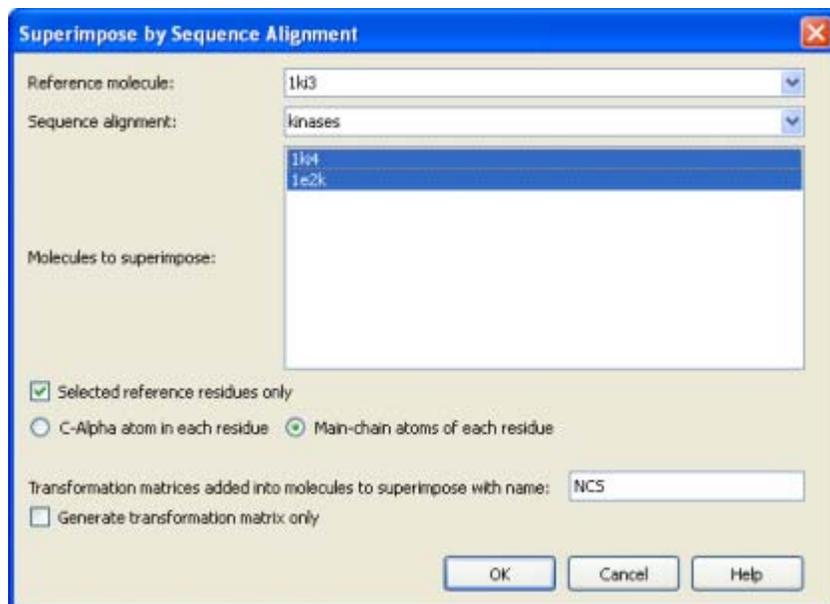
From the dialog, select 1ki3 as the **Reference molecule** from the list.

Select 1ki4 and 1e2k as **Molecules to superimpose** by clicking and dragging over the molecule names.

Check the **Selected reference residues only** checkbox.

Check **Main-chain atoms of each residue**.

Click **OK**.



This superimposes 1ki4 and 1e2k over the 1ki3 structure based on the selected residues in the 1ki3 protein. The residues in 1ki4 and 1e2k, which are aligned to the selected residues in 1ki3, will be used to superimpose the three proteins.

As in the previous step, a html report is generated that provides a summary of the RMSD of the superimposition and the number of residues used to superimpose the proteins. There are fewer residues used in this step than the previous step and the RMSD values are slightly lower as well.

Superimpose By Sequence Alignment

Main-chain atom RMSD to reference protein: 1ki3 based on selected reference residues

Protein	RMSD	Residues Used	Transformation Matrix		
1ki4	0.2547	13	0.999970	0.000004	-
			-0.000004	0.999999	0.001405
			0.000907	0.001405	0.999975
			-0.287713	-0.138962	0.430632
1e2k	0.2326	13	0.999998	-0.001740	0.001089
			0.001740	0.999997	-
			-0.001089	0.001984	0.999999
			-0.078878	-0.012491	0.114977