Introduction

In biochemistry, a kinase is an enzyme that catalyzes the transfer of phosphate groups from high-energy, phosphate-donating molecules to specific substrates, a process known as phosphorylation. The phosphorylation state of a molecule, whether it be a protein, lipid, or carbohydrate, can affect its activity, reactivity, and ability to bind other molecules. Therefore, kinases are critical in metabolism, cell signalling, protein regulation, cellular transport, secretory processes, and many other cellular pathways, which make them very important to the human body.
ABL1 gene encodes a kinase that implicates the cell differentiation, cell division, cell adhesion, and stress response. Mutations in the ABL1 gene are associated with chronic myelogenous leukemia (CML), a cancerous disease that allows the cells to proliferate without regulation. Therefore, ABL1 kinase inhibitors are vastly investigated and studied, and several have been approved for medical uses.
In this laboratory, we will investigate a series of kinase inhibitors with the ABL1 kinase receptor and try to get a glance of how to use MD tools to obtain binding energy, binding mode, interaction patterns and etc... The prepared 15 kinase inhibitor complex structures include 5 non-binders (IC50>10mM), 10 binders (IC50<10mM), and 15 kinase structures. PDB crystal data of ABL1 kinase are used to build the complex structures of Ex1 to Ex9. While the ligand poses of the rest complexes are adopted from the docking calculations.

Downloading input package

After login the Mac desktop, click Launcher and type Terminal <Enter> in the search window. In the terminal window, use curl -O (it is a Oh) on the following URL to download the tgz archive bindinglab.tgz into your current directory.

```
Curl -O
https://ronlevygroup.cst.temple.edu/courses/2017_spring/bio5412(bindinglab.tgz
```

Extract the contents by tar -xzvf bindinglab.tgz.

Run simulations on background

Enter the created folder from the extracting via cd bindinglab. Run preparations and simulations using a script as below. Please replace the exfolder as your assigned one.

```
nohup ./prepAndRunMD.sh exfolder &> prepAndRunMD.nohup &
```
prepAndRunMD.sh is a bash script containing all three individual runs for the minimization, thermalization, and production processes. You can use vi or other text editors to open it and take a look. exfolder is the folder containing two maegz files for the starting ligand and receptor structures. After the submission, you can find the *.inp files and *.out files in the exfolder. Please use a desired text editor to read the contents inside the *.inp and *.out files for all three processes and make sure the thermalization and production run is running before your leave (The ex*_mdrun.out exists). nohup is a unix command to put the job on the background and keep the job running even after you logout the desktop. We will extract simulation data and perform analysis in the next week.

Extract simulation data

Enter the previously created folder bindinglab. Run the following script to extract the time series of related quantities and conformational snapshots.

```
./extractData.sh exfolder
```

The resulting time series of data are written into *.dat files under the exfolder. There are six columns corresponding to print ID, instantaneous temperature, total energy, total potential energy, lambda, and binding energy. All extracted conformations are saved to a single *_rdtrj.maegz file. Here the last column “binding energy” represents the interaction energy (lower indicates better interaction) between the inhibitor and the kinase.

Maestro analysis tools

All prepared 15 complexes can be loaded into Maestro from the folder bindinglab/ligands. The simulated structures (in *maegz format) for individual complexes can be found under each ex* folder. Some important analysis tools in Maestro are listed as follows. You are also encouraged to use the online help documentation.

1. Measurements of distances, angles, and dihedrals: first, click Tools then go to Measurements and pick the desired option. Then go back to the workspace and select related atoms. The resulted number will appear on the screen. You can delete the measurement by Tools->Measurements->Delete Measurements.
2. Contacts information: After including both the ligand and the receptor in the workspace, select Contacts under the Tools menu. In a separate window, check or uncheck related boxes for the types of contacts to show them. The image can be saved in the Workspace menu.
3. Protein-ligand interaction diagram: After including both the ligand and the receptor in the workspace, select Ligand Interaction Diagram in the menu of Tools. The interactions such as H-Bonds, Salt Bridges, and Interfacial Water will be generated automatically. This diagram can be saved as a screenshot in the File menu.
4. Structural superposition: After including or selecting structure entries, enable it as Tools->Superposition. In the tab of “ASL”, specify the atoms to be superposed using Atom Selection Language such as “(atom.ptype "CA")”, which can be built by click the Select button.

Data analysis and questions

1. Plot the time series of instantaneous temperature, potential energy, and binding energy including the data from all three processes. Please explain their trends and fluctuations during the minimization, thermalization, and production processes.

2. Using excel or other tools to obtain the statistical quantities such as the minimum, maximum, averaged value and standard error, from the time series of the potential energy and the binding energy from the production run. If -35 kcal/mol is the cutoff value of averaged binding energy for distinguishing the binder or nonbinder state, what is your prediction for the assigned system?

3. Analyze the related structures using maestro tools and compare the final result (ex*_mdrun.maegz) to that before the minimization, thermalization, and production run. Any changes from the receptor, the ligand, and the interface?

4. Some of the receptor structures are DFG-in and some of them are DFGout, check which state is your studied in. Could you explain briefly how different the DFGin and DFGout conformation interacts with inhibitors?

5. Please explicitly explain the pi-pi interaction, hydrogen bond and hydrophobic interaction for your ligand. Please also analyze the trajectories and answer the questions: Are those hydrogen bonds well maintained in all three processes? Explain why?

6. Hinge(Phe317, Met318) are the residues which binds to ATP’s adenine group for phosphorylation. Some inhibitors are designed to pre-occupy this position to prevent ATP binding. How strong is your inhibitor binds to the hinge region, is there any other group in your ligand also well interacting with the kinase? What are those groups and how well do they interact?

7. Activation loop, residue 381-409 in ABL1 kinase, behaves different with different types of inhibitors and it is one of the important feature to separate DFG-out from DFG-in. Could you explain briefly why you think that the activation loop behaves different in different conformation?

8. Please fill in the shared google document (you will get the url via email) the information of your inhibitors during the class and compare your results with your classmates. Could you explain why the ligand you studied is a binder (or non-binder) based on the information provided by you and your classmates? If you are interested, you might search the related articles for more information.