

CHEM 436 / CHEM 630: Molecular Modelling of Proteins TUTORIAL #2c: Comparison/Validation of models

INTRODUCTION

The purpose of tutorial #2c is to examine the overall quality of the homology model produced by MODELLER, in particular the protonation states of titratable residues, the placement of hydrogen atoms, and the choice of some of the side chain rotamers.

This independent validation will be done using the MolProbity server. Although the MolProbity toolset was developed to help protein crystallographers refine structural models of their X-ray diffraction data, it can be used to assess the quality of *any* protein model, including homology models.

The homology model will also be examined using the PROPKA server, which empirically predicts the pK_a of all titratable residues of a protein structure.

REQUIRED PRE-LAB READING

The MolProbity papers:

<http://dx.doi.org/10.1093/nar/gkm216>

<http://dx.doi.org/10.1107/S0907444909042073>

PRE-LAB REPORT

None

ADDITIONAL READING

The PROPKA3 paper:

<http://dx.doi.org/10.1021/ct100578z>

PROCEDURE

STEP 11: Examine the metal coordination

Check that the divalent metal atoms you have included in your model (Mg^{2+} , Ca^{2+} , Zn^{2+}) are coordinated by oxygen, nitrogen, and sulfur atoms only—not carbon—and are not clashing with superfluous hydrogen atoms. Take note of any problem, so that you can correct it in STEP 14.

STEP 12: Predict the protonation states of titratable residues

Upload the PDB file of the “all-hydrogen” model you obtained in STEP 9 to the PROPKA server (<http://propka.org>), and submit the calculation.

Note:

If the official PROPKA server is down, you may have to use the PROPKA formula implemented on the PDB2PQR server (http://nbc-222.ucsd.edu/pdb2pqr_2.0.0/). On the PDB2PQR server, make sure you change the following options before submitting:

- Force field: CHARMM (instead of the default)
- Output naming scheme: CHARMM (instead of the default)
- Options: Uncheck "Create an APBS input file"

Once the calculation has finished, the PROPKA results will be in an output file called "?????.propka".

- ◆ Compare the predicted pK_a 's for all titratable residues (Asp=D, Glu=E, His=H, Tyr=Y, Lys=K, and Arg=R) with their standard values. Report any acidic residue (D, E, H) that has a predicted pK_a above ~ 7.0 , and any basic residue (Y, K, R) that has a predicted pK_a below ~ 7.0 . Pay particular attention to buried residues.
- ◆ If you think some D/E residues should be protonated, specify on which oxygen atom the proton is likely to be (OD1 or OD2 for D, OE1 or OE2 for E). Explain your decision in terms of possible hydrogen bonding with surrounding residues.
- ◆ For H residues, compare the PROPKA results to your own observations from STEP 10.

Note:

Protonated D and E are unusual, so you might want to ask for a confirmation from the instructor before deciding on them. Also, because most molecular dynamics software recognize only D/E residues protonated at the "O2" atom (OD2 or OE2), you may have to swap the XYZ coordinates of the two (equivalent) oxygen atoms in the PDB file—so that the proton is at its right place, on the "right" oxygen atom.

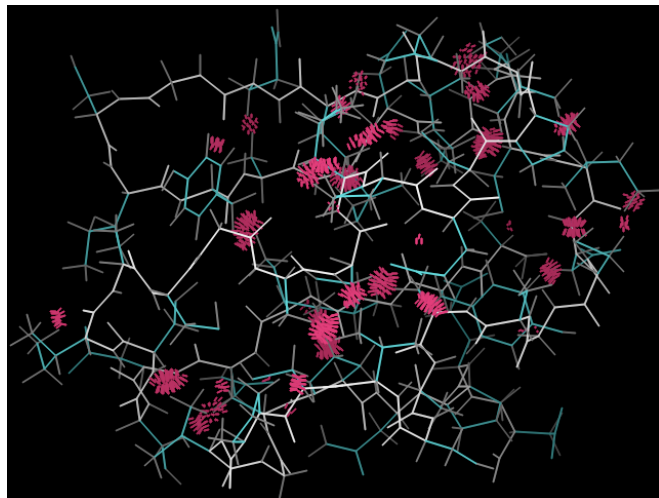
STEP 13: Check and refine the model with MolProbity

Upload the PDB file of the "all-hydrogen" model you obtained in STEP 9 to the MolProbity server (<http://molprobity.biochem.duke.edu>).

Note:

On the MolProbity server, do not use the "Back" button of your web browser. The website is designed such that you will always come back to where you started by choosing either "Continue" or "Cancel".

- ◆ Select the "Add hydrogens" tool of MolProbity and allow "Asn/Gln/His flips" (which is the default option). Once the calculation is done, report any flip suggested by MolProbity. For a detailed view of the results, click on the "Show all Asn/Gln/His" link. Write down all the "UNSURE" flips as well.
- Click on button at the bottom of the table ("Regenerate H, applying only selected flips") to regenerate the hydrogen positions with the flips. (You can download the new PDB file created by MolProbity at this point if you want.)
- Click on button "Continue".
- ◆ Back on the Main page, select the "Analyze all-atom contacts and geometry" tool of MolProbity. Perform the analysis on the new file (called "????_1FH_reg.pdb") using the pre-selected outputs. Report the "clashscore" and the MolProbity score of the model.
- ◆ View the atomic clashes in your model by clicking on the "Multi-criterion kinemage" icon (or on the "View in KiNG" link below). Check the "bad overlap" box and uncheck the "Rama outliers", "length dev", "angle dev", and "Cbeta dev" boxes. You should see something like this:

**Note:**

It is normal to see a larger number of clashes for a homology model than for a real protein structure. To identify which of these clashes are involving main-chain (“Mc”) versus side-chain (“Sc”) atoms, use the boxes “McMc contacts”, “ScSc contacts”, and “McSc contacts”.

- Close the “KiNG” window and click on button “Continue”.
- ◆ Back on the Main page again, perform the same analysis on the original file you uploaded (called “?????.pdb”). Compare your original model to that prepared by MolProbity. By how much have the clashscore and MolProbity score improved?
- ◆ From the section “Popular Downloads” of the Main Page, download the coordinates of the refined structure and compare the content of that file with that of the original “all-hydrogen” model from MODELLER. Comment on any new hydrogen atom or flipped residue. (Superimpose the two structures in PyMOL to better see the flipped residues.)

STEP 14: Prepare the PDB file for molecular dynamics simulations

Since MolProbity does not use the same convention for naming atoms and tends to permute the order of atoms within a residue, it is safer to prepare the PDB file for molecular dynamics simulations by making a copy of the “all-hydrogen” model you obtained in STEP 9 and by incorporating the modifications done by MolProbity by hand.

With the two PDB files open side-by-side in a text editor (the PDB file downloaded from the MolProbity server and the copy of the PDB file from STEP 9), make the following changes to the copy:

Asn/Gln/His flips:

- For each Asn/Gln/His residue flipped by MolProbity, copy the X,Y,Z coordinates of each atom involved in the flip:

Atoms involved for Asn: OD1, ND2, HD21, HD22

Atoms involved for Gln: OE1, NE2, HE21, HE22

Atoms involved for His: ND1, HD1 (if present), CE1, HE1, NE2, HE2 (if present), CD2, HD2

Histidines:

- Change the residue names of the histidines to either “HSD” or “HSE”, depending on whether the proton should be on the delta-nitrogen (“HSD”) or the epsilon-nitrogen (“HSE”).

- If the protonation state of a histidine has changed from the default “HSD” to “HSE”, delete the original HD1 atom and insert a new HE2 atom after atom NE2, located at the X,Y,Z position generated by MolProbity. Make sure the atom numbers are consecutive.

Aspartates/Glutamates:

- If any of these residues should be protonated, make sure the oxygen atom carrying the proton is called either “OD2” or “OE2”. (See note above for details.)

Visualize the structure of the edited PDB file in PyMOL, to make sure all the flips and protonation states are correct. Keep that file preciously. You will need it in Tutorials #4 and #5.

STEP 15: Examine the influence of the refinement on the pK_a 's of titratable residues

Run PROPKA on your new “all-hydrogen” model and compare the new pK_a profile to that of the original “all-hydrogen” model (from STEP 12).

- ◆ Discuss the changes in predicted pK_a for all residues that were modified by MolProbity.
- ◆ Prepare PyMOL images showing the coordination structure of each metal atom (if you have any). Render the metal atom as a “sphere” and show all protein ligands as “sticks”. Comment on any difference with the coordination structures observed in STEP 11.

Back up your files

Back up your files as explained in Tutorial #1c.