

Please use ray-tracing feature for all the images you are submitting. Use either the Ray button on the right side of the command window in PyMOL or variations of the commands below:

```
ray 300, 500      #creates a ray-traced image of width (300) and height (500)
set ray_shadow, 0 #removes shadow in the ray-traced image
```

## 1. Basics of protein structure (18 points)

Open the structure of murine Major Histocompatibility Complex Class I (MHC I) in complex with a peptide from Influenza A (IAV) (PDB ID: 3BUY) in PyMOL either by downloading PDB file and opening it or by simply using the command below:

```
fetch 3BUY #retrieves a protein structure from the PDB, saves it in your home directory, and
           loads it into PyMOL
```

- a. **The Peptide Bond:** Make a stick figure of the residues 208-210 (FYP) from Chain A (Chain A is at beginning of sequence) and indicate all peptide bonds. Explain how the planarity of the bonds is achieved and briefly list relevant data that support this. Measure all  $\omega$  torsion angles by using the dihedral measurement tool in the Wizard menu. Explain why proline takes cis-peptide conformation ( $\omega \approx 0^\circ$ ) in Xaa-Pro (Xaa: any amino acid) more often than the non-proline residues. **(8 points)**

To show the primary sequence of the protein, check Sequence On under the Display menu. Locate the residues 208-210 (FYP) in the primary sequence and click on them. Doing so will highlight the residues in the graphics window below. Enter the following commands in the command line.

```
hide everything #hides everything
```

```
bg white #makes the background white
```

```
show sticks, sele #shows the selected residues in stick representation
```

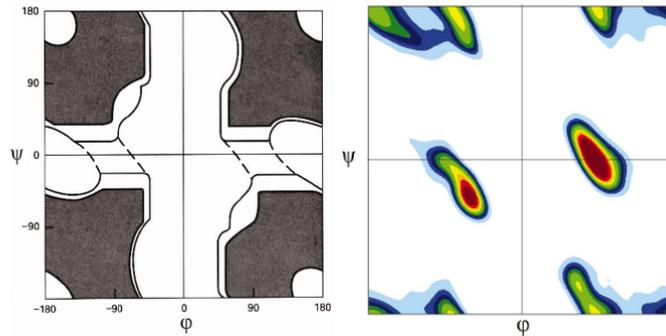
Or,

```
select FYP, (resi 208-210) #selects and names residues 208-10 as 'FYP,' which appears
                           as a new object in the menu to the right side of the
                           graphics window
```

```
show sticks, peptide #shows 'YFP' in stick representation
```

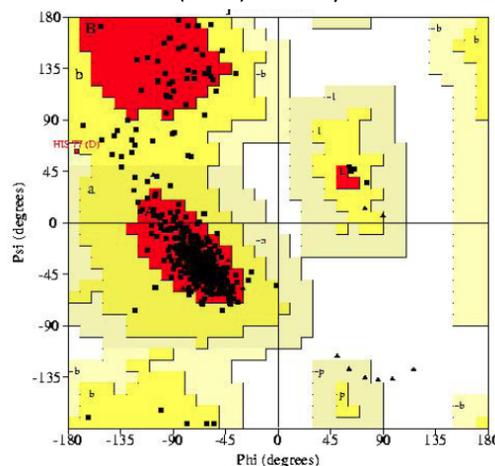
You may also find it helpful to show the polar hydrogens. You can do this by going to the side bar and selecting Action > hydrogens > add polar.

- b. Ramachandran Plot:** The classical version (1968) and refined version (2002) of the Ramachandran plot of glycine are given below (Figure 1). Why are the allowed regions determined by the  $\phi$  and  $\psi$  values much smaller in the refined version compared to the classical version? Nowadays, Ramachandran plot is used to evaluate the quality of three-dimensional protein structures. In Figure 2, you notice that several dots reside outside of the “allowed” regions on the plot. What is your guess on the identity of this residue? In what types of regions would you expect a residue of this nature to appear in the protein? Explain. **(6 points)**



**(Figure 1)**

(Hovmöller *et al.* (2002) *Acta Cryst D*58: 768-776)



**(Figure 2)**

- c. The Hydrophobic Effect:** A large part of proper protein folding depends on the hydrophobic effect. This effect, however, is not caused by the attractive forces between hydrophobic side chains. Describe the nature of the hydrophobic effect and the significance of entropy as the driving force. **(4 points)**

## 2. Secondary structure (38 points)

According to the Ramachandran plot, proteins have a limited variability of polypeptide backbone conformation (based on  $\phi$  and  $\psi$ ). The two main allowed regions in the plots correspond to the two main types of regular secondary structure ( $\alpha$  helices,  $\beta$  sheet, turn). Individual amino acid side chains also have strong preference for certain conformations, called rotamers (based on  $\chi_s$ ).

- a. Helix Formation:** Navigate to the first domain in chain A which contains a beta sheet framed by two alpha helices. Notice that the helices are not perfect. Now look at the sequence for these regions. What residues are causing these kinks? Why? **(5 Points)**

- b. **Helical Wheel Diagrams:** One section of a helix in our murine MHC I is:

163 - ECVEWLHRYLKN -174

Navigate to the Helical Wheel Projections website (<http://rzlab.ucr.edu/scripts/wheel/wheel.cgi?sequence=ABCDEFGHIJKLMNPO&subm>) and submit this sequence to generate a helical wheel diagram shaped and color coded according to the chemical properties of each residue. This particular helix is placed externally on the protein such that one side faces the hydrophobic beta sheet below it and the other is solvent exposed. Label the two sides in your helical wheel diagram. What pattern in the primary sequence hints that this helix may have a buried face and a solvent exposed face? **(5 points)**

- c. **Helices:** In your MHC I structure, find the helix from part b. Verify your findings by generating a **figure** displaying the hydrophobic patch of the helix as transparent orange spheres, and polar (both uncharged and charged) residues as transparent marine spheres. Use the following commands and variations thereof to accomplish this:

```
select helix, (resi 163-174)           #selects and names the residues as 'helix' and an object is
                                       created in the menu on the right side of the graphics
                                       window
```

```
select hydrophobes, (resn phe+tyr+ala+leu+val+ile) and helix
#selects and names the hydrophobic residues in the helix as 'hydrophobes,' in the menu
```

```
show spheres, hydrophobes           #shows residues in the object in spheres
```

```
color orange, hydrophobes           #shows residues in the object in orange
```

Use the same commands above for the polar residues (lys+thr+arg) in the helix, but name them as 'polar' and color them 'cyan'

```
set sphere_transparency=0.5         #sets sphere transparency at 50%
```

Once you have made your figure showing hydrophobic residues, hide everything and show the helix as sticks. Make **another figure** that shows hydrogen-bonding pattern holding the helix backbone together. What kind of helix is this? Explain your reasoning. (Note: Three kinds are covered in class.) **(10 points)**

- d.  **$\beta$  Sheets:** Make a **cartoon figure** by selecting and highlighting the two adjacent  $\beta$ -strands from Chain B (sequences: 36-EIQMLK-41/ 78-YACRVK-83) with a different color in your MHC I structure.

In **another figure**, isolate the two strands and show them in both cartoon and stick representations. Color the residues by element using the menu on the right side of graphics

window (C(Color)→by element). Show all hydrogen bonds stabilizing the two  $\beta$ -strands. What type of this  $\beta$ -sheet is this? (Note: Three kinds are covered in class.) Explain why this type of  $\beta$ -sheet is generally twisted. **(8 points)**

- e. **Secondary Structure Prediction:** Navigate to the three different secondary structure prediction servers listed below:

**PSIPRED** (<http://bioinf.cs.ucl.ac.uk/psipred/>)

**RaptorX Property** (<http://raptorx.uchicago.edu/StructurePropertyPred/predict/>)

**SSpro** (<http://scratch.proteomics.ics.uci.edu/>)

Go to the 3BUY PDB page and copy the FASTA Sequence under the dropdown menu 'Download Files'. Submit your sequence for Chain A to each of the servers. **Briefly describe** how well each prediction worked by comparing the predictions to the 3-D crystal that was obtained experimentally using X-ray crystallography. What is the principle behind each prediction that you think causes variation? **(10 points)**

### 3. Tertiary/quaternary structure (16 points)

- a. **Structural motif:** While the peptide binding subunit of MHC class I (Chain A in our structure) has considerable variability, the smaller subunit of the complex called  $\beta_2$  microglobulin (B2M). B2M is required for MHC I surface expression. What structural fold/motif is B2M an example of? **(2 points)**
- b. **Structural Alignment:** Copy the B2M structure from your original 3BUY structure into a new object. Rename it 3BUY\_B2M. Open one of the homologous structures of murine B2m, 5CKG, in PyMOL. This PDB ID corresponds to a human B2M structure with two stabilizing mutations. Align the structures using the following command:

i. align 3BUY\_B2M, 5CKG #aligns 3BUY\_B2M and 5CKG

\*note that there are two molecules in the 5CKG structure. This is related to how the crystal structure was solved. You can either delete the second chain or ignore it.

Generate a cartoon figure of the structural alignment between the two. Launch PDBeFold (<http://www.ebi.ac.uk/msd-srv/ssm/>) with the two PDB IDs. When selecting chains, use Chain B from 3BUY and Chain A from 5CKG. Report the root mean square deviation (RMSD) value. What does this value tell you? What is the benefit of aligning two structures using the secondary-structure matching (SSM) algorithm in PDBeFold? **(8 points)**

- c. **Sequence Alignment:** Navigate to the EMBL MUSCLE website (<http://www.ebi.ac.uk/Tools/msa/muscle/>)

Download the sequences for 5CKG and 3BUY from their PDB pages as in problem 2e and perform a sequence alignment of the Chain B sequence from 3BUY and the Chain A sequence from 5CKG using MUSCLE. Note that your sequences must be in FASTA format (Here's an

example if you've never used it before

[http://prodata.swmed.edu/mummals/info/fasta\\_format\\_file\\_example.htm](http://prodata.swmed.edu/mummals/info/fasta_format_file_example.htm).

Paste your alignment in your answer.

Based on your alignment, are your results for part 3b surprising? Where are most of the differences? **(6 Points)**

#### 4. Protein-Protein Interactions (28 points)

- a. **Electrostatics vs. hydrophobicity:** In class, you talked about a number of different weak forces that underlie interactions between macromolecular complexes. Briefly describe the difference between electrostatic vs. hydrophobic protein-protein interactions. **(3 Points)**
- b. **Surface Charges:** MHC I molecules are receptors that display peptides on the surface of cells to immune cells. This is a key way that the immune system can recognize intracellular pathogens. The binding groove for these peptides is located in between the helices packed against the  $\beta$ -sheet. The structure you have been working with has a peptide from influenza A bound in the binding pocket.

**Make a figure** showing the surface charges of the binding groove with the peptide bound. To do this, first you should extract the peptide into its own object. Select the residues for the peptide (Chain C). Go to the Action button on the right for selection and choose extract object. Feel free to rename the new peptide object.

To Generate the electrostatic surface map, go to the Action (A) button for the new protein structure without the peptide (3BUY). Select generate, then vacuum electrostatics. To change the surface transparency, go to the Setting drop down menu, select Transparency, and Surface. Set it to 40%. Show the peptide in stick form with sidechains and the main structure in cartoon beneath the surface map. What do the blue and red patches represent? **(8 Points)**

c. **Models and Programs:**

When you made the electrostatic map, a warning came up in the final dialog box:

```
Vacuum Electrostatics:
protein contact potential (local)
NOTE: Due to short cutoffs, truncations, and
lack of solvent "screening", these computed
potentials are only qualitatively useful.
Please view with skepticism!
```

Why do you think the program developers chose to include this warning? Why is it important to take into consideration the limitations of models and programs? **(5 points)**

- d. **Peptide Binding:** One MHC I protein is able to bind lots of different peptides 8-9 residues long with very different sequences. T cell receptors that bind to MHC I with peptides identify the

latter based on the way the peptide bulges out of the pocket. Look closely at the binding pocket. What are the biochemical properties of the residues that make up the peptide? Based on the biochemical character of the peptide and your electrostatic map of the binding groove, and what kind of protein-protein interaction is this? **(5 points)**

- e. **Salt Sensitivity:** Would you expect the interaction between this MHC I and peptide to be sensitive to small changes in salt concentration? Propose an experiment you could do to check salt sensitivity of the interaction. **(7 points)**