Project 4B: Investigation of Protein Structure

Part 1: Web-Based Protein Structure Programs

Part 2: Thermal Denaturation Probed by FTIR Spectroscopy

LABORATORY REPORT: Oral Presentation

PRE-LAB ASSIGNMENT

- Read the entire laboratory project described in the following pages, as well as any assigned additional reading. These include the two source papers: Leon et al (1998)[1], and Olchowitz et al (2002)[2], as well as textbook sections stipulated in the Introduction section.
- Prepare, on a typed sheet of paper, the Project Objectives for Parts A and B of this lab; on the same sheet, complete the assignment below:

1) Secondary vs. tertiary structure:
   a. α-helices and β-sheets are examples of _____structure, whereas helical bundles, β-barrels, and β-saddles are examples of _____structure.
   b. Considering these two levels of protein structure, it is easier to obtain reliable predictions (up to 80% correct) of _____structure, because this level of structure is based mostly on interactions between amino acids that are located _____.
   c. What "trick" does Swiss Model use to improve reliability of the less reliable type of prediction?

2) Regarding FTIR absorbance peaks,
   a. What types of molecular changes cause FTIR absorbance peaks?

[1] Leon, D. et al. (1998) "Structural Analysis and Modelling of Proteins on the Web" *J. Chem. Educ.* 75, 731-734.
[2] Olchowitz, J.C. et al (2002) "Using IR Spectroscopy to Investigate Protein Stucutre" *J. Chem. Educ.* 79, 369-371.

b. Why do α-helix, β-sheet, and denatured/aggregated protein have different FTIR peak locations?

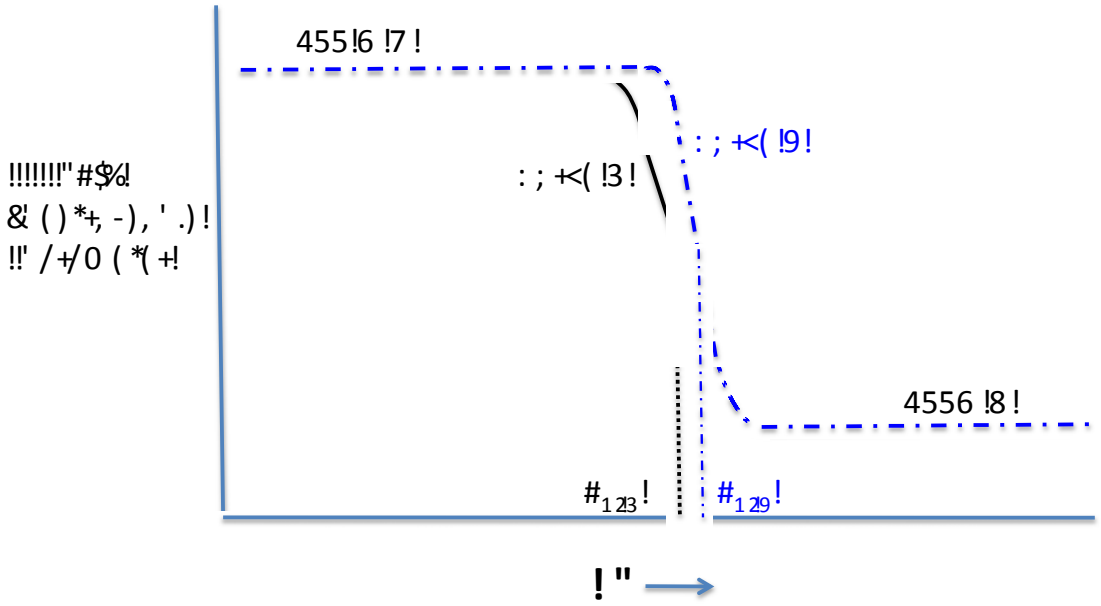
3) Using results from Olchowitz *et al. (*2*)*, summarize the structural and spectroscopic differences between
   a. native cytochrome c vs. trypsinogen, and
   b. native vs. denatured myoglobin

4) Thermodynamically, protein denaturation, or unfolding, resembles melting.
   a. What signs (+ or -) do you expect for the enthalpy and entropy of denaturation: $\Delta H°_U$ should be _____; $\Delta S°_U$ should be _____.
   b. In the table below, for the three temperature range columns,
      i. specify the relative equilibrium concentrations of the native (N) and denatured/unfolded (U) forms of the protein: Is $[N]_{eq}$ less than, equal to, or greater than $[U]_{eq}$?
      ii. and specify the sign of $\Delta G°_U$: is it +'ve, 0, or –'ve?

	At T < T_U	**At T = T_U**	**At T > T_U**
i. $[N]_{eq}$ vs. $[U]_{eq}$	$[N]_{eq}$ $[U]_{eq}$	$[N]_{eq}$ $[U]_{eq}$	$[N]_{eq}$ $[U]_{eq}$
ii. sign of $\Delta G°_U$			

   c. Calculate T_U with respect to $\Delta H°_U$ and $\Delta S°_U$.

5) Denaturation is completely irreversible only at $T >> T_U$. Therefore, at $T \leq T_U$, when the protein solutions are removed from the heat block, denatured proteins *might* refold into the native conformation prior to spectroscopic determination. Below are plotted two curves that show the change in an FTIR spectroscopic parameter (e.g., absorbance at a particular wavenumber, or location of a particular peak) with temperature.
   a. Curve ___ represents the *true* value of the N/U equilibrium, whereas curve ___ represents the *observed* N/U equilibrium after partial refolding.
   b. From this we can conclude that, due to refolding, the *observed* value for T_U could be _____er than the *true* value of T_U.


455!6 !7 !
: ; +<(!9 !
: ; +<(!3 !
!!!!!!!"#$%!
&! ( ) *+, -), ' .) !
!!' /+/0 (*(+!
4556 !8 !
#123 !
#129 !
! "

INTRODUCTION

Proteins can only function if they are folded into the correct native conformation. To understand or manipulate protein function (e.g., to create a pharmaceutically effective product), one must be familiar with protein structure. Although the most powerful computer programs that characterize the energetics and dynamics of protein structure are expensive commercial products, a number of simpler programs are available as freeware on the web. In Part 1 of this lab, which is based on the Leon *et al.* (1998) source paper (ref. 1), you will be introduced to a few useful protein structure programs. Be sure to read the Leon *et al*. source paper before the pre-lab session.

For many proteins, three-dimensional structures[3] have been solved and their atomic coordinate files deposited in the Protein Data Bank (.pdb). These files include not only the final three-dimensional fold (i.e., the tertiary structure), but also the amino acid sequence (primary structure), and secondary structures like helices, sheets, and loops that are formed from *contiguous stretches* of the amino acid sequence. Given the amino acid sequence of any novel protein, a prediction can be made regarding its structure, as long as a homologue of the known structure is available as a "**template**." In deriving a structure prediction, programs like Swiss Model start with the template, and then add in any amino acid differences between the novel homologue and the known protein template. Algorithms are employed to optimize the putative structure by maximizing attractive interactions (e.g. H-bonds, hydrophobic force, salt bridges) and minimizing repulsive interactions (e.g., steric hindrance, electrostatic repulsion).

In addition to solving protein structure via crystallography (especially x-ray diffraction), spectroscopy can also be used to characterize protein structure, especially NMR, UV-Vis, and FTIR. Infrared spectroscopy monitors the molecular vibrations of molecules. If a vibration results in a change in dipole moment, then this vibration is infrared-active. In Part 2 of this project (based on Olchowitz *et al.* (2002), ref. 2), we will see that the two dominant bands in aqueous protein IR absorbance spectra are the amide I and II peaks, around 1650 and 1550 cm^{-1}, respectively. Both arise from vibrational motions in the polypeptide backbone amide groups: The amide I band is due mainly to C=O stretching, whereas the amide II band arises from N-H

[3] For background information on protein structure, please consult the chapter in your biochemistry text entitled "The Three-Dimensional Structure of Proteins."

bending and C-N stretching. The amide I peak is especially useful in characterizing α-helix, β-sheet, and disordered regions.

Qualitative IR scans are often done measuring transmittance (0-100% T) on the *y*-axis, but more quantitative measurements are typically done in IR absorbance mode. IR absorbance obeys Beer's Law: $A_\lambda = \varepsilon_\lambda c\, \ell$, where A_λ represents the absorbance at wavelength λ, ε_λ is the molar absorptivity at that wavelength; ℓ is the path length of the cell, and *c* is the concentration of the protein. In this experiment, the concentration of protein solutions and the path length between the calcium fluoride disks are both fixed.

One problem that biochemists encounter is the strong infrared absorbance of water in the amide I region. We will minimize this problem by lowering the path length (the two CaF_2 disks are separated only by a thin Teflon spacer) and by using D_2O as solvent, thus minimizing the concentration of water. The O-D vibration is down-shifted 400 cm^{-1} from water's O-H vibration. Additionally, in D_2O the protein's amide II vibration downshifts about 100 cm^{-1} (to 1450 cm^{-1}), while the amide I frequency remains in roughly the same region.[2] The use of D_2O solvent allows one to obtain reliable absorbance spectra even if the path length ℓ is larger than expected. D_2O also provides good amide I peak separation between α-helices (1650-1660 cm^{-1}), β-sheets (1620-1640 cm^{-1}), and denatured/aggregated protein (1610-1625 cm^{-1}).[2] The difference in the amide I peak locations for different protein secondary structural motifs derives from the different C=O ····· H-N hydrogen-bonding patterns within these structures. Please review FTIR spectroscopy as described in Olchowitz *et al.*[2] and also in your Organic Chemistry text.

In Part 2 of this lab, the proteins we will be studying have been selected because they contain an excess of either helix or sheet. Part 2.1 is a qualitative study that allows you to determine whether your native protein has a larger percentage of helix or sheet. (In order to quantitatively determine the amount of each secondary structure present in a given protein, more sophisticated instrumental methods are needed, e.g., circular dichroism). In part 2.2, you will assay the thermal unfolding of helices and determine the denaturation temperature (T_U) of myoglobin or lysozyme. T_U is defined as the temperature at which the equilibrium between the native (N) and the denatured/unfolded (U) forms of the protein features a 50/50 mixture; at $T = T_U$,

$K_{den} = [U]_{eq}/[N]_{eq} = 1$, and $\Delta G°_U = 0$. We can follow changes in the denaturation equilibrium by identifying spectroscopic differences between the N and U forms, e.g., IR absorbance at a specific wavenumber, or the location of a specific IR peak ($\tilde{\nu}_{max}$). Following such IR changes upon thermal denaturation yields a plot such as that found in pre-lab question 5, allowing us to determine the value of T_U (see Appendix 4 for data analysis suggestions).

Proj. 4B, Part 1: Investigation of Protein Structure Using Web-based Programs

Please note that students work in **pairs** on Experimental parts 1-5 below,
and **alone** on part 6 ("Compare Structures of Homologous and Known Proteins").

EXPERIMENTAL (1): OBTAIN 3D COORDINATES FOR YOUR KNOWN PROTEIN

Go to the Protein Data Bank (pdb), a repository of 3D coordinates for solved protein structures:

http://www.rcsb.org/pdb/

There are several methods for finding protein structures using the PDB. Below is a suggested method for locating structures that will be useful for this lab. Feel free to explore the PDB site for interesting biomolecules in addition to the protein that you will analyze in this lab.

At the top right of the screen, click on "Advanced Search". Under "Choose a Query Type" select "Macromolecule name" or "Text Search" and type in the name of your assigned **known** protein (myoglobin or lysozyme). Below the search box, click on the link for "Add Search Criteria" and select "Number of Chains (Asymmetric Unit)" under the pull down menu. Enter "Between: **1** and **1**" to limit the search to structures with only one polypeptide chain. Feel free to use other search options, if necessary, to narrow your search. Click "Submit Query."

We will use "DsRed" (a red fluorescent protein) as an example. The pdb site will return a list of related proteins whose structures have been solved. They may differ in terms of species source, mutations inserted, number/types of subunits, ligands bound, etc. Scroll down to "Query Refinements"; under "release date," select "before 2000." Scroll down and find a protein that is as close to the active wild-type form as possible: mutations should be minimal (at most one residue), ligands should be substrates or substrate analogs, not inhibitors. (For myoglobin, avoid the "met", or ferric form.) Click on the pdb filename (e.g., 2VAD) or on the protein name (e.g., monomeric red fluorescent protein, DsRed.M1). If there is only one result, you will automatically be taken to the page for your chosen protein.

The pdb page that comes up has three main sections: (a) At the top are page tab options (e.g., "Summary", "Sequence", "Seq. Similarity", "Literature", "Biol. & Chem.", etc.); (b) In the

upper right hand corner are listed "Display Files", "Download Files", etc; (c) At the right side is a 3D image of the protein with different display options.

On the "Summary" page you can find, under "Molecular Description", more information including alternative names of your protein under the "Molecule" field (e.g. Red Fluorescent Protein) and number of polypeptide "Chains" (should be A only). Just below the "Organism" source should be listed a "Uniprot KB" number; if there is none, then you must select a different protein. Under "Ligand Chemical Component," the number and type of bound cofactors (metal ions, organic groups, etc.) are listed. More detailed information on many of these components is listed on the "Biology & Chemistry" page. Based on your search parameters, your chosen protein should have only one chain (i.e., just Polymer 1, or chains: A); if it has more chains, then select another protein from your search results. Using a protein with just a single polypeptide chain will facilitate later steps in this lab. Select the "Sequence" tab to find the amino acid sequence of your protein, and the secondary structures (α-helices, β-sheets, Ω-loops, etc.) for contiguous sections of the sequence. Examine the 1° and 2° structures, and print this page.

Create a folder in your H: drive to hold all of your protein structure files. Click on the "Download Files" menu at the upper right corner of the screen and select the "PDB File (Text)" link to download the 3D coordinates of the atoms in the protein. Save the pdb file (e.g., 2VAD.pdb) to the protein structure folder on your H: drive. Finally, under "Download Files" select "FASTA sequence" to save the AA sequence of your known protein (e.g., 2VAD_A.fasta.txt).

EXPERIMENTAL (2): CHECK THAT YOUR KNOWN PROTEIN IS IN THE UNIPROT DATABASE

Now, under the "Links" tab at the far right (above "display files"), select "PDBsum" about halfway down the "Structure Summary" menu. This opens a new window for the UniProt data base (EMBL-EBI.ac.uk) entry for your known protein. In the box in the middle, just below "Date" and "Release Date", is a "UniProt" link to your "Protein Chain A". Click on the colored UniProt file number for your known protein (e.g. Q9U6Y8), if it's there. (If it's not there, then your known protein is not in the UniProt database and you'll have to choose another one).

A new window will open with the UniProt page describing your protein; record the UniProt file number for your protein (e.g., Q9U6Y8), and print out the important pages. For example, you need NOT print references, cross-references, databases, etc. Scroll toward the bottom of the page and you'll see a color-coded depiction of the protein's secondary structure, including where the different domains are, where the cofactors are bound, and also the locations of individual helices (blue), beta strands (green), loops/turns (pink), etc. (Click on "details" to see residue numbers.) Finally, at the bottom find the AA sequence: Copy the one-letter AA sequence of your protein (or click on "FASTA"), and save this AA sequence in the same .txt document that has the PDB AA sequence. Check to make sure that the AA sequence given by UniProt is the same (or close) to that given by PDB. If not, holler.

EXPERIMENTAL (3): FIND NOVEL PROTEINS HOMOLOGOUS TO YOUR KNOWN PROTEIN

At the top right of the UniProt FASTA sequence is the "BLAST" tool. A minute or so after clicking "Go," you should get a list of proteins that are homologous to your known protein. The list is ordered with the most closely-related (green bar) homologues at the top. The table below the list includes, for each protein: a score for AA identities (%), a homology "score" from 0-800, and an E-value describing the quality of alignment of the amino acid sequence with your protein.

At the upper right, use either the "next page" button, or specify a later page number to find a protein that is between 25% and 45% identical to your known protein. If your lowest ID score exceeds 45%, select "more hits: 1000" at the top center and select a page number to view lower-score homologues. Print out the portion of this list of homologues that contains your selected protein. Click on the UniProt file number for this protein (e.g., A7UAL1).

Check if the 3D structure of your chosen novel homologue has been solved yet by scrolling down its UniProt page to "Cross-Ref./3D Structure Databases". If there is an entry there for PDB, with a pdb number listed, then your novel homologue's structure has already been solved; choose another novel homologue. If there is no listing here for PDB, then you're clear to proceed. You now have your **known** protein and a structurally uncharacterized (or **novel)** homologue; print out the UniProt page for your novel homologue. For later use, save a copy of the

AA sequence of your **novel** homologue (under sub-heading "Sequences") to a WORD or .txt document.

Go back to the full BLAST list of homologues: At the top of the page, under the left-hand column entitled "Alignments" (the one containing the color-coded homology bars), click on the colored bar that corresponds to your chosen homologue. This will display the sequence alignment between your known protein and its novel homologue. Copy and paste this alignment into a Word document (you can use a fixed-width font like `Courier`, or a screen-capture {Mac command-shift-4}); print it out. The top AA sequence is your known protein, and the bottom one is the selected novel homologue. Identical amino acids and conserved mutation/similarities (+) are shown between the two aligned sequences. Record both the % sequence identity and the % "positives" (i.e., the % amino acids that are either identical or similar). Make sure that the "positives" score is less than 60% for your novel homologue.

EXPERIMENTAL (4): COMPARE SECONDARY STRUCTURES OF YOUR TWO PROTEINS

Check whether your novel homologue is likely to have a secondary structure similar to that of your known protein. Submit the AA sequence of your novel homologue for a prediction of its likely secondary structure. One of the best places to do this is:

http://www.compbio.dundee.ac.uk/www-jpred/

Click on the "Advanced" link on the right hand side of the sequence box. If there is a sequence already in the box entitled "Paste your sequence data here:", delete that sequence before pasting your own. Put a check mark in the box for "Skip searching PDB before prediction" and enter your email address in the specified box. Click the "Make Prediction" button at the bottom of the page to submit your query.

The results should appear within a few minutes (the website claims that ~98% of submissions complete within an hour, and of those, >90% complete within seven minutes). Long sequences will take longer to process, so you may have to choose a shorter protein or break it up into separate domains if the prediction takes too long. The results from the website under "View Simple (HTML)" are clearest, so include these in your report. Print out a copy of the secondary structure prediction using either screenshots (command/shift-4 on a Mac) or a fixed-width font

like `Courier` in WORD. Use the sequence alignment that you got from BLAST to compare in detail the secondary structures predicted for your novel homologue to those that occur in your known protein. Carefully line up each α-helix (H) and β-strand (E, extended) in the two proteins; compare lengths and locations in the sequence; identify extra or missing helices or strands.

EXPERIMENTAL (5): OBTAIN A 3D STRUCTURAL MODEL FOR YOUR NOVEL HOMOLOGUE

You will now request a prediction of the 3D structure of your novel homologue from Swiss Model at:

http://swissmodel.expasy.org/

On the left hand side of the web page, under the "Modeling" menu, click on the "Automated Mode" link. Next type your email, a project title, and the UniProt file number for your novel homologue (e.g., A7UAL1) in the box provided (or paste the AA sequence). At the bottom of the screen, after "Use a specific template," type the PDB-ID of your known protein (e.g., 2VAD) and chain "A"; then click "Submit Modeling Request".

Once the modeling is complete (30-60 minutes), the Swiss Model page should refresh to a results page. If you submitted an email address with the modeling request, a link to the results page will be emailed to you upon completion of the modeling job. On the results page, under "Model Summary," click "download model as [pdb]" to save a .pdb file containing the predicted 3D coordinates for your modeled novel homologue. Save your Swiss Model results page (Print/Save this page in upper right) and print out a copy.

EXPERIMENTAL (6): COMPARE STRUCTURES OF MODELED HOMOLOGUE & KNOWN PROTEIN

Working by yourself, you will now compare your known protein structure and the predicted tertiary structure of the novel homologue; you will use PyMOL, a free biological structure viewing program. In the "Biology and Chemistry" folder on the desktop, double-click the "PyMOL" icon: a window will open. The top of the window contains the graphical user interface (GUI) with various menus and a command line. The bottom of the window is the viewer, which contains the molecule and buttons for other useful commands. To make the proteins easier to print, turn the background white: click on the "Display" menu (menu bar at the very top of the monitor screen, in the middle), select "Background" → white.

Upload the known and novel homologue protein structures:

Select "File" → "Open," then open the pdb file for your **known** protein. The file name (or ID) for your protein should appear in the panel on the right of the viewer. At the very bottom right corner of the screen, there is a button labeled "S" (Sequence) (not the "S" next to your protein's ID, but at the bottom right of the screen). Click on that button to display the sequence of your protein. You can select residues either by left-clicking on the residue now listed at the top of the window, or on the structure itself. Selecting residues will automatically create a new object called "sele" (selection), located at the upper right of the viewer screen, just below your protein ID. Selected objects can be renamed and modified (color, display, etc).

Your protein structure can be manipulated with the mouse in the following manner:

rotate:	**left**-click-drag
zoom:	**right**-click-drag
x-y position:	**middle**-click-drag (click-drag with the mouse wheel)
"slab":	roll the **mouse wheel** (cuts away slices of protein, starting from the front)

If you would like to reset the center of the structure for easier rotation, click on an atom with the middle button (mouse wheel). Other options are listed in the lower right-hand corner of the display window; however, the commands are difficult to decipher.

Next to your protein's ID, locate the "A" (action), "S" (show), "H" (hide), "L" (Label), and "C" (color) buttons. Click on the "H" (hide) button and select "everything" to hide all of the atoms and lines. Next click the "S" (show) button and select "cartoon".

In the same session, open up the pdb file for your novel **homologue** ("File" → "Open"). Hide everything and then display the cartoon view as described above. Your original protein structure might no longer be visible on the screen; never fear, it's simply been bumped lower in the viewer window. {Optional: By zooming out you can adjust the window so that you can see both structures, then drag the structures so they're next to each other. Use the mouse mode command MovO to drag one structure at a time; you may have to click in the "Mouse Mode" box to see the MovO command.}

Align and overlap the two proteins by clicking next to your known protein ID on A(ction)/align/to molecule/the ID of your novel homologue. When PyMOL is done aligning the two structures, a list of statistics will appear above the command line. Copy the following information and include it in your report: the number of aligned atoms, the number of atoms rejected during each cycle, and the Root Mean Square (RMS) alignment factor. Lower RMS values indicate a better alignment; your value should be below 1.

Compare the two structures: disulfides and ligands

For both of your structures, show the disulfide bonds by clicking on "S" (show) → "disulfides" → "sticks" or "spheres". If your proteins do not contain any disulfide bonds, your structures might not change. To make your disulfides easier to see, you can select the participating Cys residues and recolor them: Click on a cysteine of interest; the selected residue is identified in the top (user interface) window in the most recent (lowest) command line: e.g., "you clicked/protein ID/CYS'108." If you then click on C108 in the amino acid sequence line, all of its atoms are marked with red dots. You can recolor the residue by clicking, on the (sele) ID line, "C" (color) → new color. Make note of any differences in disulfide bonding patterns between the known and novel proteins.

If your known protein has a ligand, a non-amino acid cofactor, or a modified amino acid, it will not be displayed in the cartoon mode. To fix this, look in the sequence at the top of the screen for a residue that is separated from surrounding residues by a space. Click on that isolated ligand "residue". If it is not displayed, do so by clicking, on the (sele) ID line, "S" (show) → "sticks" or "spheres". Ligands and cofactors are sometimes listed after the end of the protein sequence; display them as well.[4] Water molecules (listed as "O") and metal ions can sometimes be important, depending on their location and function. You may click on them at the end of the

[4] Please note that ligands and cofactors will only be present in the pdb file of the known protein, as ascertained from the X-ray or NMR structure. Your unknown homologue will lack any ligands or cofactors, because you only sent the AA sequence to Swiss Model. That does NOT mean that your homologue necessarily lacks ligands or cofactors, it just means that Swiss Model could not put them in because it lacked sufficient information to do so.

AA sequence and determine their location. However, your novel homologue will lack these waters, metals, cofactors, etc., so they cannot be part of your comparison.

Use the mouse controls to manipulate your aligned proteins. Your proteins should have contrasting colors so that it is easy to tell them apart (click on "C" (color) button next to the protein name to change the color; don't be afraid to make your proteins look good!). If you want to hide one protein at any point, click on its ID to deselect it; clicking again shows it again. Once you find a view that highlights an interesting aspect of your alignment (ligand, cofactor, disulfide, etc.), capture an image of it in the following manner: Type *ray* into the PyMOL> command line and press return; this will do a ray trace that will make the structures look better. If you click on the structure at this point, you will lose the ray trace. To save an image of the protein, click "File" → "Save Image" → "PNG…". Save to your H: drive. If you wish to be able to return later to this view in PyMOL, select "File" → "Save Session."

Compare the two structures: surface views

NOTE: For the following three surface view comparisons, you will be using the "3-Button Editing" and "3-Button Viewing" mouse controls in the box at the bottom right of the PyMOL page. Below we introduce to you just a few of these controls; for a more detailed explanations of all of the various features afforded by "3-Button Editing and Viewing," please see Appendix 1.

Without changing the alignment of your structures, display the calculated surfaces of each protein by clicking "S" (show) → "surface". Now separate the two proteins and place them side-by-side, while maintaining their alignment, as follows: First deselect one of the proteins by clicking on its ID. Then click in the command box in the lower right of the viewer screen (green "Mouse Mode"). The Mouse Mode will now say "3-Button Editing". The "buttons & keys" column shows that if you hold down the "shift" key on the keyboard, the mouse buttons allow you to: L(eft)/RotO, M(iddle)/MovO, R(ight)/MvOZ, and Wheel/MovS. Hold down the shift key and use the mouse M(iddle)click button to drag (MovO) the selected protein to the side. Then reselect the second protein, and they should both be side-by-side, but with the identical alignment. Note the alignment (or lack thereof) of the side chains between your two proteins. Identify key regions that show similarities and differences in surface topology (e.g., bulges, crevices). If you click on

regions of interest, the selected residue is identified in the top (user interface) window in the most recent (lowest) command line: e.g., "you clicked/protein ID/VAL'109." If you then click on the identified residue in the amino acid sequence line, all of its atoms are marked with red dots. In this way you can determine which amino acids cause the differences. Save pictures and/or sessions as above, and repeat the process for multiple views of your alignment. If at any point you wish to remove the "surface" view, click "H" (hide)/everything, then "S" (show) the next view.

Next, examine the **electrostatic surface** of the two proteins. For each protein, select A(ction)/generate/vacuum electrostatics/protein contact potential. This will show a calculation of the surface charge density of each protein, with red being negative, blue positive, and white neutral. Use the mouse to rotate the pair of proteins so you can get several views that show differences and similarities in the locations of patches of negative, positive, and neutral surface areas in the two proteins. To determine which amino acids cause the differences, mark residues by clicking on the amino acid sequence line. For example, click on all H, K, and R residues to mark positive patches (blue); or, click on all D, and E residues to mark negative patches (red). When clicking on an individual residue, you may have to rotate the proteins in order to see its actual position, because in the "surface" view, residues on the back surface project through and look like they're in the front. For example, if you click on a particular aspartate residue and it doesn't show up in a red patch, rotate the proteins until the dots align with a red patch. That will be the actual surface location of that particular aspartate.

Finally, use the script "color_by_restype" to color all of the **major residue types**. In addition to coloring positive side chains blue and negatives red, as you have already done, this script will color hydrophobic side chains white, aromatics gold, polar-neutral side chains pink, prolines green, and cysteines cyan. First, click on the "ID_e_chg" to deselect the electrostatic surface structure, then click on the protein ID to select its initial structure. In your web browser, go to the website http://pldserver1.biochem.queensu.ca/~rlc/work/pymol/ , scroll down to "color by residue type" and click on the script program "color_by_restype" on the left to download the script to your desktop. The following directions are for PyMOL on PCs; if you're using a Mac,

see the footnote below.[5] In the PyMOL menu at the top, select File/Run, and click the script "color_by_restype.py" on the desktop and click "open." The script is now loaded and running. Next, in the command line type: color_by_restype <protein-ID>, hydrophobic=white, aromatic=gold, polar=pink, cysteine=cyan, proline=green, where <protein ID> is the pdb filename for the protein loaded into PyMOL (e.g., 2VAD or Model_1). When you press 'return', the protein will be recolored according to the residue type color scale specified. Do the same recoloring for both your known protein and novel homologue, and note similarities and differences. Again, you can attribute differences to specific residues by clicking on residues in the AA sequence line at the top, and seeing where the residue is by noting the appearance of the red dots. As above, you may have to rotate the aligned proteins in order to find the actual surface location of the residues.

[5] On a Mac, drag the desktop icon for the "color_by_restype.py" script into the PyMOL> command line, click in the command line to unselect the location description, and insert at the beginning of the command line the command "run" followed by a space, followed by the script location; the command line should now read: Run /Users/loginID/Desktop/color_by_restype.py. Press "return".

RESULTS AND ANALYSIS

Use Word, Power Point, or another program to fit your images on a few pages, with the known and novel proteins adjacent. Label the images: Identify the two proteins, use a key to explain colors and shapes, and include an informative caption pointing out the interesting features of the structures that are visible. The pictures should be of sufficient size to see the pertinent features clearly. You should have at least three pairs of pictures, with each pair comparing different structural aspects (e.g., surface topology, surface charge, polarity/hydrophobicity, key salt bridges or H-bonds, disulfide bridges, etc.). If you wish, you may print them out on the color printer in Olin 303.

POST-LAB: QUESTIONS

1. *Identify a number of key structural differences between your known protein and your novel homologue, and speculate in some detail as to the sources[6] of these differences. Can you identify a substrate or ligand binding cavity? What does this say about steric hindrance?*
2. *Highlight a few specific amino acid differences between your known protein and your novel homologue and discuss their influence on secondary structure[6], hydrophobicity, and/or surface charge density.*
3. *Why were you instructed to select a homologue with 25% < identity < 50%?*

POST-LAB: INFORMAL REPORT (NONE)

POST-LAB: FORMAL REPORT (ORAL REPORT)

You and your partner will have fifteen minutes to give a PowerPoint presentation. Roughly half of this time will be devoted to Part 1 of this lab, your web-based analysis: Describe the background, function, and structure of your known protein, along with a detailed comparison to the structure of your novel homologue. The introductory background/function section of your presentation should take no more than 2-3 minutes; spend the bulk of your time on your comparative structures. Include three or more types of informative comparisons (hint: a picture is worth 1000 words). Clearly explain the important similarities and differences in the structures.

[6] Chou-Fasman Rules state that the strongest (a) α-helix formers are glu, lys, gln; met, ala, leu, phe; (b) β-sheet formers are val, ile, tyr, trp, phe, cys, leu; (c) β-reverse turn/Ω-loop formers are gly, asn, pro, ser, asp, tyr; and (d) random coil formers are arg, his, thr, gln, and lys.

Proj. 4B, Part 2: Investigation of Protein Thermal Denaturation Using FTIR[7]

Morning of your FTIR day: For the denaturation part of the protein-FTIR lab, each team will incubate at several temperatures, as specified below. To attain your temperatures, we have access to a number of dry heat blocks, most of which are digitally controlled. Use the old analog block for a mid-range temperature; **allow it <u>an hour</u> (!!) or so to reach its final temperature**. You will probably have to adjust the heat setting to get within 1 °C of your target T. Use one of the digital heat blocks for the lowest temperature, then when that sample is done incubating during lab, raise the setting, allow time for temperature equilibration, then incubate a second sample at the higher temperature.

The experimental portion of this lab has two parts: (2.1) You will take FTIR spectra of two native proteins at room temperature (one protein is mainly α-helix, the other is mainly β-sheet); (2.2) you will study the thermal denaturation of either myoglobin or lysozyme at temperatures ranging from 25-90 °C. Teams A-D will heat myoglobin at 40, 66, 70, 75, 80, and 90 °C; teams E-H will heat lysozyme at 40, 66, 72, 80, 85, 90 °C. At room temperature, teams A-D will study myoglobin plus one of the following: β-lactoglobulin, RNAase, trypsin or chymotrypsin; teams E-H will study lysozyme plus one of the following: β-lactoglobulin, RNAase, trypsin or chymotrypsin.

[7] Based on Olchowitz *et al.* (2002), ref. 2.

PART 2.1: NATIVE PROTEIN SECONDARY STRUCTURE AT ROOM T

EXPERIMENTAL: MATERIALS

- Each team will use two proteins from this list (see above for your team's proteins):
  myoglobin (Mb), lysozyme (Ly), cytochrome c, peroxidase, β-lactoglobulin, ribonuclease A, or trypsin. (These are all predominantly either helix or sheet.)
- Solutions: concentration= 60 mg protein per mL of D_2O; final volumes either 50 or 400 μL.
- Eppendorf tubes: 0.6 mL *safe-lock* (use only tubes with special Safe-Lock caps)
- Deuterium Oxide (99.9 atom % D): pour out about 0.5 mL for your own use
- FTIR: 0.015 mm Spacers, two CaF_2 discs, 25 × 4 mm
- Spectra Tech press lock sample holder

NOTE: Keep all protein stock solutions on ice until they are either heated or used to obtain the IR spectra.

EXPERIMENTAL: METHODS

An infrared spectrometer will be used to obtain the IR spectra of protein samples. A protein solution made in deuterium oxide is applied to a calcium fluoride cell (CaF_2 is relatively water-insoluble and transparent to infrared radiation). A continuous IR spectrum is observed, expressing infrared absorbance as a function of wavenumber ($\tilde{\nu}$, in cm^{-1}).

Preparation of Protein solutions

You will take room temperature FTIR spectra of two proteins, one of which you will also study at several elevated temperatures. The concentration of your stock solution of each protein will be 60 mg/mL, dissolved in D_2O (99.9 atom % D). For the protein studied at room T only, you will need approximately 50 μL of stock solution; for the other protein, myoglobin or lysozyme, you will need 400 μL. Calculate the masses that you will need of the two proteins, and check with your instructor before making these solutions. Using a small square of weigh paper, weigh out the two proteins, and transfer to a 0.6 mL *safe-lock* Eppendorf tube (use only tubes with special Safe-Lock caps); dissolve/vortex in the appropriate volume of deuterium oxide. Native protein samples must be **KEPT ON ICE AT ALL TIMES**.

Preparation of FTIR Cell

Make sure that the Transmission chamber is closed on top, and locked down. At least 6 minutes (but no more than 10 minutes) before you are ready to collect your first spectrum, turn on the adjacent $N_{2(g)}$ tank to 12 psi. This purges the chamber to remove most of the atmospheric H_2O and CO_2, both of which absorb in the region of interest. The IR cell comprises a metal holder with two calcium fluoride disks, separated by a thin Teflon ring spacer; use gloves when handling the CaF_2 disks. First take a background spectrum of D_2O: Prepare a CaF_2 disk with spacer, then add a sample of 16 µL of deuterium oxide (99.9 atom % D). The other disk is placed on top. The windows are then placed inside the metal disk-holder and compressed ***very lightly***. (Do not press too hard or the disks may crack!!) Make sure that the liquid covers the disk surface and that there are no air bubbles in the center of the disk where the incident IR beam is focused.

The cell assembly must be quickly placed in the FTIR instrument through the top door; afterward, purge the chamber for approximately 6 minutes before a spectrum is taken, allowing gaseous nitrogen flow to purge water vapor and CO_2 out of the spectrometer. Protein solution spectra are taken by wiping off the previous sample, rinsing 2-3 times with D_2O, then adding the next 16 µL sample of protein solution to the prepared window. Again, check for the absence of air bubbles at the center of the disk.

Detailed instructions for taking IR spectra are available in the drawer underneath the FTIR instrument, on the desktop of the FTIR computer, and also in Appendix 2 at the end of this lab.

PART 2.2: MONITORING PROTEIN THERMAL DENATURATION

In order to estimate the denaturation temperature of a protein, samples incubated at different temperatures must be analyzed. Teams A-D will study myoglobin; teams E-H, lysozyme. Each team will incubate at six assigned elevated temperatures, using the dry heating blocks. Prepare six 50 µL protein solution aliquots in 0.6 mL microfuge *safe-lock* tubes (use only tubes with special Safe-Lock caps); label the tubes with their intended incubation temperature. Heated samples will be incubated in heat blocks at various temperatures between 40-90°C (see exact temperatures, first page of section 5B) for at least 15 minutes, then removed from the block **only**

when ready for IR data collection. Measure the incubation temperature (including room T) just before the end of incubation.

Please note that samples incubated at high temperatures may have a significant portion of the D_2O solvent condensed on the bottom of the tube cap; be sure to pipet any condensed solvent back into the tube, and re-mix the protein solution. Collect FTIR spectra as instructed in steps 2-20 and 22-25 in the detailed instructions.

When finished with the FTIR for the day, please gather all team members and the lab instructor to:

- **Turn off the N_2 Tank: main valve 1st, then drain gas from regulator.**
- **Turn off the FTIR power supply** (behind the computer monitor), **and turn off the computer.**

Finally,

- **Return clean, dry CaF_2 windows *to their packages* in the special EBI desiccator.**

RESULTS AND ANALYSIS: PART 2.1

Print out spectra of your two proteins at room temperature. Use the center frequency of amide I band to decide which proteins are primarily α-helix, or β-sheet. Support your conclusion using literature values for (a) the amide I peak location, and (b) the known helix vs. sheet content of the two proteins (from PDB, Uniprot, etc.).

RESULTS AND ANALYSIS: PART 2.2

After all FTIR spectra have been collected and saved, find a time when the instrument is not needed by other students for data collection. Teams will analyze their spectra, with one student working on the FTIR program (OMNIC) and the other recording values in an Excel spreadsheet on a laptop. After the first few spectra have been analyzed in this fashion, students should switch jobs (OMNIC vs. Excel) for the last few spectra. Analyze the spectra for your protein according to the detailed instructions found in the drawer underneath the FTIR instrument, on the desktop of the FTIR computer, or in Appendix 3 at the end of this lab.

POST-LAB: QUESTIONS

1. *Part 2.1, room temperature: Using the center frequency of amide I peaks, decide which native proteins are primarily α-helix, or β-sheet. Support your conclusion using literature values for the amide I peak location. Does your* ***qualitative*** *conclusion match the* ***quantitative*** *structural information available from Protein Data Bank or UniProt?*

2. *Part 2.2, thermal denaturation: Using the instructions for data analysis, identify relevant peaks and obtain corrected absorbance values for the peaks of interest. Construct a plot of (corrected) absorbance* ***ratios*** *vs. T. If you wish, you may also plot the shift in the amide I FTIR peak with T; depending on noise, you may have to locate the* $\overline{\nu}_{\max}$ *by hand.*

3. *From your thermal denaturation plots, determine the inflection point[8],* T_U*, as well as the uncertainty in your inflection point determination, and explain how you came by this value. Report this inflection point as the denaturation temperature,* T_U*,* ± *uncertainty, to the correct number of significant figures. Discuss uncertainty and error involved in determining* T_U*; you may wish to consult your uncertainty lab and aspartame lab reports regarding experimental error.*

4. *How well do your experimental* T_U *values match the literature? Calculate % error, if possible. What are your greatest sources of error in this determination of* T_U*?*

POST-LAB: INFORMAL REPORT (NONE)

POST-LAB: FORMAL REPORT (ORAL REPORT)

The last half of your oral report will be devoted to Part 2 of this lab, FTIR spectroscopic analysis of protein structure. In your introduction, discuss FTIR spectroscopy (especially as it pertains to α-helices, β-sheets and unfolded protein structures), and the denaturation equilibrium.

[8] There are several ways to determine T_U from the spectroscopic parameter vs. T plot: (a) estimate by eye the midpoint between the low T (N) and high T (U) asymptotes; (b) draw horizontal lines where you think $y_{\min}$ and $y_{\max}$ are, draw a horizontal line half-way between these two asymptotes, and determine where this midline intersects the curve defined by your data points; (c) plot the slope of each pair of points vs. T, and find the minimum of this first derivative plot; (d) use nonlinear regression to fit your data as outlined in Appendix C. The advantage of this last method is that you get a best-fit value not only for T_U, but for $\Delta H°_U$ as well.

Remember that organizing and labeling your results (e.g., spectra, tables with reported values ± uncertainties, plots) is important.

FURTHER READING

"Structural Analysis and Modelling of Proteins on the Web"
Leon, D. et al. (1998) *J. Chem. Educ.* 75, 731-734.

"Using IR Spectroscopy to Investigate Protein Stucutre"
Olchowitz, J.C. et al (2002) *J. Chem. Educ.* 79, 369-371.

"Thermal stability determinants of chicken egg-white lysozyme core mutants: Hydrophobicity, packing volume, and conserved buried water molecules"
P. Shih, D. R. Holland **AND** J. F. KIRSCH (1995) *Protein Sci.* **4,** 2050-2062.

"Critical Temperature of Secondary Structural Change of Myoglobin in Thermal Denaturation up to 130 °C and Effect of Sodium Dodecyl Sulfate on the Change"
Y. Moriyama and K. Takeda (2010) *J. Phys. Chem. B* **114,** 2430-2434.

"Thermal Denaturation of Myoglobin. I. Kinetic Resolution of Reaction Mechanism"
E. S. Awad and D. A. Deranleau (1968) *Biochem.* **7,** 1791-1795.

APPENDIX 1: PYMOL "3-BUTTON EDITING" AND "3-BUTTON VIEWING INSTRUCTIONS

By clicking in the box in the lower right hand corner you can switch between two different modes: "3-Button **Editing**" or "3-Button **Viewing**".

1. The "3-Button **Editing"** Mode is the best for moving your proteins as units.

3-Button Editing Commands

Buttons& Keys	L	M	R	Wheel
Mouse Only	Rotate	Move	Zoom (drag cursor up and down)	"Slab": Cut away slices of protein, start from the front
Mouse + Shft[9]	Rotate one of the proteins	Move one protein	Zoom one protein	Cut away at one protein
Mouse + Ctrl	Move a specific atom[10]	Select specific atoms, residues, etc.[11]	Selects an amino acid as "pk1" for editing[12]	Combines Zoom and "Slab" feature
Mouse + Ctrl +Shft	Pick and rotate structures around a specific atom[10]	Allows you to change the origin of rotation	?	"Super" Zoom (not very useful)
Single Click	Select an amino acid (will show in command box)	Center the protein at selected point	Gives the menu feature where cursor is	-
Double Click	Move a specific atom[10]	screen says "Dragging '#' atoms in object"	Selects an amino acid as "pk1" for editing[12]	-

[9] When there are two proteins on screen together

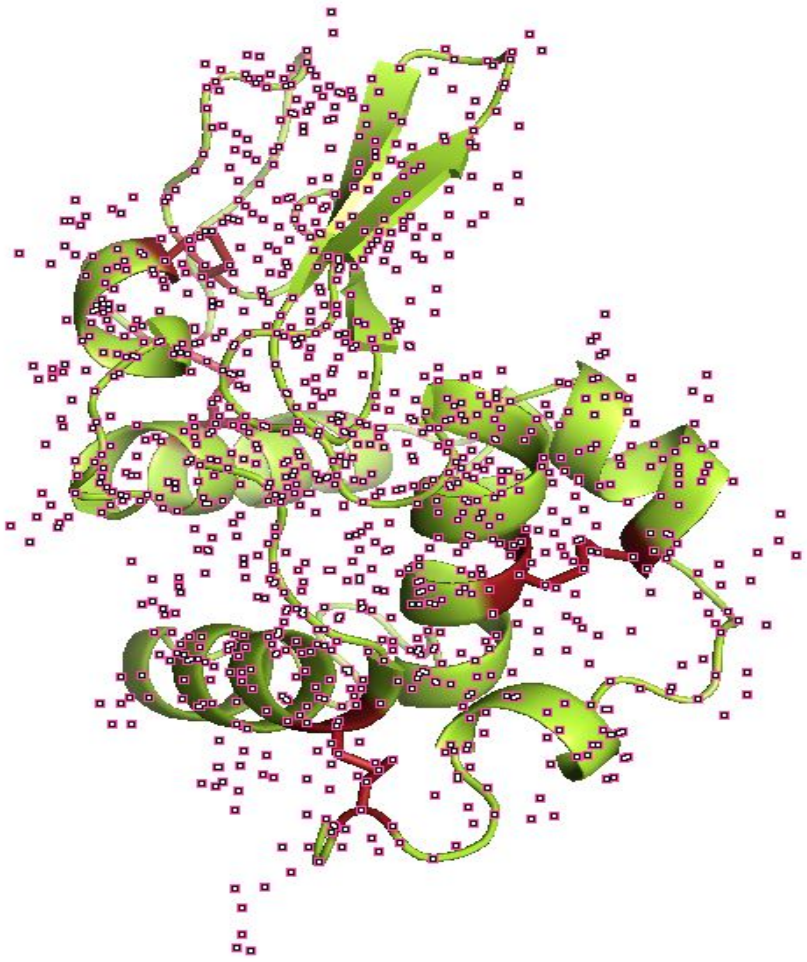
[10] Distorts the protein structure, so beware.

[11] Switch to "3-Button Viewing" by clicking in the lower R box. Choose the "selecting" object (e.g., residues, atoms, C-alphas, etc.); switch back to "3-Button Editing." Ctrl + mouse-middle-button allows you to select object.

[12] With Ctrl + mouse-right-button, click on spot of interest; the reside selected is specified in the last line in the Command line box (upper left, e.g., SER-88). This residue is now editable under "pk1" in the upper right corner, e.g., **A**ction, **S**how, **L**abel, **C**olor.

2. The "3-Button **Viewing"** mode allows you to choose what specific part of the protein you want to select. On the "Selecting" line in the "3-Button Viewing" box (lower R), you can specify: molecules, C-alphas, objects, segments, chains, residues, or atoms. What you select will change what happens in the 3-Button Editing screen when you click on the amino acid sequence at the top of the page.

For examples, if the 3-Button Viewing box (above) says you are Selecting Molecules, then when you try to select an amino acid from the sequence at the top of the screen the whole molecule will "light" up (see below).



3-Button Viewing Commands:

Buttons & Keys	L	M	R	Wheel
Mouse Only	Rotate	Move	Zoom (drag cursor up and down)	Cuts away slices of protein, starting from the front
Mouse + Shft	?	?	?	Causes proteins to disappear (?)
Mouse + Ctrl	Selects whatever object is designated in the "Selecting" line of the 3-Button Viewing box	Selects an amino acid as "pk1" for editing[12]. Clicking again deselects.	Selects an amino acid as "pk1" for editing[12]. Clicking again does not deselect.	Allows you to zoom in and out on the protein
Mouse + Ctrl +Shft	Selects a specific amino acid and places in the right hand column as "sele"	Allows you to select where the origin of the molecule is (point where molecule revolves around)	?	Allows you to zoom in and out on the protein
Single Click	Selects amino acids	Will center the protein at that point	Gives a menu box for the amino acid located where the cursor is	-
Double Click	Gives a menu box for the amino acid located where the cursor is	Tells which amino acid you clicked without selecting	Selects an amino acid as "pk1" for editing[12]. Clicking again deselects.	-

APPENDIX 2: FTIR INSTRUCTIONS, TAKING SPECTRA

NOTE: When changing samples, quickly use the **top door** of the sample compartment. The lever on the front of the sample compartment should remain in the **locked** position for the duration of the experiment. Keeping the sample chamber in a nitrogen-enriched environment as much as possible allows you to purge for shorter times between samples. If you open up the entire chamber rather than using the top slider door, you will have to purge for twice as long between running samples.

0. If the N_2 tank is low (below 1000 psi on the guage closest to the valve stem), make sure that a spare tank is available in the lab, in case you run out.
1. Ensure that the Transmission apparatus is in place in the sample chamber; ask for help if the ATR device is inserted and you are unsure how to remove it.
2. Ensure that the N_2 tank is turned on, and set to ≈ 10 psi, with knurled knob lined up with mark; purge the sample chamber for about 6 minutes.
3. Start up Program: OMNIC
4. Under "Experiment," at the top left of the screen, from the drop-down menu select "EB protein structure lab." This gives you the appropriate settings for this experiment such as the correct spectral window; because D_2O absorbs significantly in the 2000-3000 cm^{-1} region, it's best to select scanning limits from 1750-1350 cm^{-1}.
5. The first group of students (Team A) should locate within the EB folder a subfolder named your "Semester_year" (e.g., Spring 2013). The file path should be: Computer/OS(C:)/My Documents/Omnic/Spectra/EB/S 20xx. Make a new subfolder for your spectra; later you will save your spectra files, with filename format "protein_year_temp" (e.g., Mb_2013_40C) to this subfolder.
6. Remove the two clean, dry CaF_2 windows, stored in packages in the dessicator. (Be sure to put them back when you're done for the day!) Place a 0.015 mm-thick Teflon spacer ring around the edge of one CaF_2 disk, and then place 16 μL of D_2O in the center of this disk. Place the disk in the window holder, and carefully lower the second disk on top of the first. Lightly press the two disks together by carefully screwing on the aluminum cover. Quickly open the top slider door, place the sample into the FTIR transmission chamber, and close the slider door.
7. Let the sample chamber purge with N_2 at 10 psi for 3 minutes. After the purge is complete, turn the N_2 pressure down to ≈ 0.
8. Choose COLLECT > Background > OK
9. When spectrum is complete, you will see the Pop-up window: "Add to Window 1?" > No
10. Even though a small contribution from the large H_2O and CO_2 peaks remains, it is minimal and does not affect the final spectra significantly. Note that the H_2O, CO_2 peaks Abs values are ~0.02, whereas the amide I/II peaks should have absorbance >10× greater.
11. When you are ready to collect a spectrum of your (next) sample, remove the window holder, and close the FTIR chamber door quickly.

12. Gently wipe D_2O from between windows with a Kimwipe. Transfer 16 µL of your sample onto the CaF_2 disc after it has been placed in the cell holder with a spacer. If protein is aggregated, dig out ½ of sample from the Eppendorf tube with a diagonally sliced (use a razor blade) end of a Q-tip or other instrument and blot it around the surface of the disc.
13. Place the other disc on top, carefully close the cell, and place in the FTIR chamber.
14. Purge with $N_{2(g)}$ for 3 minutes at 10 psi. After the purge, lower the N_2 pressure to ≈ 0.
15. COLLECT > Collect Sample. Enter a helpful, informative name for your spectrum, e.g., myoglobin 40C (this can be more expansive/informative than the filename)
16. Confirmation Window > OK, and scan occurs.
17. Pop-up window – "Add to Window 1?" > Yes
18. Your spectrum will most likely have a lot of noise in outlying regions, and the *x* axis will most likely give you the full spectral window: 4000-400 cm^{-1}. To fix this, select: VIEW > Display Limits > 1750 cm^{-1} – 1350 cm^{-1}.
19. If you need to, change *y* axis limits (VIEW > Display Limits) so that peaks are maximized in the screen but not cut off.
20. For **PART 2.1** (room temperature spectra) only:
    a. Select ANALYZE > Find peaks
    b. Click on the "Replace" box (to the right of the name box)
    c. Under "report" (top menu item), select "Print Report" to list peaks
    d. Print the spectrum: FILE > Print (Olin 303 printer); highlight the spectrum title (under "Find Peaks").
21. Save your spectral data, in your year's subfolder in the EB Folder (and in a sub-sub folder for your group). Use the filename format: protein_year_temp, e.g., "Mb_2013_40C", "cyt c_2011_23C".
22. If you don't want to have the next spectrum show up along with this one (recommended), select VIEW > Hide Spectra.
23. Rinse both CaF_2 discs in between samples with 1 drop D_2O and then wipe gently w/ Kimwipe – do this 3× total for denatured samples, 2× for native samples.
24. For your next sample, return to step 12. You may use the same background; one D_2O background at the beginning of the set of experiments seems to be sufficient, but it is fine to repeat the background if necessary (e.g., the computer crashes).

APPENDIX 3: FTIR INSTRUCTIONS, ANALYZING ABSORBANCE VALUES FROM SPECTRA

1. Open OMNIC program.
2. Open all spectra for all of the temperatures so that all of the spectra are in the same window.
3. Zoom into the region of interest: VIEW > Display Limits > 1750 – 1350 cm^{-1}.
4. In the drag & drop menu bar that shows which spectra have been opened, use the <shift> key with a mouse click to select all spectra.
5. Click on the "Aut Bsln" tool (upper tool bar, near the middle) to perform baseline correction.
6. In the drag & drop bar, select all original (unstarred) spectra, and hide them using: VIEW > Hide Spectra.
7. Select all auto-baselined (starred) spectra, and fit all of them on the *y*-axis using: VIEW > Display Limits. Inspect the region between 1750 – 1700 cm^{-1} to ensure that at some wavenumber the spectra have roughly the same absorbance and thus Auto Baseline was done correctly.
8. Print a copy of these overlaid spectra to include in your report. This will give you a rough idea of how the spectrum changes as thermal denaturation occurs. Compare the room temperature spectrum to the highest temperature spectrum: carefully describe all spectral changes due to denaturation. (Note that due to small variances in path length, the increases/decreases in absorbance values may not necessarily fall in the order you would expect.)
9. Select all starred spectra and hide them (VIEW > Hide Spectra)
10. Now you will individually select each hidden baseline-corrected (starred*) spectrum and obtain absorbance values at several wavenumbers of interest. First, open the room temperature spectrum. In the lower left blue toolbar panel, click on the "spectral cursor tool" (cross-hair, third from left). Use the cursor to determine the wavenumber and the value of the Abs_{max} for the native amide I peak (1620-1660 cm^{-1}).
11. Open the highest temperature starred spectrum next; find a new peak that appeared on the lower energy side of the amide I peak (between 1550-1600 cm^{-1}) and use the cursor tool to determine the wavenumber and Abs_{max} of this "denaturation peak".
12. Set up an Excel spreadsheet with three columns to record all of the Abs_{max} values at the two pertinent peaks you determined above for each of the seven temperatures. (Using the cursor tool, these can be seen as *y*-values at the lower left.)
13. In a fourth column, calculate absorbance ratios, dividing each Abs_{max} value for the native amide I peak by the Abs_{max} value for the "denaturation peak". This is your absorbance ratio that will be the *y* axis on your denaturation plot.
14. Plot the absorbance ratios vs. T. Determine the inflection point as outlined in footnote 8 above. Estimate the uncertainty in your inflection point determination, and explain how you came by this value. Report this inflection point as the denaturation temperature, T_U, ± uncertainty, to the correct number of sig. figs. If you wish, you may also plot the shift in the location of the amide I peak with T; this will give you a second data set from which you can determine T_U.

APPENDIX 4: PROTEIN DENATURATION CURVE-FITTING

Protein denaturation curves plot some spectroscopic y-value (e.g., $\tilde{\nu}_{max}$, normalized absorbance, etc.) vs. T. The denaturation mid-point temperature (T_U) is the temperature at which the equilibrium mixture of protein is 50% native (N) and 50% denatured/infolded (U). Above this temperature the equilibrium shifts toward more U, and below this temperature the equilibrium shifts toward more N.

For the denaturation reaction, N → U, $K_{eq} = [U]_{eq}/[N]_{eq} = f_U/f_N$, where f_U, the fraction of protein that is denatured/unfolded, $= [U]_{eq}/[P]_{tot} = [U]_{eq}/([U]_{eq} + [N]_{eq})$.

Some denaturation plots fall as T rises, that is, y_{max} is achieved at low T and characterizes the N form, and y_{min} is achieved at high T and characterizes the U form. For such a plot,

$f_N = (y - y_{min})/(y_{max} - y_{min})$, $f_U = (y_{max} - y)/(y_{max} - y_{min})$, and $K_{eq} = (y_{max} - y)/(y - y_{min})$.

From thermodynamics, $K_{eq} = e^{-\Delta G°(U)/RT} = e^{-(\Delta H°(U) - T\Delta S°(U))/RT}$

At $T = T_U$, $K_{eq} = 1$ and $\Delta G°_U = 0$, so $\Delta H°_U = T_U \Delta S°_U$, and $\Delta S°_U = \Delta H°_U/T_U$.

Using this equation for $\Delta S°_U$, we get: $K_{eq} = e^{-(\Delta H°(U)/R)(1/T - 1/T(U))}$.

Combining, we get, $K_{eq} = (y_{max} - y)/(y - y_{min}) = e^{-(\Delta H°(U)/R)(1/T - 1/T(U))}$

Finally, solving for y,

$$y = (y_{max} + y_{min} * e^{-(\Delta H°(U)/R)(1/T - 1/T(U))})/(1 + e^{-(\Delta H°(U)/R)(1/T - 1/T(U))})$$

A good Kaleidagraph fit equation would then be:

(m1+m2*exp(-m3/0.001987*(m0^-1-m4^-1)))/(1+exp(-m3/0.001987*(m0^-1-m4^-1)));
m1=50;m2=10;m3=30;m4=340;

where m0 = T(K); m1 = y_{max}; m2 = y_{min}; m3 = $\Delta H°_U$ in kcal/mol; and m4 = T_U in K.

If you have such a "falling" denaturation plot[13], you can obtain fitted values for $\Delta H°_U$ and T_U from this fit.

[13] If on the other hand you have a denaturation plot which *rises* with T, then y_{max} is achieved at high *T* and characterizes the U form, and y_{min} is achieved at low *T* and characterizes the N form. Such a plot can be fit to the same equation as above, except that the fit will return a negative value for $\Delta H°_U$, which must then be reported as positive.