

Assignment 2

Your name:

Your email address:

Once you are done with the exercise, email your answers as an attachment to gogarten@uconn.edu and daniel.s.phillips@uconn.edu

Goals:

- Have an understanding of the content of a protein data bank file
- Be able to save individual subunits into distinct pdb files
- Align structures of divergent proteins
- Use the structure based alignment to align the linear sequences
- Align structures of a catalytic subunit during the catalytic cycle
- Appreciate that even 80% sequence divergence (or more) can leave the protein structure very, very similar.
- Appreciate that for important proteins substitutions occur so rarely that proteins remain recognizable similar in structure AND sequence.

We will use chimeraX for today's exercise. If you have problems with your personal laptop, please avail yourself to the laptops provided in class.


Exercise 1:

Do the following:

Connect to wifi:

- Go to System settings (in the apple pull-down menu on the top left)
- select wifi
- select UConn Guest or UConn Secure (the latter is preferred, but you need to log in with your netID)
- Find the chimeraX program in the Application Folder

Start ChimeraX:

open finder , select applications, select chimera, OR use the spotlight (the magnifying glass on the top right) and search for chimera.

Start the Chimera program through double clicking the chimera icon (or right click and select open). Chimera is a program to visualize and analyze protein (and other molecular) structures.

If you are the first to use the program on the laptop, you get the message that the program cannot be run, because it is not from an authorized source.

- Go to System settings (in the apple pull-down menu on the top left),
- *Select* Privacy and security
- *Scroll down* and give permission to run chimera

If you manage to obtain a beautiful display of a structure, save the image as a jpg image and save the session (from the file menu) and put an image into your class-note-book.

ATPase subunits

Objectives:

- Have at least a rough understanding of the content of a protein data bank file.
- Save individual subunits into distinct pdb files
- Align structures of divergent proteins
- Align structures of a catalytic subunit during the catalytic cycle.

Go to the [protein databank](#), download the file [1bmf](#) (as pdb file). This protein is the head group of the beef-heart mitochondrial ATP synthase.

Open the file 1bmf.pdb in a text editor ([notepad++](#) or atom or wordpad).

Scroll down the file in your texteditor, notice the sections with Remarks, Sequence Residues, heteroatoms, secondary structures, and the atom coordinates. Notice of the journal citation, and in line 294 ff, **take a note of which chain (A,B,C,D,E,F, and G) corresponds to which subunits.** Both the alpha and beta subunits have an ATP binding site. The abbreviations E(mpty), DP (diphosphate), and TP (triphosphate) refer to the occupation of the beta subunit.

The alpha subunits always have an ATP, but they are named after the neighboring beta subunit to whose ATP binding site they contribute.

[Which chain correspond to alphaE, alphaDP, alpha TP, betaE, beta DP, beta TP?](#)

Your answer --->

- alphaE:
- alphaDP:
- alphaTP:
- betaE:
- betaDP:
- betaTP:

The ATP synthase (aka as proton pumping ATPase) consists of ring of proteolipids that are integrated into the membrane, a head group (which is the structure in 1bmf), and a

stator that keep the non-rotating parts fixed. The head group, known as F1, consists of 6 ATP binding subunits (3 alpha and 3 beta subunits). The beta subunits bind and hydrolyze ATP, if the enzyme works as a proton pumping ATPase. These catalytic subunits rotate the central gamma subunit. In the intact enzyme, the gamma subunit is linked to the proteolipids, which than rotate relative to the stator. When they pass the stator, the proteolipids (proteins that behave like a lipid, but they do NOT contain a lipid) undergo a motion that moves a glutamate or aspartate residue into a different environment, where it picks up or dissociates a proton.

Why is the ATPsynthase important? Make a guess as to how much ATP are synthesized in the human body per day (use google if you are unsure).

Your answer --->

How is ATP synthesis coupled to the electron transport chain? (If in doubt, check wikipedia on [chemiosmosis](#) and [ATP synthase](#))

Your answer --->

The beta and alpha subunit evolved from a very ancient gene duplication (this duplication had already occurred in the common ancestor of Bacteria, Archaea, and the eukaryotic nucleocytoplasm); this duplication had already occurred over 3.5 billion years ago (most recent estimate 4.2 billion years ago). This means that the two subunit types (alpha and beta) are separated by for over 8 billion years of evolutionary history (8,000,000,000 years - to appreciate this time scale, note that the universe is estimated to be only 13,000,000,000 years old).

- 1) Open chimeraX and open the 1bmf file (File > fetch by ID 1bmf)
Look at the structure in the Ribbon and Space filling preset modes. Note the central gamma subunit (consisting mainly of alpha helices). [aside: in the related structure of a transcription termination factor, which unwinds a newly synthesized mRNA from the DNA template, the six ATP binding subunits have a similar arrangement and the place of the gamma subunit is taken by the RNA DNA duplex].
- 2) Color the ribbon display according to beta sheet (yellow), helices (red), and coils (forest green).
To do this: *Select > Structure > Secondary Structure > Coil*
Then: *Action > Color > Forest Green*
Then: Select > Structure > Secondary Structures > Helix, then Actions > color > red (NOTE: actions apply only to the items selected!)

Select > secondary structures > strand then
Actions > color > yellow

- 3) Look at the structure of one of the ANP molecules. "*Select > Residue > ANP*"; "*Select > invert*"; "*Action > Cartoon >hide*", "*Select > invert*". (You might also need to hide the sidechains in the binding pocket, if they are visible.)
Display the ANP molecule as Ball-Stick (you can use the Molecule Display Panel on top of the chimera window)

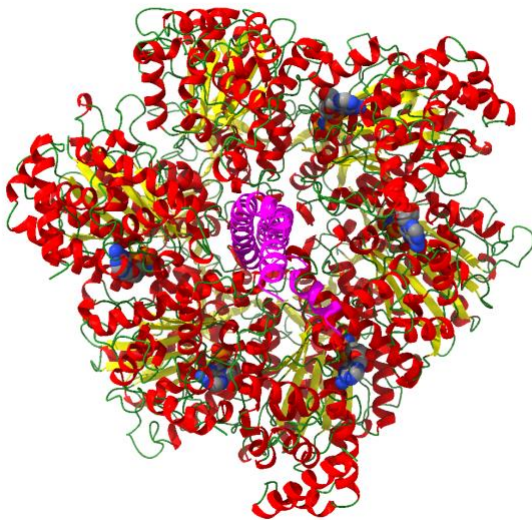
How is the PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (ANP) different from ATP?

Your answer --->

- 4) Determine **which chain**, via (Select > chain), **does not have an ATP or ATP analog bound?** (give both the Letter identifier and the descriptive name)

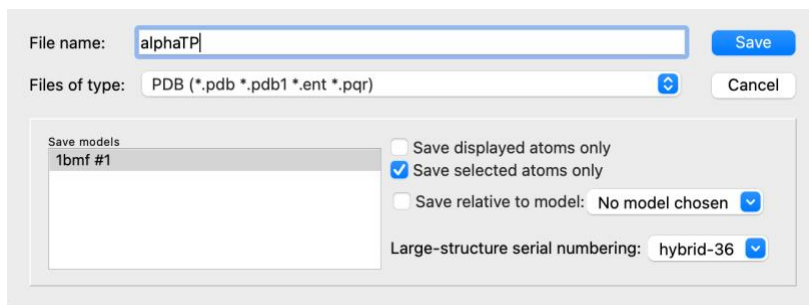
Your answer --->

Once your structure is displayed resembling this (I colored the chain G in magenta, you do NOT need to do this)



Select each of the chains in turn and save it as a pdb file (Select > chain . . . , then File > Save > select pdb as format -- in the window, enter a name (use the names of the subunits: alphaE, alphaDP, ...), select an appropriate directory, and **place a check mark into save selected atoms only !!!** Note, the menu does not remember the previously placed check marks.

This is a screenshot of how this should look like:

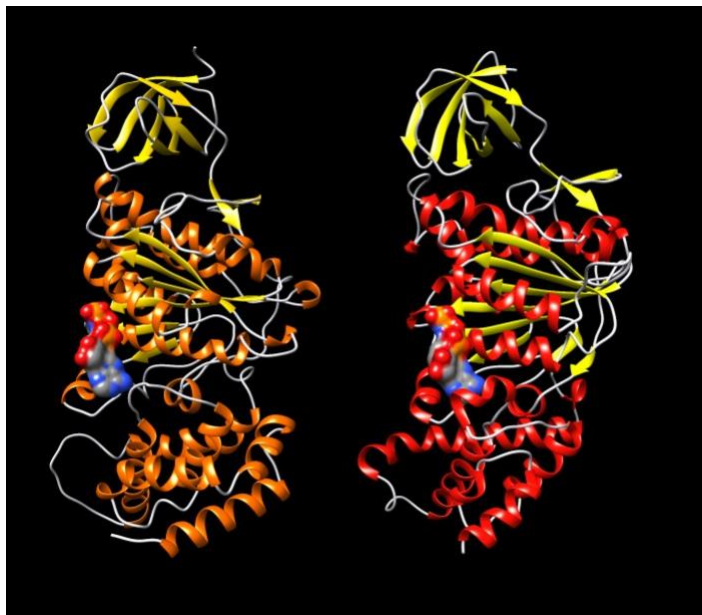


Once you saved the 6 chains (A to E), **Close your session !!!**

In chimera, open the previously saved files for alphaTP and betaTP (i.e., the subunits that contribute to the binding of ATP in the catalytic site).

Open the model panel and select one of the structures. In the top menu select “right Mouse” and then the rotate/turn icon and then use the right mouse button to move the structures side by side. **NOTE: Depending on the settings for your mouse, the right click on your trackpad may be rather difficult to use. Go to the apple menu on top, select system settings, move down in the bar on the left to “trackpad”. Under point and click, under secondary click, select “Click or Tap with 2 fingers”.** If you now select “move model” under right mouse, you can move the selected model when you use two fingers on the trackpad.

Choose a display that helps you to recognize structural similarity. If you found a good-looking display, save the session (careful, if you want to save many times along the way, give the session new names, else it overwrites the previously saved session)



Example of what your image might look like. The beta subunit is in slightly more orange colors.

Hopefully you are impressed by the degree of structural conservation over >8 billion years of evolution.

Does the similar arrangement of secondary structure elements relative to the ATP analog convince you that the two structures evolved from the same ancestral protein?

Your answer --->

To generate an automated alignment, use Tools > Structure comparison > Matchmaker. In the matchmaker window, select one of the structures as reference, the other as structure to match.

We will go with the default options, except for placing a checkmark in "Show pairwise alignments".

Both subunits bind ATP via the [Rossman fold](#). Conserved parts of this ATP binding site are the so-called [Walker motifs](#). The Walker A motif is G-x-x-x-x-G-K-[TS].

Can you locate this sequence in the structure-based alignment?

Are the Walker A motifs aligned between the two structures?

What amino acid follows the GK[T/S] in the beta subunit?

With which part of the ATP molecule do the Lysin (K) and Threonine (T) sidechains interact? (Select the GKT motif in the pairwise alignment using the mouse, then "action > Atom/Bonds > show sigchain/base"; giving the sidechains a different color also might help.)

Your answers --->

Visualize the catalytic cycle of the beta subunits:

Start a new Chimera session (**close the previous one first**).

Load the three beta subunits from the three separately saved files structures.

Apply a "good-looking" coloring scheme.

Save session as *Your_Name_for_the_session* (in case something goes wrong, you can recover from the saved session)

To align the subunits, Tools > Structure comparison > Matchmaker

Select betaDP as reference, and align the other two structures (control click adds to the selection).

Use the "model panel" to click on/off the visibility of the different chains.

Where during the catalytic cycle (betaE > betaDP > betaTP > beta E does the structure change most? How would you describe this movement?

Your answer --->

Morphing and movie making

We will morph the three structure into one another using the command line “morph” command. (Which is morphed into which depends on the order in which you loaded the subunits).

Open the commandline (under tools), enter

`morph #1,2,3`

If you use `morph #2,3,1 wrap true` morphing goes back to the beginning (#2 is the empty, three betaTP)

In the model panel, a new entry is created. Make this model visible, and color it according to secondary structure. You can save a movie of the morphing by clicking on the red button.

One problem is that the ATP (or rather ANP) molecule is absent in betaE, and thus is lost from the morphing. You can add the ANP from betaTP to the morphing: In the model panel, select betaTP, then “*select > Residues >ANP*”. Then *selection > invert*, then *Action > cartoon > hide*. If you have sidechains selected in the binding pocket, you need to hide these as well.

THE FOLLOWING IS INTERESTING, BUT NOT REQUIRED !

Comparing other divergent proteins with similar structures

A) the GRASP nucleotide binding site in enzymes that use ATP to synthesize something

Glutathione synthetase and D-Alanine D-Alanine Ligase were long ago (1990) recognized by Jim Knox (MCB emeritus faculty) to be so similar in structure that there could be no doubt about their common evolutionary origin. Later these and other enzymes were discovered to have a novel ATP binding site (the GRASP domain). These domains were identified using profile alignments (will be covered later in this course). A description of the GRASP domain family is at <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3249071/>.

Save the following files to your computer and load them into ChimeraX (new session).

[2DLN.pdb](#) (D-Alanine D-Alanine Ligase - D-ala is an important part of the bacterial cell wall, more [here](#),

[1GSA.pdb](#) (glutathione synthetase from E. coli, [glutathione](#) is the biological equivalent of mercaptoethanol),

[cpsBfrag.pdb](#), load [cpsFfrag.pdb](#) ([Carbamoyl phosphate synthetase](#) is an enzyme consisting of several domains. These are the front and the back of (1BXR download the fragments from the link, not the whole pdb file).

Based on your first impression, are these structures similar? homologous?

Your answer --->

Use Tools > Structure comparison > Matchmaker to align these structures?

Your answer --->

See [here](#) for an illustration of the structures in similar orientation.

Comparing other divergent proteins with similar structures

B) Inteins and mini inteins (i.e., inteins that have lost their homing endonuclease domain)

Most inteins are composed of two domains: one is responsible for protein splicing, and the other has endonuclease activity. A few inteins have lost the endonuclease domain completely and retain only the self-splicing domain and activity. The latter inteins are called **mini-inteins**.

The structures of several inteins have been determined through X-ray crystallography. Today we will use the following:

- *Saccharomyces cerevisiae* intein (PMID: 9160747, [1VDE](#)),
- the *Mycobacterium xenopi* mini intein (PMID: 9437427, [1AM2](#))

To do:

1. Open 1VDE in chimera. This structure has two chains. Select **chainA** and save the **selected residues only** into their own pdb file (file > save pdb > fill out the form, check save selected residues only, save). Close the session.
2. Reopen the saved chain A in chimera. Open *Mycobacterium* mini intein 1AM2. Depict the structures as ribbons and color them according to the secondary structure (if you press the shift key while executing an action, it should act on both chains; but there also is a check mark in the "Tools > Depiction > Color Secondary Structure". However, it also might look nice to use slightly different colors for the two chains. Rotate the two structures until you can see the similarities between mini intein and the large intein. (Use the "favorites > model panel" to rotate one or the other structure). Which part of the structures appears to be similar?

Your answer --->

3. Align two structures using Tools > structure comparison > matchmaker. Does the alignment correspond to your expectation?

Your answer --->

4. You can find out which part of *Saccharomyces cerevisiae* intein corresponds to the endonuclease domain by comparison of the two structures. The splicing domain matches between the two structures. The part that does not match is a candidate for the homing endonuclease domain.

Using "tools > compare structure > match -align" create a pairwise alignment of the two inteins based on the match between the two structures (you need to have had run matchmaker first!).

Color the putative self-splicing (i.e. the part that is present in 1VDE and in 1AM2) and endonuclease domains of 1VDE (no corresponding part in 1AM2) in two different colors

(selecting consecutive residues works easily via the alignment window; if you press shift, you can add to the selection.) Do you get the same result as above?

Your answer --->

In Tools > Structure Comparison > Match align you obtain a sequence alignment. In that alignment window, click on info, and then percent identity. This returns the % identity in the matched part of the sequences (at the bottom of the window displaying the structure)

What is the percent identity between the aligned parts of the splicing domain?

Your answer --->

Save your project.

5. Find and select the N and C terminals (first a.a. and the last a.a.) in both structures. If you hover over the beginning or end, the name of the residues pops up in a little window. CTRL click selects the amino acid or atom, and shift control click adds to the selection. Under actions>atoms/bonds>show side-chains make the side chains of the first and last amino acid visible.

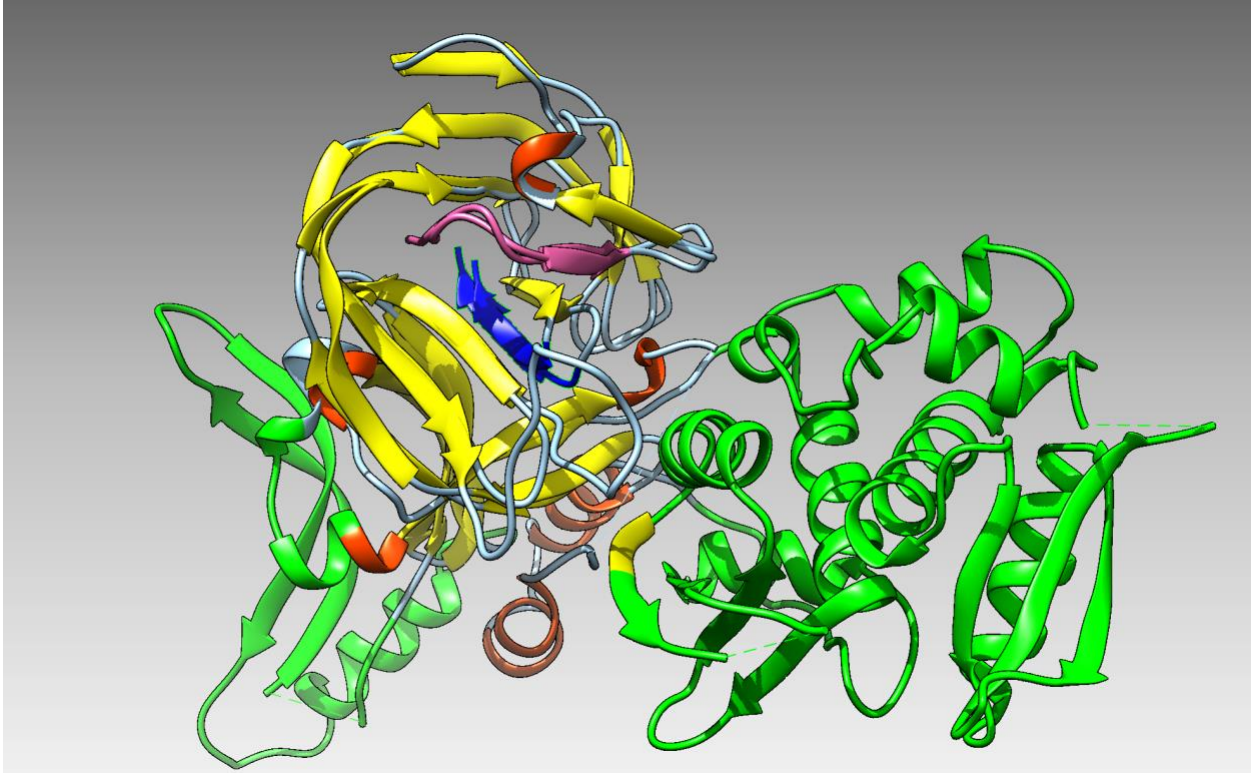
Optional: hide the rest of the structure, if it distracts you.

Rotate the structure of the sidechains of the first and last aa and decide which atoms are closest. Select these atoms (ctrl click and shift ctrl click), then go to tools>Structure analyses> distances and click on the **create** button in the window that pops up. Repeat this for a few atoms from the first and the last aa. How close are beginning and end (in Ångström and in nanometers)?

Your answer --->

in 1VDE:

in 1AM2:



Alignment between 1VDE chainA and 1AM2. The non-matching part is in green, the end of the intein in blue and the beginning in purple.

For more things to try out with divergent protein sequences go to the bottom of the assignment 3 page at

https://j.p.gogarten.uconn.edu/mcb3421_2021/labs/assign3.html