Chemical Biology 03 – Fall 09

**LAB 6: The Beta Globin Gene and some Variants**

**(using the Central Dogma)**

**October 21, 2009**

Today we will delve into the vast treasure troves of information available on hte beta globin gene and its encoded protein and its many human variants. You are going to build a report on this gene and several specific genetic variants (you will work in pairs and each pair will study a different set of mutations - there are more than 700 known mutations in this gene!). You will begin today by exploring several web-based tools where you can collect, copy, manipulate information, and you will initiate a powerpoint file in which you can collect your work and relevant information. Next week, each pair will give a presentation to the class to teach us about the different beta globin mutations that you studied. Your analysis will draw from a variety of material that we have covered in the class – ranging from amino acids and protein 3D structure to the components and workings of a gene. By next week, as a class, we will have learned about a great number of mutants and we will have really worked the Central Dogma and Protein biochemistry.

1) Begin a powerpoint file in which you will dump your work as you go along.

2) Go to the National Center for Biotechnology Information (NCBI) web site:

<http://www.ncbi.nlm.nih.gov/>

3) Search in “gene” database for beta globin

the first link should be for homo sapiens beta globin, so click there

bottom of page will show a map of the gene;

recall that beta globin has three exons:

Questions:

Q1) how are exons depicted?

Q2) How are introns depicted?

Q3) Explain what is represented by the blue region and how this differs

from the red region.

4) Click on the red link just to the right of the map that says “NP”, this will take you to the protein sequence in single letter amino acid designation; choose the “FASTA” format: Copy this sequence into your powerpoint file.

5) Click on the blue link just to the left of the map; choose the “FASTA” notation; this will take you to the nucleotide sequence.

6) Open a new window and go to: Expasy translation tool (<http://www.expasy.ch/cgi-bin/dna_aa>)

This is a convenient tool that will translate any nucleotide sequence for you; go ahead and copy the beta globin sequence and paste it in the window.

You have three options for viewing the result. For now, choose “includes nucleotide sequence”

Questions: Study this page and figure out what is going on.

Q4) Why do you get six options?

Q5) Why don’t any of these start with the met-val-his-leu (MVHL) that you know begins the beta globin protein?

*more questions on next page*

Q6) How does the ribosome that encounters this sequences (assume it is RNA), generate only one type of translation (ie. how does it choose between these different options)?

Q7) Which of these options yields the beta globin protein?

7) Click on the red link just to the right of the map that says “CCDS”. You should see two sequences, one nucleotide and one protein.

Q8) How does this nucleotide sequence differ from the one you saw before?

(you can repeat the exercise above, using the Translation Tool to help you).

Q9) Can you figure out what the color scheme corresponds to in the

sequence (some parts in black and one part in blue) relative to the beta

globin gene? Explain.

8) Now try something really cool – click on a codon or two… or click on an a.a.

9) Copy the nucleotide sequence and paste it into your power point presentation. You can later use colored fonts or shading or blocks to highlight the different regions of the gene.

10) Going back to the original map of your gene, click on the black link above the gene; you should get a nucleotide sequence.

Q10) What does this sequence correspond to? How is it different from the

others we have looked at? (you can use Translation Tool again if you like).

11) Cut and paste this sequence into your ppt file. Use alternate font colors or shading to highlight the exons; it will take some work to find them (if you prefer to first try this on hard copy with a highlighter, I will have printed copies of the sequence available in class).

***Time to find some mutants!***

12) Go to the Hemoglobin Variants Database (a web-based catalog of all mutations known to be in any of the globin family genes): http://globin.bx.psu.edu/cgi-bin/hbvar/counter

On the page of “summaries of mutations” click on “entries involving the beta gene.” This is a long collection of human mutations identified in the beta globin gene. Most have names that correspond to the town/country from which the patient came. They are listed in order of the position of the change to the DNA. Look at the list of mutants provided in class and sign up for two of the requested mutations and two of the “your choice” mutations. On the long list of 740 mutations you can find a particular one by typing its name in the “find” window of the web page. Otherwise, cruise around the page and choose something that sounds/looks interesting. By clicking on the link you will get an information page for each mutation. This is full of information that you will use as your resource for building your presentation. Your goal will be to characterize each mutation in terms of DNA, RNA, protein, protein structure and any known clinical, demographic, or other interesting information associated with this mutation. In order to consider the effect of a change in protein sequence, you will use another online database and its convenient manipulations. You will explore protein structure using the URL below to allow you access to the National archive of structural coordinates, the Protein Data Bank, or PDB. The next page will take you step by step through the use of this site. When you are finished, don’t forget t come back to step #13 below.

13) In addition to the four mutations listed above, sign up for one of the thalassemia mutations. This mutation will also be part of your presentation, and should be characterized in detail, however you will not do any protein characterization on this mutation.

**Additional guidelines for Beta Globin Gene Report**

to be presented in lab on Wed. Oct 28.

Together with your partner, you will build a Beta Globin Gene Report in powerpoint format and present this next week:

Here are some guidelines and questions you should be sure to address:

1) Where is the gene located? Provide a chromosome number and map position (each chromosome has a number and a “p” half and a “q” half which are each further divided into numbered sections; thus a number such as 9q34 designates the chromosomal “address” of the gene that determines your ABO blood group and is found on the q half (or “arm”) of chromosome #9 in region #34). You should be able to find this information just below the original map of the gene.

2) schematic map of beta globin gene including exons and introns.

3) sequence corresponding to primary RNA transcript of beta globin gene with exons demarcated

4) sequence corresponding to mature mRNA transcript of beta globin gene and sequence of translated protein.

5) For each of your mutations:

a) name

a) DNA change

b) what type of mutation is this?

c) effect on gene expression and/or protein sequence

d) speculation and rationale for effect on protein structure/function. Use what you learned about the structure and features of different amino acids as well as the details of beta globin secondary, tertiary, and quartenary hemoglobin structure to provide as complete a story as you can for each mutation.

e) clinical information on this mutation

f) frequency and demographics of this mutation if available

g) any other information you have come across that is interesting

6) Note that your thalassemia mutation analysis (step 13) does not require step **d**, but be sure that step c is thoroughly explained.

**Using the Protein Data Bank (PDB) to explore protein structure**

The following URL will allow you to access to the National archive of structural coordinates, the Protein Data Bank, or PDB: <http://www.pdb.org/pdb/home/home.do>

On this site, there is an excellent online tutorial about hemoglobin that I would urge you to look through. Read through quickly now, but I’d suggest you read it through in detail at another time. For now, copy this URL into the browser and see the information provided here for hemoglobin, which has been featured on this site as a molecule of the month. There are four pages to this tutorial, the first page with an intro, the second page with a wonderful animation of hemoglobin cooperativity, a third page highlighting some troubled hemoglobins, i.e. sickle cell hemoglobin, and a last page on exploring the structure. <http://www.pdb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/pdb41_1.html>

To check out these structures for yourself, find the search box next to the words “pdb id or keyword” at the top of the screen and type in the code for hemoglobin “2hhb” and then clicking on SEARCH. After a few seconds, a window will appear with a cartoon image of hemoglobin in the right side. There will be the option to view the structure using jmol – which is a JAVA application. Select the “view in jmol” option, and be patient. It will take a few minutes for Java to start and the application to load. When it finally displays, you will see a tetramer displayed on your screen, with subunits shown in blue (A chain), green (B chain), pink (C chain) and yellow (D chain). You are looking at a backbone representation, so you don’t see all the side groups and hydrogen bonds and so forth. Notice that you can use the mouse to spin the structure around and view it from several angles. Play a bit with the structure.

1. What type of secondary structure do you see predominates in hemoglobin?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. The four heme groups are displayed as small balls and sticks. These atoms are color coded by element (black is carbon, blue is nitrogen, red is oxygen). Also notice that you can zoom in and out on the structure by holding down the shift key AND the left button while moving the mouse forward and back. This will allow you to find the iron atom at the center of the heme. What color is the iron atom? Is this oxy or deoxy hemoglobin?

Fe color:\_\_\_\_\_\_\_\_\_\_\_ , oxy or deoxy \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Two sets of chains make little to no contact. What are they? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Also notice that if you let the mouse linger over a position, an atom identifier will pop up with the name and number of the amino acid. Find one amino acid from each chain. Write their identifiers here.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ , \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_,

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ , \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Notice that there is a Jmol Script box underneath the picture. You can use this box to type in commands. To select a particular residue, say, amino acid 15 on chain B, type “select 15:B” and hit enter. Once you have done this, this amino acid will be selected. But you still won’t be able to see it until you type in “color red” and hit enter and then “spacefill” and hit enter one final time. Now, the amino acid you have selected will be displayed with large red balls (hard to miss). What is the identity of this amino acid? What side of the helix is it on (inside or outside) and does that make sense with what you know about the polarity of this residue?
2. Notice that there is a sequence tab button on the top of the page. Use the information on this tab to identify which chains (A, B, C, or D) are the two alpha subunits and which chains (A, B, C, D) are the two beta subunits.

A: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, B: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, C:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, D: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Notice also that you can access a whole host of commands by clicking on the red “Jmol\_S” at the bottom right corner of the image. While its tons of fun to explore all these options, we will leave that for you to explore on your own and in the future. Just confirm that you do see the box and that you have clicked it and been impressed by the hundreds of commands that are hidden underneath this little icon.
2. Working in groups of two, we invite each of you to explore two Hb mutations at position 15\* that you have highlighted on the previous page. These two mutants are Hb Belfast (Trp 🡪 Arg) and Hb Randwick. (Trp 🡪 Gly). See the accompanying page for the structures of these amino acids. Go back to the window with the jmol structure for the reference hemoglobin. In the Jmol script box, once again locate position 15 in beta globin (if you’ve clicked away from that screen) by typing in the select commands as previously directed. Remember, if your mutation is at position 15 in the beta chain, you should type “select 15:B” then hit enter, then type “color red” then hit enter, and finally “spacefill” and enter. This will highlight the position of the mutation. Once you have done this, you should consider how each of these mutations might affect the structure or function of this protein. Things to consider are how similar the mutant amino acids are to the wild type amino acid. An acidic residue like a glutamate in the reference sequence being changed for an aspartate won’t be too bad because they are both acidic. But if you change to a radically different amino acid, how might that change the function? Also, if your side group is exposed to the solvent, that might be important if you change from a polar residue to a non-polar residue. Similarly, an amino acid side group that is buried and nonpolar that is mutated to another nonpolar group isn’t too bad, but changing it to a charged group is horrible! A comment about each of these mutations should be part of the powerpoint presentation you make next week. If you wish to illustrate the slide with a screen capture of the image, make it as big as possible on the computer screen, and then simultaneously press the keys Ctrl, ALT, Print Screen, and then Paste (Ctrl V or right click paste) into your powerpoint document and crop with the picture edit tool. Also include on your power point slide the name and ID of the mutation is and what you might suggest would be the impact of these changes on the structure or function of this protein.

[Normally, in most proteins, you would need to make sure that the sequence of your mutant protein aligns with the reference protein. That is not necessary here because hemoglobin is such a well studied protein, everyone automatically aligns their mutants to the reference sequence. But for the future, if, for example, your mutation is in amino acid 53 in chain A, make sure that position 53 in your mutant sequence is also position 53 in the reference sequence. The numbers don’t always line up exactly if there have been some amino acid insertions or deletions. How do you do this? Use the adjacent two or the amino acids on either side of your mutation to make sure your sequence is in register with this reference sequence. If it is, then the same numbering applies, and if it doesn’t, keep searching along the reference sequence until you find the amino acids that match on either side of your mutation site. Then when you will model your mutation on the reference protein, you will use that number and not the number given in your mutated hemoglobin. Most of the time, these numbers will be identical so don’t fret. ]

1. Maybe it didn’t seem as though the mutation would have a great effect on this deoxy form of the hemoglobin but it might have a large effect on the oxy form. So go back to the search box at the top of the page, and next to the pdbID or keywords, insert 1hho and hit the SEARCH button. Again, in a matter of seconds, the oxyhemoglobin structure will appear. Select View in Jmol button and wait as Java starts this application. While you are here, pause to find the bound oxygen, and see if you can appreciate that this structure is more “relaxed” , more compact than the deoxy “Tense” state was.

Repeat the series of jmol entries that will let you select, color and spacefill your mutation site, and comment on if there are any effects this mutation will have on the oxy state that were not apparent in the deoxy state. Write any observations or questions in your powerpoint presentation. Again, an image can be captured using the screen capture mode and pasted and cropped into powerpoint.

1. Sign up for your mutation on the class list and use Jmol to explore the implications of this mutation on the structure and function of the protein. This info should go into your class presentation.

Name\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Chemical Biology 03 – Fall 09

**LAB 6: The Beta Globin Gene and some Variants**

**(using the Central Dogma)**

**October 21, 2009**

**Questions from Lab Work**

due Monday Oct. 26

Q1) how are exons depicted?

Q2) How are introns depicted?

Q3) Explain what is represented by the blue region and how this differs

from the red region.

Q4) Why do you get six options?

Q5) Why don’t any of these start with the met-val-his-leu (MVHL) that you know begins the beta globin protein?

Q6) How does the ribosome that encounters this sequences (assume it is RNA), generate only one type of translation (ie. how does it choose between these different options)?

Q7) Which of these options yields the beta globin protein?

Q8) How does this nucleotide sequence differ from the one you saw before?

(you can repeat the exercise above, using the Translation Tool to help you).

Q9) Can you figure out what the color scheme corresponds to in the

sequence (some parts in black and one part in blue) relative to the beta

globin gene? Explain.

Q10) What does this sequence correspond to? How is it different from the

others we have looked at? (you can use Translation Tool again if you like).