**Assignment-4**

**(Full marks = 100)**

**(Sample answer)**

Answer:

Hemoglobin subunit beta [Homo sapiens]

NCBI Reference Sequence: NP\_000509.1

>gi|4504349:7-146 hemoglobin subunit beta [Homo sapiens]

EEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKY

In the FASTA header, you'll see the numerical region, for example (282-381) shows that the domain is positions 282-381 of the full protein.

\*\*\* Questions \*\*\*  
1) In this protein, what is the amino acid range of the conserved domain? [4 points]

Answer: The amino acid of the conserved domain was (7-146) aa.

2) What's known about the biological role of that domain in your protein?  For example, you can discuss how mutations in that domain cause interesting phenotypes.  Or, if that domain is essential to the function, for example, it's the ATP-binding domain in a kinase. [5 points]

Answer: Hemoglobin (Hb) is an allosterically modulated heterotetramer that plays a pivotal role in oxygen transportation of red blood cells. It has a very high affinity for oxygen in its relaxed state, but in its tense state Hb has as low oxygen affinity. Binding of various effector molecules can affect the conformational state of Hb. Molecules such as hydrogen ions, carbon dioxide, chloride ions, and 2,3-bisphosphoglycerate. There other variations of Hb such as HbA2 (expressed in adults at low levels) and Hb Gower-2 (alpha2epsilon2) (mostly present in embryonic period. There are many variations of Hbs that come along variations in oxygen binding affinity**.**

3) Paste the FASTA AA sequence of the domain into your homework document. [4 points]

Hemoglobin subunit beta [Homo sapiens]

NCBI Reference Sequence: NP\_000509.1

GenPept Identical Proteins Graphics

>gi|4504349:7-146 hemoglobin subunit beta [Homo sapiens]

EEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKY

=====================================================================  
PART I: PDB query for homologous structures  
=====================================================================

- Open PDB at http://www.pdb.org  
- Select Sequence Search (top center of page)  
- Paste your domain's FASTA sequence.  
\*\*\* Questions \*\*\*

1) Did PDB find your \*exact\* protein in the results? [2 points]

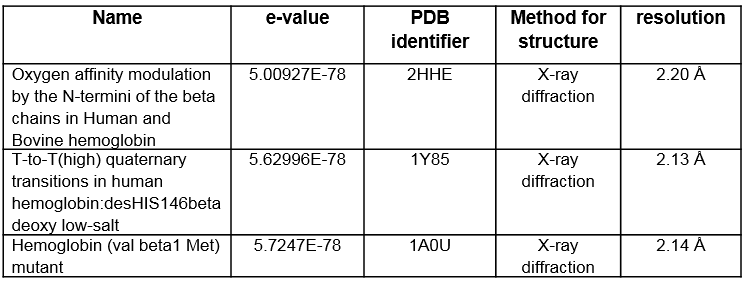
**Answer: PDB was not able to find an exact protein because there were many possible homologs that were generated. However, it is strongly implied, these are homologs to the original protein of interest.**

2) Did PDB find any plausible homologs to your domain? [5 points]

**Answer: Yes, there were some homologs with 100% identities (with zero gaps), along with a low e-value.**

1. **Oxygen affinity modulation by the N-termini of the beta chains in Human and Bovine hemoglobin**
2. **T-to-T(high) quaternary transitions in human hemoglobin:desHIS146beta deoxy low-salt**
3. **Hemoglobin (val beta1 Met) mutant are the top 3 plausible homologs to my domain.**

3) For \*each\* of the \*top 3\* plausible homologs to your domain do these questions: [12 points]  
a) What is the e-value of that hit?  
b) What is the PDB identifier (a 4-character/number code like 1A23)  
c) How was the structure elucidated (X-ray, NMR, etc)?  
d) Was the resolution given?  If so, what is the resolution?  
**Answer: The box below answers all the questions asked.**

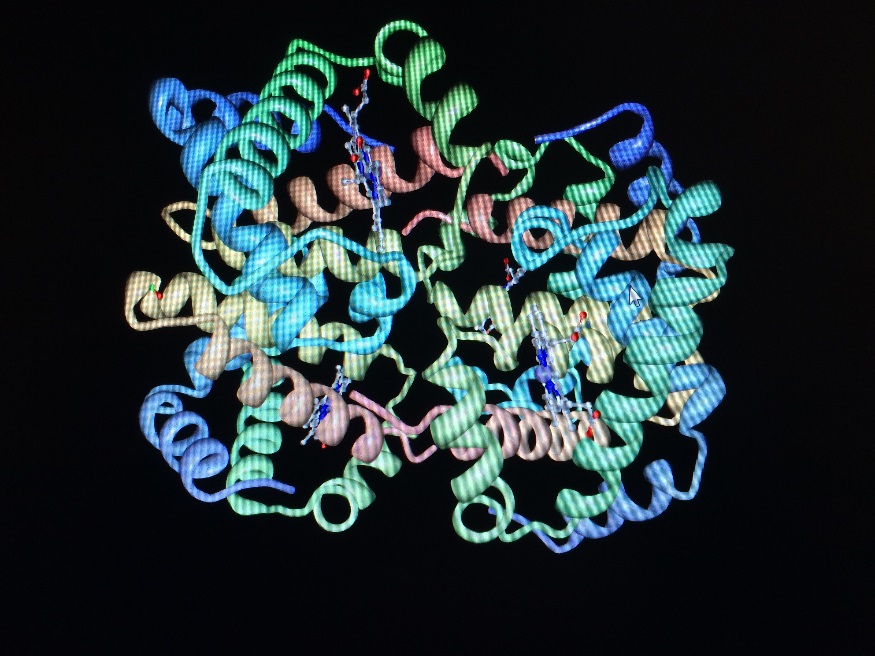
  
4) Choose one of the likely homologs (doesn't matter which), and click on its title. [5 points]

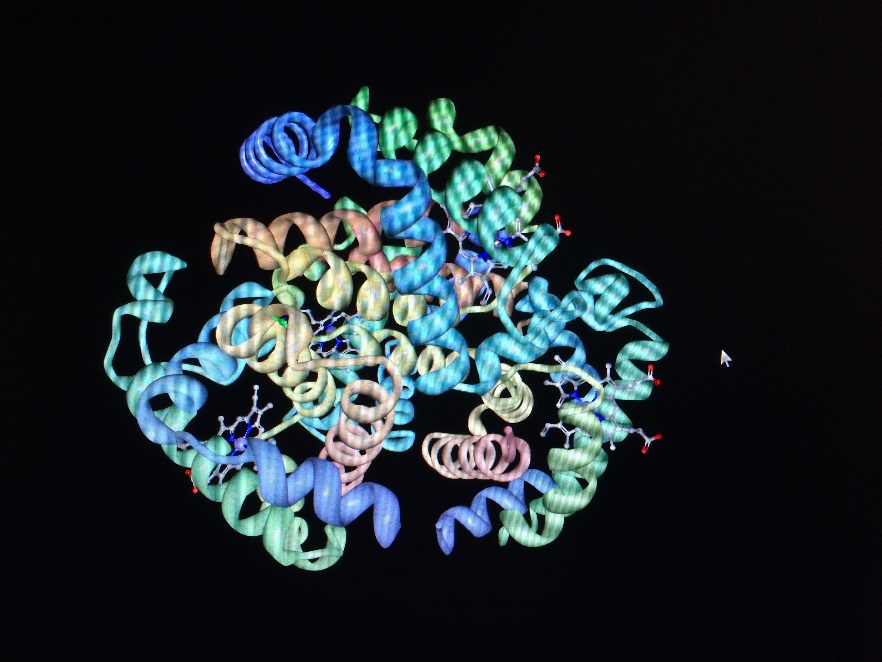
a) On the right of the screen, under the protein visualization, click "Protein Workshop"

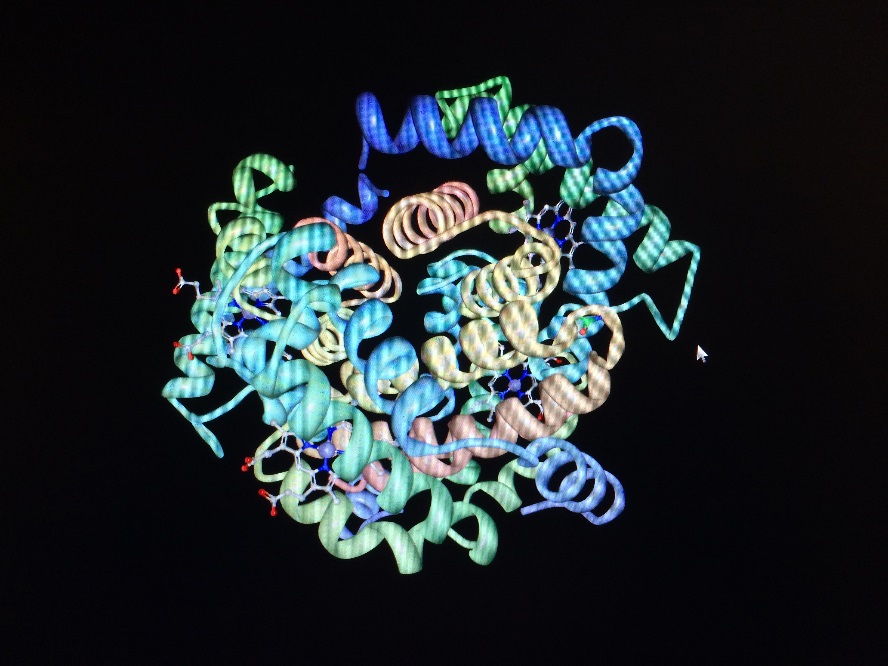
b) In Protein Workshop, save the .jpg or .png image to a file (use File in top menu), and put that .jpg or .png into your homework word processor document. If you cannot save the image with "File" -> "Save image...", then use print screen to create a snapshot of your video display. To invoke print screen on Windows use "control-printScreenButton" and on the Mac use "command-shift-3" or "command-shift-4".

c) Rotate the protein to a very different orientation, save that .jpg, and put it into your word processor file. (leave the Protein Workshop window open, you'll use it in the next section.)

**Hemoglobin (val beta1 Met) mutant**



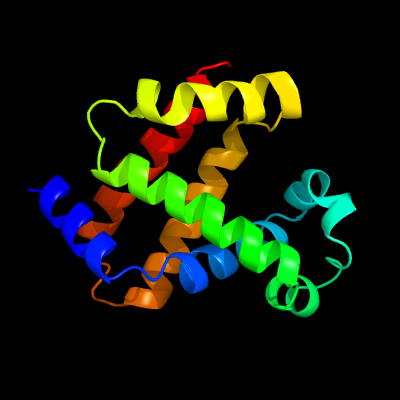


  
  
=====================================================================  
PART II: Homology modeling using Phyre2  
=====================================================================  
Read the Phyre paper (Kelley2009 on Blackboard).  This is a protocol paper, which just means it's cookbook recipe with detailed instructions.

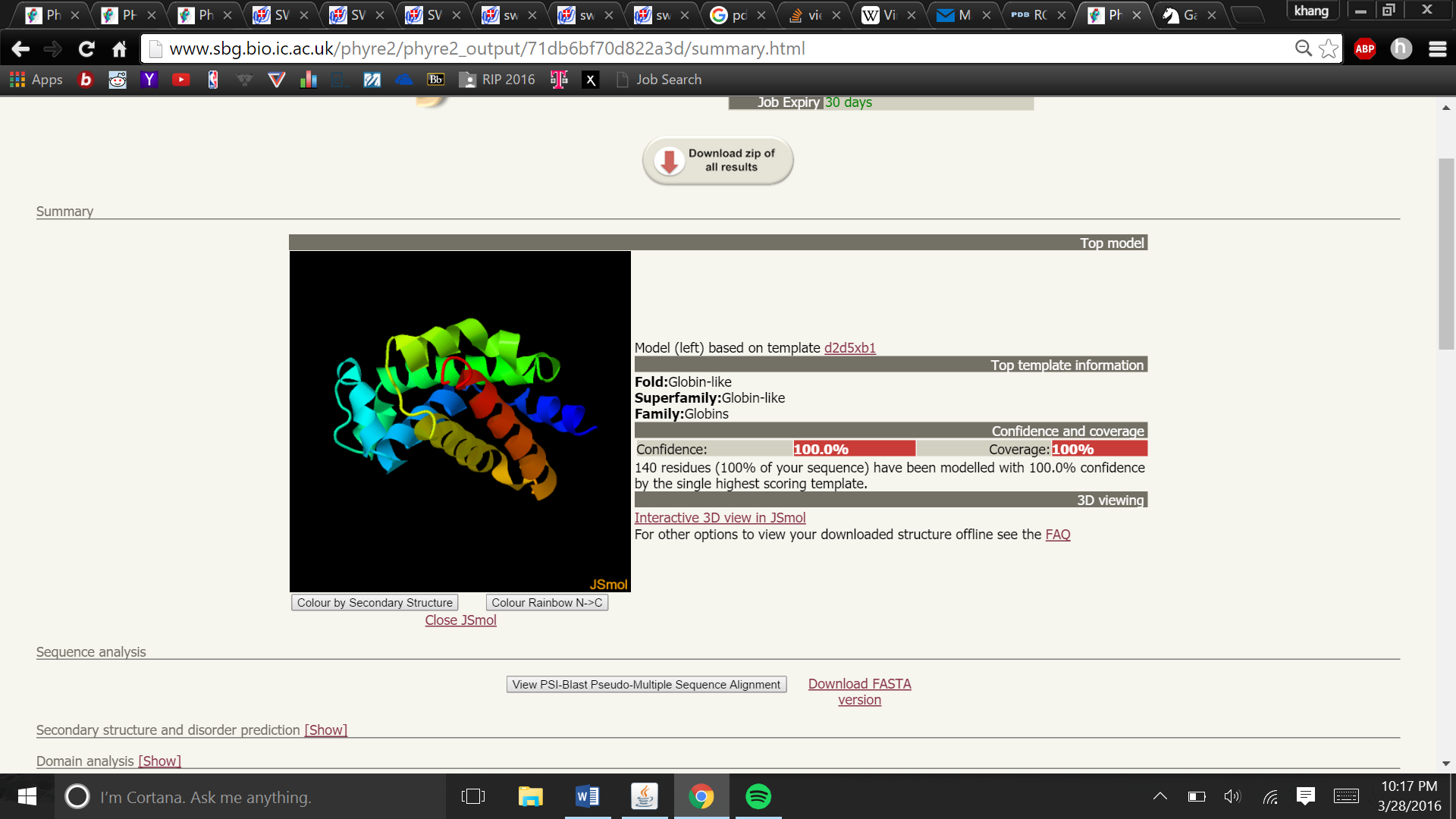
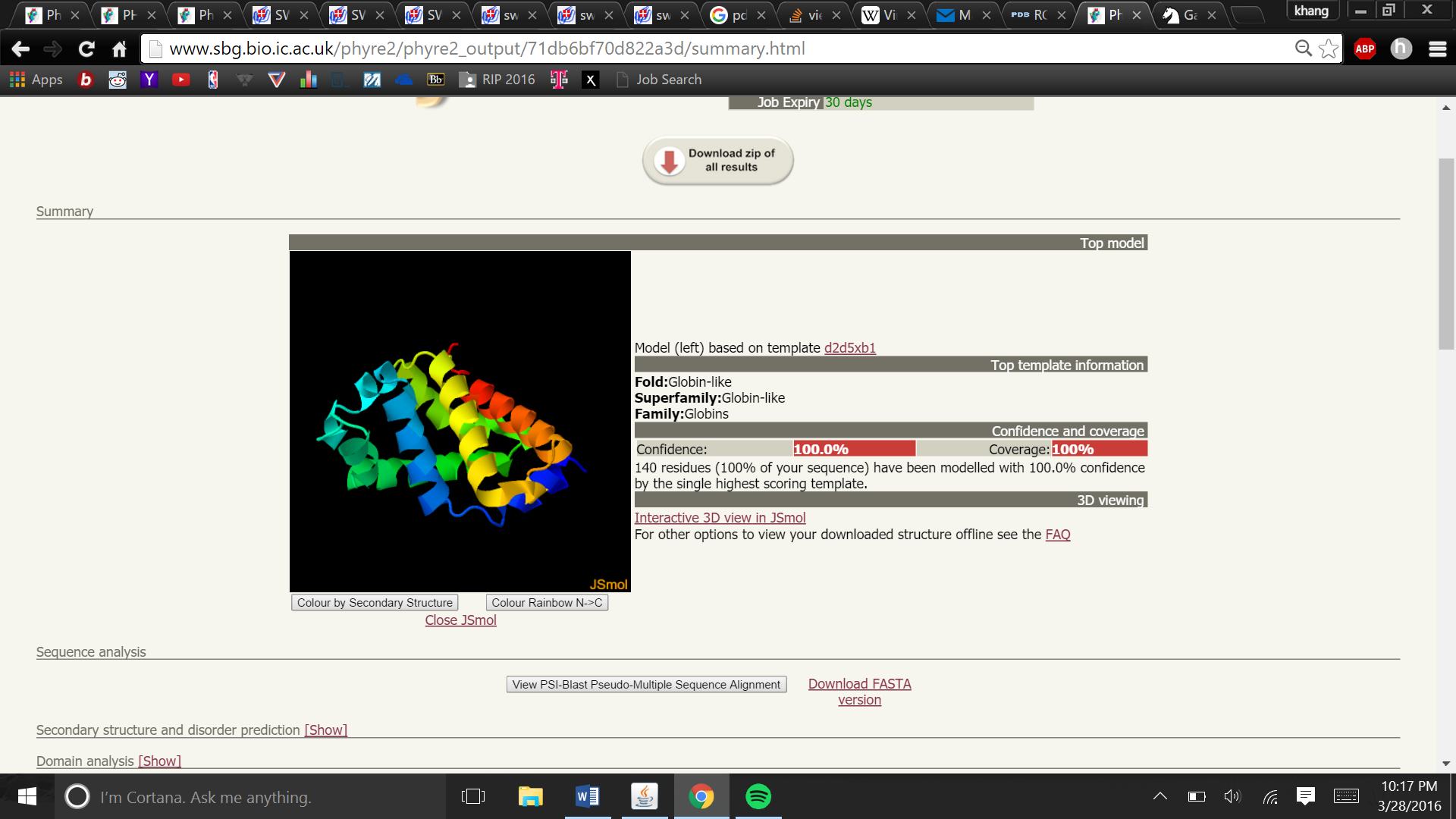
Read the textbook section on protein structure prediction and especially homology modeling

Watch the Phyre2 YouTube video at **https://www.youtube.com/watch?v=Adm8JQZMmj4**

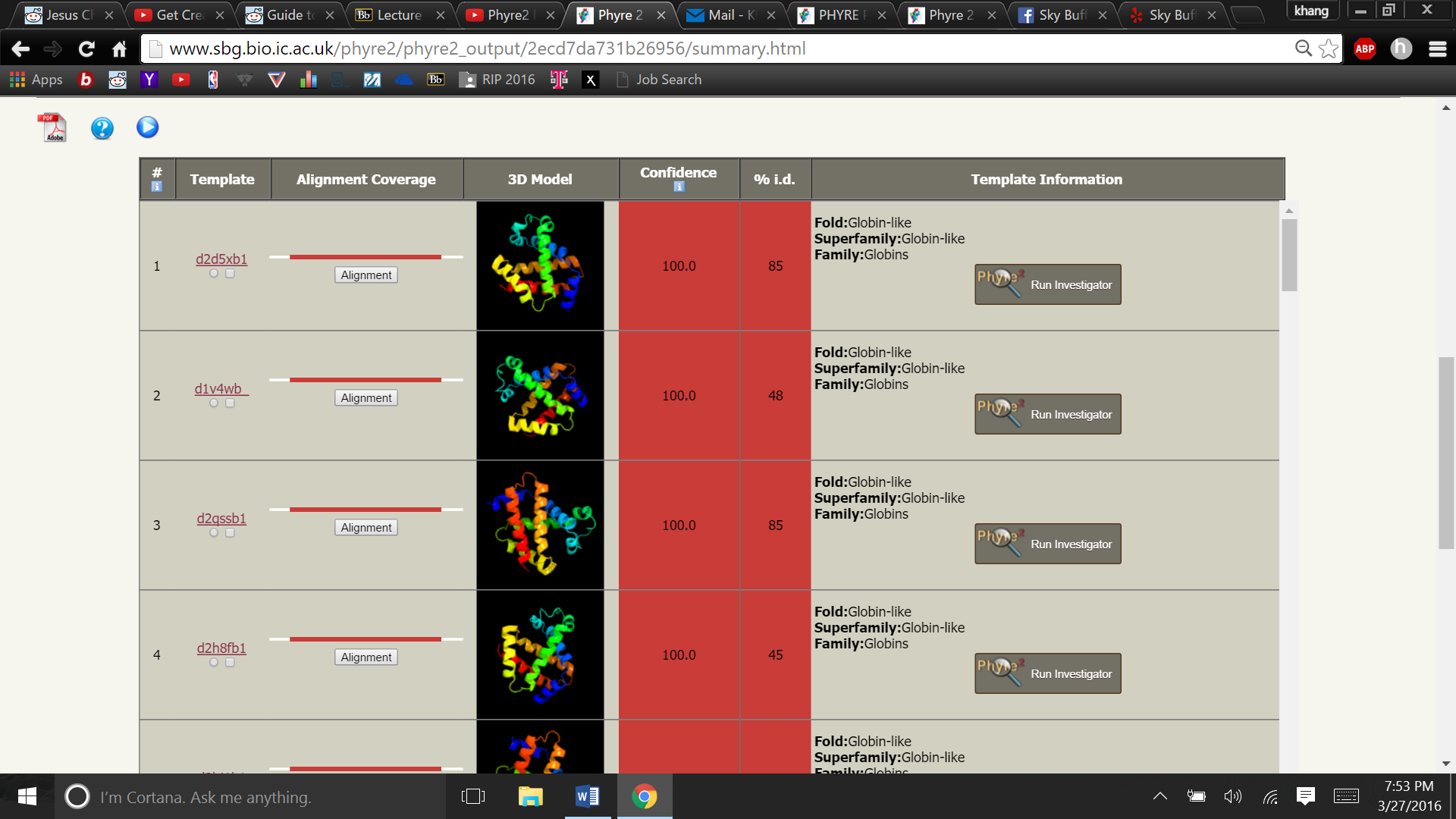
Go to the Phyre-2 server (don't use obsolete original Phyre) http://www.sbg.bio.ic.ac.uk/phyre2  
    - Paste your FASTA domain AA sequence in.  Enter your email address and hit Phyre Search.  
    - Leave the web page open.  It will automatically update the status during the modeling.  
    - When the job is done, SAVE EVERYTHING for your final project.  Be sure to save each of the PDF icons in the results web page.  
\*\*\* Questions.    
    1) Click on the 3D image to save the file of atomic coordinates.    
       - Using Protein Workshop, open that file  
       - Save the .jpg or .png. Again, if you cannot save the image with "File" -> "Save image...", then use print screen to create a snapshot of your video display. To invoke print screen on Windows use "control-printScreenButton" and on the Mac use "command-shift-3" or "command-shift-4".  
       - Put the .jpg or .png into your homework document. [5 points]



2) Rotate the image to a very different orientation, save that .jpg, and put it into your word processor file. [5 points]



3) For \*each\* of the top three templates used in building the homology model, answer: [each question carry 3 points].   
a) What was the template?  (For example, "period circadian protein homolog 2")



b) What was the confidence score of that template?

The confidence score of Template 1 (d2d5xb1) globin-like protein had a confidence score of 100.0.

The confidence score of Template 2 (d1v4wb) globin-like protein had a confidence score of 100.0.

The confidence score of Template 3 (d2gssb1) globin-like protein had a confidence score of 100.0.

c) What was the % identity of that template?

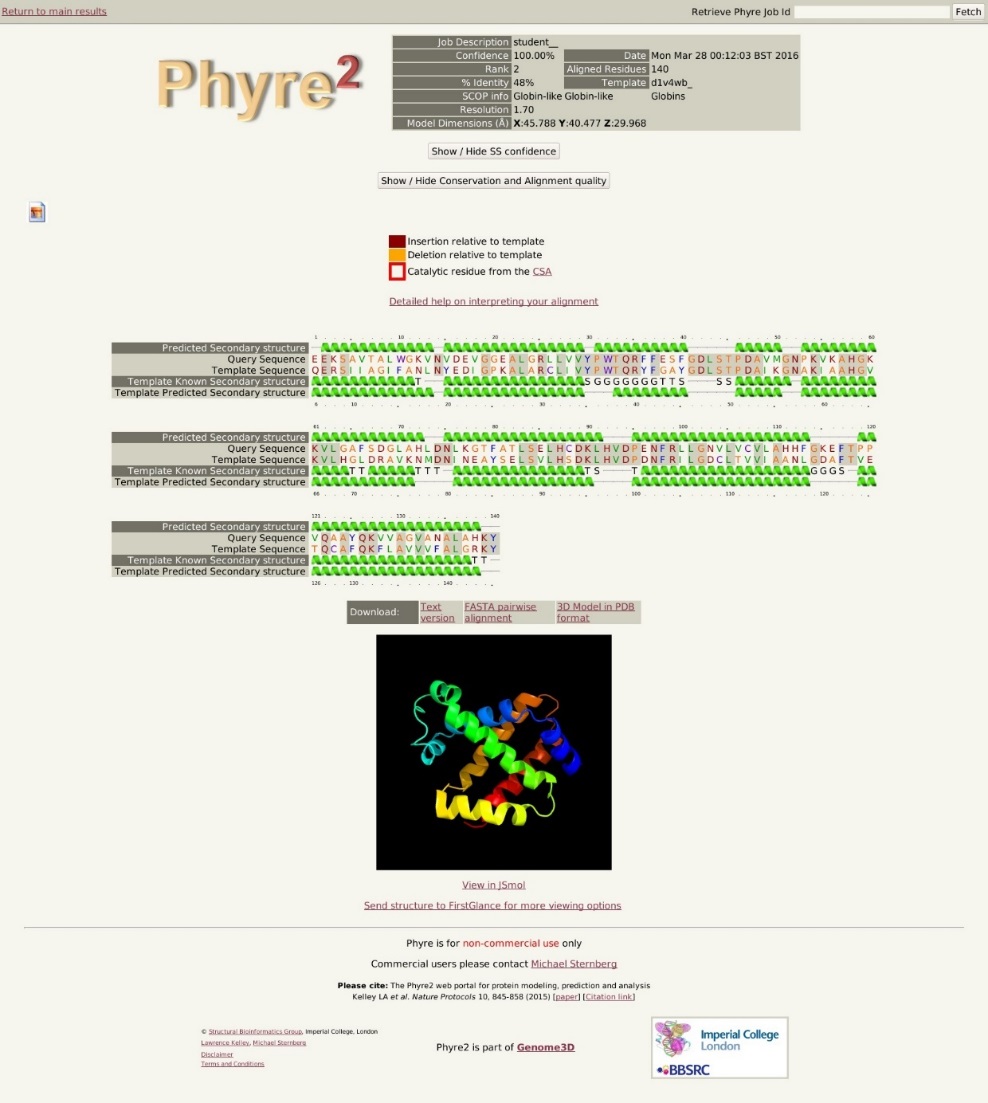
The % identity of Template 1 (d2d5xb1) globin-like protein had % identity of 85.

The % identity of Template 2 (d1v4wb) globin-like protein had % identity of 48.

The confidence score of Template 3 (d2gssb1) globin-like protein had % identity of 85.

4) For the second template used in the modeling, click Alignment. [5 points]

a) On the left of the page, click the "Generate JPEG image" button, save that .jpg, and paste it into your homework document. If the "Generate JPEG image" button does not work for you, then use print screen to create a snapshot of your video display. To invoke print screen on Windows use "control-printScreenButton" and on the Mac use "command-shift-3" or "command-shift-4".  
b) Explain and interpret the differences between your sequence and the template.



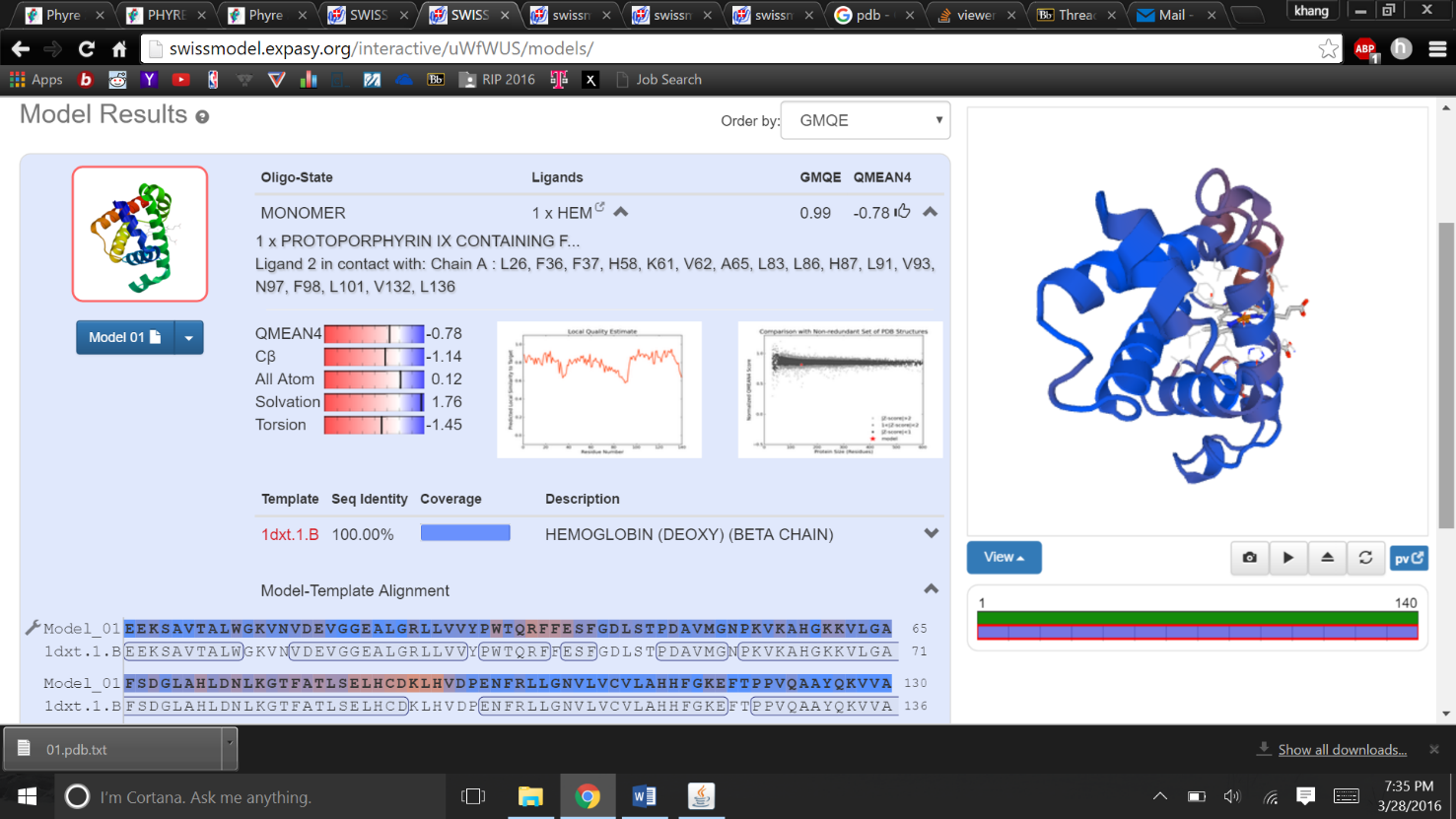
When comparing the second template to the first, I noticed there were a lot of amino acid differences between the two. This would make sense considering that Phyre2 predicted that there was going to be a 48% aa identity. When analyzing the secondary structure, I also noticed that the alpha helices largely consistent between the template sequence and the query sequence. However, it was not always perfect. There were instances of false negatives within the template predicated secondary structure (of alpha helices) that were present within the predicated secondary structure of the query sequence (query sequence 31-32, 75, and 91).

=====================================================================  
PART III:  Using SWISS-MODEL for homology modeling   
=====================================================================  
Use SWISS-MODEL to perform Homology modeling on our conserved domain in Part 1.

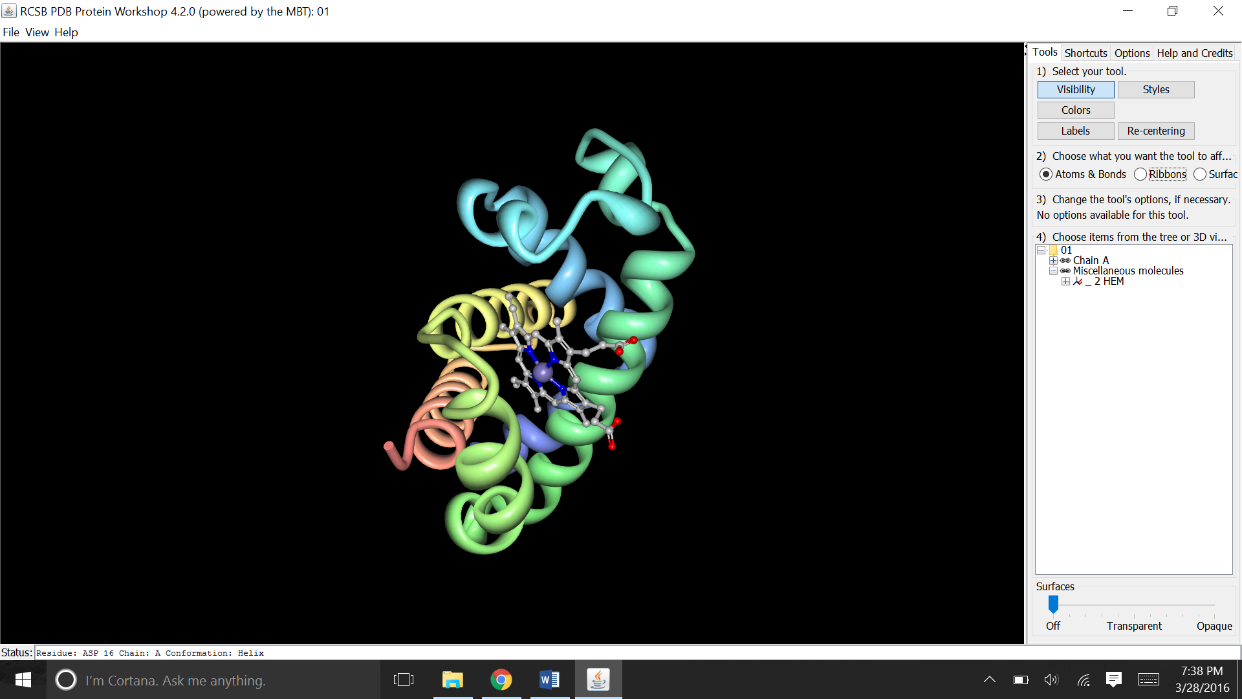
1) Watch a video about SWISS-MODEL at https://www.youtube.com/watch?v=QB7W6s7n0bU

2) Go to the SWISS-MODEL URL at http://www.swissmodel.expasy.org/  
3) Create an account so that you can build models and keep track of your project.

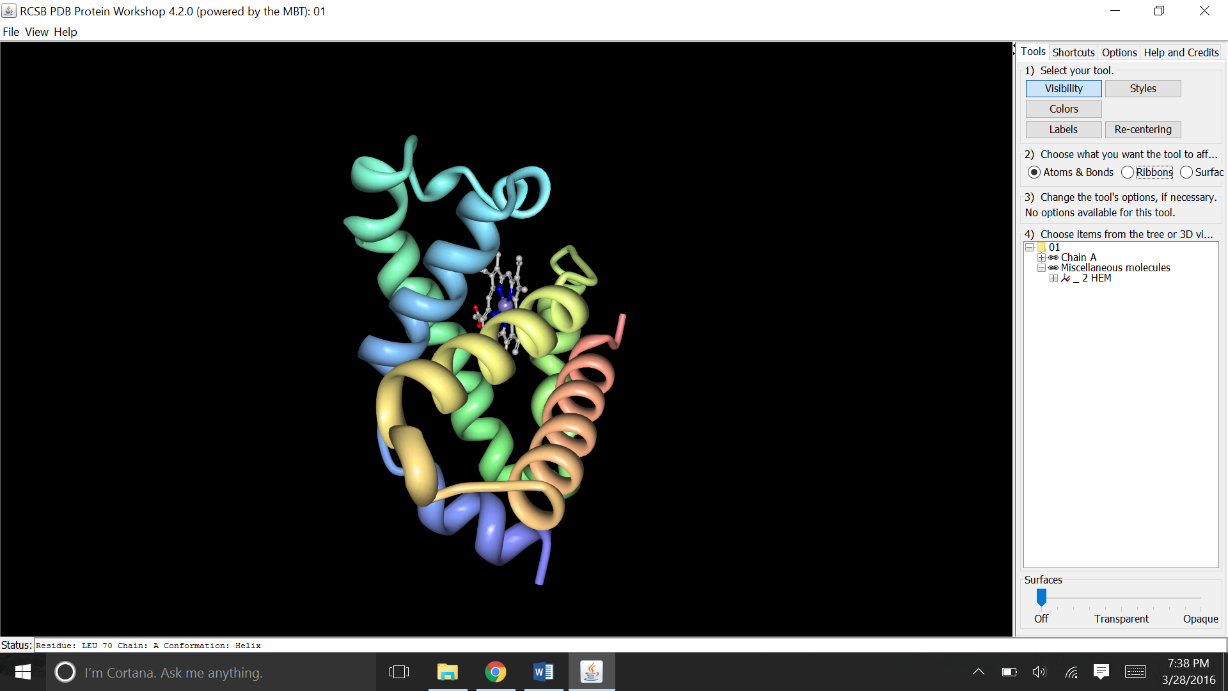
4) Paste your domain's FASTA sequence and perform a "template search". When the search is finished, you should pick the template sequence that has the highest percent identity and best resolution. [5 points]



**This was the template that I chose that has the highest percent identity and best resolution.**  
5) With your chosen template sequence, now "build models".  
6) On the blue MODEL 01, click the down arrow.  Save the PDB file to your computer.  
     - Using Protein Workshop, open that file  
     - Save the .jpg or .png  
     - Put the .jpg or .png into your homework document.  
7) Rotate the image to a very different orientation, save that .jpg or .png, and put it into your word processor file. [10 points in total]







8) On the blue MODEL 01, click Model Report to get detailed information about the model.  
Choose three different templates with different percent identity to our query domain. For \*each\* of the top three \*different\* templates used in building the homology model, answer: [each question carries 8 points. So in total 24 points]

a) What was the template?  (For example, "period circadian protein homolog 2")  
b) What was the similarity score and coverage of that template?  
c) How was that template built (e.g., x-ray, NMR, etc), and what is its resolution?

The generated box below contains all the answers for the questions.

