Instructions of how to complete your FINAL bioinformatics project. Your report, as pdf is to be emailed to jcorkill@ewu.edu by Sunday 16th June 2019 at 3:00pm. Late penalties apply. **Edit out any instructional details in orange italics. As always, style counts!**

- To start the final report, open your report for part 1 and first edit the title so it’s **Enzyme Bio-Informatics project** (and not INTRO Pt1 and …).

Then, deal with any comments added to your report by the instructor for your part I (further details, reaction diagram (use Google images if it’s not at PDBSum…most are!), cleaning up Ligand section by eliminating data on non required ligands, such as SO₄, EDO, anions, small organic compounds that co-crystalised with the protein.

Now open the MS Word template for Part 2, select it all, copy it and then paste it at the end of your part 1 right after the homology and **Your turn(#5)** Summary:

(i.e on intro… handed back on 6/4) report. Now using the instructions for part 2, you are ready to collect more data on your enzyme.

26. Now you are ready to start so go to BRENDA and start completing the second half of the project.

26. **BRENDA** data for kinetics data (KM, turnover # etc)

   First you need you EC #: if you don’t know that go to [enzyme-database.org](http://enzyme-database.org) Or [http://www.chem.qmul.ac.uk/iubmb/](http://www.chem.qmul.ac.uk/iubmb/) and use the search engine with your enzyme’s name.

27. **BRENDA** Now go to [Brenda-enzymes.info](http://Brenda-enzymes.info) and first click on **Word maps** in lowerleft under **Visualisation** (see below) and enter your EC #. Copy and paste the Word Map into your report right underneath the PDB and EC boxes right at the beginning of your report. Reduce its size to about 7' across and 3' down.

   Now click on **Functional Parameters Statistics** in **Visualisation** box.

![Word Map and Functional Parameters Statistics](image)

Then at the next page, enter your EC # in the bottom box. The top box will start with KM_Value, then hit search and you will get a histogram with a range of reported KM values for your enzyme. C & P this image into your report i.e.

![KM Values for the enzymes](image)

list the most common range under the image. Note the x axis is logarithmic! so put, for example from \(10^3\) to \(10^1\) mM. Use the correct form of superscripts!
Now do the same for Range of $K_{\text{cat}}/K_M$ and $pI$ values, Turnover #, pH optimum and temperature optimum. If there’s no data, state that in your report.

Use proper style for indices: the image is from the Home:toolbar in MSWord. Obviously the one on the left is for superscripts! Cut this image (above) and these words out of the final report.

28. For data, http://www.kegg.jp/dbget-bin/www_bget?ec:x.x.x.x (ie your EC #) C&P the Class and Comment sections. This can also be accessed via PDBsum... there’s also a link in the reaction image [KEGG] see earlier

Then on this same page as the reaction page, click on the top Pathway (rxnxxxxx) C&P this into your report (shrink to fit width of page) You will see your enzyme in red. Highlight this if possible. Click back and now click on Orthology: at this page click on Disease. i.e Hxxxx . C&P the Name and Description sections into your report.

Now go to metacyc.org and enter your enzyme name in the search engine at the top right and hit Quick Search. Now scroll down to EC Numbers and click on yours. At the next page, cut and paste (1) the parent class line (that decided the EC#), (2) the systematic name and (3) the Summary.

29. Now search VVP, OMIM, Proteopedia, Google, or Wiki for diseases associated with your enzyme using the enzyme’s name or enzyme’s name disease or EC # and describe your findings on the report

Use this data in your Executive Summary at the end.

Now the best but trickiest part i.e the production of images of your enzyme’s overall and active site structures.

First one has to download the RasMol, program. There are PC and Mac versions. (see #14)

32. At PDBSum/Protein Chain page using your chosen colored ball (#11) and click on the RasMol icon under the image on left hand side. (Note: Once you have RasMol, you can click on this icon at many other place on several of these sites). There are actually two windows that you will use (one may be in the dock along the bottom of the screen and is tricky to ‘see’ the first time). The obvious one is the image of your protein and the other is the Command line programming window. Hey! you’re a biochemical programmer!

33. Use the Display menu and explore the options: use the mouse to rotate in real time you are using what used to be called a supercomputer!). Get three (but significantly different) ‘nice’ images (from different angles) and paste them into your report (or save them until later) (a) Display:CartoonColours (this is a British software!) “Group- this shows the N terminus as blue and the follows the colours of the spectrum through to the red C terminus) (b) Display:CartoonColours:Structure (this shows the helices in red and the sheets in yellow) (c) your choice of Display and Colours but different from the first two!

34. Now using Display:Cartoon& Colours: Monochrome, you will pick out and display the AA residues in the active site (and in some cases the ligand(s) [if your paper crystallized the enzyme with ligands and/or metals]

35. To do this you will need to know the data you collected earlier about the AAs in the active site or ‘ligand interaction’ (see # above) Make sure you use the data from the ligand and active site section of the Protein View Feature mentioned in #11 as well as the LigPlot (#18) image.
36. Ok for the first AA residue (say # 135), use the other window (Command Line) which should be blinking RasMol> type (don’t type type! Just) select 135 and hit enter. It will give you the # of atoms and the name of the AA. Now go type color red and hit enter. You see a red patch somewhere. You need to (a) center on the red patch and then (b) zoom in on this by using Shift left mouse button. Now go to the image window and use Display: Sticks. This will show the AA “R” group as bendy stick (Look at VVP Fig 11-25 for an example) You can use the other Display options (try spacefill or Ball and Stick) but leave it as sticks. Go to Setting and choose ‘pick label” and the AA reside will be labeled (hard to see though!) In your report, note down the residue #, name and color you gave it. Whew! Now repeat this operation by choosing another residue by using RasMol> new select # but this time use a different color (available are red, pink, orange, cyan, blue, green, magenta and ? don’t use yellow..that’s for the ligand later) The idea is to color 2-6 AA residues in the active site and the ligand if it is there and then rotate and zoom your image until you think you have the ‘best’ view of the active site. Save this and paste it into your report you can paste in several views if you wish: it is tricky finding just the right view) If some cases, you might have found the sequence of neighbouring AA residues, say, 128-135. You can enter into RasMol> select 128-135 and color it and ‘stick’ it as above. There are many other options; one to try is use Settings:Pick Distance and then click on two atoms in different AA. The distance is display both on the mage and in the Command line window. H=bonds are about 2 Angstrom (20pm).

37. To get the orientation of the ligand in the active site enter RasMol> select ligands (or ligand) then enter then a color say yellow enter then space fill. A bunch of yellow balls will appear. If you click on them, the Command Line will tell you what they are (in many cases ions but I found ADP for one so I put RasMol> select ADP enter , colored it (blue) and spacefilled it , and then really went overboard for K+ in cyan, Gln (Q) red, Ornithine (green). There were some Mn++ as well but not close to these. Use a different colour for the cations. The possibilities are

endless and time consuming.
Here’s another image showing a substrate (yellow) and 2 Zins (green) in the active site

The final **Your Turn:**

**Executive Summary.** Under your RasMol images in your report, you now need to write an account of what you learnt about (1) the role of your enzyme *in vivo*, i.e reaction, pathway, etc. & (2) structural features (i.e single/multiple sub units, mainly α or β secondary structures, motifs/domains, etc) & (3) the role and contents of the active site from all the other data and these images. The general idea here is to tie together the information you have (i) collected and (ii) visualized about the active site of your enzyme with respect to its biological function (remember Str->Act->Biol Function is the theme of 480)

A significant number of points for this project will depend on both the quality of your RasMol images & this summary...don’t scrimp on it nor ignore it! The general idea is to see if you can relate all this enzyme data to what you have learnt in this course about proteins but enzymes in particular. You should be pleased and proud of your time and effort put into preparing your report!

Good luck!