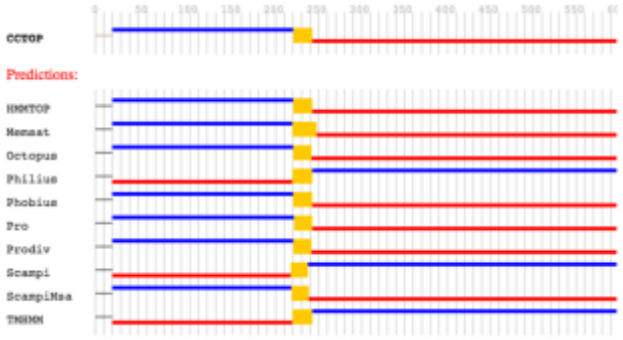
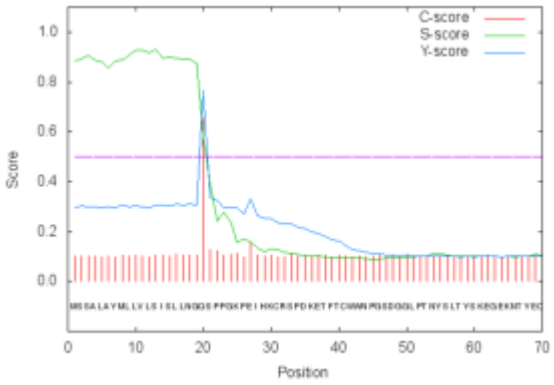
Intro to Bioinformatics/Spring 2017 - Individual Assignment #3
Julie Garcia - April 7, 2017

Using PHD, I found that my protein, Mus Musculus Prolactin Receptor, is made up of 17% alpha helices, in 11 distinct helices. Helices are located at positions 2, 127, 234, 284, 300, 424, 439, 464, 529, 551 and 576. The remaining protein contains 64% random coils, and 19% extended strands.

To find transmembrane sequences, I used TMHMM and CCTOP. The protein appears to have one transmembrane domain, with ~23 amino acids, between and including amino acids 231 and 253. Phobius reported amino acids 1 to 230 as extracellular and 254 to 608 as intracellular. TMHMM showed that the TM is a helix, which coincided with what PHD showed for the helix located at position 234. The image on the left below shows transmembrane location results.

I used SignalP, TargetP and Phobius to find a signal sequence between 1 and 19 aa, and cleavage site between position 19 and 20. The following sequence represents the signal sequence: MSSALAYMLLVLSISLLNGQ, which you can see starts with a charged amino acid, followed by a section of hydrophobic amino acids and ending with the cleavage site of a glycine-glutamine bond. The image on the right below shows SignalP results with clearly indicated signal sequence. The N-Region of the signal sequence is found between amino acids 1 and 3, the H-region between 4 and 14, and the C-region between 15 and 19. Combining all of the data, you can see that prolactin receptor is an integral membrane bound protein. ProtFun shows that the prolactin receptor is a transport and binding protein, that is classified as an enzyme, with a gene ontology class of receptor.





In addition to the transmembrane domain, using PROSITE, I found two Fibronectin type-III repeat domains between 22-122 and 124-224. Comparing these results with Phobius shows that these domains are in the extracellular region. Another pattern was found between 149 and 219, which PROSITE describes as a long hematopoietin receptor, which appears to connect to the transmembrane domain. InterPro describes this as a cytokine receptor that we know binds prolactin and is located on the outside of the cell.

GEO profiles showed data that was collected from various experiments involving the Prlr gene and tissues of the heart, prostate, skin, brain, kidneys, liver, muscle and stem cells. I found 14 matches for differential expression data. One entry showed the Prlr gene expression almost knocked down to zero when there was an HDAC3 deficiency in the heart. Another showed down regulation of the Prlr gene when IKK2ca was knocked out in prostate epithelium cells. I also searched for Prlr using the SAGE tool, in which I found links to all the coding and non-coding transcript variants of my gene. I found links to gene expression data in a SAGE Expression matrix, Digital Northern Results and Virtual Northern Results, or the Cancer Genome Anatomy Project. No Microarray data was found. The SAGE Expression data showed tags for expression during development in various cell types. I found that the Prlr gene is expressed in the brain during most stages of development, and highly expressed in heart during early development and muscle and skin during later development.

References:

https://npsa-prabi.ibcp.fr/cgi-bin/secpred_phd.pl
http://www.sacs.ucsf.edu/cgi-bin/hmmtop.py
http://cctop.enzim.ttk.mta.hu/?_=/jobs/result/bdd6efd2d4fe0e3e21cc8f9668db35e3
http://www.cbs.dtu.dk/cgi-bin/webface2.fcgi?jobid=58E7F1A900000DD769AD13EF&wait=20
http://phobius.sbc.su.se/cgi-bin/predict.pl
http://prosite.expasy.org/cgi-bin/prosite/PSScan.cgi
https://www.ebi.ac.uk/interpro/entry/IPR003528