**Database & Bioinformatic Tools for Beginners - Exercise PobA from *P. putida***

Your tasks:

1. What kind of enzyme is PobA? Look up PobA from *Pseudomonas putida* KT2440 in Uniprot (uniprot.org) and in BRENDA (brenda-enzymes.org). Below, write a brief description of the enzyme and show the reaction it performs:
2. <https://www.youtube.com/watch?v=rDhElW5ox6w>. Find the FASTA sequence of the PobA protein from *P. putida* KT2440 using the NCBI Protein database. The accession number should be AAN69138.1. Paste it below and save it as a text file on your computer.
3. <https://www.youtube.com/watch?v=q4yHvpfsRpM>. Run a protein-protein BLAST (blastp) on the PobA FASTA sequence to find similar sequences. Limit your query to non-redundant protein sequences. Using the “Organism” field, limit your results to include eubacteria and exclude pseudomonas. Submit the BLAST (it will take a few min to run). Select 8 similar sequences and save them in FASTA format. Use the BLAST results to make a table below that includes:
   1. The source (organism) of the similar sequence
   2. Description of the similar sequence
   3. The degree of similarity (% identity, query cover)
4. Look for the presence of conserved domains in PobA. Look up the protein again in the NCBI Protein database. On the right hand side of the page for AAN69138.1, you will see a link to “Identify Conserved Domains”. Describe the results below:
5. Perform multiple sequence alignments with Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) to compare PobA with each of the 8 similar sequences you selected in step #3. You’ll paste 6 sequences into the query box and then submit. Click “show colors” to highlight amino acid types in your results, then paste the alignments below:
6. Construct a phylogenetic tree with your sequence alignments from Clustal Omega (use the “Phylogenetic Tree” tab) and paste it below:
7. Predict the secondary structure of PobA using one of the tools from ExPASy (<https://www.expasy.org/resources/search/keywords:secondary%20structure%20prediction>). The input for these tools is usually your FASTA sequence file. Describe your results below. What are the limitations of secondary structure prediction software?
8. <https://www.youtube.com/watch?v=3mphYAgd60Q>. Retrieve the 3D structure of PobA from the PDB. What is the PDB code for this protein? You can use the Jmol applet on the PDB webpage to view the structure interactively (click 3D View > Structure under the image of the protein). Paste an image of the structure below. How many solvent (water) atoms are in the structure (hint: look at the PDB format file using “Display Files” tab in the upper right hand corner)?