

Bioinformatics: Life Sciences on Your Computer

Week 3 Peer Assignment

Due July 1, 2014

1. Go to the RCSB PDB database (<http://www.pdb.org>) and search for an author named Y.S. Moroz.
 - a. How many hits? _____
 - b. Find the most recent entry. What is its PDB number?
 - c. On what date was the structure released? _____
 - d. From what organism is polymer 1? _____
 - e. From the Molecular Description graphic, how many alpha helices?
 - f. From the record title, there is a mutation at amino acid 144 of calmodulin. What are the original and mutated amino acids at that position (use amino acids' full names, not one-letter abbreviations)?

2. Use the following protein. Run the [compute pI/Mw](#) tool.

```
> DNA helicase [Rhodococcus sp. BCP1]
MTVTDQTDLDRLSADCATPDADTETLLLCALLFSTRAAADIVDLLTTDDFHQPLHAELFDAIAALMS
TGQPHDPAMVLAHLERHGLAGHHGRRLAYALTATTAGADPTAAPYYAHAVISAAAYRRSFRTAGAAITE
AAETLPEHDLFDHMVAIGRTQRTATRRLEHVRARIGATPTVGNIK
```

What is this protein's molecular weight and isoelectric point?

3. Use the following protein and scan Prosite (<http://prosite.expasy.org/>). Uncheck the box by "Exclude motifs with a high probability..."

```
> adhesin [Helicobacter pylori]
MIKKNRTLFLSLALCASISYAEDDGGFFTVGYQLGQVMQDVQNPGGAKSDELARELNADVNTNNILNNNTG
GNIAGALSNAFSQYLYSLLGAYPTKLNQSDVSANALLSGAVSGTCAAAGTAGGTSLSNTQSTCTVAGYYW
LPSLTDRILSTIGSQTYNYGTNTNFPNMQQQLTYLNAGNVFFNAMNKALENKNGTSSASGTSGATGSDGQT
YSTQAIQYLQGGQNILNNAANLLKQDELLLEAFNSAVAANIGNKEFNAAFTGLVQGIIDQSQAVYNELT
KNTISGSASVISAGINSNANAVQGRASQLPNALYNAQVTLDKINALNNQVRSMPLYLPQFRAGNSRSTNIL
NGFYTKIGYKQFFGKKRNIGLRYYGFFSYNGASVGFRSTQNNVGLYTYGVGTDVLYNIFSRSYQNRSDVM
GFFSGIQLAGETFQSTLRDDPNVKLHGKINNTHFQFLDFGMRMNFGLDGLKSNRHNQHTVEFGVVVPTI
YNTYYKSAGTTVKYFRPYSVYWSYGYSF
```

- a. Find all short (high probability of occurrence) patterns found in the sequence. Look for either the zinc finger or leucine zipper motif? Which of the two is found and at what numerical location in the protein?
- b. At how many different location ranges in this protein are matches to the Protein kinase C phosphorylation site pattern. List their numerical locations.

4. Use the following sequence to run an HNN secondary structure prediction (http://npsa-pbil.ibcp.fr/cgi-bin/secpred_hnn.pl).

```
MVSVNMMTPSEMQRVSRARDLRLELNLSQQTLEKSGVSYGSLKKFEQTGQISLESLLKLAVILGCMDFKALFALKSAEEALSLDELIDDGKRKRGRK
```

- a. How many alpha helices are predicted by HNN?
- b. What percentage of the amino acids are predicted to be alpha-helical?
- c. Give the exact numerical range of the predicted alpha-helix that is closest to the N-terminus?