Practical course 3
Secondary structure prediction – conformational diseases

1. **Analysis of secondary structure prediction programs**

The programs that will be used to predict the secondary structure are GOR4, HNN and Sopma. They are provided on the website "http://npsa-pbil.ibcp.fr/" ("secondary structure prediction" section). Describe briefly the approach used by each program.

2. **Comparison of the performances of secondary structure prediction programs**

A] Search for the human thymidylate kinase (PDB code 1E2F) in the protein databank (www.rcsb.org). Search for the secondary structure limits of the protein in PDBSum (http://www.ebi.ac.uk/pdbsum; "Protein" tab, "7 strands" and "11 helices" in the menu on the left). This secondary structure assignation is provided in the file 1e2f.xls (on the Virtual University). A secondary structure assignation is also provided in the PDB file, sections "HELIX" and "SHEET" (from the PDB website, menu "Display file", "PDB file", find the "HELIX" and "SHEET" sections). Compare both secondary structure assignations (from PDBSum and from the PDB file). Why are they slightly different?

B] Perform a secondary structure prediction of 1E2F by using the three programs GOR4, HNN and Sopma. Show in the .xls file the results. Analyse and comment these results.

3. **Analysis of two sequences with particular properties**

A] Predict the secondary structure of the sequences inconnu1.fasta and inconnu2.fasta (available on the Virtual University), by using the 3 programs used in the previous sections. Record the predictions in the files inconnu1.xls et inconnu2.xls that already contain the secondary structure of the experimental structure of these proteins. Use BLAST (http://www.ncbi.nlm.nih.gov/blast/ ; "Protein Blast") to identify the proteins corresponding to these sequences. Analyse the results, make assumptions and interpret the results.

4. **Propensity of amino acids to adopt a given local structure**

A] Analyse the propensity of the 20 amino acids to adopt a local structure. Use for that purpose the Fugue program (http://babylone.ulb.ac.be/Prelude_and_Fugue/Fugue/index.php). Submit to this program 20 peptides composed by 20 amino acids of the same type (AAAAAAAAAAAAAAAAAAAA, …). Read carefully the description of Fugue and how to

interpret its results (http://babylone.ulb.ac.be/Prelude_and_Fugue/). You can divide this work in groups.

B] Use the results obtained in section *A* to test several mutations (single-site and multiple) in the sequence inconnu2.fasta, with the aim of increasing the propensity of the second helix to be predicted as an helix. Use the HNN program to make the predictions on the mutated sequences. Analyse the results. Is it possible to increase the propensity to adopt an helix by mutating some sequence positions?