Essential Bioinformatics for Biochemists

We are going to explore the bioinformatics resourceses that are available through WWW. If you are a full time student/scientist, you can access a well-organized package of GCG, which you need an account from the central facility, e.g. NHRI. The following sites are the major bioinformatic sites which can be good starting points. Please spend some time to get familiar with the tools inside these magor sites.

1.European Bioinformatics Institute http://www.ebi.ac.uk/

  A key site in Europe, near Cambridge UK. Databases, analysis tools and sequence submission.

2. ExPASy Swiss Bioinformatics Institute: http://www.expasy.ch/ or http://expasy.nhri.org.tw/

SWISS-PROT, PROSITE, 2D-PAGE, Proteomics Tools etc. Key site for protein sequence/structure researchers

3. NCBI http://www.ncbi.nlm.nih.gov

US equivalent of EBI. Home of Genbank, PubMed (http://www.ncbi.nlm.nih.gov/PubMed/), Entrez etc

Among the various bioinformatics providers, you will be able to find tools to analyze a protein in terms of primary structure, secondary structure and tertiary structure.

1. Primary structure of protein

You can identify a protein from its N-terminal sequence information derived from your experiments. Using FASTA or BLAST, its neighbor can be found from the database. Most of such jobs can be done in the database of SwissProt. If it is an enzyme, it's good to explore the ENZYME (http://expasy.nhri.org.tw/enzyme/)

2. Secondary structure

Most of the secondary structure prediction methods rely on multiple sequence alignment of homologous proteins, e.g. PhD program (http://www.embl-heidelberg.de/predictprotein/predictprotein.html)

3. Tertiary structure

For those proteins that their 3D structures are available, you can view and analyze the structural detail by display software like RasMol (http://www.bernstein-plus-sons.com/software/rasmol/ or http://www.umass.edu/microbio/rasmol/). Another good molecular graphic software that can run on a PC is Swiss-Pdb viewer (http://expasy.nhri.org.tw/spdbv/). The principal three-dimensional structure database is PDB (protein data bank, can be accessed through its Taiwan mirror http://pdb.life.nthu.edu.tw/). An excellent structure classification can be found in SCOP (http://scop.mrc-lmb.cam.ac.uk/scop/)