Structural Bioinformatics (TDA506)

Course Organiser: Graham Kemp

http://www.cse.chalmers.se/~kemp/teaching/TDA506/

Motivation

- *Structural biology* gives an understanding of biological function at the molecular level.
- These functions are ultimately due to interactions between molecules.
- Ideally, we want experimentally determined 3D structures of molecules and complexes.
- Sometimes we have to rely on computer models of molecules and their interactions.
- *Structural bioinformatics!*

What is Biology?

Ecosystem	Rain forest, desert, fresh water lake, digestive tract of animal
Community	All species in an ecosystem
Population	All individuals of a single species
Organism	One single individual
Organ System	A specialised functional system of an organism, e.g. nervous system or immune system
Organ	A specialised structural system of an organism, e.g. brain or kidney
Tissue	A specialised substructure of an organ, e.g. nervous tissue, smooth muscle
Cell	A single cell, e.g. neuron, skin cell, stem cell, bacteria
Molecule	e.g. protein, DNA, RNA, sugar, fatty acid, metabolites, pharmaceutical drugs

Structural bioinformatics (TDA506)

In this course we consider "structural bioinformatics" to be the development and application of computational methods (i) to analyse and predict the conformations of biological macromolecules and (ii) to study relationships between macromolecular structure and function. Protein molecules will be in focus, but other biological molecules will also be studied.

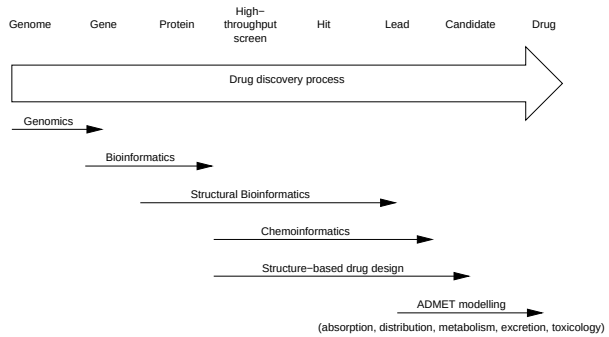
Aims

- to present some of the computational challenges in structural biology;
- to describe computational methods for analysing and predicting macromolecular conformations and interactions;
- to give practice in programming techniques for structural bioinformatics.
- to give practice in the use of molecular graphics and modelling software;
- to emphasise the relationship between macromolecular shape and function.

Some challenges in structural bioinformatics

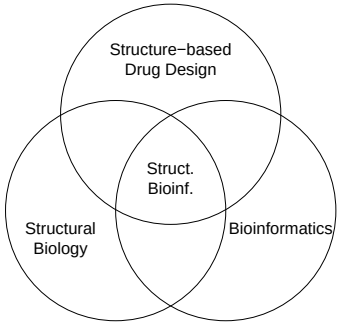
- Given the sequence of a protein, what is its three-dimensional structure?
- Given the three-dimensional structures of two macromolecules, will they associate with one another, and what will be the docking orientation?
- Given the three-dimensional structure of a macromolecule, what can be inferred about its biological function?
- Given the three-dimensional structures of a set of proteins, what can be inferred about their evolutionary relationship?
- Given the three-dimensional structure of a macromolecule, can a molecule be designed that will affect its function?

Structral bioinformatics in drug discovery



(Adapted from Fauman, Hopkins and Groom, 2003)

Structral bioinformatics in drug discovery



(Adapted from Fauman, Hopkins and Groom, 2003)

Structural bioinformatics (TDA506)

Content

three-dimensional structures of biological macromolecules; contact maps and distance maps; domain assignment; homogeneous transformation matrices; structure superposition; structure comparison; comparative protein modelling; protein fold recognition; Monte Carlo methods and simulated annealing; ab initio protein structure prediction; protein shape representation; protein-ligand interactions and applications in drug design; conformational analysis; protein-protein docking; modelling transmembrane proteins, carbohydrates and RNA; experimental protein structure determination using nuclear magnetic resonance (NMR) and X-ray crystallography; applications of structural bioinformatics.

Levels of protein structure

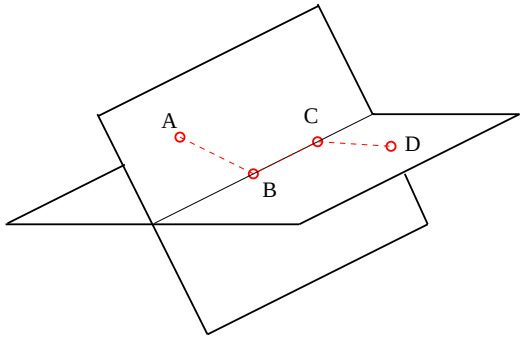
Primary structure
- amino acid sequence

Secondary structure
- assignment of helices and strands

Tertiary structure
- the 3D structure
- assembly and interaction of helices and sheets

Quaternary structure
- assembly of subunits

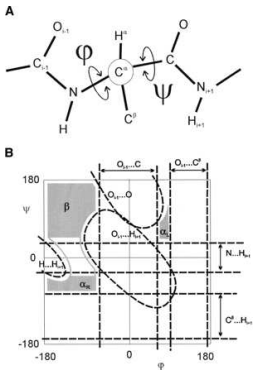
Dihedral angle



Protein Data Bank entry (extract)

```
COMPND    TRIOSE PHOSPHATE ISOMERASE (E.C.5.3.1.1)
SOURCE    CHICKEN (GALLUS GALLUS) BREAST MUSCLE
AUTHOR    D.W.BANNER,A.C.BLOOMER,G.A.PETSKO,D.C.PHILLIPS,
AUTHOR   2 I.A.WILSON
  :
JRNL        AUTH   D.W.BANNER,A.C.BLOOMER,G.A.PETSKO,D.C.PHILLIPS,
JRNL        AUTH 2 I.A.WILSON
JRNL        TITL   ATOMIC COORDINATES FOR TRIOSE PHOSPHATE ISOMERASE
JRNL        TITL 2 FROM CHICKEN MUSCLE
JRNL        REF    BIOCHEM.BIOPHYS.RES.COMM.     V.  72   146 1976
JRNL        REFN   ASTM BBRCA9  US ISSN 0006-291X                146
  :
REMARK   2 RESOLUTION. 2.5 ANGSTROMS.
  :
SEQRES   1 A  247  ALA PRO ARG LYS PHE PHE VAL GLY GLY ASN TRP LYS MET
SEQRES   2 A  247  ASN GLY LYS ARG LYS SER LEU GLY GLU LEU ILE HIS THR
  :
ATOM      1  N   ALA A   1      43.240  11.990  -6.915  1.00  0.00
ATOM      2  CA  ALA A   1      43.888  10.862  -6.231  1.00  0.00
ATOM      3  C   ALA A   1      44.791  11.378  -5.094  1.00  0.00
ATOM      4  O   ALA A   1      44.633  10.992  -3.937  1.00  0.00
ATOM      5  CB  ALA A   1      44.722  10.051  -7.240  1.00  0.00
ATOM      6  N   PRO A   2      45.714  12.244  -5.497  1.00  0.00
ATOM      7  CA  PRO A   2      46.689  12.815  -4.561  1.00  0.00
ATOM      8  C   PRO A   2      46.042  13.601  -3.411  1.00  0.00
ATOM      9  O   PRO A   2      46.030  13.141  -2.267  1.00  0.00
  :
```


Ramachandran steric map



[Ho, K.H., Thomas, A. and Brasseur, R., Protein Science, 2003, 12:2508-2522]