Introduction to Structural Bioinformatics

Lecture 1 — Aims

- To introduce the basic principles of protein conformation.

Lecture 1 — Objectives

After this lecture you will:

- be familiar with the basic features of protein conformation, and the abbreviations and symbols used in describing these;
- be aware of the twenty amino acid residues that are commonly found in proteins, and some of their properties;
- know about different levels of protein structure;
- understand how distances and angles can be calculated;
- be able to recognise common protein secondary structure elements and understand how these can be recognised automatically.

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

Some challenges in structural bioinformatics

The protein folding problem

- given the sequence, what is the structure?

The docking problem

- given two structures, will they associate?
- what is the docking orientation?

Predicting function from structure

Designing new functionalities

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
  :
```

DSSP bridges

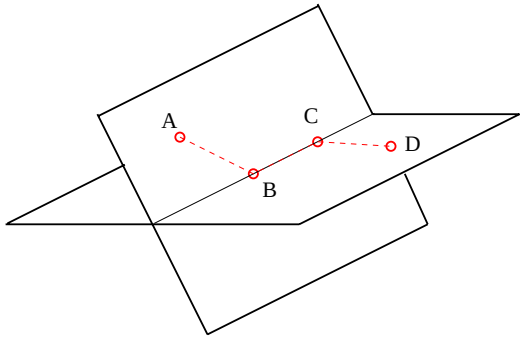
Antiparallel bridge:

[hbond(i,j) and hbond(j,i)]
or
[hbond(i-1,j+1) and hbond(j-1,i+1)]

Parallel bridge:

[hbond(i-1,j) and hbond(j,i+1)]
or
hbond(j-1,i) and hbond(i,j+1)]

Dihedral angle



DSSP summary codes

H 4-helix (α-helix)
B residue in isolated β-bridge
E extended strand, participates in β-ladder
G 3-helix
I 5-helix
T H-bonded turn
S bend

Crambin (1CRN)

```
TTCCPSIVARSNFNVCRLPGTPEAICATYTGCIIIPGATCPGDYAN
 EE SSHHHHHHHHHHHHTT  HHHHHHHHS EE SSS   TTS
```