

Structural clues from multiple sequence alignments

- Residues at highly conserved positions often have important functional or structural roles.
- Insertions and deletions can be accommodated most easily in surface loops.
- Conserved patterns of hydrophobic residues can suggest secondary structure.
- The root mean square deviation between pairs of homologous proteins generally increases as the percent residue identity decreases.

Protein stability

- good stereochemistry; no steric clashes;
- buried charged atoms must be paired;
- enough hydrophobic surface must be buried, and the interior must be sufficiently densely packed, to provide thermodynamic stability.

Modular proteins

- multi-domain proteins, often with many copies of related domains;
- domains recur in many proteins in different structural contexts.

Heuristics for manual secondary structure prediction

- Many α -helices are amphipathic. Conserved hydrophobic residues at positions i , $i+3$, $i+4$, $i+7$, etc. are highly indicative of an α -helix.
- Half-buried strands will tend to have hydrophobic and hydrophilic residues at alternate positions.
- In proteins containing both α -helices and strands the strands are often completely buried and tend to contain only hydrophobic residues.

For more details and references, see:

<http://www.bmm.icnet.uk/people/rob/CCP11BBS/secstrucpred.html>

“HSSP-curve”

- Shows the length-dependent threshold for significant sequence identity.
- Proposed by Sander and Schneider (1991) and revised by Rost (1999).
- Above the curve, identifying true positives is easy.
- Just below the curve, the number of false positives rises rapidly; distinguishing between true and false positives in the “twilight zone” is difficult.

(HSSP stands for “Homology-derived Secondary Structure of Proteins”)