Fold recognition

The idea behind "threading":

Imagine a wire wound into the shape of a known protein's main chain "fold".

Imagine next that our new sequence is represented by beads that are "threaded", in order, onto the wire, and are pushed along the wire.

At each step, a score is calculated based on which residues are adjacent in space, which residues are buried, etc.

Repeat this process for each different known fold.

A high score indicates that the sequence is compatible with that fold.

Threading the peptide through the groove

```
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
...GACPKYVKQNTLKLATGMRNVPEKQTRGLFGA...
```


Approaches to fold recognition

Profiles

e.g. Bowie et al. (1991) Science, 253:164-170

Pairwise potentials

e.g. Sippl and Weitckus (1992) Proteins, 13:258-271
e.g. Jones et al. (1992) Nature, 358:86-89
e.g. Jones (1999) J. Mol. Biol., 287:797-815 — GenTHREADER

pairwise pseudo-energy terms

$$\Delta E_k^{ab} = RT \ln(1 + m_{ab}\sigma) - RT \ln\left(1 + m_{ab}\sigma \frac{f_k^{ab}(s)}{f_k(s)}\right)$$

solvation potentials

$$\Delta E_{solv}^{a}(r) = -RT \ln\left(\frac{f^{a}(r)}{f(r)}\right)$$

Scoring function

- pairwise frequencies
- steric overlap and quality of fit
- hydrogen bonds
- positive and negative charges
- buried hydrophobic side chains
- exposed hydrophilic side chains