**Homework 4**

Xiaotong Zuo

02/29/2016

1. The assumption is that the lowest energy decoys are most likely to be similar to the native structure. The amino acid sequence completely specifies its 3-D structure. Hence the native structure of the protein is likely to correspond to the global energy minima.

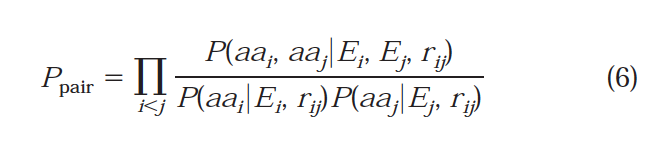
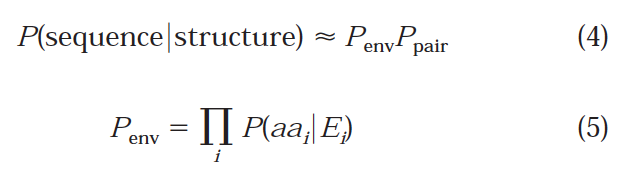
I think this assumption is not always correct. There are other factors which can influence the native structure besides its sequence. For example, the solvent environment of the protein, and the interaction among proteins will influence the protein structure. In general, the secondary structure of protein is easier to be predicted compared with 3D structure.

1. The failure of the ‘1r69’ and ‘2cro’ modeling is due to the inaccuracies in the score function. These two structures are both alpha helices, and fixing secondary structure provides stronger topological constraints for beta-sheets than for alpha-helices. Especially for ‘2cro’, we can find that the native structure has a very high energy compared with the low energy decoys with high rmsd values.

Also, as the paper mentioned, there are 30000 decoys generated for each protein. And for each decoy, there are more than 10000 moves. I think the samplings are sufficient enough for de novo modeling.

1. Ppair: physically refers to electrostatic interactions and disulfide bonds.

Penv: physically refers to the partitioning of hydrophobic residues to the interior and polar residues to the surface.



Where Ei is defined solely in terms of residue burial.

Reference: Simons, Kim T., et al. "Improved recognition of native‐like protein structures using a combination of sequence‐dependent and sequence‐independent features of proteins." *Proteins: Structure, Function, and Bioinformatics* 34.1 (1999): 82-95.