Self-consistent Estimation of Inter-residue Protein Contact Energies Based on an Equilibrium Mixture Approximation of Residues

S. Miyazawa[1], R. L. Jernigan[2]
[1]Gunma Univ., Fac. Tech., [2]NIH, NCI, LECB
[1]miyazawa@smlab.sci.gunma-u.ac.jp, [2]jernigar@exchange.nih.gov

Pairwise contact energies for 20 types of residues are estimated self-consistently from the actual observed frequencies of contacts with regression coefficients that are obtained by comparing "input" and predicted values with the Bethe approximation for the equilibrium mixtures of residues interacting. This is premised on the fact that correlations between the "input" and the predicted values are sufficiently high although the regression coefficients themselves can depend to some extent on protein structures as well as interaction strengths. Residue coordination numbers are optimized to obtain the best correlation between "input" and predicted values for the partition energies. The contact energies self-consistently estimated this way indicate that the partition energies predicted with the Bethe approximation should be reduced by a factor of about 0.3 and the intrinsic pairwise energies by a factor of about 0.6. The observed distribution of contacts can be approximated with a small relative error of only about 0.08 as an equilibrium mixture of residues, if many proteins were employed to collect more than 20,000 contacts. Including repulsive packing interactions and secondary structure interactions further reduces the relative errors. These new contact energies are demonstrated by threading to have improved their ability to discriminate native structures from other non-native folds.

Reference: *Proteins*, 34:49-68 (1999)