How effective for fold recognition are relative orientations between contacting residues in proteins?

Sanzo Miyazawa[1], Robert Jernigan[2]

[1]Graduate School of Engineering, Gunma University, 1-5-1 Tenjin, Kiryu, Gunma, 376-8515, Japan, [2]Laurence H. Baker Center for Bioinformatics and Biological Statistics, Iowa State University, 112 Office and Lab. Bldg., Ames, IA, 50011-3020

We estimate the statistical distribution of relative orientations between contacting residues from a database of protein structures and evaluate the potential of mean force for relative orientations between contacting residues. Polar angles and Euler angles are used to specify two degrees of directional freedom and three degrees of rotational freedom for the orientation of one residue relative to another in contacting residues, respectively. A local coordinate system affixed to each residue based only on main chain atoms is defined for fold recognition. To evaluate the fully-anisotropic distributions of relative orientations as a function of polar and Euler angles, we use a method in which the observed distribution is represented as a sum of δ functions each of which represents the observed orientation of a contacting residue, and is evaluated as a series expansion of spherical harmonics functions. The sample size limits the frequencies of modes whose expansion coefficients can be reliably estimated. High frequency modes are statistically less reliable than low frequency modes. Each expansion coefficient is separately corrected for the sample size according to suggestions from a Bayesian statistical analysis. As a result, many expansion terms can be utilized to evaluate orientational distributions. Also, unlike other orientational potentials, the uniform distribution is used for a reference distribution in evaluating a potential of mean force for each type of contacting residue pair from its orientational distribution, so that residue-residue orientations can be fully evaluated. It is shown by using decoy sets that the discrimination power of the orientational potential in fold recognition increases by taking account of the Euler angle dependencies and becomes comparable to that of a simple contact potential, and that the total energy potential taken as a simple sum of contact, orientation, and (ϕ, ψ) potentials performs well to identify the native folds.