

Supporting Text

In this supplement, we outline the variational model developed in refs. 1 and 2.

Variational Model. A configuration of the protein is modeled by the N position vectors of the α carbons of the polypeptide backbone. Partially ordered ensembles of polymer configurations are described by a reference Hamiltonian corresponding to a harmonic chain inhomogeneously constrained to the native structure $\{\mathbf{r}_i^N\}$

$$\beta\mathcal{H}_0 = \frac{3}{2a^2} \sum_{ij} \mathbf{r}_i \cdot \Gamma_{ij}^0 \cdot \mathbf{r}_j + \sum_i C_i (\mathbf{r}_i - \mathbf{r}_i^N)^2. \quad [1]$$

The first term enforces polymeric constraints, with $a = 3.8\text{\AA}$ and $\Gamma_{ij}^{(0)}$ determined by the correlations of a freely rotating chain (3). The values of the harmonic constraints, $\{C\}$, control the magnitude of the fluctuations of each monomer about the native structure (i.e., the temperature factors). Ensembles of partially ordered configurations are represented by monomer densities $n_i(\mathbf{r}) = \langle \delta(\mathbf{r} - \mathbf{r}_i) \rangle_0$ described as Gaussian distributions with variance $G_{ii} = \langle |\delta\mathbf{r}_i|^2 \rangle_0 / a^2$ about the mean position of the i th monomer, $\mathbf{s}_i = \sum_j G_{ij} C_j \mathbf{r}_i^N$. Here, the correlations $G_{ij} = \langle \delta\mathbf{r}_i \cdot \delta\mathbf{r}_j \rangle_0 / a^2$ depend on both the polymeric and structural constraints through $G_{ij}^{-1} = \Gamma_{ij} + C_i \delta_{ij}$.

The population of a partially ordered ensemble specified by the constraints $\{C\}$ is controlled by the free energy $F[\{C\}] = E[\{C\}] - TS[\{C\}]$. Here, T is the temperature, $S[\{C\}]$ is the entropy loss due to localizing the residues around mean positions, and $E[\{C\}]$ is the energy associated with the partially ordered ensemble. The values of the variational constraints are determined by the critical points in the free energy surface. For each local minimum or saddlepoint, there corresponds to a set of N variational constraints that solve $\partial_{C_i} F[\{C\}] = 0$. Transition state ensembles are identified as the saddlepoints of $F[\{C\}]$ that connect local minima (in a steepest descents sense). A folding route is the series of minimum-saddlepoint-minimum which connect the globule and native minima of $F[\{C\}]$.

Barrier Crossing Dynamics. The formation of local order along the folding route is characterized by the degree of localization about the native positions

$$\rho_i[\{C\}] = \left\langle \exp \left[-3\alpha^N (\mathbf{r}_i - \mathbf{r}_i^N)^2 / 2a^2 \right] \right\rangle_0. \quad [2]$$

We refer to ρ_i as the native density. The prefactor, in turn, is determined by the growth rate of $\rho_i(t)$ along the unstable mode of the free energy. In the formalism developed in ref. 2, the growth rate is

developed through the polymer dynamics of the constrained chain. Since the reference chain is harmonic, the correlation function between monomers i and j , $G_{ij}(t) = \langle \mathbf{r}_i(t) \cdot \mathbf{r}_j(0) \rangle_0$, can be expressed as a sum over normal (Rouse) modes

$$G_{ij}(t) = \sum_p \frac{Q_{ip} Q_{jp}}{\lambda_p} e^{-\sigma \lambda_p t} \quad [3]$$

where the coefficients Q_{ip} and relaxation rates λ_p are determined by

$$(\Gamma_{ij}^0 + C_i^* \delta_{ij}) Q_{jp} = \lambda_p Q_{ip}. \quad [4]$$

Here, the monomer relaxation rate $\sigma = 3D_0/a^2$ is set by the effective bond length a and monomer diffusion coefficient D_0 .

The effective diffusion matrix corresponding to the native density dynamics, $\mu_{ij}(t)$, can be defined through the Laplace transform of the correlation functions $C_{ij}(t) = \langle \rho_i(t) \rho_j(0) \rangle_0 - \langle \rho_i \rangle_0 \langle \rho_j \rangle_0$:

$$\hat{\mu}(\omega) = \mathcal{C}(0) \cdot \hat{\mathcal{C}}^{-1}(\omega) \cdot \mathcal{C}(0) - \omega \mathcal{C}(0). \quad [5]$$

Here, $\mathcal{C}(t)$ is determined by the polymer dynamics through the monomer correlations $G(t)$. Since $\mu_{ij}(t)$ depends on both the structure of the constrained ensemble as well as time, this formalism gives a microscopic realization of the effective diffusion coefficient used in more general formulations of landscape theory which includes the kinetics of trapped configurations (4,5).

Finally, the prefactor is given by $k_0 = |\omega|/2\pi$, where the growth rate ω is the negative eigenvalue of

$$\hat{\mu}_{ij}(|\omega|) \Gamma_{jk} u_k = -|\omega| u_i \quad [6]$$

where $\Gamma_{ij} = \partial^2 \beta F[\{C^*\}] / \partial \rho_i \partial \rho_j$ is the curvature of the free energy with respect to the native density evaluated at the saddlepoint.

1. Portman JJ, Takada S, Wolynes PG, (2001) *J Chem Phys* 114:5069–5081.
2. Portman JJ, Takada S, Wolynes PG, (2001) *J Chem Phys* 114:5082–5096.
3. Bixon M, Zwanzig R, (1978) *J Chem Phys* 68:1896–1902.
4. Bryngelson JD, Wolynes PG, (1989) *J Phys Chem* 93:6902–6915.
5. Lee C, Stell G, Wang J, (2003) *J Chem Phys* 118:959–968.