Coarse-grained (continuum, implicit solvent, C_α) models for proteins

J. D. Honeycutt and D. Thirumalai, "The nature of folded states of globular proteins," Biopolymers 32 (1992) 695.

T. Veitshans, D. Klimov, and D. Thirumalai, "Protein folding kinetics: timescales, pathways and energy landscapes in terms of sequence-dependent properties," Folding & Design 2 (1996)1.

3-letter C_α model: $B_9N_3(LB)_4N_3B_9N_3(LB)_5L$

B=hydrophobic

N=neutral

L=hydrophilic

$N_{sequences} = 3^{N_m} \sim 10^{22}$ Number of sequences for $N_m=46$

$N_p \sim \exp(aN_m) \sim 10^{19}$ Number of structures per sequence

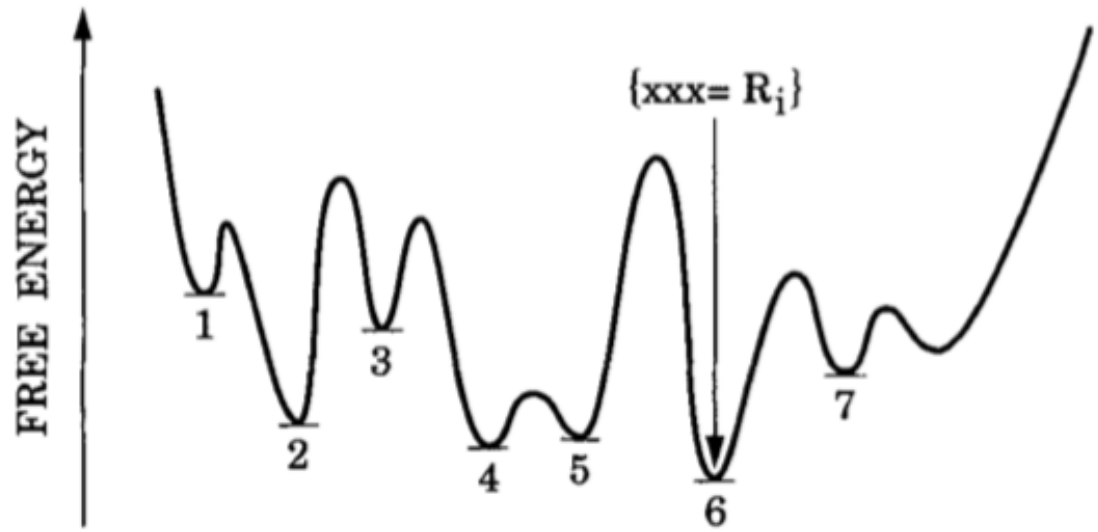


Figure 1. Schematic sketch of the free energy (or energy) profile corresponding to the folded region in the conformational space of a protein molecule. The sketch is used to illustrate the content of the metastability hypothesis. The various minima are separated by bottlenecks of differing amplitude such that the transition from one to another is unlikely to occur on the time scale of the folding process. The typical time scale for the folding process ranges from milliseconds to seconds. The particular minimum into which the protein falls depends on the initial condition. For example, in this figure the initial configurations of the protein marked by { xxx } map onto the minimum labeled 6.

and dynamics

different mapping?

Molecular Dynamics: Equations of Motion

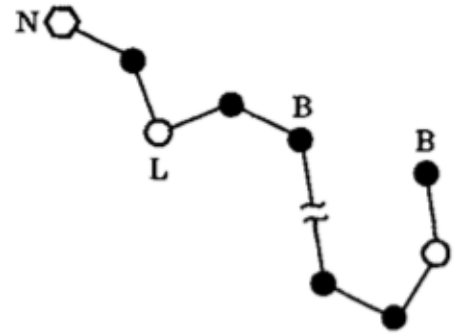
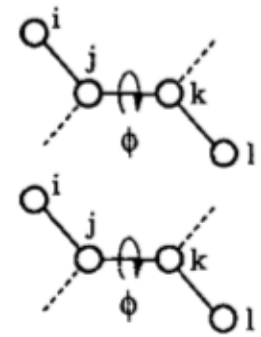
$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2 \vec{r}_i}{dt^2}$$

Coupled 2nd order Diff. Eq.

$$\vec{F}_i = -\frac{\partial V}{\partial \vec{r}_i}$$

How are they coupled?

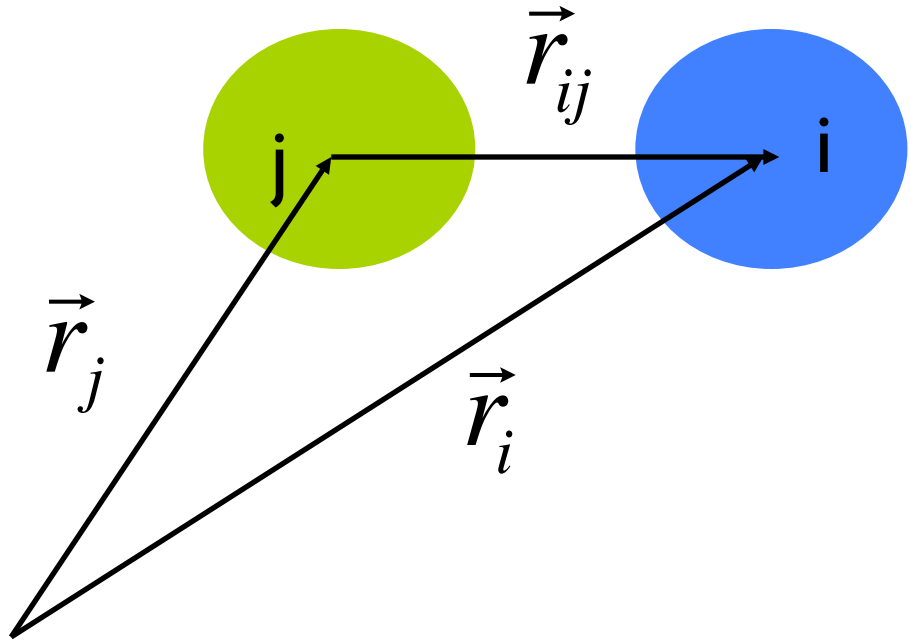
$$\vec{r}_i(t) \quad \text{for i=1,…N}_{\text{atoms}}$$

INTERACTION TYPE	FUNCTIONAL FORM
(i)  Long Range Interaction	$V_{BB}(r)= 4\varepsilon_h[(\sigma/r)^{12}-(\sigma/r)^6]$ $V_{L\alpha}(r)= 4\varepsilon_L[(\sigma/r)^{12}+(\sigma/r)^6]$ ← $\varepsilon_L= 2/3\ \varepsilon_h$ (α= L or B) $V_{N\alpha}(r)= 4\varepsilon_h(\sigma/r)^{12}$
(ii)  Bond Angle Potential	$V(\theta)=1/2k_\theta\ (\theta-\theta_0)^2$ $k_\theta= 20\varepsilon_h/\ (rad)^2$; $\theta_0= 105^\circ$
(iii)  Dihedral Angle Potential	(i, j, k≠ N) $V(\phi)=A\ (1+\cos\ \phi) + B\ (1+\cos\ 3\phi)$ $A= B= 1.2\varepsilon_h$ (i, j,= N ; (i, j, k= N) k, l,≠ N) $V(\phi)= B\ (1+\cos\ 3\phi)$ $B= 0.2\varepsilon_h$

(iv) Bond length potential

Figure 2. A sketch of the various interactions that are allowed in our model. The associated functional form of the potential and the parameters are also provided.

Pair Forces: Lennard-Jones Interactions



$$\vec{r}_j + \vec{r}_{ij} = \vec{r}_i$$

Parallelogram rule

$$\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$$

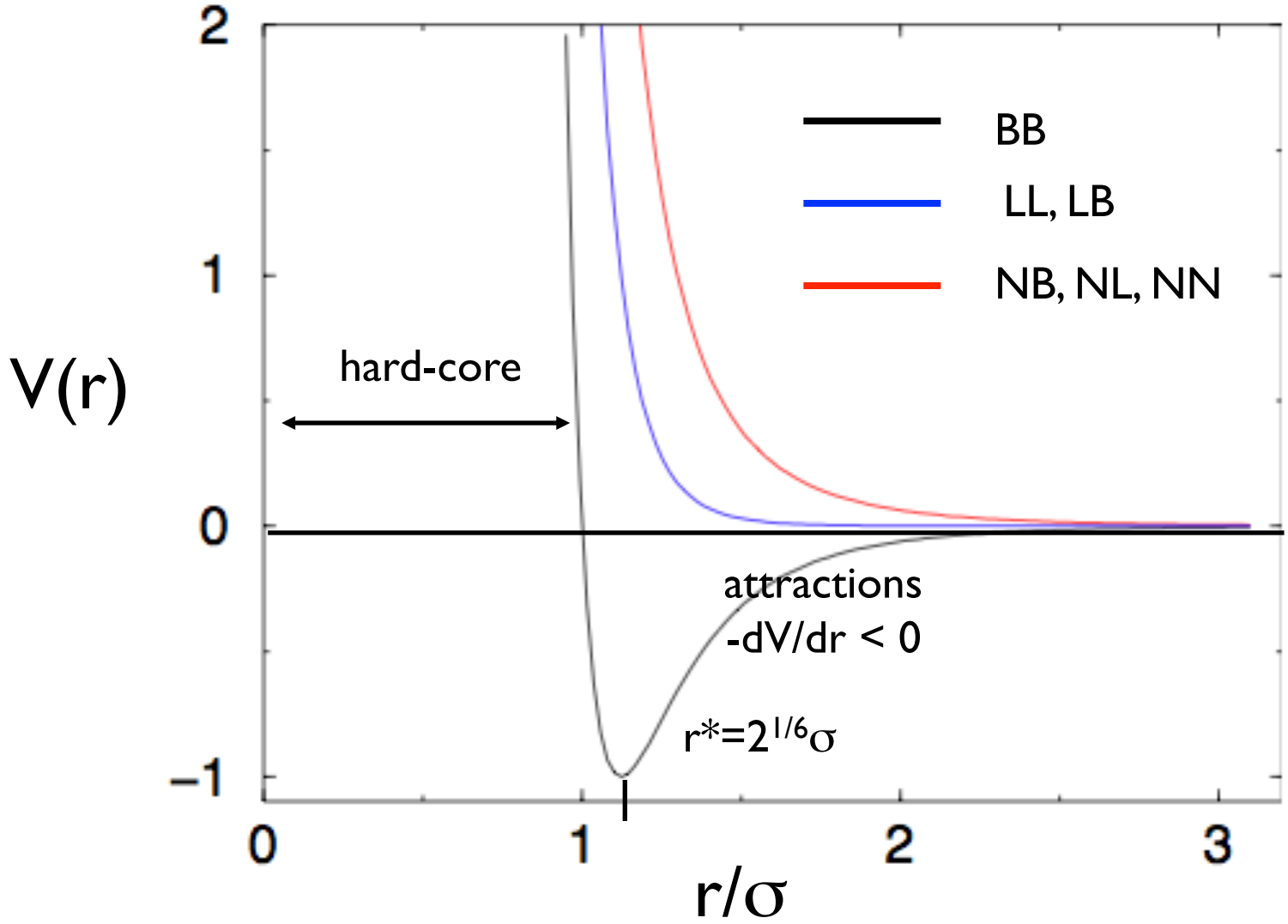
force on i due to j

$$\vec{F}_{ij} = -\frac{dV}{dr_{ij}}\hat{r}_{ij}$$

$-dV/dr_{ij} > 0$; repulsive
$-dV/dr_{ij} < 0$; attractive

$$\vec{F}_i = \sum_j \vec{F}_{ij}$$

‘Long-range interactions’



Bond Angle Potential

$$V_b\left(\theta_{ijk}\right) = \frac{k_0}{2}\left(\theta_{ijk} - \theta_0\right)^2 \qquad \theta_0 = 105°$$

ALA N-Cα-C Angles

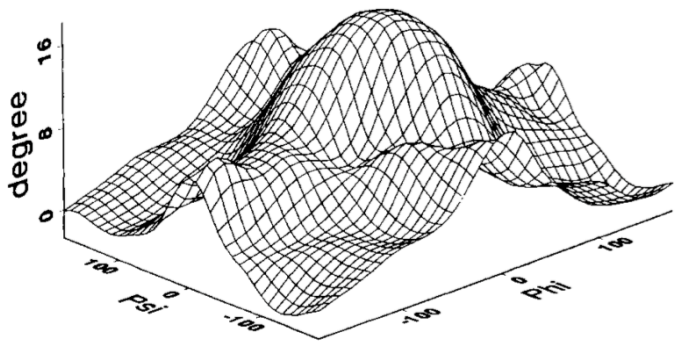
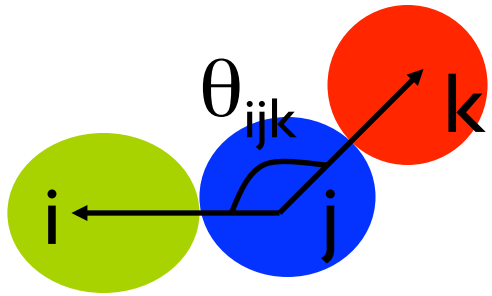


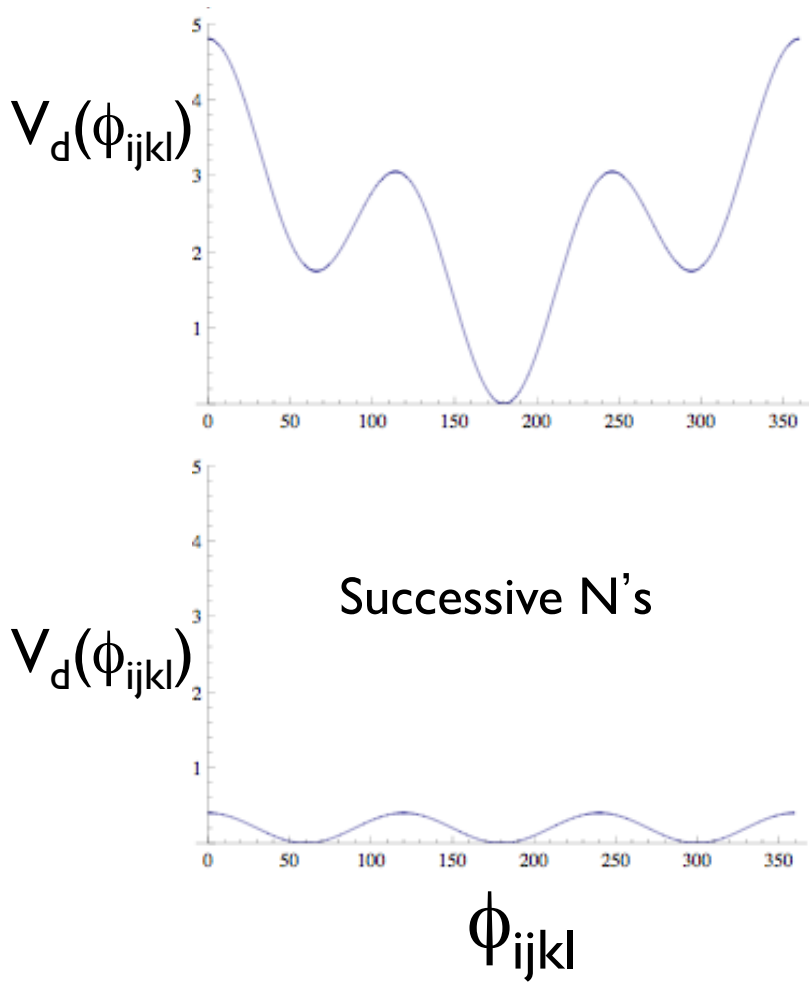
FIGURE 2 The variation of the N-C$^{\alpha}$-C′ backbone bond angle in ALA in ϕ(N-C$^{\alpha}$)/ψ(C$^{\alpha}$-C′) space. The values plotted (in degrees) are relative to the value (105.6°) of N-C$^{\alpha}$-C′ at (ϕ = −165°, ψ = −165°). They were obtained by HF/4-21G ab initio geometry optimizations at constant values of ϕ and ψ, forming a 30° grid in the region −180° < ϕ < +180° and −180° < ψ < +180°.



$$\cos\theta_{ijk} = \hat{r}_{ji} \bullet \hat{r}_{jk} \qquad \theta_{ijk} = [0,\pi]$$

$$\vec{F}_j^b = -\frac{dV_b}{d\theta_{ijk}}\frac{d\theta_{ijk}}{d\vec{r}_j} = -k_0\left(\theta_{ijk} - \theta_0\right)\frac{d\theta_{ijk}}{d\vec{r}_j}$$

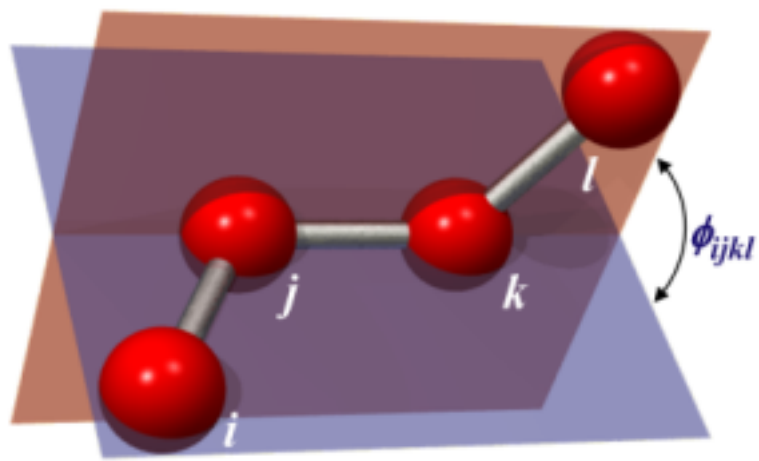
Dihedral Angle Potential



$$V_d\left(\phi_{ijkl}\right) = A\left(1+\cos\phi_{ijkl}\right) + B\left(1+\cos 3\phi_{ijkl}\right)$$

$$\cos\phi_{ijkl} = \frac{\left(\vec{r}_{ij}\times\vec{r}_{kj}\right)\bullet\left(\vec{r}_{jk}\times\vec{r}_{lk}\right)}{\left|\vec{r}_{ij}\times\vec{r}_{kj}\right|\left|\vec{r}_{jk}\times\vec{r}_{lk}\right|}$$

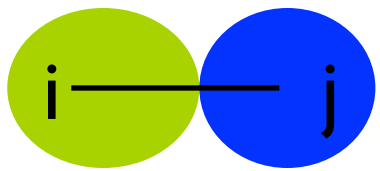
$$\vec{F}_j^d = -\frac{dV_d}{d\phi_{ijkl}}\frac{d\phi_{ijkl}}{d\vec{r}_j} = \left(A\sin\phi_{ijkl} + 3B\sin 3\phi_{ijkl}\right)\frac{d\phi_{ijkl}}{d\vec{r}_j}$$



Bond Stretch Potential

$$V_{bs} = \frac{k_{bs}}{2}\left(r_{ij} - \sigma\right)^2$$ for i, j=i+1, i-1

$$\vec{F}_{ij}^{bs} = -k_{bs}\left(r_{ij} - \sigma\right)\hat{r}_{ij}$$



Equations of Motion

$$\vec{F}_i^{tot} = \vec{F}_i^{lr} + \vec{F}_i^{b} + \vec{F}_i^{d} + \vec{F}_i^{bs} = m_i \vec{a}_i$$

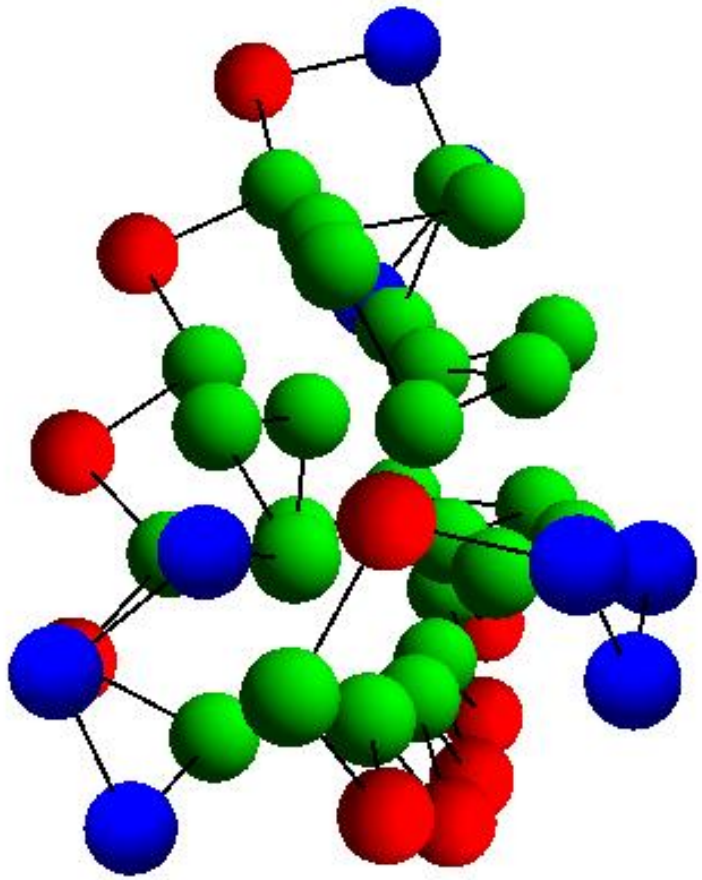
$$x_i(t+\Delta t) = x_i(t) + v_i(t)\Delta t + \frac{1}{2} a_i(t)(\Delta t)^2$$

$$v_i(t+\Delta t) = v_i(t) + \frac{a_i(t) + a_i(t+\Delta t)}{2}\Delta t$$

velocity verlet algorithm

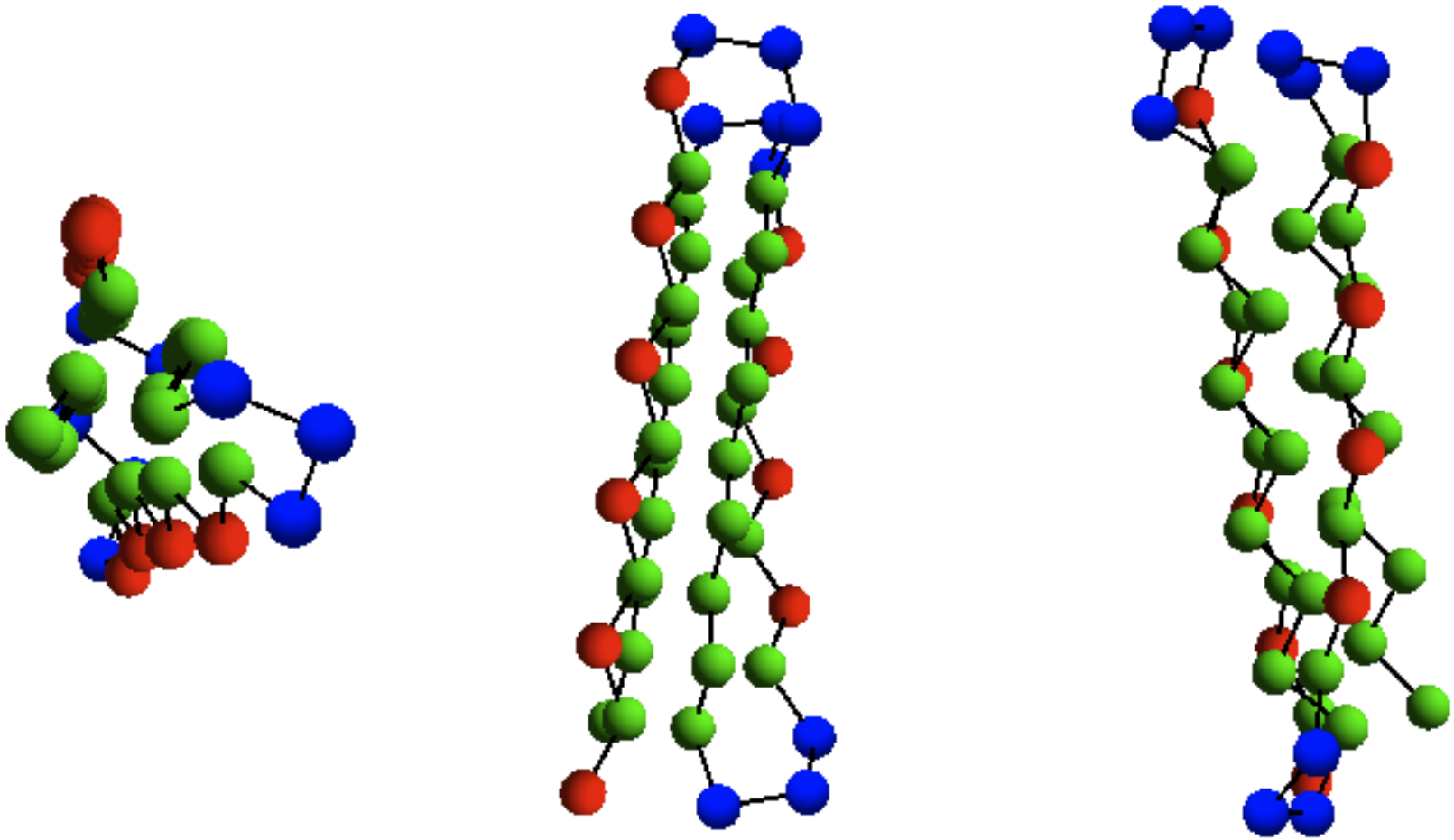
Constant Energy vs. Constant Temperature
(velocity rescaling, Langevin/Nosé-Hoover thermostats)

Collapsed Structure

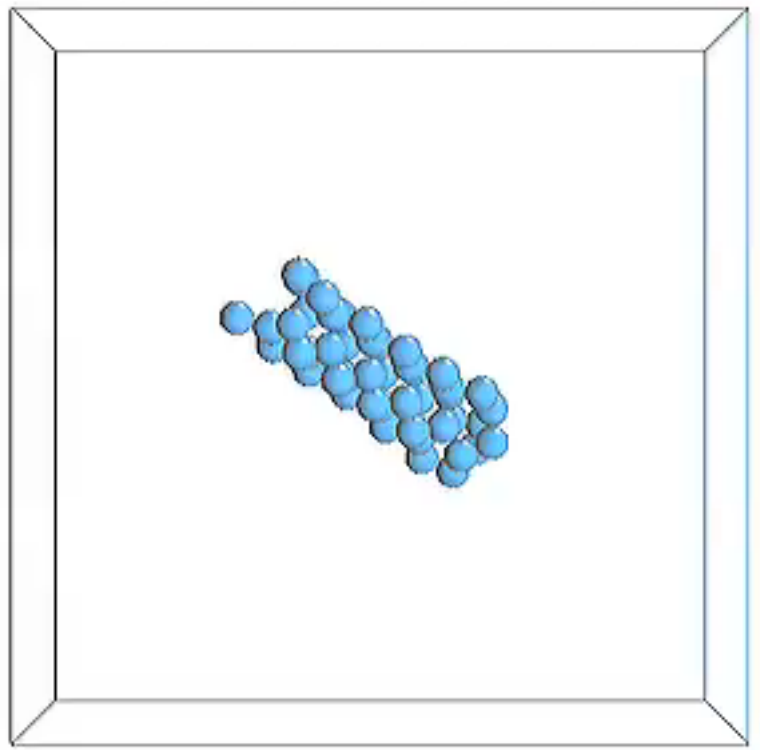


$T_0=5\varepsilon_h$; fast quench; $(R_g/\sigma)^2= 5.48$

Native State



$T_0=\varepsilon_h$; slow quench; $(R_g/\sigma)^2= 7.78$



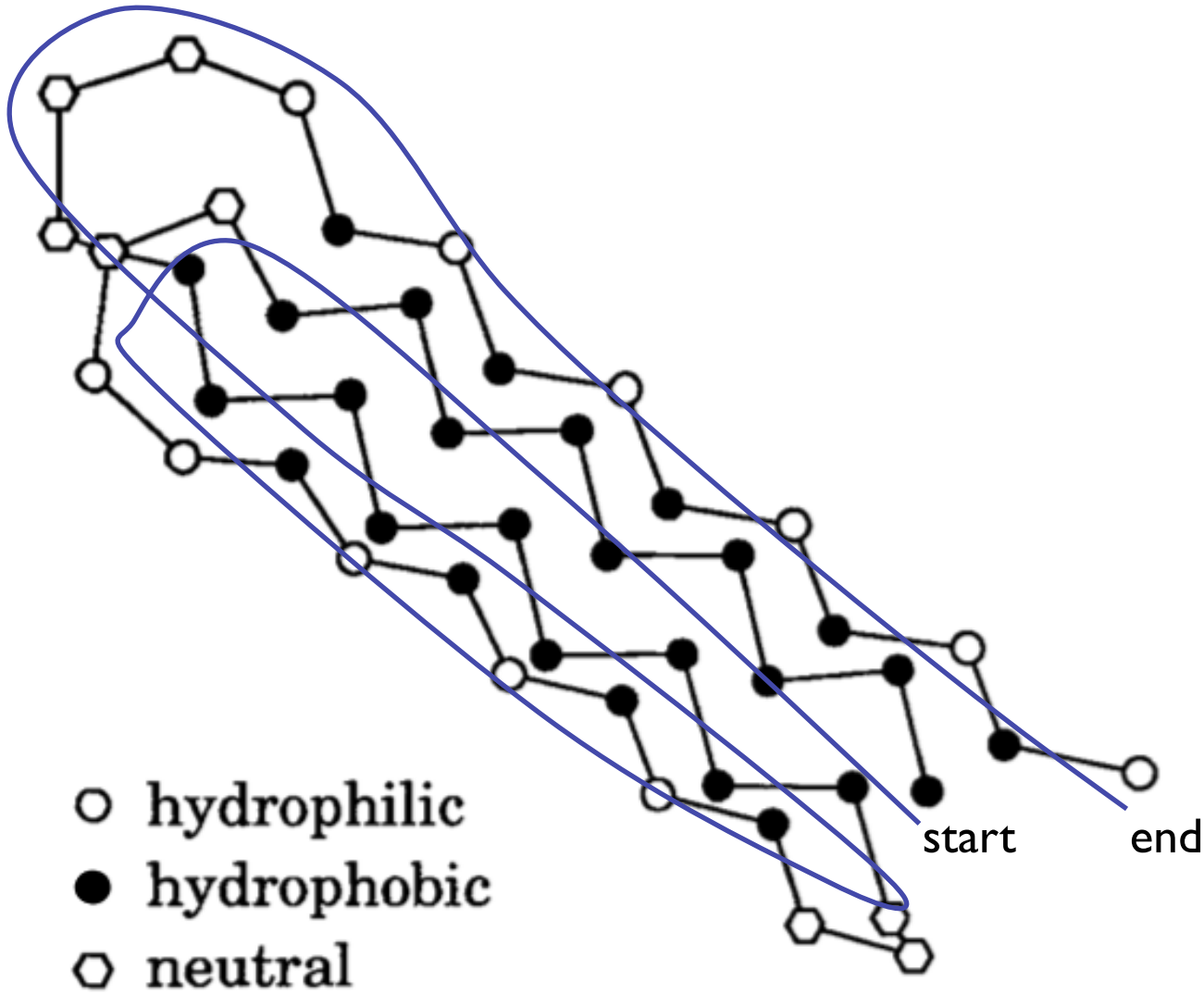


Figure 3. Typical β-barrel structure of the chain below the unfolding–folding transition temperature. The configurations were obtained by a process of slow cooling. The nature of the various residues for this chain consisting of 46 beads is shown by the various symbols.

Total Potential Energy

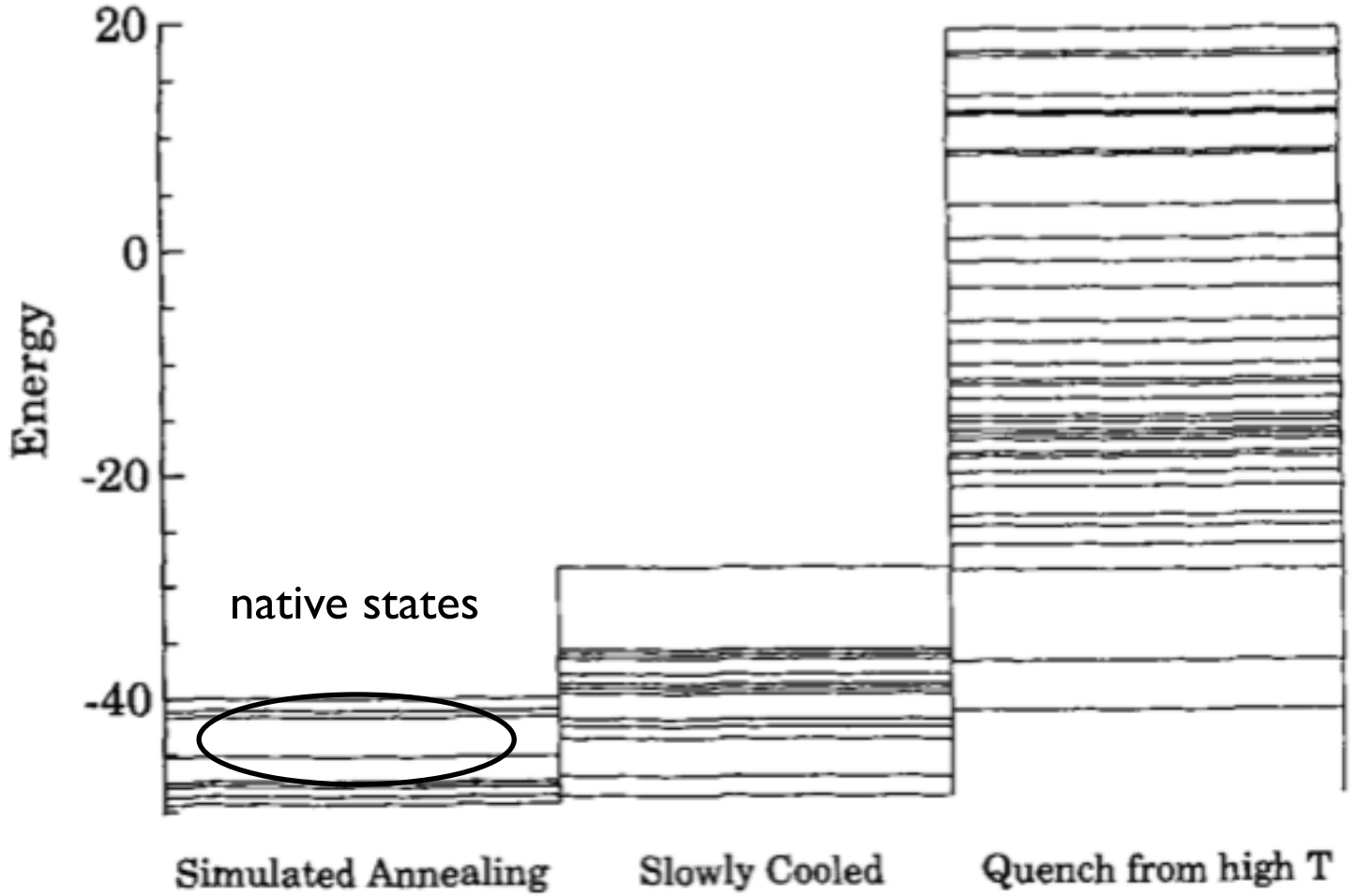


Figure 4. Spectrum of energies of the various quench structures obtained using several simulation techniques. The explanation of the three columns is given in the text. T is the temperature.

Radius of Gyration

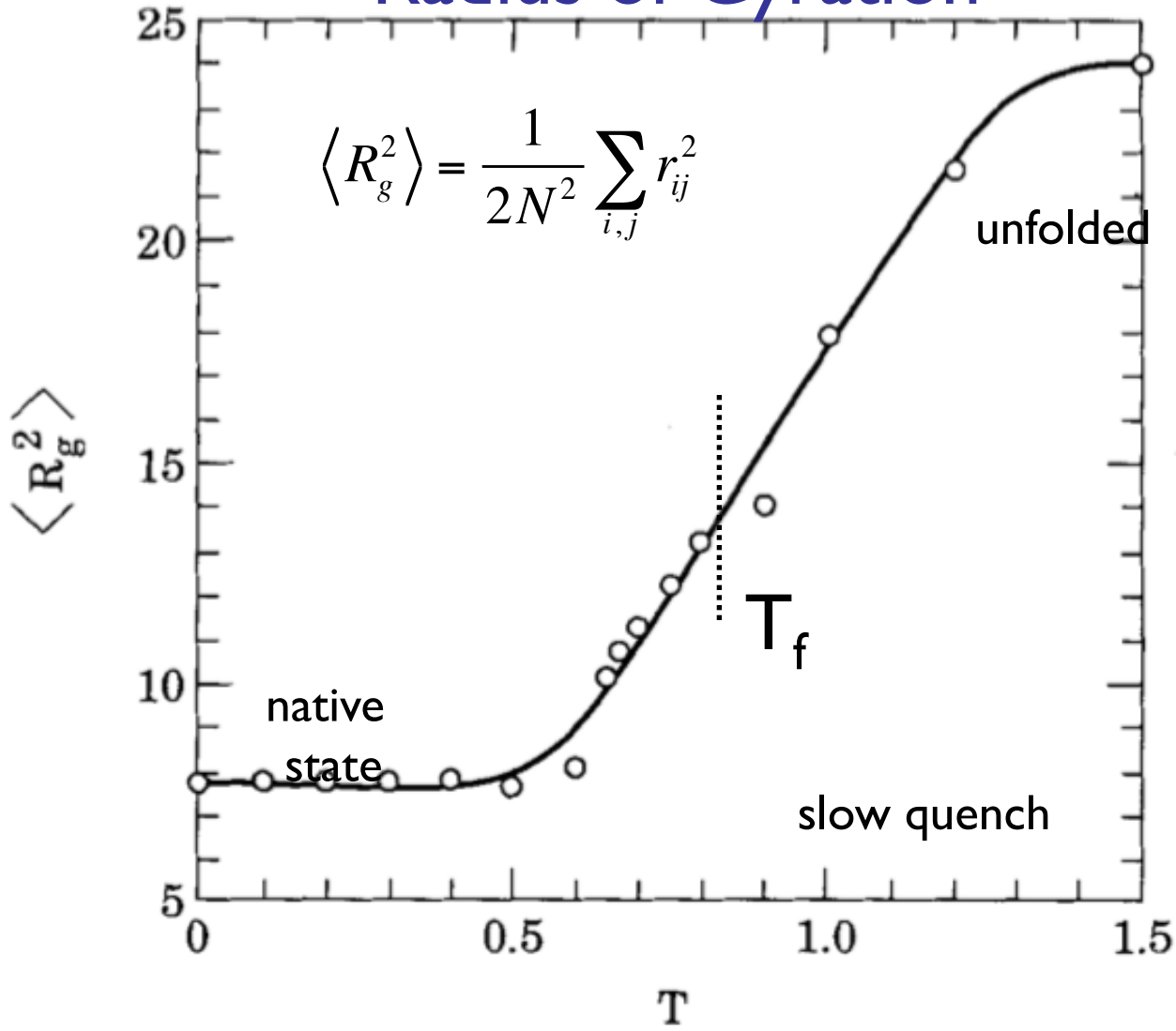


Figure 5. A plot of the radius of gyration $\langle R_g^2 \rangle$ as a function of temperature for the $n = 46$ case. The solid line is drawn by smoothly connecting all the points. The transition from an extended to a compact structure is at $T_F \sim 0.65$.

Reliable Folding at Low Rate