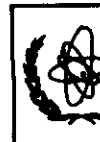




UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL ATOMIC ENERGY AGENCY
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE



SMR.961 - 18

**WORKSHOP ON:
PROTEINS, MEMBRANES and their INTERACTIONS**

22 JULY - 2 AUGUST 1996

***"Energy landscape theory
of protein folding"***

PART I

**Zaida LUTHEY-SCHULTEN
University of Illinois
School of Chemical Science
505 South Mathews Avenue
IL 61801 Urbana
U.S.A.**

These are preliminary lecture notes, intended only for distribution to participants.

Energy Landscape Theory of Protein Folding

I. Phenomenological Theory: Few Parameters

1. **Features common to both random heteropolymers (RHP) and proteins**

2. **Simplest Viable Protein Folding Funnel**

- slope $\frac{\delta E}{S}$ (biased), stratified, ΔE (ruggedness)

3. **Reaction Coordinates / Order Parameters**

- Q (native contacts), Correct Angles, % Secondary Structure.

4. **Thermodynamic and Kinetic Foldability**

- $\text{Max} \left(\frac{\delta E}{S} / \Delta E \right) \longleftrightarrow \text{Max } T_f / T_g$ Basis of Optimization

II. Protein Structure Prediction:

1. (Threading) Results using local energy function

- Discrimination and Partial Ordering
- Partial Unknowns: Signatures, M. Adenylate Kinase, and Cu-A

2. (Molecular Dynamics) Results using Associative Memory Energy Function

- Myoglobin and folding subunits "foldons"
- HIV1 Rev

III: Experimental Clues about the Topography of the Protein Folding Funnel

1. Helix-Coil Transition of a Collapsed Protein

2. Thermodynamics of Folding-Unfolding Experiments

3. Fast Folding Experiments - Multiple Time Scales.

Acknowledgments:

(Illinois) Peter Wolynes, J. Bryngelson (NIH)
M. Friederichs (SQUIB),
R. Goldstein (Mich), A. Pauchenko*,
R. Cole*, K. Koretke*

(CalTech) H. Gray, B. Ramirez

(Support): NIH, NSF, FMC, Carver
NCSA, San Diego (SDSC)

(San Diego) J. Onuchic, N. Socci

(Illinois) P. Haney, J. Konisky
M. Gruebele, R. Ballew

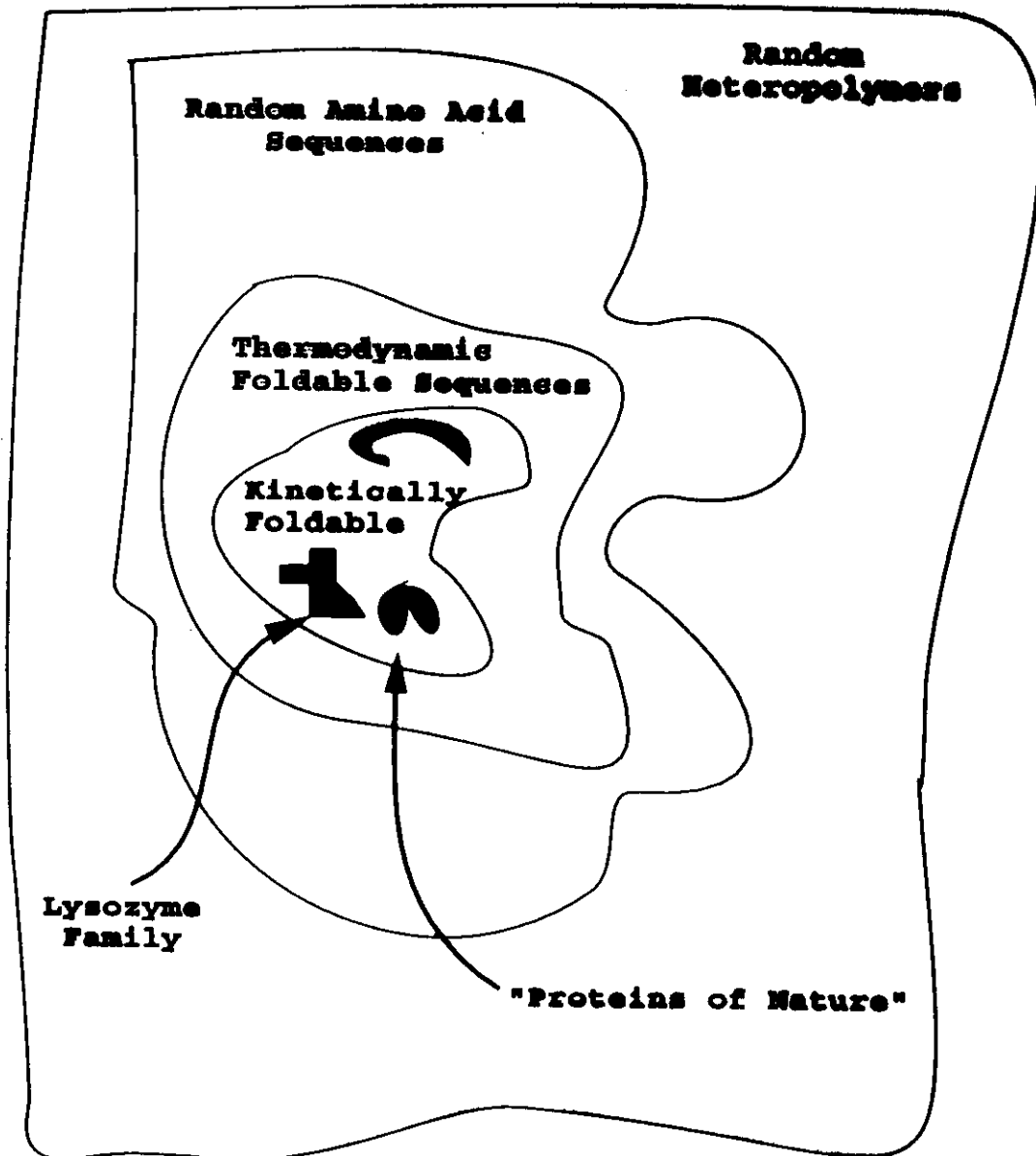
71

lyc_chick	DYGILQINER	WNCNDGRTFS	SWELCHIFGS
lyc_macmu	DYGIFQINSE	YWCNDGRTFS	AVRACHISGN
lyc_human	DYGIFQINER	YWCNDGRTFS	AVRACHISGN
lyc1_rat	DYGIFQINER	YWCNDGRTFR	AMRACHISGN
lyc_bovin	DYGIFQINER	WNCNDGRTFN	AVRACHISGN
lyc1_tacac	DYGILQINER	YWCNDGRTFS	SWELCHISGS

131

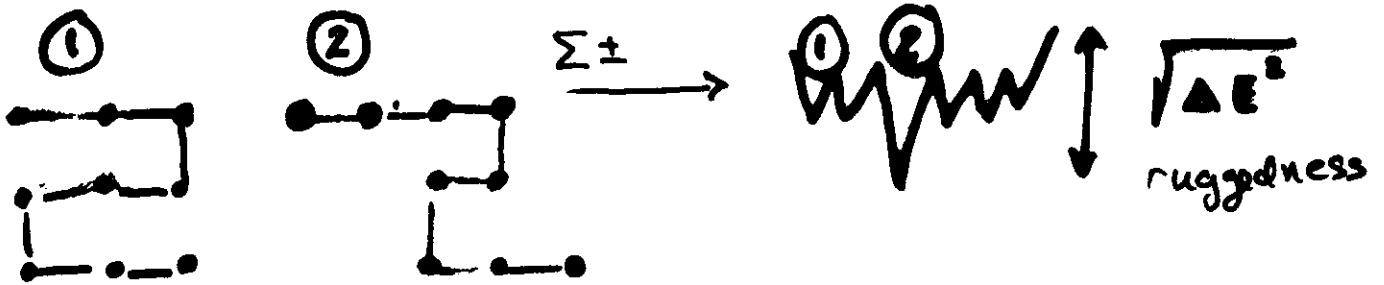
lyc_chick	ALLSSDITAS	VNCAKRVSD	GGGMIANVAN
lyc_macmu	ALLQDNIADA	VTCARVVSD	PQGIRANVAN
lyc_human	ALLQDNIADA	VACARVVSD	PQGIRANVAN
lyc1_rat	ALLQDDITQA	IQCARKVVSD	PQGIRANVAN
lyc_bovin	ELMENDIACA	VACARVISE	.QGITANVAN
lyc1_tacac	KLLDDITDD	LKCAKRIAGE	AKGLTPVAN

< 30%



Protein Sequences Look Random, Are They?

Postulate A: (Yes...) When non-native contacts are made, the energy contributions are random.



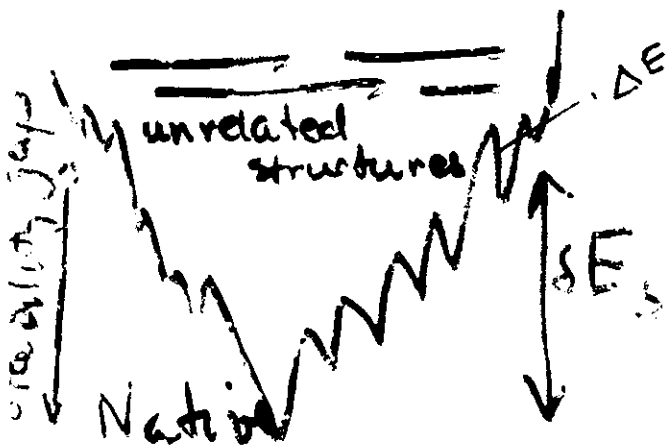
⇒ RHP has a REM landscape

Bryngelson & Wolynes, Garel & Orlando, Shakhnovich & Gutin.

Postulate B: (Not so fast...) There is a smooth overall slope because of the harmonious cooperativity of interactions:

Principle of Minimal Frustration

cf: Go

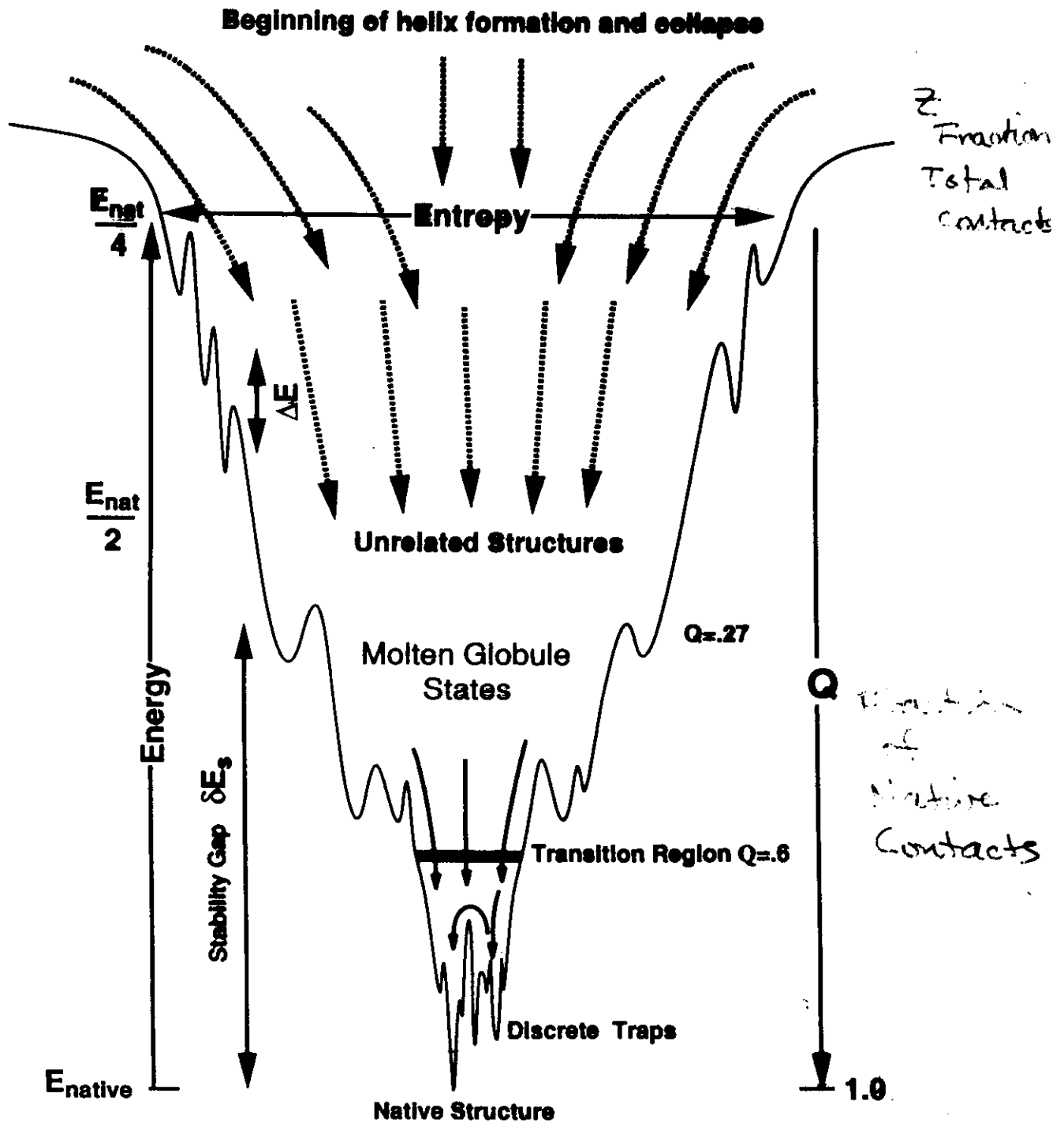


As native contacts are made energy goes down more than usual for

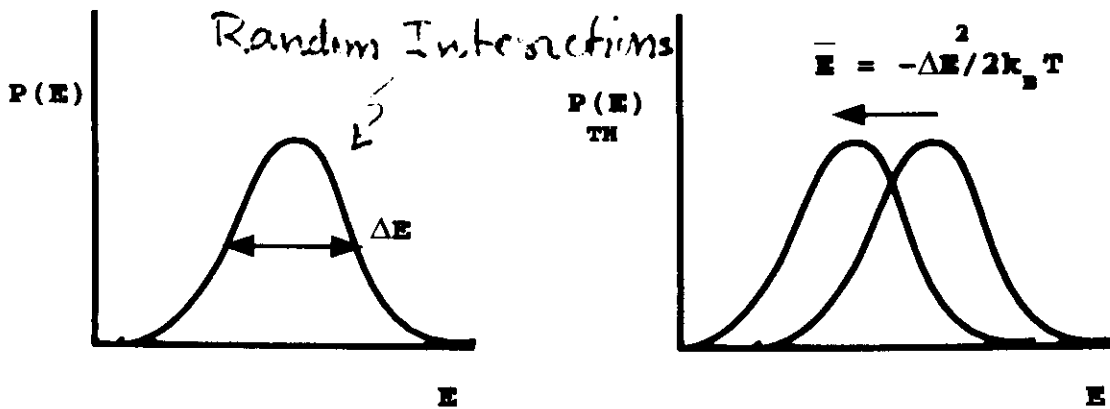
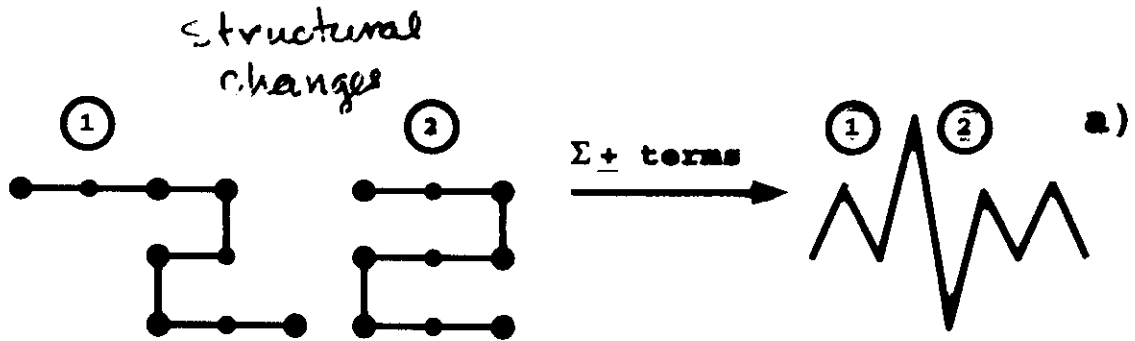
Onuchic, et al. RHP

Computer studies lattice: Thirumalai/Honig, Shakhnovich, Karplus, Sali

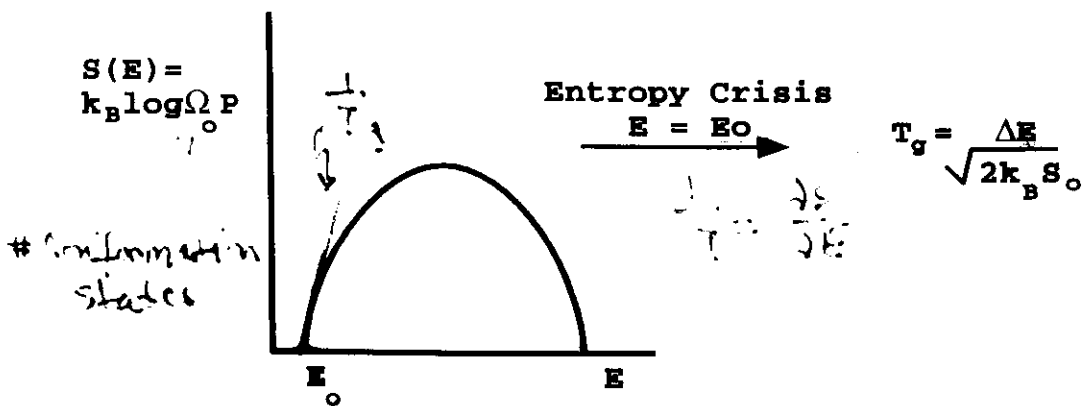
Simplest viable Protein Folding Funnel



Energy Landscape of a Random Heteropolymer



$$P_{TH}(E) = \frac{e^{-E/KT} \cdot e^{-E^2/\Delta E^2}}{Z}$$



Postulate A: Probability distribution of misfolded states is gaussian

REM

$$P(E)\delta E = \frac{1}{\sqrt{2\pi\Delta E^2(Q)}} \exp\left\{-\frac{(E - \bar{E}(Q))^2}{2\Delta E^2(Q)}\right\} \delta E \quad (1)$$

Density of conformational states with energy E and similarity Q:

$$\Omega(E, Q) = \Omega_0(Q)P(E) \quad (2)$$

Total entropy:

$$S(E, Q) = S_0(Q) - k_B \frac{(E - \bar{E}(Q))^2}{2\Delta E^2(Q)} - k_B \log\left(\frac{\sqrt{2\pi\Delta E^2(Q)}}{\delta E}\right) \quad (3)$$

Most Probable Energy:

$$E_{m.p.}(Q) = \bar{E}(Q) - \frac{\Delta E^2(Q)}{k_B T} \quad (4)$$

Corresponding Entropy:

$$S(E_{m.p.}, Q) = S_0(Q) - \frac{\Delta E^2(Q)}{2k_B T^2} \quad (5)$$

Free-energy of the misfolded structures with configurational similarity Q and at a fixed temperature:

$$\begin{aligned} F(Q, T) &= E_{m.p.}(Q) - TS(E_{m.p.}, Q) \\ &= \bar{E}(Q) - \frac{\Delta E^2(Q)}{2k_B T} - TS_0(Q) \end{aligned} \quad (6)$$

At the folding temperature T_f , $S_{native} = 0$ and $F_{native} = F(Q_{min}, T_f)$,

$$T_f = \frac{\delta E_s + \sqrt{\delta E_s^2 - 2S_0 \Delta E^2(Q_{min})/k_B}}{2S_0} \quad (7)$$

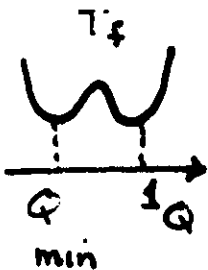
Local glass transition in each stratum is characterized by an entropy crisis where $S(T_g, Q) = 0$

$$\begin{aligned} T_g(Q) &= \sqrt{\frac{\Delta E^2(Q)}{2k_B S_0(Q)}} \\ \frac{T_f}{T_g} &\approx 2\sqrt{\frac{k_B}{2S_0} \left(\frac{\delta E_s^2}{\Delta E^2}\right)} \end{aligned} \quad (8)$$

For a protein to fold, T_f/T_g must be greater > 1 .

$$\frac{T_f}{T_g} = \sqrt{\lambda} + \sqrt{\lambda - 1} \quad \text{where } \lambda = \frac{k_B}{2S_0} \left(\frac{\delta E_s^2}{\Delta E^2}\right)$$

$$\frac{S}{E} = \frac{1}{T}$$



$$S E_s = \bar{E} - E_{min}$$

$$\begin{aligned} &> \text{Max } T_f/T_g \\ &> \text{Max } \frac{\delta E_s}{\Delta E} \end{aligned}$$

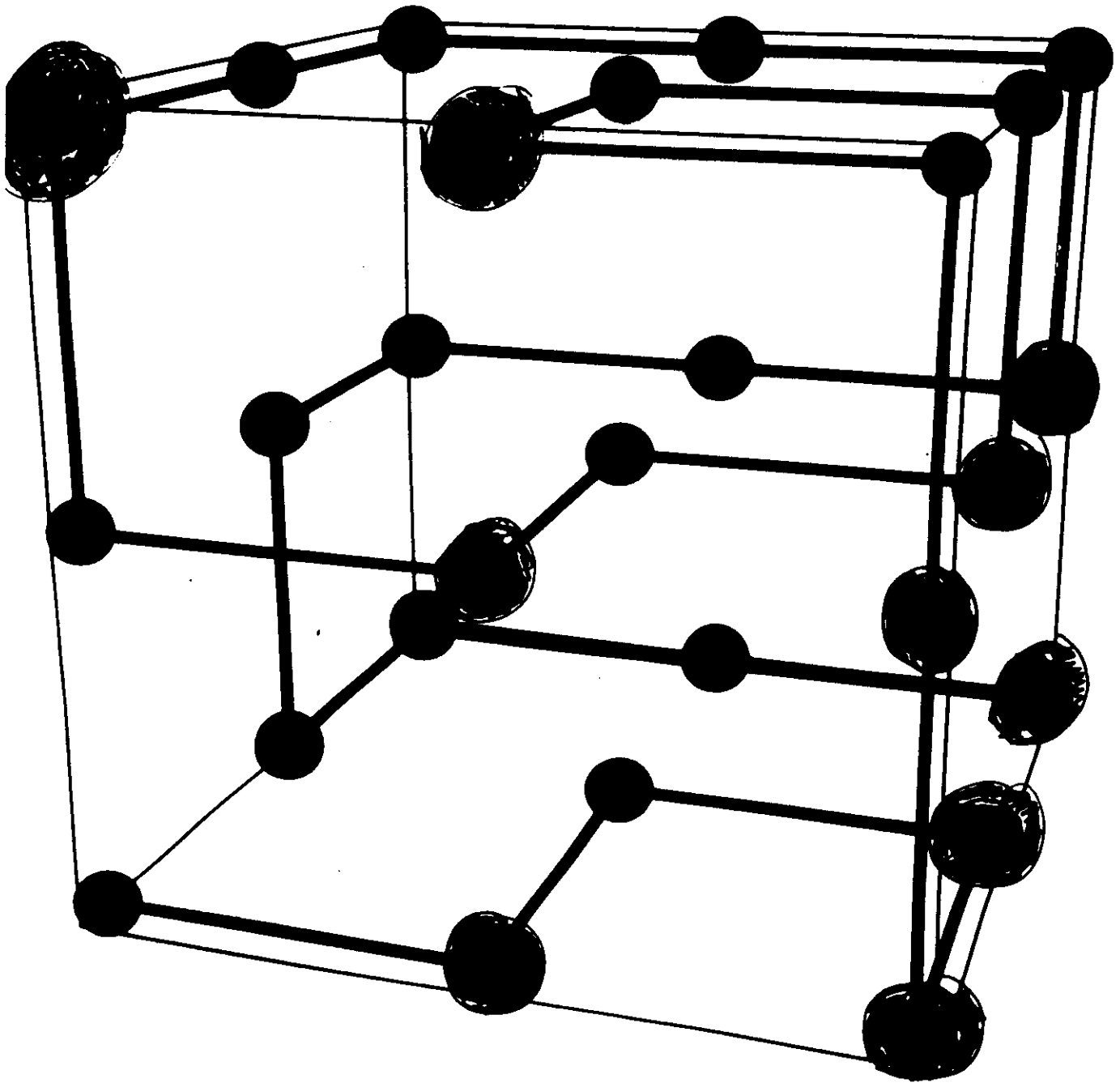
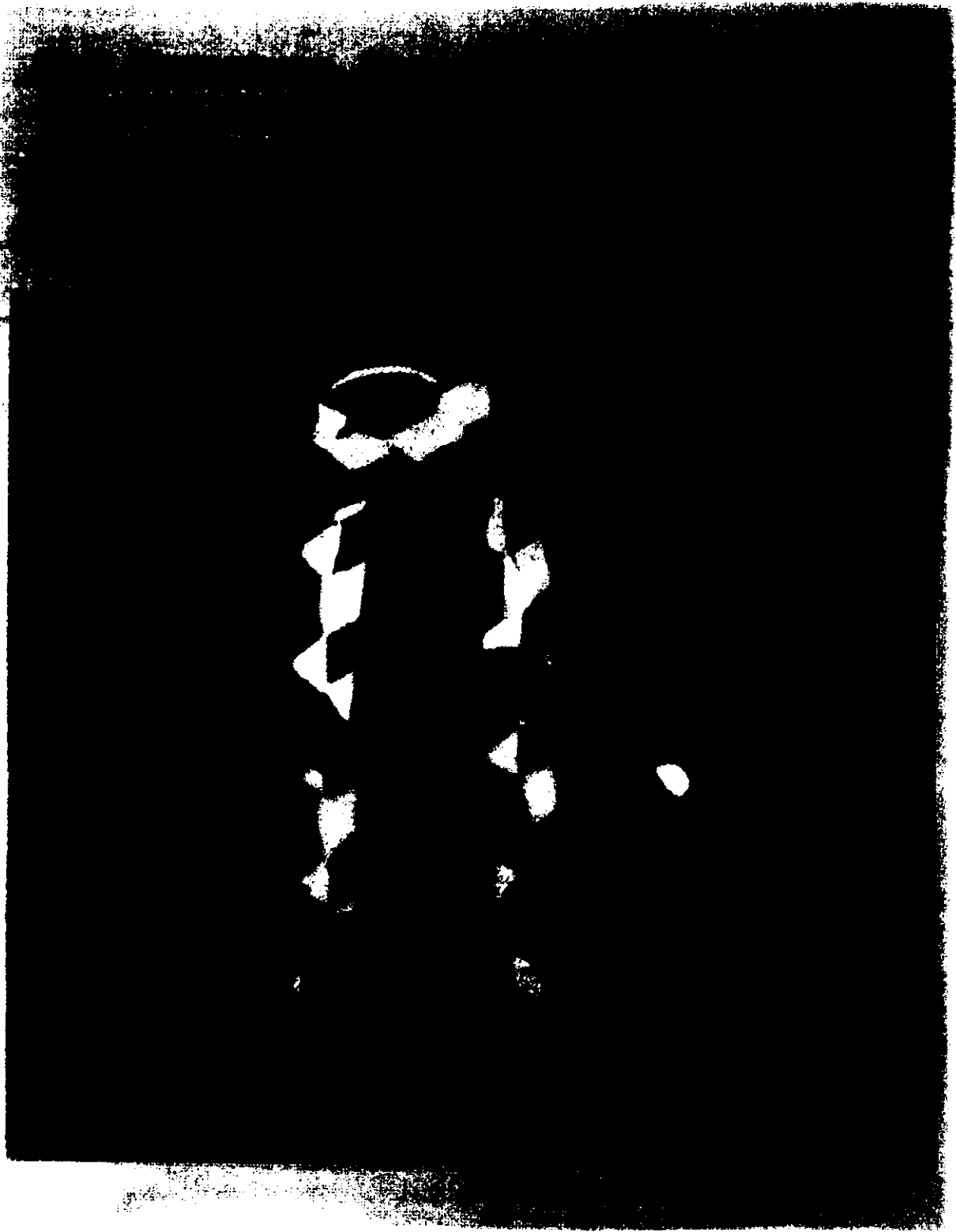
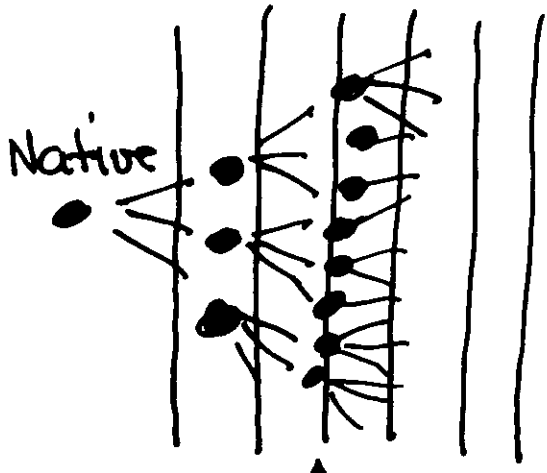


Figure 1: The native state (minimum energy) cube for the sequence studied in this work.

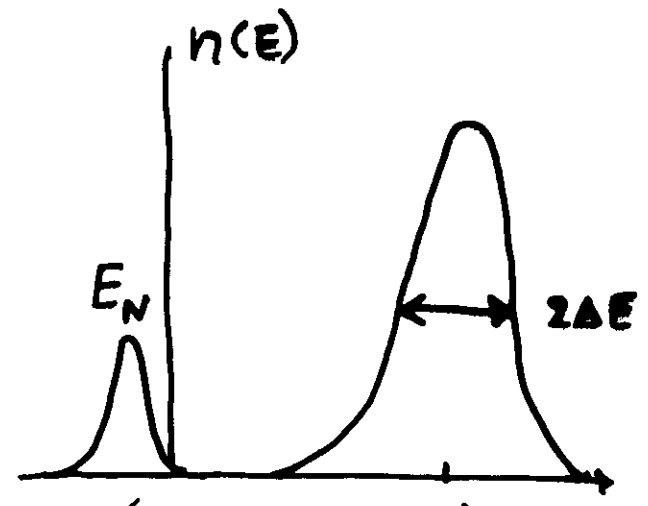




$Q=1$
 $A=1$
 $E=E_{nat}$
 $S=0$

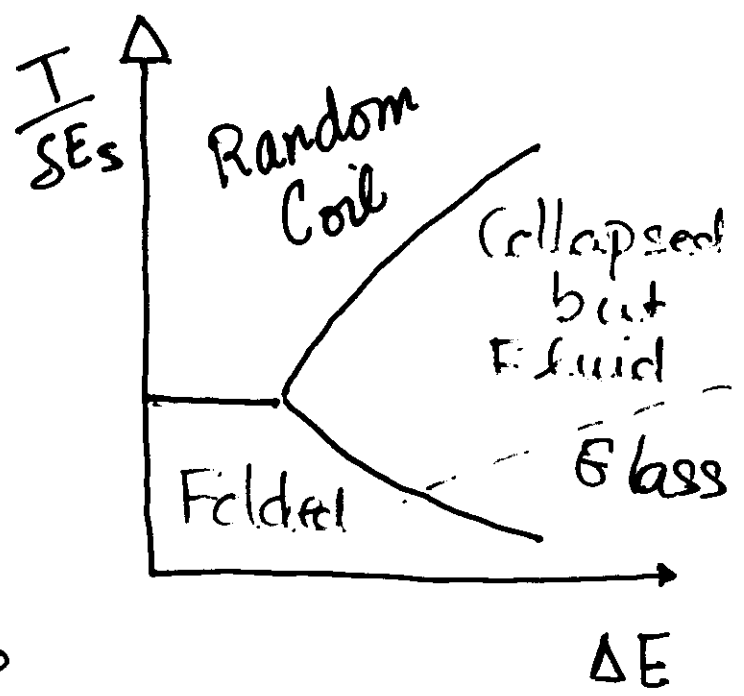
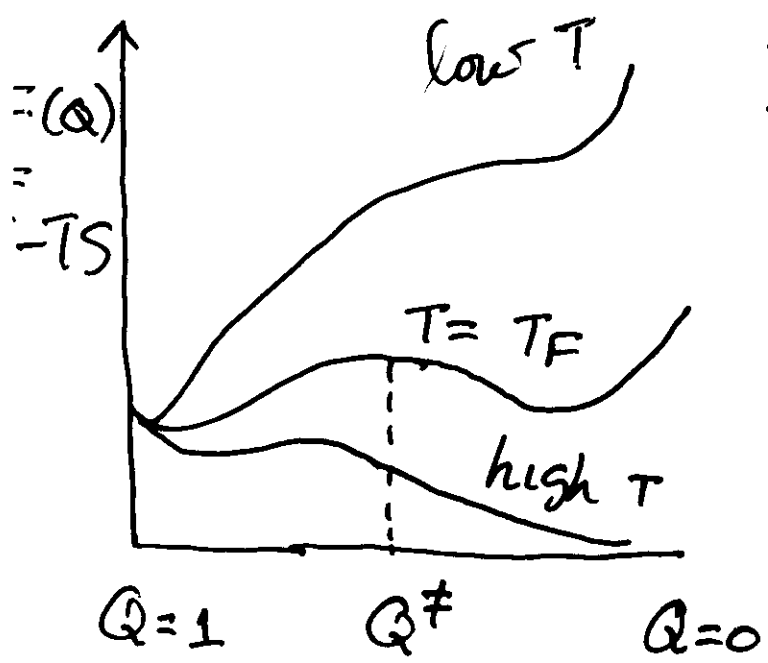
Each layer
 is a REM
 Hopping Rate
 $\sim e^{-\Delta E^2(Q)/2T^2}$

$Q=0$
 $A=0$
 $E=0$
 $S=S_0$



Stability Gap
 $\Delta E_S = E_N - \bar{E}_{MG}$

$$\frac{T_F}{T_G} \propto \frac{\Delta E_S}{\Delta E} > 1$$



Kinetics of a Minimally Frustrated Heteropolymer

I. Diffusion Eq. for Probability Flow between Strata:

$$\frac{\partial P(Q,t)}{\partial t} = \frac{\partial}{\partial Q} \left\{ D(Q,t) \left[\frac{\partial P(Q,t)}{\partial Q} + P(Q,t) \frac{\partial \beta F(Q)}{\partial Q} \right] \right\}$$

\uparrow
 $D_0 e^{-\beta \Delta E^2(Q)}$

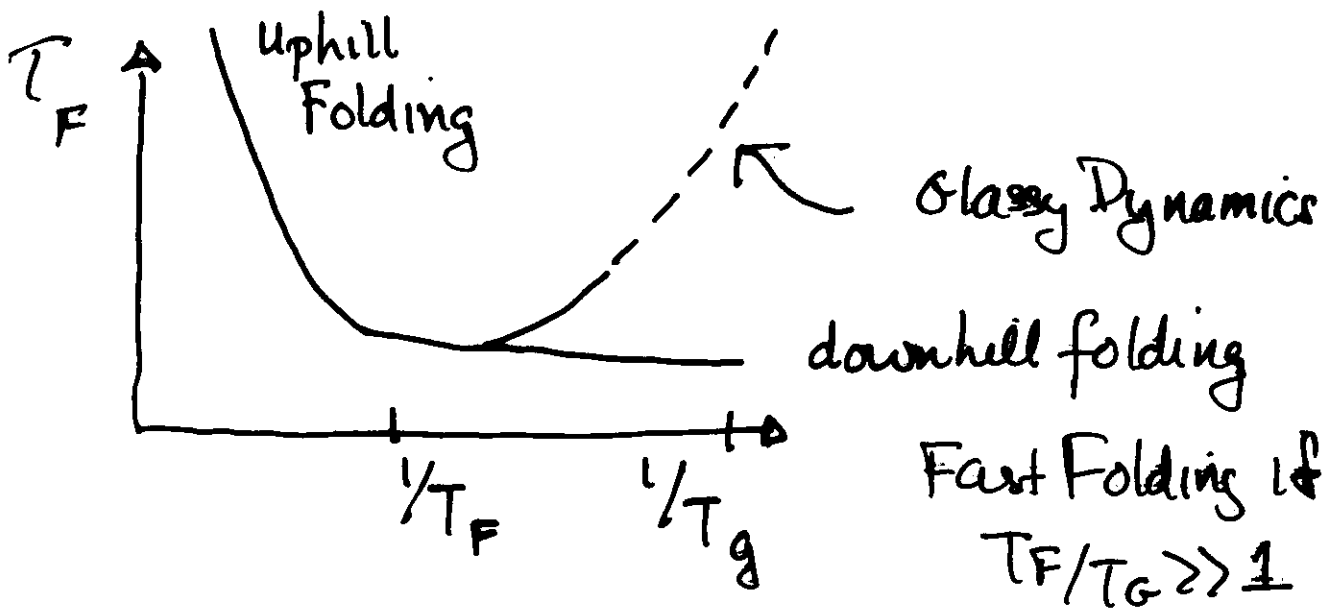
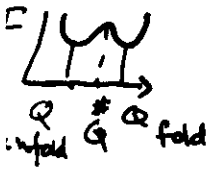
\uparrow
 Direction of Flow

"Folding Time"

$$\tau_F = \int_{Q_{unf}}^{Q_{fold}} dQ \int_0^Q dQ' \frac{\exp[\beta F(Q) - \beta F(Q')]}{D(Q)}$$

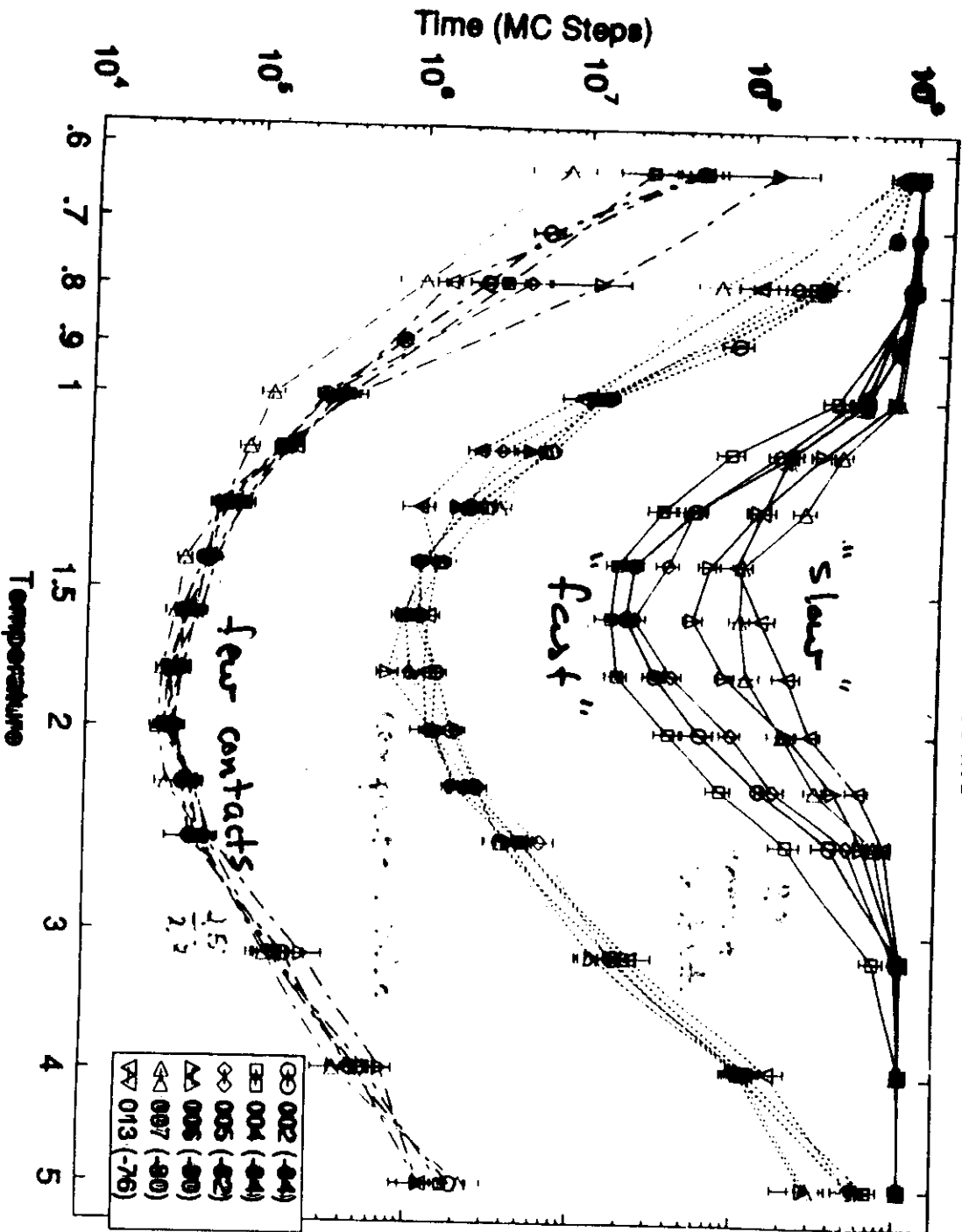
$$\approx \left(\frac{2\pi}{\beta}\right)^{1/2} \frac{1}{D(Q^*) \omega_{unf} \omega^\ddagger} e^{\beta [F(Q^\ddagger) - F(Q_{unf})]}$$

$$\approx \tau_0 e^{-\beta \Delta E^\ddagger} e^{\beta \Delta E^*}$$



Time Scales vs Temperature

3x3x3 Lattice MC

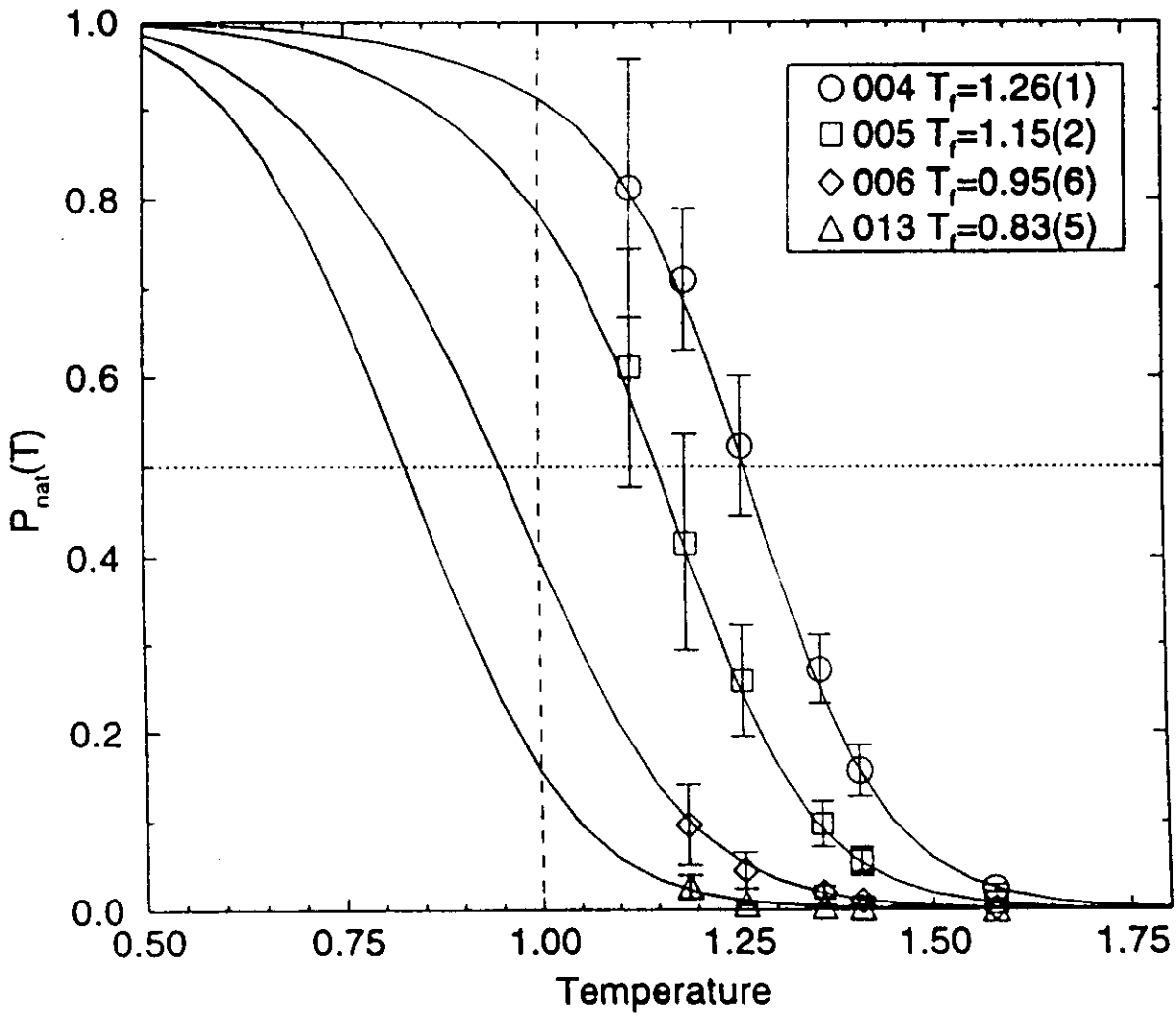


Thu Jun 2 13:57:58 1994

Secci & Onuchic

JCSO

Thermodynamic Study of $P(T)$ Native



1/2

Monte Carlo Histogram Technique

Ferrenberg
Swendsen
Onuchic
Socci

Histogram: $h(E, T')$ = # times E occurs in simulation = $\frac{n(E) e^{-E/T'}}{Z(T')} = P(E)$

$$Z(T') = \sum_E n(E) e^{-E/T'}$$

\uparrow density of states?

$$n(E) = h(E, T') e^{E/T'} \underbrace{Z(T')}_{\text{constant}}$$

Thermodynamic Quantities: $\langle \dots \rangle(T)$

$$\langle O \rangle(T) = \frac{\sum_E O(E) n(E) e^{-E/T}}{\sum_E n(E) e^{-E/T}} = \frac{\sum_E O(E) h(E, T') e^{E/T' - E/T}}{\sum_E h(E, T') e^{E/T' - E/T}}$$

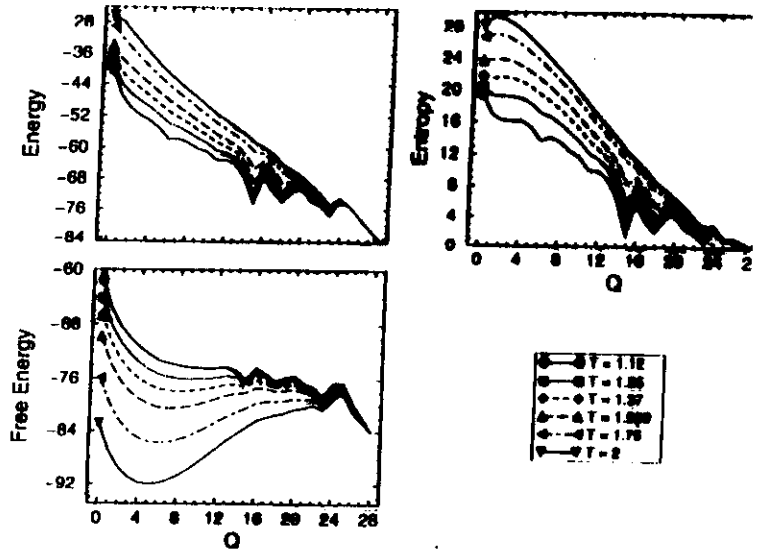
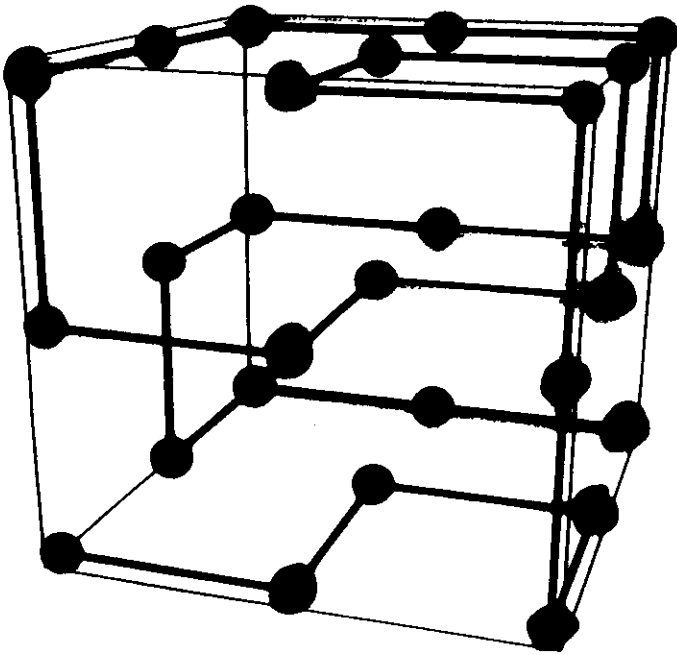
Free Energy: $F(T) = -T \log Z(T)$ $\rightarrow Z(T')$

$$F(T) = -T \log Z(T) = -T \ln \left[\sum_E h(E, T') e^{E/T' - E/T} \right] - T \ln Z(T')$$

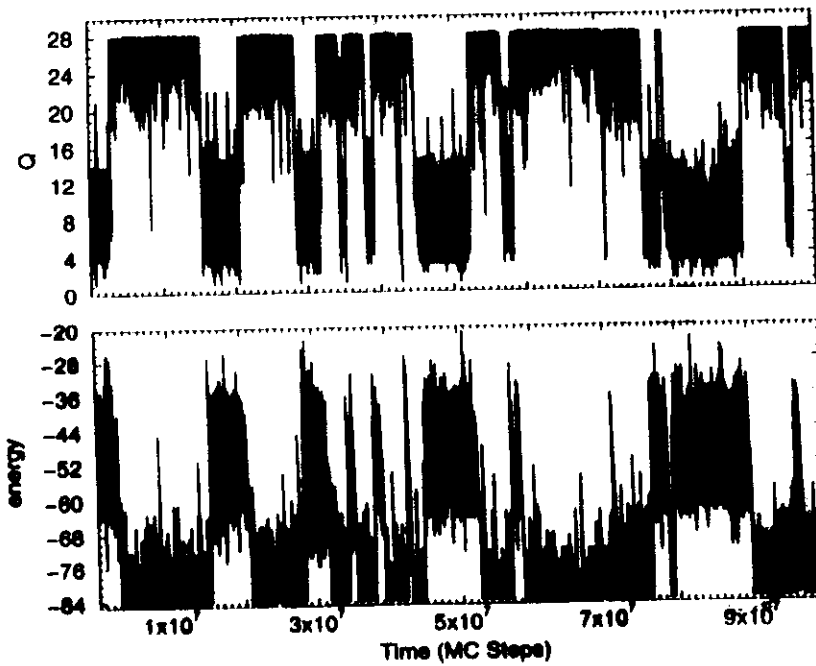
•
•
•
•

* Use simulation at T' to calculate thermodynamic quantities at T' and T .

$E(Q), F(Q), S(Q)$ vs. Q



2-state approximation



$T = T_f$

$\langle n \rangle \approx 7$

$E_{MC} \approx -50$

Socci, Onuchic, Wolynes
JCP 1996

Time \rightarrow

A-Q Trajectory

Seq012 (3 Letter Code)

$T=1.509 (T_f)$

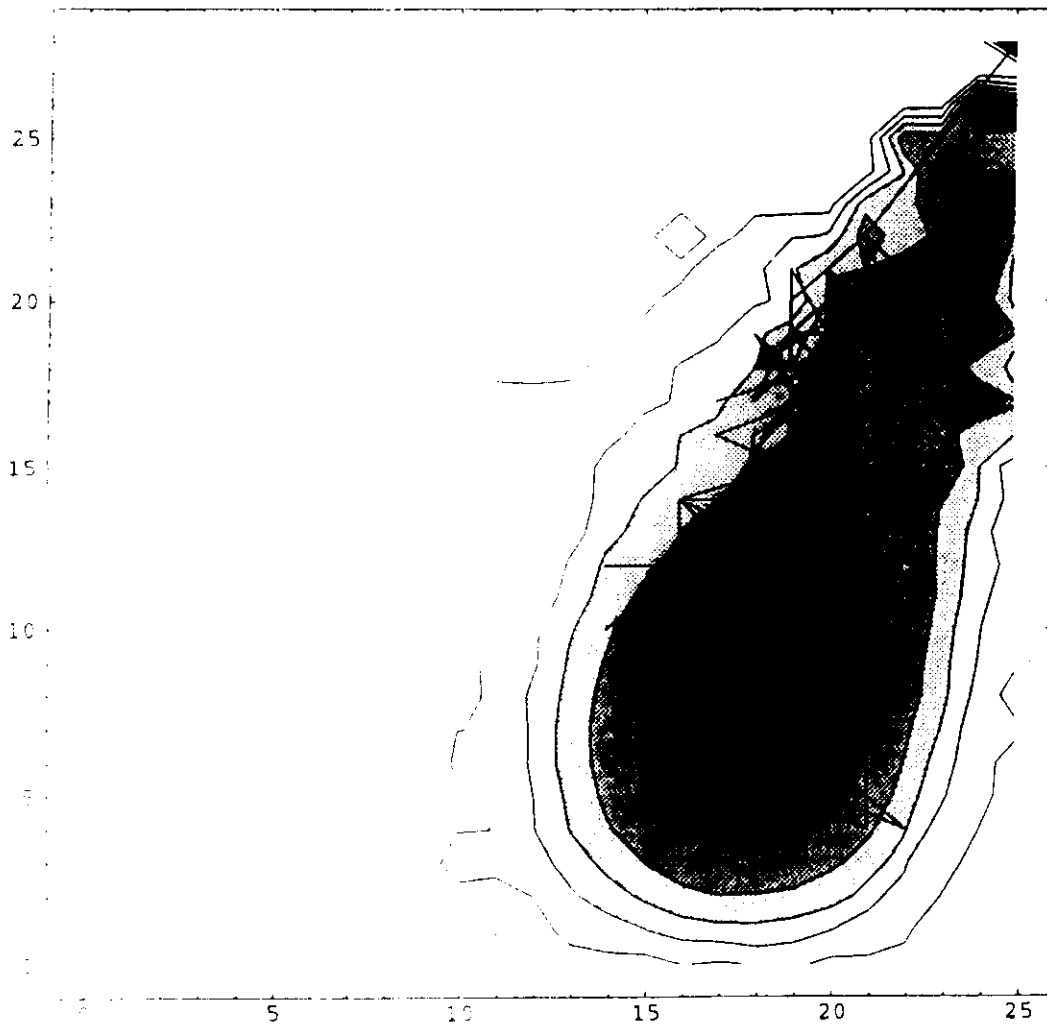
Trag 004 (dt=135, steps=1024)

contours = (-82, -80, -78, -76, -74, -72, -70, -68)

In[20]:=

```
Show[%3, %19, PlotRange->((-1, 26), (-1, 29))];
```

Q

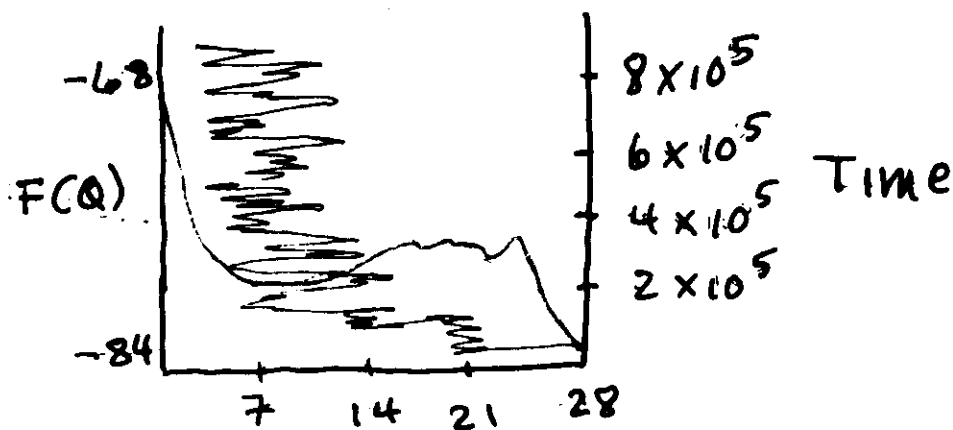


A

NDSocci

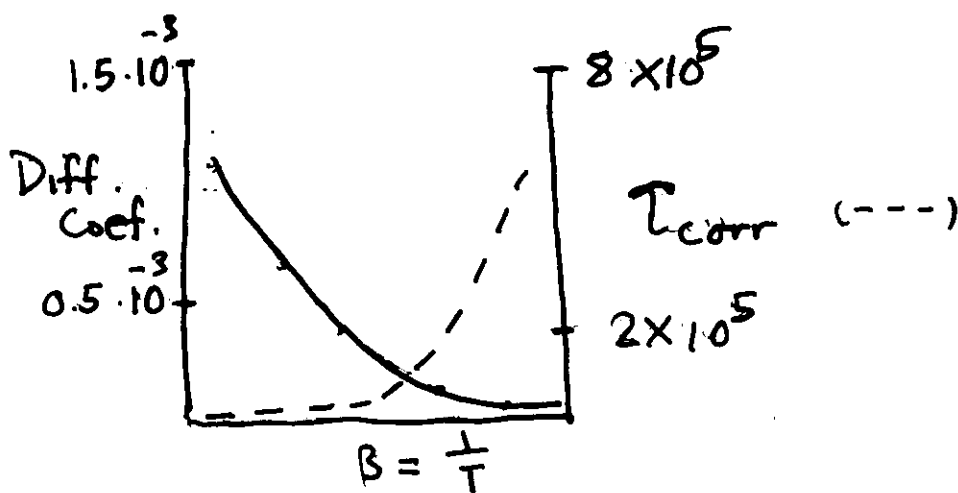
Interval before transition to Native

A Folding Trajectory $Q(t)$ at T_f



③ Native Contacts

Autocorrelation Time and Diffusion Coefficient



$$D = \frac{\Delta G^2(T)}{\tau_{corr}(T)}$$

Autocorrelation Function (Q)

Seq012 T=1.509

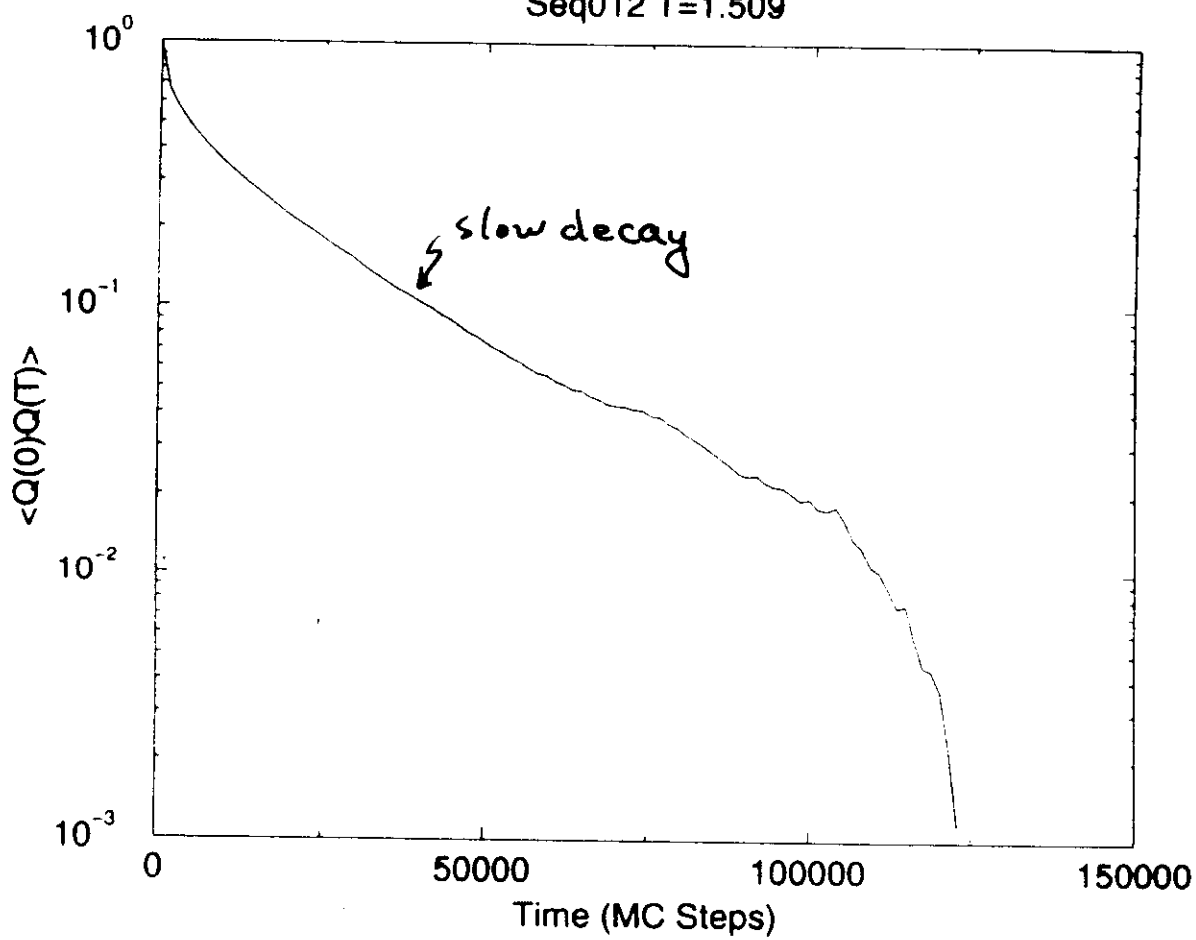
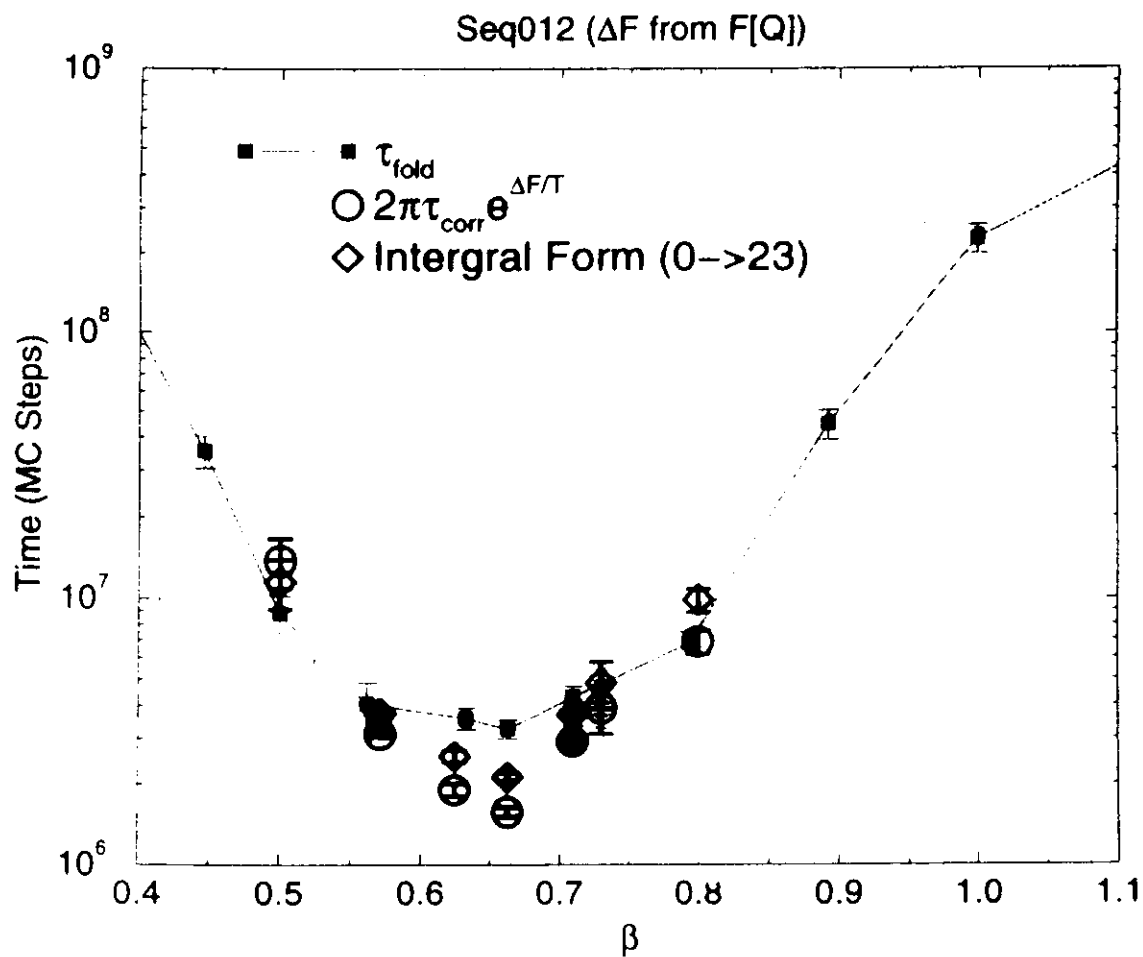


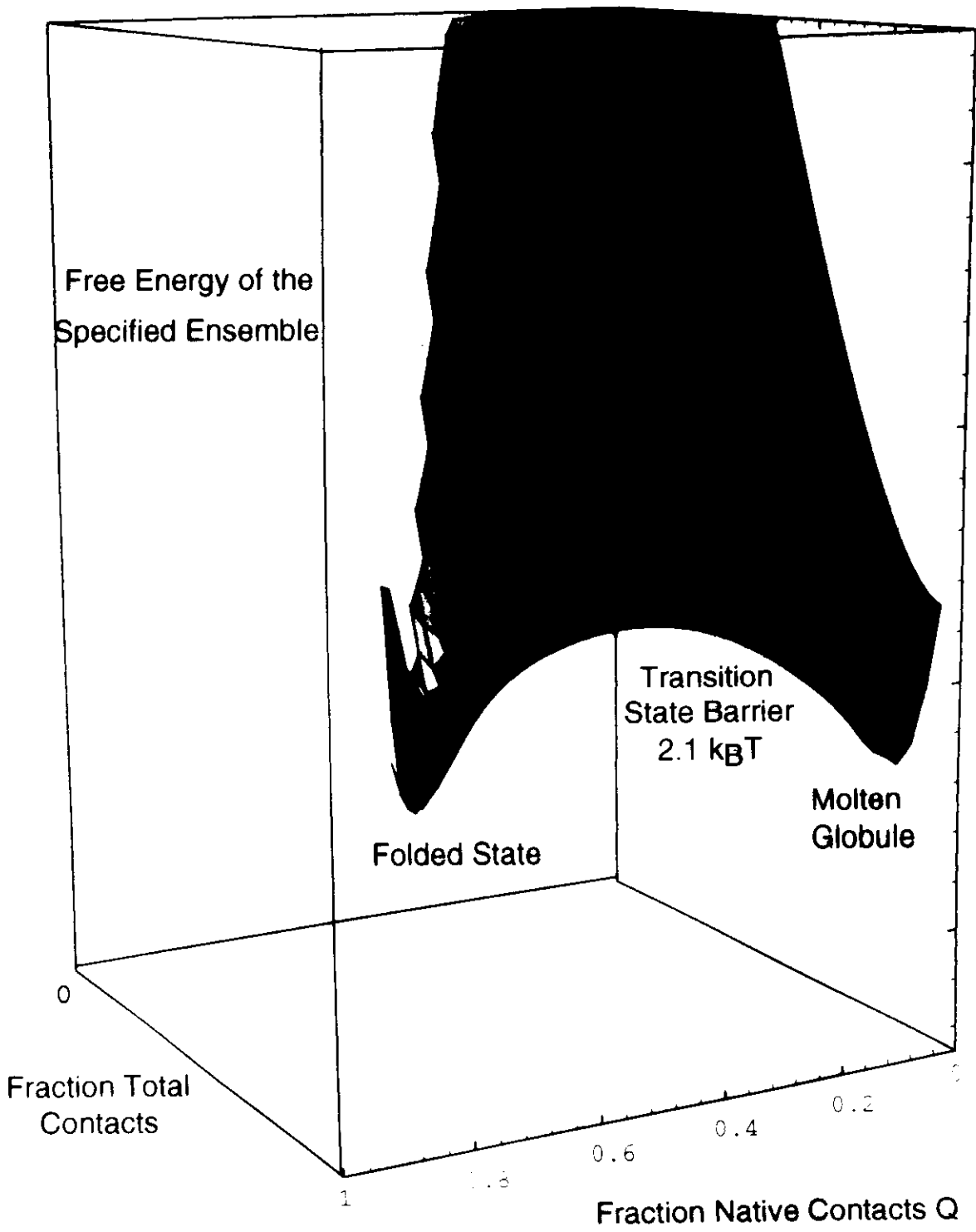
Figure 4: Autocorrelation function for Q ($\langle Q(0)Q(t) \rangle$) plotted on a semi-log scale. $T = 1.509$ which is the folding temperature for this sequence.

Kramers Rate vs MC MFPT

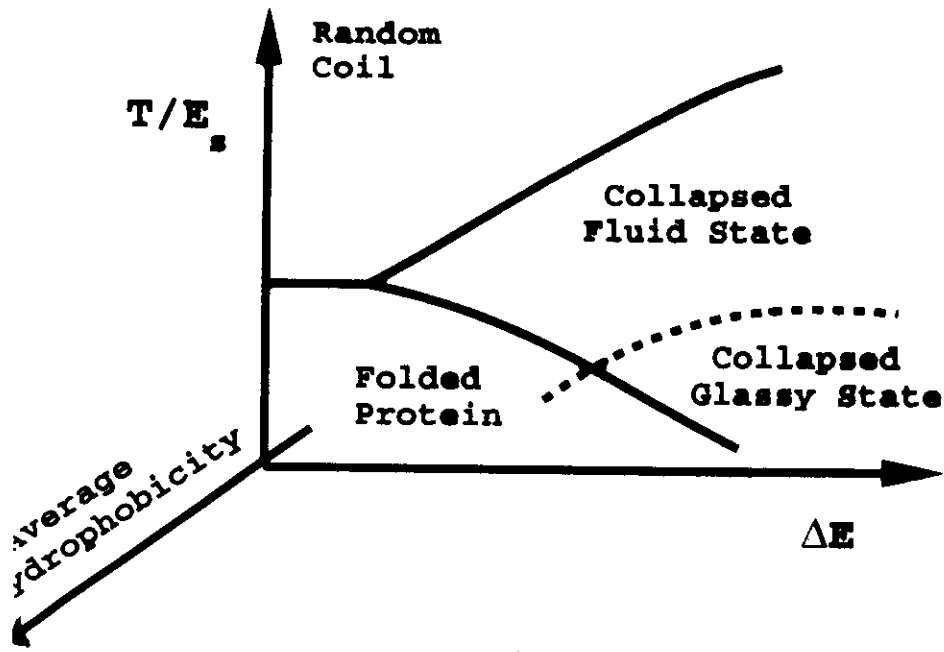


NDSocci

Figure 6: Comparison of mean first passage times from monte carlo simulations with folding times calculated using the Kramers formula.



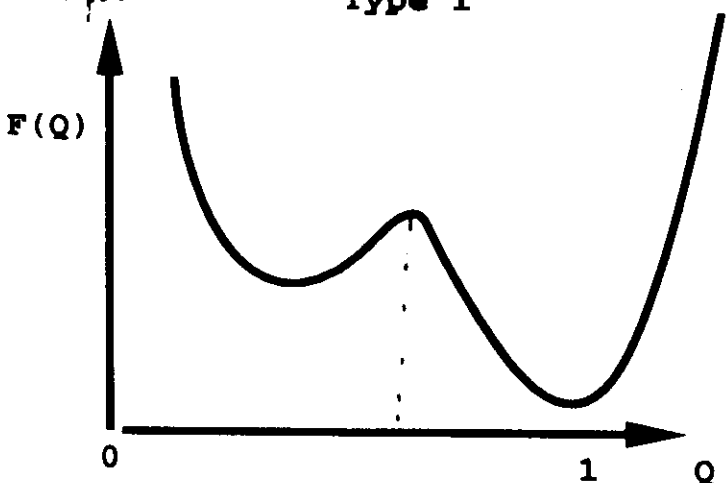
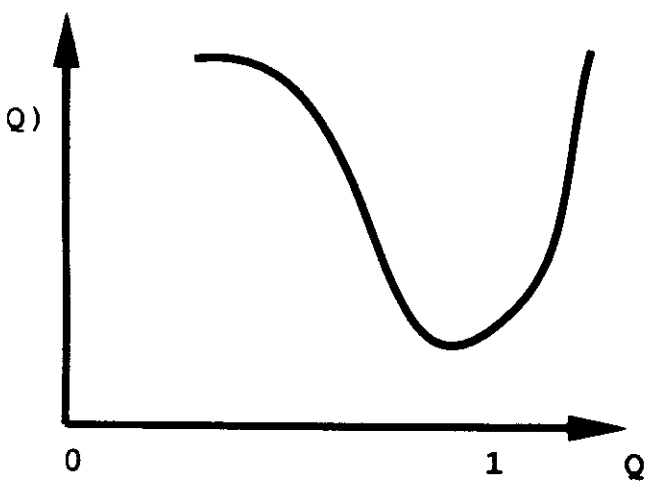
Scenarios for Protein Folding



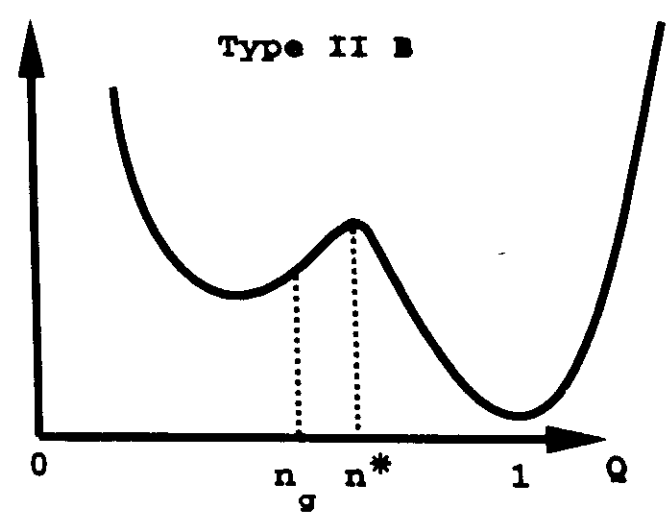
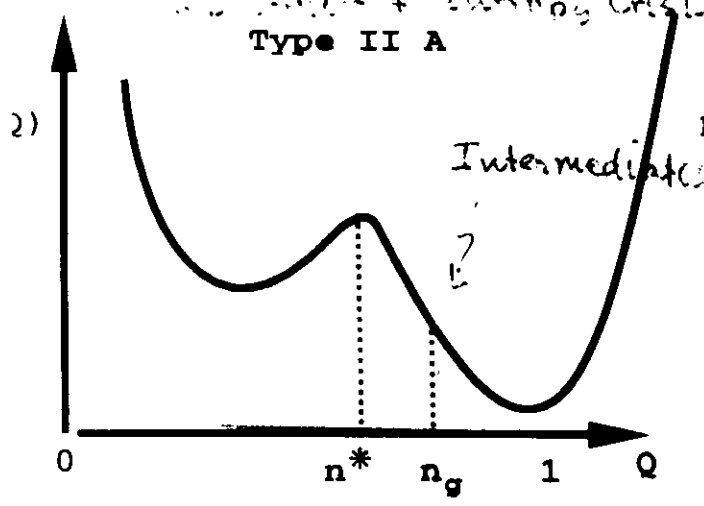
Changing solvent, pH, T shifts scenarios

Type 0 fold ~ collapse

"Barrier" Type I $t_f \approx D(n^*) e^{\beta \Delta F^\ddagger}$



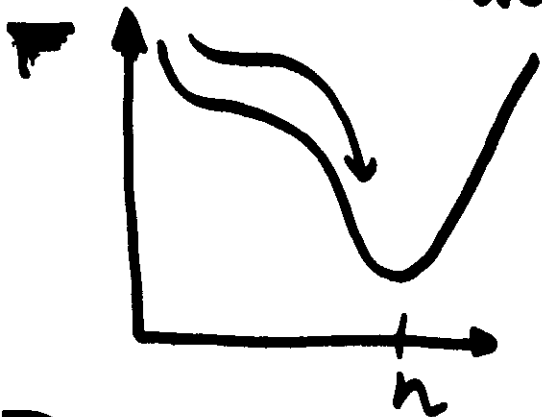
"Intermediate + Substrate Crisis" Type II A



Scenarios for Protein Folding

JSB/MSJ
Mammi/MS

Case 0. "Downhill" Folding



A. $T > T_g(n) \forall n$

$\tau_{Fold} \approx \tau_{Collapse}$
at Roughly similar times

B. $T > T_g(n^*)$ for $n^* < 1$

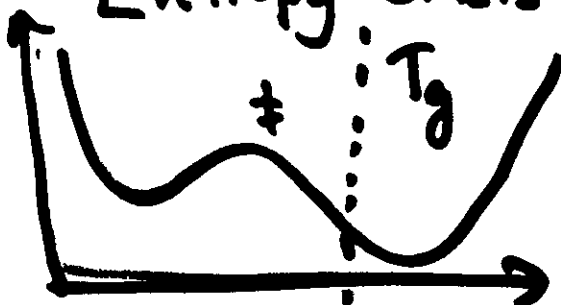
slow collapsed search $\tau_{Fold} \approx \tau_{Levintal}(n^*)$
"Off pathway" traps

Case I. Barrier for Folding: Nucleation



$\tau_p \approx D(n^*) e^{+\beta(F^\ddagger - F)}$
 $T > T_g(n^*), \forall n$

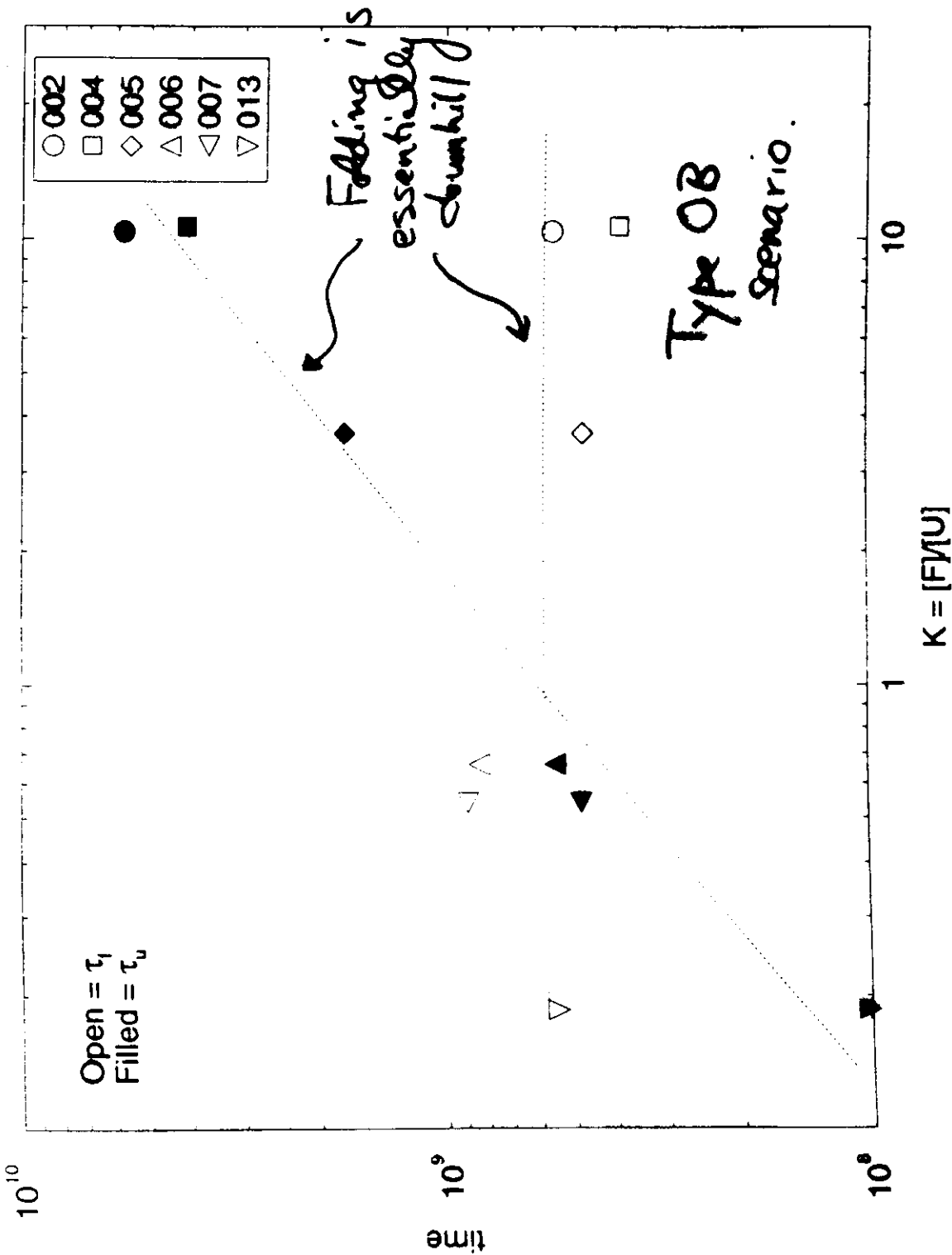
Case II. Barriers for Folding and
Entropy Crisis



Nucleation, and
very Slow Growth
= Intermediates, "Unfolded"

τ_f, τ_u VS K

T = 1.00

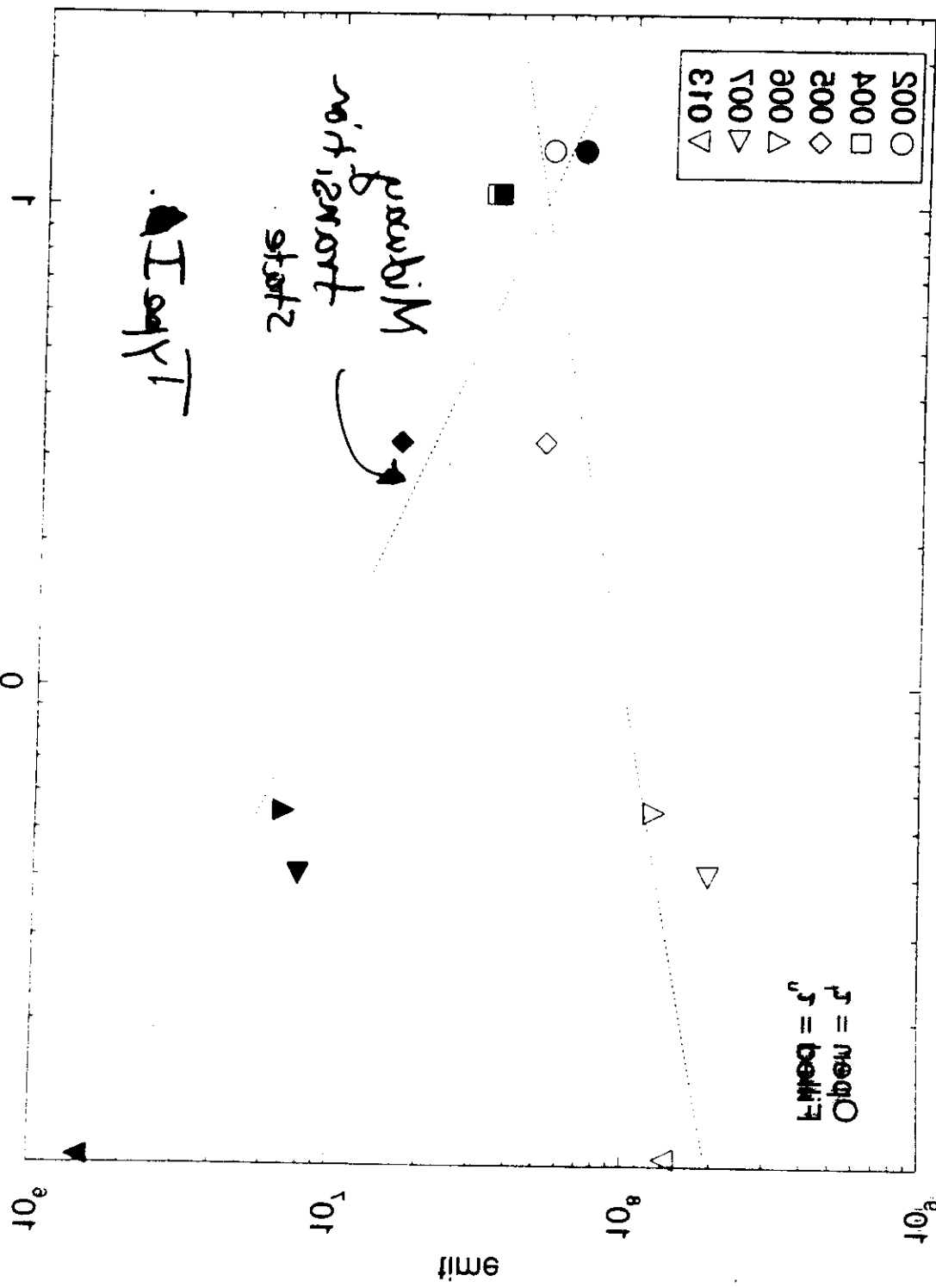


Socci & Onuchic
UCSD

Thu Jun 2 15:45:27 1994

$\lambda_{2V} \mu_{J,q} J$

85.1 = T



UCCSD & Onurpic

49er 85:04:ar 5 nul. nrt

Requirements of Tunnel Geometry

I. Levinthal entropy is Geometrical.



Observed MG $\sim 65\%$ helical

$$\left. \begin{array}{l} \epsilon_{MG}/T_F \\ \downarrow \\ S_L(T_F) = .65 k_B \end{array} \right\} \Rightarrow \begin{array}{l} 27 \text{ cube} \\ \text{"} \\ \sim 50 \text{ residue} \\ \text{protein.} \end{array}$$

II. Ruggedness follows from fluctuation time scale
MG's can fold in msec.

$$\tau_F > \tau = \tau_{RZ} \frac{\Delta E^2}{k_B T^2}$$



Helix-Coil Rouse-Zimm time 10^{-9} sec

$$\frac{\Delta E^2}{2T_F^2} \approx 15$$

III. Slope follows from Thermodynamic Stability

$$\frac{\delta E}{-\frac{\Delta E^2}{T_F^2}} = T_F S = T_F \left(S_L - \frac{\Delta E^2}{2T_F^2} \right)$$

$$\frac{\delta E_s}{T_F} \approx 50$$

$$\Rightarrow \frac{\delta E_s}{\Delta E} = 10 \Rightarrow \frac{T_F}{T_0} \approx 1.6$$

Then
Cubic

Postulate A: Probability distribution of misfolded states is gaussian

$$P(E)\delta E = \frac{1}{\sqrt{2\pi\Delta E^2(Q)}} \exp\left\{-\frac{(E - \bar{E}(Q))^2}{2\Delta E^2(Q)}\right\} \delta E \quad (1)$$

Density of conformational states with energy E and similarity Q :

$$\Omega(E, Q) = \Omega_0(Q)P(E) \quad (2)$$

Total entropy:

$$S(E, Q) = S_0(Q) - k_B \frac{(E - \bar{E}(Q))^2}{2\Delta E^2(Q)} - k_B \log\left(\frac{\sqrt{2\pi\Delta E^2(Q)}}{\delta E}\right) \quad (3)$$

Most Probable Energy:

$$E_{m.p.}(Q) = \bar{E}(Q) - \frac{\Delta E^2(Q)}{k_B T} \quad (4)$$

Corresponding Entropy:

$$S(E_{m.p.}, Q) = S_0(Q) - \frac{\Delta E^2(Q)}{2k_B T^2} \quad (5)$$

Free-energy of the misfolded structures with configurational similarity Q and at a fixed temperature:

$$\begin{aligned} F(Q, T) &= E_{m.p.}(Q) - TS(E_{m.p.}, Q) \\ &= \bar{E}(Q) - \frac{\Delta E^2(Q)}{2k_B T} - TS_0(Q) \end{aligned} \quad (6)$$

At the folding temperature T_f , $S_{native} = 0$ and $F_{native} = F(Q_{min}, T_f)$,

$$T_f = \frac{\delta E_s + \sqrt{\delta E_s^2 - 2S_0 \Delta E^2(Q_{min})/k_B}}{2S_0} \quad (7)$$

Local glass transition in each stratum is characterized by an entropy crisis where $S(T_g, Q) = 0$

$$\begin{aligned} T_g(Q) &= \sqrt{\frac{\Delta E^2(Q)}{2k_B S_0(Q)}} \\ \frac{T_f}{T_g} &\approx 2\sqrt{\frac{k_B}{2S_0} \left(\frac{\delta E_s^2}{\Delta E^2}\right)} \end{aligned} \quad (8)$$

For a protein to fold, T_f/T_g must be greater > 1 .

Figures

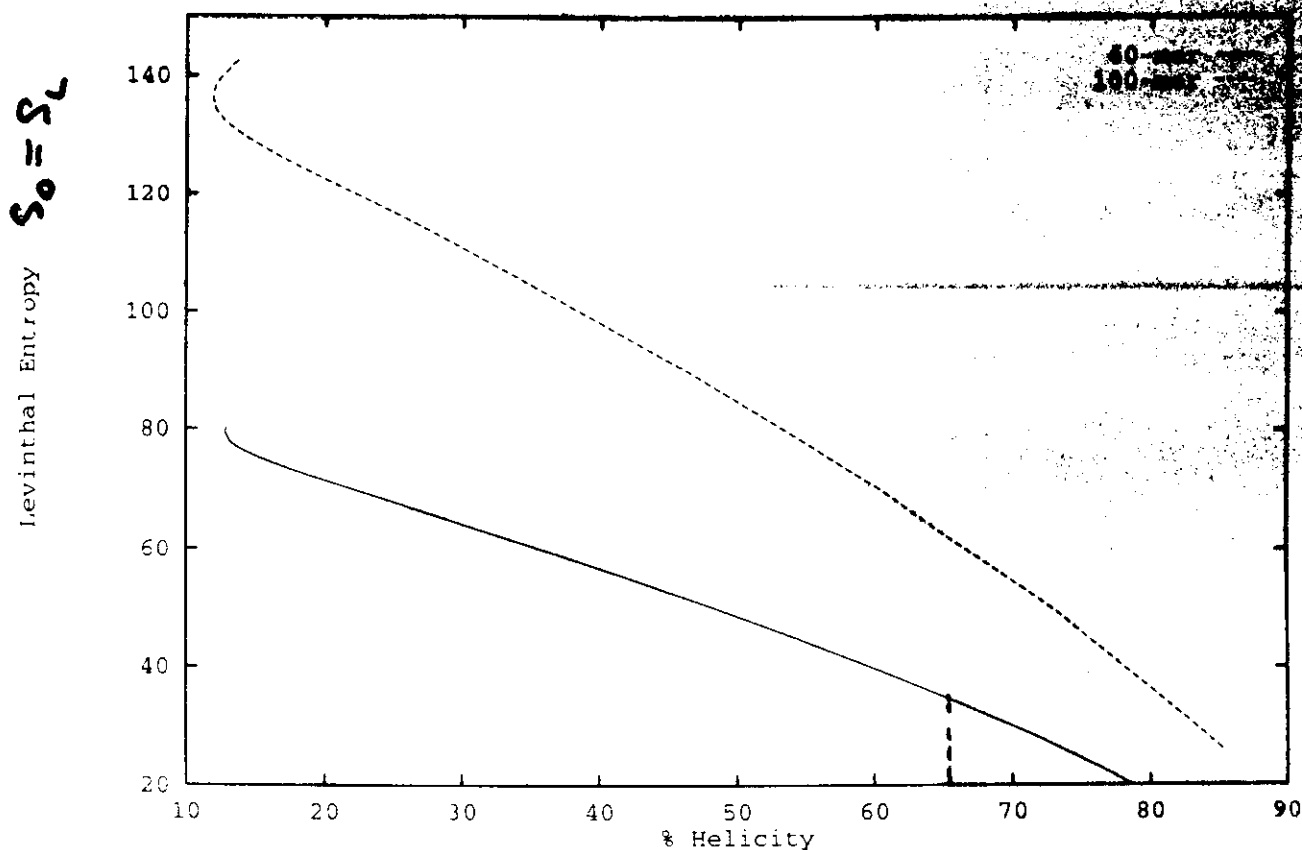
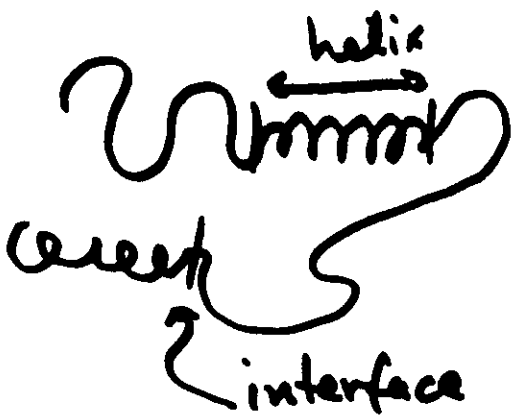


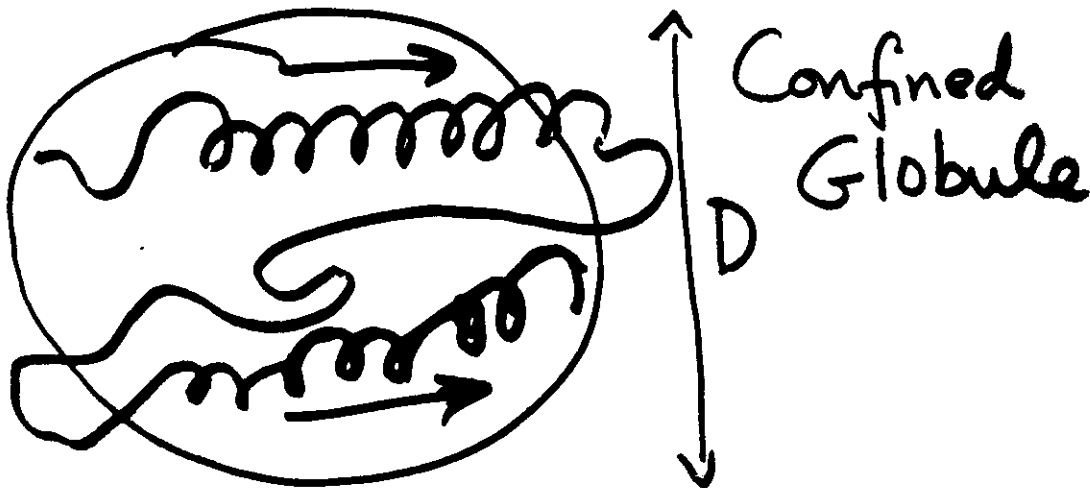
Figure 1: Configurational (Levinthal) entropy versus helicity according to the theory of Luthey-Schulten et.al. Both quantities depend parametrically on the effective hydrogen bond strength divided by $k_B T$.

More Order Parameters for Protein

Z. Lichtig Schultz
B. Ramirez (CIT)
PGW



$$E^{(2)} = -2\epsilon_{H_3} N^H + 4\epsilon_{H_2} N^I$$

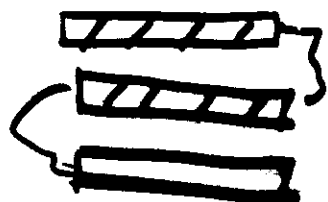


Liquid
Xtallinity

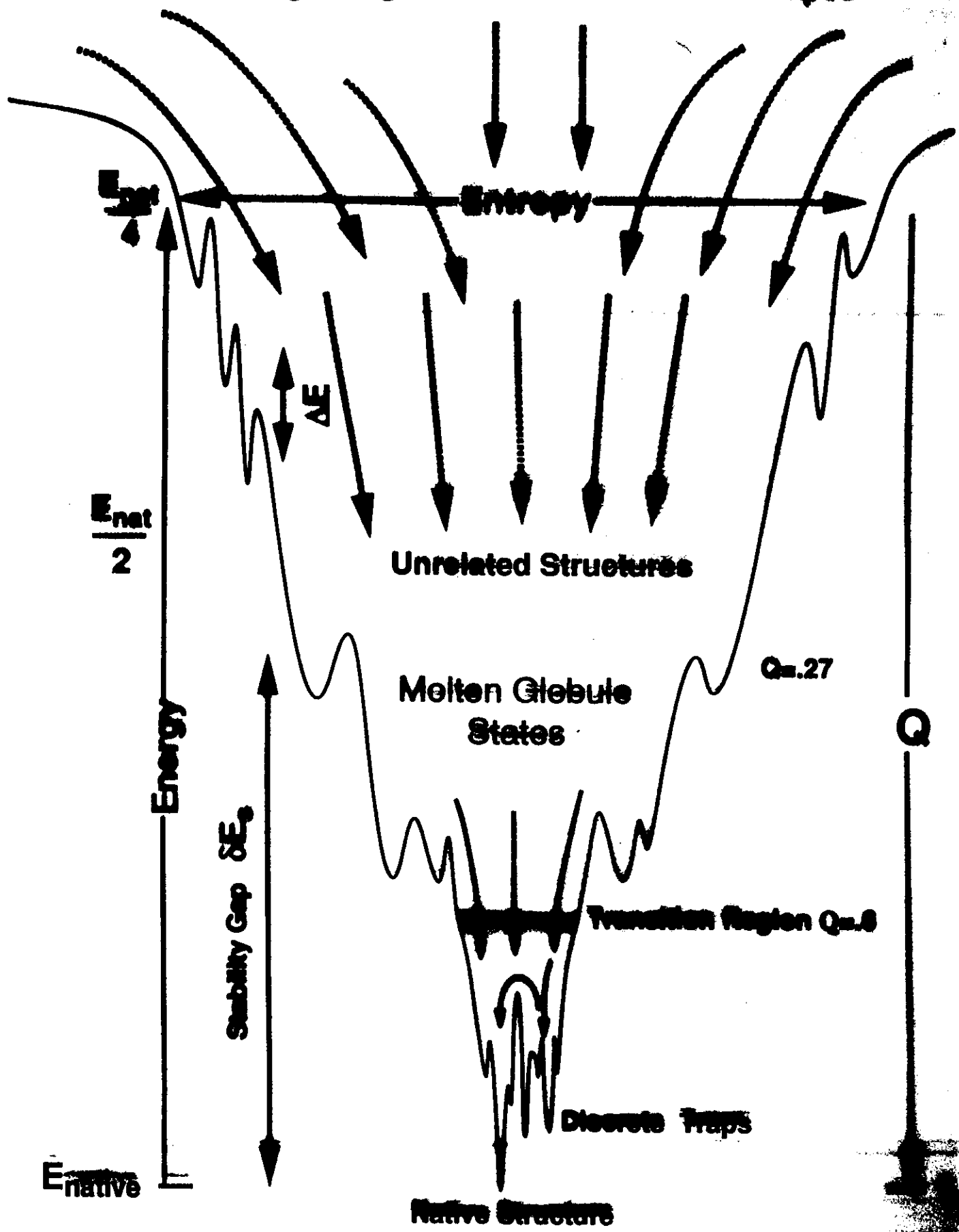
$$\Omega(n_H, n_I; D) = \mathcal{C}(N_H, N_I) \sum_C^{N_C} \sum_I^{(2N_I/2)} \dots$$

$$\exp\left\{ - \frac{\pi^2 R_0^2}{6 D^2} - (N_C + \alpha \frac{2N_H}{N_I}) \left[\frac{(1-\eta)}{\eta} \frac{R_H^2}{R_0^2} + 1 \right] \right\}$$

$$R_0^2 = N_C R_C^2 + \frac{2N_H^2}{N_I \cdot 2} R_H^2 \Rightarrow 2 \dots \text{structure} \dots \text{fractal}$$



Beginning of helix formation and collapse



How Smooth are Real Protein Folding Funnel

{ Onuchic
Wolynes
Luthy-Schulz
Soci

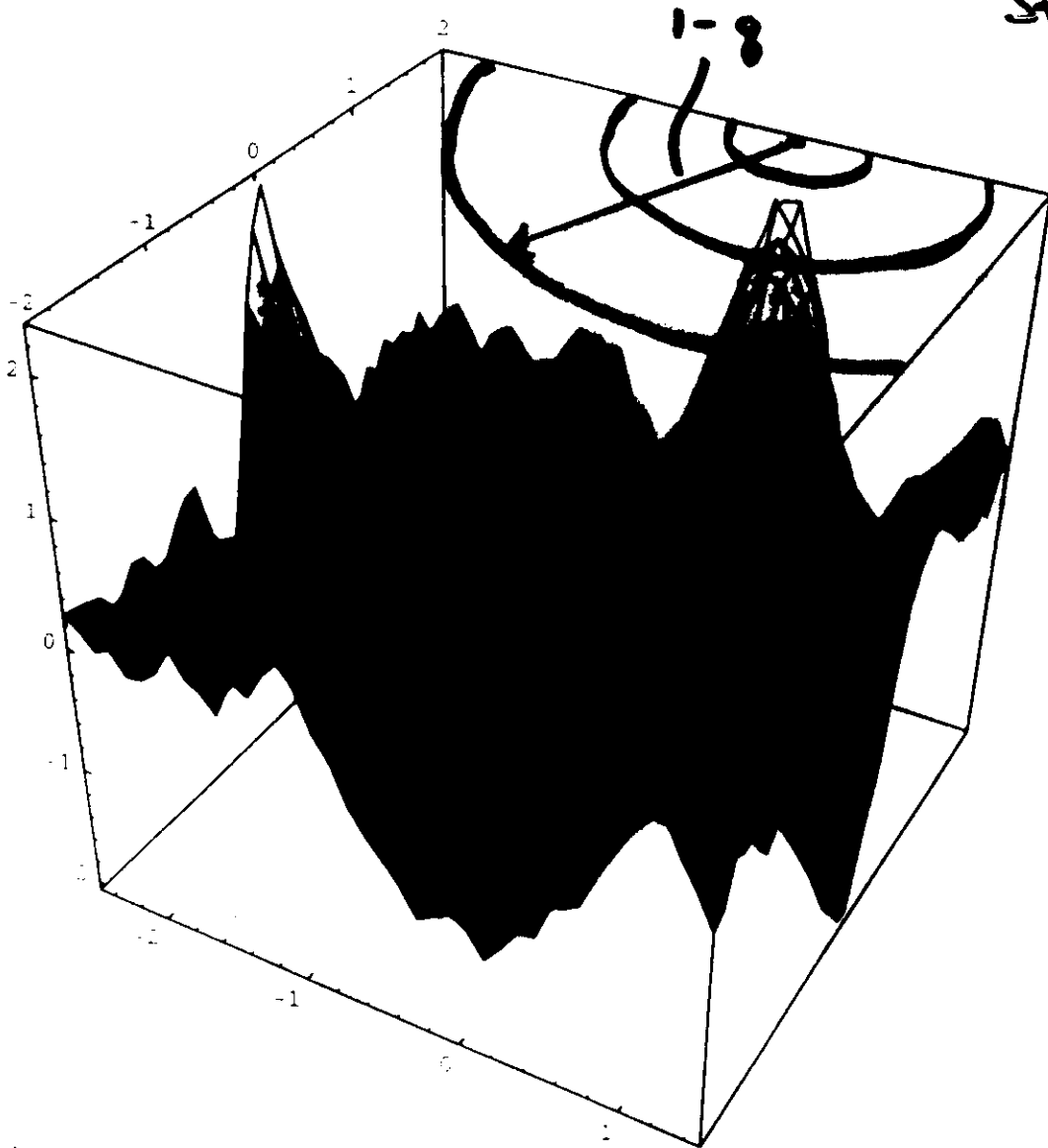
What is the law of corresponding states for proteins? What are the characteristic parameters?

1. The size of the Landscape \propto Levinthal Entropy
2. Ruggedness of Landscape ΔE^2
3. The slope of the Funnel ΔE_s

Estimates can be made for helical proteins using theory \equiv "Quasi-static" Experiment.

Correlated Energy Landscapes for Random Heteropolymers

Jim Wang
Jup Sun
Steve Plotkin
P&J.



Polymer physics determines

$$\Omega(q) = e^{S(q)/k_B}$$

$$\Delta E^2(q) = \frac{\Delta E^2(1-q)}{1-q}$$



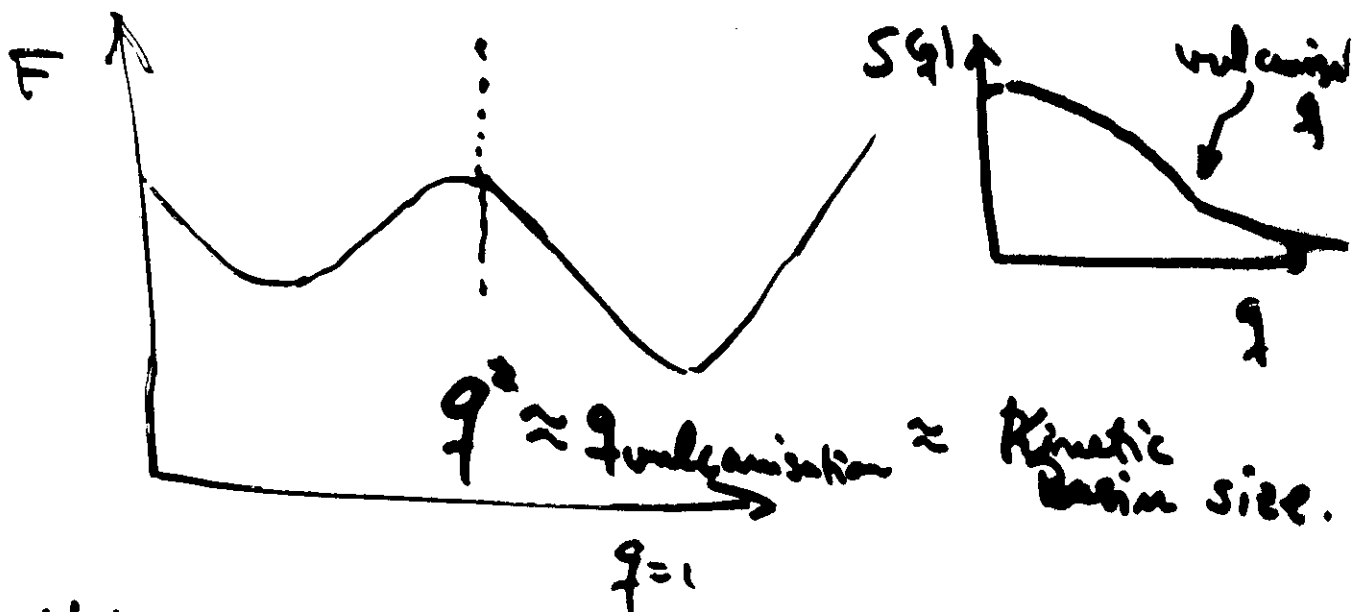
Generalized Random Energy Model
à la Derrida, Gardner

RHP-GREM and Geometry of Attraction Basin

$$P(E_f, E, q) = e^{-\left[\frac{(E_1 + E_f)^2}{8\Delta^2(1+q)} + \frac{(E_1 - E_f)^2}{8\Delta^2(1-q)} \right]}$$

$$\Omega(q) = e^{S(q)/k_B}$$

$$F(q) = q E_f - \frac{\Delta E^2 (1-q^2)}{8k_B T} - T \left[S(q) - \frac{2E_f^2}{\Delta E^2} \right]$$



① Notice in GREM, ^{selected} minimally frustrated polymers are those with low E_f !

② Hopping rate modified, but still gets small

③ Levinthal problem is smaller but still open

Toward an outline of the topography of a realistic protein-folding funnel

J. N. ONUCHIC*, P. G. WOLYNES†, Z. LUTHEY-SCHULTEN†, AND N. D. SOCCI*

†School of Chemical Sciences, University of Illinois, Urbana, IL 61801; and *Department of Physics-0319, University of California, San Diego, La Jolla, CA 92093-0319

Proc. Natl. Acad. Sci. U.S.A.
Vol. 92, pp. 3626–3630, April 1995
Biophysics

Toward an outline of the topography of a realistic protein-folding funnel

J. N. ONUCHIC*, P. G. WOLYNES†, Z. LUTHEY-SCHULTEN†, AND N. D. SOCCI*

†School of Chemical Sciences, University of Illinois, Urbana, IL 61801; and *Department of Physics-0319, University of California, San Diego, La Jolla, CA 92093-0319

Contributed by P. G. Wolynes, December 28, 1994

ABSTRACT Experimental information on the structure and dynamics of molten globules gives estimates for the energy landscape's characteristics for folding highly helical proteins, when supplemented by a theory of the helix-coil transition in collapsed heteropolymers. A law of corresponding states relating simulations on small lattice models to real proteins possessing many more degrees of freedom results. This correspondence reveals parallels between "minimalist" lattice results and recent experimental results for the degree of native character of the folding transition state and molten globule and also pinpoints the needs of further experiments.

Recently a framework for understanding biomolecular self-organization using a statistical characterization of the free-energy landscape of protein molecules has emerged (1–5). Based on the physics of mesoscopic, disordered systems, it can capitalize on the ability to simulate "minimalist" models of proteins, to characterize the folding mechanism through a few energetic and entropic parameters describing the free-energy surface globally. The energy landscape of a foldable protein resembles a many-dimensional funnel with a free-energy gradient toward the native structure. The funnel is also rough, giving rise to local minima, which can act as traps during folding. Most random heteropolymers have numerous funnels to globally different low-energy states just as do glasses and spin glasses. The search through the energy minima of a rough landscape is slow and becomes more difficult as the glass transition is approached. Typically a random heteropolymer will not fold to its lowest free-energy minimum in times less than that needed to explore completely the configuration space if there were no barriers. This supposed difficulty for a natural protein has been called the Levinthal paradox (5). For most random heteropolymers, the search problem of the Levinthal paradox is real, but the guiding forces engineered by molecular evolution can overcome the Levinthal paradox provided they are strong enough, in accordance with the "principle of minimal frustration" (1–4). Most simply, the landscape of a protein funnel is characterized by three parameters: the mean square interaction energy fluctuations, ΔE^2 , measuring ruggedness; a gradient toward the folded state, δE ; and an effective configurational entropy, S_L , describing the search problem size. Our goal here is to use experiments, theory, and simulations to estimate these topographic parameters that determine the folding mechanism. Bryngelson *et al.* (5) classify several regimes of folding. In part of the protein's phase diagram, folding is entirely downhill in a free-energy sense; i.e., as the ensemble of intermediate structures becomes progressively more native-like, the energy gradient completely overcomes the entropy loss. This occurs for folding funnels with very large δE , and is called type 0 folding. Under thermodynamic conditions near the folding transition midpoint, entropy and energy do not completely compensate each

other; thus, intermediates are not present at equilibrium (i.e., a free thermodynamic energy barrier intercedes). In a type I transition, activation to an ensemble of states near the top of this free-energy barrier is the rate-determining step. Type I transitions occur when the energy landscape is uniformly smooth. When the landscape is sufficiently rugged, in addition to surmounting the thermodynamic activation barrier, at an intermediate degree of folding, unguided search again becomes the dominant mechanism. At this point, a local glass transition has occurred within the folding funnel. If the glass transition occurs after the main thermodynamic barrier, the mechanism is classified as type IIa. Here specific kinetic intermediates occur late in folding but are native-like. On the other hand, if the glass transition occurs before the main thermodynamic barrier, intermediates are misfolded traps, which can be described by a multistep chemical kinetic scheme. The details of this type IIb mechanism are very sensitive to the thermodynamic state, interaction potentials, and the specific sequence.

Starting with Levitt and Warshel (6), a variety of simple models of protein folding called minimalist have been developed. Recent studies with continuum models by Honeycutt and Thirumalai (7) and others have been interpreted using energy landscape ideas. Another class of minimalist models studies folding of heteropolymers on a lattice using Monte Carlo kinetics (4, 8–11). Both these studies and the exact enumeration schemes pioneered by Dill and co-workers (12) provide a characterization of the energy landscape for such minimalist models. The simplicity of these models strikes some as being terribly unrealistic, since real proteins possess many details not present in most minimalist models such as hydrogen-bonded secondary structure and side chain conformational degrees of freedom important for packing. Each of these has a different energy scale. Can these features be at all taken into account when making the connection between minimalist models and experiments on real proteins without studying highly complex models?

The energy landscape philosophy and the analogy to phase transitions provide the key. The broad mechanism of phase transitions depends only on gross features of the energy function. When appropriately scaled, the part of the phase diagrams relevant to boiling of liquids as disparate as water and xenon can be superimposed. At the empirical level, this remarkable similarity is known as the law of corresponding states (13), and it is also the basic idea of the renormalization group (14).

The energy landscape picture suggests that there is also a law of corresponding states mapping the phase diagram and kinetic mechanisms of real proteins onto those for minimalist models. If separate phase transitions for ordering the additional degrees of freedom possessed by real proteins intervene during folding, a multistep mechanism can still result, but in a major part of the more complex phase diagram, the effect of the extra degrees of freedom in real proteins will be to "renormalize" energy and entropy scales for the protein-folding funnel.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Here we explore a correspondence of real proteins with minimalist protein-folding models that uses an analytic theory of helix-coil transitions in collapsed heteropolymers to effectively renormalize out secondary structure formation. When combined with experimental measurements of the amount of secondary structure, the theory quantifies the effective number of degrees of freedom of a helical protein through the configurational entropy. Dynamical measurements on the molten globule state crudely characterize the energy landscape ruggedness. The funnel's slope is then inferred using the thermodynamics of the molten globule to folded transition.

The reduction of configuration entropy through helix formation in the collapsed state yields an energy landscape comparable in extent or complexity with that for minimalist models with fewer residues but lacking explicitly secondary structure. The corresponding energetic topography for an optimized three-letter code minimalist lattice model roughly corresponds with the energy gradient and ruggedness of a realistic folding funnel.

The parallel between the gross features of the landscape of real proteins and the three-letter code lattice models allows us to quantify aspects of the folding mechanism in real proteins by using computer simulations of the model. By simulating many folding trajectories and characterizing the free energy as a function of several order parameters, we can identify the location of the relevant thermodynamic free-energy barrier, which is rather small, and determine the position of glass transition within the folding funnel. While the broad transition state occurs early, the glass transitions occur rather late in the folding processes for this model. At the denaturation midpoint, folding occurs via a type **IIa** scenario but is rather close to the downhill type **0b** scenario. The details of a protein's folding after the glass transition late in the funnel cannot be studied using the corresponding states principle, but the earlier events can.

Establishing the Correspondence Between Minimalist Models and Real Proteins

Collapsed states have been established as rather general intermediates in folding (15). Some compact intermediates contain a substantial percentage of helical secondary structure. At least two views of the collapsed states are prevalent. Some argue that the molten globule state has a specifically defined tertiary structure comparable to the native protein. Others view the equilibrium collapsed state as still conformationally fluid in terms of the backbone structure resembling a polymer below its θ point (16). Both Kallenbach and co-workers (17) and Engelman and co-workers (18) have found compact states with varying degrees of helical structure, thus suggesting its lability. The two pictures may not be so clearly separable since the guiding forces of the funnel do induce a significant amount of fluctuating tertiary order in the disordered globule. The relevant collapsed states are the dynamic ones of the early stages of folding and not necessarily the equilibrium states found elsewhere in the phase diagram (e.g., the "acid molten globule").

The degree of helical content of equilibrium collapsed states has often been measured to be quite high. The relation between helicity and the conformational entropy shown in Fig. 1 can be found from the theory developed by Luthey-Schulten *et al.* (19), since both depend parametrically on the effective hydrogen bond energy. The theory due to Basile *et al.* (20) could also be used with appropriate modification. Taking 65% helicity as a reasonable estimate, our theory gives a conformational entropy of $0.6k_B$ per monomer unit. The effective number of states to be searched is related to this Levinthal entropy, e^{S/k_B} . Though the states differ in character, the number of states in the mechanism here is comparable to that for the framework model (ref. 21 and references therein; ref.

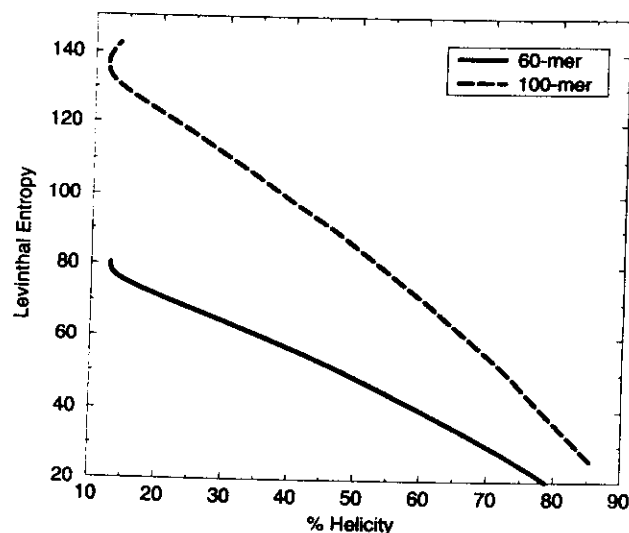


FIG. 1. Configurational (Levinthal) entropy versus helicity according to the theory of Luthey-Schulten *et al.* (19). Both quantities depend parametrically on the effective hydrogen bond strength divided by $k_B T$.

22). Diffusion-collision calculations assume only the correct helices can be formed and direct construction of the fold from the high entropy random coil. A free chain has an entropy of $2.3k_B$ per monomer unit (23). Unlike the framework picture, the dramatic reduction in entropy upon collapse arises from confinement, indiscriminate helix formation, and the constrained orientation of the helices, not from strong local biases toward correct secondary structure. The molten globule from which further guided searches take place in our mechanism is a collapsed liquid-crystalline polymer. The compact configurations of simple lattice model polymers have an entropy of $\approx 1.0k_B$ per unit considering only the reasonably compact states after fast collapse. Thus the renormalized entropy or scope of configuration space of a 60-amino-acid helical protein is a bit bigger than the 27-mer lattice model often studied.

Comparing the dynamics of free and collapsed chains yields the ruggedness of the landscape. In a free chain, flickering secondary structural elements reconfigure in roughly $\tau_0 = 1$ nsec (24). This time is similar to interdiffusion times over distances of the size of the molten globule diameter as calculated using the Rouse-Zimm theory (25). Thus, to first order, we can be agnostic as to the nature of the underlying move set in comparing real proteins and minimalist models. The dynamics of a condensed molten globule is slower than that for free chains because of transient trapping in low-energy states. The reconfiguration time $\tau_{reconfig}$ in a rough energy landscape (5, 26) is given by $\tau_{reconfig} = \tau_0 \exp(\Delta E^2/2T^2)$.

Few experiments directly measure reconfiguration times within the globule. For lactalbumin, Baum *et al.* (27) observe field-dependent broadening of 1H NMR resonances, suggesting reconfiguration rates slower than 1 per millisecond. Wand and co-workers (28) interpret their NMR studies on the apocytochrome b_{562} molten globule with similar times. In the fastest folding, the downhill scenario (type **0**), folding takes only a few times the typical reconfiguration time. Thus an upper bound on the ruggedness is known since collapsed states can completely fold in times ranging from a millisecond to a second. These estimates for $\tau_{reconfig}$ suggest the ruggedness of the energy landscape, at the folding temperature, $\Delta E^2/2T_f^2$, ranges from 11 to 18. The typical size of hydrophobic forces needed for protein collapse gives directly a similar estimate (29). The actual entropy of the molten globule state is lower than S_L , since low-energy states are preferentially occupied. For the 60-amino-acid chain at 60% helicity, the random energy model gives an entropy $S(T) = S_L - \Delta E^2/2T^2 = 21k_B$ to $28k_B$. At the folding transition, the energy loss in falling

down the funnel must equal the temperature times the entropy loss. Thus the stability gap or energy gradient of the funnel is $\delta E_s/T_f = S_L + \Delta E^2/2T_f^2$. The stability gap δE_s measures the difference in energy of the native protein and the average compact state. The dimensionless ratio of the energy gradient of the funnel to the overall ruggedness is then $\{\delta E_s/T_f\}/\{\sqrt{\Delta E^2/2T_f^2}\} \approx 14$. Using the configurational entropy estimate, the thermodynamic glass transition temperature is $T_g = \Delta E/\sqrt{2S_L} \approx 0.6T_f$. If interaction strengths were temperature independent, the thermodynamic glass transition temperature for a compact denatured state would be 160 K. The dynamical glass transition in folded myoglobin actually depends strongly on solvent and occurs in glycerol at 180 K (30). The coincidence might support the ideas of Frauenfelder and Wolynes (31) and of Honeycutt and Thirumalai (7) that some taxonomic substates of folded proteins correspond with the final protein folding intermediates.

Comparing the estimate of gradient-to-ruggedness ratio with the 27-mer simulations shows that the landscape is smoother than landscapes generated for optimally designed sequences using a two-letter code. The ratio between T_f and the kinetic T_g (relevant to the folding time scale) is about 1.3 for the two-letter code sequences (32). The thermodynamic glass temperature for the collapsed states of the two-letter code 27-mers calculated using the random energy model estimate is close to this kinetic T_g . Two-letter code lattice models in the bulk limit usually exhibit ground-state degeneracy, probably connected with microphase separation (33). Yue *et al.* (34) suggest that designing foldable two-letter code proteins is nontrivial.

Simulations have been performed with three-letter codes—i.e., strong interactions for residues of the same kind and weak interactions for different. When the values of the couplings are the same as for the two-letter code, an optimized three-letter code folded configuration still has only correct strong contacts, but most three-letter code compact configurations have fewer wrong strong contacts than for the two-letter code. Optimized three-letter code proteins have a larger T_f and a smaller kinetic T_g than two-letter ones. The T_f/T_g ratio increases to 1.6, close to the ratio for realistic folding funnels. The mechanisms of folding of the two- and three-letter code results differ since the two-letter code model is closer to its global glass transition.

The Folding Scenario for a Realistic Folding Funnel

The corresponding state analysis allows us to sketch the following folding scenario, based on the three-letter code simulations, and to picture a folding funnel whose main features are shown in Fig. 2. We have rendered this folding funnel reasonably accurately to scale. The width is a measure of the entropy, whereas the depth is illustrated with both an energy and two correlated structural scales. Although no one-dimensional scale reflects properly the multidimensionality of the funnel and the multiple minima, the barrier heights in the figure represent ΔE , whereas the total depth is scaled to the energy of the folded state. The molten globule region, a transition state region representing an ensemble of structures that acts as a bottleneck, and a *locally* glassy region are identified based on detailed examination of many folding trajectories for the three-letter code model coupled with numerical measurements of density of states, free energies, and related quantities. Defining several collective coordinates compresses much information about the trajectories into a simple form. Our characterization of the transition state region for a realistic folding funnel differs from the results of Sali *et al.* (35), which apparently model proteins near the border of kinetic foldability.

Two coordinates examined are ones like that of Bryngelson and Wolynes (26), the fraction of angles in their native configuration, A , and the fraction of native contacts, Q . The

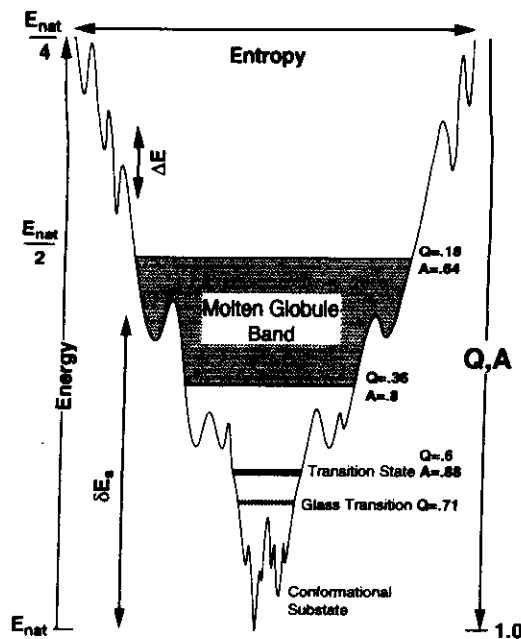


FIG. 2. The schematic funnel for a realistic 60 amino acid helical protein corresponding to the three-letter code. This shows the position of the molten globule, the transition state ensemble, and the local glass transition where discrete trapping states emerge as a function of the order parameters described in the text, the energy E , the fraction of native contacts Q , and the fraction of angles in their native configurations A . Q and A have been normalized to their maximal values for a 27-mer lattice model, 28 and 25, respectively. For the three-letter code, the molten globule is stabilized by nearly half the native energy (E_{nat}) relative to the random coil. The stability gap δE_s quantifies the specificity of the native contacts.

reaction coordinate A only changes by a small amount on each elementary step, so local gradients of free energies are meaningful. For a random coil, the value of A should be ≈ 0.68 . Q is intimately connected with the interaction energy function and is useful in describing overall topology. Care must be used in interpreting gradients of free energies with respect to Q since the elementary moves can lead to large Q changes. Both coordinates refer to overall structure features. For larger heteropolymers, additional coordinates describing distinct parts of the chain are needed to define critical nuclei (36, 37) for large single-domain proteins or independent folding of parts of multidomain proteins.

Time series of the reaction coordinates and interaction energy are two-state-like with fast transitions between the folded and unfolded regions. The mean folding time for this sequence of $\approx 3 \times 10^6$ time steps corresponds to 3 msec of real time using the estimates for τ_0 . As for simple reactions, the rate for transition between the two main regions depends on short time events.

The duration of a transition event fluctuates, but most events are over in less than 10,000–50,000 time steps. Q and A vary in a correlated manner through the transition. Sometimes they quickly traverse between the stable regions, whereas in other cases they are transiently trapped during the crossing. The duration of the trapping events is much shorter than the average folding time, indicating two-state kinetics.

The Monte Carlo histogram technique (38) was used to determine the density of states as a function of energy and the order parameters. This technique is similar to that used by Hansmann and Okamoto (39) and Hao and Scheraga (40) to determine overall thermodynamics. The densities of states yield the free-energy plots, precisely locate the folding temperature, and determine a local thermodynamic glass transition region where discrete intermediates appear.

At T_f , the free-energy function for the optimal three-letter code 27-mer is plotted as a function of Q and A . Projected onto these plots are two illustrative trajectories out of 86 examined (Fig. 3). The free-energy function is bistable. The disordered globule has $Q \approx 0.28$ and $A \approx 0.73$. The globule has a much greater amount of native structure (Q) than that expected for a random coil, but clearly an enrichment of pair contacts by a factor of 10 does not by itself imply a "unique" structure for a molten globule (41). The two-dimensional free energy has a saddle at $A \approx 0.88$ and $Q \approx 0.6$. The Q value of 0.6 means that each native contact is made three-fifths of the time in an ensemble of configurations of the transition state. This is in harmony with recent observations on chymotrypsin inhibitor folding where the transfer coefficient for mutations at each site ϕ varies between 0.3 and 0.7 (42). Since chymotrypsin inhibitor has a good deal of β -sheet as well as a helix, the agreement may be fortuitous. The superimposed trajectories agree with assigning a transition state region encompassing Q values from ≈ 0.57 to ≈ 0.64 and A values from ≈ 0.84 to ≈ 0.92 . Late barriers depending on Q alone are kinetically meaningless, since reactive trajectories jump across such barriers in the Q direction through crankshaft moves in which an entire arm of the protein is retracted into its native position. The A coordinate on the other hand varies only by one or two units per elementary move and is a more appropriate reaction coordinate (1). Because of the flatness of the free energy, the thermodynamic barrier from the free-energy plot is small but broad. There are numerous recrossings of the transition state region caused by the trapping due to the landscape's ruggedness. Thus, as Bryngelson and Wolynes (26) suggest, folding times must be computed using a diffusive picture instead of standard transition state theory, which neglects recrossings.

Monitoring correlated fluctuations of the collected coordinates gives the diffusion constants for Q in the molten globule $D \approx 3.5 \times 10^{-4}$ (correct contacts)² per time step. A crude diffusive rate theory that assumes the free energy well is harmonic and that the barrier top curvature equals the well's gives a folding time $\tau_F = 2\pi\tau_{\text{corr}}\exp\{F^*/k_B T\}$, where F^* is the activation barrier of $2.4k_B T_f$ from the two-dimensional plot and τ_{corr} is the correlation time for the harmonic fluctuations. τ_{corr} for both A and Q is approximately 20,000 time steps. The resulting $\tau_F \approx 1.4 \times 10^6$ is a bit shorter than the simulated

value. Since this system becomes glassy only after the transition region is traversed, landscape ruggedness should be well accounted for by the diffusion picture. The folding time from the Bryngelson–Wolynes (26) approximation is good, but there are more near-ballistic trajectories through the transition region than expected, suggesting the relevance of the frequency dependence of the structural diffusion or of additional geometrical variables.

After leaving the transition region, the protein progresses to become more native-like. Occasionally, the trajectory becomes caught in a few longer lived native-like states whose lifetime is shorter than the average folding time. These discrete states arise from a local glass transition, which can be located by computing $Y = \sum_i P_i^2$, where P_i is the Boltzmann occupation of a microstate. Y measures the inverse number of the thermally accessible states and reveals the replica symmetry breaking of spin glasses (43) and of random heteropolymers (44). To define the local glass transition, we compute $Y(Q)$ using only states with a given value of the coordinate Q . Since the protein is of finite size, $Y(Q)$ never vanishes but instead varies from the inverse density of states at Q up to unity. At T_f a rapid rise of $Y(Q)$ occurs at $Q \approx 0.7$, defining a local glass transition (Fig. 4).

At T_f the transition state region occurs before the local glass transition, so folding conforms to a type **IIa** scenario. Kinetic constraints, which vary from sequence to sequence, are encountered after the transition state is reached for strong folders just as in simulations of Honeycutt and Thirumalai (7) and of Chan and Dill and co-workers (10, 12).

The small size of the thermodynamic barrier as opposed to kinetic barriers from transient trapping suggests that proteins are not just overall marginally stable but that a realistic folding funnel describes a marginally stable system even for intermediate degrees of order—surprisingly much like a system near a critical point. *In vivo*, proteins are not poised at T_f but are stable by several $k_B T$. The additional slope to the funnel's energy gradient should suffice to make folding occur by a downhill type **Ob** scenario where the only intermediates are near native kinetic traps. Since folding does not dramatically speed up with increasing stability once a downhill scenario is reached, perhaps there is no evolutionary drive to greater stability. The combination of marginal stability and proximity

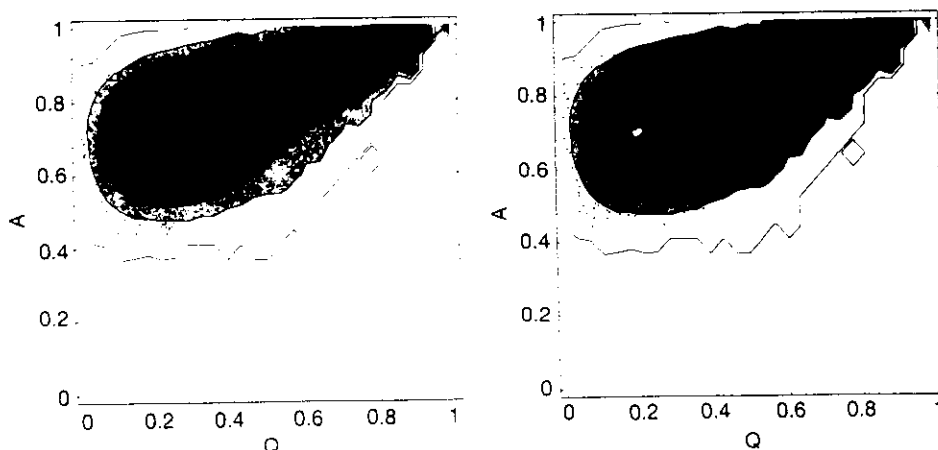


FIG. 3. Two transition trajectories projected onto the Q - A plane. The time span is roughly 25% of the folding time, which is $\approx 3 \times 10^6$ Monte Carlo steps. (Left) The transition event occurs in $\sim 10^5$ Monte Carlo steps. For this trajectory, there is some trapping in the transition region. In the early part of the trajectories, the individual points are not connected, whereas in the latter segments the points are connected. The trajectories are superimposed on a contour plot of the free energy with levels spanning the range from -67.5 to -82.5 in increments of 2.5. (Right) A very fast event in which the system moves almost ballistically through the transition region. The last event occurs in roughly 3000 Monte Carlo steps. The trajectories shown were chosen at random from a sample of 86. The sequence used was ABABBBBCBACBABABACACBACAACAB and was studied at the folding temperature ($T_f \approx 1.509$). The model is a three-dimensional cubic lattice heteropolymer with a contact potential. If the two monomers are of the same type, then the energy for the contact is $E_1 = -3$ and if the monomers are not the same the energy is $E_u = -1$. The above sequence was designed to have an unfrustrated nondegenerate native states; i.e., in the native state all contacts are between monomers of the same type.

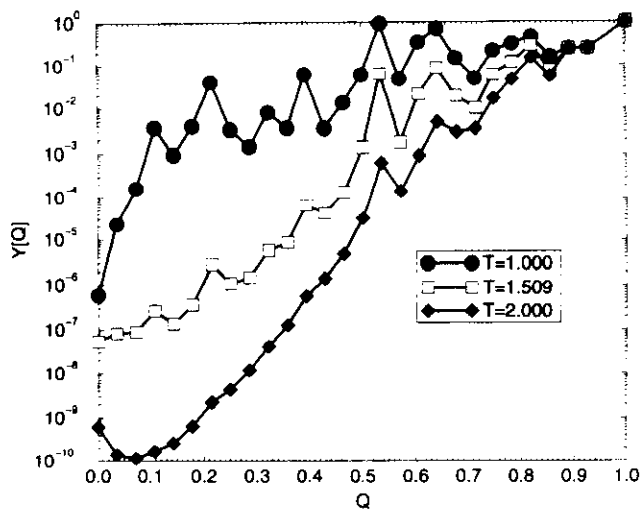


FIG. 4. A plot of $Y(Q) = \sum_i |P_i(Q)|^2$ vs. Q for three temperatures. While at the global T_g discrete states are apparent even for small degrees of nativeness, at $T_f = 1.509$ the discrete intermediate is highly native-like.

to the local glass transition may explain the mutation sensitivity of some collapsed globules (17).

Discussion

Experimental data along with simple geometrically based statistical mechanics help locate small helical proteins in their phase diagram, allowing an estimate of the parameters needed to describe folding by statistical energy landscape analysis. A law of corresponding states relates simple lattice models to the laboratory situation, leading to an outline of the topography of a realistic folding funnel, which can serve as a starting point for other investigations. On the experimental side, our analysis pinpoints a great need for more dynamic measurements on the molten globule state itself, one of the weaker points in the numerical estimates. Also the order of collapse and secondary structure formation still needs resolution. The quantitative features of the funnel should help guide and can be refined by fast folding experiments made possible by laser-induced initiation of folding (45). For the folding funnel of the three-letter code model, the minimal frustration of the protein results from harmony between tertiary contacts. Direct local biases like those in the framework picture (46) can be included as an additional slope to the funnel through A rather than Q . Similarly secondary structure may be more directly coupled to the landscape if effective pair interactions depend specifically on the helicity of segments. These considerations require a still more multidimensional view of the funnel, but the low dimensional picture here can serve as a zeroth order starting point.

The vigorous debates among members of the minimalist folding community including J. Bryngelson, H. S. Chan, K. Dill, E. Shakhnovich, and D. Thirumalai helped us crystallize our thoughts. We also thank W. Eaton and H. Frauenfelder for reading the manuscript. N.D.S. is a University of California at San Diego Chancellor Fellow. This work was supported by the National Institutes of Health (Grant IR01 GM44557), the Beckman Foundation, and the National Science Foundation (Grant MCB-93-16186).

1. Bryngelson, J. D. & Wolynes, P. G. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7524–7528.
2. Goldstein, R. A., Luthey-Schulten, Z. A. & Wolynes, P. G. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 4918–4922.
3. Goldstein, R. A., Luthey-Schulten, Z. A. & Wolynes, P. G. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9029–9033.
4. Leopold, P. E., Montal, M. & Onuchic, J. N. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 8721–8725.
5. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., & Wolynes, P. G. (1994) *Proteins*, in press.

6. Levitt, M. & Warshel, A. (1975) *Nature (London)* **253**, 694–698.
7. Honeycutt, J. & Thirumalai, D. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 3526–3529.
8. Abe, H. & Gó, N. (1980) *Biopolymers* **20**, 1013–1031.
9. Skolnick, J. & Kolinski, A. (1991) *J. Mol. Biol.* **221**, 499–531.
10. Chan, H. S. & Dill, K. A. (1991) *Annu. Rev. Biophys. Biophys. Chem.* **20**, 447.
11. Shakhnovich, E., Farztdinov, G., Gutin, A. M. & Karplus, M. (1991) *Phys. Rev. Lett.* **67**, 1665–1668.
12. Miller, R., Danko, C. A., Fasolka, M. J., Balazs, A. C., Chan, H. S. & Dill, K. A. (1992) *J. Chem. Phys.* **96**, 768–780.
13. Lifshitz, E. M. & Pitaevskii, L. P. (1980) *Statistical Physics* (Pergamon, Oxford), 3rd Ed.
14. Wilson, K. G. & Kogut, J. (1974) *Phys. Rep.* **12**, 75–200.
15. Pitsyn, O. B. (1992) in *Protein Folding*, ed. Creighton, T. E. (Freeman, New York), p. 243.
16. Griko, Y. V., Privalov, P. L., Venyaminov, S. Y. & Kutyschenko, V. P. (1988) *J. Mol. Biol.* **202**, 127–138.
17. Lin, L., Pinker, R. J., Forde, K., Rose, G. D. & Kallenbach, N. R. (1994) *Nat. Struct. Biol.* **1**, 447–452.
18. Flanagan, J. M., Kataoka, M., Fujisawa, T. & Engelman, D. M. (1993) *Biochemistry* **32**, 10359–10370.
19. Luthey-Schulten, Z. A., Ramirez, B. E. & Wolynes, P. G. (1995) *J. Phys. Chem.* **99**, 2177–2185.
20. Bascle, J., Garel, T. & Orland, H. (1993) *J. Phys. (Paris)* **3**, 245–253.
21. Bashford, D., Karplus, M. & Weaver, D. L. (1990) in *Protein Folding*, eds. Gierasch, L. M. & King, J. (Am. Assoc. Advance. Sci., Washington, DC), pp. 283–290.
22. Kim, P. S. & Baldwin, R. L. (1990) *Annu. Rev. Biochem.* **59**, 631–660.
23. Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules* (Wiley, New York).
24. McCammon, J. A. & Harvey, S. C. (1987) *Dynamics of Proteins and Nucleic Acids* (Cambridge Univ. Press, New York).
25. Doi, M. & Edwards, S. F. (1986) *The Theory of Polymer Dynamics* (Oxford Univ. Press, Oxford).
26. Bryngelson, J. D. & Wolynes, P. G. (1989) *J. Phys. Chem.* **93**, 6902–6915.
27. Baum, J., Dobson, C. M., Evans, P. A. & Hanly, C. (1989) *Biochemistry* **28**, 7–13.
28. Feng, Y., Sliagar, S. G. & Wand, A. J. (1994) *Nat. Struct. Biol.* **1**, 30–36.
29. Bryngelson, J. D. & Wolynes, P. G. (1990) *Biopolymers* **30**, 177–188.
30. Frauenfelder, H., Alberding, N. A., Ansari, A., Braunstein, D., Cowen, B., Hong, M., Iben, I., Johnson, J., Luck, S., Marden, M., Mourant, J., Ormos, P., Reinisch, L., Scholl, R., Shyamsunder, E., Sorensen, L., Steinbach, P., Xie, A.-H., Young, R. & Yue, K. (1990) *J. Phys. Chem.* **94**, 1024–1037.
31. Frauenfelder, H. & Wolynes, P. G. (1994) *Phys. Today* **47**, 58–64.
32. Socci, N. D. & Onuchic, J. N. (1994) *J. Chem. Phys.* **101**, 1519–1528.
33. Sfatos, C. D., Gutin, A. M. & Shakhnovich, E. I. (1994) *Phys. Rev. E* **50**, 2898–2905.
34. Yue, K., Fiebig, K., Thomas, P. D., Chan, H. S., Shakhnovich, E. I. & Dill, K. A. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 325–329.
35. Šali, A., Shakhnovich, E. & Karplus, M. (1994) *Nature (London)* **369**, 248–251.
36. Thirumalai, D. & Guo, Z. (1995) *Biopolymers* **35**, 137–140.
37. Abkevich, V. I., Gutin, A. M. & Shakhnovich, E. I. (1994) *Biochemistry* **33**, 10026–10036.
38. Ferrenberg, A. M. & Swendsen, R. H. (1988) *Phys. Rev. Lett.* **61**, 2635–2638.
39. Hansmann, U. H. E. & Okamoto, Y. (1993) *J. Comput. Chem.* **14**, 1333–1338.
40. Hao, M.-H. & Scheraga, H. A. (1994) *J. Phys. Chem.* **98**, 4940–4948.
41. Peng, Z.-Y. & Kim, P. S. (1994) *Biochemistry* **33**, 2136–2141.
42. Otzen, D. E., Itzhaki, L. S., ElMasry, N. F., Jackson, S. E. & Fersht, A. R. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 10422–10425.
43. Mezard, M., Parisi, G. & Virasoro, M. A. (1987) *Spin Glass Theory and Beyond* (World Scientific, Singapore).
44. Shakhnovich, E. & Gutin, A. (1990) *J. Chem. Phys.* **93**, 5967–5971.
45. Jones, C. M., Henry, E. R., Hu, Y., Chan, C.-K., Luck, S. D., Bhunya, A., Roder, H., Hofrichter, J. & Eaton, W. A. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11860–11864.
46. Kim, P. S. & Baldwin, R. L. (1982) *Annu. Rev. Biochem.* **51**, 459–489.

