Biophysical Chemistry, 34 (1989) 187–199
Elsevier

BIOCHE 01380

Formation of unique structure in polypeptide chains

Theoretical investigation with the aid of a replica approach

E.I. Shakhnovich and A.M. Gutin

Institute of Protein Research, Academy of Sciences of the U.S.S.R., 142292 Puschino, Moscow Region, U.S.S.R.

Received 3 January 1989
Revised manuscript received 1 June 1989
Accepted 5 June 1989

Polymer, inhomogeneous; Protein folding; Replica trick; Spin glass; Multiple-minima problem

A replica approach analogous to that used in spin glass systems is implemented to study the configurational space of a heteropolymeric model of protein with a quenched, disordered sequence of links in the limit of a large number of link types. It is shown that there exists a threshold value of chain heterogeneity which separates two qualitatively different types of behavior. For a low degree of heterogeneity the protein globule is like a homopolymer in a collapsed state without definite chain folds: an exponentially large number of folds make a significant contribution to the partition function in this regime. After the threshold heterogeneity has been overcome, the chain freezes drastically but without latent heat; few (approx. 1) frozen states with definite chain folds are thermodynamically dominant in this state. The relation of these results to thermodynamic aspects of protein folding is discussed.

1. Introduction

The problem of protein folding remains a major issue in the field of molecular biology [1]. The general aim of any theory of protein folding is to obtain a prediction scheme which derives the tertiary structure of a protein from the sequence of its links. The protein folding story (see review in ref. 2) shows unambiguously that the physical understanding of this process must precede the determination of any reasonable predictive algorithm.

The starting point of the problem of protein folding is the well-known argument of Levinthal [3] which stresses the enormous number of conformational states in proteins and thus the impossibility of scanning these states to choose the native one. Such an argument leads one to consider a naive thermodynamical approach to structure prediction as being invalid and gave rise to discussions of an alternative – kinetic – approach [4,5]. It has been argued, however, that Levinthal has overestimated the number of relevant states, since obvious physical restrictions (such as compactness due to the screening of nonpolar groups from water, secondary structure formation, etc.) decrease the number of energetically favorable and, thus, statistically relevant conformations [6]. However, a considerable cost is involved in the reduction of phase space. As a result of this reduction, the well-known multiple-minima problem arises [7], since such definite states (free energy minima) become separated by higher barriers and, is favored by the decreasing number of relevant states, that a protein loses with time due to the overriding barriers between these states.

Thus, for proteins, elucidation of the real situation would be very interesting. To be more specific,

Correspondence address: E.I. Shakhnovich, Institute of Protein Research, Academy of Sciences of the U.S.S.R., 142292 Puschino, Moscow Region, U.S.S.R.

one should carry out investigations on the existence, number and distribution of frozen states (energy minima) of a protein.

A distinguishing feature of proteins compared with such simple objects as homopolymeric molecules deserves emphasis. Homopolymers may also condense into a compact, globular, state [8]. However, such globules do not possess a definite fold; the set of its conformations is almost as random as that of coil conformations. * It is clear from the fact that the entropy jump at the coil-globule transition is very small (this is a surface effect and falls off according to $N^{-1/3}$ per link when the number of links $N \to \infty$ [9]). Therefore, a homopolymeric globule possesses no definite chain fold, i.e., it does not have any frozen states. As a result, when considering a protein, it is necessary to take into account its heterogeneity, i.e., to treat it as a heteropolymer.

Heteropolymeric molecules may possess the same thermodynamic states as a homopolymer and, furthermore, frozen states with definite chain folds with small fluctuations.

This is similar to the case of spin glasses (SG) where analogous phases, viz., paramagnetic, ferromagnetic and frozen, have been observed. Extension of this analogy to the application of the ideas in SG theory is the logical consequence regarding investigations of proteins. The relation between the SG theory and that of native, completely folded proteins has been discussed in ref. 10, application of SG theory to the case of protein folding being dealt with in ref. 11. In both studies, however, phenomenological models were investigated where a particular structure of the set of states and the distribution of their energies were postulated but not derived from considerations of a microscopic protein model. Thus, numerous frozen states were assumed for a protein chain in ref. 11 and a phenomenological phase diagram was suggested. Investigation of a microscopic heteropolymer model with the aid of the replica approach was suggested in refs. 12–15. A model and basic relations were introduced in refs. 12 and 13. In two of the studies [12,14], results were obtained for the case where polymeric bonds were ignored; treatment of the polymeric bonds was carried out according to a crude approximation [12] such that it was impossible to deal with the case of real protein folding; furthermore, the random coil-random globule transition was investigated only within the limits of the approximation used [12,14]. A special case of a low-dimensional heteropolymer where polymeric bonds play essential roles was described in refs. 13 and 15. The present paper is devoted to the investigation of a three-dimensional microscopic model of globular protein chains with the use of a replica approach. (The replica approach for studies on heteropolymers was employed in refs. 16 and 17, in which most attention was paid to the coil state of a polymer.)

In this paper we consider a microscopic model which treats a protein as a heteropolymeric chain with a quenched disordered sequence of links. This involves the taking of any particular chain as being representative of a sufficiently large set of randomly synthesized chains and conclusions concerning the general features of such chains. Hence, we are dealing with values that are self-averaging, i.e., values that are typical for any particular representative of set of sequences. Such an approach is justified by the observation that the statistical properties of sequences of globular proteins are indistinguishable from those of a random uncorrelated set of sequences [18] (see ref. 19 for a detailed discussion of such an approach in protein theory).

We shall investigate the coil, homopolymer-like globular and frozen states of a polymer and outline those conditions under which they occur. The heterogeneity of a protein is taken into account explicitly, i.e., the major focus of attention is that of the heteropolymeric aspects of the problem. It is shown that a sufficiently high degree of heterogeneity gives rise to a small number of thermodynamically dominant states with definite three-di-

* It is necessary, however, to distinguish between unfolded (i.e., coil) and globular states even without definite chain folds. As stressed in ref. 9, coil and globule are two different macroscopic phases of a chain which differ drastically in fluctuational regime: the density of a globule has a well-defined value. Therefore, a coil-globule transition takes place when only the average density changes, and a real folding transition when the pathway of the chain becomes established is possible only for heteropolymeric chains.

mensional structures. This may justify a thermodynamic approach to the problem of protein folding.

Our consideration of the microscopic model confirms the main assumption made in a phenomenological investigation [11] of frozen states and makes it possible for one to outline the conditions under which a frozen phase as well as a random globule and coil can exist, i.e., constructing a phase diagram for the model.

2. The model and qualitative investigation

An elaborated replica approach for investigation of the set of states in quenched disordered systems has been developed for spin glasses (for a review, see ref. 20). Within the framework of the replica method, one considers n copies (replicas) of the same disordered system, followed by averaging over the disorder and finally takes $n \to 0$ to obtain the free energy of the system. During this procedure order parameters appear for which physical values of different replicas become mixed. These order parameters are the well-known overlaps between replicas in spin glasses: $q_{\alpha\beta} = \sum_i s_i^\alpha s_i^\beta$ where α and β designate the replicas and $\{s_i^\alpha\}$ the spin configuration of replica α. The cornerstone of the physical interpretation of the replica approach involves the identification of replicas with the physical states of a system. This form of interpretation was demonstrated to be well justified for spin glasses [21]. The equilibrium values of $q_{\alpha\beta}$ may be either replica symmetric or nonsymmetric [22,23]. The replica symmetric solution signifies that there exists only one equilibrium state with given (for spin system) local magnetization $M_i = \bar{s}_i$ and $q_{\alpha\beta} = q_0$ for all pairs α,β (where the overbar denotes thermal averaging). The nonsymmetric solution for $q_{\alpha\beta}$ assumes that $q_{\alpha\beta}$ is now an $n \times n$ matrix with unequal matrix elements. This means [21,22] that there exist numerous distinctly differing states which correspond to different sets of spin configurations $\{s_i\}$ and are separated by infinite (in the thermodynamic limit, $N \to \infty$) barriers. The appearance of these infinite barriers shows that the system exhibits nonergodic behavior which depends strongly on its prehistory.

In this work, we investigate the set of states of a heteropolymeric protein model. The configuration (microstate) of a molecule is now characterized by the set of vectors $\{r_i\}$ $(i = 1, \ldots N)$ which describe the spatial positions of all N links of a chain. The lattice model will be taken for the sake of convenience, i.e., links are assumed to be positioned on a three-dimensional lattice.

The replica trick mentioned above is applied to investigation of the thermodynamics of such a polypeptide chain. It means that we consider n chains with microstates $\{r_i^\alpha\}$ $(\alpha = 1, \ldots, n)$. As a result of the application of standard replica procedures (detailed below) the following overlap parameters emerge:

$$Q_{\alpha\beta}(R_1, R_2) = \sum_i \Delta(R_1 - r_i^\alpha)\Delta(R_2 - r_i^\beta) \qquad (1)$$

where Δ denotes the Kronecker delta which has the meaning:

$$\Delta(X) = \begin{cases} 1 & \text{for } X = 0 \\ 0 & \text{otherwise} \end{cases}$$

The order parameters are more complicated than the analogous overlaps in SG theory, since $Q_{\alpha\beta}(R_1, R_2)$ is the $n \times n$ matrix of functions of the two spatial coordinates R_1, R_2. It is convenient to introduce a simpler overlap parameter

$$q_{\alpha\beta} = \frac{1}{N}\sum_i \Delta(r_\alpha^i - r_\beta^i) = \frac{1}{N}\sum_R Q_{\alpha\beta}(R, R) \qquad (1a)$$

The functions $Q_{\alpha\beta}(R_1, R_2)$ obey the normalization conditions:

$$\sum_{R_1, R_2} Q_{\alpha\beta}(R_1, R_2) = N \qquad (2)$$

$$\sum_{R_2} Q_{\alpha\beta}(R_1, R_2) = \rho_\alpha(R_1) \qquad (2a)$$

where

$$\rho_\alpha(R) = \sum_{j=1}^{N} \Delta(R - r_j^\alpha)$$

is the usually defined average density of monomers of replica α. Summation in eq. 2 is performed over all lattice sites.

One readily observes that the order parameter introduced into eq. 1 are the correlators of chain folds for replicas α and β. These order parameters are extremely useful in studies on the existence and number of frozen chain folds. For example, consider a globular chain with a density of links ρ. Such a chain occupies a lattice region of volume $V = N/\rho$. Two qualitatively different situations are possible a priori for this chain:

(1) The globule is like a collapsed homopolymer, i.e., it has no definite fold. In this case

$$Q_{\alpha\beta}(R_1, R_2) \cong N/V^2 = \rho^2/N; \quad q_{\alpha\beta} = \rho/N \qquad (3)$$

i.e., it is negligibly small when N is large.

(2) The globule possesses numerous frozen chain folds. This means that a set of thermodynamically definite coordinates $\{r_i^\tau\}$ exists, where τ denotes the fold, i.e., pure state, using the terminology for spin glasses [21]. This signifies that, in this case (for $N \gg 1$),

$$Q_{\alpha\beta}(R_1, R_2) = \begin{cases} \rho\Delta(R_1 - R_2) \\ \quad \alpha,\beta \text{ in the same state} \\ 0 \quad \alpha,\beta \text{ in different states} \end{cases} \qquad (4)$$

and

$$q_{\alpha\beta} = \begin{cases} 1 & \alpha,\beta \text{ in the same state} \\ 0 & \alpha,\beta \text{ in different states} \end{cases} \qquad (4a)$$

In particular, when the globule possesses one thermodynamically stable chain fold, i.e., only a single pure state, then $Q_{\alpha\beta}(R_1, R_2) = \rho\Delta(R_1 - R_2)$; $q_{\alpha\beta} = 1$ for all pairs α,β. In this case, the solution for $Q_{\alpha\beta}(R_1, R_2)$ and $q_{\alpha\beta}$ must be replica symmetric.

Resuming this discussion, we note that the qualitative features of the order parameter $Q_{\alpha\beta}(R_1, R_2)$ provide complete information on the existence and number of pure states with definite chain fold: for $Q_{\alpha\beta}(R_1, R_2) \to 0$ when $N \to \infty$, the globular state of a protein is then similar to a collapsed homopolymer; for $Q_{\alpha\beta}(R_1, R_2) = \rho\Delta(R_1 - R_2)$ for all pairs α,β, then only one fold exists; and finally, provided eq. 4 is valid, numerous frozen states with large barriers between them exist.

The absolutely (up to the microscopic scale) frozen states have been considered in the above discussion. The intermediate situation is possible when the chain fold is not completely frozen but fluctuates around some fold (fig. 1). The scale of fluctuations, R, represents the average distance of the deflection from this fold for each residue. In this case, for a sufficiently large globule (i.e., when $V \gg a^3$, a being the size of a bond between links) $Q_{\alpha\beta}(R_1, R_2)$ depends only on the difference between R_1, R_2 and must be a localized function of scale R:

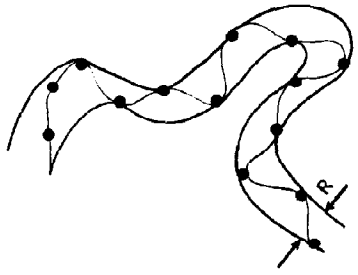

Fig. 1. 'A chain in a tube' model representing a fold defined to scale R.

$$Q_{\alpha\beta}(R_1, R_2) = \rho Q^1_{\alpha\beta}\left(\frac{R_1 - R_2}{R}\right)/R^3 \qquad (5)$$

where $Q^1_{\alpha\beta}(x)$ denotes a normalized function on a lattice with a unit scale of:

$$\sum_x Q^1_{\alpha\beta}(x) = 1$$

Section 3 examines details of the microscopic model of a protein and the calculation of its free energy. The above-discussed order parameters appear automatically during the process of calculation, and the mean-field approximation is applied to obtain the equilibrium solution for $Q_{\alpha\beta}$.

It is shown that the solution is given by eq. 3 for the case of a low degree of heterogeneity of a chain and by eq. 4 for heterogeneity of sufficiently large extent. For increasing chain heterogeneity, a drastic transition for the equilibrium $Q_{\alpha\beta}(R_1, R_2)$ from eq. 3 to eq. 4 occurs, which signifies that drastic freezing of a chain into a thermodynamic state with numerous frozen folds and nonergodic behavior takes place. Further increase in heterogeneity leads to a decrease in the number of equilibrium folds, i.e., a more replica-symmetric solution for $Q_{\alpha\beta}$.

Readers not interested in the details of calculations may skip section 3 and proceed to section 4.

3. Basic relations and the mean field approximation

Consider a linear heteropolymer chain on a three-dimensional cubic lattice. Monomers of a chain have their own excluded volume v and it is natural to assume the volume of a lattice site to be of the order v.

The partition function for a chain has the following form:

$$Z = \sum_{\{r_i\}} \prod_{i=1}^{N-1} g(r_{i+1} - r_i) \exp\left(- \frac{H\{r_i\}}{k_B T} \right) \tag{6}$$

where k_B represents Boltzmann's constant and T the absolute temperature. Summation is taken over all configurations of a chain $\{r_i\}$, i the number of a current link, and N the number of links; functions $g(r_i, r_{i+1})$ describe the chemical structure of polymeric bonds; it was introduced in ref. 9 for the simplest, Gaussian, model of a chain and was described for different models of chain rigidity in ref. 24. This function has the sense of the conditional probability that link $i+1$ is characterized by coordinates r_{i+1} provided that the i-th link has coordinates r_i. All information on the details of short-range interactions (between neighboring links along the chain) is contained in g. We take this function in the simplest form:

$$g(r_{i+1} - r_i) = \frac{1}{(2\pi a^2)^{3/2}} \exp\left(- \frac{(r_{i+1} - r_i)^2}{2a^2} \right) \tag{7}$$

where the characteristic size of a bond a is introduced.

The energy of interactions $H\{r_i\}$ is expressed in the form of the two- and three-particle approximation:

$$H\{r_i\} = \tfrac{1}{2} \sum_{i \neq j} B_{ij} \Delta(r_i - r_j) + \tfrac{1}{6} C \sum_{\substack{i \neq j \\ j \neq k}} \Delta(r_i - r_j) \Delta(r_k - r_j) \tag{8}$$

where C denotes the three-particle interaction constant; it describes the excluded volume effect and is assumed to be equal for all links. It is introduced for the purpose of stability of the model, in order to prevent its collapse into a one-lattice site. The disordered character of a chain is taken into account in the random two-particle term B_{ij}. We assume a Gaussian probability distribution for the two-particle interactions:

$$P(B_{ij}) = \frac{1}{(2\pi B^2)^{1/2}} \exp\left[- \frac{(B_{ij} - B_0)^2}{2B^2} \right] \tag{9}$$

This is a very important assumption which is valid if the number of types of interaction between links N_i exceeds the number of links N. N_i is correlated with the number of types of link N_t via the obvious relation: $N_i = N_t(N_t - 1)/2$, thus the condition $N_t^2 \geq N$ must be fulfilled to make the distribution, eq. 9, valid.

We are now in a position to calculate the free energy of a chain. This value coincides with the free energy averaged over the sequence distribution eq. 9 due to self-averaging of the free energy.

The averaging of free energy, defined as the logarithm of the partition function eq. 6, may be carried out using the replica trick:

$$\langle F \rangle_{av} = -k_B T \cdot \langle \ln Z \rangle_{av} = -k_B T \cdot \lim_{n \to 0} \frac{\langle Z^n \rangle_{av} - 1}{n} \tag{10}$$

where $\langle \ \rangle_{av}$ denotes averaging over disorder, i.e., over all sets $\{B_{ij}\}$ with weight defined by eq. 9.

Averaging of the terms on the right-hand side of eq. 10 is straightforward; for integer n it yields:

$$\langle Z^n \rangle_{av} = \int \sum_{\{r_i^\alpha\}} \prod_{\alpha=1}^{n} \prod_{i=1}^{N-1} g(r_{i+1}^\alpha - r_i^\alpha) \times \exp\left[- \frac{1}{2} \frac{B_{ij}}{k_B T} \sum_\alpha \sum_{i \neq j} \Delta(r_i^\alpha - r_j^\alpha) - \frac{1}{6} \frac{C}{k_B T} \sum_\alpha \sum_{\substack{i \neq j \\ j \neq k}} \Delta(r_i^\alpha - r_j^\alpha) \times \Delta(r_k^\alpha - r_j^\alpha) \right] \prod_{i,j} P(B_{ij}) \, dB_{ij} \tag{11}$$

where α designates the current replica index. Integration over disorder can be performed explicitly, since integrals over B_{ij} are independent and Gaussian. We obtain:

$$\langle Z^n \rangle_{av} = \sum_{\{r_i^\alpha\}} \prod_{\alpha=1}^{n} \prod_{i=1}^{N-1} g(r_{i+1}^\alpha - r_i^\alpha) \times \exp\left[- \frac{H_{eff}(\{r_i^\alpha\})}{k_B T} \right] \quad (12)$$

summation in eq. 12 represents configurational summation over all positions of the $(N-1)n$ links of n replicas. (To exclude the translational entropy of whole chains, we assumed the first link of all n replicas to be fixed at the origin.) $H_{eff}(\{r_i^\alpha\})$ denotes the effective Hamiltonian, which can be presented in the following form:

$$H_{eff}(\{r_i^\alpha\}) = H_1(\{r_i^\alpha\}) + H_2(\{r_i^\alpha\})$$

where

$$H_1 = \tfrac{1}{2}\tilde{B} \sum_\alpha \sum_{i \neq j} \Delta(r_i^\alpha - r_j^\alpha) + \tfrac{1}{6} C \sum_\alpha \sum_{\substack{i \neq j \\ j \neq k}} \Delta(r_i^\alpha - r_j^\alpha)\Delta(r_k^\alpha - r_j^\alpha) \quad (13)$$

is the one-replica contribution of the effective Hamiltonian, describing the average density of a chain and $\tilde{B} = B_0 - B^2/2k_B T$.

$$H_2 = - \frac{B^2}{4k_B T} \sum_{\alpha \neq \beta} \sum_{i \neq j} \Delta(r_i^\alpha - r_j^\alpha)\Delta(r_i^\beta - r_j^\beta) \quad (14)$$

is the two-replica component which describes interactions between replicas; due to this term all the nontrivial effects occur.

One can now observe that the order parameter average density ρ and $Q_{\alpha\beta}$ appear explicitly: eqs. 13 and 14 can be rewritten directly in terms of these order parameters:

$$H_1 = \frac{\tilde{B}}{2} \sum_\alpha \sum_R \rho_\alpha^2(R) + \frac{C}{6} \sum_\alpha \sum_R \rho_\alpha^3(R) \quad (15)$$

$$H_2 = - \frac{B^2}{4k_B T} \sum_{\alpha \neq \beta} \sum_{R_1, R_2} Q_{\alpha\beta}^2(R_1, R_2) \quad (16)$$

The order parameters ρ_α and $Q_{\alpha\beta}$ are interrelated through the normalization condition, eq. 2a, which is fulfilled automatically if one takes $Q_{\alpha\beta}$ in the form eq. 5.

It follows that evaluation of the order parameters may be performed independently. The simplest problem is the determination of density ρ_α. This can be carried out identically to the procedure described in ref. 9 for a homopolymer where it was shown that a sufficiently large chain forms a globule with a core of constant density and a sharp boundary where the density falls to zero. The Hamiltonian is symmetric relative to the permutation of replica indices and thus the density, which is the one-replica value, must be the same for all replicas. The equilibrium density in the core is obtained, according to ref. 9, from eq. 15 together with the normalization condition:

$$\sum_R \rho(R) = N$$

$$\rho_\alpha = -3\tilde{B}/2C \quad (17)$$

The remaining (and least trivial) problem is now the evaluation of the equilibrium overlaps $Q_{\alpha\beta}$. The overlaps are determined from the mean-field equation

$$\delta\left(H_2\{Q_{\alpha\beta}\} - k_B T S\{Q_{\alpha\beta}\}\right)/\delta Q_{\alpha\beta} = 0 \quad (18)$$

together with the normalization conditions, eqs. 2 and 2a. $S\{Q_{\alpha\beta}\}$ in eq. 18 denotes the conformational entropy which determines the number of chain pathways corresponding to the given set $\{Q_{\alpha\beta}\}$:

$$S\{Q_{\alpha\beta}\} = \ln\left[\sum_{\{r_i^\alpha\}} \prod_{\alpha=1}^{n} \prod_{i=1}^{N-1} g(r_{i+1}^\alpha - r_i^\alpha) \times \delta\left(Q_{\alpha\beta}(R_1, R_2) - \sum_i \Delta(R_1 - r_i^\alpha)\Delta(R_2 - r_i^\beta) \right) \right] \quad (19)$$

4. Phase diagram

Assuming that $Q_{\alpha\beta}(R_1, R_2)$ is of the form given by eq. 5, it is not difficult to ascertain the behavior

of this order parameter. The energy expressed by eq. 16 is obtained in the form:

$$H_2 = -N\frac{\rho B^2 A_1}{4k_B T R^3} \tag{20}$$

where

$$A_1 = \sum_{\alpha \neq \beta} \sum_{X} \left(Q^1_{\alpha\beta}(X)\right)^2 \tag{21}$$

Investigation of entropy is more complicated and has been left to the appendix. However, the dependence of the entropy on the scale R of the order parameter can be established in a simple manner for $R \gg a$. Consider a chain confined within a 'curved' tube of radius R (fig. 1). The tube mimicks the existence of a frozen fold. The entropy of a chain in a tube can be readily derived (e.g., see ref. 25):

$$S(R) \cong S_0 - \frac{a^2 A_2}{R^2} \tag{22}$$

where $A_2 \sim 1$ and S_0 represents the entropy of a free coil. A_1 and A_2 in eqs. 20 and 22 change sign when $n < 1$, therefore, it is necessary to maximize $F(R) = H_2(R) - k_B T S(R)$ in order to obtain R. A plot of $F(R)$ for this case is shown in fig. 2. It is clear from fig. 2 that only two regimes for R correspond to stable states: (1) $R = \infty$ which implies $Q_{\alpha\beta} = 0$ and (2) $R \sim v^{1/3} \ll a$. The latter regime corresponds to chain freezing on a scale up to the microscopic range.

The existence of only two stable values of R means that the solution for $Q_{\alpha\beta}$ takes the form of eq. 4. In order to investigate the number of 'freezing patterns', i.e., the dependence of $Q_{\alpha\beta}$ on α, β, one must divide n replicas into n/x_0 groups, each group containing x_0 replicas. To this end

$$Q_{\alpha\beta}(R_1, R_2) = \begin{cases} \rho\Delta(R_1 - R_2); q_{\alpha\beta} = 1 & \text{for } \alpha,\beta \text{ in the same group} \\ 0; q_{\alpha\beta} = 0 & \text{for } \alpha,\beta \text{ in different groups} \end{cases} \tag{4b}$$

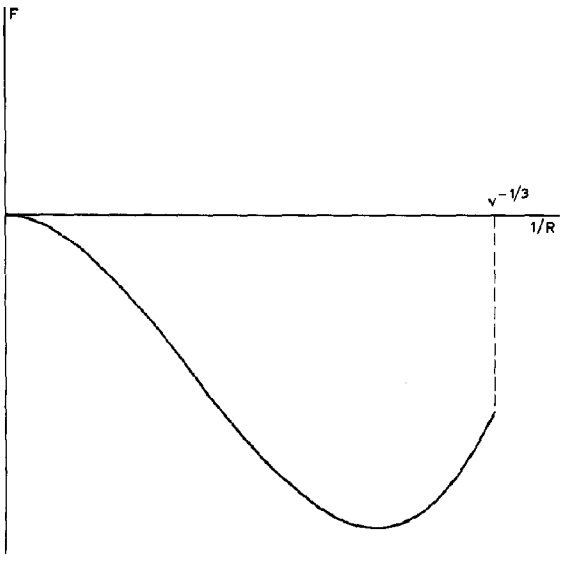

Fig. 2. Free energy of n replicas plotted vs. reciprocal scale $(1/R)$ of the overlap function (for $n < 1$). $v^{1/3}$ is the limiting microscopic scale. Two limiting values of R correspond to stable states, namely, $R = \infty$ and $R = v^{1/3}$. Intermediate values of R correspond to unstable states.

Within the framework of the replica method, the limit $n \to 0$ must be taken after derivation of the free energy. The order parameters $Q_{\alpha\beta}$ and $q_{\alpha\beta}$ instead of being in the form of an $n \times n$ matrix become functions of the replica parameter x: $Q_{\alpha\beta}(R_1, R_2) \to Q(x, R_1, R_2)$; $q_{\alpha\beta} \to q(x)$ $(0 \leq x \leq 1)$. This involves the application of the ansatz of Parisi [22] for replica-symmetry breaking (RSB) during the course of taking the limit $n \to 0$.

One readily observes that the above division of replicas into independent groups corresponds to the first stage of RSB in the Parisi ansatz and this is the only possible RSB for this system, since only two values of $Q_{\alpha\beta}$ and $q_{\alpha\beta}$ ($Q(x)$ and $q(x)$ for $n \to 0$) given by eq. 4b are stable.

In this case $Q(x, R_1, R_2)$ and $q(x)$ are step functions of x (see fig. 3 for $q(x)$). The position of the breakpoint x_0 must be obtained from maximization of the free energy $F_2(x_0) = H_2(x_0) - k_B T S(x_0)$ where the energy of replica interactions, H_2, is obtained simply from eq. 16.

$$H_2(x_0) = -N\frac{n}{x_0}(x_0 - 1)x_0\frac{B^2}{4k_B T}\rho \tag{23}$$

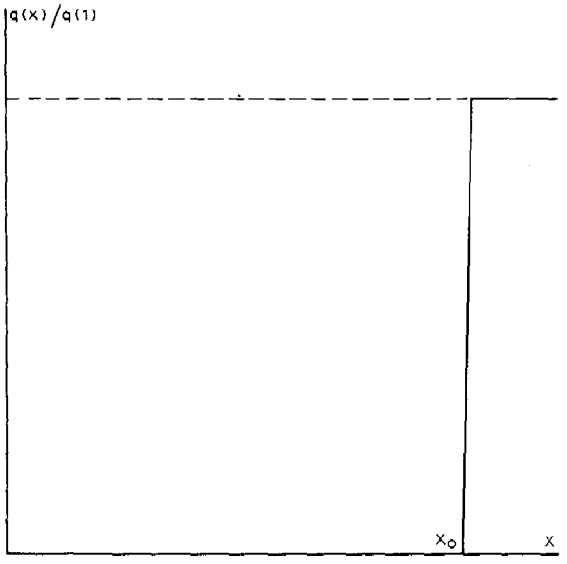

Fig. 3. Plot of order parameter $q(x)$.

Calculation of the entropy is detailed in the appendix (eq. A20):

$$S(x_0) = N\frac{n}{x_0}(x_0 - 1)\ln\frac{v}{a^3} \tag{24}$$

The condition $0 \le x_0 \le 1$ must be fulfilled. The result of maximization is:

$$x_0 = \left(\frac{-4k_B^2T^2\ln\frac{v}{a^3}}{\rho B^2}\right)^{1/2}$$

$$\text{when } \frac{-4k_B^2T^2\ln\frac{v}{a^3}}{\rho B^2} < 1$$

and in the opposite case:

$$x_0 = 1. \tag{25}$$

$x_0 = 1$ corresponds to the absence of frozen states: $F_2 \equiv 0$ and $Q(x) \equiv 0$.

Thus, there exists the threshold value:

$$B_{tr} = \left(\frac{-4k_B^2T^2\ln\frac{v}{a^3}}{\rho}\right)^{1/2} \tag{26}$$

at which the drastic transition connected with the appearance of frozen states occurs. This transition is quite peculiar: it displays the features of both first- and second-order phase transitions. There is a jump of the order parameter at the transition point (e.g., $q(1)$ jumps from 0 to 1). It is well known that $q(1)$ represents the self-overlap of a replica, i.e., it is the Edwards-Anderson order parameter [20]. This jump denotes the occurrence of drastic freezing of a chain pathway after the threshold value of chain heterogeneity eq. 26 has been attained. This transition, however, has no latent heat, since $x_0 = 1$ at the transition point and there is no jump of energy H_2 at the transition. Therefore, this is a true glass transition similar to those observed in structural glasses [26].

The results obtained enable one to derive a general phase diagram for this protein model. It is reasonable to use variables (B_0, B) which denotes the mean and standard variance of the distribution of interaction energies eq. 9. As discussed below, three phases must be defined: coil ($\rho_\alpha = 0$),

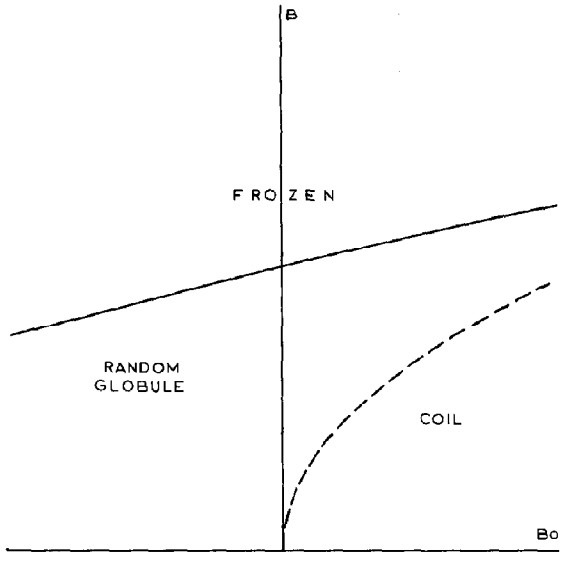

Fig. 4. Phase diagram of a chain plotted in the coordinates of mean interaction (B_0) and the standard variance of interaction (heterogeneity, B). The dashed line denotes the second-order phase boundary, the continuous line corresponding to the glass transition with a jump of the order parameter and without latent heat. The random globule is similar to the collapsed homopolymer; the frozen globule is discussed in the text.

homopolymer-like, or random, globule ($\rho > 0$, $Q(x) = 0$) and frozen phase ($\rho > 0$, $Q(x) > 0$ for $x_0 < x < 1$). The random coil-random globule transition line is determined by the condition of effective θ-point $\tilde{B} = 0$ which yields:

$$B_0 - B^2/2k_BT = 0 \tag{27}$$

This transition is of the second-order type for large globules [9].

The random globule-frozen phase transition is expressed by the equation:

$$B = B_{tr}(\rho(\tilde{B})) \tag{28}$$

where B_{tr} is given by eq. 26 and $\rho(\tilde{B})$ is determined from eq. 17 for $\rho v \leq 1$ and $\rho = v^{-1}$ for extremely dense globules. The resulting phase diagram is shown in fig. 4.

5. Discussion

We have investigated the possible frozen states and freezing transitions in the model which treats a protein as being a heteropolymer with random independent interactions between links.

The frozen phase in the present model is very similar to the Potts glass phase investigated previously [27,28]. This is not a chance occurrence: the energy of interaction for our model (eq. 8) is very similar to that in the Potts model and our links are analogous to Potts spins with coordinates of links analogous to a state of Potts spins. However, a difference is evident in the polymeric bonds $g(r_{i+1} - r_i)$ which are absent in the Potts model. This term leads to essential corrections: first, due to this term, the transition in our model is always discontinuous: intermediate values of scale R for the order parameter are always unstable. This contrasts with the situation for Potts glass where discontinuous behavior of the order parameter is observed only when the number of components of Potts spins $p > 4$ [27]. Second, polymeric bonds reduce the number of available states for a monomer from V/v to a^3/v, therefore the effective number of states of a link is reduced to a^3/v which is reflected explicitly by eqs. 25 and 26.

Taking these corrections into account, we may use the advanced degree of knowledge concerning the Potts glass state [27,28] to gain deeper insights into the results obtained.

The first question which arises relates to the number of frozen states in the model. This may be formulated in terms of the complexity, or configurational entropy of states:

$$S_c = \sum_s P_s \ln P_s$$

where $P_s = \exp(-E_s/k_BT)/\Sigma_s \exp(-E_s/k_BT)$ designates the Boltzmann weight of state s and the summation is performed over all frozen states. It has been shown previously [21] that the number of states with weights in the interval $(P, P + dP)$ is $f(P)dP$ where $f(P)$ is given by:

$$f(P) = \frac{P^{(-1-x_0)}(1-P)^{x_0-1}}{\Gamma(1-x_0)\Gamma(x_0)}$$

The number of states $M = \exp(S_c)$ is obtained directly from the latter equations. The asymptotes are:

$$M \sim \begin{cases} \exp\left(\dfrac{1}{1-x_0}\right) & \text{for } 1 - x_0 \ll 1 \\ 1 + x_0 & \text{for } x_0 \ll 1 \end{cases} \tag{29}$$

with x_0 being derived from eq. 25. We note that the number of frozen states is exponentially large at the transition point where $x_0 = 1$. This is the reason for the absence of latent heat of the transition: the frozen states at the transition point are almost equally disordered to the random state, i.e., the chain freezes into one out of a considerable number of available frozen states.

It should be emphasized, however, that the number of relevant states is $O(1)$ and not $O(N)$ immediately after the transition. The significance of this factor is that when the heterogeneity of the chain increases or the temperature falls, only a few states dominate. This reflects one of the most remarkable structural features of proteins [29].

Our model, being equivalent to the model of Potts glass with $p \gg 1$, is equivalent to the random energy model (REM) [20]. This equivalence is clearly observed from the behavior of the order

parameter $q(x)$ (fig. 3) which is analogous to that in REM. Therefore, the microscopic parameters for this model can be transformed directly into REM parameters. As a consequence, there exist $(a^3/v)^N$ energy levels corresponding to different chain folds, which are independent random variables with a Gaussian distribution. The density of energy levels is given by:

$$n(E) = \left(\frac{a^3}{v}\right)^N \frac{1}{(\pi N\sigma)^{1/2}} \exp\left(-\frac{(E-E_0)^2}{N\sigma^2}\right) \tag{30}$$

where $\sigma = B\rho^{1/2}$ and $E_0 = \tilde{B}/3C$.

Eq. 30 provides an extremely useful qualitative picture. The vast majority of levels occur within the interval $(E_0 - \sigma N^{1/2}, E_0 + \sigma N^{1/2})$. The density of energy levels in this range is exponentially large, and the chain scans several such states during the process of undergoing thermal motion for which the energy fluctuation is approx. $k_B T N^{1/2}$. Another part of the energy 'spectrum' is that relating to energies $E^* \leq E_0 - N\sigma(\ln a^3/v)^{1/2}$. The density of levels in this region is so low (or gaps in the spectrum so large) that $n(E)k_B T N^{1/2} \leq 1$ for $E < E^*$. This corresponds to the real freezing which occurs when a chain does not pass from state to state in the process of thermal fluctuation. This provides the origin of the freezing transition at temperatures sufficiently low such that $E \leq E^*$.

It is worth mentioning that our study at the microscopic level confirms the main assumption made in the phenomenological treatment of Bryngelson and Wolynes [11] that the energies for different states of a protein have independent Gaussian distributions.

'The principle of minimal frustration' was introduced phenomenologically by the above authors [11], and the special case of the folded state emerged. We see, however, that there is no particular need to introduce such factors as minimal frustrations, since even quasirandom interactions lead to the domination by only a few states providing the possibility in terms of thermodynamics for the self-organization of proteins.

The final question which now arises concerns the accuracy of the description and the degree of validity of such a model of independent interactions. The independence of interactions signifies the following: take three links designated i, j, k with the two-particle interaction constants B_{ij}, B_{ik}, B_{jk}. For independent interactions, $\langle B_{ij} B_{ik} B_{jk} \rangle_{av} = \langle B_{ij} \rangle_{av} \langle B_{ik} \rangle_{av} \langle B_{jk} \rangle_{av}$ and the same must hold true for all higher moments. The two-body interaction constants were previously presented [14] in the form:

$$B_{ij} = \sum_{p=1}^{M} A^p \xi_i^p \xi_j^p$$

where the sequence is characterized by the M-component random vectors $\{\xi_i\}$ and M represents the number of link types; A^p are effective energetic parameters of interactions between links of different types. One can now see that

$$\frac{\langle B_{ij} B_{ik} B_{jk} \rangle_{av}}{\langle B_{ij} \rangle_{av} \langle B_{ik} \rangle_{av} \langle B_{jk} \rangle_{av}} - 1 = O(1/M) \tag{31}$$

It follows that the approximation of independence is justified if the number of types of links is large. As regards proteins, 20 different types of amino acid side chains occur. If we take a large block (e.g., dipeptide) as an 'interaction unit' then the number of types will become 210 when the mutual orientations of interacting dipeptides are neglected. Such estimations show that this approximation may be justified for real proteins.

6. Conclusion

We have demonstrated the principle possibility for the self-organization of proteins as governed by thermodynamic rules. This arises from the fact that at sufficiently low temperatures, very few states with a definite (up to the microscopic scale) structure are dominant. This makes it reasonable for one to consider the folded state of a protein as being the global minimum of free energy which coincides with the global minimum of energy due to a negligibly small contribution by the entropy.

Another unresolved question concerns the height of the barriers between different free energy minima. Elucidation of this problem is very

important for gaining a better understanding of the kinetics of folding, but cannot be achieved within any thermodynamically based, analytical approach. The numeric investigation of spin glasses [30] has shown that barriers between states vary according to $N^{1/4}$. Protein molecules are large but far from being at the thermodynamic limit. A relatively small size for a protein may cause overriding barriers to occur on the way to a truly stable state. The numerical simulation of the dynamics of the protein folding process (either by a Monte-Carlo approach or by molecular dynamics) is the only way to clarify this crucial point.

Appendix: Evaluation of configurational entropy

We evaluate the configurational entropy of n replicas (eq. 19) by using the method suggested in ref. 9. Consider n replicas with interaction potential between replicas expressed as $\theta_{\alpha\beta}(r_i^\alpha, r_i^\beta)$. The partition function for such a system is:

$$Z\{\theta_{\alpha\beta}\} \equiv \exp\left(-\frac{F\{\theta_{\alpha\beta}\}}{k_B T}\right) = \sum_{\{r_i^\alpha\}} \prod_{\alpha=1}^{n} \prod_{i=1}^{N-1} g(r_{i+1}^\alpha - r_i^\alpha) \times \exp\left[-\frac{1}{k_B T} \sum_{\alpha,\beta} \sum_i \theta_{\alpha\beta}(r_i^\alpha, r_i^\beta)\right] \quad \text{(A1)}$$

On the other hand, this partition function can be expressed through $Q_{\alpha\beta}$:

$$Z\{\theta_{\alpha\beta}\} = \int DQ_{\alpha\beta}(R_1,R_2) \times \exp\left[-\frac{1}{k_B T} \sum_{\alpha,\beta} \sum_{R_1,R_2} \theta_{\alpha\beta}(R_1,R_2) \times Q_{\alpha\beta}(R_1,R_2) + S\{Q_{\alpha\beta}\}\right] \quad \text{(A2)}$$

The last integral can be evaluated by the saddle-point method which yields:

$$\theta_{\alpha\beta}(R_1,R_2) = \frac{\delta S\{Q_{\alpha\beta}\}}{\delta Q_{\alpha\beta}(R_1,R_2)} k_B T \quad \text{(A3)}$$

As a result we obtain:

$$F\{\theta_{\alpha\beta}\} = \sum_{\alpha,\beta} \sum_{R_1,R_2} \theta_{\alpha\beta}(R_1,R_2) Q_{\alpha\beta}(R_1,R_2) - k_B T \cdot S\{Q_{\alpha\beta}\}$$

and

$$S\{Q_{\alpha\beta}\} = \frac{1}{k_B T}\left[\sum_{\alpha,\beta} \sum_{R_1,R_2} \theta_{\alpha\beta}(R_1,R_2) \times Q_{\alpha\beta}(R_1,R_2) - F\{\theta_{\alpha\beta}\}\right] \quad \text{(A4)}$$

In the final equation, the assumption is made that the 'external field' $\theta_{\alpha\beta}(R_1,R_2)$ forms a particular 'configuration' $Q_{\alpha\beta}(R_1,R_2)$; this is expressed by eq. A3. Thus, the remaining problem concerns the calculation of $F\{\theta_{\alpha\beta}\}$ in the external field $\theta_{\alpha\beta}(R_1,R_2)$. In order to do so, let us introduce the Green function (propagator), the significance of which is that of a partition function for the case of n chains with fixed ends:

$$Z\left(\begin{matrix}1\\ \vec{\eta}\end{matrix}\middle|\begin{matrix}N\\ \vec{\xi}\end{matrix}\right) = \sum_{\{\vec{r}_i\}} \prod_{\alpha=1}^{n} \prod_{i=1}^{N-1} g(r_{i+1}^\alpha - r_i^\alpha) \times \exp\left[-\frac{1}{k_B T} \sum_{\alpha,\beta} \sum_i \theta_{\alpha\beta}(r_i^\alpha, r_i^\beta)\right] \times \Delta(\vec{r}_1 - \vec{\eta}) \Delta(\vec{r}_N - \vec{\xi}) \quad \text{(A5)}$$

where $\vec{r} = (r^1, r^2, \ldots, r^n)$ denotes the set of coordinates for all replicas. The recurrence relation connecting the Green function for a chain of $N+1$ links with that (eq. A5) corresponding to a chain with N links can be written directly from eq. A5:

$$Z\left(\begin{matrix}1\\ \vec{\eta}\end{matrix}\middle|\begin{matrix}N+1\\ \vec{\xi}\end{matrix}\right) = \exp\left(-\frac{1}{k_B T} \sum_{\alpha,\beta} \theta_{\alpha\beta}(\xi_\alpha, \xi_\beta)\right) \times \sum_{\{\vec{\xi}'\}} \prod_{\alpha=1}^{n} g(\xi_\alpha - \xi'_\alpha) Z\left(\begin{matrix}1\\ \vec{\eta}\end{matrix}\middle|\begin{matrix}N\\ \vec{\xi}'\end{matrix}\right) \quad \text{(A6)}$$

further on we introduce a shorter form of notation $\hat{g}$ for the integral operator:

$$\hat{g}\psi = \sum_{\{\vec{\xi}'\}} \prod_{\alpha=1}^{n} g(\xi_\alpha - \xi'_\alpha) \psi(\vec{\xi}') \quad \text{(A7)}$$

The general solution of eq. A6 can be written as a bilinear expansion as follows:

$$Z\left(\begin{array}{c|c} 1 & N \\ \vec{\eta} & \vec{\xi} \end{array}\right) = \sum_k \Lambda_k^N \psi_k(\vec{\xi}) \psi_k^+(\vec{\eta}) \qquad \text{(A8)}$$

where ψ_k and Λ_k respectively designate the eigenfunctions and eigenvalues of the integral equation:

$$\hat{g}\psi = \Lambda \exp\left(\frac{1}{k_B T} \sum_{\alpha,\beta} \theta_{\alpha\beta}(\xi_\alpha, \xi_\beta)\right)\psi \qquad \text{(A9)}$$

If eq. A9 possesses a discrete spectrum (as occurs for for globules [9]), the largest eigenvalue dominates in the expansion, eq. A8, for $N \gg 1$. Consequently, we obtain:

$$Z\left(\begin{array}{c|c} 1 & N \\ \vec{\eta} & \vec{\xi} \end{array}\right) = \Lambda^N \psi(\vec{\xi})\psi(\vec{\eta}) \qquad \text{(A10)}$$

with Λ and ψ being the largest eigenvalue and eigenfunction, respectively, of eq. A9. Eq. A10 yields the following expression for the free energy:

$$F\{\theta_{\alpha\beta}\} = -k_B T N \cdot \ln \Lambda\{\theta_{\alpha\beta}\} \qquad \text{(A11)}$$

The density of monomers of all replicas $\rho(\vec{x}) = \sum_{i=1}^N \prod_{\alpha=1}^n \Delta(r_i^\alpha - x_\alpha)$ is expressed via the eigenfunction:

$$\rho(\vec{x}) = \psi^2(\vec{x}) \exp\left(\frac{1}{k_B T} \sum_{\alpha,\beta} \theta_{\alpha\beta}(x_\alpha, x_\beta)\right) \qquad \text{(A12)}$$

$\rho(\vec{x})$ obeys the normalization condition:

$$\int \rho(\vec{x})\, d\vec{x} = N$$

Finally, we obtain from eqs. A12, A9 and A4:

$$S\{Q_{\alpha\beta}\} = \sum_{\vec{x}} \rho(\vec{x}) \ln \frac{\hat{g}\psi}{\psi} \qquad \text{(A13)}$$

This expression (eq. A13) represents the natural generalization of the corresponding expression for entropy obtained in ref. 9 for homopolymers. The conditions required to be met for self-consistency must be added to make the system, eqs. A9, A12 and A13, closed. Eq. 18 together with eq. A3 yields:

$$\theta_{\alpha\beta}(R_1, R_2) = -\frac{B^2}{4k_B T} Q_{\alpha\beta}(R_1, R_2) \qquad \text{(A14)}$$

We may express the overlaps $Q_{\alpha\beta}(R_1, R_2)$ through the density $\rho(\vec{x})$ and ψ:

$$\begin{aligned} Q_{\alpha\beta}(R_1, R_2) &= \sum_{\vec{x}} \rho(\vec{x}) \Delta(x_\alpha - R_1)\Delta(x_\beta - R_2) \\ &= \sum_{\vec{x}} \psi^2(\vec{x}) \exp\left(\frac{1}{k_B T} \sum_{\mu,\nu} \theta_{\mu\nu}(x_\mu, x_\nu)\right) \\ &\quad \times \Delta(x_\alpha - R_1)\Delta(x_\beta - R_2) \end{aligned} \qquad \text{(A15)}$$

Eqs. 16, 18, A9 and A12–A15 constitute the closed system from which the free energy and equilibrium $Q_{\alpha\beta}(R_1, R_2)$ can be derived.

We shall analyze the configurational entropy with respect to the two limiting cases: i.e., for the scale $R \gg a$ and thus $Q_{\alpha\beta}(R_1, R_2)$, $\theta_{\alpha\beta}(R_1, R_2) \ll 1$ and the opposite case when $R - v^{1/3} \ll a$.

For $R \gg a$ the integral operator $\hat{g}$ can be simplified to:

$$\hat{g} = \hat{1} + a^2 \sum_\alpha \frac{\partial^2}{\partial x_\alpha^2} \qquad \text{(A16)}$$

Taking $\theta_{\alpha\beta} \ll 1$, we obtain the following simple expression for the entropy:

$$S\{Q_{\alpha\beta}\} = -a^2 \sum_{\alpha=1}^n \sum_{\vec{r}} \left(\frac{\partial \psi}{\partial r_\alpha}\right)^2 \qquad \text{(A17)}$$

Assuming that $Q_{\alpha\beta}(R_1, R_2)$ has the scale R (also valid for ψ, see eq. A15), we obtain eq. 22 directly from eq. A17.

For the opposite case, i.e., $R \ll a$, the simple presentation, eq. A16, fails. In this situation, we must take $\psi(\vec{x})$ in the form:

$$\psi(\vec{x}) = \psi_0(x_1)\chi(x_2 - x_1)\ldots\chi(x_n - x_1) \qquad \text{(A18)}$$

where $\psi_0(x) = \sqrt{\rho(x)}$ and the functions χ are localized on the scale $v^{1/3}$, even and obey the normalization condition:

$$\sum_x \chi^2(x) = 1$$

This is the only form of ψ which is consistent with the assumption that $Q_{\alpha\beta}(R_1, R_2)$ has a small scale, $v^{1/3} \ll a$, and obeys the normalization conditions, eq. 2. The core of g can be taken in the form of eq. 7, and we readily obtain:

$$\hat{g}\chi = \frac{v}{a^3}\chi \tag{A19}$$

We obtain immediately from eqs. A13 and A19, and the normalization condition for $\rho(\vec{x})$:

$$S = N(n-1)\ln\frac{v}{a^3} \tag{A20}$$

which coincides with eq. 24 if one assumes the number of replicas in each group to be x_0 (instead of n in eq. A20) and the number of groups is n/x_0.

Acknowledgements

We would like to thank O.B. Ptitsyn, A.Yu. Grosberg and A.V. Finkelstein for many fruitful discussions on many aspects of this work.

References

1 D. Wetlaufer, The protein folding problem (West-View, Boulder, CO, 1984).
2 T.E. Creighton, J. Phys. Chem. 89 (1985) 2452.
3 C. Levinthal, J. Chim. Phys. 65 (1968) 44.
4 O.B. Ptitsyn and A.V. Finkelstein, Q. Rev. Biophys. 13 (1980) 339.
5 M. Karplus and D.L. Weaver, Biopolymers 18 (1979) 1421.
6 K.A. Dill, Biochemistry 24 (1985) 1501.
7 G.H. Paine and H.A. Scheraga, Biopolymers 24 (1985) 1391.
8 E.L. Slagowsky, B. Tsai and D. McIntyre, Macromolecules 9 (1978) 687.
9 I.M. Lifshitz, A.Yu. Grosberg and A.R. Khokhlov, Rev. Mod. Phys. 50 (1979) 683.
10 D.L. Stein, Proc. Natl. Acad. Sci. USA 82 (1985) 3670.
11 J.B. Bryngelson and P.G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. 84 (1987) 7524.
12 T. Garel and H. Orland, Europhys. Lett. 6 (1988) 307.
13 E.I. Shahnovich and A.M. Gutin, Europhys. Lett. 8 (1989) 327.
14 T. Garel and H. Orland, Europhys. Lett. 6 (1988) 597.
15 E.I. Shahnovich and A.M. Gutin, J. Phys. A 22 (1989) 1647.
16 S.P. Obukhov, J. Phys. A 19 (1986) 3655.
17 A. Fernandez and H. Rabitz, Biophys. Chem. 28 (1987) 89.
18 O.B. Ptitsyn and M.V. Volkenstein, J. Biomol. Struct. Dyn. 4 (1986) 137.
19 A.Yu. Grosberg and E.I. Shakhnovich, Biofizika 31 (1986) 1045.
20 K. Binder and P. Young, Rev. Mod. Phys. 58 (1986) 801.
21 M.G. Mezard, G. Parisi, N. Sourlas, G. Toulouse and M. Virasoro, J. Phys. (Paris) 45 (1984) 843.
22 G. Parisi, J. Phys. A 13 (1980) L155.
23 G. Parisi, Phys. Rev. Lett. 50 (1983) 1946.
24 A.Yu. Grosberg and A.R. Khohlov, in: Problems in solid state physics, ed. A.S. Prokhorov (Mir Publ., Moscow, 1984) p. 217.
25 F.T. Wall, W.A. Seitz, J.C. Chin and P.G. de Gennes, Proc. Natl. Acad. Sci. U.S.A. 75 (1978) 361.
26 J. Jäckle, Rep. Prog. Phys. 49 (1986) 171.
27 D.J. Gross, I. Kantor and H. Sompolinsky, Phys. Rev. Lett. 55 (1985) 304.
28 T.R. Kirkpatrick and P.G. Wolynes, Phys. Rev. B 36 (1987) 8552.
29 A.V. Finkelstein and O.B. Ptitsyn, Prog. Biophys. Mol. Biol. 50 (1988) 171.
30 N.D. Mackenzie and A.P. Young, Phys. Rev. Lett. 49 (1982) 301.