Journal of Statistical Physics, Vol. 89, Nos. 1/2, 1997

A Nonequilibrium Approach for Random Amphiphilic Copolymer Model

E. G. Timoshenko,[1] Yu. A. Kuznetsov,[1] and K. A. Dawson[1]

Received September 1, 1996; final February 7, 1997

We propose a self-consistent method for studying equilibrium and kinetics of a random copolymer model with a Gaussian distribution of quenched disorder characterized by a generic covariance matrix. The complete phase diagram of the model is obtained and it contains five different states separated by the collapse, glass, and folding "transitions." We analyze the thermodynamic limit for the high-density collapsed globule that permits a simplified analytical study. We argue that the approach may be applied in a variety of situations including biopolymers, gels, and other soft random materials.

KEY WORDS: Quenched disorder; self-consistent method; kinetics.

1. INTRODUCTION

Understanding of the conformational properties of polymeric systems possessing frustration due to the presence of a quenched disorder in monomer interactions has been one of the significant challenges of statistical mechanics. The diversity of applications that are contingent on the resolution of this fundamental problem is quite impressive, varying from random synthetic copolymers, randomly cross-linked polymer and neural networks and gels, to numerous biopolymers such as proteins, RNA and DNA. The analytical methods that have been traditionally used for studying such problems originated from the theory of spin glasses, and rely on the replica formalism.[1] However they have turned out to be rather difficult to apply to polymers, particularly if dynamical properties and kinetics of conformational transitions are concerned. Such transitions may be accompanied by a change of not only the characteristic order parameters describing phase

[1] Theory and Computation Group, Irish Centre of Colloidal Science and Biomaterials, Department of Chemistry, University College Dublin, Dublin 4, Ireland.

0022-4715/97/1000-0347$12.50/0

separation, freezing and so on, but also of the fractal dimensions of the system both in the extrinsic space, in which it is embedded, and in the internal metric of the chain. Whilst excellent methods have been developed to deal with simple polymers, there are few that can be applied to more complex systems, and the method we propose here is quite rare in that it is capable of dealing also with kinetics and general non-equilibrium phenomena. This Gaussian selfconsistent approach, which resembles the time-dependent Hartree approximation, has the advantage of being straightforward, tractable and of general applicability to many problems. At equilibrium this method reduces to the Gibbs–Bogoliubov variational estimate with a quadratic trial Hamiltonian widely used in many different fields. For polymers the variational treatment was proposed by des Cloizeaux[2] and Edwards[3] and later generalised in the works such as those of ref. 4. Unfortunately, it is known to yield an incorrect Flory exponent for good solvent conditions due to an improper probability distribution at small distances. Nevertheless, this drawback of the Gaussian theory is currently well understood and may be resolved either fundamentally by improving the distribution function,[5] or ad hoc by enforcing a short-range cut-off. Despite this limitation even the original version of the method turns out to be quite adequate not only around the theta-point, but also in the dense globular state. Generalisations of the method to kinetics of homopolymer have been considered by us[6, 7] and others.[8] We have found that the scaling of the collapse time agrees with Monte Carlo simulations[9] and with the phenomenological "sausage" model due to Gennes.[10]

Extension of this technique to copolymers, polymers whose interactions (hydrophobicity, stiffness, charge etc.) depend on the relative monomer positions along the chain,[11] can be achieved quite naturally.[12] It would be important for many applications both in the area of synthetic and biological macromolecules.

Many biomolecules are examples of copolymers with rather bizarre properties. For instance, the primary sequence of amino acids of a protein in the native state, which is the biologically active form of these molecules within living cells, determines its unique conformation (tetriary structure).[13] An important reservation in trying to apply the methods of statistical mechanics to proteins is that protein sequences are very special since they have been carefully selected by the long biological evolution. As a preliminary, it is crucial to understand what are the factors that govern the conformational changes of arbitrary polypeptide sequences, and perhaps, prebiotic proteins. One can consider classes of polypeptides in the framework of the random copolymer model. The latter have been extensively studied using the replica formalism of the spin glass theory.[14, 15] Kinetics of copolymer folding has been analyzed in numerous computer

simulations,[16–20] often stimulated by the interest coming from the protein research. A fundamental challenge that remains is to construct a unified analytical approach that could combine and relate the equilibrium and kinetic phenomena in random copolymers.

In this paper we generalise the method proposed in ref. 21 for an Edwards-type model of random amphiphilic copolymer. We specify whether a monomer m is hydrophilic or hydrophobic by the variable $\Lambda_m = +1$ or -1 respectively. Further approximations from this so-called binary disorder model are generally made, leading finally to a continuous distribution of disorder variables, Λ_m. This could be a Gaussian distribution with covariance matrix $\Gamma_{mm'} = \overline{\Lambda_m \Lambda_{m'}}$. The method presents a version of the Gaussian self-consistent approach for arbitrary heteropolymers[12] with a disorder-dependent effective potential. It yields a set of self-consistent equations, which are further directly averaged over the quenched disorder in the lowest order of the dispersion keeping the "fully dressed" quantities and applying an enforced closure for the infinite chain of equations.

In ref. 21 we have numerically solved the resulting self-consistent equations for a few examples of kinetics after a quench from the extended coil to globular states. Here we complete this analysis by elucidating the phase diagram of the model. The numerical procedure is currently applied only to rather short chains, so one would like to consider the thermodynamic limit, in which one would hope to obtain simpler equations amenable to analytical study. The thermodynamic limit serves as an important preliminary for understanding the most universal features of the system under consideration.

2. THE METHOD

For clarity we shall denote the monomer spatial positions by capital characters, $\mathbf{X}_m$, and their Fourier transforms by the lowercase ones, $\mathbf{x}_q$. The Fourier transformations for a ring polymer are defined as,

$$\mathbf{X}_m = \sum_{q=0}^{N-1} f_m^{(-q)} \mathbf{x}_q, \qquad \mathbf{x}_q = \frac{1}{N} \sum_{m=0}^{N-1} f_m^{(q)} \mathbf{X}_m, \tag{1}$$

$$f_m^{(q)} \equiv \exp\left(\frac{2\pi i q m}{N}\right), \tag{2}$$

where N is the degree of polymerization. In the absence of the hydrodynamic effect the exact Langevin equation for the sequence model of

a random copolymer may be written in terms of the Fourier modes as follows,

$$\zeta \frac{d}{dt} \mathbf{x}_q(t) = -\frac{\partial H}{\partial \mathbf{x}_{-q}} + \boldsymbol{\eta}_q(t), \tag{3}$$

$$\langle \eta_q^{\alpha}(t)\, \eta_{q'}^{\alpha'}(t') \rangle = 2k_B T \zeta\, \delta_{q+q',0}\, \delta^{\alpha\alpha'}\, \delta(t-t'), \tag{4}$$

where $\zeta = N\zeta_b$ and ζ_b is the bare friction constant. The effective free energy functional, $H = \bar{H} + H_{dis}$, consists of the homopolymeric, $\bar{H}$, and the disordered, H_{dis} parts respectively,

$$\bar{H} = \frac{\kappa}{2} \sum_n (\mathbf{X}_{n+1} - \mathbf{X}_n)^2 + \sum_{L>2} \bar{u}_L \sum_{\{m\}} \prod_{i=1}^{L-1} \delta(\mathbf{X}_{m_i} - \mathbf{X}_{m_{i+1}}), \tag{5}$$

$$H_{dis} = \frac{1}{2} \sum_{m_1 m_2} (\Lambda_{m_1} + \Lambda_{m_2})\, \delta(\mathbf{X}_{m_1} - \mathbf{X}_{m_2}). \tag{6}$$

Here κ is the spring constant, $\bar{u}_L$ are the virial coefficients of the excluded volume interactions and summation over $\{m\}$ includes all values of indices $m_1, \ldots, m_L$ with $m_i \neq m_{i+1}$. Also Λ_m are random variables that are independent on the half-period with the Gaussian distribution of disorder described by an arbitrary translationally invariant covariance matrix,

$$\overline{\Lambda_m \Lambda_{m'}} = \Gamma_{m-m'}. \tag{7}$$

Henceforth we use the brackets $\langle A \rangle$ to denote the statistical averages over the noise and initial ensemble of monomer positions $\{\mathbf{X}(t=0)\}$ and the bar $\bar{A}$ to denote averages over the quenched distribution of disorder $\{\Lambda\}$. The Fourier transforms $\{\lambda\}$ are likewise independent random Gaussian variables with zero mean value and dispersion $\tilde{\Gamma}_q$,

$$\overline{\lambda_q \lambda_{q'}} = \tilde{\Gamma}_q\, \delta_{q+q',0}, \quad \tilde{\Gamma}_q = \frac{1}{N} \sum_k f_k^{(q)}\, \Gamma_k. \tag{8}$$

For a constant covariance matrix of the disorder distribution we shall use the standard notation $\Gamma_k \equiv \Delta^2$. We believe that treatment of a sequence dependent matrix is important for modelling biopolymers and proteins in particular.

For a given complexion of disorder Eq. (3) is exactly identical to the Langevin equations for arbitrary heteropolymer, with the two-body virial coefficients that are given by the formula,

$$u^{(2)}_{m_1 m_2} = \bar{u}_2 + \tfrac{1}{2}(\Lambda_{m_1} + \Lambda_{m_2}). \tag{9}$$

In our earlier work[12] we have shown that such a system could be successfully studied in the framework of the Gaussian self-consistent method with a nondiagonal self-consistent potential. Since there are no a priori symmetry properties along the chain, when we use the Fourier variables (1) the selfconsistent potential is non-diagonal and denoted by $V_{qp}(t)$. Thus, we replace the exact Langevin Eq. (3) by a linear stochastic ensemble,

$$\zeta \frac{d}{dt} \mathbf{x}_q = -\sum_p V_{qp}(\{\lambda\}, t)\, \mathbf{x}_p + \boldsymbol{\eta}_q(t), \tag{10}$$

where the potential V_{qp} is to be determined self-consistently from the exact equations. The potential has a homopolymeric diagonal part and a non-diagonal part describing the disorder that, according to Eq. (6), should be taken as a linear combination of the disorder variables,[21]

$$V_{qp}(\{\lambda\}, t) \equiv V_q(t)\, \delta_{qp} + U_{qp}(t)\, \lambda_{q-p}. \tag{11}$$

Finally, having derived as many self-consistent equations as there are unknown functions, one has to average over the quenched disorder. This can be accomplished for the Gaussian disorder perturbatively by application of the Wick theorem,

$$\overline{A(\{\Lambda\})} = \exp\left(\frac{1}{2} \sum_{nn'} \Gamma_{n-n'} \frac{\partial^2}{\partial \Lambda_n\, \partial \Lambda_{n'}}\right)\Bigg|_{\{\Lambda\}=0} A(\{\Lambda\}). \tag{12}$$

Let us introduce one of the observables of interest,

$$\mathscr{F}_q(t) \equiv \overline{F_q(t)},\ F_q(t) = \tfrac{1}{3}\langle |\mathbf{x}_q|^2(t)\rangle, \tag{13}$$

which is the mean squared amplitude of q-th normal mode. The mean squared radius of gyration is given simply by $R_g^2 = \sum_{q \neq 0} F_q$. Multiplying Eq. (3) by $\mathbf{x}_{-q}(t)$ and performing evaluations of ref. 21 one can derive,

$$\frac{\zeta}{2} \frac{d}{dt} F_q(t) = k_B T - \frac{1}{3} \left\langle \mathbf{x}_{-q} \frac{\partial H}{\partial \mathbf{x}_{-q}} \right\rangle, \tag{14}$$

where the latter average may be recovered by a differentiation of the mean energy with respect to a set of parameters $\{\gamma\}$:

$$\left\langle \mathbf{x}_{-q} \frac{\partial H}{\partial \mathbf{x}_{-q}} \right\rangle = \frac{\partial \langle H \rangle}{\partial \gamma_{-q}}\Bigg|_1. \tag{15}$$

Here we have used the notation,

$$D_{mm'} = \tfrac{1}{3} \sum_{qp} \gamma_q \gamma_p c^{(q)}_{mm'} c^{(p)}_{mm'} \langle \mathbf{x}_q \mathbf{x}_p \rangle, \tag{16}$$

$$c^{(q)}_{mm'} \equiv f^{(-q)}_{m} - f^{(-q)}_{m'}, \tag{17}$$

and we have introduced the auxiliary parameters γ_q, that are set equal to unity at the end of our calculations. Then the quantities (16) acquire a transparent meaning,

$$D_{mm'}(\gamma_q = 1) \equiv \frac{1}{3} \langle (\mathbf{X}_m - \mathbf{X}_{m'})^2 \rangle = \sum_q d^{(q)}_{mm'} F_q, \tag{18}$$

$$d^{(q)}_{mm'} = 2 \left(1 - \cos \frac{2\pi q (m - m')}{N} \right). \tag{19}$$

Finally, the three-body correlations $D_{mm'm''}$ are defined according to

$$D_{mm'm''} = \tfrac{1}{3} \sum_{qp} \gamma_q \gamma_p c^{(q)}_{mm'} c^{(p)}_{m''m'} \langle \mathbf{x}_q \mathbf{x}_p \rangle, \tag{20}$$

and the higher order terms may be found in ref. 7. Note that for $\gamma_q = 1$ there are simple reduction relations,

$$D_{mm'm''}(\gamma_q = 1) \equiv \tfrac{1}{3} \langle (\mathbf{X}_m - \mathbf{X}_{m'})(\mathbf{X}_{m''} - \mathbf{X}_{m'}) \rangle = \sum_q d^{(q)}_{mm'm''} F_q, \tag{21}$$

$$d^{(q)}_{mm'm''} = \tfrac{1}{2} (d^{(q)}_{mm'} + d^{(q)}_{m''m'} - d^{(q)}_{mm''}). \tag{22}$$

We also introduce the disorder correlation functions,

$$\varphi_{qp}(t) \equiv \overline{\phi_{qp}(t)}, \; \phi_{qp}(t) = \tfrac{1}{3} \lambda_{q-p} \langle \mathbf{x}_{-q}(t) \, \mathbf{x}_p(t) \rangle, \tag{23}$$

and the cumulants of monomer spatial correlations,

$$\mathscr{D}_{mm'} \equiv \overline{D_{mm'}}, \tag{24}$$

$$\overline{D_{mm'} D_{m''m'}}^{(c)} \equiv \overline{D_{mm'} D_{m''m'}} - \overline{D_{mm'}} \; \overline{D_{m''m'}}. \tag{25}$$

The quenched disorder averages can be performed using the Wick theorem (12) in the approximation in which we neglect contributions of order higher than $\tilde{\Gamma}$ and combine the resulting terms into the "fully dressed" averages $\mathscr{F}_q$ and also apply the closure relations of the higher order disorder correlation functions $\overline{\langle \mathbf{x}_q \mathbf{x}_p \rangle \, \lambda_{s_1} \cdots \lambda_{s_k}}$ via φ_{qp}. Carrying out the Fourier transformation and some simple algebra we can prove that the

cumulants (25) depend only on the differences of their indices: $k_1 = m - m'$ and $k_2 = m'' - m'$,

$$\overline{D_{mm'} D_{m''m'}}^{(c)} \equiv P_{k_1 k_2}, \mathscr{D}_{mm'} \equiv \mathscr{D}_k. \tag{26}$$

The functions $P_{k_1 k_2}$ may be expressed via quantities (23) as,

$$P_{k_1, k_2} = \sum_s \tilde{\Gamma}_s^{-1} P_{k_1}^{(s)} P_{k_2}^{(s)} \tag{27}$$

$$P_k^{(s)} = \sum_p d_k^{(p, p+s)} \varphi_{p, p+s}, \tag{28}$$

where the coefficients are

$$d_k^{(q, p)} = \tfrac{1}{2}(d_k^{(q)} + d_k^{(p)} - d_k^{(q-p)}). \tag{29}$$

The cumulant of the squared radius of gyration is related only to the fluctuations of the "composition," λ_0,

$$\overline{R_g^2 R_g^{2(c)}} = \tilde{\Gamma}_0^{-1} \sum_{q,\, p \neq 0} \varphi_{qq} \varphi_{pp}, \varphi_{qq} = \overline{\lambda_0 F_q}, \tag{30}$$

and it may be considered as a spin glass order parameter.[22] Another order parameter of interest,

$$\Psi = \frac{1}{6N^2} \sum_{mm'} \overline{(\Lambda_m + \Lambda_{m'} - 2\lambda_0) D_{mm'}} = \sum_{q \neq p,\, q,\, p \neq 0} \varphi_{qp} \tag{31}$$

is actually related to the phase separation. Indeed, for just two types of monomers hydrophobic "A" ($\Lambda_A = -\Delta$) and hydrophilic "B" ($\Lambda_B = \Delta$) with equal concentrations $n_A = n_B = 1/2$ this reduces simply to

$$\Psi = \Delta(R_g^2(B) - R_g^2(A))/2. \tag{32}$$

By averaging the self-consistent equations for $\langle \mathbf{x}_q \mathbf{x}_p \rangle$ we derive a system of two closed kinetic equations,

$$\frac{\zeta}{2} \frac{d}{dt} \mathscr{F}_q(t) = k_B T - V_q \mathscr{F}_q - \sum_p U_{qp} \varphi_{qp}, \tag{33}$$

$$\zeta \frac{d}{dt} \varphi_{qp}(t) = -(V_q + V_p) \varphi_{qp} - \tilde{\Gamma}_{q-p} U_{qp} (\mathscr{F}_q + \mathscr{F}_p). \tag{34}$$

These are the fundamental equations of our theory. The effective potentials are to be found as derivatives,

$$V_q = \frac{2}{3}\frac{\partial \mathscr{E}}{\partial \mathscr{F}_q}, \quad U_{qp} = \frac{2}{3}\frac{\partial \mathscr{E}}{\partial \varphi_{qp}}, \tag{35}$$

of the mean energy $\mathscr{E} = \overline{\langle H \rangle}$,

$$\begin{aligned}\frac{\mathscr{E}}{N} = {} & \frac{3\kappa}{2}\mathscr{D}_1 + \hat{u}_2 \sum_k \frac{1}{\mathscr{D}_k^{3/2}} + \hat{u}_3 \sum_{k_1 k_2} \frac{1}{Y_0(k_1, k_2)^{3/2}} \\ & - \frac{3}{2}\hat{l} \sum_k \frac{\Phi_k}{\mathscr{D}_k^{5/2}} + \hat{u}_2 \frac{15}{8} \sum_k \frac{P_{k,k}}{\mathscr{D}_k^{7/2}} \\ & + \hat{u}_3 \frac{15}{8} \sum_{k_1 k_2} \frac{Y_2(k_1, k_2)}{Y_0(k_1, k_2)^{7/2}} - \hat{u}_3 \frac{3}{2} \sum_{k_1 k_2} \frac{Y_3(k_1, k_2)}{Y_0(k_1, k_1)^{5/2}},\end{aligned} \tag{36}$$

where $\hat{u}_L \equiv (2\pi)^{-3(L-1)/2} \bar{u}_L$ and $\hat{l} \equiv (2\pi)^{-3/2}$. Here we have used the following set of definitions,

$$Y_0(k_1, k_2) = \mathscr{D}_{k_1} \mathscr{D}_{k_2} - \mathscr{D}_{k_1 k_2}^2, \tag{37}$$

$$\begin{aligned}Y_2(k_1, k_2) = {} & \mathscr{D}_{k_1}^2 P_{k_2, k_2} + \mathscr{D}_{k_2}^2 P_{k_1, k_1} + 4\mathscr{D}_{k_1 k_2}^2 P_{k_1 k_2, k_1 k_2} \\ & + 2\mathscr{D}_{k_1} \mathscr{D}_{k_2} P_{k_1, k_2} - 4\mathscr{D}_{k_1 k_2}(\mathscr{D}_{k_2} P_{k_1, k_1 k_2}\end{aligned} \tag{38}$$

$$Y_3(k_1, k_2) = P_{k_1, k_2} - P_{k_1 k_2, k_1 k_2}, \tag{39}$$

$$\Phi_k = \sum_{qp} d_k^{(q,p)} \varphi_{qp}, \tag{40}$$

with $d_{k_1 k_2}^{(q,p)}$ expressed via $d_k^{(q,p)}$ by Eq. (22).

The equilibrium free energy $\mathscr{A}[V_q, U_{qp}] = \mathscr{E} - T\mathscr{S}$ can be obtained by disorder averaging its λ-dependent Gibbs–Bogoliubov estimate.[21] Alternatively, it is easy to see that the entropy $\mathscr{S} = \sum_{i=0}^{\infty} \mathscr{S}_i$ can be found by resolving the recurrence relation

$$\frac{\partial \mathscr{S}_{i+1}}{\partial \varphi_{qp}} = -\frac{\varphi_{qp}}{\tilde{\Gamma}_{q-p}(\mathscr{F}_q + \mathscr{F}_p)} \left(\frac{\partial \mathscr{S}_i}{\partial \mathscr{F}_q} + \frac{\partial \mathscr{S}_i}{\partial \mathscr{F}_p} \right), \tag{41}$$

starting from the homopolymer entropy $\mathscr{S}_0 = (3k_B/2) \sum_{q \neq 0} \log \mathscr{F}_q$, and satisfying the integrability condition,

$$\mathscr{F}_q \frac{\partial \mathscr{S}_i}{\partial \mathscr{F}_q} + \sum_p \varphi_{qp} \frac{\partial \mathscr{S}_i}{\partial \varphi_{qp}} = 0, \qquad i > 1. \tag{42}$$

In the lowest order this reproduces the formula,

$$\frac{\mathscr{S}}{k_B}=\frac{3}{2}\sum_{q\neq 0}\log \mathscr{F}_q-\frac{3}{4}\sum_{q,\,m\neq 0}\frac{\varphi_{qp}^2}{\mathscr{F}_q\tilde{\Gamma}_{q-p}\mathscr{F}_p}+O(\tilde{\Gamma}^2), \tag{43}$$

earlier obtained in ref. 21 in the Gibbs–Bogoliubov scheme. The equations of motion (33, 34) now may be rewritten up to terms of order $\tilde{\Gamma}^2$ via the gradients of the free energy,

$$\frac{\zeta}{2}\frac{d}{dt}\mathscr{F}_q(t)=-\frac{2}{3}\left(\mathscr{F}_q\frac{\partial \mathscr{A}}{\partial \mathscr{F}_q}+\sum_p \varphi_{qp}\frac{\partial \mathscr{A}}{\partial \varphi_{qp}}\right), \tag{44}$$

$$\frac{\zeta}{2}\frac{d}{dt}\varphi_{qp}(t)=-\frac{2}{3}\left(\varphi_{qp}\left(\frac{\partial \mathscr{A}}{\partial \mathscr{F}_q}+\frac{\partial \mathscr{A}}{\partial \mathscr{F}_p}\right)+\tilde{\Gamma}_{q-p}(\mathscr{F}_q+\mathscr{F}_p)\frac{\partial \mathscr{A}}{\partial \varphi_{qp}}\right). \tag{45}$$

We have made a preliminary study of the kinetic behaviour implied by these equations in ref. 21. However, we may note that these equations possess another interesting application. That is, when the time derivatives are set to zero we find equilibrium solutions for quenched disorder, but without having to apply the replica trick. This represents at least an interesting alternative to the replica approach.

3. PHASE DIAGRAM

In Fig. 1 we present the "phase diagram" of the system for a constant disorder covariance matrix. However, let us note carefully that there is at present no guarantee that these are in any sense genuine macroscopic phases, or the boundaries between them are genuine phase transitions. Nevertheless, the various order parameters of the theory undergo rapid transitions from one state to another, and where discontinuous transitions are involved we can locate distinct minima in the free energy corresponding to them. The region (I) of positive $\bar{u}_2$ and comparatively small dispersions of disorder Δ corresponds to the extended Flory coil state akin to the analogous homopolymer state. Decreasing $\bar{u}_2$ leads to globular states via a continuous collapse transition. There are three different globular states. Phase (II) describes the liquid-like globule analogous to the homopolymer one. Phase (IV) is the "glassy" phase characterised by large values of the spin glass order parameter $\overline{R_g^2 R_g^{2(c)}}$. Phase (V) corresponds to the "folded" globule characterised by smaller values of $\overline{R_g^2 R_g^{2(c)}}$ and larger values of the phase separation order parameter Ψ. The glass transition is first-order-like for sufficiently negative $\bar{u}_2$ and it becomes continuous after the tricritical point. Note that, above the collapse transition curve, the continuation of

the glass transition separates the Flory coil (I) from phase (III). One can see that on passing from the Flory coil to it R_g^2 decreases somewhat, whilst the degree of "glassiness" as exhibited by $\overline{R_g^2 R_g^{2(c)}}$ increases significantly. However, it is only on passing across the collapse curve that the coil really condenses. Therefore we believe that phase (III) is composed of relatively open coils with numerous loops formed by attractive sites. However at present we have not studied any order parameter that would establish this conjecture, although we have observed this behaviour in related Monte Carlo simulation.[9]

The "folding" transition is first order-like in the whole region. By "spinodal" curves here we understand the curves separating regions of distinct final states of kinetics. Thus, these also are manifested in a slowing-down of kinetics on approaching them from an adjacent phase. Unlike the case for the homopolymer[23] the kinetic Eqs. (33), (34) do not represent motion against the free energy gradients in the space of averaged dynamical variables $\mathscr{F}_q$, φ_{qp}. Therefore, the "spinodal" lines here are by no means special points of the Hessian matrix of the free energy. It is worthwhile placing emphasis that there is a pronounced region in the lower part of the diagram between the "folding" transition curve and its "spinodal" depicted

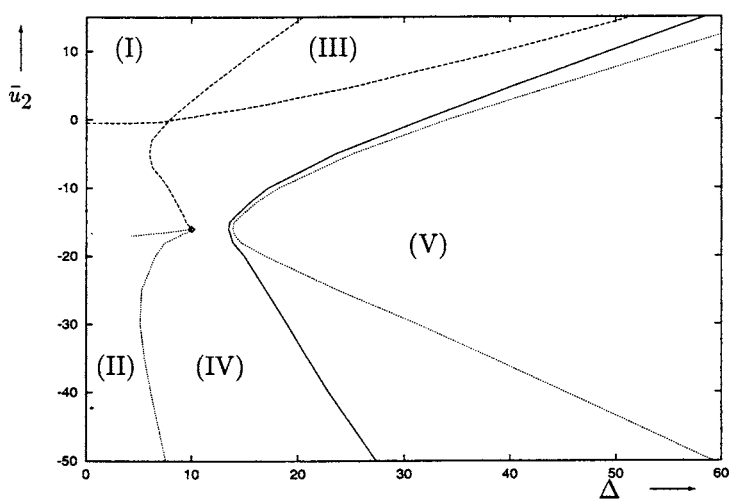


Fig. 1. The phase diagram of the model in terms of the second virial coefficient, $\bar{u}_2$, and the dispersion of disorder, Δ, obtained from the complete self-consistent Eqs. (33), (34). Solid lines represent first order-like transitions, dashed lines—continuous transitions, and dotted lines—"spinodal" curves. The Roman numerals correspond consequently to: Flory coil, liquid-like globule, random coil, "glassy" phase and folded globule. Continuous transition curves are determined by the points of the fastest change of respective order parameters ($\overline{R_g^2}$ for the collapse transition and $\overline{R_g^2 R_g^{2(c)}}$ for the glass transition). Here $N = 30$, $\kappa = 1$ and $\bar{u}_3 = 10$.

in Fig. 1. On quenching across the boundary between phases (IV) and (V) to the region where (IV) is still metastable the system will be trapped for a long time in this metastable glassy state.[2]

Finally, let us discuss the problems of the current version of the Eqs. (33), (34). The exact determination of the transition curves of first order-like transitions requires comparison of the free energy of all local minima and location of the global one. One of the problems is that the entropy (43), and consequently the free energy, corresponds to the kinetic Eqs. (33), (34) only up to $\tilde{\Gamma}^2$. Since for the glass transition the difference of the "approximate" free energy in phases (II) and (IV) appears to be of order $\tilde{\Gamma}^2$, its exact location cannot be reliably established in the present approximation. This deficiency, however, is absent for the "folding" transition, where the free energy differs between phases (IV) and (V) by order of $\tilde{\Gamma}$, what is well within the grasp of the approximation. For this reason our predictions are less accurate for the "glass" transition and its "spinodals." The "spinodal" curves in principle should be unaffected by this, as they are determined by the kinetic equations themselves (33), (34). They might though be shifted by higher order corrections in the internal energy $\mathscr{E}$. Another limitation here is that we did not account for $\bar{u}_4$ and the higher virial coefficients that should be included for high Δ and $-\bar{u}_2$ to prevent instability caused by strong two-body attraction.

4. THERMODYNAMIC LIMIT FOR DENSE GLOBULE

In this section we analyse the limit of high density globule (i.e., our further considerations are valid only in the region $\hat{u}_2 < 0$), so that $\rho^{2/3} \sim (|\hat{u}_2|/\hat{u}_3)^{2/3} \gg \kappa$, and hence one may neglect the spring term. To be precise we would like to investigate the following thermodynamic limit: $N \to \infty$ and $\kappa \to 0$ in the complete set of Eqs. (33), (34). We start by noting that for a homopolymer in this limit the equations possess only a constant solution $\mathscr{F}_q \equiv \mathscr{F} = const$, and $\mathscr{D}_m \equiv D = 2N\mathscr{F} = (4/3)(2\hat{u}_3 N/|\hat{u}_2|)^{2/3}$ for conformational modes q, $m \neq 0$. For a random copolymer we shall seek an analogous constant solution by requiring in addition that $\varphi_{qp} \equiv \varphi = const$. The latter solution though does not have the status of the exact solution because the diagonal and nondiagonal elements of this matrix require separate considerations. Nevertheless, for large N the contribution of the diagonal elements becomes negligible compared to the that of the

[2] In certain applications, such as the protein folding problem, one is interested in disorder distributions possessing better kinetic accessibility of the "folded" state. To improve it, $\tilde{\Gamma}_q$ should be optimised so that the "folding" transition and the "spinodal" are shifted towards smaller Δ and closer to each other thereby minimising the barrier height.

nondiagonal ones as may be seen from the exact numerical solution. Therefore one may still use the above Ansatz in the thermodynamic limit.

It is instructive to appeal first to the solution of the exact Eqs. (33), (34) for sufficiently small κ. In Fig. 2 we draw the behaviour of the quantity $r^2 \equiv 2N^{-2/3}R_g^2$ versus the dispersion of disorder. In total there are three branches of solutions: (II) corresponding to the liquid-like globule, (IV) corresponding to the "glassy" phase and (V)—to the "folded" globule. Analysis of the spatial correlations $\mathscr{D}_m$ presented in Fig. 8 in ref. 21 shows that this function is most convex in phase (IV), remaining non-constant even in the limit $\kappa \to 0$, as opposed to solutions (II) and (V). This allows one to expect that the constant Ansatz may not describe this "glassy" phase (IV). Evidently, as the spring constant vanishes no information remains about positions along the chain. We comment that phase (IV) has been interpreted by us in ref. 21 as existing due to the frustration on short distances along the chain after formation of locally phase separated clusters comprised by neighbouring monomers in the chain. Such an interpretation is indeed supported by the behaviour of the internal modes and order parameters. It seems natural that positions along the chain do matter for the "glassy" phase (IV). On the other hand, it is also important for this phase to distinguish the diagonal and non-diagonal elements of φ because

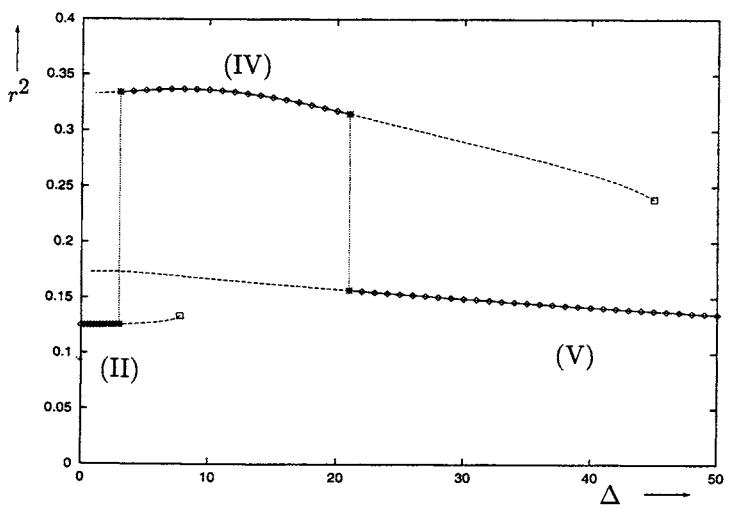


Fig. 2. Behaviour of the quantity $r^2 \equiv 2N^{-2/3}R_g^2$ vs Δ determined from the complete self-consistent Eqs. (33), (34). Solid lines correspond to the values of r^2 in the global minimum of the free energy (denoted by the same numbers as in Fig. 1), dashed lines—in metastable minima; vertical dotted lines correspond to the points of discontinuous transitions and quadrangles—to the "spinodal" points. Here $N = 31$, $\kappa = 0.1$, $\bar{u}_2 = -40$ and $\bar{u}_3 = 10$.

the former characterise the "glass" phase[21] and the latter describe the degree of phase separation, Ψ, according to Eq. (31).

Thus, given the limitations of the current procedure, let us study its consequences and compare the results here with those for the complete set of equations. It is natural to introduce the rescaled variables,

$$r^2 = \frac{D}{N^{2/3}}, \qquad \chi = \frac{N^2 \varphi}{\Delta D}. \tag{46}$$

The first definition obviously reflects the scaling of the compacted globule size on the polymer length N. The second variable could be also understood as the dimensionless degree of the phase separation. Indeed, from the definition $\Psi = N^2\varphi$ and in the case of the binary distribution and equal concentration using Eq. (32), we obtain

$$\chi = \frac{1}{2} \frac{R_g^2(B) - R_g^2(A)}{R_g^2(B) + R_g^2(A)}, \tag{47}$$

and, evidently, this is bound by the conditions

$$0 \leqslant \chi < 1/2. \tag{48}$$

Let us write out the specific energy, entropy and free energy

$$\varepsilon = \frac{\mathscr{E}}{N}, \qquad s = \frac{\mathscr{S}}{N}, \qquad a = \frac{\mathscr{A}}{N}, \tag{49}$$

as well as the mobility per monomer and the characteristic time-scale

$$\zeta_b = \frac{\zeta}{N}, \qquad \tau = \frac{t}{N^{2/3}}. \tag{50}$$

Then, the specific energy and entropy reduce to the following expressions,

$$\varepsilon = r^{-3}(\hat{u}_2(1 + \tfrac{15}{8}\chi^2) - \tfrac{3}{2}\hat{\Delta}\chi) + r^{-6}(\tfrac{4}{3})^{3/2}\hat{u}_3(1 + \tfrac{9}{2}\chi^2), \tag{51}$$

$$s = \tfrac{3}{4}k_B(4 \log r + \log(1 - 4\chi^2)). \tag{52}$$

Note that this entropy is exact in the sense that it can be derived by the summation in all orders of the recurrence relations (41), (42).

The kinetic equations following from Eqs. (33), (34) take form,

$$\frac{\zeta_b}{2}\frac{dr}{d\tau}=\frac{k_BT}{r}-\frac{1}{3}\frac{\partial\varepsilon}{\partial r}=-\frac{1}{3}\frac{\partial a}{\partial r}, \tag{53}$$

$$\frac{\zeta_b}{2}\frac{d\chi}{d\tau}=-\frac{\chi}{r^2}\left(2k_BT+\frac{1-4\chi^2}{3\chi}\frac{\partial\varepsilon}{\partial\chi}\right)=-\frac{1-4\chi^2}{3r^2}\frac{\partial a}{\partial\chi}, \tag{54}$$

and they turn out to be N-independent. It is encouraging that the time derivative of the free energy is non-positive due to the bound (48) and the relation,

$$\zeta_b\frac{da}{d\tau}=-\frac{2}{3}\left(\left(\frac{\partial a}{\partial r}\right)^2+\frac{1-a\chi^2}{r^2}\left(\frac{\partial a}{\partial\chi}\right)^2\right). \tag{55}$$

By setting the time derivatives to zero in Eqs. (53), (54) we obtain the equilibrium equations. The resulting phase diagram is presented in Fig. 3. Here (A) and (B) correspond respectively to the regions of unique and stable liquid-like globule (II) in Fig. 1; (C) and (D) correspond to the

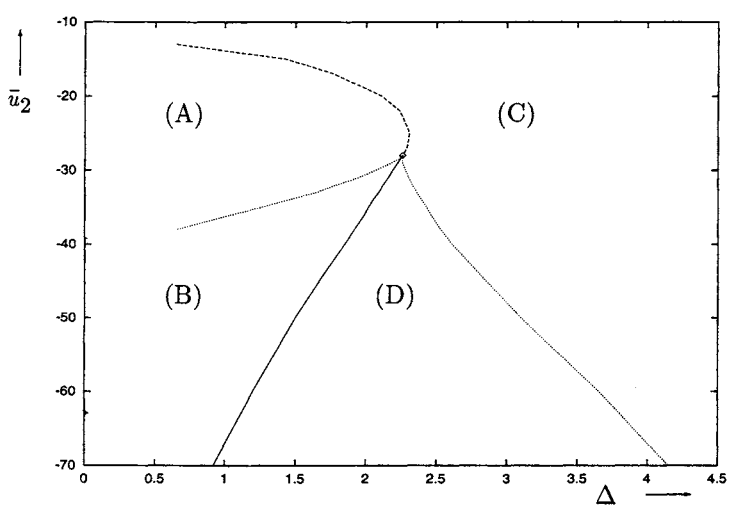


Fig. 3. The phase diagram of the model in terms of the second virial coefficient, $\bar{u}_2$, and the dispersion of disorder, Δ, obtained from the reduced self-consistent Eqs. (53), (54). Solid lines represent first order-like transitions, dashed lines—continuous transitions, and dotted lines—"spinodal" curves. The letters (A)–(D) correspond consequently to: unique liquid-like globule state, coexistence of a stable liquid-like globule and a metastable frozen globule, unique frozen globule state, and coexistence of a stable frozen globule and a metastable liquid-like globule. Here and below $\bar{u}_3 = 10$.

regions of unique and stable "folded" globule (V) in Fig. 1. In this limit increasing the dispersion of disorder causes a transition from the liquid-like globule directly to the "folded" phase (without entering the "glassy" phase (IV)). The "folded" phase, as we know, is characterized by a non-zero "glass" order parameter and large phase separation parameter. In this context the transition between (B) and (D) we shall call the "freezing" transition and the phase (D) "frozen" in agreement with the standard terminology. Really, only this transition occurs in the so-called constant density approximation widely used in the literature,[14, 15] which in some sense is analogous to our constant Ansatz. Our foregoing more general treatment elucidates the limitations of such an approximation in that it looses the intermediate "glassy" phase (IV).

Similar to Fig. 1 we have a continuous transition above the tricritical point and first order-like transition for higher two-body attraction. Thanks to the availability of a valid entropy, we now can clearly identify the transition line and both "spinodals." Still, for too large $|\hat{u}_2|$ one needs to account for the four-body and higher order excluded volume interactions. The left spinodal appears to be rather sensitive to these terms and we will not attempt to obtain the concrete laws for it. As to the second spinodal it can be easily calculated in the simple asymptotic regime of small $k_B T/|u_2|$. In this case the extremum conditions yield the equations,

$$\chi^3 - \frac{4}{45}\chi + \frac{8}{45}\frac{\hat{\Delta}}{|\hat{u}_2|} = 0, \tag{56}$$

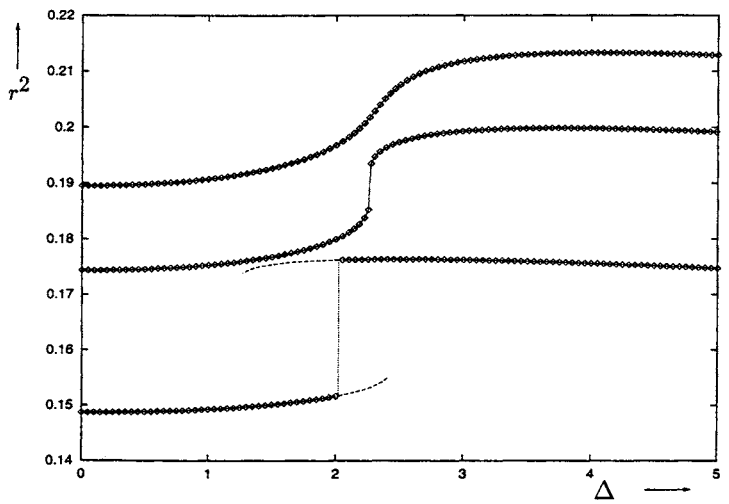
$$r^{-3} = \frac{|\hat{u}_2|\,(1 + \frac{15}{8}\chi^2) + \frac{3}{2}\hat{\Delta}\chi}{2(\frac{4}{3})^{3/2}\,\hat{u}_3(1 + \frac{9}{2}\chi^2)}. \tag{57}$$

By setting the discriminant of the cubic equation to zero we obtain the law for the right spinodal in Fig. 3,

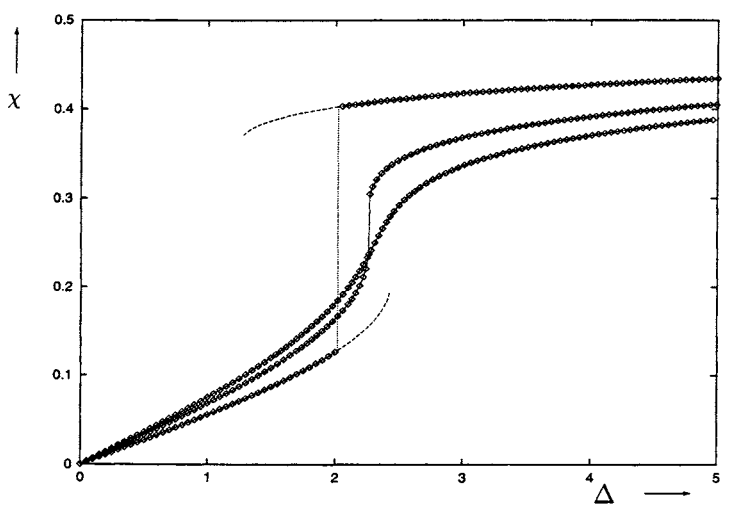
$$\hat{\Delta} = \frac{2}{9\sqrt{15}}|\hat{u}_2| \simeq 0.057\,|\hat{u}_2|, \tag{58}$$

and this estimate becomes very accurate away from the tricritical point.

As one can see from Fig. 4 the "freezing" transition here describes the transformation of a homogeneous liquid-like mixture of species to the phase separated and considerably frozen state. Freezing is manifested in the sharp decrease of the entropy, s, during the transition. The size, r, of the globule undergoes an abrupt significant growth at the "freezing" transition and then slowly decreases with Δ.

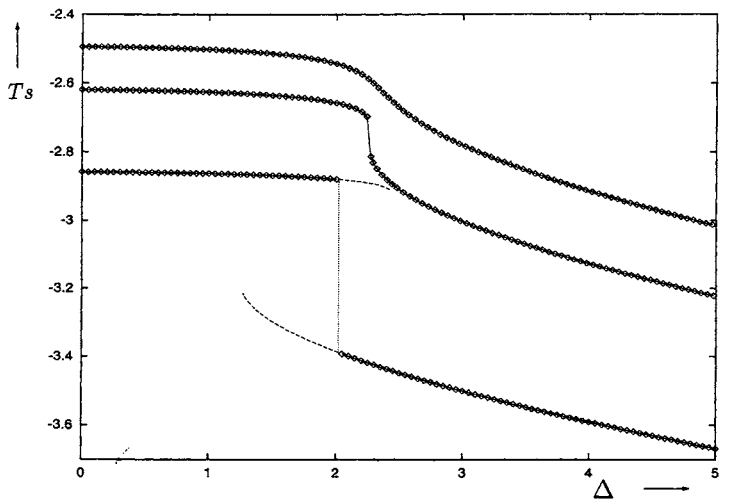


a



b

Fig. 4. Plots of the quantities r^2 (Fig. 4a), χ (Fig. 4b) and Ts (Fig. 4c) vs the dispersion of disorder Δ for three values of $\hat{u}_2$ (from top to bottom on the left-hand side of pictures): -25 (continuous freezing transition), -28 (close to the tricritical point) and -35 (discontinuous freezing transition). Solid lines correspond to the values of observables in the global minimum of the free energy, dashed lines—in metastable minima; vertical dotted lines correspond to the points of discontinuous transitions.



c

Fig. 4. (*Continued*)

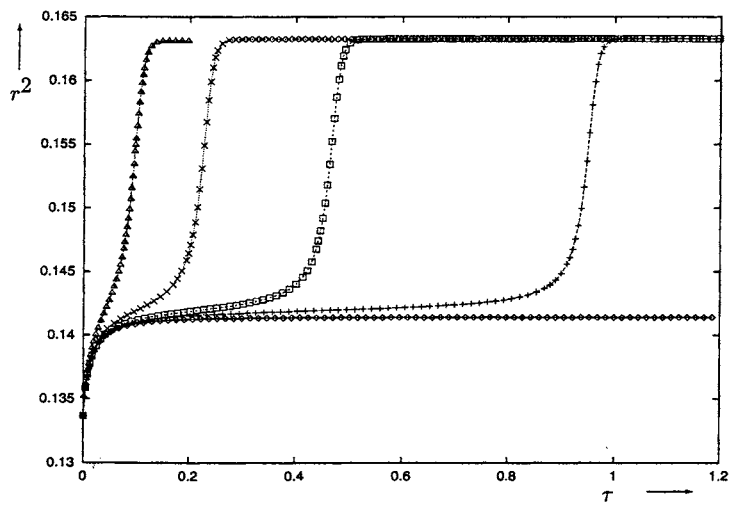
Finally, in Fig. 5 we draw the time evolution of the same quantities in kinetics after an instantaneous quench from $\Delta = 0$ (liquid-like globule) to the "frozen" phase in the vicinity of the right spinodal line. This kinetics possesses a characteristic slowing down on approaching the spinodal from the right. Numerically we find that the timescale of slowing-down diverges as,

$$\tau^{spinod} \sim |\hat{u}_2 - \hat{u}_2^{spinod}|^{-1/2}. \tag{59}$$

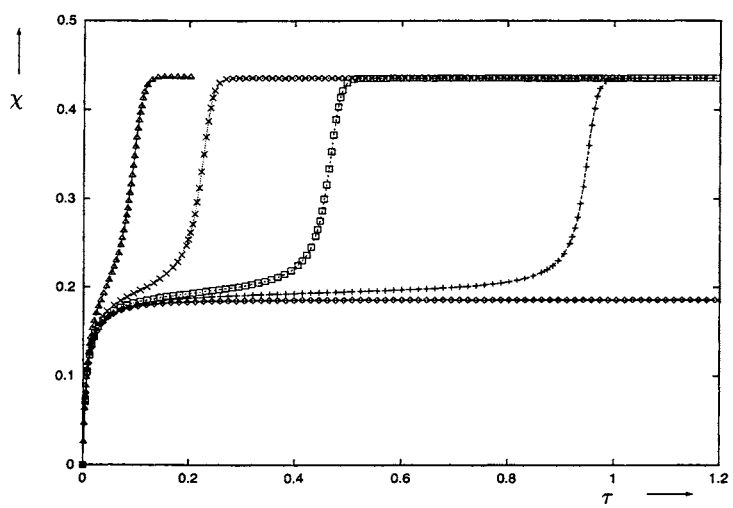
5. CONCLUSION AND DISCUSSIONS

In this paper we have generalized the Gaussian self-consistent method for random copolymers of ref. 21 to a Gaussian disorder with arbitrary translationally invariant covariance matrix and obtained the phase diagram of the model in the simplest case of constant disorder covariance matrix. The method permitted us to describe qualitatively both the extended coil (Flory and "random") and the compacted globular conventional (liquid-globule-like) and disordered ("glassy" and "folded") phases.

Now, using a simplified Ansatz for the high density globule in the thermodynamic limit we have managed to obtain some analytical predictions and to clarify the fundamental issue about the relationship between the kinetic equations and the equilibrium free energy in this method. The

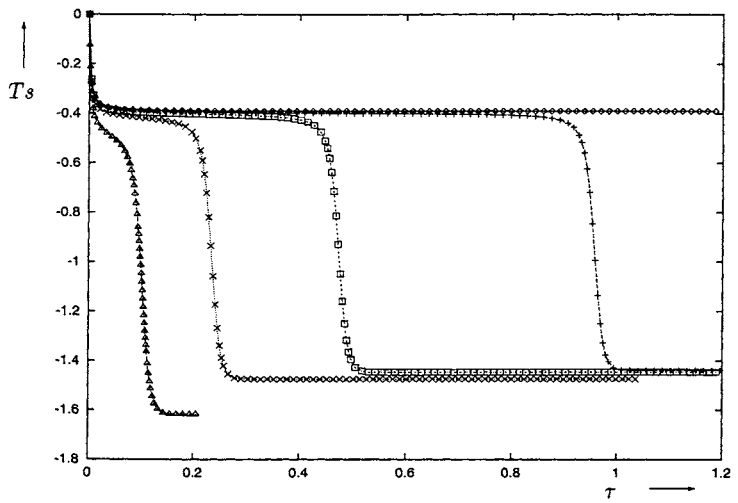


a



b

Fig. 5. Time evolution of the quantities r^2 (Fig. 5a), χ (Fig. 5b) and $T(s(\tau)-s(0))$ (Fig. 5c) for $u_2 = -40$ and different final values of Δ: 2.90 (triangles), 2.70 (crosses), 2.66 (quadrangles), 2.65 (pluses), 2.64 (diamonds). The latter corresponds to the quench to the metastable liquid-like globule.



c

Fig. 5. (*Continued*)

validity and meaning of such a simplified treatment has been subjected to scrutiny by comparing it to the numerical solution of the complete set of equations. The conclusions are mixed. On one hand, the Ansatz may be viewed as an adequate asymptotic solution for two of the globular phases. On the other hand, it cannot describe the "glassy" intermediary, for which a more general treatment is desirable. Yet, the Ansatz allows one to resolve a difficult question regarding the thermodynamic stability of the "folded" ("frozen") phase and it gives a qualitatively correct description of the "freezing" transition. In some sense our analysis indicates that one should bear in mind certain reservations while trying to apply the widely used constant density. The enforcement of such an idealization results really in a picture of freezing of a homogeneous liquid droplet rather than of a compacted polymer chain, and hence has a limited relevance to the original problem.

To reliably obtain the analytical laws for all of the transition lines and kinetics of corresponding conformational transformations we will still have to overcome the weakest point of the method. Namely, it is important to calculate higher orders in the weak disorder expansion for the energy and entropy. This seems feasible, at least, in the reduced form of the theory using the constant Ansatz. In this case, we hopefully should be able to construct a proper Flory-type theory for copolymers with quenched disorder. In any case, we have currently progressed to a higher level of understanding

of the approach and can see its advantages and limitations. Although the primary focus of this paper was methodological, we believe that this work may have a number of practical ramifications for synthetic copolymers and complex biopolymers, such as proteins. If one is to proceed further in the latter direction, however, it will be necessary to focus on more realistic disorder distributions relevant to proteins. In addition we have argued that one can fit the frustrated coil with loop structure into the picture, another interesting advance.

Finally, we now wish to point out that our whole approach can be readily generalised to the study of randomly substituted (chemically modified) cross-linked networks. This is important because this has become of very great interest recently.[24, 25] we believe that many of the states in our study have analogues in networks, and that the application of our approach would be of considerable interest in that field.

ACKNOWLEDGMENTS

The authors acknowledge interesting discussions with Professor A. Yu. Grosberg, Professor A. R. Khokhlov, Professor P. Pincus, Dr R. V. Polozov, Dr Yu. A. Rochev and our colleagues Dr A. V. Gorelov and A. Moskalenko.

REFERENCES

1. M. Mezard, G. Parisi, M. Virasoro, *Spin glass theory and beyond*, World Scientific, Singapore (1987).
2. J. des Cloizeaux, *J. Phys.* (Paris), **31**:715 (1970).
3. S. F. Edwards and P. Singh, *J. Chem. Soc. Faraday Trans.* **2**:75, 1001 (1979).
4. G. Allegra and F. Ganazzoli, *J. Chem. Phys.* **83**: 397 (1985); *ibid* **87**:1817 (1987).
5. J. des Cloizeaux and G. Jannink, *Polymers in Solution*, Clarendon Press, Oxford (1990).
6. E. G. Timoshenko, Yu. A. Kuznetsov, and K. A. Dawson, *J. Chem. Phys.* **102**:(4), 1816 (1995).
7. Yu. A. Kuznetsov, E. G. Timoshenko, and K. A. Dawson, *J. Chem. Phys.* **104**:(9), 3338 (1996).
8. F. Ganazzoli, R. La Ferla, and G. Allegra, *Macromolecules* **28**:(15), 5285 (1995).
9. Yu. A. Kuznetsov, E. G. Timoshenko, and K. A. Dawson, *J. Chem. Phys.* **103**:(11), 4807 (1995).
10. P. G. de Gennes, *J. Phys. Lett.* **46**:L639 (1985).
11. T. Garel and H. Orland, *Europhys. Lett.* **6**:(4), 307 (1988); *ibid* **6**:597 (1988); G. H. Fredrickson, S. T. Milner, and L. Leibler, *Macromolecules* **25**:6341 (1992).
12. E. G. Timoshenko, Yu. A. Kuznetsov, and K. A. Dawson, *Phys. Rev. E* **53**:(4), 3886 (1996).
13. H. Frauenfelder, in *Structure and dynamics of nucleic acids, proteins and membranes*, Ed. by E. Clementi, S. Chin, Plenum, (1986); P. G. Wolynes, In *Spin glass ideas in biology*, edited by D. Stein, World Scientific, Singapore (1991); T. E. Creighton (ed.), *Protein*

Folding, Wiley, New York (1992); R. Elber, (ed.), *New Developments in Theoretical Studies of Proteins*, World Scientific, Singapore (1994).
14. V. S. Pande, A. Yu. Grosberg, and T. Tanaka, *Phys. Rev. E* **51**:(4), 3381 (1995); *J. Chem. Phys.* **101**:(9), 8246 (1994).
15. E. I. Shakhnovich and A. M. Gutin, *J. Phys. A* **22**:1647 (1989); *J. Chem. Phys.* **93**:(8), 5967 (1990); C. D. Sfatos, A. M. Gutin, and E. I. Shakhnovich, *Phys. Rev. E* **48**:(1), 465 (1993).
16. J. D. Bryngelson and P. G. Wolynes, *Proc. Natl. Acad. Sci. USA* **84**:7524 (1987); P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, *Science* **267**:1619 (1995).
17. K. F. Lau and K. A. Dill, *Macromolecules* **22**:3986 (1989); H. S. Chan, and K. A. Dill, *J. Chem. Phys.* **95**:3775 (1991).
18. C. J. Camacho and D. Thirumalai, *Phys. Rev. Lett.* **71**:2505 (1993).
19. J. D. Honeycutt and D. Thirumalai, *Biopolymers* **32**:695 (1992); D. Thirumalai, *J. de Phys. I* **5**:1457 (1995).
20. N. D. Socci and J. N. Onuchic, *J. Chem. Phys.* **103**:(11), 4732 (1995).
21. E. G. Timoshenko, Yu. A. Kuznetsov, and K. A. Dawson, *Phys. Rev. E* **54**:(4) 4071 (1996).
22. M. Mezard and G. Parisi, *J. Phys. I* **1**:809 (1991).
23. Yu. A. Kuznetsov, E. G. Timoshenko, and K. A. Dawson, *J. Chem. Phys.* **105**:(16) 7116 (1996).
24. S. Panyukov and Y. Rabin, *Phys. Rep.* **269**:(1, 2) 1 (1996); S. Panyukov, Y. Rabin, and A. Feigel, *Europhys. Lett.* **28**:149 (1994).
25. A. R. Khokhlov and E. Yu. Kramarenko, *Macromol. Theory Simul.* **3**:45 (1994).