

Structural transformation in the collapse transition of the single flexible homopolymer model

Wenbing Hu

Department of Macromolecular Science and Laboratory of Molecular Engineering of Polymers, Fudan University, Shanghai 200433, People's Republic of China

(Received 30 January 1998; accepted 28 May 1998)

The structural transformation in the coil-globule transition of a single flexible lattice chain has been investigated using dynamic Monte Carlo simulations. The results based upon ensemble averaging illustrated that for the homopolymers with limited chain length, an intermediate state with a dense-core and molten-shell structure reversibly occurs in the transition region. It was attributed to a special microphase separation behavior in an isolated coil, performing with densifying the dense core and contracting the thin shell. The continuous appearance of the size transition and its tendency to discontinuity at the theta temperature with the chain length approaching infinity were illustrated by the coexistence curves of the monomers with limited chain length. A possible explanation and its implications to the general mechanism of protein folding are also discussed. © 1998 American Institute of Physics. [S0021-9606(98)50933-6]

I. INTRODUCTION

A flexible homopolymer is a chainlike entity consisting of many identical chemical residues (called monomers) covalently bonded together and suspended singly in a polymer dilute solution. It is a small system in comparison with the concentrated polymer solutions or their bulk state, but is still complex due to its high molecular weight and chain flexibility that can generate a tremendous number of probable conformations. Change in the solvent thermodynamic condition or, equivalently, the system temperature will make the single chain collapse from a coil state to a globule state.^{1,2} Up to now, almost all of the reports on this phenomenon described it to be a direct transition between these two states.³⁻⁵ Some early scientists regarded it as a first-order phase transition due to its similarity to the gas-liquid transition in thermodynamic nature, but later, scientists became convinced that it should be a second-order or quasi-second-order phase transition. This conviction is supported not only by mean-field theory, but also by the experimental fact that the size transition is not really sharp for limited chain length systems. However, the coil-globule transition is just a phenomenological definition based upon the change in coil size, rather than a thermodynamic definition of phase transition. So this transition process needs further investigation from the viewpoint of structural transformation within the coil.

In the experiments, the stable globule state of the isolated chains is very difficult to approach without any aggregation or precipitation of the solute molecules in the poor solvent. Recently, a stable intermediate state has been observed in the vicinity of the size transition region by light scattering, which was explained as a model containing a dense-core and molten-shell structure,⁶ and as the so-called "disordered molten globule" state for an isolated chain.⁷ Generally, the molten globule state has been readily found in protein folding and is of importance for some physiological processes,⁸ but its thermodynamic basis is not yet clear.

In this paper, we have performed dynamic Monte Carlo simulations of a single homopolymer chain on the lattice space, in order to illustrate the special structure of the intermediate state. As we shall see in Sec. III, the results show that this equilibrium state containing the distinct core-shell structure can only occur after ensemble averaging, and can be explained as the result of a microphase separation behavior in an isolated coil. The variability of the chain length for the sample system was explored in order to understand the importance of the theta point. Furthermore, the implication of the microphase separation behavior to the protein folding is also discussed in Sec. III. Section IV gives our conclusions.

II. SIMULATION TECHNIQUE

Lattice chain models have been widely and successfully used to represent flexible chain systems. A definite number of consecutive occupations on the lattice sites can represent a homopolymer chain, where the occupied sites represent identical monomers and the other empty sites represent voids or solvent molecules.

A dynamic Monte Carlo simulation method was employed to move the chain along the lattice sites without self-crossing, and each step of the motion is generated through a microrelaxation model⁹ which obeys periodic boundary conditions in a cubic box. The model we have used can be described as a single-site bond fluctuation model with the addition of a slithering diffusion terminated by extending the nearest kink conformation along the chain (as illustrated in Fig. 1). This model has proven to be highly efficient in relaxing local chain conformation on the lattice. Since the monomers are always located on the lattice sites with the coordination number 26, the distance between sites *A* and *B* should be defined by the minimum walk steps, which may either be along the lattice grid or along the diagonal lines. On the Cartesian reference frame, this distance will be the maxi-

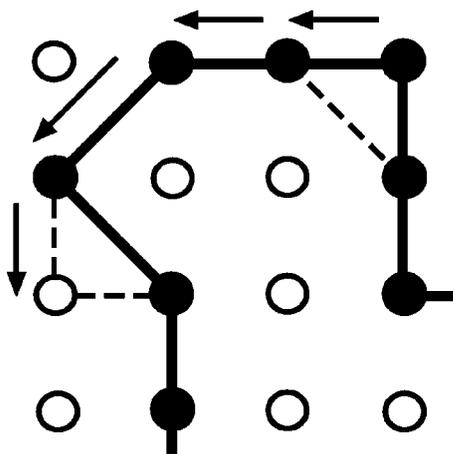


FIG. 1. Schematic presentation of the microrelaxation model. The solid circles represent monomers connected by solid bonds, the dashed lines are new bond positions produced in a step of model evolution, and the arrows are the directions of monomer jumping.

imum absolute value of $(x_A - x_B)$, $(y_A - y_B)$ and $(z_A - z_B)$. So the sites which have the same distance to the same original site form a cubic box on the lattice space. If the mass center of a coil is approximated to the nearest-neighbor site, the radial distribution of the monomers and the radius of gyration of this coil can thus be properly calculated.

The sample system was generated by presetting a regularly folded chain containing 512 (or 256, 128, or 64 for variability of the chain length) monomers on a 64-sized cubic lattice,¹⁰ then permitting it to relax to the equilibrium state under athermal solvent conditions. The dynamic relaxation time is about 148 Monte Carlo steps (MCS), which was calculated⁹ from the integration of the nonlinear relaxation curve of the square radius of gyration up to 1000 MCS, with the curve averaged over 10 000 individual observations. Even though the relaxation is so fast, about 6×10^5 MCS was still done to make sure of the equilibrium for the sample system. Here, 1 MCS was defined as the averaging once attempt of each monomer for jumping.

The athermal equilibrium state can be regarded as a state with infinitely high temperature. Through a step cooling process starting from this state, the chain began to collapse. A well-known METROPOLIS importance sampling method¹¹ was employed which sets the movable probability as 1 if the energy barrier ϵ of each microrelaxation step is less than zero, or as $\exp(-\epsilon/kT)$ otherwise, while k is the Boltzmann constant and T is the temperature. The energy barrier was determined by the cohesive energy change B for each monomer-solvent contact multiplied by the pure number of such contacts generated in each microrelaxation step, while $B = E_{ms} - (E_{mm} + E_{ss})/2$ according to the quasichemical approximation.¹² Here, E represents the two-bead cohesive energy and the subscripts m and s represent the monomer and solvent molecule, respectively. So the van der Waals interaction between monomers was expressed as a hard core repulsion plus the nearest-neighbor attraction. If we assume that B is always positive and insensitive to the temperature, the stepwise increase of the reduced temperature B/kT in units of 0.005 from zero with about 1.6×10^6 MCS relax-

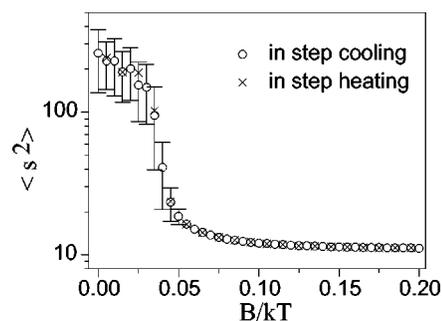


FIG. 2. The heating and cooling curves of the mean square radius of gyration vs reduced temperatures. The error bars that are less than the symbol size are not shown.

ation at each step will correspond to a very slow cooling process. To provide further details, we first let the system relax for 6×10^5 MCS at each step, then the structural properties are calculated as an ensemble average result over a subsequent 10 000 observations of the sample system with a 100 MCS interval for each. The reversed heating process was also performed with a stepwise decrease of B/kT in units of 0.01 from 0.195 in order to check the reversibility of the states obtained in the cooling process.

III. RESULTS AND DISCUSSION

A. Collapse transition of the single chain with 512 monomers

The coincidence of the cooling and heating curves in the coil size has been obtained in Fig. 2, which implies that at each temperature step the coil size has approached its equilibrium value. Just as with the observations of past experiments and simulations, the size transition is not really sharp.

The structural information of the coil includes the spatial distribution of the monomers and their exact physical states. The latter can no longer be properly described by the density or monomer concentrations because the monomers are distributed inhomogeneously on the space. The physical states of the monomers can be divided into their energy states and their momentum states. Here, we only pay attention to the exact energy state of each monomer, which is the occupation number q of the nonbonded monomers on its coordination sites and can range from 0 to 24 for the middle monomers or from 0 to 25 for the chain-end monomers.

Furthermore, the “phase” terminology should be reconsidered. It is ‘a portion of a physical system (liquid, gas, solid) that is homogeneous throughout, has definable boundaries, and can be separated physically from other phases’.¹³ Here, the homogeneity refers to the exact physical states of the monomers, but not the density.

We take three typical states into consideration for the collapse transition process. The results are shown in Fig. 3. The snapshot of the coil state at the reduced temperature $B/kT=0$ clearly shows the inhomogeneity of the monomer distribution and the anisotropy of the coil shapes. But the energy state distributions of the monomers located at the different radial distance r are quite similar after ensemble

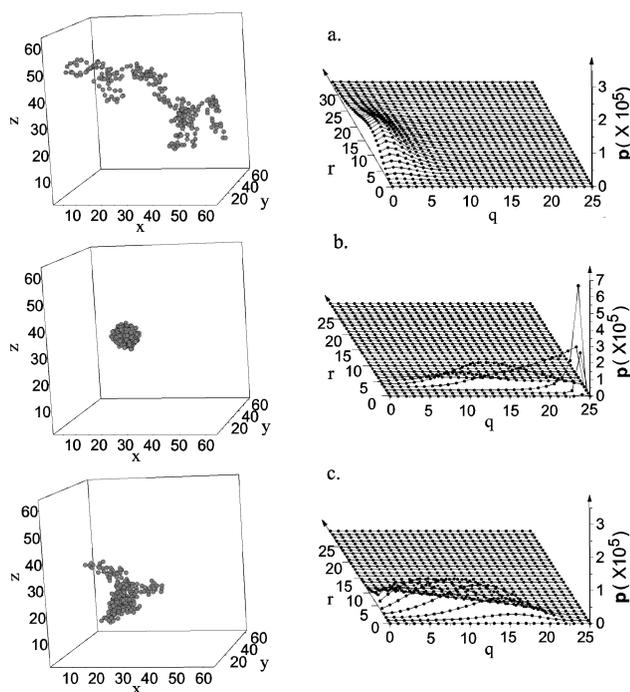


FIG. 3. The snapshot and the energy state (q) distribution along the radial distance (r) of the sample coil for three typical states, p is the monomer population counted from 10 000 observations, the solid lines are drawn to induce the eyes: (a) at $B/kT=0$; (b) at $B/kT=0.190$; (c) at $B/kT=0.045$.

averaging. So we can say that the energy state of the monomers is homogeneous within the coil, and this state is called the coil phase.

The snapshot of the globule state at $B/kT=0.190$ clearly shows the chain conformation as a small dense sphere. The energy states of the monomers within this sphere are also homogeneous except for the thin surface layer of the sphere. This state can be called the globule phase.

The snapshot of the coil state at $B/kT=0.045$, which is just located at the size transition region, shows some cilia or loops of the partial chains surrounding a relatively dense core, and sometimes looks like a crab. After ensemble averaging, the energy state distributions of the monomers are quite similar to each other in the vicinity of the coil mass center, but apparently different from those far from the mass center, the latter seem to keep still in the coil state. We can thus define these two distinct distributions of the physical states along the radial distance as coexisting two phases.

The change of the thermodynamic condition, which drives the single chain to collapse, also drives the phase separation behavior in the concentrated polymer solutions. If the phase separation behavior is considered to be the performance of the monomers and solvent molecules, it is possible that this behavior can also exist in the single coil. In that case, the size of each phase may be so small that, referring to the similar case of diblock copolymers, it would be better to refer to it as a microphase separation behavior.

The spatial distribution of the monomers within the molten shell is very inhomogeneous and sparse just like the monomers in the coil state. So the stretched cilia or loops may contribute little to the radius of gyration R_g for the whole coil, but have apparent contribution to its hydrody-

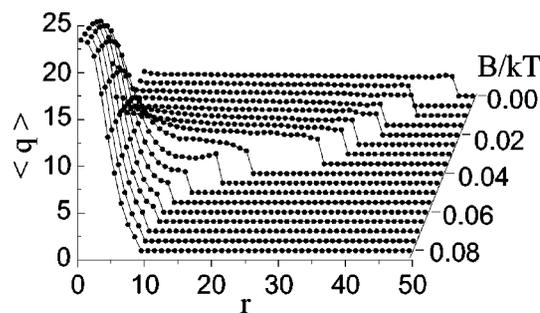


FIG. 4. The radial distribution of the mean energy states ($\langle q \rangle$) of the monomers at variable reduced temperatures. At the radial distance r far from the mass center, the monomer population which is less than 1 000 is treated as zero in the calculation. The solid lines are drawn to induce the eyes.

namic radius R_h . Therefore one can observe a valley of the ratio R_g/R_h at the size transition region by light scattering experiment.⁶

If mean values are employed to represent the exact energy states of those monomers at each radial distance, the radial distribution of the energy states changing with the system temperatures can thus be obtained in Fig. 4. The coexisting phases with distinct energy states clearly occur in the size transition region.

Along the radial distance, the wide transition range between coexisting phases does not mean a thick interface. This is due to the ensemble average over the anisotropical shapes of the dense cores with reference to the mass center of the individual coils.

From Fig. 4, one can also find that the size of the dense core is not sensitive to the system temperature, and the microphase separation behavior is performing mainly through the core being densified and the shell contracted simultaneously. This may be due to the following mechanism: When a coil collapses below the theta temperature, it will have lower conformational entropy of the chain. Parts of the chain that locate at the coil surface may expand to release the conformational entropy. They are quite stable and generate a dilute shell after ensemble averaging. When the system temperature becomes lower, the driving force to the monomer condensation is larger. It makes the core be further densified and only the shorter part of the chain can expand on the coil surface, others will be attracted into the dense core. Thus the thickness of the shell becomes smaller, until almost all the monomers are attracted into the dense core.

If we choose the mean values of q at the radial distance $r=1$ to represent the physical state of the dense core and that at $r=10$ to represent the physical state of the molten shell, the coexistence curves can thus be drawn in Fig. 5. The coexistence states are stable over the whole temperature range due to their overlap with the results obtained from the heating process.

From Fig. 5, three typical kinds of chain states can be observed over the whole temperature range. At the first stage, the monomers within the isolated coils remain in a homogeneous physical state until the second stage where the microphase separation process takes place within each coil; after all the monomers have moved into the dense core, they

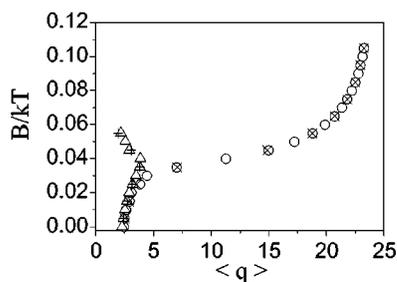


FIG. 5. The coexistence curves of the monomers in the sample coil. The open circles and triangles denote the mean values of the energy state at radial distance $r=1$ and $r=10$, respectively, in the cooling process; the crosses denote the corresponding results in the heating process.

become homogeneous again within the dense spheres (the third stage).

From Fig. 5, one can also find that the beginning of the distinct two-state coexistence seems to originate in the vicinity of the critical point of the coexistence curves. This may be the reason why the size transition of a flexible chain always appears to be continuous or to be a second-order phase transition.

B. Collapse transition of the single chain with different chain length

In the classical mean-field lattice theory, the theta point, where the excess entropy of the chain is just compensated by its excess internal energy, has been found to be just the critical point of the condensation transition of the polymer solution when the chain length approaches infinity. Does it still have any physical importance to the collapse transition of the limited chain length systems? We can study this problem by changing the chain length doubly.

The experimental definition of the theta point is the state while the second virial coefficient vanishes. But this definition has caused some disputes because the third virial coefficient is not really vanishing at this time.¹⁴ The theory¹⁵ and the simulations¹⁶ have also found that the theta state temperature will be changed with a chain length according to the experimental definition. In fact, as we have discussed above, this definition is not so good for the dilute polymer solution due to the spatial inhomogeneous distribution of the mono-

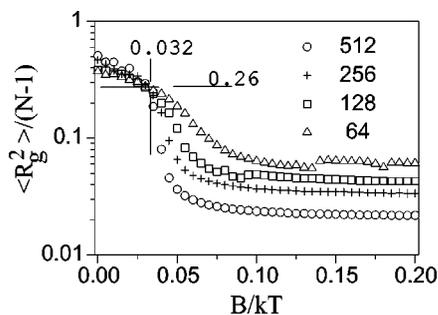


FIG. 6. The cooling curves for the ratio of the mean square radius of gyration to the chain length with the different chain length. The denoted numbers are the numbers of the monomers N on the chain. The crossed lines indicate the position of the theta point.

TABLE I. Athermal coil sizes on the 32-sized cubic lattice with the density 0.9375.

N	$\langle R^2 \rangle$	$\langle R_g^2 \rangle$	$\langle R^2 \rangle / \langle R_g^2 \rangle$	$\langle R_g^2 \rangle / (N-1)$
4	4.25	0.793	5.36	0.264
8	10.2	1.82	5.60	0.260
16	22.7	3.88	5.85	0.259
32	47.4	8.00	5.93	0.258
64	99.6	16.5	6.04	0.262
128	196	32.7	5.99	0.257

mers. Here, we use an alternative definition which considers the theta state of the chain as the unperturbed state according to their equality in the coil size.

The unperturbed chains contain a linear scaling law of their mean square radius of gyration to the chain length. Figure 6 clearly shows this linear scaling relation existing at the state point with the reduced temperature 0.032 and the ratio 0.26. This scaling relation has been demonstrated not only by the experiments but also by the recent off-lattice Monte Carlo simulations.¹⁷ The coil sizes along each cooling curve in Fig. 6 are stable according to their overlap with the corresponding heating curve.

We also performed the simulations of the athermal multiple chain systems with the density up to 0.9375. The mean square end-to-end distance $\langle R^2 \rangle$ and the mean square radius of gyration $\langle R_g^2 \rangle$ are calculated from enough observations of the multiple chain systems to make sure that the statistical uncertainty of $\langle R^2 \rangle$ should be less than 1%. Each observation has a 500 MCS interval to the next in order to eliminate their system correlation. The results averaged over more than 8000 chains are shown in Table I, where the ratio $\langle R^2 \rangle / \langle R_g^2 \rangle$ and the size scaling of the chains has demonstrated their Gaussian-like unperturbed state in the bulk system. The mean value of the ratio $\langle R_g^2 \rangle / (N-1)$ is 0.26, which identifies the above simulation result of the single chain system.

The coexistence curves with different chain length can be obtained in the same way. In Fig. 7, only the dense side of the coexistence curves are drawn for clarity. One can find that, above the theta temperature, the energy states of the monomers with different chain length are similar in the coil state, but below it, the diversity immediately occurs. This implies that the theta point still has critical importance to the collapse transition of the limited chain length systems.

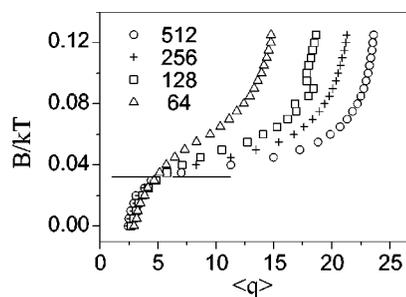


FIG. 7. The dense side of the coexistence curves for the coil with the different chain length. The data are read at the radial distance $r=1$ and the line indicates the position of the theta point.

From Fig. 7, one can also find that, the longer the chain, the narrower the transition region. Can we conclude that the collapse transition will tend to be really sharp if the chain length approaches to infinity? This may be possible if we think that the microphase separation behavior in the isolated coil originates from its limited chain length and its interface to the pure solvent according to our above explanation.

When the supercooling with reference to the theta temperature is not high, the critical size of the nuclei will be so large that nucleation in the small coil becomes impossible. An alternative path may occur before nucleation, that is the microphase separation behavior within the coil. This behavior generates a stable intermediate state in the collapse transition process, while partial chains expand on the surface in order to release the conformational entropy. The thickness of the molten shell depends on the density of the monomers in the core. Shorter chains will have difficulty densifying the smaller core and then contracting the thin shell unless under higher supercooling, so it performs the collapse transition over a wider temperature range.

When the chain length approaches to infinity, no interface of the densified coil exists. So the collapse transition may tend to perform like the normal gas-liquid condensation transition. The theta temperature will be the critical point of this tendency. Most of the theories,³⁻⁵ which can describe the collapse transition of a flexible chain, have predicted a second-order or quasi-second-order phase transition for the infinite chain length system. But they are more or less based upon the mean-field assumption or smooth distribution of the matter on the space, without any detailed analysis of the coil structure. Further theoretical construction should consider this speciality.

C. Implication to the mechanism of protein folding

It is well-known that phase transition behavior under small supercooling and small size of the system is a very slow process. Though a kinetic process may be quite different from the thermodynamic process, the above three-stage mechanism seems to correspond to three kinetic stages of the collapse transition, where the middle stage is relatively slow.^{18,19} For the folding process of the protein molecules, this kind of slow process and the special structure of the intermediate state makes some fast processes become possible to form an unified tertiary structure, such as the pair selection of the sequences due to their strong special interactions, and the parallel or antiparallel arrangement of the rigid helices due to their anisotropic hydrodynamic interaction, the latter may be through a liquid-crystallization process. Based upon the above consideration, one may construct a sequential description later for the formation of the molten globule states in the protein folding.

IV. CONCLUSION

Three stable stages clearly occur in the structural transformation of the collapse transition process—the coil state, the disordered molten globule state and the globule state.

The intermediate state contains a distinct core-shell structure after ensemble averaging, and is attributed to a microphase separation behavior in the coil. The coexistence curves of the monomers are quite unusual and demand further consideration from the thermodynamics of the small systems without any assumption of the smooth spatial distribution of the matter.

The microphase separation behavior in the single coil can be explained as the conformational entropy release of the partial chains on the coil surface. It stabilizes the intermediate state of the coil and makes the size transition appear continuous.

The critical importance of the theta point to the collapse transition of the limited chain length systems may exist. At this critical point, the collapse transition tends to be really sharp when the chain length approaches infinity.

The above results may supply us a new way of analyzing the thermodynamics of the molten globule state in protein folding.

ACKNOWLEDGMENTS

The author is very indebted to Professor Chi Wu for stimulating this study, Professor Renyuan Qian for helpful discussions and Professor O. B. Ptitsyn for kindly supplying a reprint of Ref. 8. He also thanks Professor Yuliang Yang and Dr. Jianming Lu for helpful guidance about the microrelaxation model. The financial support of CNNSF for youngster is appreciated too.

¹P. J. Flory, *J. Chem. Phys.* **17**, 303 (1949).

²W. H. Stockmayer, *Makromol. Chem.* **35**, 54 (1960).

³I. M. Lifshitz, A. Yu Grosberg, and A. R. Khokhlov, *Rev. Mod. Phys.* **50**, 683 (1978), and reference therein.

⁴C. Williams, F. Brochard, and H. Firsich, *Annu. Rev. Phys. Chem.* **32**, 433 (1981), and reference therein.

⁵A. Yu. Grosberg and D. V. Kuznetsov, *Macromolecules* **25**, 1970 (1992), and reference therein.

⁶C. Wu and S. Zhou, *Phys. Rev. Lett.* **77**, 3053 (1996).

⁷O. B. Ptitsyn, in *Protein Folding*, edited by T. E. Creighton (Freeman, New York, 1992), p. 265. This terminology is employed to distinguish the present case from the native molten globule state which contains quasitertiary structure of proteins. Other disordered molten globule states of the heteropolymer may be caused by some special interactions such as the hydrophobic interactions in the proteins.

⁸O. B. Ptitsyn, *Adv. Protein Chem.* **47**, 83 (1995).

⁹K. Kremer and K. Binder, *Comput. Phys. Rep.* **7**, 259 (1988).

¹⁰This lattice size is large enough because the properties of the athermal systems have been calculated on the same lattice space without boundaries, and the results did not show apparent diversity. But this kind of check becomes more than ten times slower due to losing the memory of the void sites on the lattice.

¹¹N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).

¹²E. A. Guggenheim, *Mixtures* (Clarendon, Oxford, 1952), p. 38.

¹³*McGraw-Hill Dictionary of Physics and Mathematics*, edited by D. N. Lapedes (McGraw-Hill Book, New York, 1978), p. 732.

¹⁴P. G. deGennes, *J. Phys. (France)* **36**, L55 (1975).

¹⁵C. Domb, *Polymer* **15**, 259 (1974).

¹⁶K. Kremer, A. Baumgartner, and K. Binder, *J. Phys. A* **15**, 2879 (1981).

¹⁷A. Milchev, W. Paul, and K. Binder, *J. Chem. Phys.* **99**, 4786 (1993).

¹⁸A. Sali, E. Shakhnovich, and M. Karplus, *Nature (London)* **369**, 248 (1994).

¹⁹B. Chu, Q. Ying, and A. Yu. Grosberg, *Macromolecules* **28**, 180 (1995).