

# Monte Carlo simulations on an equilibrium globular protein folding model

(protein conformational transition/native state/ $\beta$ -barrel/phase-space reduction/denatured state conformation)

ANDRZEJ KOLINSKI\*, JEFFREY SKOLNICK, AND ROBERT YARIS

Institute of Macromolecular Chemistry, Department of Chemistry, Washington University, St. Louis, MO 63130

Communicated by S. I. Weissman, June 2, 1986

**ABSTRACT** Monte Carlo simulations were performed on a diamond lattice, globular protein model in which the trans conformational state is energetically favored over the gauche states (thereby perhaps favoring a  $\beta$ -sheet secondary structure) and in which nonspecific nonbonded nearest-neighbor attractive interactions are allowed. If the attractive interactions are sufficiently weak that the molecule possesses a relatively high fraction of trans states in the denatured state, then on collapse, a  $\beta$ -barrel tertiary structure, highly reminiscent of the "native" structure seen in  $\beta$ -proteins, spontaneously forms. If, however, the attractive interactions are dominant, a coil-to-random globule collapse transition is observed. The roles of short-, medium-, and long-range interactions and topological constraints in determining the observed tertiary structure are addressed, and the implications and limitations of the simulations for the equilibrium folding process in real globular proteins are explored.

The development of an equilibrium statistical mechanical theory capable of producing the native conformation of a globular protein starting from the primary sequence is a long-sought objective of biophysical chemistry (1, 2). The native structure may possess ordered regions composed of  $\alpha$ -helices and/or  $\beta$ -sheets connected by loops or bends juxtaposed in a specific and complicated three-dimensional arrangement. Any successful theory must incorporate the role of short-, medium-, and long-range interactions as well as topological constraints inherent in a structure formed from stiff linear regions ( $\alpha$ -helices or  $\beta$ -sheets) joined by bends. Thus, the development of even a phenomenological statistical mechanical theory of the globular protein native-to-denatured transition is very complicated and is unfortunately beyond existing capabilities. In this paper, we introduce a series of Monte Carlo simulations of well-defined simple models designed to guide the development of a qualitative understanding of various aspects of globular protein folding as well as to provide a proving ground for the testing of statistical mechanical theories suggested by the simulation results.

In the past, equilibrium Monte Carlo simulations have been used for the folding from a specific native conformational state to the denatured state or vice versa (3, 4). However, to reduce the amount of computer time, specific biasing toward the native structure is used, rather than having the interactions themselves independently dictate the native structure. For complicated models of real proteins, this is necessary to keep the requisite computer time within practical bounds. On the other hand, with this approach one forgoes the insights that could be obtained if the system were allowed to freely hunt through all of phase space to ultimately find the native structure. In the simulations described below, we opt for the

latter approach and forgo the former inherently more detailed description of specific proteins.

Another theoretical approach models the thermodynamics of the globular protein conformational transition by a coil-to-molten globule transition (5, 6) and proceeds in the spirit of Flory's theory (7). The denatured state is assumed to be a voluminous random coil and the native state is treated as a compact random coil. That is, the distribution of chain segments is assumed to be essentially Gaussian. Hence, a most essential feature of globular proteins, the presence of highly ordered regions composed of  $\alpha$ -helices and/or  $\beta$ -sheets, is entirely neglected. (Such neglect is understandable given the difficulty of *a priori* treating the native state.) However, this by no means implies that these studies are not extremely useful. They allow one to qualitatively assess the importance of local-chain stiffness on the character of the conformational transition (8, 9). Moreover, these model calculations are consistent with the two-state model that well characterizes the equilibrium folding of small, single-domain globular proteins and can be generalized to treat proteins containing more than one domain (5). Thus, these treatments present an important zeroth order picture.

It should be pointed out that the simulation results presented below have applications to finite single-chain polymer conformational statistics; a detailed study will be presented elsewhere.

## THE MODEL

Consider a polymer chain composed of  $n$  beads confined to a tetrahedral lattice in a periodic Monte Carlo box of volume  $L^3$ . Each bead schematically represents an amino acid residue, and to include the effect of excluded volume multiple occupancy of a given lattice site is prohibited. For simplicity here all of the beads are taken to be identical, although this is not required.

A given overall chain configuration requires the specification of  $n - 3$  rotational conformational states, each of which may be either the lowest energy trans (t) or one of the two higher energy gauche ( $g_+$  or  $g_-$ ) states. The *a priori* statistical weights of trans and gauche states are given by

$$\omega_t = 1 \quad [1a]$$

$$\omega_{g_+} = \omega_{g_-} = \exp\{-\varepsilon_g/k_B T\}, \quad [1b]$$

where  $\varepsilon_g \geq 0$ . In addition, we include the possibility of an attractive interaction between nearest-neighbor, nonbonded pairs of beads of strength  $\varepsilon_a = -A\varepsilon_g$ , with  $A > 0$ . Since the polymer is confined to a tetrahedral lattice, the interaction is directed along a bond and in a sense mimics hydrogen bonding. This fact combined with the lowest-energy conformational state being trans might allow for the formation of

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

\*Permanent address: Department of Chemistry, University of Warsaw, 02-093 Warsaw, Poland.

antiparallel  $\beta$ -sheets in the ordered, low-temperature state. Thus, the present simulation focuses on the properties of a hypothetical, homopolymeric  $\beta$ -protein (2, 10). It should be pointed out that by choosing a different set of parameters (e.g.,  $\epsilon_{g_+} < 0$ ,  $\epsilon_{g_-} > 0$ ), one might conceivably produce compact structures containing right-handed  $\alpha$ -helices.

The simplicity of the model also carries with it certain limitations. In the present formulation, the *a priori* statistical weight of a gauche state of an individual bond is independent of the conformation of the nonbonded nearest neighbors. As the temperature is reduced, trans states become more favorable, and we would expect, and indeed do find, a more ordered unfolded structure (i.e., one containing a higher fraction of trans states) than is actually present in real proteins (1, 11). Nevertheless, as demonstrated below, the present model exhibits numerous qualitative features of a globular protein folding transition, and thus the study of it should hopefully prove to be useful and informative.

A successful Monte Carlo algorithm must be able to quickly and efficiently sample the phase space accessible to the chain in both the expanded, essentially random coil and the compact structures. We have used two Monte Carlo sampling algorithms, both of which produce essentially identical equilibrium properties. Basically, a cycle consists of (i) a "reptation" step [which is the most efficient part of the algorithm (12, 13)]—in one of the algorithms, one of the chain ends is randomly chosen, and the end bead is snipped off and added (at a random orientation) to the other end while, in the other algorithm, clipping proceeds from one end until the move is unsuccessful and then proceeds from the other end; (ii) end flips in which one or two bonds at a chain end are randomly rotated; and (iii) three-bond, kink motions involving (for a randomly chosen bond)  $g_{\pm} \rightarrow g_{\mp}$ .

For each particular set of parameters, long runs consisting of from  $0.3$  to  $5 \times 10^7$  reptation steps plus a corresponding number of steps (ii) and (iii) were performed. Great care was taken to ensure proper equilibration of the systems, after which equilibrium properties, described below, were calculated every 250–500 cycles; thus, a total of  $2$ – $6 \times 10^4$  states contribute to each equilibrium average. If  $n \leq 400$ ,  $L = 100$  and for the more expanded chain conformations when  $n = 400$ , we set  $L = 124$ . Additional details describing the specifics of the simulations will be published elsewhere. In the case of stiff polymers, a series of both "cooling" and "heating" runs were performed near the collapse transition to explore the existence of metastable states and to estimate the statistical uncertainty in the reported results.

## RESULTS

We begin by examining, for a chain containing  $n = 200$  beads (Fig. 1A), a plot of  $\langle R^2 \rangle / (n-1)l^2$  vs.  $k_B T / \epsilon_g$  as a function of the ratio,  $A$ , of the absolute value of the attractive interaction energy between nonbonded nearest-neighbor pairs of beads to the energy of a  $g$  state. Smaller values of  $A$  imply that local chain stiffness is relatively more important. Here,  $\langle R^2 \rangle$  is the mean square end-to-end distance obtained from the Monte Carlo simulation and  $l$  is the length of a single bond. If the temperature is infinite, then the probability of being in a gauche or a trans state is equal. The open arrow shows the value for an ideal chain (nonreversing random walk) at infinite temperature in a system entirely devoid of attractive interactions between nonbonded nearest-neighbor beads (14). The plot for a system obeying ideal chain statistics—i.e., chain dimensions obtained from the rotational isomeric model (15)—is represented by the dotted line. The solid arrow shows the corresponding chain dimensions for the athermal system—i.e., one in which  $\epsilon_a = \epsilon_g = 0$ —but where multiple occupancies of all lattice sites are prohibited. The curve having solid squares joined by a dot-dashed line

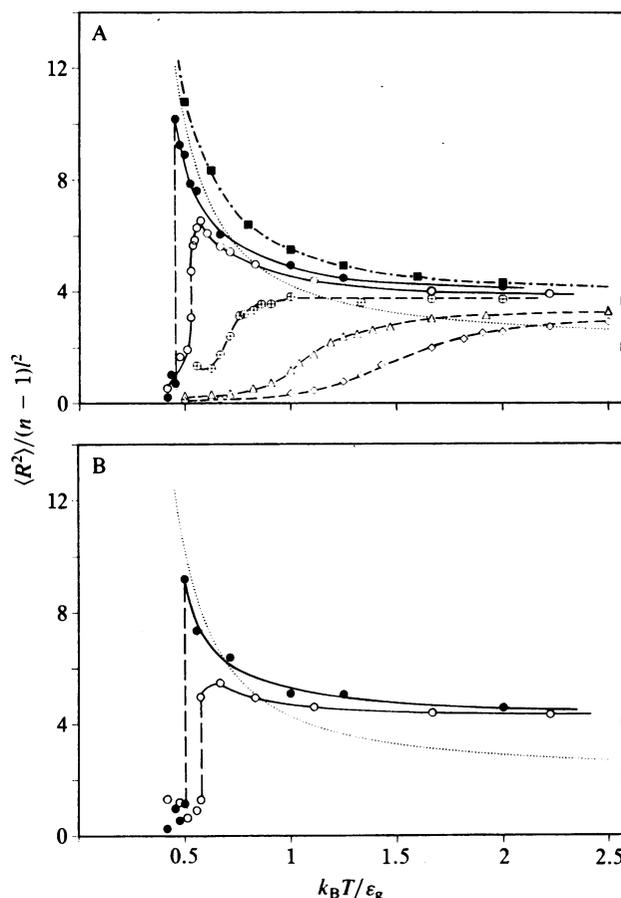


FIG. 1.  $\langle R^2 \rangle / (n-1)l^2$  vs.  $k_B T / \epsilon_g$  for chains containing  $n = 200$  (A) and  $n = 400$  (B) beads.

presents the Monte Carlo data for a system lacking attractive interactions ( $A = 0$ ) but whose fraction of trans states and concomitantly intrinsic stiffness increases with decreasing temperature. As expected, chains possessing excluded volume (dot-dashed line) but lacking attractive interactions have larger dimensions than those predicted by ideal chain statistics (dotted line).

Next we investigate (Fig. 1A) the role of increasing attractive interaction for systems possessing a given intrinsic stiffness. The dependence of  $\langle R^2 \rangle / (n-1)l^2$  on the value of  $A$  for values of  $1/4$ ,  $1/3$ ,  $1/2$ ,  $3/4$ , and  $1.0$  is represented by the curves defined by the filled circles, open circles, open octagons with a cross, open triangles, and diamonds, respectively. In Fig. 1B, we present a representative selection of plots  $\langle R^2 \rangle / (n-1)l^2$  vs.  $k_B T / \epsilon_g$  for chains containing  $n = 400$  beads (all symbols are identical to those used in Fig. 1A).

In Fig. 1A, the system having  $A = 1/2$  (open circles) clearly divides the behavior of  $\langle R^2 \rangle$  as a function of temperature into two distinctly different regimes. We continue by discussing the behavior of systems having  $A > 1/2$ . Here, the attractive interactions dominate, and we observe a continuous collapse to a compact, disordered globule. Representative high- and low-temperature configurations obtained from the simulation are shown in Fig. 2A. Physically, because of the relatively large value of  $A$ , the chain collapse occurs when there are a large number of independent statistical segments, each of length  $P$  (in units of bond length), per chain. Thus, since these chains always remain relatively flexible, the possibility of a collapse to a compact disordered structure is not at all surprising and in fact has been recognized for  $>25$  years (15). Furthermore, because of the relatively high degree of chain flexibility, the transition is continuous. It is of interest that

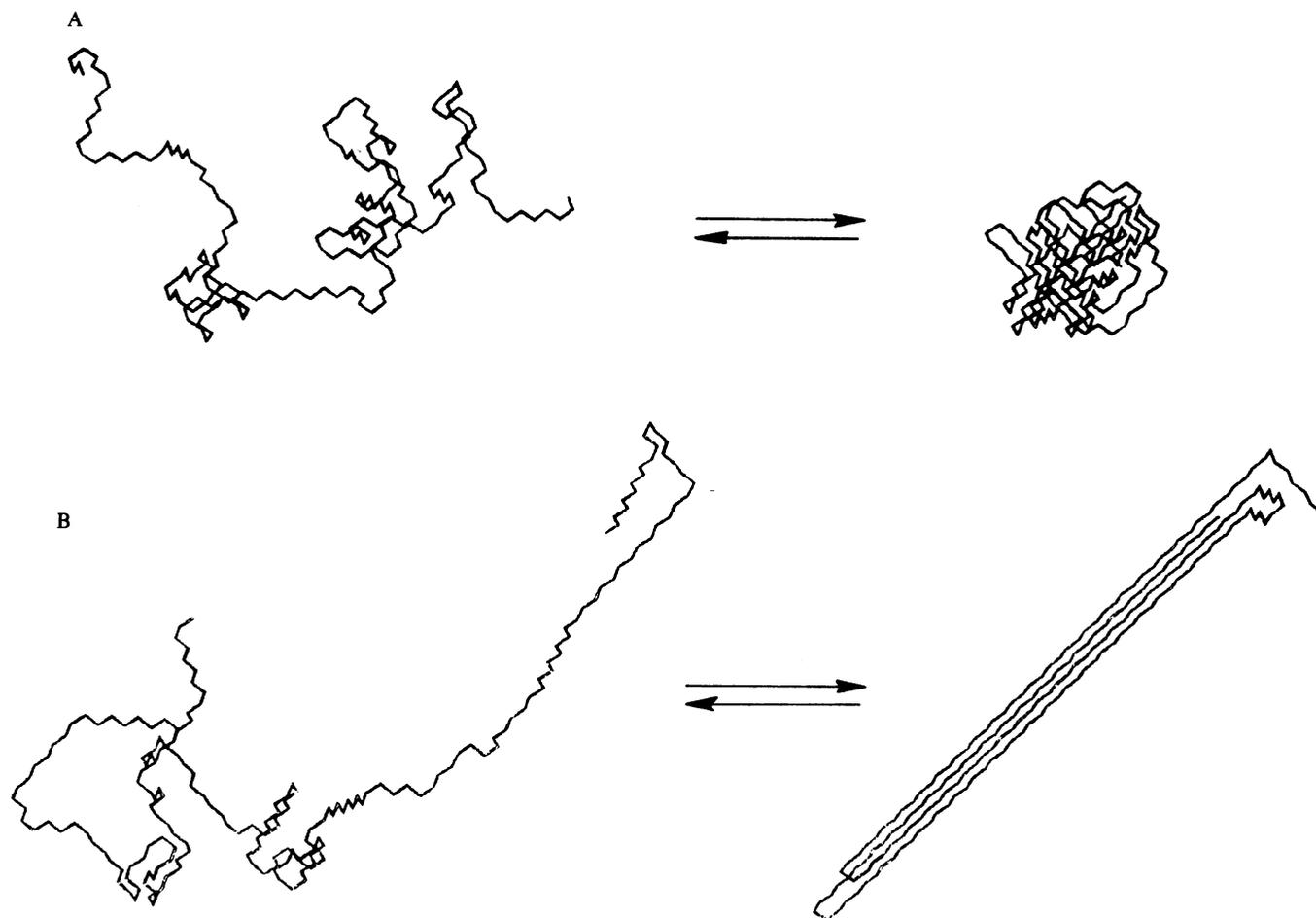


FIG. 2. (A) Representative expanded and compact configurations of a flexible polymer. (B) Representative expanded and compact configurations of a locally stiff polymer.

once the collapsed state is reached, the fraction of gauche states,  $f_g$ , becomes frozen in because of the high density of the compact structure, is essentially temperature independent, and lies considerably above the ideal chain value obtained by using the statistical weights from Eqs. 1a and b,  $f_g^0$ . For example, at  $k_B T/\epsilon_g = 0.5$ , when  $A = 1$ , the system with  $n = 200$  has  $f_g = 0.5485$ , whereas  $f_g^0 = 0.2130$ . This provides additional evidence that the attractive interactions dominate over the effect of increasing chain stiffness. In fact, by freezing the value of  $f_g$  close to that at which the coil-to-globule transition occurs, the collapsed, disordered state is locally more flexible than that predicted for the hypothetical expanded disordered state at the same temperature. Finally, if we define the  $\theta$  temperature as the temperature at which  $\langle R^2 \rangle$  from the Monte Carlo simulation equals that calculated by ideal-chain statistics (where the dotted and dashed lines in Fig. 1A cross), we see the expected monotonic decrease in chain dimensions on decreasing  $T$  below  $\theta$  (16, 17).

When  $A = 1/2$  and  $n = 200$ , over a wide temperature range,  $\langle R^2 \rangle$  is essentially independent of  $k_B T/\epsilon_g$ ; i.e., as  $k_B T/\epsilon_g$  decreases, the increase in local chain stiffness that acts to increase  $\langle R^2 \rangle$  is almost completely compensated by the attractive interactions between segments that serve to decrease  $\langle R^2 \rangle$ . Ultimately, as  $k_B T/\epsilon_g$  further diminishes, the attractive interactions dominate and this system, too, collapses to a globule that may (with the inherent uncertainties of the simulation we cannot say with absolute certainty) contain some locally ordered regions.

We turn next to the behavior of systems having  $A < 1/2$ , where over a large temperature range, the increasing local

stiffness with decreasing temperature that tends to expand the coil dominates over the attractive interactions between segments that tend to contract the coil. Hence, as shown in Fig. 1A,  $\langle R^2 \rangle$  increases as  $k_B T/\epsilon_g$  decreases, a trend that is evident even below the  $\theta$  temperature. However, these molecules are still far from the rigid rod limit that, for example in the  $n = 200$  case, gives  $\langle R^2 \rangle / (n-1)l^2 = 132.6$ . The system having  $A = 1/4$  appears to possess a discontinuous transition to a collapsed state; the less stiff system with  $A = 1/3$  has a very cooperative, but apparently continuous, collapse transition.

Representative structures for the expanded coil and the highly ordered system obtained from the simulation for the case  $n = 200$  and  $A = 1/4$  at the phase transition temperature are shown in Fig. 2B. The "native" state contains from four to six (depending on the run) antiparallel  $\beta$ -strands whose conformation is highly reminiscent of the cylindrical arrangement of  $\beta$ -pleated sheets seen in globular proteins (2). The particular conformation displayed here appears to resemble the Greek key (18) with one of the internal strands eliminated. What is remarkable is that such a highly ordered structure emerges from a system lacking any site-specific interactions. We describe below a simple physical rationalization of these observations.

Unlike the case in which attractive interactions dominate prior to the coil-to-globule collapse, prior to the transition to the ordered phase there are a relatively small number of independent segments of length  $P$ ; i.e., these molecules behave not like polymers in which  $n \gg P$  but like "small"-to "moderate"-size molecules in which  $n$  is about an order of

magnitude greater than  $P$ . (For example, in the case  $n = 200$  and  $A = 1/4$ , there are  $\approx 20$  independent segments per chain.) Because these molecules are locally stiff, prior to the collapse transition they experience a considerable reduction in accessible phase space available in the disordered state relative to a highly flexible Gaussian chain; the preformed trans stretches are indicative of a successive partitioning of configuration space that greatly expedites the search for the ordered,  $\beta$ -sheet phase of the native state. This study suggests that such partitioning might occur in the equilibrium folding process of "real" globular proteins (5).

Further insight is obtained from Table 1, which presents the mean length of a trans sequence under a variety of conditions. Comparison of column iv with column vi and of column v with column vii indicates that  $n_t$  and  $\langle n_t \rangle$  are essentially molecular weight independent. Together with the fact that the persistence length  $P$  depends on  $A$  and not on molecular weight, this implies that as  $n \rightarrow \infty$  at fixed  $A$ , the transition should become continuous and produce a collapse to a globule that lacks global ordering but possesses considerable local parallelism of constant-length trans stretches. Evidence indicative of this behavior is provided by comparing in Fig. 1 *A* and *B* the two curves having filled or open circles; the latter case with  $A = 1/3$  is most prominent. The maximum in  $\langle R^2 \rangle / (n-1)l^2$  is reduced on going from  $n = 200$  to  $n = 400$ , and the slope of  $\langle R^2 \rangle$  vs.  $k_B T / \epsilon_g$  also decreases in magnitude with increasing  $n$ .

Another striking characteristic of the conformational transition is evident from Table 1. At the transition temperature, the mean length of a trans stretch in the ordered collapsed state *increases* over that in the disordered state. In other words, the collapse itself induces additional ordering. This study points out how intersegmental interactions can perturb the conformational equilibrium between trans and gauche states. Further evidence comes from the direct examination of the values of  $f_g$  that are appreciably *less* than that in an ideal chain. For example, the  $n = 200$  chain with  $A = 1/4$  in the ordered state has  $f_g = 0.0888$  whereas  $f_g^0$  is 0.1814. While the short-range interactions (the preference for trans over gauche states) dictate that the tertiary structure is a  $\beta$ -sheet, it is in fact the "long-range" nonbonded attractive interactions that dictate the linear dimensions of the  $\beta$ -sheet. Otherwise, the mean length of a trans stretch would be essentially the same both above and below the transition, and this is not the case.

Physically, the excess stiffening of the chain in the ordered state resulting in longer trans stretches is due to the desire of the system to minimize the number of bends and maximize the number of contacts between nonbonded beads. On a diamond lattice, it takes a minimum of four bonds to reverse

Table 1. Average length of a trans sequence

| $k_B T / \epsilon_g$ | $n = 200,$<br>$A = 1$ |                       | $n = 200,$<br>$A = 1/4$ |                       | $n = 400,$<br>$A = 1/4$ |                       |
|----------------------|-----------------------|-----------------------|-------------------------|-----------------------|-------------------------|-----------------------|
|                      | $n_t$                 | $\langle n_t \rangle$ | $n_t$                   | $\langle n_t \rangle$ | $n_t$                   | $\langle n_t \rangle$ |
| $\infty$             | 1.528                 | 2.069                 | 1.528                   | 2.069                 | 1.529                   | 2.073                 |
| 2                    | 1.815                 | 2.633                 | 1.852                   | 2.723                 | 1.853                   | 2.725                 |
| 1.25                 | 1.985                 | 2.867                 | 2.117                   | 3.243                 | 2.131                   | 3.284                 |
| 1                    | 2.445                 | 3.751                 | 2.360                   | 3.746                 | 2.376                   | 3.767                 |
| 0.667                | 2.705                 | 3.856                 | 3.200                   | 5.408                 | —                       | —                     |
| 0.5                  | 2.681                 | 3.920                 | 4.661                   | 8.297                 | 4.716                   | 8.526                 |
| (0.5)                | —                     | —                     | —                       | —                     | (14.16)                 | (34.63)               |
| 0.455                | —                     | —                     | 5.593                   | 10.39                 | —                       | —                     |
| (0.455)              | —                     | —                     | (22.48)                 | (35.09)               | (15.80)                 | (35.19)               |

$\bar{n}_t$ , arithmetic average of mean length, in numbers of bonds, of an all-trans sequence.  $\langle n_t \rangle$ , weight average:  $\langle n_t \rangle = (1/f) \sum_{i=1}^n i \cdot f_{t,i}$ , where  $f_{t,i}$  is average fraction of trans states in sequence of length  $i$ . Data in parentheses are for the collapsed, ordered structure.

direction, and the bonds in the bend do not experience any attractive interactions. Moreover, bends or loops are disfavored by entropic considerations. Since the ends of the loop are constrained to be at the correct distance and orientation that allows the two  $\beta$ -strands to interact, they perform have a reduction in allowed phase space relative to the case in which any distance and orientation of the ends are allowed. This effect is known as loop entropy (19, 20) and generally acts to keep the bends fairly tight (allowing of course for lattice constraints) (21). Thus, the fraction of gauche states is reduced in the native state over what it was in the expanded, globally "denatured" state.

The above discussion suggests a simple procedure for calculating the phase transition temperature  $k_B T / \epsilon_g$ . For the denatured state, let us assume that the Helmholtz free energy,  $A_d$ , is well approximated by that of a locally stiff ideal chain, corrected for attractive interactions between nonbonded nearest neighbors. (Actually, the latter contribution is very small, but it is included for completeness.) This approximation is based on the fact that these polymers are locally quite stiff and thus excluded-volume effects should be relatively unimportant ( $\langle R^2 \rangle$  is somewhat less than ideal chain dimensions). Thus, we set

$$A_d = -k_B T (n-3) \ln[1 + 2 \exp(-\epsilon_g/k_B T)] + n\nu_d \epsilon_a. \quad [2a]$$

Furthermore we approximate the Helmholtz free energy of the ordered phase,  $A_0$ , by assuming it is merely the average attractive interaction energy between nonbonded pairs of beads and entirely neglect any entropic contribution. Since the ordered states are highly specific structures confined to a lattice, this should be a reasonable approximation. Thus,

$$A_0 = n\nu_o \epsilon_a. \quad [2b]$$

In Eq. 2 *a* and *b*,  $n\nu_d$  ( $n\nu_o$ ) is the average number of pairs of interacting beads per molecule in the disordered (ordered) phase at the transition temperature and is obtained from the Monte Carlo stimulation. Equating Eqs. 2 *a* and *b* allows us to obtain  $k_B T_i / \epsilon_g$ . In Table 2 we present results for the  $n = 200$  and  $n = 400$  systems with  $A = 1/4$  and  $1/3$ . As expected, the results improve when  $A$  is decreased (increasing the relative importance of local chain stiffness) and when  $n$  is increased at fixed  $A$  (the latter is probably a finite-size effect). Clearly for the  $A = 1/4$  case, the approximations embodied in Eqs. 2 *a* and *b* work remarkably well. For the  $A = 1/3$  case, the value of  $k_B T_i / \epsilon_g$  predicted for the  $n = 400$  case is fairly good while, for the  $n = 200$  case, in which the transition between ordered and disordered phase is continuous, the predicted value of  $k_B T_i / \epsilon_g$  is in poor agreement with the simulation. Thus, we conclude that the simple picture developed above appears to work when one has an abrupt, apparently first-order transition between a disordered, but locally stiff, coil and a very well defined  $\beta$ -pleated sheet, globular phase.

For the system with  $A = 1/3$  undergoing a continuous transition to the ordered state, we cannot tell from our Monte Carlo simulations containing a single molecule whether or not it is adequately described by the two-state model. However,

Table 2. Comparison of the transition temperatures  $k_B T_i / \epsilon_g$

| $n$ | $A$ | $\nu_d$ | $\nu_o$ | $k_B T_i / \epsilon_g$ |             |
|-----|-----|---------|---------|------------------------|-------------|
|     |     |         |         | Calc.*                 | Monte Carlo |
| 200 | 1/4 | 0.0373  | 0.463   | 0.482                  | 0.455       |
| 400 | 1/4 | 0.0395  | 0.499   | 0.494                  | 0.500       |
| 200 | 1/3 | 0.139   | 0.294   | 0.383                  | 0.529       |
| 400 | 1/3 | 0.069   | 0.472   | 0.523                  | 0.575       |

\*Obtained by equating Eqs. 2 *a* and *b*.

since the properties of these systems continuously change as a function of  $A$  for  $A < 1/2$ , and since there is a pseudo-first-order phase transition with  $A = 1/4$ , where the two-state model certainly holds, we expect it to hold for the  $A = 1/3$  case also.

### DISCUSSION

In the present series of Monte Carlo simulations, we have examined a class of model systems that possesses many qualitative aspects of the equilibrium folding transition in globular proteins. Provided that the molecule possesses sufficient local stiffness because of a relatively high fraction of trans states prior to a very cooperative transition to a condensed state, we have found that the collapsed conformation is remarkably similar to the  $\beta$ -barrel conformation seen in some globular proteins. However, caution should be exercised in overinterpreting these results in that the disordered state probably possesses greater secondary structure than in the denatured state of a globular protein. It is in fact the presence of such a residual structure that makes the transition pseudo-first-order rather than continuous (compare for example the  $A = 1/4$  case vs. the  $A = 1/3$  case). Nevertheless, even with this caveat the ordered state of these systems is highly reminiscent of the native globular protein. First, the collapse itself is seen to induce substantial additional secondary structure. Furthermore, the folding appears to proceed from states that successively sample smaller and smaller regions of phase space and that are funneled into the native state. In our case, the reduction in phase space proceeds by the preformation of local trans stretches. Finally, since even in the absence of site-specific interactions,  $\beta$ -barrel-like native structures are formed, this implies that an important factor determining the conformation of the globular protein may be the location of the loops and/or bends (22–24), with the tertiary structure then mediated by the site specificity of the subsequent long-range interactions that then induce additional tertiary structure (25).

This research was supported in part by Grant PCM-82-12404 from the Biophysics Program and Grant DMR 83-03197 from the Polymer

Program of the National Science Foundation as well as by Grant BRSG S07 RR07054-20 from the Biomedical Research Support Grant Program of the Division of Research Resources, National Institutes of Health. Acknowledgement is made to the Monsanto Company for a research grant for the purchase of a  $\mu$ VAX-II computer. J.S. is an Alfred P. Sloan Foundation Fellow.

1. Creighton, T. E. (1985) *J. Phys. Chem.* **89**, 2452–2459.
2. Ptitsyn, O. B. & Finkelstein, V. A. (1980) *Q. Rev. Biophys.* **13**, 339–386.
3. Miyazawa, S. & Jernigan, R. L. (1982) *Biopolymers* **21**, 1333–1363.
4. Levitt, M. (1982) *Annu. Rev. Biophys. Bioeng.* **11**, 251–271.
5. Dill, K. A. (1985) *Biochemistry* **24**, 1501–1509.
6. Sanchez, I. C. (1979) *Macromolecules* **12**, 980–988.
7. Flory, P. J. (1953) *Principles of Polymer Chemistry* (Cornell Univ. Press, Ithaca, NY), pp. 495–540.
8. de Gennes, P. G. (1975) *J. Phys. (Paris) Lett.* **36**, 55–57.
9. Post, C. B. & Zimm, B. H. (1979) *Biopolymers* **18**, 1487–1501.
10. Levitt, M. & Chothia, C. (1976) *Nature (London)* **261**, 552–557.
11. Tanford, C. (1968) *Adv. Protein Chem.* **23**, 121–282.
12. Binder, K., ed. (1984) *Applications of the Monte Carlo Method in Statistical Physics* (Springer, Heidelberg).
13. Wall, F. T. & Mandel, F. (1975) *J. Chem. Phys.* **63**, 4592–4595.
14. Whittington, S. G. (1982) *Adv. Chem. Phys.* **51**, 1–48.
15. Stockmayer, W. H. (1960) *Macromol. Chem.* **25**, 54–74.
16. Flory, P. (1969) *Statistical Mechanics of Chain Molecules* (Interscience, New York), pp. 49–129.
17. Kremer, K., Baumgartner, A. & Binder, K. (1981) *J. Phys. A: Math. Gen.* **15**, 2879–2897.
18. Richardson, J. S. (1977) *Nature (London)* **268**, 495–500.
19. Poland, D. & Scheraga, H. A. (1969) *Theory of Helix Coil Transitions in Biopolymers* (Academic, New York), pp. 188–191.
20. Lin, S., Konishi, Y., Denton, M. & Scheraga, H. (1984) *Biochemistry* **23**, 5504–5512.
21. Skolnick, J. (1985) *Biochem. Biophys. Res. Commun.* **129**, 848–853.
22. Venkatachalam, C. M. (1968) *Biopolymers* **6**, 1425–1436.
23. Chou, P. Y. & Fasman, G. D. (1977) *J. Mol. Biol.* **115**, 135–175.
24. Rose, G. D., Gierasch, L. M. & Smith, J. A. (1985) *Adv. Protein Chem.* **37**, 1–109.
25. Skolnick, J. & Holtzer, A. (1982) *Macromolecules* **15**, 303–314.