

# Simulations of $\beta$ -hairpin folding confined to spherical pores using distributed computing

D. K. Klimov\*<sup>†</sup>, D. Newfield<sup>‡</sup>, and D. Thirumalai\*<sup>†</sup>

\*Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742; and <sup>‡</sup>Parabon Computation, 3930 Walnut Street, Suite 100, Fairfax, VA 22030-4738

Communicated by George H. Lorimer, University of Maryland, College Park, MD, April 12, 2002 (received for review December 18, 2001)

**We report the thermodynamics and kinetics of an off-lattice Go model  $\beta$ -hairpin from Ig-binding protein confined to an inert spherical pore. Confinement enhances the stability of the hairpin due to the decrease in the entropy of the unfolded state. Compared with their values in the bulk, the rates of hairpin formation increase in the spherical pore. Surprisingly, the dependence of the rates on the pore radius,  $R_s$ , is nonmonotonic. The rates reach a maximum at  $R_s/R_{g,N}^b \approx 1.5$ , where  $R_{g,N}^b$  is the radius of gyration of the folded  $\beta$ -hairpin in the bulk. The denatured state ensemble of the encapsulated  $\beta$ -hairpin is highly structured even at substantially elevated temperatures. Remarkably, a profound effect of confinement is evident even when the  $\beta$ -hairpin occupies less than a 10th of the sphere volume. Our calculations show that the emergence of substantial structure in the denatured state of proteins in inert pores is a consequence of confinement. In contrast, the structure of the bulk denatured state ensemble depends dramatically on the extent of denaturation.**

**F**olding in cells occurs in the presence of lipids, carbohydrates, and other biological molecules, i.e., in a crowded environment (1, 2). Therefore, many proteins spontaneously fold to their native states in the environment, which geometrically restricts their conformational space (3). There are many reasons for considering folding in confined spaces. The recognition and encapsulation of a substrate protein by the GroEL molecule places it in a cylindrical cavity. Much of the annealing action takes place in a dynamic cage that is built by the interactions of GroEL with GroES (4). In the course of protein synthesis the nascent polypeptide chain is confined to a narrow “tube” (5). Such a narrow space subjects the protein to a tensile force ( $f \sim k_B T/D$ , where  $D$  is the diameter of the pore), which may play a role in mechanically squeezing the newly synthesized protein out of the ribosome. These considerations and other studies (6–9) probing the effect of confinement on protein stability have motivated the present work.

Recently, Eggers and Valentine examined the structure of four proteins encapsulated in a silica matrix (8). They showed that, in general, confinement leads to enhanced stability of the folded protein. For example, they surmised that the encapsulation of  $\alpha$ -lactalbumin results in the increase in the melting temperature,  $T_m$ , by  $\approx 30^\circ\text{C}$ . However, the relationship between confinement and enhanced stability of a folded state is not simple, because changes in the water structure in pores could alter the interactions stabilizing the native state (9). Nevertheless, it seems that confinement generally increases the stability of the native structure.

Folding in a crowded cellular environment may be modeled by confining a polypeptide chain to narrow pores. The enhanced stability in pores was explained recently by Zhou and Dill using polymer physics concepts (6). If the denatured state of a polymer chain is assumed to be a random coil, then the free energy cost to localize it in a region of size  $D$  (e.g., sphere, slit, or tube) is  $\Delta F_U \approx RTN(a/D)^{1/\nu}$  (10), provided that  $D < N^{\nu}a$ , where  $a$  may be taken to be the average distance between the  $C_\alpha$  atoms ( $\approx 3.8 \text{ \AA}$ ), and  $N$  is the number of residues. (This estimate is accurate up to a logarithmic factor in  $D$ .) The Flory exponent  $\nu$  relates the

radius of gyration of a chain  $R_g$  at  $D \rightarrow \infty$  (the size of a chain in bulk solution) to  $N$  according to  $R_g \approx aN^\nu$ . If the chain is ideal, then  $\nu = 0.5$  and  $\Delta F_U = RTN(a/D)^2$ . Because of the reduction in the translational entropy, confinement also increases the free energy of the native state, i.e.,  $\Delta F_N > 0$ . If  $\Delta F_N/\Delta F_U \ll 1$ , then localization of a protein in a confined space stabilizes the native state compared with the bulk. It also follows that there is a range of  $D$  values over which stability is maximized. These arguments and the explicit calculations of polymers confined to slits (11) suggest that excluded volume effects should not qualitatively alter the conclusions reached by Zhou and Dill (6). Because confining an unfolded chain with excluded volume leads to higher entropy costs (provided  $a < D < N^{\nu}a$ ), the change in stability estimated by assuming that the denatured state is ideal should be a lower bound.

In this article we examine the effects of confinement on the nature of the denatured state ensemble (DSE), folding rates, and stability by studying  $\beta$ -hairpin formation in spherical pores of different radii  $R_s$  (Fig. 1). We find that confinement, which may be considered as a model for the hydrophilic cavity in GroEL (12), leads to enhancement in the folding rates and stability compared with their values at  $R_s \rightarrow \infty$ . Throughout the paper we refer to the bulk hairpin when  $R_s \rightarrow \infty$  and the terms “confined” and “encapsulated” hairpins are used interchangeably. Superscripts b and c refer to bulk and confined cases, respectively.

## Methods

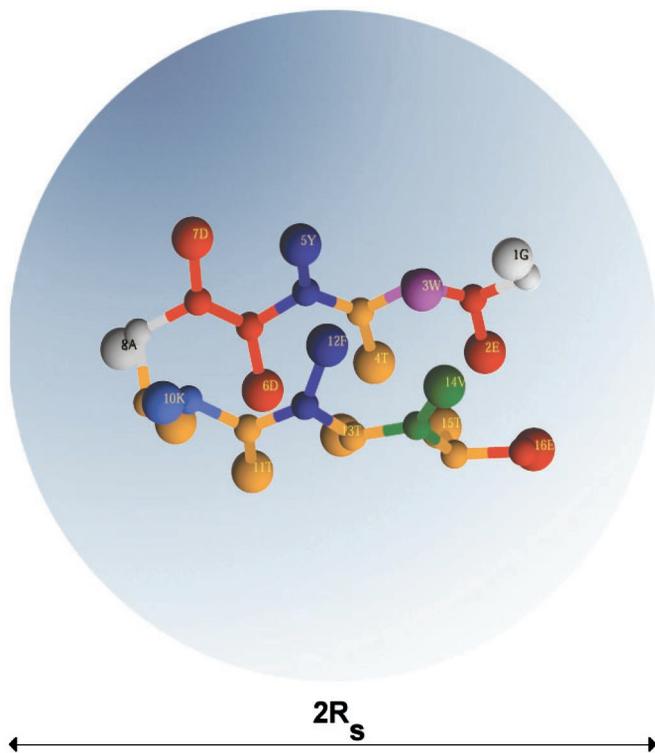
**Model.** To study confinement effects we use a 16-residue C-terminal  $\beta$ -hairpin from protein G (Fig. 1), for which experimental (13, 14) and theoretical (15) studies have been reported. We provide a brief overview of the model and simulation technique (see ref. 15 for details) focusing on the relevant changes caused by confinement.

The off-lattice coarse-grained protein model explicitly includes the backbone  $C_\alpha$  carbons and side chains, which are represented as spheres of appropriate van der Waals radii positioned at the centers of mass of amino acids. The model includes backbone hydrogen bonds (HBs) between NH and CO groups, which are treated as virtual moieties between adjacent  $C_\alpha$  atoms (15). The potential energy of a bulk conformation of polypeptide chain  $V_b$  is given by the sum of bond-length potential, side chain–backbone connectivity potential, bond-angle potential, dihedral angle potential, HB and nonbonded long-range potentials (15).

Because our goal is to study the effect of confinement on the  $\beta$ -hairpin formation, the Go model, which includes only native interactions, should suffice. Previous studies have shown that the Go models provide a qualitative description of two-state folding sequences (16–18). Therefore, we consider amino acid-dependent interactions only for the side chains that are in contact in the wild-type native structure of C-terminal hairpin

Abbreviations: DSE, denatured state ensemble; HB, hydrogen bond; HC, hydrophobic cluster.

<sup>†</sup>To whom reprint requests may be addressed. E-mail: thirum@glue.umd.edu or klimov@glue.umd.edu.



**Fig. 1.** Encapsulation of 16-mer C-terminal  $\beta$ -hairpin from the protein G in a spherical pore of the radius  $R_s$  is sketched. The surprising outcome of this work is that even confinement to relatively large pores ( $R_s/R_{g,N}^b \geq 2.0$ ) has a profound effect on the hairpin stability and folding rates. The amino acids in the  $\beta$ -hairpin are color-coded and annotated with a one-letter code.

from protein G. Side chains are in contact if their centers of mass are less than 7.6 Å apart. We also consider only native HBs.

**Confining Potential.** To study the effects of confinement, a polypeptide chain is encapsulated into a sphere of radius  $R_s$ . Confinement is achieved by assigning short-range repulsive interactions between hairpin atoms and the sphere's inner walls. The interaction between a unit surface element of the sphere located at  $\vec{R}_s$  and the hairpin atom ( $C_\alpha$  carbon or side chain) located at  $\vec{R}$  is taken to be

$$v_w(r) = 4\varepsilon_w \left(\frac{a}{r}\right)^{12}, \quad [1]$$

where  $r = |\vec{R}_s - \vec{R}|$  and  $\varepsilon_w = \varepsilon_h = 1.25$  kcal/mol ( $\varepsilon_h$  is the energy unit in the model), and  $a = 3.8$  Å is the average distance between successive  $C_\alpha$  carbon atoms. The total interaction energy of a hairpin atom with the confining sphere is

$$V_w = \int_S \rho v_w ds, \quad [2]$$

where  $\rho$  is the sphere surface density ( $\rho a^2 \equiv 1$ ) and the integral is taken over the entire inner surface of the sphere  $S$ . By integrating  $v_w$  over  $S$ , we obtain the confining potential for a hairpin atom located at  $\vec{R}$

$$V_w(R) = \frac{4\pi\rho a^2 R_s \varepsilon_w}{5R} \left[ \left(\frac{a}{R_s - R}\right)^{10} - \left(\frac{a}{R_s + R}\right)^{10} \right], \quad [3]$$

where  $R = |\vec{R}|$ . A similar approach has been used to describe monomer-wall interactions in the simulations of grafted poly-

meric brushes in polymer melts (19). The form of the confining potential in Eq. 3 was chosen because it is well defined (no singularities) for all values of  $R < R_s$ . The precise functional form of  $V_w(R)$  is not important for the conclusions of this study as long as  $V_w(R)$  is a short-ranged repulsive potential. For example, we tested the case in which Eq. 3 is replaced with Eq. 1 with  $r = R_s - R$  and found that the conclusions of the study do not change qualitatively. Furthermore, it has been shown (11) that the size of a homopolymer confined to slits is the same irrespective of the interaction potential (soft or hard) with the wall. In general, a soft wall can always be replaced by a hard wall with a slightly larger range of interactions (20). However, the physics of folding would be altered if the confining potential is long-ranged or contains attractive terms.

The potential energy of the confined hairpin is the sum of  $V_b$  and the energies  $V_w$  for all hairpin atoms. In this study we consider  $R_s \geq R_{s,0} = 4.475a = 1.40R_{g,N}^b$ , where  $R_{g,N}^b$  is the radius of gyration of the bulk native conformation. The value of  $R_{s,0}$  is the smallest radius of the confining sphere, which minimally perturbs the bulk native structure.

**Numerical Simulations.** Langevin simulations based on the velocity form of the Verlet algorithm and the multiple histogram technique were used to compute the equilibrium properties of confined hairpins (15). Folding kinetics in a sphere was obtained by using Langevin dynamics at approximately water viscosity. Hundreds of folding trajectories were generated at each value of  $R_s$  to calculate the folding rate  $k_F = 1/M \sum_{i=1}^M \tau_{i1}^{-1}$ , where  $M$  is the total number of trajectories and  $\tau_{i1}$  is the first passage time to the native state. To assess the heterogeneity of the folding pathways we computed the fraction of unfolded molecules at time  $t$   $P_u(t) = 1 - \int_0^t P_{fp}(s) ds$ , where  $P_{fp} = 1/M \sum_{i=1}^M \delta(s - \tau_{i1})$  is the distribution of first passage times (21).

**Characterization of DSE.** To characterize changes in the DSE caused by confinement, we sample the conformational space in long equilibrium simulations at  $T_s = 1.2T_F^b$ , where  $T_F^b$  is the folding transition temperature in the bulk. To assess how the DSE changes with the degree of denaturation, which is determined by the temperature in this work, the equilibrium states also were analyzed at  $T_s = 1.6T_F^b$ . The heterogeneity of the DSE was determined by applying a pattern-recognition algorithm to the DSE structures (15). This method allows us to assign DSE conformations to structural clusters based on their similarity to the native conformation. The resulting clusters represent structurally distinct denatured states. Our focus is to distinguish between the DSEs of bulk and confined hairpins.

**Distributed Computing.** To calculate folding trajectories for a  $\beta$ -hairpin, we used distributed computing. This computational approach capitalizes on the recent explosive growth in the number of internet-connected computers. Traditional computer simulations are performed on clusters of workstations or by using supercomputers. In both cases, the number of available CPUs is limited. In dramatic contrast, distributed computing harnesses scattered world-wide heterogeneous resources for concerted work. FRONTIER, the system developed by Parabon Computation on which we ran some of our simulations, is capable of distributing hundreds of thousands of individual tasks to run independently. It also is useful in handling all the scheduling, distribution, checkpointing, resumption, and collation aspects of computations. The number of individual tasks sent to the FRONTIER server for simultaneous execution is limited by the number of PCs connected to the server. This number is on the order of  $10^5$ . Since the number of individual folding trajectories (i.e., elementary individual tasks) used for computing  $k_F$  at a given  $R_s$  is  $\sim 10^3$ , virtually all trajectories were executed in parallel. In practice, it requires  $\sim 200$  h to complete 1,000

individual folding trajectories on the workstation with a single 600-MHz alpha processor. For comparison, the FRONTIER server allowed us to perform the same computations in  $\sim 2\text{--}4$  h, which results in a dramatic speedup of otherwise extremely long and intensive simulations. Distributed computing is a powerful and general computational tool suitable for a wide variety of computationally intensive applications (including those targeting biomolecules).

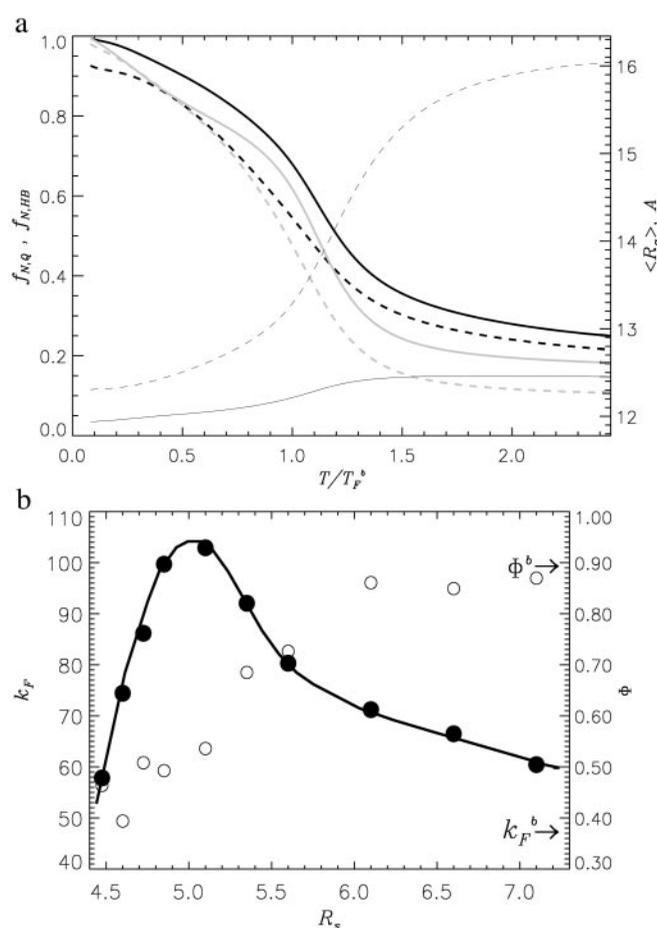
## Results

**Equilibrium Folding in Confined Space.** We describe the folding thermodynamics of the  $\beta$ -hairpin confined to a sphere with  $R_s = 4.6a$ . For the bulk hairpin the largest distance between its center of mass and any of the backbone or side chain atoms is  $3.3a$  (the native radius of gyration in the bulk  $R_{g,N}^b = 3.2a$ ). Thus, the sphere with  $R_s = 4.6a = 1.43R_{g,N}^b$  can accommodate the polypeptide with relatively small perturbations of the structure. (The value  $V_w$  for the atom closest to the sphere wall is  $0.3$  kcal/mol). The size of the native structure of the confined hairpin is  $R_{g,N}^c = 3.1a$ . Confinement introduces an additional native contact, thus increasing  $Q$ , the number of native contacts, to 22. All native contacts of the bulk hairpin are present also in the confined native conformation. The energy of interactions with the wall is merely  $1.0$  kcal/mol, and it makes minor contribution to the total energy of the native conformation ( $E_N^c = -24.4$  kcal/mol and  $E_N^b = -25.9$  kcal/mol). The overlap,  $\chi$  (22), of the confined native conformation with the bulk exceeds  $0.9$  ( $\chi = 1$  implies identical conformations). Thus, encapsulating the hairpin to a sphere with  $R_s = 4.6a$  only marginally affects the hairpin native conformation.

In Fig. 2a we show several thermodynamic functions that probe the formation of the native state for the confined and bulk  $\beta$ -hairpins. The temperature scale (horizontal axis) is given in the units of the folding temperature  $T_F^b = 308$  K (15). The thermal fractions of native contacts  $f_{N,Q}^c$  and HBs  $f_{N,HB}^c$  (solid thick black and gray lines) reveal a broad weakly cooperative folding transition. Both quantities change most rapidly (i.e.,  $df_N/dT$  reaches maximum) at  $1.1T_F^b$ , which is identified with  $T_F^c$ , the folding transition temperature for the confined hairpin. Remarkably,  $\langle R_g^c \rangle$  (solid thin black curve) changes with temperature by merely 3%. The collapse transition may be associated with the peaks in  $d\langle R_g^c \rangle/dT$  or, alternatively, specific heat  $C_v$ . Both quantities indicate that the collapse temperature is  $T_F^c \approx T_F^b = 1.1T_F^b$ . Confinement to a sphere severely limits the available conformation space and almost eliminates any variation in  $\langle R_g^c \rangle(T)$  (Fig. 2a). In effect, nearly all hairpin structures accessible in a spherical cavity are now *compact collapsed species*.

Comparison of  $f_N$  for the confined and bulk hairpins (solid and dashed lines in Fig. 2a) shows that confinement stabilizes the hairpin. The folding transition temperature  $T_F^c$  increases by  $\sim 10\%$ . The increase in the stability of the native state due to confinement is reflected in the free energy of stability  $\Delta G = -T_s \ln[P_{NBA}/(1 - P_{NBA})]$ , where  $P_{NBA}$  is the thermal probability of occupancy of the native basin of attraction (NBA). The NBA is the set of conformations, for which  $\chi \geq \langle \chi \rangle(T_F)$ . At  $T_s = 0.82T_F^b$ ,  $\Delta G$  for the bulk hairpin is  $-0.9$  kcal/mol, whereas for the confined hairpin it decreases by a factor of 2.3 to  $-2.1$  kcal/mol. The large increase in  $|\Delta G|$  is remarkable given the finite size of the model system! A qualitatively similar effect of stabilization of the native state is observed at other values of  $R_s$ . However, as expected the gain in stability decreases as the sphere radius increases (at  $R_s = 4.85a$  and  $5.1a$   $\Delta G = -1.9$  and  $-1.6$  kcal/mol, respectively). Nevertheless, even when  $R_s/R_{g,U}^b > 1.2$  [ $R_{g,U}^b$  is the radius of gyration of the bulk hairpin at  $T > 2T_F^b$  (Fig. 2a)] there is substantial gain in the stability.

**Confinement Effect on the Kinetics of  $\beta$ -Hairpin Formation.** We have computed the folding rates  $k_F$  at  $T_s = 0.82T_F^b$  for  $4.475a \leq R_s \leq$



**Fig. 2.** (a) Thermal refolding of the  $\beta$ -hairpin in a spherical pore with the radius  $R_s = 4.6a$ . The black and gray thick solid lines represent the fractions of the native contacts  $f_{N,Q}^c$  and the HBs  $f_{N,HB}^c$ , respectively. The black thin solid line refers to the radius of gyration  $\langle R_g^c \rangle$ . Dashed curves correspond to the quantities of the bulk hairpin. Comparison of the curves for the encapsulated and bulk  $\beta$ -hairpins shows that confinement eliminates expanded conformations and stabilizes the folded state (see text for details). (b) Dependence of the folding rates  $k_F$  (●) and the fraction of fast trajectories  $\Phi$  (○) on  $R_s$ . Encapsulation generally accelerates folding rates. A nonmonotonic dependence of  $k_F$  on  $R_s$ , reflects changes in the mechanisms of  $\beta$ -hairpin formation in small pores. The solid line is drawn as a guide to the eye. The arrows mark the bulk values  $k_F^b = 47 \mu\text{s}^{-1}$  and  $\Phi^b = 0.89$ .

$7.1a$  and  $R_s \rightarrow \infty$ . Initial conformations were equilibrated at  $T_h = 2.4T_F^b$ . By using the distribution of first passage times  $P_{fp}$ , the fraction of unfolded molecules as a function of time,  $P_u(t)$ , is calculated (see *Methods*). Both  $k_F$  and  $P_u(t)$  are useful probes of folding kinetics.

At the smallest value of  $R_s$  the ratio  $(R_{g,N}^b/R_s)^3 \approx 1/3$ , which would suggest that the confining sphere should not alter the  $\beta$ -hairpin formation significantly. Thus, for the range of  $R_s$  considered here,  $k_F(R_s)$  is expected to increase compared with  $k_F(R_s \rightarrow \infty)$ , because the search for the native state takes place among the restricted set of compact structures. Surprisingly, we find that the dependence of folding rate  $k_F$  on  $R_s$  is not monotonic (Fig. 2b, closed circles). As  $R_s$  is decreased from its bulk value, the folding rates increase compared with  $k_F(R_s \rightarrow \infty) \equiv k_F^b = 47 \mu\text{s}^{-1}$ . At  $R_s = 5.1a$ ,  $k_F$  reaches a maximum of  $103 \mu\text{s}^{-1}$ , which is more than twice the bulk value. Strikingly, the folding rate decreases when  $R_s < 4.85a$ . The value of  $k_F$  at the smallest  $R_s = 4.475a$  is marginally higher than  $k_F^b$ . Hence, encapsulation in a sphere generally accelerates  $\beta$ -hairpin formation over a broad range of  $R_s$  values.

Although encapsulation leads to rate acceleration, the mechanism of  $\beta$ -hairpin formation changes, especially for  $R_s/R_{g,N}^b \leq 1.5$ . In this range of  $R_s$  folding trajectories partition into slow and fast kinetic phases, such that  $P_u(t) \approx \Phi \exp(-k_{\text{fast}}t) + (1 - \Phi) \exp(-k_{\text{slow}}t)$ . The overall folding rate  $k_F(R_s) = \Phi k_{\text{fast}} + (1 - \Phi)k_{\text{slow}}$ . For large cavities the fast phase amplitude  $\Phi$  is comparable to the bulk value  $\Phi^b = 0.89$  (Fig. 2b). However,  $\Phi$  decreases for  $R_s \leq 6.0a$ . The opposing behavior of  $k_{\text{fast}}(R_s)$  and  $\Phi(R_s)$  leads to an optimal value of  $R_s$ , for which the overall rate  $k_F$  is maximum. For  $R_s \leq 4.8a$  a certain fraction of trajectories becomes transiently trapped. The transition to the native state requires partial unfolding, which becomes difficult in a tight, confining space.

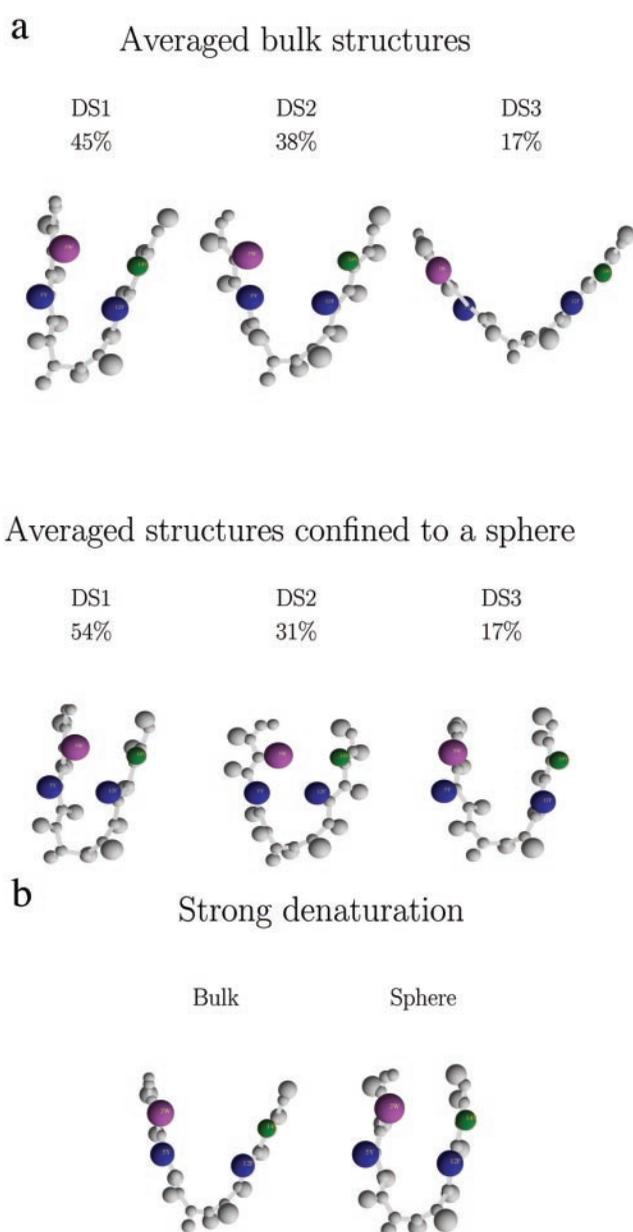
**Effect of Confinement on DSE.** Encapsulation has a profound effect on the kinetics and thermodynamics of  $\beta$ -hairpin formation. At most values of  $R_s$  the energy of the native state (and conformations belonging to the NBA) is changed only moderately. In contrast, encapsulation severely restricts the conformational space of the DSE. It follows that the enhanced stability of the  $\beta$ -hairpin at finite  $R_s$  is due to the decrease in the entropy of the DSE compared with its bulk value. To investigate changes in DSE upon confinement we analyzed the unfolded conformations for the bulk and confined hairpins by using the pattern-recognition method (see *Methods* and ref. 15). The ensemble of unfolded conformations was generated in long equilibrium simulations at  $T = 1.2T_F^b$ . We choose  $R_s = 4.85a$ , at which  $k_F \approx k_F^{\text{max}}$ . In all, for each hairpin we collected 10,000 distinct conformations.

The DSE of the bulk hairpin consists of three distinct clusters, DS1–DS3 (Fig. 3a). The bulk DS1 is compact (its radius of gyration exceeds that of the native state by 10%), and on average one-third of interstrand contacts and half of HBs are formed. The turn region is structured, and the HC is partially formed. DS2 and DS3 are unstructured and expanded ( $R_g/R_g^N$  is 1.2 and 1.4, respectively). In both DS2 and DS3 only a few interstrand contacts and HBs are formed in the turn region, while the HC is not formed. Overall, the DSE of the bulk hairpin consists of two nearly equally populated types of conformations: the structured compact conformations with significant native content and open unstructured ones, which do not contain native interactions.

The DSE of the confined hairpin can be partitioned into three distinct conformational clusters (Fig. 3a), all of which are compact ( $R_g/R_g^N$  ranges from 1.03 to 1.07). Similar to the bulk DSE, DS1 has substantial native content and is well structured. About half of native interstrand contacts and HBs are present. More significantly, the HC is formed. Although DS2 is overall less structured, we find that the HC is formed with the probability of 0.5. As in the bulk DS2 and DS3 clusters, the confined DS3 has a significantly lower fraction of native interactions, which are localized near the turn region. HC is not formed in DS3.

The DSEs for the bulk and confined hairpins have common characteristics. Both partition (in terms of their native content) into two types of clusters. One has a significant amount of native interactions and is composed of structured conformations, in which the HC is formed or partially formed (as in the bulk DS1). The other constitutes unstructured conformations with low native content. These observations are highly significant, because they suggest that elements of the native structure are present even in the denatured states of the bulk and encapsulated hairpins.

There are also *crucial differences* between both DSEs. All conformations in the confined DSE are compact, whereas in the bulk only half are compact. More importantly, confinement strongly induces the *formation of native interactions*. In the bulk DSE only 45% of all conformations have high probability for the HC formation, but in the confined DSE their fraction almost doubles and reaches 83%! The fraction of open unstructured species in which HC is not present drops from 55% in the bulk case



**Fig. 3.** (a) The distribution of DSE structures for the bulk and confined  $\beta$ -hairpins at  $1.2T_F^b$  (mild denaturation). The percentage numbers indicate the distribution of DSE conformations over the structural clusters. Bulk DS1 has significant probability (0.43) of hydrophobic cluster (HC) formation, whereas other bulk DSs are unstructured. The DSE of the encapsulated hairpin has two distinct structural clusters, DS1 and DS2, in which HC is formed. More than 80% of all confined DSE conformations belong to these clusters. The third cluster, DS3, is unstructured. Thus, confinement to a sphere induces native interactions in the DSE. A DSE structural cluster is represented as an averaged conformation computed over a given cluster. The HC side chains are color-coded and labeled. The sizes of the amino acids correspond to their actual van der Waals radii. (b) Comparison of DSE clusters for the bulk and encapsulated hairpins at  $1.67T_F^b$  (strong denaturation). The clusters shown have the largest fraction of native interactions, yet the difference between them is dramatic. Under strong denaturation the bulk DSE loses almost all native interactions, whereas the encapsulated DSE still retains significant native content (see text for details).

to a mere 17% for the confined hairpin. Thus, confinement of a hairpin to a sphere significantly alters the distribution of states in the DSE by forcing the formation of the HC. We surmise that confining a hairpin *biases the DSE toward the native state*.

To determine how the DSE changes with temperature (i.e., extent of denaturation) we generated unfolded structures for the bulk and confined hairpins at  $T = 1.6T_F^b$  (Fig. 3b). The bulk DSE falls into two clusters that differ only in their size. Strikingly, almost *no native interactions are present* in either of these, and the probability of forming HC is negligible. On the other hand, confined DSE resembles that at lower  $T = 1.2T_F^b$  and consists of three clusters. In two (72% of all DSE structures) the probability of HC formation is significant ( $\geq 0.3$ ), while the third has no traces of HC. All three DSE clusters are compact. These computations show that the nature of DSE of the encapsulated hairpin weakly depends on the degree of denaturation, whereas the bulk DSE shows very strong dependence on the temperature.

## Conclusions

Substantial reduction in the entropy of the DSE leads to enhanced stability of proteins encapsulated in inert cavities. It is remarkable that *even modest confinement* ( $R_s/R_{g,N}^b \geq 2.2$ ) leads to rate enhancement and increased stability of the native conformation. Confinement makes the open unstructured conformations inaccessible. Thus, the increase in  $k_F(R_s)$  is due to the restricted search for the native state among a subset of compact structures. Our calculations also show that for  $R_s/R_{g,N}^b \leq 1.5$  the mechanism of  $\beta$ -hairpin formation changes. A large fraction of folding  $\beta$ -hairpins becomes transiently trapped in compact misfolded conformations. The transition from such states can occur only through partial unfolding, which becomes less probable in restricted spaces.

There are several implications of this study:

1. Chan and Dill argued that the internal architecture of proteins can arise by chain compaction alone (23–25). The steric constraints that arise in the collapsed states of polypeptide chains force them to adopt elements of regular structure. The present work shows that confinement leads to compact structures (Fig. 2a). A detailed analysis of the equilibrium structures obtained even at elevated temperatures shows that the peptide is forced to adopt substantial elements of native structure (Fig. 3), which is in accord with the studies by Chan and Dill.
2. The structural characterization of the denatured states of proteins is receiving increasing attention in recent years (26). In this context, few lessons from the present work are pertinent. (i) There are only few structural clusters in the bulk DSE at temperatures slightly above  $T_F$ , which would correspond to mild denaturation. In the dominant structural cluster there is significant residual native structure. However, at more elevated temperatures, corresponding to strong denaturation, the residual native structure is present only marginally. Hence, the amount of native content in the bulk DSE depends on the degree of denaturation. (ii) The situation is dramatically different when the polypeptide chain is encapsulated in a pore. In this case, substantial native structure in the DSE is retained both at mild and strong denaturation and is determined mostly by the pore's size. Thus, confinement alone confers significant residual native structure that only *weakly depends* on the extent of denaturation.

Shortle and Ackerman (26) recently examined dipolar couplings of certain residues in  $\Delta 131\Delta$  fragment of the  $\alpha + \beta$  staphylococcal nuclease. A remarkable finding of that work is that long-ranged order persists in  $\Delta 131\Delta$  even

in 8 M urea. Their experiments also suggest that the presence of such ordering is more or less independent of the urea concentration. The experiments were performed in a polyacrylamide gels, in which  $\Delta 131\Delta$  is oriented slightly. In an inert gel the protein should be localized in the largest pores. In the bulk denatured state  $\Delta 131\Delta$  is expanded by a factor of 1.3–1.5 (26). The approximate size of  $\Delta 131\Delta$  in the native state may be estimated by using the Flory formula  $R_g \approx aN^{1/3} \approx 19 \text{ \AA}$  (for comparison  $R_g$  for the 149-mer wild-type staphylococcal nuclease is 20  $\text{\AA}$ , which is 25% larger than the measured value of 16  $\text{\AA}$ ). Although the gel confines  $\Delta 131\Delta$  to pores, some of the properties of a protein may not change due to confinement. For example, we expect small changes in the rotational relaxation time  $\tau_R \sim R_g^3$  as long as the protein is in large enough voids. Hence, encapsulation in a cavity that can accommodate  $\Delta 131\Delta$  should produce little change in  $\tau_R$  compared with the bulk (26).

In contrast to the properties that only depend on the overall chain dimensions, others can be altered dramatically by confinement. Theoretical (27, 28) and experimental (29) studies have shown that homopolymers in inert gels are collapsed. The present work shows that such confinement-induced collapse states can have substantial residual structure even at high temperatures for the pores, for which  $R_s/R_{g,N}^b \geq 2.2$ ! Thus, moderate confinement, which would not alter  $\tau_R$ , could induce long-range order in  $\Delta 131\Delta$ . If long-ranged order is purely due to the confinement, it implies that the structure would be insensitive to the extent of denaturation, i.e., the concentration of urea, which would be consistent with our finding that the structures of the DSE for the encapsulated  $\beta$ -hairpin at  $T = 1.2T_F^b$  and  $1.6T_F^b$  are similar. The agreement between the simulation and experimental findings (26) allows us to speculate that the presence of significant native interactions in  $\Delta 131\Delta$  even at strongly denaturing conditions is caused by the conformational restrictions introduced by the surrounding gel.

3. The argument that confinement stabilizes proteins is based on the assumption that the entropy of the DSE is reduced so greatly that  $\Delta F_N/\Delta F_U \ll 1$  (see the Introduction). If  $\Delta F_N \geq \Delta F_U$ , then encapsulation can destabilize the protein. If the residual entropy of the NBA is negligible, then  $\Delta F_N \geq \Delta F_U$  only if the entropy of the native interactions increases in restricted space, which would imply that the interactions that stabilize the native state (e.g., hydrophobic effect) become less favorable upon confinement. This reasoning might explain the observed destabilization of apomyoglobin in the silica matrix (8). Our arguments also lead to a *testable prediction* that if the experiments were performed in 2,2,2-trifluoroethanol solution, that is known to render stability to helical proteins, then the extent of destabilization of apomyoglobin in pores can be reversed or minimized.

We are grateful to Paragon Computation for making available the FRONTIER server that harnesses the power of distributed computing. The use of supercomputing resources of the Center for Scientific Computation and Mathematical Modeling at the University of Maryland is also gratefully acknowledged. We also thank Dr. D. Eggers for useful comments on the manuscript. This work was supported in part by National Science Foundation Grants CHE99-75150 and National Institutes of Health Grant IR01 NS41356-01.

1. Minton, A. P. (2000) *Curr. Opin. Struct. Biol.* **11**, 34–39.
2. Ellis, R. J. (2001) *Curr. Opin. Struct. Biol.* **11**, 114–119.
3. Minton, A. P. (1992) *Biophys. J.* **63**, 1090–1100.
4. Xu, Z. & Sigler, P. B. (1999) *J. Struct. Biol.* **123**, 129–141.

5. Ramakrishnan, V. & Moore, P. B. (2001) *Curr. Opin. Struct. Biol.* **11**, 144–154.
6. Zhou, H. X. & Dill, K. (2001) *Biochemistry* **40**, 11289–11293.
7. Minton, A. P. (2000) *Biophys. J.* **78**, 101–109.
8. Eggers, D. K. & Valentine, J. S. (2001) *Protein Sci.* **10**, 250–261.

9. Eggers, D. K. & Valentine, J. S. (2001) *J. Mol. Biol.* **314**, 911–922.
10. de Gennes, P. G. (1985) *Scaling Concepts in Polymer Physics* (Cornell Univ. Press, Ithaca, NY).
11. Cordeira, C., Molisana, M. & Thirumalai, D. (1997) *J. Phys. II [French]* **7**, 433–447.
12. Betancourt, M. R. & Thirumalai, D. (1999) *J. Mol. Biol.* **287**, 627–644.
13. Munoz, V., Thompson, P. A., Hofrichter, J. & Eaton, W. A. (1997) *Nature (London)* **390**, 196–199.
14. Munoz, V., Henry, E., Hofrichter, J. & Eaton, W. A. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 5872–5879.
15. Klimov, D. K. & Thirumalai, D. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 2544–2549.
16. Clementi, C., Jennings, P. & Onuchic, J. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 5871–5876.
17. Clementi, C., Nymeyer, H. & Onuchic, J. N. (2000) *J. Mol. Biol.* **298**, 937–953.
18. Klimov, D. K. & Thirumalai, D. (2001) *Protein Struct. Funct. Genet.* **43**, 465–475.
19. Grest, G. S. (1996) *J. Chem. Phys.* **105**, 5532–5541.
20. Weeks, J. D., Selinger, R. L. B. & Broughton, J. Q. (1995) *Phys. Rev. Lett.* **75**, 2694–2697.
21. Bryngelson, J. D. & Wolynes, P. G. (1989) *J. Phys. Chem.* **93**, 6902–6915.
22. Camacho, C. J. & Thirumalai, D. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6369–6372.
23. Chan, H. S. & Dill, K. A. (1989) *J. Chem. Phys.* **90**, 493–509.
24. Chan, H. S. & Dill, K. A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6388–6392.
25. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D. & Chan, H. S. (1995) *Protein Sci.* **5**, 561–602.
26. Shortle, D. & Ackerman, M. (2001) *Science* **293**, 487–489.
27. Baumgartner, A. & Muthukumar, M. (1987) *J. Phys. Chem.* **87**, 3082–3088.
28. Honeycutt, J. D. & Thirumalai, D. (1990) *J. Phys. Chem.* **93**, 6851–6858.
29. Briber, R. M., Liu, X. & Bauer, B. J. (1995) *Science* **268**, 395–397.