

Monte Carlo Simulations of Proteins in Cages: Influence of Confinement on the Stability of Intermediate States

Pedro Ojeda,[†] Martin E. Garcia,^{†*} Aurora Londoño,[‡] and Nan-Yow Chen[§]

[†]Theoretische Physik, FB 18, and Center for Interdisciplinary Nanostructure Science and Technology, Universität Kassel, Germany;

[‡]Department of Molecular Biology, Instituto Potosino de Investigación Científica y Tecnológica, San Luis Potosí, Mexico; and [§]Institute of Physics, Academic Sinica, Nankang, Taiwan

ABSTRACT We study the folding of small proteins inside confining potentials. Proteins are described using an effective potential model that contains the Ramachandran angles as degrees of freedom and does not need any a priori information about the native state. Hydrogen bonds, dipole-dipole-, and hydrophobic interactions are taken explicitly into account. An interesting feature displayed by this potential is the presence of metastable intermediates between the unfolded and native states. We consider different types of confining potentials to describe proteins folding inside cages with repulsive or attractive walls. Using the Wang-Landau algorithm, we determine the density of states and analyze in detail the thermodynamical properties of the confined proteins for different sizes of the cages. We show that confinement dramatically reduces the phase space available to the protein and that the presence of intermediate states can be controlled by varying the properties of the confining potential. Cages with strongly attractive walls destabilize the intermediate states and lead to a two-state folding into a configuration that is less stable than the native structure. However, cages with slightly attractive walls enhance the stability of native structure and induce a folding process, which occurs through intermediate configurations.

INTRODUCTION

Protein folding is one of the most intensively studied and still unsolved problems in biology. Many diseases such as Alzheimer and Parkinson are believed to be caused by the misfolding and aggregation of certain proteins (1–3). Although in the last years several aspects related to the Levinthal's paradox have been clarified with the help of lattice models and other approaches (4–7), many questions regarding details of the folding and misfolding mechanisms still remain open.

In this article, we focus on the problem of protein folding assisted by Chaperones, which is one of the mechanisms present in nature to avoid aggregation and misfolding. Chaperones are molecules in the form of a cage inside which proteins fold correctly. Recently, some progress has been achieved in the understanding of the folding of proteins inside chaperones. These studies have shown that stability and folding kinetics are strongly correlated with the geometry and the degree of confinement inside the cage (8–13). However, many details of the folding under confinement still remain uncovered.

In this work, we focus on the folding of the peptide V3-loop, Protein Data Bank ID 1NJ0, and analyze it under two kinds of time-independent confining potentials. The first potential simulates a cage being composed by rigid walls, while the second potential describes a cage with an attractive inner surface. The influence of both potentials is reflected in the thermodynamical properties, which we calculate using the Wang-Landau algorithm (14,15).

As one of the main results of this work, we obtain that the folding process of V3-loop occurs through metastable intermediate states (16), and that the presence of those states can be controlled by the confining potential.

For the description of the protein, we use a force field that does not depend on the previous knowledge of the native structure and is also able to describe folding of proteins into both helices and β -sheets with the same set of parameters (17). In addition to this improvement, two new features not reported previously are included: 1), the dipole-dipole interaction between the CO-NH pairs lying on the amide plane; and 2), the local hydrophobic interaction between neighboring residues, which takes into account the hydrophobic and hydrophilic properties of the side chains. The sequence of the amino acids is the only input of the force field.

The article is organized as follows. In Theory, we describe the model used and the Monte Carlo method applied to calculate the thermodynamical properties of the protein. In Results and Discussion, we present our results and make a careful analysis of our simulations and finally, we present a summary.

THEORY

The model

The structure of the protein is simulated using the reduced off-lattice model developed in Chen et al. (17). The amino acids are represented by means of backbones. Each backbone contains the atoms N, C $_{\alpha}$, C', O, and H. The residues are modeled as spherical beads, R , attached to the C $_{\alpha}$ values. The only remaining degrees of freedom are the

Submitted November 6, 2007, and accepted for publication September 12, 2008.

*Correspondence: magarcia@physik.uni-kassel.de

Editor: Ruth Nussinov.

© 2009 by the Biophysical Society
0006-3495/09/02/1076/7 \$2.00

doi: 10.1529/biophysj.107.125369

Ramachandran angles ψ and ϕ . The values for the bond lengths and angles are given in Solomons and Fryhle (18).

The force field containing all relevant interactions in the protein is given by

$$E_{\text{Protein}} = E_{\text{Steric}} + E_{\text{HB}} + E_{\text{DD}} + E_{\text{MJ}} + E_{\text{LocalHP}}, \quad (1)$$

where E_{Steric} represents a hard-core interparticle-potential to avoid unphysical contacts, E_{HB} accounts for the hydrogen bonding, and E_{DD} describes the dipole-dipole interactions. E_{MJ} is a distance-dependent version of the Miyazawa-Jerningan (MJ) matrix (19), which describes the interactions between residues. E_{LocalHP} accounts for local hydrophobic effects. The role of the presence of water molecules is taken into account by the terms E_{MJ} and E_{LocalHP} . It is important to point out that E_{MJ} partially includes the effect of water polarization (20). The values of the parameters of this potential are given in the original work by Chen et al. (17).

In addition to E_{Protein} we add a term to simulate the confinement of the protein within the cage. This is accomplished in the present work by using two different kinds of spherically symmetric potentials depending on a radius R_c , which is a measure of the size of the cage. In a first approach, we use an external potential $V_1(r)$ which allows the protein to fold freely for distances r smaller than R_c , but has a strongly repulsive part for larger distances, simulating the presence of the walls of the cage. The potential $V_1(r)$ reads (11)

$$V_1(r) = \frac{0.01}{R_c} \left[e^{r-R_c} (r-1) - \frac{r^2}{2} \right], \quad (2)$$

where $r = |\vec{R}|$ denotes the position of each residue.

Since $V_1(r)$ might represent a too-simple description of the confining potential of a chaperon, we also investigate the effect of an external potential $V_2(r)$ simulating attractive walls (21), which reads

$$V_2 = 4\delta_h \frac{\pi R_c}{r} \left(\frac{1}{5} \left[\left(\frac{\sigma}{r-R_c} \right)^{10} - \left(\frac{\sigma}{r+R_c} \right)^{10} \right] - \frac{\delta}{2} \left[\left(\frac{\sigma}{r-R_c} \right)^4 - \left(\frac{\sigma}{r+R_c} \right)^4 \right] \right). \quad (3)$$

The physical meaning of the different parameters in Eq. 3 can be described as follows. A uniform distribution of beads spreads out on the surface of the cage with a number density $1/\sigma^2$. The parameter ϵ is used to simulate the degree of attraction of the inner surface of the cage. A wall with a purely attractive lining has a value of $\epsilon = 1$ whereas a purely repulsive lining has a value $\epsilon = 0$. In Eq. 3, we set $\epsilon_h = 1.25$ kcal/mol and $\sigma = 3.8$ Å. The external potential $V_1(r)$ has the only effect of confining the protein inside the cage whereas the external potential $V_2(r)$ interacts with the protein by slightly reducing its energy as ϵ increases. As a consequence, the residues tend to be far apart of each other in the region close to the walls of the cage.

Simulations

Various methods based on Monte Carlo (MC) simulations have been proposed to compute the thermodynamical properties of finite systems. They include, for instance, multicanonical simulations (22) and simulated annealing (23). In this work we use the Wang-Landau algorithm (14), also including a recent improvement introduced by Pereyra et al. (15). One of the main advantages of Wang-Landau simulations is that they allow us to obtain directly the density of states (DOS) of the system, which is, of course, independent of the simulation temperature. Once the DOS is known, one can obtain all the thermodynamical properties of the system at any temperature. Within this framework, the transition probability between two conformations before and after a MC trial move, \mathbf{X}_1 and \mathbf{X}_2 respectively, is calculated as

$$P(\mathbf{X}_1 \rightarrow \mathbf{X}_2) = \min \left[1, \frac{g(\mathbf{X}_1)}{g(\mathbf{X}_2)} \right], \quad (4)$$

where $g(\mathbf{X})$ is the DOS of the system and \mathbf{X} is a generalized coordinate, which in our case is represented by a vector with two entries $\mathbf{X} = (E, Q)$, being E the configurational energy and Q the end-to-end distance of the protein structure. Note that Q can be interpreted as an order parameter for the folding (unfolding) transition.

The original scheme developed by Wang and Landau (14) can be briefly described as follows: one sets the initial function $g(\mathbf{X})$ together with an auxiliary histogram $H(\mathbf{X})$ to be equal to 1. Then, each time the bin \mathbf{X} is visited, one updates the histogram $H(\mathbf{X})$ and modifies $g(\mathbf{X})$ as $g(\mathbf{X}) \rightarrow g(\mathbf{X}) \times f$, with $f = e = 2.718281\dots$. This procedure is continued until a flat histogram (with a certain significance, i.e., 80%) is obtained. At this step, the histogram $H(\mathbf{X})$ is reset and the factor f is reduced. The usual way to perform this reduction is by taking $f_{i+1} = \sqrt{f_i}$. Convergence is achieved when a value for f_{i+1} close enough to 1 is obtained. The last step must be compatible with the desired accuracy, for example $f = \exp(10^{-7})$.

We use the modified Wang-Landau approach proposed in Berlardinelli and Pereyra (15), which has been shown to speed up simulations and to partially avoid the problem of saturation error. According to the new scheme, one does not need to wait until the histogram $H(\mathbf{X})$ is flat, but it is enough to require that all the entries of $H(\mathbf{X})$ are visited. Then $H(\mathbf{X}) = 0$ is reset, and $f_{i+1} = \sqrt{f_i}$ is updated.

Following Berlardinelli and Pereyra (15), we employ a second histogram $H_2(\mathbf{X})$ which is never reset during the whole simulation and define the Monte Carlo time-step as $t = j/N$, with N being the number of points in the energy axis and j the number of trial moves performed. If $f_{i+1} \leq t^{-1}$ then $f_{i+1} = f(t) = t^{-1}$, and from this point on, $f(t)$ is updated at each Monte Carlo time step. $H(\mathbf{X})$ is not used during the rest of the calculation. Convergence is achieved when $f(t) < f_{\text{final}}$. In the present simulations we used $f_{\text{final}} = \exp(10^{-7})$. Finally, the thermodynamical properties of the

system such as the free energy $F(T)$, internal energy $U(T)$, entropy $S(T)$, and specific heat $C(T)$ can be calculated from $g(\mathbf{X})$. For the specific case treated in this work, where $\mathbf{X} = (E, Q)$, the corresponding definitions read as

$$F(T) = -k_B T \ln \left(\int dE \int dQ g(E, Q) e^{-\beta E} \right), \quad (5)$$

$$U(T) = \langle E \rangle_T = \frac{\int dE \int dQ E g(E, Q) e^{-\beta E}}{\int dE \int dQ g(E, Q) e^{-\beta E}}, \quad (6)$$

$$S(T) = \frac{U(T) - F(T)}{T}, \quad (7)$$

$$C(T) = \frac{\langle U^2 \rangle_T - \langle U \rangle_T^2}{k_B T^2}, \quad (8)$$

where $\beta = 1/k_B T$ and k_B is the Boltzmann constant. The free energy landscape as a function of E and Q can be computed as

$$F(E, Q) = -k_B T \ln \left(\frac{g(E, Q) e^{-\beta E}}{\int dE \int dQ g(E, Q) e^{-\beta E}} \right), \quad (9)$$

and

$$F(E) = -k_B T \ln \left(\frac{\int dQ g(E, Q) e^{-\beta E}}{\int dE \int dQ g(E, Q) e^{-\beta E}} \right). \quad (10)$$

RESULTS AND DISCUSSION

We focused our attention on a peptide composed of 16 amino acids with PDB code 1NJ0 to study the folding mediated by confining potentials. This peptide conforms to the V3-loop of the exterior membrane glycoprotein of the Human Immunodeficiency Virus type 1.

To explore the relevant part of the phase space of the protein we have chosen an energy window between -132.0 kcal/mol and -30 kcal/mol and the end-to-end distance Q ranging from 5 \AA to 50 \AA . This region is enough to cover both the highly ordered structures (present at $T \sim 0$) and the fully disordered random coils (stable for $T \sim \infty$). The MC search was generated by changing each pair of Ramachandran angles ψ_i and ϕ_i at each MC step using cutoffs with values $|\Delta\psi_c| \leq 40^\circ$ and $|\Delta\phi_c| \leq 40^\circ$. To reach $f_{\text{final}} = \exp(10^{-7})$, 8×10^9 trial moves were necessary.

We first analyze the properties of the peptide without confinement (bulk case). The obtained ground-state structure of the V3-loop is depicted in Fig. 1. It consists of a β -sheet structure with energy ~ -132.0 Kcal/mol and an end-to-end distance of $\sim 5.5 \text{ \AA}$.

A new feature described by our force field is the presence of intermediate structures between the native (N) and the unfolded (U) states as shown in Fig. 2 (16). We obtain two intermediate states in the free energy profile $F(E)$ at the transition temperature. The intermediates, denoted as I_1 and I_2 in

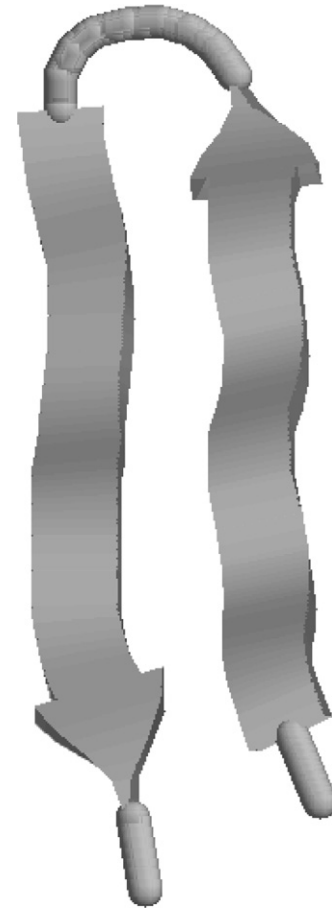


FIGURE 1 Ground-state structure (β -sheet) of the peptide 1NJ0 ($E_g \sim -132.0$).

the figure, appear as local minima of $F(E)$ versus E . To analyze the nature of the intermediate states we have split the energetic and entropic parts of the free energy. Results indicate

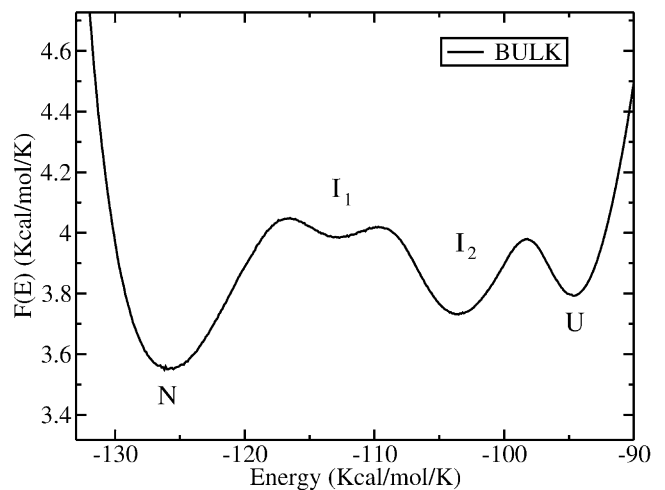


FIGURE 2 Free energy as a function of the configurational energy E for the bulk ($R_c \rightarrow \infty$) showing the presence of the native (N), the intermediate (I_1, I_2), and the unfolded (U) states.

that the intermediates are mainly stabilized by their energy, but that there is a nonnegligible entropic contribution.

The next problem to be addressed is the influence of confinement on the free energy landscapes and folding behavior of the V3-loop. For this purpose, we calculated first the DOS and the specific heat of the protein assuming the rigid-wall confining potential $V_1(r)$ described above. We considered different cage-diameters ($R_c = 15 \text{ \AA}$, 20 \AA , and 25 \AA). In Fig. 3 *a*, we show influence of $V_1(r)$ on the behavior of the DOS. Note that, due to confinement, $\log [g(E)]$ considerably decreases at high energies compared to the bulk case ($R_c \rightarrow \infty$). For energies close to the ground state, $\log [g(E)]$ does not exhibit any noticeable change because the protein is almost folded. Since its gyration radius in the ground state is $R_g \sim 13 \text{ \AA}$, barriers of radii equal or larger than 15 \AA do not affect folded structures. This result is consistent with the intuitive picture that cages, for instance chaperones, restrict the otherwise huge phase space for high energies, making the number of available structures, and consequently the entropy, considerably smaller than in absence of a cage. The effect of confinement can be also observed in the specific heat of the V3-loop, which we show in Fig. 3 *b*. Here, we plot the specific heat for different values of the cage radius (15 \AA , 20 \AA , 25 \AA) and for the bulk case ($R_c \rightarrow \infty$) as a function of T/T_f^0 , where $T_f^0 = 321 \text{ K}$ is the transition (unfolding) temperature in absence of a cage. The effect of the rigid-wall potential $V_1(r)$ is to increase the transition temperature (see Table 1) and to make the curve of the specific heat broader as the radius of the cage decreases. A broader curve means that there are more structures with energies close to the native state than in the bulk-case, where only the native state

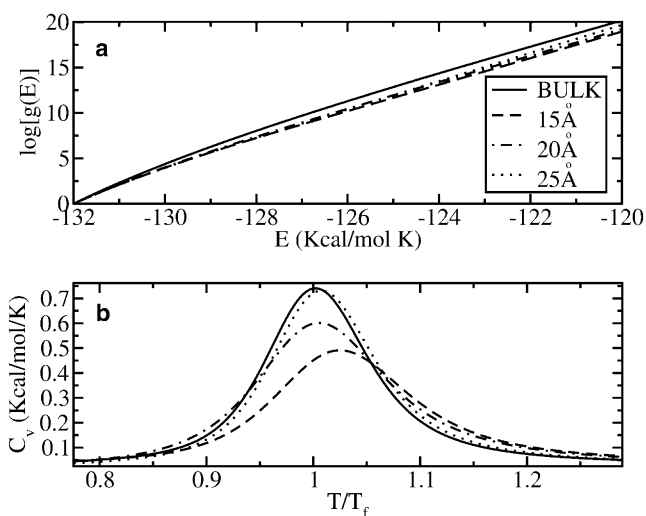


FIGURE 3 (a) Logarithm of the density of states (DOS) $g(E)$ of the protein inside the confining potential $V_1(r)$ and for different values of R_c (15 \AA , 20 \AA , 25 \AA) as well as for the bulk case. One notices the remarkable decrease of the DOS for decreasing R_c . (b) Specific heat for the bulk case and for confining potentials with radii 15 \AA , 20 \AA , and 25 \AA . $T_f = 321 \text{ K}$ is the transition temperature in the bulk case. T_f increases as the radius R_c decreases. The confining potential in panels *a* and *b* is purely repulsive.

TABLE 1 Transition temperatures T_f for different values of the radius R_c of the potential $V_1(r)$ (see main text)

| Temperature (K) | Radius (\AA) |
|-----------------|-------------------------|
| 329.2 | 15 |
| 323.4 | 20 |
| 323.2 | 25 |
| 321.0 | ∞ |

Note that T_f decreases for increasing R_c .

is the most important compact structure. For radii larger than 25 \AA , the transition temperatures are equal to T_f^0 within the statistical error of our simulations. We conclude that the protein is more stable as the radius of the cage decreases. This results are in agreement with Rathore et al. (11), in which Monte Carlo simulations were used, and with Takagi et al. (9) and Lu et al. (21), where Langevin simulations were performed. It is important to mention, however, that in those cited simulations a simplified Gō-type force field was used. Thirumalai et al. (24) made a considerable improvement to the force field by introducing the effect of the nonnative interactions. However, important interactions such as dipole-dipole and hydrogen bonds were not taken into account. The presence of intermediates was not reported either in those studies.

The main goal of this article is the study of the influence of confinement on the potential landscape and, consequently, on the stability of the native and intermediate states. Note that the intermediates can be better characterized by analyzing the free energy F as a function of both the energy E and the order parameter Q . In Fig. 4 *a*, we show the contour plot of $F(E, Q)$ for the bulk case. Clearly, the two intermediate structures, which we denote as I_1 and I_2 , can be identified as local minima of $F(E, Q)$. It is important to point out that the values of the end-to-end distance Q in the intermediates I_1 and I_2 are larger than in the native structure, but smaller than in the unfolded state. Fig. 3 *a* shows the importance of choosing the adequate order parameters to plot the free energy.

To study the effect of a cage with purely repulsive walls (potential $V_1(r)$) we have determined $F(E, Q)$ for different values of the cage-radius R_c . In Fig. 4, *b–d*, we show the corresponding contour plots for $R_c = 25 \text{ \AA}$, $R_c = 20 \text{ \AA}$, and $R_c = 15 \text{ \AA}$, respectively. The different minima of $F(E, Q)$ are shown together with representative snapshots of the corresponding structures. The main effect of the cage of $R_c = 25 \text{ \AA}$ is to restrict the size of the unfolded states (U), which is reflected in a shift of the local U -minimum to a smaller value of Q (see Fig. 4 *b*). Further decrease of R_c leads to a stronger reduction of the size of the unfolded states. For example, for $R_c = 15 \text{ \AA}$ the local U -minimum is situated at $Q \sim 20 \text{ \AA}$, i.e., which means that the end-to-end distance of the unfolded states has been halved in value with respect to the bulk case. This can also be observed by the change in the form of the unfolded structures shown in the figure. Interestingly, there is also a small shift of the U -minimum to lower

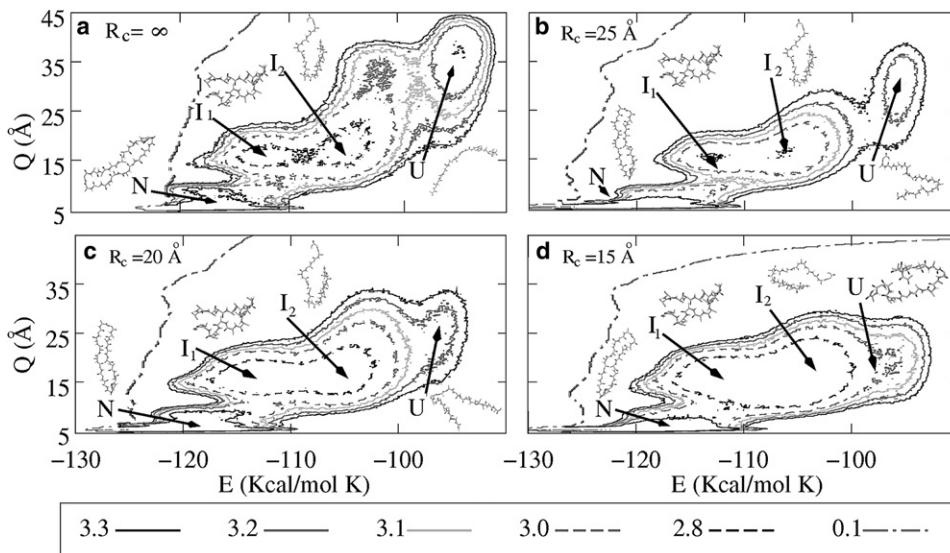


FIGURE 4 Contour plots of the free energy landscape $F(E, Q)$ as a function of the configurational energy E and the end-to-end distance Q for a purely repulsive confining potential. Plots *a–d* correspond to the bulk case and cages of radius 15 Å, 20 Å, and 25 Å, respectively. The unfolded states are strongly affected when the size of the cage decreases. The native state and the intermediates are only slightly modified. The contour lines represent the free energy difference with respect to the native state and are given in Kcal/mol K.

energies for decreasing R_c . This is simply because strong spatial confinement necessarily leads to the formation of contacts which were not present in the bulk. The shift of the U -minimum to lower energies also explains the reduction of the DOS upon confinement shown in Fig. 3 *a*.

In contrast to the unfolded states, the native structure is practically not affected by confinement, at least up to $R_c = 15$ Å, and the position of the N -minimum remains almost unchanged (see Fig. 4).

Different and interesting features are observed in the behavior of the intermediate states I_1 and I_2 upon repulsive confinement. In the bulk, the minima corresponding to the intermediates are well defined and separated by a potential barrier. Although both the position of the minima and the structure of the intermediates remain unchanged when the radius of the cage is reduced, the depth of the energy minima and the potential barrier between them decrease. Already for $R_c = 20$ Å, both minima start to merge and form an extended and shallow minimum. This effect is even stronger for $R_c = 15$ Å. Note that this happens when the radius of the confining potential becomes comparable to the end-to-end distance of the intermediate states in the bulk.

Now, we report on the influence of hydrophobic effects in the inner surface of the cage. Attractive cage-walls were considered by using the confining potential $V_2(r)$ (Eq. 3) with radius $R_c = 30$ Å. The degree of attraction is described by the coefficient ϵ . A completely attractive cage-wall is obtained when $\epsilon = 1.0$, whereas $\epsilon = 0.0$ corresponds to a completely repulsive or neutral inner surface of the cage. The effect of ϵ can be visualized in the following way: as ϵ increases from 0 to 1, the walls of the cage tend to attract the residues because of the relative minimum generated by the potential $V_2(r)$. The deepest minimum of $V_2(r, \epsilon)$ is reached when $\epsilon = 1.0$ and corresponds to $V_2^{\min} \sim 5$ Kcal/mol. This energy is comparable to the energy required to break one hydrogen bond, $\Delta E_{HB} \sim 4.8$ Kcal/mol. Therefore, for $\epsilon \sim 1.0$ the poten-

tial is able to destroy the structure of the protein (denaturation).

The influence of the confining potential $V_2(r)$ on the DOS of the protein is shown in Fig. 5 *a*, where different degrees of attraction and the bulk case are considered. Two clear features can be distinguished. First, the DOS at energies close to the native state increases for increasing ϵ . This means that the potential landscape is changed near the global minimum. Furthermore, for large energies one can clearly observe a dramatic reduction of $g(E)$ by up to ~ 13 orders of magnitude as ϵ goes from 0 to 1. However, this remarkable

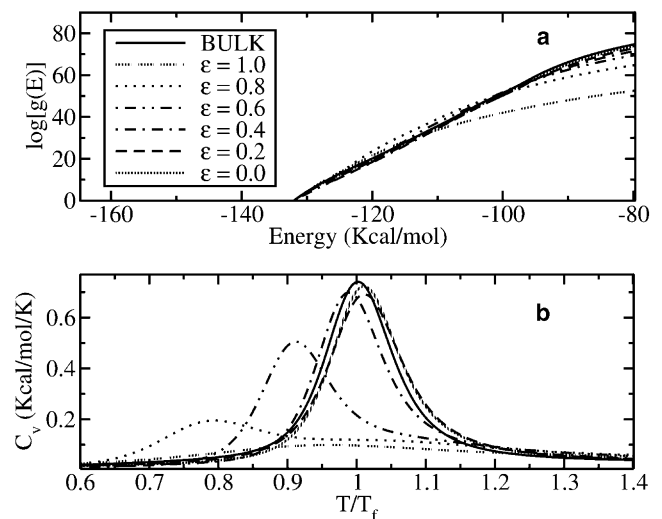


FIGURE 5 (*a*) Logarithm of the DOS $g(E)$ for different degrees of hydrophobicity ($\epsilon = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$) and for the bulk case. Notice the abrupt decay of $g(E)$ by ~ 13 orders of magnitude as ϵ goes from 0.0 to 1.0. For high values of ϵ , the protein tends to be in the unfolded state. (*b*) Specific heat of the protein for different values of ϵ , 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0, compared to the bulk case. $T_f = 321$ K is the transition temperature for the bulk. Notice how T_f and the peak of the specific heat decrease as ϵ goes from 0 (purely repulsive wall) to 1 (strongly attractive wall).

TABLE 2 Transition temperatures T_f for the confining potential $V_2(r)$ (see main text) for different degrees of hydrophobicity, $\epsilon = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$, and for the bulk case

| Temperature (K) | Radius (\AA) |
|-----------------|-------------------------|
| 321.0 | Bulk |
| 324.2 | $\epsilon = 0.0$ |
| 324.1 | $\epsilon = 0.2$ |
| 314.5 | $\epsilon = 0.4$ |
| 292.1 | $\epsilon = 0.6$ |
| 253.5 | $\epsilon = 0.8$ |
| — | $\epsilon = 1.0$ |

Notice that in general T_f decreases as ϵ increases. For $\epsilon = 1.0$ it is not possible to define T_f because the specific heat is almost completely attenuated.

reduction of the phase space in this case does not help the protein to fold correctly but forces it to acquire a denatured conformation. This effect occurs because the peptide decreases its energy by placing some of the residues close to the border of the cage. Then, the number of accessible states at those energies decreases and residues are no longer allowed to be far apart from the border, since it would cost much energy. As a consequence, the peptide sticks to the wall of the cage.

The influence of the potential $V_2(r)$ on the specific heat $C(T)$ is shown in Fig. 5 b. As ϵ increases, the curve $C(T)$ becomes broader. The transition temperatures for different values of ϵ and for the bulk case are presented in Table 2. Interestingly, for $\epsilon = 0-0.4$ we obtain an increase of the transition temperature compared to the bulk case. The range $0.3 \leq \epsilon \leq 0.4$ seems to be the optimal one regarding stability. For that range of ϵ , the protein is more stable than in the absence of a cage. For higher values of ϵ , the transition temperatures become lower. For $\epsilon = 1.0$, the curve of the specific heat is extremely broad and attenuated, reflecting the fact that the protein is almost denatured.

One of the main results of this article is illustrated in Fig. 6, where we show the contour plots the free energy

$F(E, Q)$ for different values of ϵ . The radius of the cage is equal to 30 \AA in all cases.

For $\epsilon = 0$ (Fig. 6 a), the presence of the native state (N), the intermediates (I_1 and I_2), and the unfolded states (U) can be clearly observed. The increase of ϵ leads to an effective increase of the confinement, and also to the presence of shallower minima for the intermediate states. For $\epsilon = 0.4$ (Fig. 6 b), the global minimum and the three local minima can still be distinguished. However, the intermediates become almost unstable and the energy landscape has practically only two well-defined minima. A further increase of ϵ completely changes the energy landscape. For $\epsilon = 0.6$ (Fig. 6 c), the native state and the intermediates I_1 and I_2 are no longer present. Instead, a new intermediate state N' appears, which has more native contacts than I_1 and I_2 , but less than N , and exhibits a much lower value of energy E . By looking at the structure corresponding to the minimum denoted by N' it is clear that is very similar to N , but not completely folded. We interpret that the N' state is a slight deformation of the native state produced by the presence of attractive wall. Note that also the U -minimum is shifted to much lower energies. This clearly indicates that the protein sticks to the wall of the cage. The net effect is that the potential landscape shows for this value of ϵ a two-state situation. Finally, for $\epsilon = 0.8$ (Fig. 6 d), only the unfolded states are present.

Jewett et al. (13) showed that, for a particular protein, one metastable state might exist in the presence of a weakly hydrophobic barrier. In this work and for the peptide V3-loop we obtain a different result, namely, that the protein shows a folding behavior through intermediates in the bulk, but an attractive barrier with an optimum degree of hydrophobicity can lead to weakening the intermediate states and inducing a quasi two-state folding process.

Summarizing, we have studied the folding of the peptide 1NJ0 under different kinds of confining potentials. We

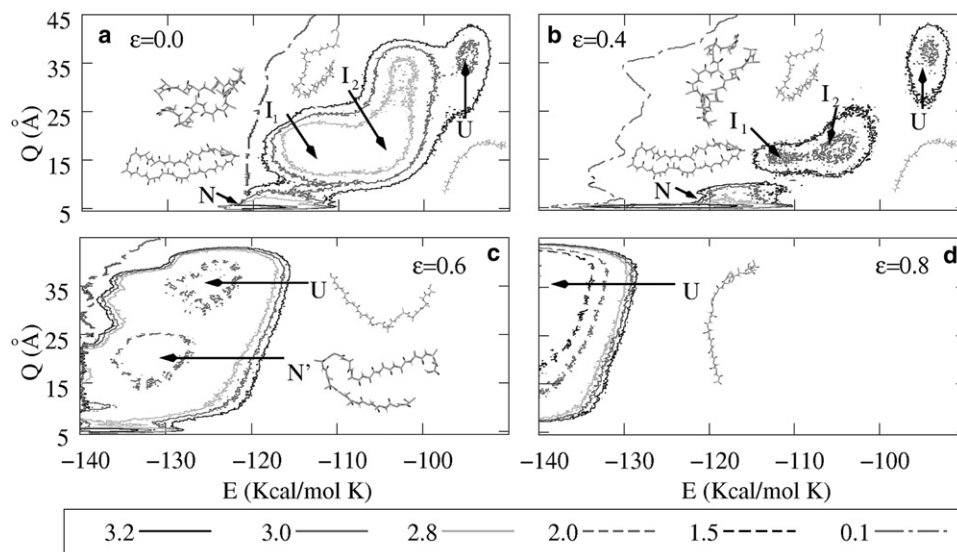


FIGURE 6 Contour plots of the free energy landscape $F(E, Q)$ for a cage with attractive inner surface. Different degrees of hydrophobicity are displayed in plots a–d, corresponding to $\epsilon = 0.0$, $\epsilon = 0.4$, $\epsilon = 0.6$, and $\epsilon = 0.8$, respectively. The native and the intermediates states are slightly modified for $0.0 < \epsilon < 0.4$ but for larger values of ϵ the intermediate states disappear and the native structure is deformed. As a consequence, $F(E, Q)$ represents a two-states landscape. The contour lines represent the free energy difference with respect to the native state and are given in Kcal/mol K.

used a force field which is independent on the native structure and includes relevant interactions such as the dipole-dipole and hydrogen bonds. We demonstrated the presence of intermediate states not reported before. These intermediates are strongly affected by the confining potentials.

P.O. thanks the Deutscher Akademischer Austausch Dienst for the financial support for his PhD.

REFERENCES

- Etienne, M., J. Aucoin, Y. Fu, R. McCarley, and R. Hammer. 2006. Stoichiometric inhibition of amyloid-protein aggregation with peptides containing alternating a,a, disubstituted amino acids. *J. Am. Chem. Soc.* 128:3522–3523.
- Kelly, J. 1998. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* 8:101–106.
- Lynn, D., and S. Meredith. 2000. Review: Model peptides and the physicochemical approach to β -amyloids. *J. Struct. Biol.* 130:153–173.
- Skolnick, J., and A. Kolinski. 1990. Simulations of the folding of a globular protein. *Science*. 250:1121–1125.
- Abkevich, V., A. Gutin, and E. Shakhnovich. 1994. Specific nucleus as the transition state for protein folding: evidence from the lattice model. *Biochemistry*. 33:10026–10036.
- Duan, Y., L. Wang, and P. Kollman. 1998. The early stage of folding of villin headpiece subdomain observed in a 200-nanosecond fully solvated molecular dynamics simulation. *Proc. Natl. Acad. Sci. USA*. 95:9897–9902.
- Guo, Z., and D. Thirumalai. 1998. Kinetics and thermodynamics of folding of ADE novodesigned four-helix bundle protein. *Proc. Natl. Acad. Sci. USA*. 263:9897–9902.
- Fang, F., and L. Szleifer. 2006. Controlled release of proteins from polymer-modified surfaces. *Proc. Natl. Acad. Sci. USA*. 103:5769–5774.
- Takagi, F., N. Koga, and S. Takada. 2003. How protein thermodynamics and folding mechanisms are altered by the chaperonin cage: molecular simulations review. *Proc. Natl. Acad. Sci. USA*. 100:11367–11372.
- Thirumalai, D., D. Klimov, and G. Lorimer. 2003. Caging helps proteins fold. *Proc. Natl. Acad. Sci. USA*. 100:11195–11197.
- Rathore, N., T. Knotts, and J. Pablo. 2005. Confinement effects on the thermodynamics of protein folding: Monte Carlo simulations. *Biophys. J.* 90:1767–1773.
- Netto, A., C. Silva, and A. Caparica. 2006. Wang-Landau sampling in three-dimensional polymers. *Braz. J. Phys.* 36:619–622.
- Jewett, A., A. Baumketner, and J. Shea. 2004. Accelerated folding in the weak hydrophobic environment of a chaperonin cavity: creation of an alternate fast folding pathway. *Proc. Natl. Acad. Sci. USA*. 101:13192–13197.
- Wang, F., and D. Landau. 2001. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.* 86:2050–2053.
- Belardinelli, R. E., and V. D. Pereyra. 2007. Fast algorithm to calculate density of states. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 75:046701–046705.
- Schnabel, S., M. Bachmann, and W. Janke. 2007. Two-state folding, folding through intermediates, and metastability in a minimalistic hydrophobic-polar model for proteins. *Phys. Rev. Lett.* 98:048103.
- Chen, N. -Y., Z. -Y. Su, and C. -Y. Mou. 2006. Effective potentials for folding proteins. *Phys. Rev. Lett.* 96:078103–078107.
- Solomons, G., and C. Fryhle. 2000. *Organic Chemistry*, 7th Ed. John Wiley & Sons, Oxford.
- Miyazawa, S., and R. Jerningan. 1996. Residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J. Mol. Biol.* 256:623–644.
- Wang, Z., and H. Lee. 2000. Origin of the native driving force for protein folding. *Phys. Rev. Lett.* 84:574–577.
- Lu, D., Z. Lu, and J. Wu. 2006. Structural transitions of confined model proteins: molecular dynamics simulation and experimental validation. *Biophys. J.* 90:3224–3238.
- Berg, B., and T. Neuhaus. 1991. Multicanonical algorithms for first order phase transitions. *Phys. Lett. B.* 267:249–253.
- Kirkpatrick, S., C. Gelat, and M. Vecchi. 1983. Optimization by simulated annealing. *Science*. 220:671–680.
- Thirumalai, D., D. Klimov, and G. Lorimer. 2006. Nanopore-protein interactions dramatically alter stability and yield of the native state in restricted spaces. *J. Mol. Biol.* 357:632–643.