Computational Studies of the Reversible Domain Swapping of p13suc1

Jorge Chahine* and Margaret S. Cheung†
*Departamento de Física, UNESP-Universidade Estadual Paulista, São José do Rio Preto-SP 15054-000, Brazil; and †Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

ABSTRACT A minimalist representation of protein structures using a Go-like potential for interactions is implemented to investigate the mechanisms of the domain swapping of p13suc1, a protein that exists in two native conformations: a monomer and a domain-swapped dimer formed by the exchange of a β-strand. Inspired by experimental studies which showed a similarity of the transition states for folding of the monomer and the dimer, in this study we justify this similarity in molecular descriptions. When intermediates are populated in the simulations, formation of a domain-swapped dimer initiates from the ensemble of unfolded monomers, given by the fact that the dimer formation occurs at the folding/unfolding temperature of the monomer (T_f). It is also shown that transitions, leading to a dimer, involve the presence of two intermediates, one of them has a dimeric form and the other is monomeric; the latter is much more populated than the former. However, at temperatures lower than T_f, the population of intermediates decreases. It is argued that the two folded forms may coexist in absence of intermediates at a temperature much lower than T_f. Computational simulations enable us to find a mechanism, "lock-and-dock", for domain swapping of p13suc1. To explore the route toward dimer formation, the folding of unstructured monomers must be retarded by first locking one of the free ends of each chain. Then, the other free termini could follow and dock at particular regions, where most intrachain contacts are formed, and thus define the transition states of the dimer. The simulations also showed that a decrease in the maximum distance between monomers increased their stability, which is explained based on confinement arguments. Although the simulations are based on models extracted from the native structure of the monomer and the dimer of p13suc1, the mechanism of the domain-swapping process could be general, not only for p13suc1.

INTRODUCTION

Domain swapping is the exchange of identical structural elements of different proteins leading to the formation of dimers or an assembly of oligomers. The extent of the exchanged portion can be a few linked residues or an element of secondary structure or a portion of the tertiary structure of the polypeptide chain. The concept of domain swapping and its possible biological function was first introduced by Eisenberg and co-workers (1,2). Since its first observation in crystals (3) 10 years ago, there has been a lot of discussion regarding its role, which includes regulating function and a mechanism to form amyloids, aggregates, and misfolded structures. It has also been proposed as part of an evolutionary process to form protein oligomers. Amyloids and protein oligomers have been associated with pathologies such as neurological diseases (4,5). A conclusive discussion regarding the biological role of domain swapping is still missing due to the lack of evidence of domain swapping in vivo as it has only been observed in vitro. Such observations include proteins as chymotrypsin inhibitor 2 or CI2 (6), SH3 (7), and many others. An extreme example of the domain swapping is the native structure of T4 endonuclease VII formed by the exchange of almost the entire chain to form the dimer (8).

This study focuses on the domain swapping of p13suc1. A cell cycle regulatory protein such as p13 was chosen not only because of its biological importance, but also because there have been interesting mutagenetic experimental studies to understand the mechanism of its dimer formation, whose detailed molecular mechanisms could be possibly investigated by computational simulations. The dimer of p13suc1 is formed by the exchange of a β-strand. In the domain-swapped structure this β-strand defines the ''hinge loops'', which are the residues connecting the exchanged structures with the rest of the protein. Its length is $\sim$6 residues where two prolines are believed to alter the stability of the conformations. A scheme has been proposed (9) by which the prolines in the hinge loop could play an important role in the process. By making the loop more rigid, the monomeric form would be under strain which could be released in the dimer conformation. And once the dimer is formed, another portion of the hinge loop (a second proline) would be under strain and thus produce reversible domain swapping.

In this work, this reversible transition is studied by molecular dynamics (MD) and the use of a simplified model for the dimer represented by its main chain C_α atoms (i.e., positions are determined in its native structure). The motivation for using models described by parameters taken from the native structure is related to the fact that proteins are sufficiently energetically minimally frustrated (10,11) that the geometrical heterogeneity observed in the transition state ensemble is mostly determined by topological effects. This observation has been confirmed for the folding events of small proteins (12,13) and even in simple dimers (14). All atomistic simulations have been performed (15) to study the unfolding pathway of p13suc1. In this study, the unfolding transition states were described by the amount of secondary

Submitted April 19, 2005, and accepted for publication July 15, 2005.
Address reprint requests to Jorge Chahine, E-mail: chahine@ibilce.unesp.br.

0006-3495/05/10/2693/08 $2.00
doi: 10.1529/biophysj.105.062679

and tertiary structures. Dimerization through domain swapping is revealed in structural details that follow a ''lock-and-dock'' mechanism.

METHODS

Models

A minimalist representation of protein structures where only C_α atoms are taken from each residue is implemented in this study. We use Suc13 in both monomer (Protein Data Bank ID code 1SCE) and domain-swapped dimer (Protein Data Bank ID code 1PUC) as the protein models to study dimerization.

The construction of the Hamiltonian is inspired by the work of Itzhaki and co-workers (16) and ϕ-value analysis (9) in which the monomer form showed more ''strain'' than the dimer (that is to be defined as one of the two prolines in the hinge loop is mutated to alanine); mutation in this residue to alanine (P90A) completely removes the ability of the protein to dimerize. In this study, the idea of such strain and the role of prolines in the hinge loop of a monomer is reflected when the structural Hamiltonian is designed to favor a dimer configuration.

We modified a standard Go-like potential (17) in which only contacts defined in the native conformation are attractive (12) in a monomer to accommodate dimer formation. The potential in Eq. 1 includes the local terms (i.e., bond, dihedral, and angles) to favor the dimer interactions, as well as allowing nonlocal terms to form contacts within the monomer itself and/or between a dimer. There are a total of 565 native contacts among residues. Each monomer has 214 contacts, and there are an additional 137 contacts among residues of different chains. Because local terms of monomers retain dimer information, although nonlocal interactions of a monomer follow a minimally frustrated contact formation, folding of monomers is under strain and frustrated by competitions to dimer formation. Because both monomer and dimer structures of p13suc1 are solved by the x-ray crystallography method, native contacts between two domain-swapped monomers are justified pragmatically in our study to reflect frustrations in dimer formation. In this aspect, the philosophy to design the contact Hamiltonian is different from a recent ''symmetrized Go-type'' potential (18), in which intrachain contacts between two domain-swapped monomers are theoretically modeled from interchain interactions in a single, folded monomer.

$$E = \sum_{\text{bonds}} K_r(r - r_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_d[1 + \cos(n(\Phi - \Phi_0))] + \sum_{|i-j|<4}^{\text{monomer}} \left\{ 5\varepsilon_{ij}\left(\frac{R_{ij}^o}{r_{ij}}\right)^{12} - 6\varepsilon_{ij}\left(\frac{R_{ij}^o}{r_{ij}}\right)^{10} \right\} + \sum_{i,j}^{\text{dimer}} \left\{ 5\varepsilon_{ij}\left(\frac{R_{ij}^o}{r_{ij}}\right)^{12} - 6\varepsilon_{ij}\left(\frac{R_{ij}^o}{r_{ij}}\right)^{10} \right\} + \sum_{i,j} \varepsilon_2\left(\frac{R_{ij}^{\text{excl}}}{r_{ij}}\right)^{12} \quad (1)$$

r_o, θ_o, and Φ_o are the bond, angle, and dihedral angles defined in the native configuration of a dimer. Nonbonded interactions follow a 10–12 Lennard-Jones potential that favors both dimer and monomer. This settlement to suit both configurations gives rise to fluctuations of either monomer or dimer formation. ε_{ij} is set to 0.5 ε and R_{ij}^o are bonding distances of native pairs in either monomer or dimer. $\varepsilon_2(R_{ij}^{excl})^{12}$ for nonnative pairs is a constant (0.002 ε) to give excluded volume effects. The mass of each bead is 100 a u, K_r is 100 ε, and K_θ is 20 ε. The dihedral term has the periodicity of $n = 2$: that with $K_d = 0.3$ ε and the other with 0.6 ε. In this work, similar to other studies (12,13) we choose $\varepsilon = 1.0$ kcal/mol. Temperatures are presented in units of T_f, which is the folding temperature of an isolated monomer, and energies are scaled in units of kT_f instead of ε.

The center of mass of two monomers is constrained by a repulsive potential set at a distance of 45 Å, which corresponds to a concentration of 0.25 M (molar). Noticeably, this is a very concentrated condition if done in experiments. Nevertheless, it is still far from the distance between the center of mass of two monomers in the native dimer (27.5 Å).

Molecular Dynamics

MD simulations are performed in the range $0.9\,T_f < T < 1.2\,T_f$ where eight sets of temperatures were chosen to run simulations by the method of replica exchange (19). This technique is based on the exchange of configurations (replicas), each at a given temperature, which is accepted or not according to a Metropolis criterion. The range of temperature is broad enough to sample the configurational space of the unfolded monomers (high temperature) and the folded dimer (low temperature). Quantities such as order parameters and specific heat as a function of temperature, as well as free energies as a function of order parameters, are calculated by the WHAM algorithm (20). By using reweighing techniques, all the thermodynamics quantities can be calculated as a function of the temperature.

Order parameters

We introduce several order parameters with various convenient definitions to label all the different phases. The order parameter $Q1$ ($Q2$) is the number of native contacts of monomer 1 (2). $Q12$ is the number of native contacts between different chains. When the letter n is used with these order parameters (e.g., $Q1n$, $Q2n$, and $Q12n$), it means that order parameters have been normalized against their maximum values. $Q_{total} = Q1 + Q2 + Q12$, is the total number of native contacts in Fig. 1. In addition, we found it convenient to introduce Q_{hybrid}, which gives insights on how two monomers interact at various circumstances. $Q_{hybrid} = Q1 + Q2$ when $Q12$ is negligible ($Q12 < 30\%$ of its maximum), and $Q_{hybrid} = Q12$ when $Q12$ is noticeable ($Q12 > 40\%$ of its maximum).

Transition states analysis

We calculated a quantity ϕ_{ij} defined by

$$\phi_{ij} = \frac{f_{ij}^{TS} - f_{ij}^{U}}{f_{ij}^{F} - f_{ij}^{U}} \quad (2)$$

where f_{ij} are the frequencies (probabilities) of contacts between residues i and j in the three ensembles: transition state (TS), folded (F), and unfolded (U). The transition state (TS) portions of the phase diagram for monomers and dimers are delimitated by two parallel arrows (see Fig. 4 *b*) and labeled by MTS and DTS.

We choose a cutoff of Φ-values >0.5 to show whether transition states are correlated or not. The qualitative description does not change if other slightly higher values of cutoff are used; the number of colored points merely decreases because the cutoff value increases (data not shown).

RESULTS AND DISCUSSIONS

Description of monomer and dimer formation

Stability of monomers and dimers

The specific heat calculated from thermodynamics simulations of dimers is plotted against temperature in Fig. 1 *a* and shows two distinctive peaks at 1.02 T_f and at 1.1 T_f. The

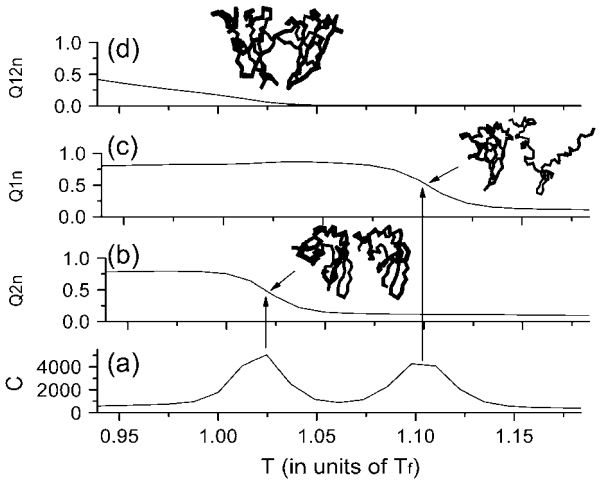

FIGURE 1 $Q12n$ is the fraction of native contacts between different chains, $Q1n$ ($Q2n$) is the fraction of native contacts among residues of the monomer 1 (2), and C is the specific heat. These quantities are plotted as a function of temperature T in units of T_f. One of the two monomers collapses at $T = 1.1\ T_f$ as shown in *b*. The second monomer collapses at the lower temperature $1.02\ T_f$ as shown in *c*. The population of domain-swapped dimers, shown in *a*, increases as temperature decreases below $T = 1.03\ T_f$.

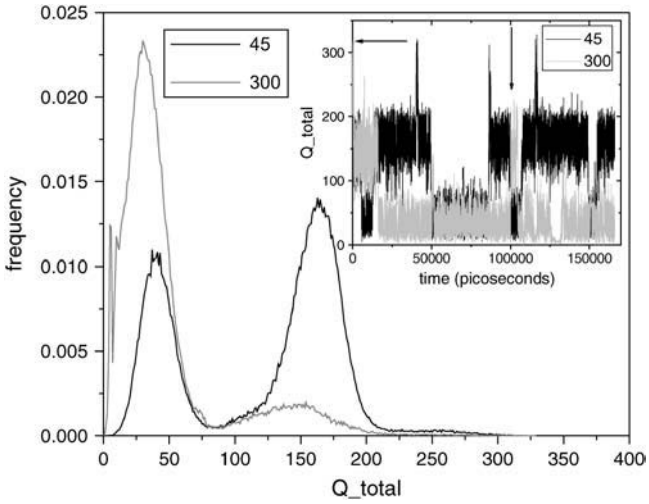

FIGURE 2 Simulations performed at $1.09\ T_f$ to investigate effects of different constrained distances on stability. The inset shows the number of total native contacts against time at two different maximum distances between two monomers, 45 Å and 300 Å. Two simulations started from the same native dimer conformation (indicated by the *horizontal arrow*). After 100,000 ps (indicated by the *vertical arrow*), simulations are terminated. Noticeably, conformations of a dimer with a 300 Å constrained maximum distance (*shaded area*) are more relaxed, reflected by a lower value of Q_total than that with 45 Å (*black area*). The histogram (frequency) of the conformations shows a shift of an equally populated state (*solid line*) when a larger constrained distance is imposed (*shaded line*) instead.

enthalpy change at $1.02\ T_f$ was a result of folding collapses of both monomers (*vertical arrow* at *left*), although there is still a small percentage of domain-swapped dimers that takes <10% of the population. Next, we analyze an increase of $Q1n$, but not in $Q12n$ at $1.1\ T_f$ and suggest that only one of the two monomers unfold at a higher temperature, $1.1\ T_f$. This latter temperature is interesting because it appears with only dimer simulations but is not found in monomer simulations.

We therefore asked what could be the reasons for the second higher collapsed temperature in the dimer simulations. It must be associated with high concentrations because in monomer simulations that project to a much diluted condition there is only one folding temperature. Indeed, current studies on dimer simulations take place in a concentrated condition ($D_{max} = 45$ Å) in which it is possible to have interchain interactions that affect the stability of dimer conformations.

This finding attracts our attention to vary the constrained distance between two monomers. We compare these results with simulations done using a larger constrained maximum distance of 300 Å in Fig. 2. Simulations start from the same native dimer conformations at $1.09\ T_f$ with two different D_{max}. After a certain number of time steps (100,000 ps), there is a larger amount of Q_total in the conformation with $D_{max} = 45$ Å (*solid line*), suggesting it to be more compact than that with $D_{max} = 300$ Å (*shaded line*).

We made a normalized histogram (frequency) from simulations in Fig. 2. It shows a shift from two equally populated states when a larger D_{max} is imposed. The onset of a second peak at a higher temperature must relate to interchain interactions (i.e., chain-chain excluded volume interactions) because as we relax D_{max}, there is an increase of configurational space that minimizes the possibility to access partially unfolded conformations (i.e., open conformations). Consequently, Q_total decreases. This finding can be related to less specific systems such as polymer models in athermal solvent (21,22) where restrictions in a conformational space of chains lead to an entropically driven collapse transition.

Mechanism of dimer formation

Because we are interested in a distribution of conformations at $1.02\ T_f$ and $1.1\ T_f$ that denote two peaks in the specific heat plot (Fig. 1), we look into the potential of mean force (PMF) as a function of some order parameters for structural information. Unfortunately, Q_total failed to differentiate two competing intermediates (i.e., one is composed of a partially formed, domain-swapped dimer structure, and the other is composed of a folded monomer accompanied with an unfolded one) because both give similar values. Therefore, we introduce another convenient order parameter, Q_hybrid, which suitably provides relative stabilities of the two. We define Q_hybrid based on the content of $Q12$: Q_hybrid $= Q1 + Q2$ when $Q12$ is <30% of its maximum value, whereas Q_hybrid $= Q12$ when it is >40% of its maximum value. The reason behind this is the formation of two monomers or a domain-swapped dimer is mutually exclusive in the ensemble. In other words, given a confor-

mation of two monomers that satisfy a domain-swapped dimer, it is not possible to be in the state of two separated monomers at the same time, and vice versa. Q_hybrid is not used for the description of a complete conformational space of monomers and dimer, because it is not a continuous parameter to correspond to the full density of states. However, it is useful to distinguish the unfolded and folded forms about the transitions. It correlates well with the continuous order parameters ($Q1$, $Q2$, and $Q12$) in a short range when close to the transitions.

In Fig. 3, the PMF is plotted against Q_hybrid. We now see that, at 1.02 T_f, there is an intermediate to be a folded monomer with an unfolded one (type 2) that populates more than the other kind of intermediate (a partially formed, domain-swapped dimer (type 1)). At this temperature, the intermediate (type 2) may favor forming two separately folded monomers because its barrier to this minimum is lower than that of a dimer. At a higher temperature, 1.1 T_f, this intermediate coexists with unfolded monomers. On the other hand, at a lower temperature, 0.94 T_f, the population of intermediates decreases and the folded form becomes more stable. The latter attracts our attention to a question of whether intermediates are necessary between transitions of folded monomers and a dimer at low temperatures. To answer this we extend our simulations using the replica exchange method to low temperatures, $0.6\ T_f < T < 1.3\ T_f$, because a single canonical MD simulation takes a prohibitively long time for a folded conformation to escape from its local minimum. The results are shown in Fig. 4. There is

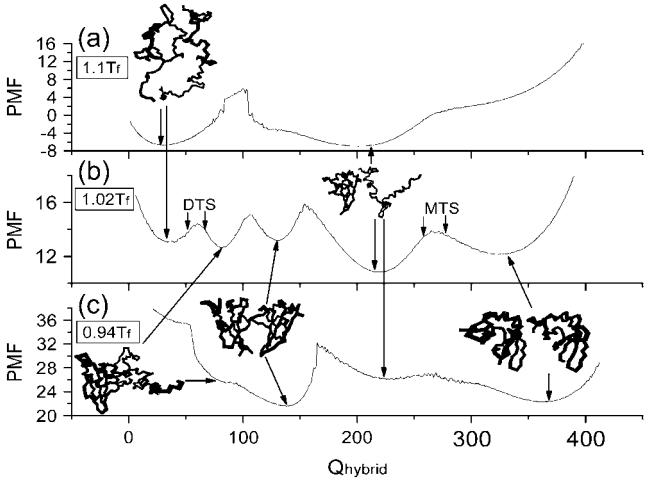

FIGURE 3 Potential of the mean force (in units of kT_f) as a function of order parameter Q_hybrid, which is defined to be $Q1 + Q2$ when $Q12 <$ 30% of its maximum value, and to be $Q12$ when >40% of its maximum value. The profile for 1.02 T_f shows all different phases. A DTS is indicated by two parallel arrows at the left, and an MTS is represented by the two parallel arrows at the right. At higher temperatures (1.1 T_f), unfolded monomers coexist with intermediates that are populated with folded monomers and unfolded ones. At a lower temperature (0.94 T_f), there is a decrease in stabilities of two kinds of intermediates: one is a partially folded dimer, and the other is composed of folded and unfolded monomers.

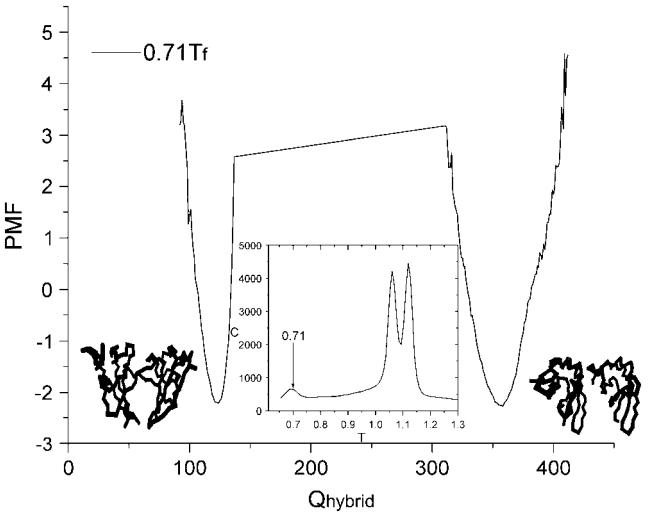

FIGURE 4 The PMF (in units of kT_f) at temperature 0.71 T_f is plotted against the order parameter Q_hybrid. The two minima correspond to the native forms of two separate monomers and a dimer, respectively. The inset is the specific heat as a function of temperature. In addition to two major peaks, there is another one that appears at a much lower temperature, 0.71 T_f.

a relatively small peak at 0.71 T_f, in addition to two major peaks in the specific heat profile, as shown by the inset. The PMF as a function of Q_hybrid at this low temperature suggests both types of intermediates no longer exist; intermediates may not be necessary for transitions between two folded forms, although the barrier between monomers and dimers is relatively high, which could make this transition a rare event.

Analysis of two transition states: monomer and dimer

Before going into the details of the transition states of monomers and dimers, it is instructive to follow some trajectories that start with either an unfolded monomer or a folded dimer and look at associated transitions in Fig. 5, *a* and *b*. The constraint between the two centers of mass (cm_cm) is 27.8 Å, and simulations are performed at 1.03 T_f.

To capture essential structural information in the system, we use intramolecular contacts ($Q12$) and cm_cm to monitor intramolecular interactions between two monomers, the number of native contacts ($Q1$ or $Q2$) to monitor formation of each monomer, and vdW that keeps track of the total van der Waals energy to monitor energetic changes in the system. Fig. 5 *a* starts with two unfolded monomers (snapshot 1), and they rapidly collapsed into a partially folded dimer. Snapshots (2–5) represent a temporal evolution of the partially folded dimer to the native dimer. Interestingly, as the dimer continues to unfold (snapshot 5–7), the number of disrupted intramolecular contacts ($Q12$) is neatly compensated by contacts formed in one of the monomers (Q1, which is the number of contacts among residues of the monomer

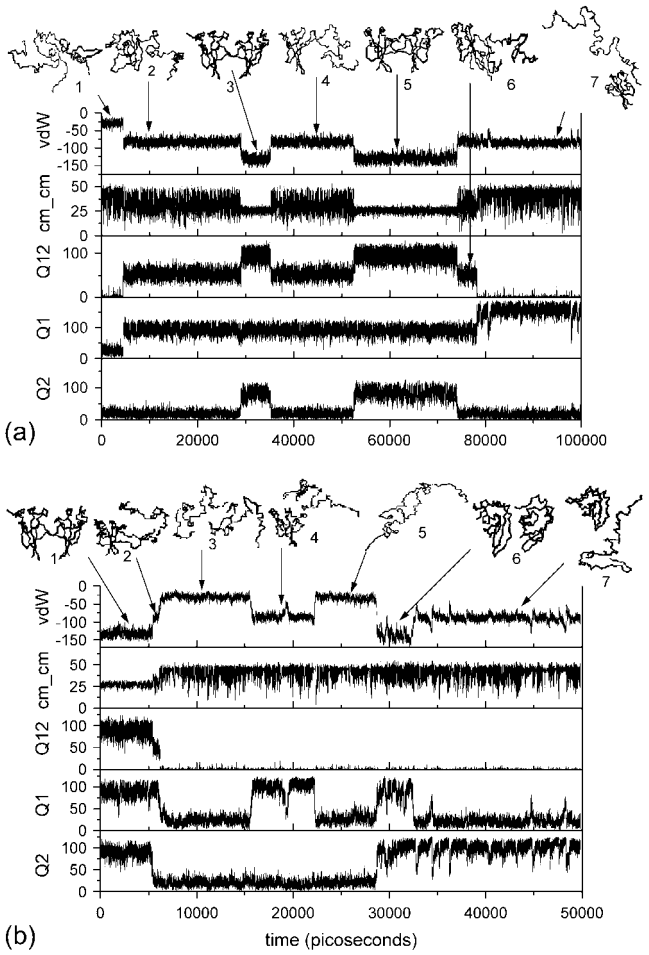

FIGURE 5 cm_cm is the distance between the centroids of two monomers. vdW is the van der Waals energy between residues. $Q12$ is the number of intramolecular contacts. $Q1$ ($Q2$) is the number of contacts in each of the two monomers. Simulations are performed at 1.02 T_f. (*a*) Snapshots 1–5 show temporal evolution of dimer formation from two unfolded monomers, and then unfold again in snapshots 6–7. (*b*) Snapshots 1–4 show a temporal evolution of a dimer to a partially folded dimer, followed by unfolding of both monomers in snapshot 5. Snapshots 6–7 show two folded monomers, and then one of them unfolds.

represented in *blue*). As a result, there is nearly no change in Vdw as the dimer unfolds. Snapshots 1–3 suggest that dimer formation could take place in the presence of the folding transition of monomers, meaning that dimerization via domain swapping occurs through the unfolded state of the monomers, around the temperature T_f. Fig. 5 *b* represents another simulation with a native dimer as initial conformation. Snapshots 1–5 show a temporal evolution of two unfolded monomers. Monomers are unfolded ($Q1$ or $Q2 \approx 0$), and very few intramolecular contacts ($Q12$) are found in snapshot 5. Next, in snapshots 6–7, both monomers fold by themselves and then one of them unfolds. Snapshots 1–6 also suggest that the two monomers, to be formed from the dimeric form, require a large opening of the chains represented by snapshot 5.

In the following, the mechanisms of dimer formation and aspects of the monomer and dimer transition states (DTS) are discussed. The conformational parameters described in Fig. 5 *a* show that monomer 1 is formed almost simultaneously with the insertion of its strand in the other monomer. This is suggested by snapshots 1 and 2, which show the almost simultaneous change in the quantities $Q12$ and $Q1$, after which monomer 2 folds after the insertion of its strand, as shown by the increase of $Q12$ and $Q2$ and as is suggested by snapshots 2–3. The almost simultaneous changes in $Q12$ and $Q1$ followed by the changes in $Q12$ and $Q2$ may indicate that the transition state of the dimer would be correlated with that of the monomer. Unfolding and folding kinetics experiments (23) have shown evidence that the structure of the dimeric transition state is similar to the monomeric transition state. This would be the explanation for the equal rates of unfolding for dimers and monomers. Those experiments also showed, by ϕ-value analysis, that the monomeric transition state is highly correlated with the dimeric transition state.

Fig. 6 shows the extent of contact formation that has $\phi_{ij} > 0.5$ (See Methods for definitions of ϕ_{ij} at two distinct transition states). (A contact is said to be made when the

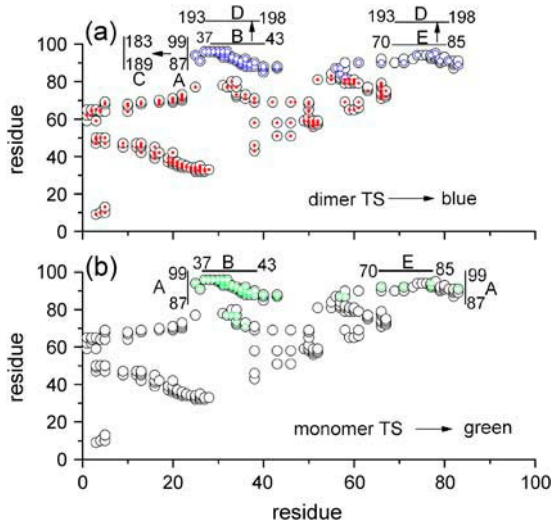

FIGURE 6 Black circles are the native contact maps of the dimer and monomer. The contacts made at the transition state of the dimer (monomer) are shown by the blue (green) circles. Circles that do not have the red dots inside are those contacts among residues of the native monomer which are disrupted in the native dimer. These contacts constitute the transition state. In the transition state of the monomer, region *A*, formed by residues 87–99, comes close to region *E* of the chain formed by residues 70–85. In the next step, region *B*, formed by residues 37–43, joins the other two. In the DTS, the contacts are made by residues of different chains, as shown by the arrows, but the residues of one chain involved in these contacts are the same as those for the MTS. Region *B* of the first chain (residues 37–43) binds to region *D* of the second chain (residues 193–198), and region *A* of the first chain (residues 87–99) binds to region *C* of the other chain (residues 183–189). In addition, region *E* of the first chain (residues 70–85) binds to region *D* of the second chain (residues 193–198).

distance between two C_α values is <7 Å. The contacts that are likely to be made at the transition state are determined by the following procedure: the two pairs of vertical arrows shown in Fig. 3 *b* delimit the DTS and MTS. The probability of contact formation is calculated in the neighborhood of the left arrow of each pair and then compared to the probability of contact formation calculated in the neighborhood of the right arrow. The highest differences between these two quantities are represented in color in Fig. 6.) The black circles are native contacts. Blue circles mark contacts of a dimer at the transition states, and the green ones mark contacts of a monomer at the transition states. Comparing the distribution of blue and green circles in two panels, we suggest that two transition states are structurally similar. Such a finding suggests that these crucial contacts responsible for domain swapping have lower frequencies in a folded monomer due to an obligatory opening to form the dimer and are the same crucial contacts that must have higher frequencies to fold the monomer.

The data represented in Fig. 6 are in qualitative agreement with experimental φ values (23) where the highest (>0.5) values for ϕ, calculated for dimer and monomer, are those in the neighborhood of residue 38 and those in the neighborhood of residue 93. These regions are marked by the horizontal and vertical segments close to the green circles of Fig. 6, where it is easy to notice the similarity of the transition state of the monomer (*green circles*) and the transition state of the dimer (*blue circles*). In Fig. 6 *b* region *A* is in contact with region *B* and also with region *E* of the chain. In Fig. 6 *a* the contacts are made between the two chains. Regions *A* and *C* are in contact as are *B* and *D* and also *D* and *E*. Fig. 6, *a* and *b*, shows similar transition states for monomers and dimers. Mutagenesis experiments (23) also showed that the monomeric transition state is highly correlated with the dimeric transition state. Such agreements between simulations and experiments suggest that the structure of the dimeric transition state should be the same as the monomeric transition state, accounting for similar unfolding rates in dimers and monomers.

Dimerization follows a "lock-and-dock" mechanism via domain swapping

To understand the mechanism of the DTS, we must know how the monomer forms in structural details. The monomer transition state (MTS) is characterized by two crucial steps described in Figs. 6 *b* and 7. At first, the strand *A* (residues 87–99) interacts with the region *E* (residues 70–85). This step is represented by the dashed line close to region *E* in Fig. 7. In the next step, region *B* (residues 37–43) joins the other two regions, as shown by the other dashed line. Once regions *E* and *B* are held together by strand *A*, the monomer proceeds to form, which is attributed to an increase in *Q*1. The contacts made in these two steps are those shown in green circles in Fig. 6 *b*: those labeled for *A* and *E* represent

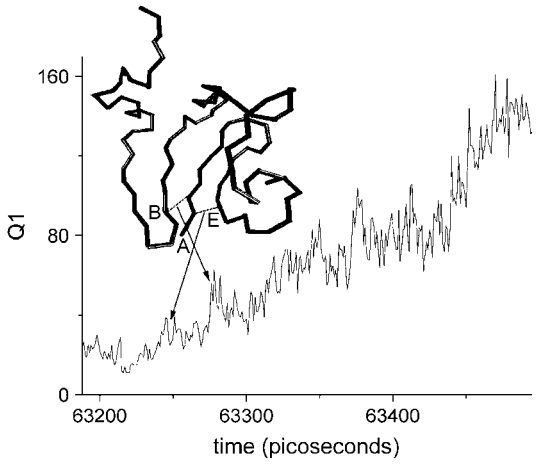

FIGURE 7 MTS is characterized by two crucial steps: first, strand *A* binds to the region labeled as *E* (*thin line close to E*); second, the region labeled as *B* joined (*thin line close to B*). Once region *E* and *B* both bind to strand *A*, the monomer proceeds to form, indicated by an increase in *Q*1. These findings are suggested by a drastic increase in the frequency of contacts in those regions at the transition state.

the contacts made in the first step, and those labeled for *A* and *B* refer to the second step.

Next we learn how dimer forms from the simulations. Figs. 6 *a* and 8 reveal structural details of the dimer formation, whose transition state is also characterized by two crucial steps. First, strand *A* of one chain interacts with

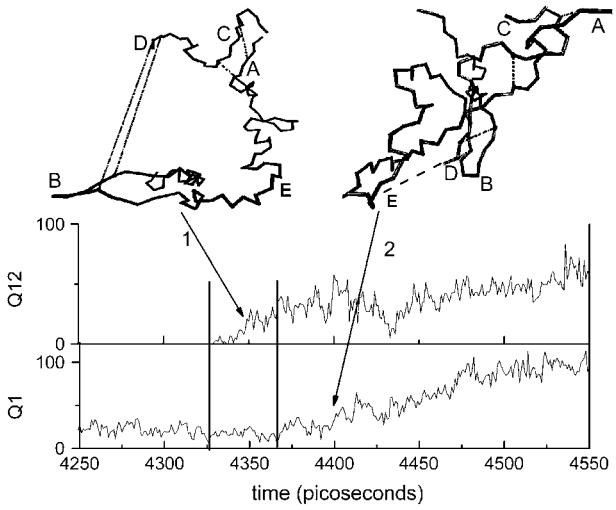

FIGURE 8 DTS is characterized by two crucial steps: first, the strand *A* of one chain binds to the region labeled as *C* of the other chain, as shown by the upper left conformation. *Q*12 increases, but not *Q*1, as indicated in the region defined by two solid parallel lines; in the next step, the distance between regions *D* and *B* decreases (*figure upper right*). This step is analogous to the bound *A*–*E* of Fig. 9 that triggers the monomer formation. The monomer continues to form (in the dimeric structure) as region *D* comes close to region *E*, which is equivalent to the formation between *A* and *B* in Fig. 9. At this point, *Q*1 starts to increase (indicated by *arrow 2*). These findings are suggested by the drastic increase in the frequency of contacts in those regions at the transition state.

region *C* (residues 183–189) of the other chain, as shown in the snapshot at left in Fig. 8, where the two shorter dashed lines represent the formed native contacts (we call this the ''lock'' step). At this point, the monomeric form is no longer possible because strand *A* is blocked and thus impeded from interacting with region *B* (which is a crucial step for the monomeric form). At this point, regions *B* and *D* are still far from each other, as shown by longer dashed lines. $Q12$, which represents interchain contacts, increases, but not the intrachain contact $Q1$, as indicated in the region defined by the two solid parallel lines. Then, the distance between region *D* and *B* decreases (shown by the *upper right* snapshot in Fig. 9) followed by the approach of region *E*. This step plays a similar role of binding regions *A* and *B* in the monomer formation described by Fig. 8. The chain acts as if its strand *A* would be indistinguishably replaced by *D*, which plays the same role to trigger the monomer formation as *D* approaches *B*, followed by the approach of region *E*. At this point, $Q1$ starts to increase (*arrow 2*) and the ''monomer'' is formed in the dimeric structure (we call the approaching of regions *B*, *D*, and *E* the ''dock'' step). This ''lock-and-dock'' mechanism is the essence of the domain-swapping process in this protein.

CONCLUDING REMARKS

The two peaks in the specific heat of Fig. 1, corresponding to two different folding temperatures for the monomers, seem to be related to the confinement of the chains. There are several results concerning the effects of confinement on the thermodynamics of proteins and on protein models (24–29). Some of these studies show that folding temperatures increase in a restricted conformational space environment by stabilizing the folded conformation against reversible unfolding. In this study this effect is manifested by increasing the folding temperature of the monomer, and once one of them folds, the other will, in a less restricted conformational space, fold closer to the temperature the two monomers would fold to if they were separated by a relatively large distance that could prevent the restriction on their conformational space.

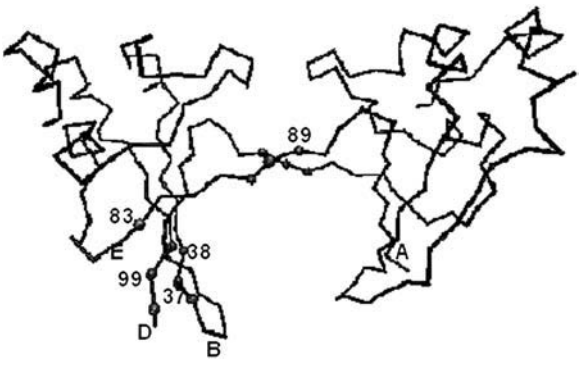

FIGURE 9 Selected monomers responsible for the lock and dock steps at the transition states are highlighted on the native dimer structure. At the transition state of the monomer formation (e.g., left monomer), strand *D* of the right monomer is swapped with strand *A*. Some monomers that have high Φ-values on strand *D*, regions *E* and *B* (as in Rousseau et al. (23)), at the dock step have to come together in the transition states of the monomeric and dimeric formations. Particularly, residue 89, which also has a high Φ-value, along with other monomers, is important for the lock step, which agrees with experimental findings in Rousseau et al. (23).

The formation of the dimeric form of p13suc1 through domain swapping is highly controlled by the presence of intermediate states. The main form of these states is characterized by a folded monomer with the other unfolded. This state will convert less to the dimeric form than to both folded monomers due to the relatively higher barrier to form the dimer. The predominance of the monomeric form was experimentally found by Itzhaki and co-workers (30), who also showed the presence and role of intermediates in the domain swapping of p13suc1. When intermediates are populated, rollover occurs on the chevron plots of p13suc1, causing the system to be under kinetic control. It is reported that the amount of dimer decreases in the presence of intermediates. The data from this study correlate with the results of that article. These simulations also showed a small peak of the specific heat at low temperatures (0.71 T_f) related to a transition where native monomers coexist with a significant population of native dimer without the presence of intermediates, although a relatively high barrier appears between the folded forms.

The fact that most of the transition state of the dimer is formed by interresidue contacts (as shown by Fig. 6 *a*) and also because of the similarity between the two transition states (as shown by Fig. 6, *a* and *b*), the following reasoning is plausible: Let us assume that monomers 1 and 2 have *N* residues each. Fig. 6 *a* shows that most of the interchain contacts belong to the transition state of the dimer which is similar to the MTS. This means that, also due to the symmetry of the domain swapping, if we take the interchain contacts and subtract the residue number of the second chain by *N*, we would obtain the intrachains contacts which would be crucial in the transition state of the monomer. Among those interchain contacts that do not contribute to the MTS are those few contacts related to the lock mechanism described above. But the majority of the interchain contacts are related to the dock mechanism that triggers the monomer formation. That is why dimer and MTS are correlated such that their interchain contacts could reveal the important contacts made at the transition state of the monomer.

Fig. 9 shows the native dimer and its transition state, represented by some residues involved in the lock and dock steps. The transition state of the left monomer is reached when strand *D* of monomer 2 is replaced by strand *A*. We compare the experimental Φ-values ($\Phi > 0.5$) of some monomers at strand *D* and region *B* in (23) with our simulation results in Fig. 6, and they also have $\phi_{ij} > 0.5$. Such an agreement encourages us to suggest that strand *D* and several monomers of region B (37–43) have to come

together at the transition states of both monomeric and dimeric formation (the dock step). Particularly, residue 89, which has a high Φ-value (23), should mainly present in the dimeric form and it is important to complete the lock step (represented by the monomers close to residue 89 in Fig. 9).

As a final comment, we believe that correlation between the DTS and the MTS may be a general result of domain-swapping phenomena and not a particular feature of p13suc1. It is likely that the ''lock-and-dock'' mechanism described in the previous section is behind the dimerization by domain-swapping processes.

We thank Dr. José Onuchic for stimulating discussions and Dr. Laura Itzhaki for carefully reading the manuscript. We are also grateful to Sichun Yang for helping us with the replica exchange method.

J.C. thanks the Brazilian Agency CNPq for financial support. M.S.C. thanks the Alfred P. Sloan Foundation for a postdoctoral fellowship. Computational resources were supported by the National Science Foundation-sponsored Center for Theoretical Biological Physics in San Diego (PHY-0216576 and PHY-0225630).

REFERENCES

1. Bennett, M. J., M. P. Schlunegger, and D. Eisenberg. 1995. 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.* 4:2455–2468.
2. Schlunegger, M. P., M. J. Bennet, and D. Eisenberg. 1997. Oligomer formation by 3D domain swapping: a model for protein assembly and misassembly. *Adv. Protein Chem.* 50:61–122.
3. Bennett, M. J., S. Choe, and D. Eisenberg. 1994. Domain swapping—entangling alliances between proteins. *Proc. Natl. Acad. Sci. USA.* 91:3127–3131.
4. Fink, A. L. 1998. Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold. Des.* 3:R9–R23.
5. Cohen, F. E., and S. B. Prusiner. 1998. Pathologic conformations of prion proteins. *Annu. Rev. Biochem.* 67:793–819.
6. Chen, Y. W., K. Stott, and M. F. Perutz. 1999. Crystal structure of a dimeric chymotrypsin inhibitor 2 mutant containing an inserted glutamine repeat. *Proc. Natl. Acad. Sci. USA.* 96:1257–1261.
7. Kishan, K. V., G. Scita, D. T. Wong, P. P. Di Fiore, and M. E. Newcomer. 1997. The SH3 domain of Eps8 exists as a novel intertwined dimer. *Nat. Struct. Biol.* 4:739–743.
8. Raaijmakers, H., O. Vix, I. Toro, S. Golz, B. Kemper, and D. Suck. 1999. X-ray structure of T4 endonuclease VII: a DNA resolvase with a novel fold and unusual domain-swapped dimer architecture. *EMBO J.* 18:1447–1458.
9. Rousseau, F., J. W. H. Schymkowitz, H. R. Wilkinson, and L. S. Itzhaki. 2001. Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc. Natl. Acad. Sci. USA.* 98:5596–5601.
10. Bryngelson, J. D., and P. G. Wolynes. 1987. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. USA.* 84:7524–7528.
11. Leopold, P. E., and J. N. Onuchic. 1992. Protein folding funnels: kinetic pathways through compact conformational space. *Proc. Natl. Acad. Sci. USA.* 89:8721–8725.
12. Clementi, C., H. Nymeyer, and J. N. Onuchic. 2000. Topological and energetic factors: what determines the structural details of the transition state ensemble and ''on route'' intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* 298:937–953.
13. Onuchic, J. N., H. Nymeyer, A. E. Garcia, J. Chahine, and N. D. Socci. 2000. The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. *Adv. Protein Chem.* 53:87–152.
14. Levy, Y., A. Caflisch, J. N. Onuchic, and P. G. Wolynes. 2004. The folding dimerization of HIV-1 protease: evidence for a stable monomer from simulations. *J. Mol. Biol.* 340:67–79.
15. Alonso, D. O. V., E. Alm, and V. Dagget. 2000. The unfolding pathway of the cell-cycle protein p13suc1: implications for domain swapping. *Structure.* 8:101–110.
16. Schymkowitz, J. W. H., F. Rousseau, L. R. Irvine, and L. S. Itzhaki. 2000. The folding pathway of the cell-cycle regulatory protein p13suc1: clues for the mechanism of domain swapping. *Struct. Fold. Des.* 8:89–100.
17. Taketomi, H., Y. Ueda, and N. Go. 1975. Studies on protein folding, unfolding and fluctuations by computer simulation. *Int. J. Peptide Res.* 6:445–459.
18. Yang, S., S. S. Cho, Y. Levy, M. S. Cheung, H. Levine, P. G. Wolynes, and J. N. Onuchic. 2004. Domain swapping is a consequence of minimal frustration. *Proc. Natl. Acad. Sci. USA.* 101:13786–13791.
19. Sugita, Y., and Y. Okamoto. 2000. Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape. *Chem. Phys. Lett.* 329:261–270.
20. Swendsen, R. H. 1993. Modern methods of analyzing Monte Carlo computer simulations. *Physica A.* 194:53–62.
21. Dijkstra, M., D. Frenkel, and J. P. Hansen. 1994. Phase-separation in binary hard-core mixtures. *J. Chem. Phys.* 101:3179–3189.
22. Luna-Bárcenas, G., G. E. Bennett, I. C. Sanchez, and K. P. Johnston. 1996. Monte Carlo simulation of polymer chain collapse in athermal solvents. *J. Chem. Phys.* 104:9971–9973.
23. Rousseau, F., J. W. H. Schymkowitz, H. R. Wilkinson, and L. S. Itzhaki. 2002. The structure of the transition state for folding of domain-swapped dimeric p13suc1. *Structure.* 10:649–657.
24. Eggers, D. K., and J. S. Valentine. 2001. Molecular confinement influences protein structure and enhances thermal protein stability. *Protein Sci.* 10:250–261.
25. Friedel, M., D. J. Sheeler, and J.-E. Shea. 2003. Effects of confinement and crowding on the thermodynamics and kinetics of folding of a minimalist β-barrel protein. *J. Chem. Phys.* 118:8106–8113.
26. Minton, A. P. 2000. Implications of macromolecular crowding for protein assembly. *Curr. Opin. Struct. Biol.* 10:34–39.
27. Minton, A. P. 2001. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J. Biol. Chem.* 276:10577–10580.
28. Zhou, H. X., and K. A. Dill. 2001. Stabilization of proteins in confined spaces. *Biochemistry.* 40:11289–11293.
29. Klimov, D., D. Newfield, and D. Thirumalai. 2002. Simulations of beta-hairpin folding confined to spherical pores using distributed computing. *Proc. Natl. Acad. Sci. USA.* 99:8019–8024.
30. Rousseau, F., J. W. H. Schymkowitz, H. R. Wilkinson, and L. S. Itzhaki. 2004. Intermediates control domain swapping during folding of p13suc1. *J. Biochem. (Tokyo).* 279:8368–8377.