Mimicking the action of GroEL in molecular dynamics simulations: Application to the refinement of protein structures

Hao Fan and Alan E. Mark

Protein Sci. 2006 15: 441-448; originally published online Feb 1, 2006;
Access the most recent version at doi:10.1110/ps.051721006

Mimicking the action of GroEL in molecular dynamics simulations: Application to the refinement of protein structures

HAO FAN[1] AND ALAN E. MARK[1,2]

[1]Groningen Biomolecular Sciences and Biotechnology Institute (GBB), Department of Biophysical Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands
[2]School of Molecular and Microbial Sciences, University of Queensland, Queensland, Australia

(RECEIVED July 23, 2005; FINAL REVISION November 24, 2005; ACCEPTED November 26, 2005)

Abstract

Bacterial chaperonin, GroEL, together with its co-chaperonin, GroES, facilitates the folding of a variety of polypeptides. Experiments suggest that GroEL stimulates protein folding by multiple cycles of binding and release. Misfolded proteins first bind to an exposed hydrophobic surface on GroEL. GroES then encapsulates the substrate and triggers its release into the central cavity of the GroEL/ES complex for folding. In this work, we investigate the possibility to facilitate protein folding in molecular dynamics simulations by mimicking the effects of GroEL/ES namely, repeated binding and release, together with spatial confinement. During the binding stage, the (metastable) partially folded proteins are allowed to attach spontaneously to a hydrophobic surface within the simulation box. This destabilizes the structures, which are then transferred into a spatially confined cavity for folding. The approach has been tested by attempting to refine protein structural models generated using the ROSETTA procedure for ab initio structure prediction. Dramatic improvements in regard to the deviation of protein models from the corresponding experimental structures were observed. The results suggest that the primary effects of the GroEL/ES system can be mimicked in a simple coarse-grained manner and be used to facilitate protein folding in molecular dynamics simulations. Furthermore, the results support the assumption that the spatial confinement in GroEL/ES assists the folding of encapsulated proteins.

Keywords: protein structure prediction; molecular dynamics; structure refinement; chaperone; GroEL

Molecular chaperones facilitate the folding of a wide range of proteins in vivo. The best characterized chaperonin, GroEL, is a symmetric tetradecamer that resembles a hollow cylinder open at both ends. The apical domains, located at each entrance to the cylinder, bind alternatively the protein substrate and the co-chaperone GroES (Braig et al. 1994; Xu et al. 1997; Wang and Boisvert 2003; Wang and Chen 2003). The folding cycle of GroEL is thought to have two basic stages (Bukau and Horwich 1998). First, non-native polypeptides bind to GroEL through an exposed hydrophobic surface (Fenton et al. 1994; Hlodan et al. 1995; Itzhaki et al. 1995; Lin et al. 1995). The binding of the misfolded proteins is believed to prevent the aggregation and may also induce partial unfolding helping some proteins escape from kinetically trapped states (Zahn et al. 1994, 1996; Ranson et al. 1995; Shtilerman et al. 1999). Second, the binding of GroES to GroEL encapsulates

Reprint requests to: Prof. Alan E. Mark, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), Department of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands; e-mail: a.e.mark@rug.nl; fax: +31-50-3634800.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.051721006.

the bound polypeptide and triggers its release into the central channel of GroEL/ES where folding occurs (Weissman et al. 1995; Mayhew et al. 1996). Based on experimental evidence and on theoretical considerations, two possible mechanisms by which GroEL may enhance folding have been proposed. One is based on the concept of iterative annealing. GroEL is assumed to optimize protein folding by multiple cycles of binding and release (Todd et al. 1996; Betancourt and Thirumalai 1999; Thirumalai and Lorimer 2001; Stan et al. 2003). Repeated binding is assumed to disrupt poorly packed regions of the protein structure helping the system escape from kinetic traps. In this way, GroEL increases the opportunity for misfolded proteins to find the most thermodynamically stable state, leading to the acceleration of folding. An alternative model proposes that a primary function of GroEL/ES is to encapsulate and confine the substrate (Kim et al. 1998; Brinker et al. 2001). This is considered to facilitate folding because spatial confinement is expected to favor the formation of compact, native-like states.

A number of simulation studies have been performed to investigate chaperone-mediated protein folding. Calculations using lattice models have supported the idea that a chaperone, by providing a hydrophobic environment, could potentially disrupt misfolded protein domains and in this way accelerate folding (Chan and Dill 1996; Sfatos et al. 1996; Betancourt and Thirumalai 1999). These calculations also indicated that the repeated oscillation of the hydrophobicity of the environment (iterative annealing) could both reduce the folding time and increase the yield of the folded protein. Off-lattice simulations have been performed to examine the effect of confinement on folding using coarse-grained models of the protein and the chaperone (Klimov et al. 2002; Baumketner et al. 2003; Takagi et al. 2003; Jewett et al. 2004). It has been suggested that the encapsulation in the chaperone cage reduces the entropy of the denatured states, leading to an enhancement of the rate of folding (Klimov et al. 2002). In a previous study, we attempted to mimic the generic action of chaperones in atomistic molecular dynamics (MD) simulations by cycling the polarity of the solvent environment (Fan and Mark 2004a). The approach was tested by attempting to refine protein models generated by Baker and coworkers using the ROSETTA algorithm (Simons et al. 1999, 2001; Bonneau et al. 2001). It was shown that an oscillating solvent environment enabled the system to cross barriers in the free energy landscape, greatly facilitating the search for the native structure. Most importantly, such studies demonstrated that it is possible to mimic the effect of chaperones on folding using very simple principles in atomistic simulations.

In this work, the combined effect of iterative annealing and spatial confinement on protein folding is examined in explicit solvent and at atomic detail using MD simulations. To mimic the action of GroEL, we explicitly incorporate a spherical cage into the simulation and cycle the cage environment. The cage initially displays a hydrophobic surface to which the protein substrate may bind. The nature of the surface is then changed to release the protein into a spatially confined cavity. The effect of including this chaperone-like cage was tested by attempting to refine protein models generated by Baker and coworkers using the ROSETTA algorithm (Simons et al. 1999, 2001; Bonneau et al. 2001), which were considered to represent compact misfolded proteins. The models chosen were the same as those used in our previous studies on refinement and deviate significantly from the experimental structures (Fan and Mark 2004b).

Results and Discussion

Effect of the chaperone cage on native folded proteins

This study consisted of two phases. First, control simulations were performed starting from the respective experimental structures of the three proteins investigated (PDBID 1vcc, 1sro, and 1afi) (Sharma et al. 1994; Bycroft et al. 1997; Steele and Opella 1997). The aim of the control simulations was to determine the effect of three refinement protocols on a native folded protein. In the first stage of the refinement protocols, each structure was subjected to 5-nsec molecular dynamics (MD) simulation in a cage primarily consisting of CH2 groups and presenting a hydrophobic surface to the proteins. The conformation obtained after 5 nsec was then subjected to the second stage of the refinement. This differed between the three protocols. In protocol I, the surface of the chaperone cage was made more hydrophilic by reducing the number of CH2 groups and increasing the number of amide groups. In protocol II, a cage constructed of repulsive particles (CY) was used. In protocol III, the cage was removed and the structure was simulated in pure water. The three proteins showed similar behavior with respect to each of the three protocols. This is illustrated in Figure 1 for the case of 1afi. The top row (Fig. 1a–c) shows the positional backbone root-mean-square deviation (RMSD) of the elements of secondary structure from the corresponding NMR structure as a function of the simulation time. During the first stage of the protocol, the protein rapidly migrated from the center to the walls of the cavity and adhered to the hydrophobic surface. The minimum distance between the protein and the cage surface progressively decreased from ~1.0 nm to 0.25 nm. This is associated with an increase in the RMSD, reflecting the fact that the protein partially unfolds on contact with the hydrophobic surface. During the second stage of the refinement protocols refolding is evident in all cases. In

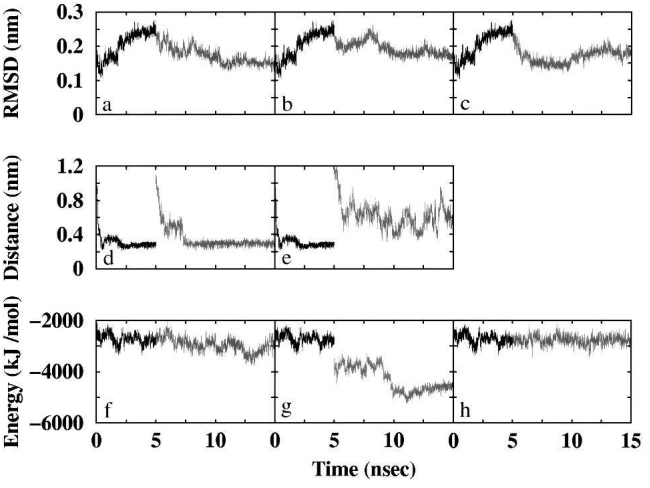

Figure 1. (*Top* row) The time evolution of the backbone positional root mean square deviation (RMSD) of the elements of secondary structure for the NMR structure, 1afi, when subjected to each of the three refinement protocols: refolding in a hydrophilic cage (*a*); refolding in a repulsive cage (*b*); refolding in bulk water (*c*). (*Middle* row) The time evolution of the minimum distance between the protein substrate and the internal surface of the cage: refolding in a hydrophilic cage (*d*); refolding in a repulsive cage (*e*). (*Bottom* row) The time evolution of the internal potential energy of the protein: refolding in a hydrophilic cage (*f*); refolding in a repulsive cage (*g*); refolding in bulk water (*h*).

the case of protocol I, the unfolded protein, which was transferred back to the center of the cavity, rebound to the hydrophilic surface. In contrast, in protocol II the protein remained mobile in the confined space formed by the repulsive particles (Fig. 1e). A notable feature that distinguishes the different protocols is a sharp decrease in the intra-protein potential energy associated with the transfer of the unfolded protein into the repulsive cage (Fig. 1g).

Refinement of the modeled structures

The control simulations established that the refinement protocols did not significantly disrupt the native structure of the proteins. The protocols induced a slight degree of unfolding on binding but the structures refolded when released. The effectiveness of the three protocols for refinement was then tested by attempting to refine ROSETTA models of the proteins 1vcc, 1sro, and 1afi. The same three refinement regimes, as applied to the experimental structures, were used. First, the proteins were encapsulated in the hydrophobic cage and allowed to bind to the internal surface of the cage. Then the bound substrates were discharged into one of three different environments: (1) a hydrophilic cavity; (2) a repulsive, spatially confined cage; and (3) pure water. This procedure was repeated 10 times. As observed for the experimental structures, all of the proteins showed a high affinity for the hydrophilic surface used during the refolding stage in protocol I. As the binding to the hydrophilic surface inhibited refolding of the protein models the following discussion will primarily focus on protocols II and III.

Figure 2 shows the backbone positional RMSD of the elements of secondary structure for the three model structures compared to the respective experimental structures and the intra-protein potential energy during 10 cycles of refinement using protocols II and III. The specific effects of the refinement protocols on the three test proteins are described below. Some general comments can, however, be made. First, the binding of the substrate to the hydrophobic surface in the first stage of each cycle, in general, results in the disruption of the structure especially in poorly packed regions. This effect was reflected by an increase in the RMSD and an increase in the intra-protein potential energy. Combined with the observation from the control simulations that the binding to the hydrophobic surface did not significantly affect the native state of the proteins tested (Fig. 1), this suggests that the binding to the hydrophobic surface is able to facilitate the escape of the structures from metastable non-native conformations characterized by weak hydrogen bonding interactions and poor hydrophobic packing. Second, while the refolding environment in both protocols II and III resulted in a decrease in both the RMSD and the intra-protein potential energy during most cycles, it is clear that refolding in the repulsive cage with spatial confinement is much more effective than refolding in pure water. In particular, this can be seen in relation to the intra-protein potential energy, which has a clear step-like appearance in protocol II but is comparatively smooth in protocol III. This suggests that spatial confinement strongly favors the formation of compact native-like states in the simulations and by analogy in GroEL/ES.

1vcc

The initial RMSD of the ROSETTA model for 1vcc from the X-ray structure is 0.54 nm. After 10 cycles of refinement using protocol II the RMSD was reduced to 0.47 nm (minimum deviation 0.42 nm). Using protocol III, the RMSD of the final conformation was 0.75 nm. The RMSD alone, however, does not fully reflect the degree of refinement observed. Figure 3 shows the initial ROSETTA model of 1vcc, the structure after refinement using protocol II and the corresponding X-ray structure all orientated in a similar manner. In the X-ray structure of 1vcc, two short, parallel α-helices lie behind a plane formed by five β strands. The first three β strands on the left side are derived from the N terminus while the fourth and fifth strands arise from the C terminus. In the ROSETTA model, the C-terminal β-sheet is missing. More importantly, the model structure deviates from the experimental structure in regard to the orientation and alignment of the N-terminal structure. In the X-ray structure the N and C

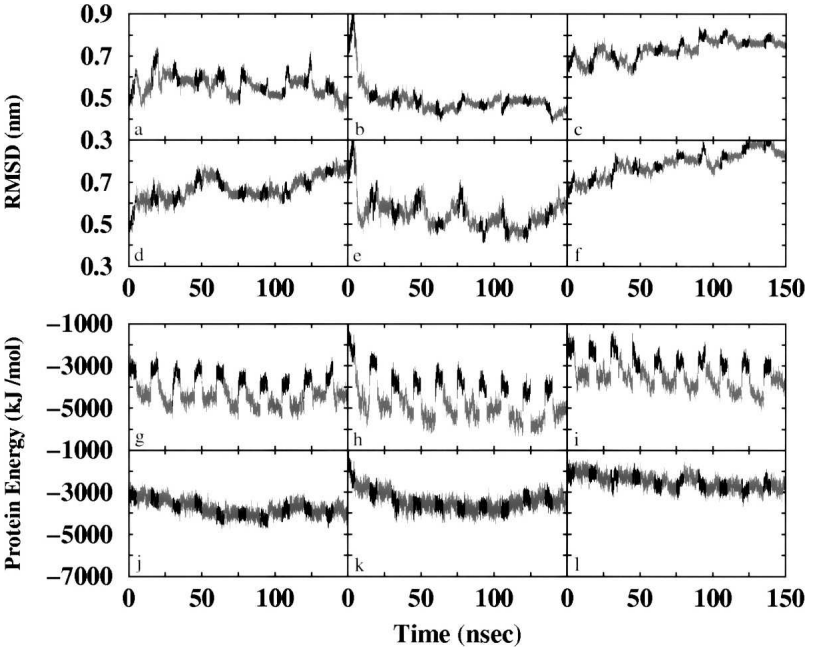

Figure 2. (*Upper* graphs) The time evolution of the RMSD of the elements of secondary structure of the three ROSETTA models for the proteins 1vcc (*a*), 1sro (*b*), and 1afi (*c*) when refolding in a repulsive cage, and for 1vcc (*d*), 1sro (*e*), and 1afi (*f*) refolded in bulk water. (*Lower* graphs) The time evolution of the internal potential energy of the three protein models 1vcc (*g*), 1sro (*h*), and 1afi (*i*) refolding in a repulsive cage, and for 1vcc (*j*), 1sro (*k*), and 1afi (*l*) refolding in bulk water.

termini are parallel, whereas they are antiparallel in the ROSETTA model. Refinement using protocol II not only resulted in the formation of the complete β-plane but also corrected the orientation of the N and C termini yielding the correct global fold. This very significant improvement was associated with a lower RMSD. Nevertheless, problems with the structure remain. In particular, one helix was lost. In addition, the β-plane is more rounded in the refined structure than that observed in the X-ray structure.

1sro

The initial RMSD deviation of the ROSETTA model of 1sro from the experimental structure is 0.82 nm. After 10 cycles of refinement, this was reduced to 0.42 nm and 0.62 nm using protocol II and III, respectively. The minimum RMSD to the experimental structure was 0.37 nm obtained using protocol II. The ROSETTA model, the final structure after 10 cycles of refinement using protocol II and the NMR structure of 1sro are presented in Figure 4. The initial model of 1sro contains a similar percentage of secondary structure as the NMR structure but has a completely different overall fold. In the initial model, the three-stranded β-sheet at the N terminus is parallel to the two C-terminal β strands. In the NMR structure, however, the two β-sheets are perpendicular to each other and are bridged by the first strand at the N terminus. During the initial few cycles

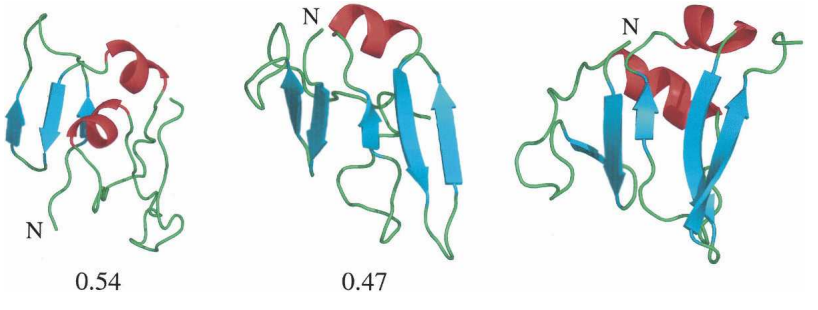

Figure 3. (From *left* to *right*) The ROSETTA model of the protein 1vcc, the final structure after 150 nsec of refinement using 10 cycles of protocol II (repulsive cage), the structure of 1vcc determined experimentally (X-ray). The number *below* each figure corresponds to the backbone RMSD of the elements of secondary structure (nm) with respect to the experimental structure.

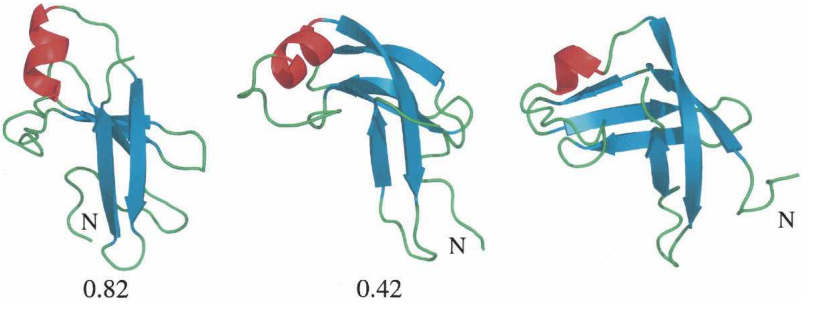

Figure 4. (From *left* to *right*) The ROSETTA model of the protein 1sro, the final structure after 150 nsec of refinement using 10 cycles of protocol II (repulsive cage), the structure of 1sro determined experimentally (NMR). The number *below* each figure corresponds to the backbone RMSD of the elements of secondary structure (nm) with respect to the experimental structure.

of refinement using protocol II, the tertiary structure rearranged accompanied by a large decrease in the RMSD. The relative orientation of the elements of secondary structure remained generally correct thereafter. The percentage of secondary structures, however, fluctuated significantly during the cycles. Although the correct global fold was achieved rapidly, the first strand at the N terminus only began to oscillate between the two regions of β-sheet after 90 nsec of simulation (6 cycles). This movement resulted in the formation of the correct bridging motif during the last few cycles.

1afi

In the case of 1afi, the model structure unfolded during both protocols. The RMSD values of the final conformations were significantly higher than that of the initial ROSETTA model (0.72 nm). Some initial improvement in the RMSD was, however, observed in the first few cycles using protocol II with a minimum RMSD of 0.60 nm. The ROSETTA model, the structure after 10 cycles of protocol II and the NMR structure of 1afi are presented in Figure 5. The initial model of 1afi has localized errors in the β-sheet and major discrepancies in the packing of the helices. The starting model, however, bound very strongly to the hydrophobic surface and experienced a large disruption in structure every cycle. Initially the refinement protocol led to a more native-like packing of the helices , which was associated with the slight decrease in RMSD during the early cycles (Fig. 2c). However, repeated cycles of binding and release led to the progressive loss of secondary structure and the adoption of a more elongated conformation.

The effect of the chaperone cage

Although significant improvements were observed in two of the three test proteins it is clear that the current cage and refinement protocol is not optimal. Different proteins interact quite differently with the cage suggesting that for refinement different types of cages might need to be used for different proteins. In particular, the specific cage used in this work proved more disruptive to regions of α-helix then to regions of β-sheet. Major helical regions were lost in both 1vcc and 1afi. This might be related to the high degree of curvature of the cage or the packing density of the hydrophobic groups (CH2). It is also clear that given the degree of disruption observed, especially in the case of 1afi, the length of time available for refolding was insufficient.

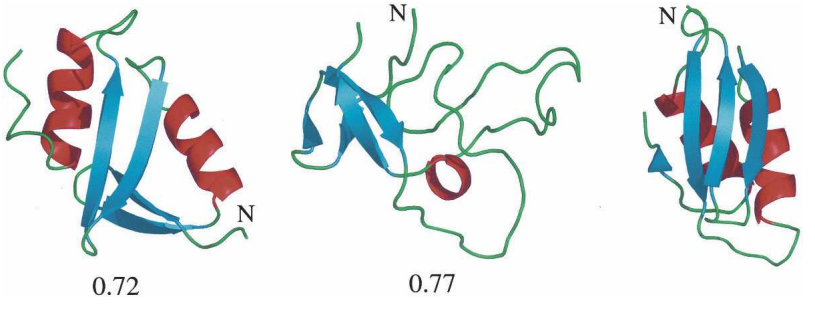

Figure 5. (From *left* to *right*) The ROSETTA model of the protein 1afi, the final structure after 150 nsec simulation using 10 cycles of protocol II (repulsive cage), the structure of 1afi determined experimentally (NMR). The number *below* each figure corresponds to the backbone RMSD of the elements of secondary structure (nm) with respect to the experimental structure.

The final question is, to what extent the simulations shed light on the mechanism of action of GroEL and related chaperones? In this regard it must be stressed that the primary aim of the work was not to model GroEL explicitly. Instead, the aim was to use what is known in regard to the mechanism of action of GroEL to design potential algorithms to accelerate folding in atomistic simulations. In addition, although the inclusion of the cage in the simulation and the cycling of the nature of the surface do clearly facilitate folding, the sample size is small and the accessible timescales limited. This means that it is not possible to quantify the relative contributions of the different effects. Nevertheless, the simulations do support the proposal that confinement plays a major role in facilitating protein folding within GroEL. The effect was primarily observed using a purely repulsive surface potential. Compared to the weakly polar surface, a purely repulsive potential will reduce the effective size of the cavity. The dimensions of the cage used in the simulations were chosen such that the size of the cavity was similar to that of the GroEL/ES complex. The test proteins are, however, significantly smaller than the average substrate of GroEL. We would predict that the effect of confinement on folding within the GroEL/ES cavity would only be significant for proteins of more than 100 amino acids.

The work also suggests that folding is enhanced by a preference for specific elements of secondary structure during the initial binding of the unfolded state. In the simulations, the formation of β-sheet was favored. Experimental evidence suggests that within GroEL peptides adopt a helical conformation on initial binding (Landry and Gierasch 1991; Landry et al. 1992). It is therefore interesting to note that most protein substrates known to interact with GroEL in vivo have an αβ structural motif (Houry et al. 1999).

Conclusions

In this study, we have attempted to refine misfolded protein structures by mimicking some aspects of the folding cavity formed by the chaperone GroEL/ES in MD simulations. Models of proteins generated using the semi ab initio structure prediction algorithm ROSETTA, were used to represent partially folded proteins. The GroEL/ES assembly was mimicked by a spherical cage just large enough to contain the protein substrates. The substrates were subjected to multiple cycles of binding and release. The effect of spatial confinement during refolding was also investigated. It has been shown that repeated cycles of binding and release significantly facilitated refolding. The effect was improved using a combination of binding and release together with spatial confinement during the refolding. Two of the three test models showed dramatic improvements in their overall fold, with the final structures being much closer to the corresponding native states than the initial structures. The third protein 1afi proved very sensitive to the hydrophobic cavity with the degree of disruption during the unfolding stage being greater than the degree of refolding achieved during 10 nsec of simulation. Much longer refolding times would be needed for this system.

While promising, the work presented is primarily a proof of principle and must now be tested on a wider set of proteins. In addition, we have shown that the use of a spherical cage with a smooth surface was highly disruptive to α-helical elements. The specific cage used in the current study favored the formation of planar structural elements such as β-sheet and thus may be appropriate only for β-structures. Alternative surfaces that also promote the formation of α-helical structures are needed. Nevertheless, the work clearly demonstrates that it is possible to greatly facilitate the productive refolding of proteins in MD simulations by using simple biologically inspired protocols. It particular, it suggests that the combination providing specific folding templates for helical or sheet structures or other approaches to avoid trapped states, together with confinement to bias the system toward compact, native-like structures, could provide a way to make folding simulations in near atomic detail more tractable. Finally, we would note that in this work the initial structures were very far from the native conformation. Changes in the global fold of the molecules were required and thus represent extremely difficult cases to test refinement protocols. Clearly, one would expect the approach to be most effective in cases where misfolding is more localized.

Materials and methods

The structural properties of the three proteins (PDB IDs 1vcc, 1sro, and 1afi) used in the current study are summarized in Table 1. The structure of 1vcc was solved by X-ray crystallography while the structures of 1sro and 1afi were solved by NMR (Sharma et al. 1994; Bycroft et al. 1997; Steele and Opella 1997). In the case of the NMR-derived structures, the first structure in the ensemble of structures deposited in the PDB was taken as the representative conformation. Each of these "native" structures has previously been shown to be stable in molecular dynamics (MD) simulations using the GROMOS96 force field (Fan and Mark 2003). For each protein, one "misfolded" structure was used to test the refinement protocols. The backbone RMSD of the secondary structure elements in the selected structures to the corresponding experimental structures are 0.54, 0.82, and 0.72 nm for 1vcc, 1sro, and 1afi, respectively. The misfolded structures were models generated by Baker and coworkers as part of a test of the ROSETTA algorithm (Simons et al. 1999, 2001; Bonneau et al. 2001) and were obtained from http://depts.washington.edu/bakerpg/decoys. The initial models only contained the peptide backbone and C_β coordinates. The coordinates of the side chains were generated using the WHATIF package (Vriend 1990). The

Table 1. *Proteins used in the simulations*

PDB	Description	Exp.	Resol.	N_{res}	N_{α}	N_{β}	N_{charge}
1vcc	Amino-terminal domain of vaccinia virus DNA topoisomerase I	X-ray	0.16	77	14	22	−1
1sro	S1 RNA binding domain from Pnpase	NMR	—	76	4	23	1
1afi	Reduced form of MERP, mercuric transport protein	NMR	—	72	22	22	3

PDB, PDB entry name.
Exp., experimental method for determining the protein structure.
Resol., structural resolution in nm for X-ray structures.
$N_{res,}$ the total number of residues in the protein.
$N_{\alpha,}$ the number of helical residues as defined in the PDB file.
$N_{\beta,}$ the number of strand residues as defined in the PDB file.
$N_{charge,}$ the net charge of the protein at pH 7.0.

models used in this study were the same as those used in our previous studies on protein refinement (Fan and Mark 2004a,b).

Simulation methodology

All simulations were performed using the GROMACS (Groningen Machine for Chemical Simulation) 3.0 package (Berendsen et al. 1995; Lindahl et al. 2001; van der Spoel et al. 2001) in conjunction with the GROMOS96 43a1 condensed-phase force field (van Gunsteren et al. 1996, 1998). The protonation state of ionizable groups was chosen appropriate to pH 7.0. The Simple Point Charge (SPC) model was used to describe the water (Berendsen et al. 1981). During the simulations, the temperature was maintained at 300 K using a Berendsen thermostat with a relaxation time of 0.1 psec (Berendsen et al. 1984). Nonbonded interactions were evaluated using a twin-range cut-off. Interactions within the shorter-range cut-off (0.9 nm) were calculated every step, whereas interactions between the shorter- and longer-range cut-off (1.4 nm) were updated every five steps, together with the pairlist. To reduce the effect of truncating the electrostatic interactions beyond the 1.4-nm cut-off, a reaction field (RF) correction was applied using a relative dielectric constant of 78 (Tironi et al. 1995). Covalent bond lengths within proteins were constrained using the LINCS algorithm (Hess et al. 1997). The geometry of the water molecules was constrained using SETTLE (Miyamoto and Kollman 1992). In addition, hydrogen atoms in the proteins were replaced by dummy interaction sites. This eliminates high-frequency degrees of freedom associated with the hydrogen atoms, allowing a time step of 4 fsec to be used to integrate the equations of motion (Feenstra et al. 1999).

Chaperone model and refinement protocols

Spherical cages of constant volume (diameter 6.0 nm) were used to mimic the chaperone cavity. These cages were in turn simulated within a periodic rhombic dodecahedron box. The minimum distance between the cage and the box wall was 0.4 nm. Six passages of 1.5 nm radius in the cage surface allowed water to circulate freely in the system (Fig. 6). Three functional groups were used to form the cages: a methylene (CH2) group (hydrophobic surface), an amide (NH + CO) group (hydrophilic surface) and a group termed CY, which differs from the CH2 group in the GROMOS force field in that it only has repulsive van der Waals interactions with other particles. To mimic the action of GroEL/ES, the ROSETTA model of each protein was initially placed in the center of the cavity. Three different refinement protocols were tested. Each protocol was composed of two stages. In the first stage, the protein was placed at the center of a hydrophobic cavity formed by a monolayer of closely packed CH2 particles (minimum distance 0.15 nm) in which 12 amide groups were embedded. The three protocols differed in the second stage. In protocol I, the wall of the cage was made more hydrophilic by embedding 18 amide groups (Fig. 6, middle) into the surface and by removing 75% of the CH2 particles. In protocol II, the cage was constructed of loosely packed, repulsive CY particles (minimum distance between particles was 0.8 nm). In protocol III, the cage was removed and the proteins were simulated in water. Each protein was subjected to 10 cycles of the above three protocols, of which the first stage was 5 nsec and the second stage was 10 nsec in length, making each simulation 150 nsec in total.

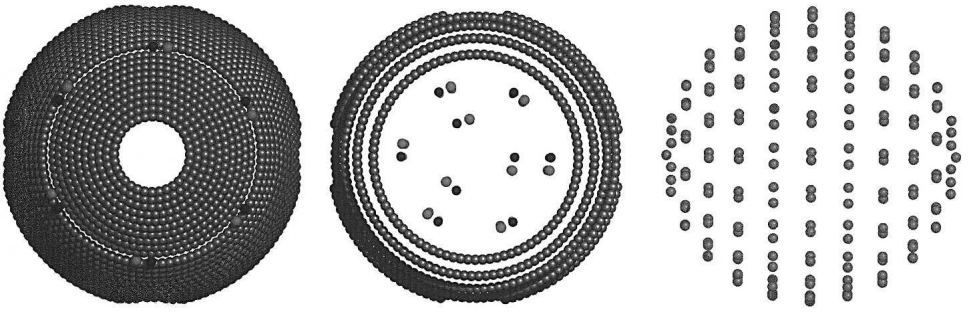

Figure 6. (From *left* to *right*) The relatively hydrophobic cage formed by a monolayer of densely packed CH2 particles incorporating only a small number of polar amide groups, the relatively hydrophilic cage consisting of a loosely packed hydrophobic surface and incorporating more amide groups, the repulsive cage formed from loosely packed repulsive particles.

References

Baumketner, A., Jewett, A., and Shea, J.E. 2003. Effects of confinement in chaperonin assisted protein folding: Rate enhancement by decreasing the roughness of the folding energy landscape. *J. Mol. Biol.* **332:** 701–713.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., and Hermans, J. 1981. Interaction models for water in relation to protein hydration. In *Intermolecular forces* (ed. B. Pullman), pp. 331–342. Reidel, Dordrecht, The Netherlands.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Di Nola, A., and Haak, J.R. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81:** 3684–3690.

Berendsen, H.J.C., van der Spoel, D., and van Drunen, R. 1995. GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* **91:** 43–56.

Betancourt, M.R. and Thirumalai, D. 1999. Exploring the kinetic requirements for enhancement of protein folding rates in the GroEL cavity. *J. Mol. Biol.* **287:** 627–644.

Bonneau, R., Strauss, C.E.M., and Baker, D. 2001. Improving the performance of Rosetta using multiple sequence alignment information and global measures of hydrophobic core formation. *Proteins* **43:** 1–11.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L., and Sigler, P.B. 1994. The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* **371:** 578–586.

Brinker, A., Pfeifer, G., Kerner, M.J., Naylor, D.J., Hartl, F.U., and Hayer-Hartl, M. 2001. Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* **107:** 223–233.

Bukau, B. and Horwich, A.L. 1998. The Hsp70 and Hsp60 chaperone machines. *Cell* **92:** 351–366.

Bycroft, M., Hubbard, T.J., Proctor, M., Freund, S.M., and Murzin, A.G. 1997. The solution structure of the S1 RNA binding domain: A member of an ancient nucleic acid-binding fold. *Cell* **88:** 235–242.

Chan, H.S. and Dill, K.A. 1996. A simple model of chaperone-mediated protein folding. *Proteins* **24:** 345–351.

Fan, H. and Mark, A.E. 2003. Relative stability of protein structures determined by x-ray crystallography or NMR spectroscopy: A molecular dynamics simulation study. *Proteins* **53:** 111–120.

———. 2004a. Mimicking the action of folding chaperones in molecular dynamics simulations: Application to the refinement of homology based protein structures. *Protein Sci.* **13:** 992–999.

———. 2004b. Refinement of homology based protein structures by molecular dynamics simulation techniques. *Protein Sci.* **13:** 211–220.

Feenstra, K.A., Hess, B., and Berendsen, H.J.C. 1999. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comp. Chem.* **20:** 786–798.

Fenton, W.A., Kashl, Y., Kurtak, K., and Horwich, A.L. 1994. Residues in chaperonin GroEL required for polypeptide binding and release. *Nature* **371:** 614–619.

Hess, B., Bekker, H., Berendsen, H.J.C., and Fraaije, J.G.E.M. 1997. LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* **18:** 1463–1472.

Hlodan, R., Tempst, P., and Hartl, F.U. 1995. Binding of defined regions of a polypeptide to GroEL and its implications for chaperonin-mediated protein-folding. *Nat. Struct. Biol.* **2:** 587–595.

Houry, W.A., Frishman, D., Eckerskorn, C., Lottspeich, F., and Hartl, F.U. 1999. Identification of in vivo substrates of the chaperonin GroEL. *Nature* **402:** 147–154.

Itzhaki, L.S., Otzen, D.E., and Fersht, A.R. 1995. Nature and consequences of GroEL-protein interactions. *Biochemistry* **34:** 14581–14587.

Jewett, A.I., Baumketner, A., and Shea, J.E. 2004. Accelerated folding in the weak hydrophobic environment of a chaperonin cavity: Creation of an alternate fast folding pathway. *Proc. Natl. Acad. Sci.* **101:** 13192–13197.

Kim, S., Schilke, B., Craig, E.A., and Horwich, A.L. 1998. Folding in vivo of a newly translated yeast cytosolic enzyme is mediated by the SSA class of cytosolic yeast Hsp70 proteins. *Proc. Natl. Acad. Sci.* **95:** 12860–12865.

Klimov, D.K., Newfield, D., and Thirumalai, D. 2002. Simulations of β-hairpin folding confined to spherical pores using distributed computing. *Proc. Natl. Acad. Sci.* **99:** 8019–8024.

Landry, S.J. and Gierasch, L.M. 1991. The chaperonin GroEL binds a polypeptide in an α-helical conformation. *Biochemistry* **30:** 7359–7362.

Landry, S.J., Jordan, R., McMacken, R., and Gierasch, L.M. 1992. Different conformations for the same polypeptide bound to chaperones DnaK and GroEL. *Nature* **355:** 455–457.

Lin, Z., Schwarz, F.P., and Eisenstein, E. 1995. The hydrophobic nature of GroEL-substrate binding. *J. Biol. Chem.* **270:** 1011–1014.

Lindahl, E., Hess, B., and van der Spoel, D. 2001. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.* **7:** 306–317.

Mayhew, M., da Silva, A.C.R., Martin, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F.U. 1996. Protein folding in the central cavity of the GroEL-GroES chaperonin complex. *Nature* **379:** 420–426.

Miyamoto, S. and Kollman, P.A. 1992. SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comp. Chem.* **13:** 952–962.

Ranson, N.A., Dunster, N.J., Burston, S.G., and Clarke, A.R. 1995. Chaperonins can catalyse the reversal of early aggregation steps when a protein misfolds. *J. Mol. Biol.* **250:** 581–586.

Sfatos, C.D., Gutin, A.M., Abkevich, V.I., and Shakhnovich, E.I. 1996. Simulations of chaperone-assisted folding. *Biochemistry* **35:** 334–339.

Sharma, A., Hanai, R., and Mondragon, A. 1994. Crystal structure of the amino-terminal fragment of vaccinia virus DNA topoisomerase I at 1.6 Å resolution. *Structure* **2:** 767–777.

Shtilerman, M., Lorimer, G.H., and Englander, S.W. 1999. Chaperonin function: Folding by forced unfolding. *Science* **284:** 822–825.

Simons, K.T., Bonneau, R., Ruczinski, I., and Baker, D. 1999. Ab initio protein structure prediction of CASP III targets using ROSETTA. *Proteins* **37 Suppl 3:** 171–176.

Simons, K.T., Strauss, C., and Baker, D. 2001. Prospects for ab initio protein structural genomics. *J. Mol. Biol.* **306:** 1191–1199.

Stan, G., Thirumalai, D., Lorimer, G.H., and Brooks, B.R. 2003. Annealing function of GroEL: Structural and bioinformatics analysis. *Biophys. Chem.* **100:** 453–467.

Steele, R.A. and Opella, S.J. 1997. Structures of the reduced and mercury-bound forms of MerP, the periplasmic protein from the bacterial mercury detoxification system. *Biochemistry* **36:** 6885–6895.

Takagi, F., Koga, N., and Takada, S. 2003. How protein thermodynamics and folding mechanisms are altered by the chaperonin cage: Molecular simulations. *Proc. Natl. Acad. Sci.* **100:** 11367–11372.

Thirumalai, D. and Lorimer, G.H. 2001. Chaperonin-mediated protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **30:** 245–269.

Tironi, I.G., Sperb, R., Smith, P.E., and van Gunsteren, W.F. 1995. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102:** 5451–5459.

Todd, M.J., Lorimer, G.H., and Thirumalai, D. 1996. Chaperonin-facilitated protein folding: Optimization of rate and yield by an iterative annealing mechanism. *Proc. Natl. Acad. Sci.* **93:** 4030–4035.

van der Spoel, D., van Buuren, A.R., Apol, E., Meulenhoff, P.J., Tieleman, D.P., Sijbers, A.L.T.M., Hess, B., Feenstra, K.A., Lindahl, E., van Drunen, R., et al. 2001. *Gromacs user manual, version 3.0*. University of Groningen, Groningen, The Netherlands. Internet: http://www.gromacs.org.

van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P.K.H.C., Mark, A.E., Scott, W.R.P., and Tironi, I.G. 1996. *Biomolecular simulation: The GROMOS96 manual and user guide*. vdf Hochschulverlag AG, Zürich, Switzerland.

van Gunsteren, W.F., Daura, X., and Mark, A.E. 1998. GROMOS force field. In *Encyclopedia of computational chemistry* (eds. von Ragué Schleyer et al.), Vol. 2, pp. 1211–1216. Wiley, New York.

Vriend, G. 1990. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.* **8:** 52–56.

Wang, J. and Boisvert, D.C. 2003. Structural basis for GroEL-assisted protein folding from the crystal structure of $(GroEL\text{-}KMgATP)_{14}$ at 2.0 Å resolution. *J. Mol. Biol.* **327:** 843–855.

Wang, J. and Chen, L. 2003. Domain motions in GroEL upon binding of an oligopeptide. *J. Mol. Biol.* **334:** 489–499.

Weissman, J.S., Hohl, C.M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H.R., Fenton, W.A., and Horwich, A.L. 1995. Mechanism of GroEL action: Productive release of polypeptide from a sequestered position under GroES. *Cell* **83:** 577–587.

Xu, Z., Horwich, A.L., and Sigler, P.B. 1997. The crystal structure of the asymmetric GroEL-GroES-$(ADP)_7$ chaperonin complex. *Nature* **388:** 741–749.

Zahn, R., Spitzfaden, C., Ottiger, M., Wuthrich, K., and Pluckthun, A. 1994. Destabilization of the complete protein secondary structure on binding to the chaperone GroEL. *Nature* **368:** 261–265.

Zahn, R., Perrett, S., and Fersht, A.R. 1996. Conformational states bound by the molecular chaperones GroEL and SecB: A hidden unfolding (annealing) activity. *J. Mol. Biol.* **261:** 43–61.