JCTC Journal of Chemical Theory and Computation

Ab Initio Protein Folding Using a Cooperative Swarm of Molecular Dynamics Trajectories

Neil J. Bruce and Richard A. Bryce*

School of Pharmacy and Pharmaceutical Sciences, University of Manchester, Oxford Road, Manchester, M13 9PT, United Kingdom

Received February 1, 2010

Abstract: The use of atomistic simulation techniques to directly resolve the protein tertiary structure from the primary amino acid sequence is hindered by the rough topology of the protein free energy surface and the resulting simulation time scales required. We explore here the use of a molecular dynamics technique based on swarm intelligence to identify the native states of two peptides and a Trp-cage miniprotein. In all cases, the presence of cooperative swarm interactions significantly enhanced the efficiency of molecular dynamics simulations in predicting the native conformation.

1. Introduction

Despite the vast conformational space available to them, globular proteins are able to fold rapidly to their unique native geometries.[1] Accurate prediction of these folded states, using only primary sequence information, is an important goal in pharmaceutical science, as knowledge of the structure of protein targets is an important step in rational drug design. Knowledge-based[2] and coarse-graining methods[3] have had some success in predicting protein structure. There are associated drawbacks, however, for example, a dependence on available geometrically similar structures in the former approach and a lack atomic resolution for the latter, which may lead to omission of important structural features.

Atomistic simulations are ideally placed to provide direct molecular level insights into the structure and dynamics of proteins. However, such approaches are hindered by the simulation time scales required to observe folding events. The free energy surface upon which protein simulations operate is rugged and characterized by a broad range of barriers, at scales both lower and higher than that of thermal energy kT. Advanced simulation methods[4−8] seek to increase the rate at which these barriers are traversed, while maintaining the representative features of the free energy surface; these methods offer the potential to increase the rate at which folding events occur during simulation, allowing the study of larger systems with more complex folding mechanisms.

One route to enhanced sampling that has proved successful in other areas of computational chemistry involves artificial intelligence methods. Genetic algorithms have had a major impact as conformational search tools in protein−ligand docking.[9,10] An alternative class of artificial intelligence methodology is based on *swarm intelligence*, the emergent behavior observed in nature when social animals, such as swarming insects, flocking birds, or schooling fish, act together cooperatively. In groups, the animals are able to show a greater searching efficiency than they would when acting individually. A concept originally applied computationally in 1989 to cellular robotic systems,[11] the swarm behavior is modeled by a set of simple rules which describe how individual agents act cooperatively within the system.

A recent swarm intelligence approach based on the behavior of ant colonies has proved successful in guiding molecular docking[12] and loop refinement.[13] Another swarm intelligence approach, particle swarm optimization,[14] has been used in the development of QSAR models[15] and molecular docking.[16,17] For the latter, the swarm algorithm exhibited significant improvements in the RMSD of pose for 37 protein−ligand complexes, when compared to GOLD (Darwinian genetic algorithm), AutoDock (Larmakian genetic algorithm), and the commonly used FlexX and DOCK methods.[16] The swarm intelligence approach has been extended to molecular dynamics simulations using model potentials[18] but to our knowledge has yet to be applied to protein folding problems. In this letter, therefore, we explore the use of a swarm algorithm to guide protein structure prediction via multicopy molecular dynamics (MD). The method is applied to two peptides and a Trp-cage miniprotein.

2. Methodology

Our approach follows the SWARM-MD method of Huber and van Gunsteren,[18] using a swarm of replica simulations that interact to cooperatively search phase space. This cooperation occurs through the incorporation of a swarm potential into the dynamics of each replica, which acts to drive each swarm member toward the mean trajectory of the entire swarm. The swarm potential is given as

$$V(\{\phi^j\})_{\text{swarm}} = \sum_{j=1}^{M} A \exp[-Bd_{\text{rms}}(\phi^j)] \qquad (1)$$

* Corresponding author tel.: (0)161-275-8345, fax: (0)161-275-2481, e-mail: R.A.Bryce@manchester.ac.uk.

10.1021/ct100060t
Published on Web 06/28/2010

where $d_{rms}(\phi^j)$ is the root-mean-square distance of swarm member j from the average location of the swarm (in dihedral space), given by

$$d_{rms}(\phi^j) = \left(\frac{1}{N} \sum_{i=1}^{N} (\phi_i^j - \langle \phi_i \rangle_{swarm})^2 \right)^{1/2} \quad (2)$$

and ϕ_i^j is dihedral angle i in swarm member j, $\langle \phi_i \rangle_{swarm}$ is the average value of dihedral angle i over M members of the swarm, and N is the number of dihedral angles used in calculation of $d_{rms}(\phi^j)$. A and B are parameters that we take to be the same for each swarm member. A defines the maximum strength of the swarm potential, while B defines the range over which it acts. The total dynamics of the system are therefore described by two sets of parameters: the social parameters, described by eq 1 above, which result in the cooperation between members of the swarm, and the cognitive parameters that describe the interatomic potentials of the molecular mechanics force field, which act independently on each swarm member.

We have implemented the SWARM-MD algorithm into a modified version of Amber 9.[19] Details of implementation (in particular, our treatment of $\langle \phi_i \rangle_{swarm}$) and associated parameters are given in the Supporting Information. Following Huber and van Gunsteren,[18] we take range B to be 0.8 rad^{-1} and explore specific values of strength A, as discussed below. On the basis of trial calculations for AEK17 in an implicit solvent, we find a swarm size of 16−20 as optimal (Supporting Information); therefore, in this study, we employ 20-replica swarms. The method is applied to the simulated annealing of polyalanine, AEK17, and Trp-cage. Computational details of these annealing simulations, including initial structure generation and subsequent secondary structure analysis, are also provided in the Supporting Information.

3. Results and Discussion

We evaluate the ability of swarm-enhanced MD to fold two small peptides, polyalanine [Ac−(Ala)$_{11}$−NH$_2$] and AEK17 [Ac−Ala−(Glu−Ala−Ala−Ala−Lys)$_3$−Ala−NH$_2$], into their known α-helical structures.[20,21] We also consider the folding of the Trp-cage miniprotein in an implicit solvent. In each of the three systems, we compare the results of a 20-replica swarm-enhanced MD simulation to those of 20 independent MD simulations. We consider each test case in turn.

3.1. Polyalanine. Starting from extended conformations, simulated annealing of gas-phase polyalanine was performed over 1.2 ns. In the absence of a swarm potential (A = 0.0 kcal/mol) during annealing, most of the 20 independent polyalanine replicas become kinetically trapped in nonhelical conformations, with only five replicas reaching a fully helical conformation. This results in a final average helicity of 29% (Figure 1a). Note that polyalanine simulation replicas are described as folding to a completely helical conformation if all nonterminal amino acid residues are assigned as helical in the final annealed conformation (equivalent to a helical content greater than 82%).

Using a swarm potential of strength A = −50.0 kcal/mol, folding performance is significantly improved: the final average helicity across polyalanine replicas is 75% (Figure 1a). Fifteen of the 20 swarm members anneal to a fully helical structure by the end of the simulation. A stronger swarm potential (A = −100 kcal/mol) was also applied in a separate multicopy MD simulation and found to further improve performance. In this case, all replicas adopted a fully helical conformation, resulting in a final average helicity of 91% (Figure 1a).

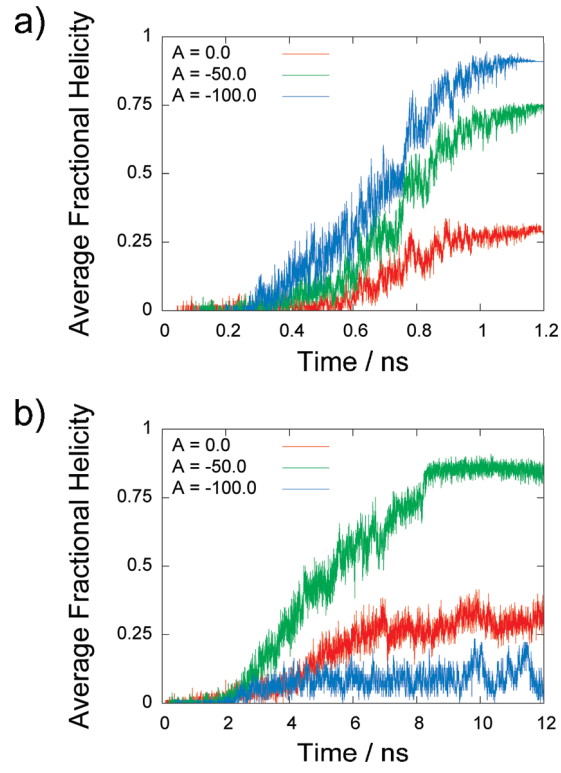

Figure 1. Average fractional helicity across 20 simulation replicas for swarm potentials of strength A = −50.0 (green) and −100.0 kcal/mol (blue) and in the absence of the swarm potential (A = 0.0 kcal/mol, red) during simulated annealing of (a) gas-phase polyalanine and (b) AEK17 in aqueous solution.

3.2. AEK17. Second, we assess the ability of swarm-enhanced MD to fold the 17-mer AEK17 peptide, with its zwitterionic Glu−Ala−Ala−Ala−Lys repeat sequence, into an α helix. The peptide was modeled in an aqueous generalized Born (GB) solvent,[22] and simulated annealing from extended conformations was conducted over 12 ns. AEK17 replicas are described as having successfully folded if they exhibit greater than 80% average helicity over the final 2 ns of constant temperature dynamics.

With no swarm potential applied, the final 20 annealed structures of AEK17 have a final average helicity of 34% (Figure 1b); only one simulation replica reaches a folded helical conformation. In the presence of a swarm potential (A = −50.0 kcal/mol), 17 swarm replicas fold to completely helical conformations, resulting in final average helicity of 82% (Figure 1b). Interestingly, this agrees well with the experimentally observed value of ∼80%, measured by circular dichroism spectroscopy at 274 K and pH 7 in 0.01 M NaCl.[21]

A stronger swarm potential was also applied (A = −100 kcal/mol), but found in this case to produce a negative effect on the observed folding rates. None of the simulation replicas was able to fold to a helical conformation during the 12 ns of simulation, resulting in an average final helicity of only 9% (Figure 1b). One possible explanation for the difference between this result and that of polyalanine, where the stronger potential was more

effective, may be found in the differing nature of the vacuum and aqueous potential energy surfaces. In a vacuum, unshielded interatomic Coulombic forces may be experienced at higher temperatures during annealing, leading to early prefolded nucleation steps in polyalanine. In contrast, in an aqueous environment, these forces are dampened; this is evidenced in independent MD simulations, where (un)folding transitions at ambient temperatures were found to occur more frequently in implicit aqueous solvent than *in vacuo*. As the temperature is lowered during annealing, the social swarm forces may dominate before cognitive MM forces in individual replicas sufficiently initiate physical folding processes. This results in the swarm force biasing the swarm members toward non-native high-energy minima, leading to premature convergence of the system. This observation underlines the importance of balancing social and cognitive potentials acting on the system.

3.3. Trp-cage. A more challenging model of protein folding is provided by the Trp-cage miniprotein (sequence: N_{20}LYIQWLKDGGPSSGRPPPS$_{39}$).[23] This small fast-folding (~4.1 μs)[24] 20-residue protein forms a globular structure in solution and consists of an α helix (residues 20−28), a short 3_{10} helix (residues 30−33), and a polyproline II helix (residues 36−38). The small size, and fast folding nature, of Trp-cage makes it an ideal test case to bridge the gap between studying small peptide systems and larger proteins. Indeed, it has already been the subject of much attention in protein folding and structure prediction studies, both through molecular dynamics simulation[25−32] and nondynamical optimization.[33−39]

Starting from extended conformations, a 20-replica 40 ns swarm-enhanced MD simulation of Trp-cage in GB solvent was performed. We adopt the strength of swarm potential ($A = -50$ kcal/mol) that gave the best folding performance for AEK17 in an implicit solvent. As with polyalanine and AEK17, we compare the performance of this simulation to that of 20 independent 40 ns molecular dynamics simulations.

For each simulation replica, the average root-mean-square deviation (RMSD) of both the backbone atoms and the heavy atoms, relative to the native NMR structure, was calculated, over the final 5 ns of simulation. In the absence of a swarm potential, the 20 independent replicas folded to conformations with an average backbone RMSD across all 20 simulations of 3.3 Å (Figure 2a) and an average heavy atom RMSD of 4.8 Å (Figure 2b). With the swarm potential applied, these values improved to 1.6 and 2.6 Å, respectively (Figure 2a and b). More specifically, of the 20 independent MD simulations, none was able to fold to within an average backbone RMSD of 1.5 Å (Figure 2a). In comparison, of the 20 swarm replicas, 16 folded to a backbone RMSD below 1.5 Å (Figure 2a). The lowest average backbone RMSD of any individual swarm member was 1.3 Å, significantly lower than the best performing independent replica, which had an average RMSD of 1.8 Å. The 16 folded swarm replicas all display an average heavy atom RMSD over the final 5 ns of below 2.3 Å (Figure 2b). Again, this is a significant improvement over nonswarm simulations, where the replica with the lowest average heavy atom RMSD over the final 5 ns differed from the NMR structure by 2.8 Å (Figure 2b).

The largest atomic deviations of the folded swarm members from the NMR structure of Trp-cage occur in its N- and

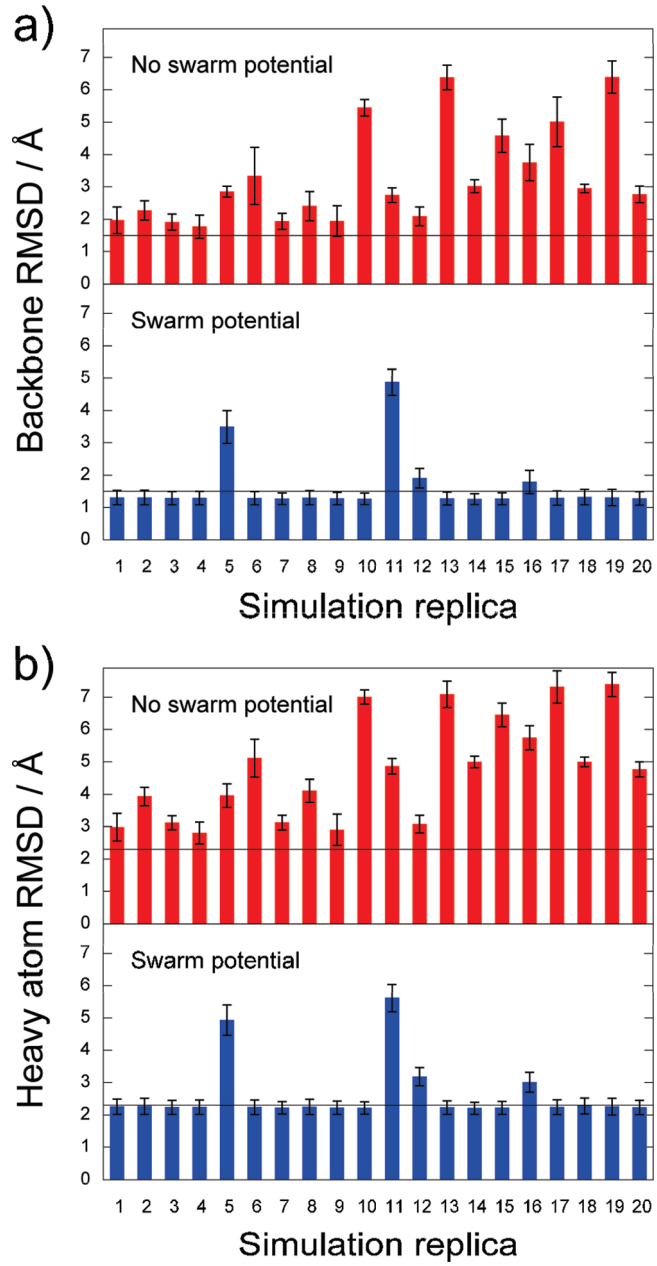

Figure 2. Mean RMS deviations of (a) backbone atom positions and (b) heavy atom positions from the native NMR structure of Trp-cage (conformer 11 from PDB code 1L2Y, which lies closest to average of NMR ensemble) during the final 5 ns of simulation for each simulation replica. Red indicates the absence of swarm potential; blue is with swarm potential present ($A = -50$ kcal/mol). Standard deviations during the final 5 ns of trajectories are shown as error bars.

C-terminal regions and the coil−3_{10} helix−coil region (residues 29−35). These deviations are caused by the presence in the simulated structures of two salt bridges: between the terminal ammonium and carboxylate groups and between the side chains of Asp28 and Arg35. The terminal salt bridge is not present in any of the 38 NMR structures, while the Asp28−Arg35 interaction is present in approximately half the NMR structures. The persistence of these two salt bridges in the simulations results in the chain ends of Trp-cage lying closer in the folded structures compared to NMR (Figures 3a and 3b). The presence

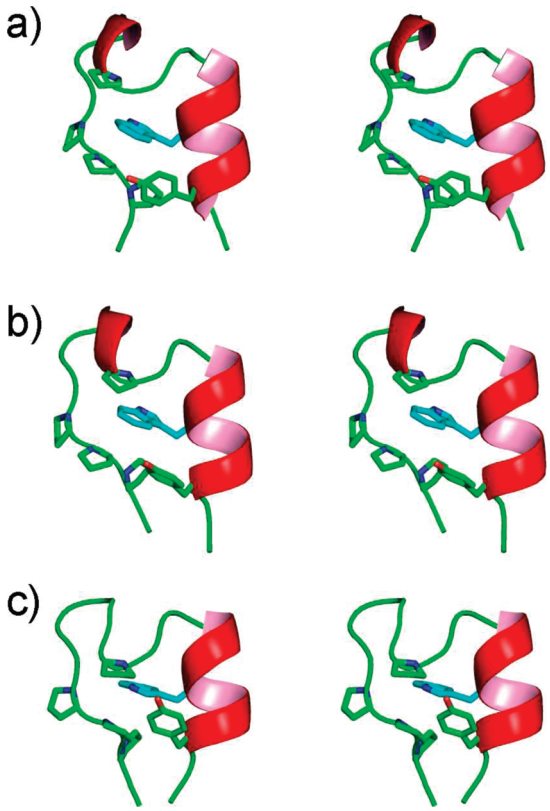

Figure 3. Stereoscopic image giving a comparison between the (a) Trp-cage NMR structure (PDB code 1L2Y; conformer 11), (b) best folded swarm member, and (c) a near-folded conformation with incorrect Trp25 orientation. The side chain of Trp25 is shown in blue; the side chains of hydrophobic cage residues Pro31, Pro36, Pro37, Pro38, and Tyr22 are shown in green.

of these interactions in 15 and 16 of the 16 folded swarm members suggests an overestimation of salt bridge strength, perhaps due to the force field−implicit solvent model combination (a known problem[30,40]), which may be hindering further improvement in the heavy atom RMSD of the folded Trp-cage structures.

Four swarm members did not fold to within a backbone RMSD of 1.5 Å (Figure 2). Two swarm replicas of Trp-cage have a fairly high backbone RMSD (>3 Å) and contain only a partial native secondary structure: in both cases, the polyproline II helix was present. In one, the α helix was fully formed, while in the other, the α helix was partially formed but the 3_{10} helix was present. It is possible that these replicas represent intermediate structures in the folding mechanism of Trp-cage. The remaining two Trp-cage replicas have a backbone RMSD of less than 2.0 Å, corresponding to near-native states. However, the Trp25 side chain displays an incorrect orientation, pointing away from the cage (Figure 3c). It has been suggested[25,26] that the rate-limiting step of Trp-cage folding is the incorporation of the Trp25 residue into the hydrophobic cage, followed by the formation of the residue's native contacts (Figure 3a). The resulting loss of degrees of freedom in this residue produces an entropic barrier in the free energy surface that must be overcome for the simulations to attain a native conformation.

It appears that the swarm-based simulations effectively lower this barrier: of the 18 swarm-enhanced Trp-cage replicas that

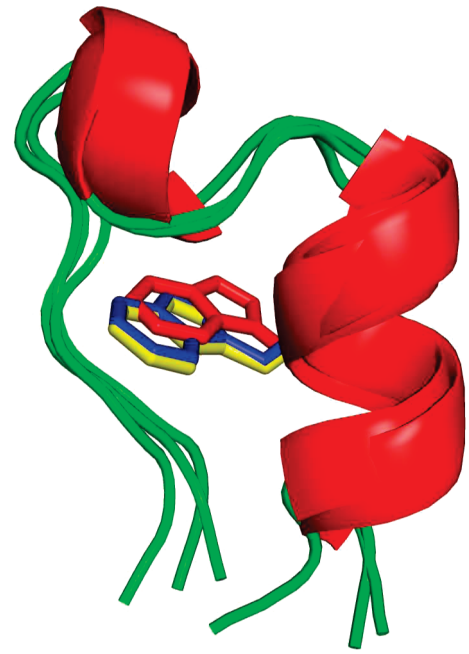

Figure 4. Overlay of NMR-derived Trp-cage and swarm-folded Trp-cage with the least NOE violations: Trp25 in NMR structure (blue) and best swarm structure before (red) and after refinement (yellow).

folded to a native backbone conformation (all with an average backbone RMSD < 2.0 Å), 16 (89%) displayed a correctly packed Trp25, pointing into the folded hydrophobic cage (Figure 3b), with just two showing the incorrect Trp25 geometry (Figure 3c). By contrast, only four of the nine (44%) independent MD replicas that produced a folded, or partially folded, hydrophobic cage (defined by an average backbone RMSD over residues 22−38 of less than 2.0 Å) displayed the correct Trp25 orientation. The addition of the cooperative swarm potential into the dynamics of the system therefore guides the simulations over the energy barrier associated with the entry of Trp25 into the cage, driving them toward the native state.

Interestingly, because of the swarm potential, the presence of a few replicas with mispacked Trp25 appears to slightly influence the orientation of that amino acid in the other replicas. While Trp25 packs correctly in these 16 replicas, the residue is subtly displaced relative to the NMR structure (Figure 4). This is clearly seen by direct comparison with the 168 NOE distance constraints used to derive the native structure ensemble of Trp-cage. In overall terms, the swarm-enhanced simulations violate only 25% of the NOE constraints, as compared to 34% from the independent simulations. However, the amino acid with the highest number of experimentally observed constraints (49 of the 168) is the core Trp25 residue. Consequently, because Trp25 tends to sit toward the back of the cage in the swarm, it is a replica from the independent MD, not swarm MD, that has the least NOE violations, with a value of 18% (replica 4, Figure 2). However, if the influence of the swarm on its replicas is tapered to zero in a subsequent refinement step, the Trp25 residue is able to assume its correct position, such that the best folded swarm member shows an NOE violation of only 14%. (This was achieved by annealing parameter A from −50.0 to

0.0 kcal/mol over 3 ns, followed by 2 ns of simulation in the absence of swarm potential.)

4. Conclusions

For polyalanine, AEK17, and Trp-cage, the incorporation of the swarm intelligence potential into the simulation dynamics is found to increase the folding performance of MD simulations. The cooperative nature of the swarm protocol prevents the swarm members from becoming trapped in local minima and helps drive the simulated structures toward the native state. Inhibition of AEK17 folding via the stronger swarm potential shows that a correct balance of social and cognitive factors is required for efficient conformational searching. However, this balance appeared suitably transferable between two implicitly solvated systems (AEK17 and Trp-cage).

For the study of Trp-cage folding, none of the 20 independent 40 ns molecular dynamics simulations anneal to a correctly folded conformation, i.e., to within a backbone RMSD of 1.5 Å of the native structure. By allowing the trajectories to interact cooperatively, 16 of the 20 trajectories adopt the native geometry within the last 5 ns of the simulations. A final refinement step, where the influence of the swarm potential was relaxed, was found to be useful in obtaining the detailed orientation of Trp25 in the miniprotein core. Both the swarm and independent MD calculations consume 800 ns of aggregate simulation time. To estimate the time scale required to correctly fold Trp-cage via unbiased molecular dynamics simulations at 300 K, we consider two recent generalized Born studies of the miniprotein: in the first study,[25] 77 independent 100 ns MD simulations of Trp-cage obtained only five conformations folded to within a backbone RMSD of 2.0 Å of the native state; a second MD study by Snow et al.[31] harvested in the region of 1000 trajectories of length 30 ns, less than 1% of which folded to below 2.6 Å RMSD of the native C_α structure. Simply considering computational cost per folded structure and recognizing that the SWARM-MD simulations also include the effect of annealing, the study of Snow et al. required $1-4\ \mu$s per folded structure, compared to 50 ns per folded structure via the swarm-enhanced calculation, an improvement of 30- to 80-fold. In terms of overall computing time such that misfolded structures are included, the swarm simulations are 10 to 40 times shorter.

Other methods have been introduced to enhance the sampling of phase space by molecular simulation methods, such as metadynamics,[41] locally enhanced sampling,[42] and replica-exchange schemes.[4] Several of these techniques have been applied to the folding of Trp-cage.[27−29,32] Most recently, a temperature REMD simulation[27] was performed on Trp-cage in generalized Born solvent using 16 40 ns replicates spanning 300−460 K; the 300 K trajectory latterly sampled mainly folded Trp-cage (within ∼2 Å heavy atom RMSD of the native structure). These calculation conditions and the method's performance are comparable to that of the SWARM-MD simulation of Trp-cage presented here. Hamiltonian-based replica exchange is also possible,[27,28,32] and these methods have proved particularly powerful, for example, obtaining folded Trp-cage structures using five[27] or six[32] replicas of sub-100 ns trajectories. An additional advantage of temperature-based and Hamiltonian-based replica exchange schemes is the generation of correct 300 K ensembles, providing information on folding pathways and intermediates. However, neither of these methods scale favorably with system size due to the constraint of the exchange condition: efficient exchange between neighboring replicas requires sufficient energy overlap between replicas. Therefore, the number of required replicas grows rapidly with the size of the simulation system, and correspondingly longer simulation times are required to allow efficient sampling of the temperature space. SWARM-MD does not incorporate an exchange move and therefore does not suffer from this exacting requirement. This difference may be of greater importance when extending the method to study explicitly solvated systems, where system size is greatly increased by the degrees of freedom of water molecules. As mentioned previously with regard to salt bridge stability in Trp-cage, the absence of an explicitly modeled solvent can affect the free energy minima predicted through simulation, due to an incorrect representation of charge shielding.[30,40] This overstabilization is not unique to Trp-cage.[43,44] Existing implicit solvent models have also been found to incorrectly predict the secondary structure preferences of peptides[45] and overestimate melting temperatures.[29,46] It has also been suggested that the absence of explicit water molecules may incorrectly describe the effects of dewetting of solvent-exposed hydrophobic residues in folded protein structures[47] and the effect of structural water molecules on the folding landscape.[48] From these observations, and the SWARM-MD algorithm's suitability for distribution of replicas over parallel architectures, a cooperative swarm of trajectories appears well-placed to predict the conformations of larger systems, particularly systems involving the incorporation of explicitly modeled bulk solvent. However, for future applications to predictive folding of larger polypeptide structures and interaction of flexible protein receptors with ligands, it will be useful to explore optimization schemes which anneal the influence of the swarm, to prevent the undue biasing effects of outlier replicas.

Acknowledgment. This work was supported by the EPSRC.

Supporting Information Available: Derivation of atomic forces due to swarm potential. Detailed simulation protocols. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Dill, K. A.; Chan, H. S. *Nat. Struct. Biol.* **1997**, *4*, 10–19.

(2) Zhang, Y.; Skolnick, J. *Proc. Natl. Acad. Sci.* **2005**, *102*, 1029–1034.

(3) Nanias, M.; Chinchio, M.; Pillardy, J.; Ripoll, D. R.; Scheraga, H. A. *Proc. Natl. Acad. Sci.* **2003**, *100*, 1706–1710.

(4) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(5) Zagrovic, B.; Sorin, E. J.; Pande, V. *J. Mol. Biol.* **2001**, *313*, 151–169.

(6) Wu, X. W.; Brooks, B. R. *Biophys. J.* **2004**, *86*, 1946–1958.

(7) Yang, L.; Grubb, M. P.; Gao, Y. Q. *J. Chem. Phys.* **2007**, *126*, 125102–125107.

(8) Wolf, M. G.; de Leeuw, S. W. *Biophys. J.* **2008**, *94*, 3742–3747.

(9) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727–748.

(10) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639–1662.

(11) Beni, G. From swarm intelligence to swarm robotics. In *Swarm Robotics*; Sahin, E., Spears, W. M., Eds.; Springer-Verlag: Berlin, 2005; Vol. 3342, pp 1−9.

(12) Korb, O.; Stutzle, T.; Exner, T. E. PLANTS: Application of ant colony optimization to structure-based drug design. In *Ant Colony Optimization and Swarm Intelligence, Proceedings*; Dorigo, M., Gambardella, L. M., Birattari, M., Martinoli, A., Stutzle, T., Eds.; Springer-Verlag: Berlin, 2006; Vol. 4150, pp 247−258.

(13) Xiang, Z.; Soto, C. S.; Honig, B. *Proc. Natl. Acad. Sci.* **2002**, *99*, 7432–7437.

(14) Kennedy, J.; Eberhart, R. Particle swarm optimization. In *1995 IEEE International Conference on Neural Networks Proceedings*; IEEE: New York, 1995; Vols 1−6, pp 1942−1948.

(15) Cedeno, W.; Agrafiotis, D. K. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 255–263.

(16) Chen, H. M.; Liu, B. F.; Huang, H. L.; Hwang, S. F.; Ho, S. Y. *J. Comput. Chem.* **2007**, *28*, 612–623.

(17) Chen, K.; Li, T. H.; Cao, T. C. *Chemom. Intell. Lab. Syst.* **2006**, *82*, 248–259.

(18) Huber, T.; van Gunsteren, W. F. *J. Phys. Chem. A* **1998**, *102*, 5937–5943.

(19) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M. *AMBER 9*, University of California: San Francisco, 2006.

(20) Levy, Y.; Jortner, J.; Becker, O. M. *Proc. Natl. Acad. Sci.* **2001**, *98*, 2188–2193.

(21) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci.* **1987**, *84*, 8898–8902.

(22) Tsui, V.; Case, D. A. *Biopolymers* **2000**, *56*, 275–291.

(23) Neidigh, J. W. *Nat. Struct. Biol.* **2002**, *9*, 425.

(24) Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 12952–12953.

(25) Chowdhury, S.; Lee, M. C.; Duan, Y. *J. Phys. Chem. B* **2004**, *108*, 13855–13865.

(26) Chowdhury, S.; Lee, M. C.; Xiong, G.; Duan, Y. *J. Mol. Biol.* **2003**, *327*, 711–717.

(27) Kannan, S.; Zacharias, M. *Int. J. Mol. Sci.* **2009**, *10*, 1121–1137.

(28) Piana, S.; Laio, A. *J. Phys. Chem. B* **2007**, *111*, 4553–4559.

(29) Pitera, J. W.; Swope, W. *Proc. Natl. Acad. Sci.* **2003**, *100*, 7587–7592.

(30) Simmerling, C.; Strockbine, B.; Roitberg, A. E. *J. Am. Chem. Soc.* **2002**, *124*, 11258–11259.

(31) Snow, C. D.; Zagrovic, B.; Pande, V. S. *J. Am. Chem. Soc.* **2002**, *124*, 14548–14549.

(32) Son, W. J.; Jang, S.; Pak, Y.; Shin, S. *J. Chem. Phys.* **2007**, *126*, 5.

(33) Carnevali, P.; Tóth, G.; Toubassi, G.; Meshkat, S. N. *J. Am. Chem. Soc.* **2003**, *125*, 14244–14245.

(34) Schug, A.; Herges, T.; Verma, A.; Lee, K. H.; Wenzel, W. *ChemPhysChem* **2005**, *6*, 2640–2646.

(35) Schug, A.; Herges, T.; Wenzel, W. *Phys. Rev. Lett.* **2003**, *91*, 158102.

(36) Schug, A.; Wenzel, W. *Europhys. Lett.* **2004**, *67*, 307.

(37) Schug, A.; Wenzel, W.; Hansmann, U. H. E. *J. Chem. Phys.* **2005**, *122*, 7.

(38) Verma, A.; Schug, A.; Lee, K. H.; Wenzel, W. *J. Chem. Phys.* **2006**, 124.

(39) Verma, A.; Wenzel, W. *Biophys. J.* **2009**, *96*, 3483–3494.

(40) Geney, R.; Layten, M.; Gomperts, R.; Hornak, V.; Simmerling, C. *J. Chem. Theory Comput.* **2005**, *2*, 115–127.

(41) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562–12566.

(42) Simmerling, C.; Fox, T.; Kollman, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 5771–5782.

(43) Zhou, R.; Berne, B. J. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12777–12782.

(44) Zhou, R. H. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 148–161.

(45) Okur, A.; Wickstrom, L.; Layten, M.; Geney, R.; Song, K.; Hornak, V.; Simmerling, C. *J. Chem. Theory Comput.* **2006**, *2*, 420–433.

(46) Yeh, I. C.; Lee, M. S.; Olson, M. A. *J. Phys. Chem. B* **2008**, *112*, 15064–15073.

(47) Daidone, I.; Ulmschneider, M. B.; Di Nola, A.; Amadei, A.; Smith, J. C. *Proc. Natl. Acad. Sci.* **2007**, *104*, 15230–15235.

(48) Rhee, Y. M.; Sorin, E. J.; Jayachandran, G.; Lindahl, E.; Pande, V. S. *Proc. Natl. Acad. Sci.* **2004**, *101*, 6456–6461.

CT100060T