Fast In Silico Protein Folding by Introduction of Alternating Hydrogen Bond Potentials

M. G. Wolf and S. W. de Leeuw
Technical University of Delft, 2628 BL Delft, The Netherlands

ABSTRACT We accelerate protein folding in all-atom molecular dynamics simulations by introducing alternating hydrogen bond potentials as a supplement to the force field. The alternating hydrogen bond potentials result in accelerated hydrogen bond reordering, which leads to rapid formation of secondary structure elements. The method does not require knowledge of the native state but generates the potentials based on the development of the tertiary structure in the simulation. In protein folding, the formation of secondary structure elements, especially α-helix and β-sheet, is very important, and we show that our method can fold both efficiently and with great speed.

INTRODUCTION

Computational studies, and in particular MD simulations, are widely applied to study protein-folding processes, providing detailed insight at atomic resolution. Unfortunately, MD can sample a system only for a time period of nanoseconds up to a microsecond, whereas protein folding generally takes microseconds to milliseconds. Attempts to overcome this sampling problem involve the use of simplified models that require many assumptions (1,2) and consequently lack microscopic detail required to monitor the various interactions that lead to folding. Alternatively, large-scale distributed computing can be used (3–6). However this approach is not accessible to everyone, and with an explicit solvent, the sampling problem cannot be overcome completely.

Folding of a protein into the native state cannot be described by a random search through all the degrees of freedom but is believed to be a guided process (7). Expressed in terms of free-energy landscapes, a protein traversing the free-energy landscape is funneled from the high-energy unfolded conformations into the low-energy native state (8–10).

The free-energy landscape of a protein is determined by two major contributions: hydrophobic interactions and hydrogen bonding. Hydrophobic interactions joining the hydrophobic side chains (hydrophobic collapse) are generally viewed as one of the driving forces of protein folding (11). The role of hydrogen bonding is more contested because of the difficulty of quantifying hydrogen bond energies. Because the formation of hydrogen bonds between protein atoms results in the loss of hydrogen bonds formed with water, it is still unclear whether protein intramolecular hydrogen bond formation in an aqueous environment contributes favorably (12–14) or is negligible to the free energy (15,16). In any case, during the hydrophobic collapse, intramolecular hydrogen bond formation is necessary to compensate for the high free-energy cost associated with burying unsatisfied hydrogen bonding groups (17). Because the number of intramolecular hydrogen bond partners is limited, the necessity for these bonds limits the number of allowable protein conformations and thus primarily provides specificity (11,18).

The free-energy landscape of a protein is rugged, displaying many valleys and energy barriers (19,20). Consequently, a protein traversing this landscape is very likely to encounter valleys corresponding to local free-energy minima before reaching the valley corresponding to the native state. Because the valleys are generated by an accumulation of individual interactions (hydrophobic and hydrogen bonds), crossing a barrier requires some of these interactions to be disrupted. Even a very simple peptide such as penta-alanine possesses a rugged free-energy landscape, with many valleys of comparable free energy (21). Studies have shown that a barrier height of only $2\ k_BT$ is sufficient to prevent the folding of a protein to occur in a downhill manner (22). The free energy barrier associated with a single intramolecular hydrogen bond is $\sim 3\ k_BT$ in a protein in water (23–25), preventing straightforward folding.

Nature resolves the slow folding problem of proteins by using chaperones to assist in the folding of proteins inside cells. Chaperones prevent association with other proteins and limit the number of accessible conformations (26,27). In addition, they can actively stimulate proteins to (partially) unfold, allowing the protein to escape from a local free-energy minimum and have another attempt at folding. In this way transitions in the free-energy landscape are facilitated, and, thus, the folding rate is increased (28). Unfortunately, these systems are too large for efficient use in MD simulation and in silico folding experiments.

Here we propose a novel computational method based on the idea that occasional (partial) unfolding of a protein enhances the frequency of barrier crossing and the folding rate

Submitted September 19, 2007, and accepted for publication November 26, 2007.

Address reprint requests to M. G. Wolf, Technical University of Delft, Julianalaan 136, 2628 BL Delft, The Netherlands. E-mail: m.g.wolf@tudelft.nl.

Abbreviations used: MD, molecular dynamics; AHBP, alternating hydrogen bond potentials; RMSD, root mean-square deviation.

Editor: Ron Elber.

0006-3495/08/05/3742/06 $2.00

doi: 10.1529/biophysj.107.122192

of proteins. We perform MD simulations during which we periodically introduce temporary additional forces that alternately stimulate unfolding and folding. These forces act on the intramolecular hydrogen bonds. The first reason for this is that distinct hydrogen bonds in a similar context contribute equally to the free energy, but a free-energy barrier separates all the possible hydrogen bonds. In other words, hydrogen bonds provide kinetic stability both in the global minimum and in local minima rather than thermodynamic stability. This has important implications: Unfolding and folding can be stimulated by reimbursing the activation energy set by the kinetic barrier of a hydrogen bond. In addition, the hydrogen bonds provide specificity rather than stability with respect to the tertiary structure of a protein, which means that the interactions that provide thermodynamic stability are unaltered and still guide the folding process of the protein into its native state, and the time in free-energy minima is decreased. A second more technical reason for influencing the intramolecular hydrogen bonds is that the number of required additional forces is minimal. This is because the number of donor-acceptor pair combinations in a protein is limited, and the hydrogen bonds are orientation dependent (29), requiring introduction of only a few relevant hydrogen bond potentials.

The manipulation of the hydrogen bonds is performed within a single MD simulation, where alternating attractive or repulsive hydrogen bond potentials are introduced in addition to the standard force-field potentials. The repulsive potential destabilizes the hydrogen bonds and lifts the protein to a higher free-energy level. The attractive potential in turn facilitates hydrogen bond formation to enable rapid identification of the conformational regions of free-energy minima. Such a local unfolding/folding mechanism would be comparable with the barrier-crossing effect of a chaperone protein. In this method we do not need a priori information on the native state; rather, we use the structure of the protein as it develops during the simulation to determine which potentials are introduced.

We show that manipulation of hydrogen bonds during an MD simulation can accelerate the folding of a protein. The two secondary structure elements appearing most, α-helix and β-sheet, can be folded efficiently. This is demonstrated by the folding of a 16-residue polyalanine to the α-helical native state and the 16-residue C-terminal of the 1GB1 protein to the β-hairpin native state.

METHOD

AHBP

We introduce a hydrogen bond potential V_{hb} as a supplement to the standard force field, which acts on the atoms involved in hydrogen bonding to accelerate protein folding in MD simulations. This is implemented as a staged MD protocol, where we distinguish three stages, the repulsive stage (R), the attractive stage (A), and the relaxation stage (E). These three stages each treat hydrogen bonds differently. In R, a potential stimulates hydrogen bond breakage, in A, a potential facilitates hydrogen bond formation, and in E, the system is allowed to relax. In our simulations, each stage is active for 0.5 ps in the order $(\mathrm{R–E–A–E})_{\mathrm{n}}$.

When a stage is active, every 0.1 ps all intramolecular donor-acceptor pairs of the protein are evaluated. The relevant pairs are selected, and potentials are introduced that will result in a force acting on the atoms. During selection, we avoid targeting strong hydrogen bonds because they can be native and try to introduce only one potential per atom to minimize the manipulation of the system. Therefore, a pair is excluded from selection if 1), it is a strong hydrogen bond (characterized by a donor-acceptor distance <0.35 nm and a donor-hydrogen-acceptor angle larger than 120°), 2), the atoms of the pair are involved in another strong hydrogen bond, and 3), the atoms in the pair are already targeted in another hydrogen bond potential (e.g., from a previous evaluation). For the remaining donor-acceptor pairs, those with the largest hydrogen bond potential energy (Eq. 1) are selected, with the rule that the atoms in a pair may be selected only once. At the end of each stage, all hydrogen bond potentials are removed.

The hydrogen bond potential $V_{\mathrm{hb}}(q,t)$ is given in (1):

$$V_{\mathrm{hb}}(q,t) = fc(q,t) \times E_{\mathrm{d}}(q(t_{\mathrm{ev}})) \times E_{\theta}(q(t_{\mathrm{ev}})). \qquad (1)$$

It is a function of time t and consists of a distance potential $E_{\mathrm{d}}(q(t_{\mathrm{ev}}))$, an angle potential $E_{\theta}(q(t_{\mathrm{ev}}))$, a gradually changing force constant $fc(q,t)$, and the positions of the atoms in the hydrogen bonds q.

In the repulsive stage, the distance potential $E_{\mathrm{d}}(q(t_{\mathrm{ev}}))$ is determined by the distance d (nm) between donor and acceptor (Fig. 1) at the evaluation time t_{ev}. Cutoff distances d_{min} and d_{max} of 0.35 and 0.40 nm, respectively, are used. For the attractive stage, the distance between hydrogen and acceptor (Fig. 1) is considered, and the cutoff distances d_{min} and d_{max} are 0.23 and 0.40 nm, respectively. The values of the cutoff distances ensure that only weak to very weak hydrogen bonds are targeted, so that, in the repulsive stage, the hydrogen bond is pushed up the last part of the free-energy barrier, and in the attractive stage, the formation of hydrogen bonds that would possibly be formed in the future is accelerated:

$$E_{\mathrm{d}}(q(t_{\mathrm{ev}})) = \begin{cases} 1 & d(t_{\mathrm{ev}}) < d_{\mathrm{min}} \\ 1 - \dfrac{d(t_{\mathrm{ev}}) - d_{\mathrm{min}}}{d_{\mathrm{max}} - d_{\mathrm{min}}} & d_{\mathrm{min}} \le d(t_{\mathrm{ev}}) < d_{\mathrm{max}} \\ 0 & d_{\mathrm{max}} \le d(t_{\mathrm{ev}}) \end{cases}. \qquad (2)$$

The angle potential $E_{\theta}(q(t_{\mathrm{ev}}))$ depends on the angle θ (degrees) of the donor hydrogen acceptor (Fig. 1) at activation time t_{ev}. The cutoff angle θ_{bound} in the repulsive stage is set to 120°, which ensures targeting all weak hydrogen bonds, and in the attractive stage to 60°, allowing generation of many hydrogen bonds:

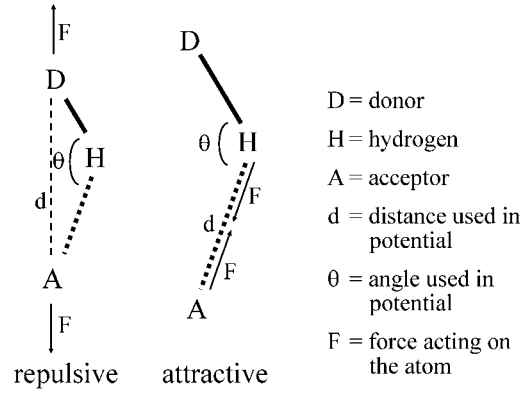

FIGURE 1 Donor-acceptor pair including the corresponding hydrogen. Schematic view of the relevant quantities in the hydrogen bond potential and the resulting forces.

$$E_\theta(q(t_{ev})) = \begin{cases} 1 & \theta(t_{ev}) \geq \theta_{bound} \\ 0 & \theta_{bound} > \theta(t_{ev}) \end{cases}. \quad (3)$$

The potential is slowly introduced and removed from the system to avoid cutoff effects at the boundaries of the distance and angle potential. This is achieved by growing or shrinking the force constant in 50 increments to a maximum force constant or zero, respectively. At each time step the counter n_{inc} is incremented by 1 if the derivative of the distance and angle potential is nonzero until the maximum value of 50 is reached, and 1 is subtracted when this derivative is zero until the counter is zero. The force constant is then given by Eq. 4, with $0 \leq n_{inc} \leq 50$. The use of 50 increments to reach the maximum force constant is chosen arbitrarily within the idea of gradually introducing the forces in the system to its maximum:

$$fc(q,t) = \frac{1}{50} \times n_{inc} \times fc_{max}. \quad (4)$$

To obtain the maximum force constant, several values were tested, and the values showing a good response, i.e., many unfolding and folding events, were used. The hydrogen bond potential leads to the introduction of the following force acting on the acceptor atom (Fig. 1):

$$F_A = fc(q,t) \times \begin{cases} \dfrac{1}{d_{max} - d_{min}} \dfrac{\mathbf{d}_{XA}(t_{ev})}{d_{XA}(t_{ev})} & d_{min} \leq d(t_{ev}) < d_{max}; \theta(t_{ev}) \geq \theta_{bound} \\ 0 & rest \end{cases}. \quad (5)$$

The balancing force is $F_X = -F_A$. In these equations, the X refers to the donor atom in the repulsive stage and to the hydrogen atom in the attractive stage (Fig. 1).

Introducing the AHBP potentials only at evaluation time as well as the gradual introduction/removal of the potential means that the energy is not conserved. The possible numerical instability is resolved in a constant-temperature ensemble by means of the heat bath.

Simulation protocol

All simulations were performed using the GROMACS (16) software package version 3.3.1 and the GROMOS96 43a1 force field (30) in combination with simple point charge water (31). A time step of 2 fs was used, with all bonds constrained using the LINCS algorithm (32). Van der Waals interactions were ignored outside a cutoff of 1.2 nm. Electrostatic interactions were treated with the particle mesh Ewald method applying a real-space cutoff of 0.9 nm (33). The system temperature was coupled to a Berendsen thermostat, and no pressure coupling was used.

The starting structures of the simulations both for the α-helix and for the β-hairpin simulations were extended conformations. The proteins were dissolved in a box of ~2000 simple point charge water to obtain a density of 0.99 kg/l, and initial velocities were generated randomly to create a system at 300 K. In the case of the α-helix, this system was equilibrated for 1 ns to allow the chain to collapse. For every production run, new random velocities corresponding to a system at 300 K were generated.

RESULTS

The method presented above aims to accelerate in silico protein folding. This is achieved by manipulating the intramolecular hydrogen bonds, leading to an increase in the number of barrier transitions. To show that this is indeed the case, the time behavior of a 16-residue polyalanine was examined with standard MD (four simulations of 30 ns) and with AHBP-MD (five simulations of 10 ns). The simulations were started from a collapsed coil, which represents a structure in a local minimum possessing many hydrogen bonds. The maximum force constants used in the MD simulation including AHBP were -600 kJ mol^{-1} nm^{-1} for the attractive potential and 450 kJ mol^{-1} nm^{-1} for the repulsive potential.

To test if the faster and broader sampling of the conformational space of a protein by the AHBP-MD simulations leads to rapid formation of secondary structure elements, two systems were tested. The polyalanine simulations used to show enhanced barrier crossing in AHBP-MD were also used to test the ability of the AHBP method to form an α-helical secondary structure. To test the β-sheet secondary structure formation, we investigated the folding of the 16-residue C-terminus of the protein G (Protein Data Bank code 1GB1), which adopts a β-hairpin conformation in an aqueous environment. We performed 10 standard MD simulations of 50 ns and 10 AHBP-MD simulations of 30 ns, which all started from an extended conformation. In these AHBP-MD simulations of the β-hairpin, we used maximum force constants of -300 and 900 kJ mol^{-1} nm^{-1} for the attractive and the repulsive potentials, respectively.

Hydrogen bond reordering

The hydrogen bond lifetime and the average number of hydrogen bonds in the polyalanine simulations were evaluated to show that the intramolecular hydrogen bond reordering is accelerated. To compare the hydrogen bond lifetime in standard MD and AHBP-MD, the autocorrelation function of the intramolecular hydrogen bonds is displayed in Fig. 2. Although the autocorrelation function does not converge within the simulated time, and thus an accurate estimate of the hydrogen bond lifetime is impossible, it is clear that this function decreases significantly faster in the AHBP-MD simulation than it does in the standard MD simulation, showing that the lifetime of hydrogen bonds is much shorter because of AHBP. The average number of intramolecular hydrogen bonds per molecule under AHBP conditions is larger than that under standard conditions, having values of 4.7 and 3.9, respectively. So a smaller hydrogen bond lifetime and a larger average number of hydrogen bonds indicate that formation of new hydrogen bonds and opening of old ones are faster in the AHBP-MD simulations than in standard MD simulations.

To show that this fast reordering of hydrogen bonds also leads to a fast and broad sampling of the conformational space, we constructed RMSD matrices for all trajectories. An element in this matrix consists of the RMSD value between the structure of the protein at time x and time y. This value is

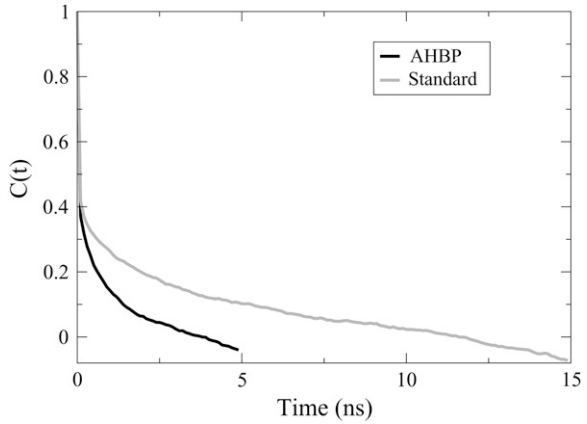

FIGURE 2 Autocorrelation function of the intramolecular hydrogen bonds in a 16-residue polyalanine. The autocorrelation function is averaged over all simulations.

calculated after a translational and a rotational fit on the backbone of the proteins. In an RMSD matrix, the similarity of each conformation in the trajectory with all the other conformations is measured by their RMSD value, with a low value indicating strong structural similarity. In these matrices, white (low RMSD) squares along the diagonal indicate time intervals where the protein structure remains very similar. Comparing the white squares in the RMSD matrix of the standard MD simulation to those of the AHBP-MD simulations shows that a standard MD simulation samples longer in a free-energy minimum, characterized by relatively similar structures, whereas the AHBP-MD simulation increases the sampling of the conformational space of a protein (Fig. 3).

α-Helix

For the polyalanine simulations, the average number of residues in an α-helical conformation is plotted versus time in Fig. 4. The N- and C-termini are not taken into account because they are too mobile. From this figure, it is clear that within the very short time of the AHBP simulation, fast formation of α-helix secondary structure occurs. The fastest formation of a full helix is observed within 6 ns (Fig. 5), and all simulations show formation of α-helical structure elements. In our four standard MD simulations, we observe only one short instance of α-helix formation (data not shown), confirming that α-helix formation is much faster and more abundant when AHBP is turned on.

Although folding of a full α-helix is very fast with the AHBP method, the structure is not stable for the rest of the simulation (Fig. 5). Clearly AHBP provides an efficient way to escape free-energy minima, but this includes the global minimum as well. Especially in the case of a polyalanine, which gains only minor stability in the global minimum (21), the protein is expected to leave this minimum quickly. On average, however, the α-helix is the most visited conformation, arising 39% of the simulation time.

β-Sheet

To test for β-sheet formation in the simulation of the folding of 1GB1 β-hairpin, we plotted the average number of residues in a β-sheet conformation versus simulation time (Fig. 6). In the AHBP-MD simulations, a steady rise of the number of residues in a β-sheet conformation is observed, whereas in the standard MD simulations this number is not as high and not as consistent. So in addition to α-helix formation, AHBP-MD simulations can also lead to fast formation of β-sheet secondary structure.

To establish if the β-sheet structures formed resemble the native state, the RMSD between the folding trajectory and the NMR structure is shown in Fig. 7. An RMSD value below 0.28 nm indicates a very good structural overlap, and simulations reaching this value are folded into the native state. Fig. 7 shows the best-performing, i.e., lowest RMSD value, AHBP and standard MD simulations. It is clear that the AHBP-MD results in fast folding to the native state, which does not occur in standard MD simulations. Four of 10 sim-

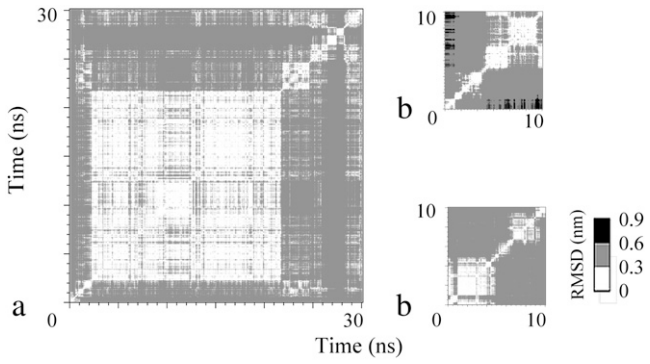

FIGURE 3 Representative RMSD matrices of standard MD (*a*) and AHBP-MD (*b*) of a 16-residue polyalanine. Matrix sizes correlate with the simulation length, and the RMSD scale is the same for all matrices.

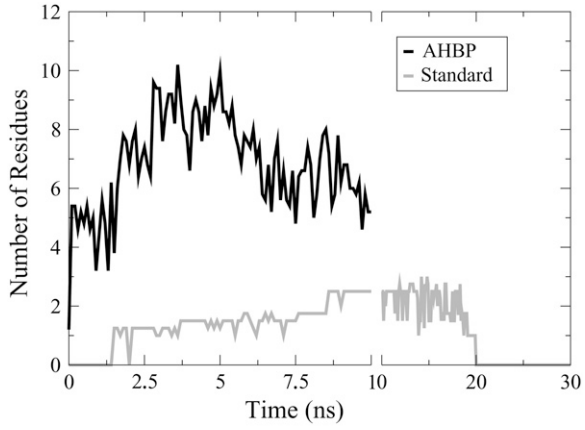

FIGURE 4 Average number of amino acid residues in an α-helix as a function of time for the AHBP-MD simulations (*solid*) and the standard MD simulations (*shaded*).

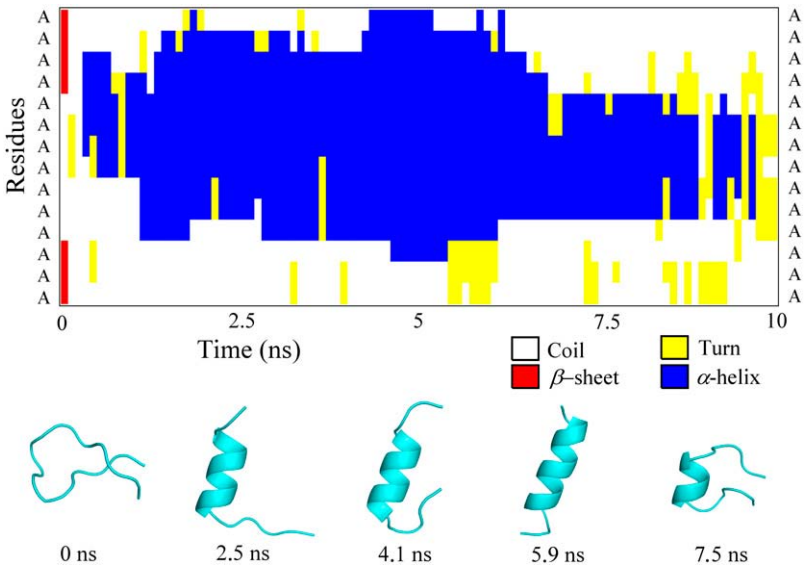

FIGURE 5 Secondary structure as a function of time for one of the polyalanine α-helix formations in an AHBP-MD simulation. Below the graph some representative structures are depicted.

ulations reach a structure similar to the NMR structure (34) within 30 ns of folding simulation. Some of the remaining folding simulations also yield a β-hairpin conformation, but nonnative side-chain interactions lead to an RMSD value of ~0.5 nm. However, these structures still show a very good structural overlap of the backbones.

DISCUSSION

We introduced AHBP in an MD simulation as a supplement to the force field to accelerate in silico protein folding. This method proves capable of rapid folding of the two most important secondary structure elements, as shown by folding of the α-helix of polyalanine and the β-hairpin of the C-terminus of protein G.

Many folding studies rely on very simple models (1,2) to sample a longer time span in a simulation. Especially implicit solvent models (35,36) are widely applied; however, the clear importance of an explicit solvent model in protein folding cannot be ignored (37–39). With our method, microscopic detail is retained, as well as an explicit description of the solvent molecules. These requirements can also be achieved with parallel MD schemes (5); however, these schemes are limited to small, rapidly folding proteins because of the slow dynamics of each individual simulation. The proposed method increases this speed and thus requires far less computer time to obtain similar folding ensembles. Other methods that increase the dynamics of individual simulations include, for instance, a self-guided MD approach (40), but this results in an irreversible folding path. AHBP is a reversible method as shown by the unfolding events in the polyalanine simulations.

Because many parameters have not been optimized, and a very simple hydrogen bond potential is used in these experiments, i.e., no COH angle and no dihedral angles are incorporated, we expect to achieve even better results by tuning

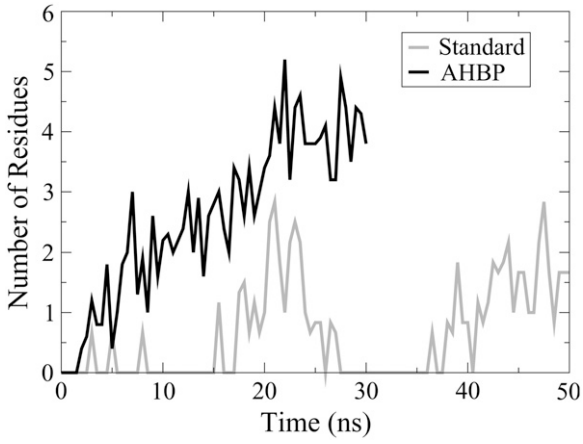

FIGURE 6 Average number of amino acid residues in a β-sheet as a function of time for the AHBP-MD (*solid*) and the standard MD (*shaded*) simulations.

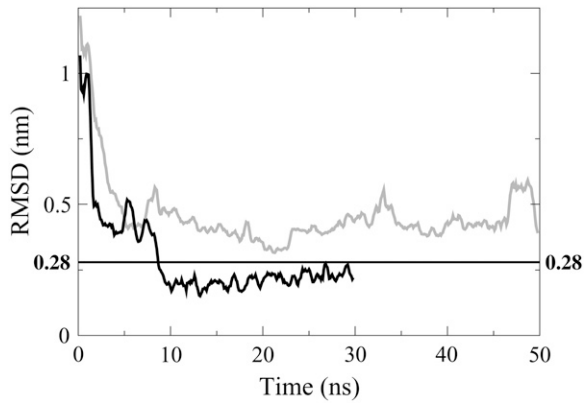

FIGURE 7 RMSD between the simulation trajectory and the NMR structure as a function of time. The best-performing AHBP-MD (*solid*) and standard MD (*shaded*) simulations are plotted.

the parameters and improving the AHBP. Work in this direction is under way with promising results.

The authors are grateful to Jocelyne Vreede, Jeroen van Gestel, Jeroen van Bemmelen, and Jon Laman for critically reading the manuscript. The authors also thank Jouke Heringa for his assistance with all computer and cluster related issues. Finally, the authors thank Jaap Flohil from the Foldyne company for sharing their software code and for many useful discussions.

This work was supported by The Netherlands Organization for Scientific Research for funding (grant No. 635.100.012, program for computational life sciences).

REFERENCES

1. Irbäck, A. 2005. Peptide folding and aggregation studied using a simplified atomic model. *J. Phys. Cond. Matt.* 17:S1553–S1564.
2. Smith, A. V., and C. K. Hall. 2001. Alpha-helix formation: discontinuous molecular dynamics on an intermediate-resolution protein model. *Proteins.* 44:344–360.
3. Snow, C. D., E. J. Sorin, Y. M. Rhee, and V. S. Pande. 2005. How well can simulation predict protein folding kinetics and thermodynamics? *Annu. Rev. Biophys. Biomol. Struct.* 34:43–69.
4. Gnanakaran, S., H. Nymeyer, J. Portman, K. Y. Sanbonmatsu, and A. E. Garcia. 2003. Peptide folding simulations. *Curr. Opin. Struct. Biol.* 13: 168–174.
5. Pande, V. S., I. Baker, J. Chapman, S. P. Elmer, S. Khaliq, S. M. Larson, Y. M. Rhee, M. R. Shirts, C. D. Snow, E. J. Sorin, and B. Zagrovic. 2003. Atomistic protein folding simulations on the submillisecond time scale using worldwide distributed computing. *Biopolymers.* 68:91–109.
6. Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314:141–151.
7. Levinthal, C. 1969. How to fold graciously. *In* Mossbauer Spectroscopy in Biological Systems. J. T. P. DeBrunner and E. Munck, editors. University of Illinois Press, Monticello, IL. 22–24.
8. Bryngelson, J. D., J. N. Onuchic, N. D. Socci, and P. G. Wolynes. 1995. Funnels, pathways, and the energy landscape of protein-folding - A synthesis. *Proteins.* 21:167–195.
9. Dill, K. A., and H. S. Chan. 1997. From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4:10–19.
10. Dobson, C. M., and M. Karplus. 1999. The fundamentals of protein folding: bringing together theory and experiment. *Curr. Opin. Struct. Biol.* 9:92–101.
11. Tanford, C. 1962. Contribution of hydrophobic interactions to stability of globular conformation of proteins. *J. Am. Chem. Soc.* 84:4240–4247.
12. Makhatadze, G. I., and P. L. Privalov. 1995. Energetics of protein structure. *In* Advances in Protein Chemistry. Academic Press, San Diego. 307–425.
13. Myers, J. K., and C. N. Pace. 1996. Hydrogen bonding stabilizes globular proteins. *Biophys. J.* 71:2033–2039.
14. Baker, E. N., and R. E. Hubbard. 1984. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* 44:97–179.
15. Honig, B. 1999. Protein folding: from the Levinthal paradox to structure prediction. *J. Mol. Biol.* 293:283–293.
16. Van der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen. 2005. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* 26:1701–1718.
17. Fernandez, A., J. Kardos, and Y. Goto. 2003. Protein folding: could hydrophobic collapse be coupled with hydrogen-bond formation? *FEBS Lett.* 536:187–192.
18. Lumb, K. J., and P. S. Kim. 1995. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry.* 34:8642–8648.
19. Brockwell, D. J., and S. E. Radford. 2007. Intermediates: ubiquitous species on folding energy landscapes? *Curr. Opin. Struct. Biol.* 17:30–37.
20. Krishna, M. M. G., and S. W. Englander. 2007. A unified mechanism for protein folding: Predetermined pathways with optional errors. *Protein Sci.* 16:449–464.
21. Mu, Y. G., P. H. Nguyen, and G. Stock. 2005. Energy landscape of a small peptide revealed by dihedral angle principal component analysis. *Proteins.* 58:45–52.
22. Gruebele, M. 2005. Downhill protein folding: evolution meets physics. *C.. R..Soc. Biol. (Paris).* 328:701–712.
23. Fersht, A. R., J. P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Y. Waye, and G. Winter. 1985. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature.* 314:235–238.
24. van der Spoel, D., P. J. van Maaren, P. Larsson, and N. Timneanu. 2006. Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media. *J. Phys. Chem. B.* 110:4393–4398.
25. Williams, D. H., M. S. Searle, J. P. Mackay, U. Gerhard, and R. A. Maplestone. 1993. Toward an estimation of binding constants in aqueous solution: studies of associations of vancomycin group antibiotics. *Proc. Natl. Acad. Sci. USA.* 90:1172–1178.
26. Young, J. C., V. R. Agashe, K. Siegers, and F. U. Hartl. 2004. Pathways of chaperone-mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell Biol.* 5:781–791.
27. Hartl, F. U., and M. Hayer-Hartl. 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science.* 295:1852–1858.
28. Thirumalai, D., and G. H. Lorimer. 2001. Chaperonin-mediated protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 30:245–269.
29. Kortemme, T., A. V. Morozov, and D. Baker. 2003. An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein-protein complexes. *J. Mol. Biol.* 326:1239–1259.
30. van Gunsteren, W. F., S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, and W. G. Tironi. 1996. Biomolecular Simulation: The GROMOS96 Manual and User Guide. Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland
31. Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. 1981. Interaction models for water in relation to protein hydration. *In* Intermolecular Forces. B. Pullman, editor. D. Reidel Publishing, Dordrecht, The Netherlands. 331–342.
32. Hess, B., H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. 1997. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18:1463–1472.
33. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98: 10089–10092.
34. Gronenborn, A. M., D. R. Filpula, N. Z. Essig, A. Achari, M. Whitlow, P. T. Wingfield, and G. M. Clore. 1991. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science.* 253:657–661.
35. Feig, M., and C. L. Brooks. 2004. Recent advances in the development and application of implicit solvent models in biomolecule simulations. *Curr. Opin. Struct. Biol.* 14:217–224.
36. Vorobjev, Y. N., and J. Hermans. 1999. ES/IS: estimation of conformational free energy by combining dynamics simulations with explicit solvent with an implicit solvent continuum model. *Biophys. Chem.* 78: 195–205.
37. Tan, C. H., L. J. Yang, and R. Luo. 2006. How well does Poisson-Boltzmann implicit solvent agree with explicit solvent? A quantitative analysis. *J. Phys. Chem. B.* 110:18680–18687.
38. Zhou, R. H. 2003. Free energy landscape of protein folding in water: explicit vs. implicit solvent. *Proteins.* 53:148–161.
39. Rhee, Y. M., E. J. Sorin, G. Jayachandran, E. Lindahl, and V. S. Pande. 2004. Simulations of the role of water in the protein-folding mechanism. *Proc. Natl. Acad. Sci. USA.* 101:6456–6461.
40. Wu, X. W., and S. M. Wang. 1998. Self-guided molecular dynamics simulation for efficient conformational search. *J. Phys. Chem. B.* 102: 7238–7250.