PROTEINS: Structure, Function, and Genetics 34:464–471 (1999)

A New Hybrid Monte Carlo Algorithm for Protein Potential Function Test and Structure Refinement

Hongyu Zhang*
Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland

ABSTRACT **A new Hybrid Monte Carlo (HMC) algorithm has been developed to test protein potential functions and, ultimately, refine protein structures. The main principle of this algorithm is, in each cycle, a new trial conformation is generated by carrying out a short period of molecular dynamics (MD) iterations with a set of random parameters (including the MD time step, the number of MD steps, the MD temperature, and the seed for initial MD velocity assignment); then to accept or reject the new conformation on the basis of the Metropolis criterion. The novelty in this paper is that the potential in MD iterations is different from that in the MC step. In the former, it is a molecular mechanics potential, in the latter it is a knowledge-based potential (KBP). Directed by the KBP, the MD iteration is used to search conformational space for realistic conformations with low KBP energy. It circumvents the difficulty in using KBP functions directly in MD simulation, as KBP functions are typically incomplete, and do not always have continuous derivatives required for the calculation of the forces. The new algorithm has been tested in explorations of conformational space. In these test calculations the KBP energy was found to drop below the value for the native conformation, and the correlation between the root mean square deviation (RMSD) and the KBP energy was shown to be different from the test results in other references. At the present time, the algorithm is useful for testing new KBP functions. Furthermore, if a KBP function can be found for which the native conformation has the lowest energy and the energy/RMSD correlation is good, then this new algorithm also will be a tool for refinement of the theory-based structural models. Proteins 1999;34:464–471.** © 1999 Wiley-Liss, Inc.

Key words: protein folding; energy function test; Monte Carlo; molecular dynamics; hybrid

INTRODUCTION

The present strategies for the prediction of protein tertiary structures could be divided into two main categories: comparative modeling and de novo prediction. In the former, one can build some successful models of proteins from their homologous analogues, but in many cases, techniques such as "threading" need to be applied for those proteins that have very low sequence similarity to the known protein structures. In de novo prediction, random initial conformations can be folded to near-native conformations using some conformational search methods and force fields. One common premise of the two methods is the energy or score function that can effectively distinguish the near-native conformation from alternate conformations.

To stringently test an energy function, we need a large number of decoys in various kinds of test sets. These decoys, or alternate conformations, are expected to be compact, globular, and not very different from native conformations. There are various methods to generate those test sets.[1] The most generally used one is based on the sequence-recognize-structure protocol developed by Hendlich et al.[2] The target sequence is modeled as a sequence fragment of all larger protein structures and is threaded through each larger protein sequence advancing one residue at a time, generating a larger number of alternate structures. The structures generated by this method are sometimes nonnative: they are usually not compact and side chains are often disregarded. In addition, the structure library generated by this method is quite limited for large proteins. Another major strategy for generating test sets is the MD trajectory method.[3–9] In this method, a significant amount of conformational space in the neighborhood of the native structure can be sampled by the MD simulation. Compared to the conformations in the threading test set, these alternate structures have both the backbone and the side-chain atoms, moreover, they have reasonable atomic level packing and interaction. Thus they are more native-like. On the other hand, the test structures can be very close to the real native structure, and therefore be more challenging for energy functions. Recently, Wang et al.[3–4] have shown that many energy functions, which were proved to perform very well in the threading test, failed significantly in the MD trajectory test. They showed that a potential based on an atomic solvation model, the WZS model, trained by neural networks performed impressively in discriminating the native structure among many molecular dynamics (MD)-generated decoys. Following this initial work, other researchers have tested various potentials on MD trajectories.[5–9] Huang et al.[5,6] demonstrated that a very simple

*Correspondence to: Hongyu Zhang, Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, MD 20850. E-mail: hyzhang@carb.nist.gov
Received 28 April 1998; Accepted 5 November 1998

hydrophobic fitness (HF) potential could impressively discriminate the native structures of several proteins among the large conformations pools generated by MD simulations. Similarly, Debolt et al.[7] developed two atomic-level potentials shown to be good in the rank identifying of the progressively less native-like structures generated by MD sampling.

Useful as the MD trajectory test set is, it is often hard for us to judge how convincing it is. The reason is that MD trajectories can vary significantly depending on the simulation conditions such as temperature, initial velocity, and environment. Therefore in the present paper we develop a new hybrid algorithm that can be used to search the conformational space of proteins under dynamical MD simulation conditions (variable temperature and starting velocity, etc.); moreover, it uses the KBP energy to direct the MD simulations in order to search for the conformational space of low knowledge-based potential (KBP) energy. This new hybrid algorithm is an extension of the Hybrid Monte Carlo algorithm (HMC) developed by Duane et al.[10] originally used in quantum chromodynamics.

The Hybrid Monte Carlo algorithm is a kind of mixture of molecular dynamics and Monte Carlo algorithm.[10–14] The essential idea of the algorithm is to carry out L steps of molecular dynamics iterations with a Monte Carlo step between each series of L dynamic iterations. Hence, the L iterations correspond to the random perturbation of the conformations in the classical MC approaches. For large L and vanishing MC contributions, the technique converges to classical molecular dynamics. The HMC algorithm was shown to be a protocol with a more powerful conformational searching ability and faster-convergence characteristic than the traditional MD method in the simulation of biopolymers.[10–14]

Although there are various new powerful algorithms like HMC being developed (see reviews of Byrne et al.[15] and Leontidis et al.[16]), they still have obstacle in sampling the conformational space for lack of good energy functions. For example, Novotny et al.[17,18] demonstrated that it is hard for the traditional molecular mechanics force fields to discriminate the native structure from the alternate structures.

In order to solve the problem, various knowledge-based potentials were derived from the protein structure database (for review of applications, see Sippl,[19–21] Jernigan and Bahar,[22] Rooman and Wodak,[23] Godzik et al.,[24] Miyazawa and Jernigan,[25] Thomas and Dill,[26] and Moult[27]). It has been demonstrated that they are able to efficiently discriminate the correct fold against other decoys. A very new method for calculating the total conformational free energy of proteins in water solvent was developed by Hermans and coworkers[28] lately. It used both dynamics simulations with an explicit solvent and an implicit solvent continuum model. The new energy function can successfully discriminate the misfolded conformations.

A natural idea to develop a new folding algorithm is to combine advanced conformational searching methods with the newly developed potentials.[29] The new HMC algorithm is one of the steps in this direction.

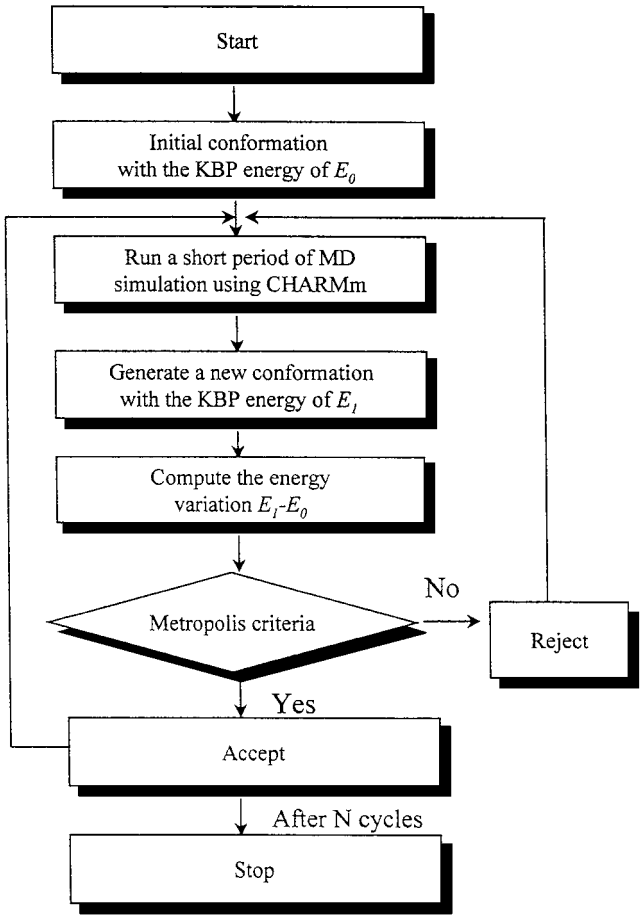

Fig. 1. Frame work of the new hybrid Monte Carlo algorithm.

METHODS

The framework of the new HMC algorithm is illustrated in Figure 1. The initial conformation of the protein has the KBP energy of E_0. A short period of MD simulation using CHARMm[30] is carried out to generate a trial conformation. The KBP energy of the newly generated trial conformation is E_1. The probability of accepting the conformation is given by the Metropolis criterion

$$P = \min\{1, e^{-\Delta E/kT}\}$$

where $E = E_1 - E_0$, and k is the Boltzmann constant. After accepting or rejecting the trial conformation, the next cycle of simulation is repeated with a new set of random parameters including the MD-time step, the number of MD steps, the MD temperature and the seed for initial MD velocity assignments. The temperature T in the Metropolis criterion is gradually decreased during simulation (simulated annealing).

In normal MC, the trial conformation would be generated by randomly perturbing the initial conformation typically by rotating some torsion angles, while in this hybrid MC algorithm, it was generated by a random MD iteration.

All the MD simulations in this work were carried out in vacuum, with a distance- dependent dielectric constant. The time step of the MD simulation is a random value between 0.5 and 5 femto-seconds (fs). Each series of MD iteration is a heating procedure. The initial velocity is assigned from a Gaussian distribution around 300 K. The system is heated from 300 K to a random temperature between 300 and MD_T. The default value of MD_T is 350. In the case where the system is trapped in one conformation for too long (e.g., staying in the same conformation in more than 20 MC cycles successively), MD_T will be increased 20 percent each time the new conformation is rejected until it reaches 600. And once the system finds a new acceptable state, MD_T is reset to the initial value of 350. In the same way, the number of MD iteration steps in one MC cycle is also set to a random number between 1 and MD_STEP. And MD_STEP has the default value of 100 and will be increased if the simulation is trapped. In some systems, a short energy minimization was carried out between the MD iterations and the Monte Carlo step to relieve bad contacts.

In this algorithm, the MD iteration basically works as a conformational searching tool. The CHARMm potential energy is not related to the Metropolis criterion except that it becomes positive, which means that the new conformation has some very unfavorable interactions, such as bad steric packing or electrostatic distribution, etc. Under this special circumstance the new conformation would be rejected absolutely.

Since there are so many KBP functions available in the literature and their performance varies a lot dependent on the test sets, this work only specifically selected those potentials already demonstrated to perform well in MD trajectory test set. They are the hydrophobic fitness (HF) potential developed by Huang et al.,[5,6] the pair-wise atomic potential (PWA) by Debolt et al.[7] and the WZS potential by Wang et al.[3,4]

The first energy function we tried is the HF score. This is a quite simplified energy function. It enumerates contacts between hydrophobic residues while weighting their sum by the total number of residues surrounding these hydrophobic residues. Thus, it prefers compact folds with the desired structural feature of a buried, intact core. In their work, Huang et al.[6] have shown that this simple energy function can impressively rank the native structures as the lowest energy conformations among thousands of decoys generated by MD simulations; moreover, it was shown to have a good RMSD/energy correlation.

The second energy function, PWA potential is a detailed atomic-level pair-wise function. It reflects the contact preference of protein atomic pairs. The energy value was shown to be correlated well to the RMSD of the structure during MD simulation.[7] The third energy function tested in this work is the WZS potential. It is based on a solvation model trained by neural networks, and was demonstrated to be able to impressively discriminate almost all of the decoys in the near native space generated by MD simulations.[3,4]

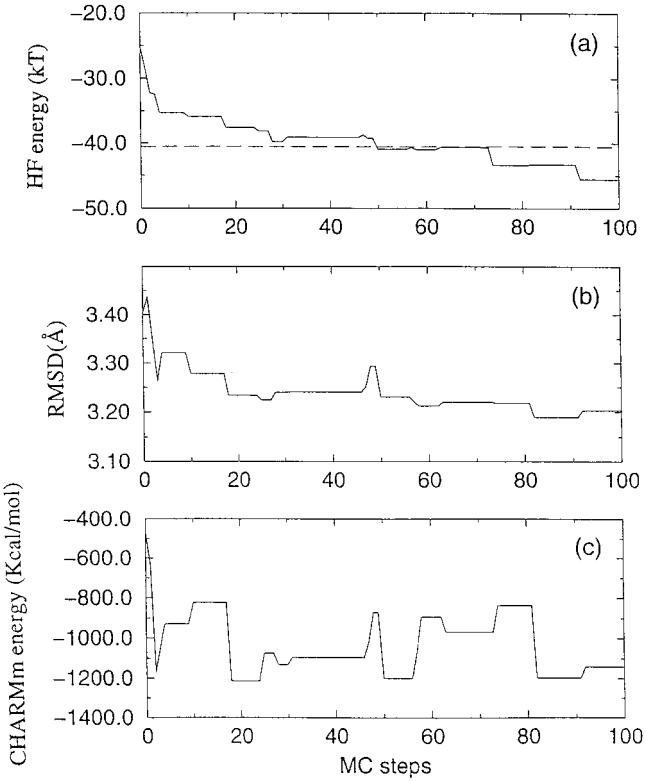

Fig. 2. Variations of the HF energy, the RMSD and the CHARMm energy during the simulation of protein 4icb. The dashed line in (a) represents the native energy. MC temperature: 1–0.9.

Three proteins were selected as the testing systems in this paper: protein 4icb, protein A and protein 2i1b. Protein 4icb is a four-helix bundle. The reason to select this protein is that it is the most successful example for the HF energy as illustrated in the paper of Huang et al.[6] In the work, it was demonstrated that the native structure of this protein is ranked as the lowest energy conformation among 2,000 decoys generated by the MD simulation, among them 1,000 decoys are around 1.96 Å and the other 1,000 are around 4.80 Å in RMSD from the native structure. In our test, the initial conformation of protein 4icb was generated from in vacuo high temperature MD simulation at 600 K. The C atom RMSD from the native structure is 3.4 Å.

The second protein, protein A, is a small protein consisting of three helix bundles. It's folding pathway has been extensively studied by several groups (Kolinski et al.,[31] Olszewski et al.,[32] Boczko and Brooks[33]). The reason to select this protein is that it is a real system to test the HMC algorithm for a refinement of the protein structure, as the initial conformation is directly from an ab initio simulation lattice model simulation by Kolinski et al.[31] The C RMSD from the native structure is 4.1 Å.

To also test the performance of the potentials in a more grossly non-native conformational space, the third protein 2i1b was selected from the EMBL deliberately misfolded set.[34] The native fold of 2i1b is an all- protein, while the

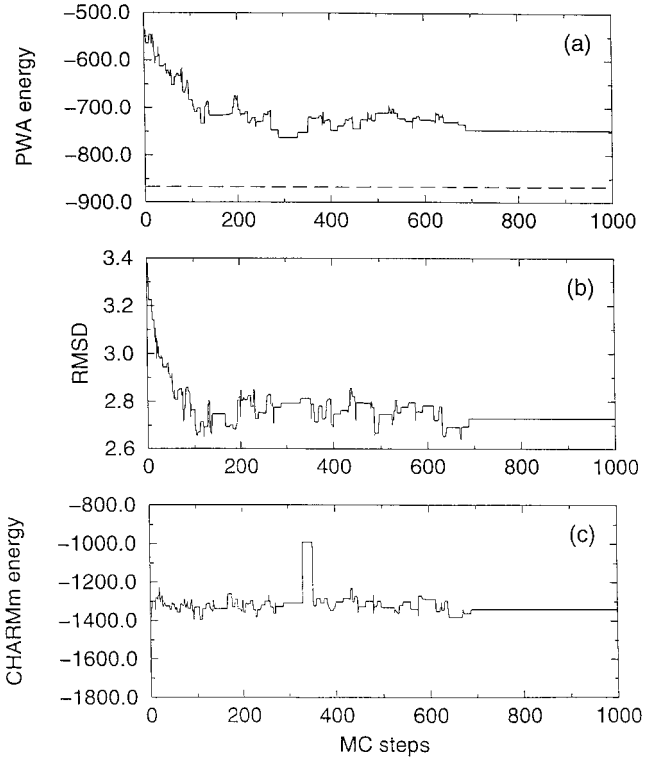

Fig. 3. Variations of the PWA energy, the RMSD and the CHARMm energy during the simulation of protein 4icb. The dashed line in (a) represents the native energy. MC temperature: 10–6.

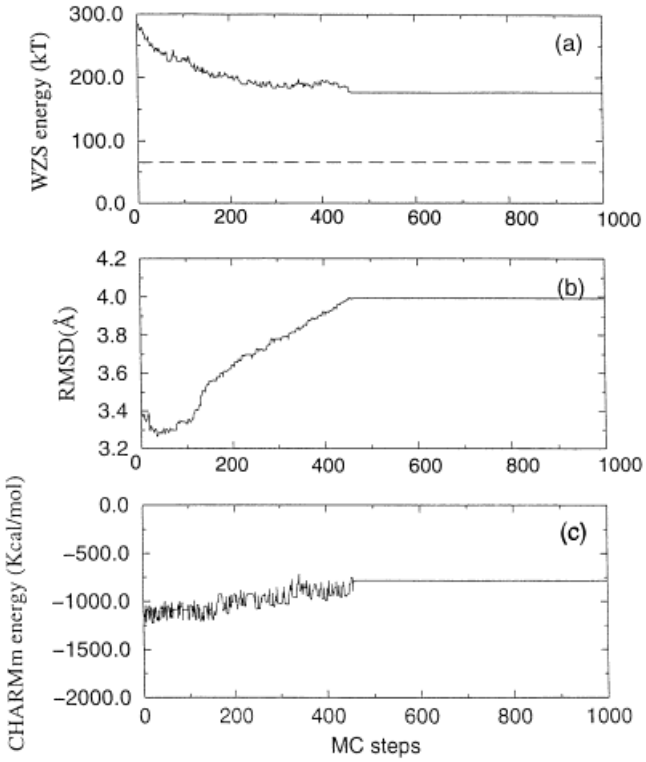

Fig. 4. Variations of the WZS energy, the RMSD and the CHARMm energy during the simulation of protein 4icb. The dashed line in (a) represents the native energy. MC temperature: 5–0.5.

misfolded conformation was generated by swapping the residues of 2i1b to another protein 1lh1, which has the same sequence length but is an all- protein. The side-chain packing was annealed using a Monte Carlo process.[34] The RMSD between the conformation and the native structure is 17.5 Å. In the test of this system, an extra short energy minimization consisting of 50 steps of Steepest Descent (SD) minimization was carried out following the MD iterations to relieve the bad contacts within the structure.

The main program of the HMC algorithm was implemented in C++, which uses "system" commands to execute the CHARMm and KBP code. The advantage here is that it is very easy to switch KBP functions, as it only needs the executable codes of the potential functions, which are simply "plugged into" the main program (Plug and Play). Thus one does not need to write a new code for a new potential.

RESULT AND DISCUSSION

The main question our work addresses is: can this algorithm be basically used as a tool for the performance test of energy functions? And more interestingly, can it refine protein structures? To address the above question we tested the KBP functions on three proteins: 4icb, protein A and 2i1b.

In the first system (protein 4icb), Figure 2 displays the variation of the HF energy during the hybrid simulation. The HF energy of the initial conformation is well above the native structure energy, which conforms to the original work of Huang et al.,[6] i.e., this potential works well in discriminating the MD decoys from the native structure. In the following HMC simulation, however, the energy was significantly minimized to lower than the native structure energy within only 100 Monte Carlo steps. The final conformation has the RMSD around 3.2 Å, which means that we have found a decoy structure with an energy below the native structure energy. In this case, the HMC algorithm is a stricter test method than the MD trajectory test used in the original reference (see Figure 2 in ref. 6).

During the energy minimization, we also investigated the variation of the RMSD. When the energy declined, the RMSD changed just slightly, from 3.4 Å to 3.2 Å. Several runs with different hybrid algorithm parameters were tested, most of which repeat this weak trend.

The CHARMm potential energy does not show obvious correlation with the RMSD variation (Figure 2c). There could be at least two factors for this, first, we have not included solvation effects directly in the MD simulation; secondly, as pointed out by a couple of works[3,4,6] ordinary molecular mechanics force fields usually do not find a good energy/RMSD correlation.

As a next step, we checked if more detailed atomic-level KBP functions would work as more sensitive tool in the hybrid algorithm. From Figure 3 we can see that the PWA energy and the RMSD were simultaneously minimized during the simulation. The energy curve and RMSD curve

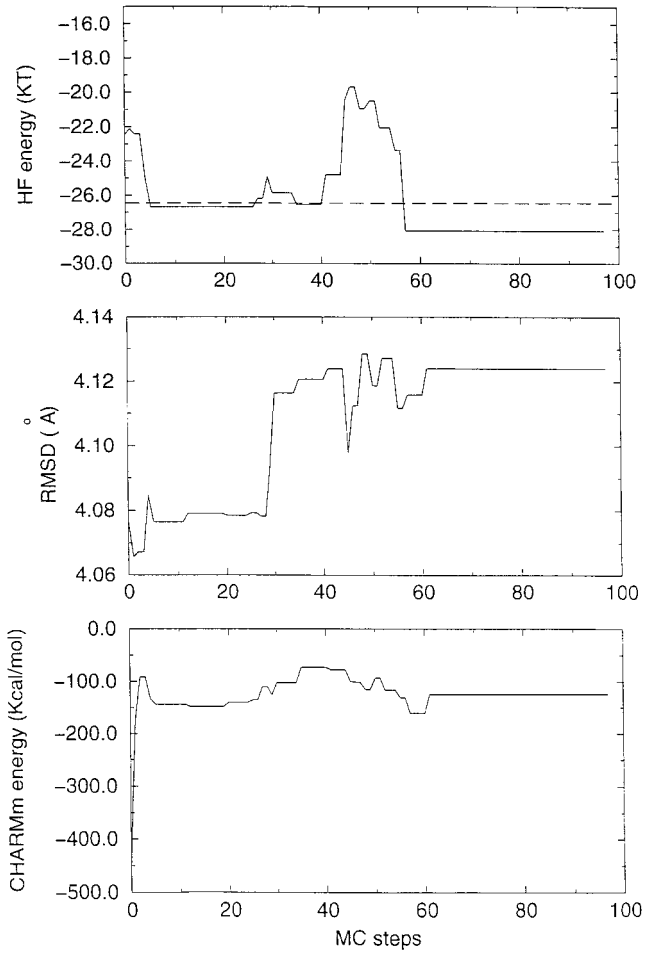

Fig. 5. Variations of the HF energy, the RMSD and the CHARMm energy during the simulation of protein A. The dashed line represents the native energy. MC temperature: 1–0.9.

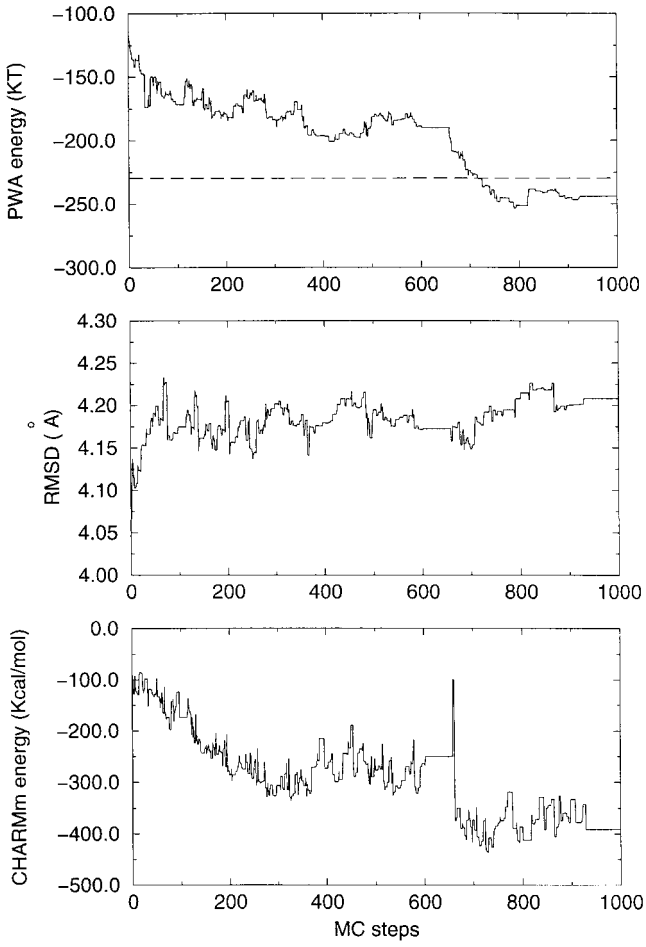

Fig. 6. Variations of the PWA energy, the RMSD and the CHARMm energy during the simulation of protein A. The dashed line represents the native energy. MC simulated annealing temperature: 5–1.

are not exactly the same, however their trends are very similar.

During the simulations, the energy of the third potential (WZS potential) stayed higher than the native energy shown in Figure 4, which suggested that this potential could well discriminate the native structure from the MD-generated conformations in this test system. However, the RMSD went up as the energy went down in the simulation, although it was shown in the original reference that the WZS potential has a good energy/RMSD correlation (Fig. 5 in ref. 4).

The second system, protein A, is not only a decoy system, but also a real problem for the structure refinement method, as the initial conformation is directly from an ab initio simulation.[31] It can be seen from Figure 5 that, like the first system, the HF potential energy quickly plunged lower than the native structure energy after only a few MC steps. Here the resultant conformation is still far away from the native structure (RMSD around 4.1 Å). We can see in Figure 6 and Figure 7 that the PWA potential and the WZS potential met the same trouble. They all can be minimized significantly, the RMSD curve, however, went up slightly.

The third system, protein 2i1b, was started from a misfolded conformation far away from the native structure. It is expected that the KBP potentials should be able to correctly identify the misfolded conformation from the native fold. As seen in Figure 8 to Figure 10, the three KBP functions did verify their discrimination ability for the initial conformations. The PWA potential, however, was minimized to be lower than the native energy after about 150 steps, when the RMSD is still very high (around 17.4 Å). The reason for the difference might result from the physical basis of the three energy functions. When browsing the misfolded conformation in a graphical interface, we noticed that there are quite a few exposed hydrophobic and buried hydrophilic residues. The major factor in the HF and the WZS potentials is the hydrophobic effect, while the PWA potential mainly reflect a pair-wise preference between residues, which is only partly related to the hydrophobic effect. Therefore, in this system, the HF potential and the WZS potential are believed to perform better than the PWA potential.

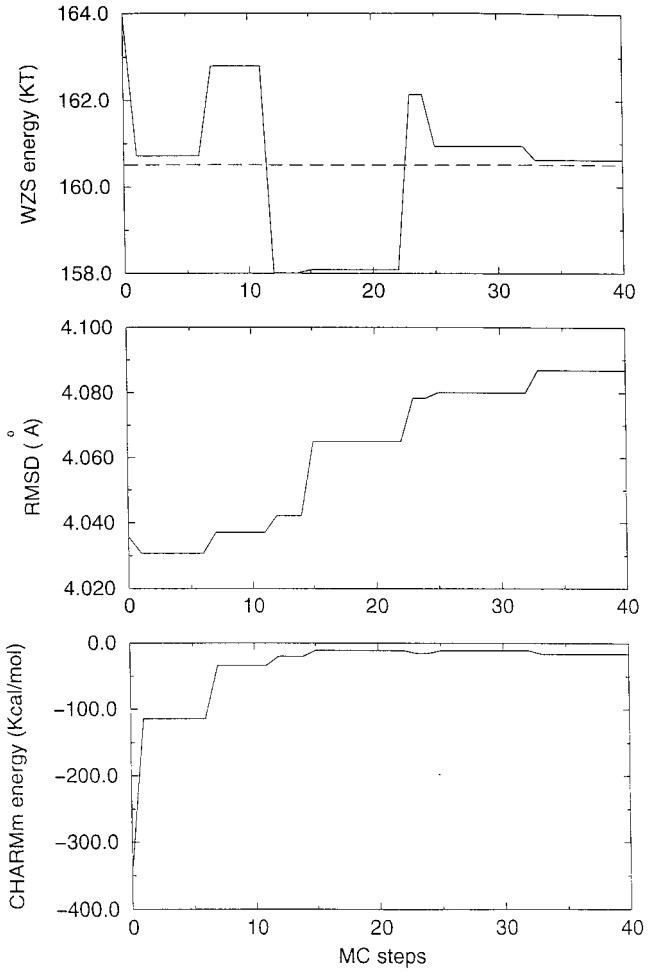

Fig. 7. Variations of the WZS energy, the RMSD and the CHARMm energy during the simulation of protein A. The dashed line represents the native energy. MC temperature: 5–1.

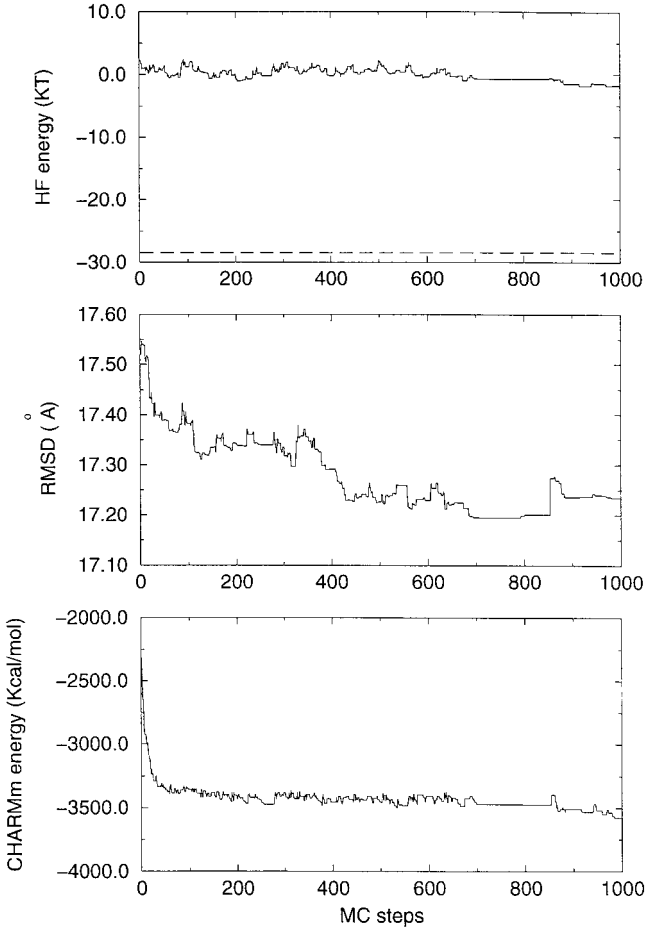

Fig. 8. Variations of the HF energy, the RMSD and the CHARMm energy during the simulation of protein 2i1b. The dashed line represents the native energy. MC temperature: 0.5–0.1.

Based on all the test results above, one can see that the new HMC algorithm is a useful tool to test the KBP functions; moreover, in many ways, this new algorithm is stricter than a classical MD trajectory method. The reason is this: in the HMC algorithm, the MD simulation dynamically searches for the conformational space of low KBP energy, therefore it is more aggressive than the static energy function test methods such as the MD trajectory method or the threading method. Actually, it was illustrated in this work that an energy minimum in the conformational space distant to the native structure can always be found for these potentials.

Although all the potential functions tested in this paper have been demonstrated to be good at energy/RMSD correlation in the MD trajectory test in literatures, the inclusion of them in the new HMC algorithm does not bring good performance in the structure refinement. Actually, this is an unfortunate fact for the present protein-folding community: there is no ideal energy function available yet. Most of the present energy functions describe only part of the interactions in protein folding, such as hydrophobic packing, pair-wise preference or solvation surface exposure. One of the possible ways to improve structure refinement results might be to use some sort of combinations of different functions. Besides that, one can argue that better simulation strategies are yet required such as carrying out the MD simulation in solvent instead of in vacuum or selecting more appropriate MD/MC parameters.

Because we are prohibited by the searching extent of MD iterations and also depend on the cooperation between the MD iterations and the KBP functions, the conformational searching ability of this new algorithm is still not broad enough to cover a pathway between a grossly non-native conformation and the native conformation. Thus we postulate that this new algorithm works as a structure refinement tool.

One could argue that using the KBP function alone in any MD or MC minimization could produce a better decoy set than using the present algorithm, however, the KBP functions are typically not continuously differentiable, making it hard to implement them in any MD simulation. Second, although there is a general way to use KBP

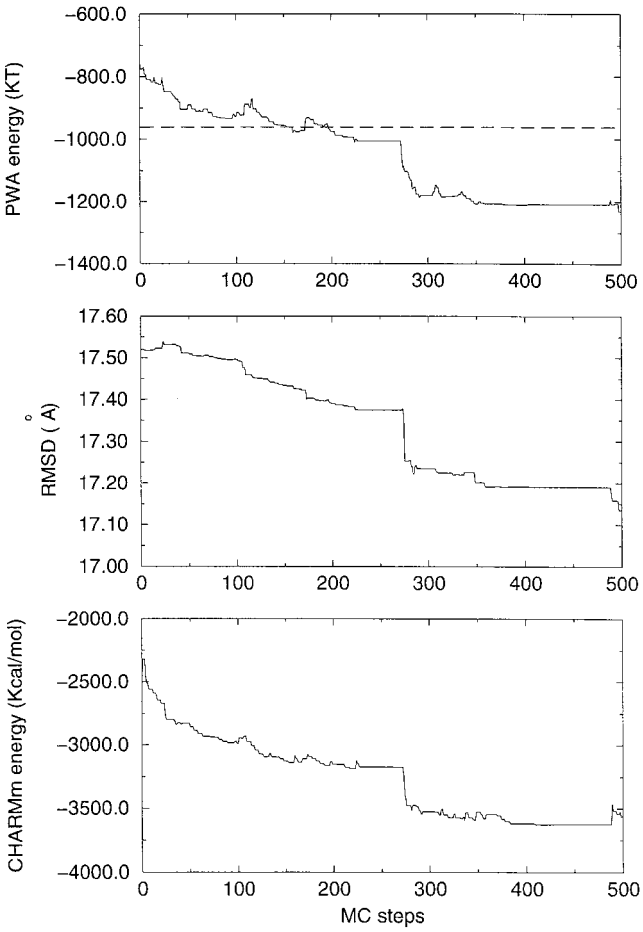

Fig. 9. Variations of the PWA energy, the RMSD and the CHARMm energy during the simulation of protein 2i1b. The dashed line represents the native energy. MC temperature: 10–5.5.

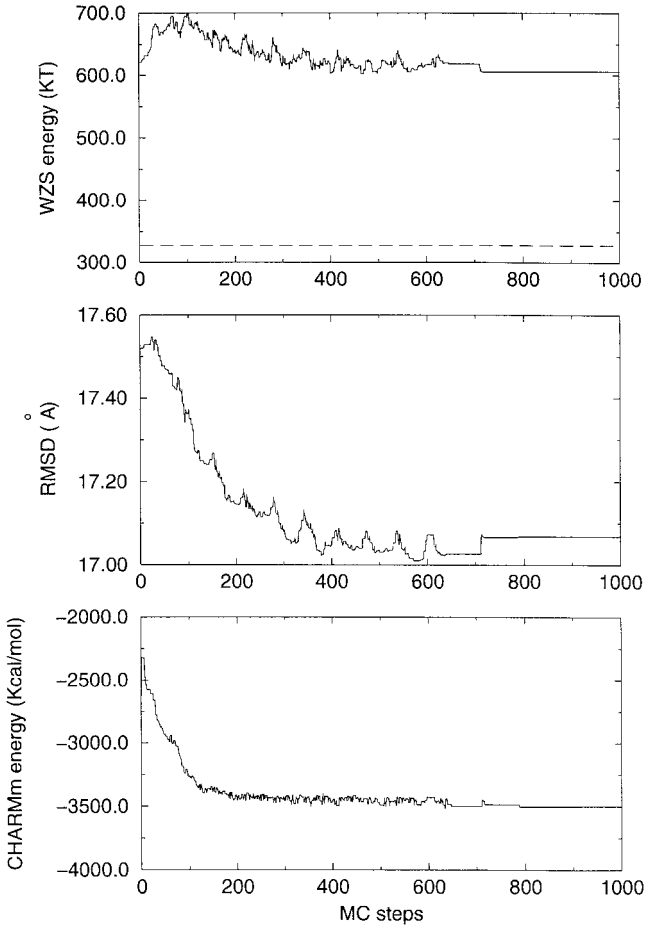

Fig. 10. Variations of the WZS energy, the RMSD and the CHARMm energy during the simulation of protein 2i1b. The dashed line represents the native energy. MC temperature: 10–1.

functions in MC or genetic algorithm (GA), one usually needs specific programming for different KBP functions. In our algorithm, however, one does not need to write a single line of new code when testing a new potential. Lastly, we have illustrated that our algorithm generated more realistic decoys, which is proved by the molecular mechanics potential energy.

CONCLUSION

The new Hybrid Monte Carlo algorithm was demonstrated to be a useful tool for energy function test. It found out structure decoys that have lower energies than the native structures, which were not revealed by other methods; and the RMSD/energy correlation is illustrated to be different from what we saw in the MD trajectory test. This algorithm is also expected to be a protocol for protein structure refinement given a better energy function.

ACKNOWLEDGMENTS

I am very grateful to Prof. John Moult in CARB for the encouragement in writing this paper. The idea of this paper was born out during my short visiting in the group of Prof. Jeffrey Skolnick in The Scripps Research Institute (TSRI) and my Ph.D. work in the group of Prof. Luhua Lai in Peking University, so I am thankful to them for the stimulating environment. I also thank Dr. Yuhong Wang for providing the WZS model code and parameters, Dr. Michal Vieth for providing the CHARMm script to generate the all atom model of protein A from the C model, and Dr. Li Zhang for helpful discussion. My last thanks are to the editor, Prof. Jan Herman, and my coworker in CARB, Dr. Susan K. Gregurick. They helped me make this paper easier to understand.

REFERENCES

1. Moult J. Pro-Star: the protein potential site. Evaluation of potentials (http://prostar.carb.nist.gov), 1998.
2. Hendlich M, Lackner P, Weitchus S, et al. Identification of native protein folds amongst a large number of incorrect models. The calculation of low energy conformations from potentials of mean force. J Mol Biol 1990;216:167–180.
3. Wang Y, Zhang H, Wei L, Scott R. Discriminating compact nonnative structures from the native structure of globular proteins. Proc Natl Acad Sci USA 1995;92:709–713.
4. Wang Y, Zhang H, Scott R. A new computational model for protein folding based on atomic solvation. Protein Sci 4: 1995;1402–1411.

5. Huang ES, Subbiah S, Tsai J, Levitt M. Using a hydrophobic contact potential to evaluate native and near-native folds generated by molecular dynamics simulations. J Mol Biol 1996;257:716–725.
6. Huang ES, Subbiah S, Levitt M. Recognizing native folds by the arrangement of hydrophobic and polar residues. J Mol Biol 1995;252:709–720.
7. Debolt S, Skolnick J. Evaluation of atomic level mean force potentials via inverse folding and inverse refinement of protein structures: atomic burial position and pair-wise non-bonded interactions. Protein Eng 1996;9:637–655.
8. Park B, Levitt M. Energy functions that discriminate X-ray and near native folds from well-constructed decoys. J Mol Biol 1996;258: 367–392.
9. Park B, Huang E, Levitt M. Factors affecting the ability of energy functions to discriminate correct from incorrect folds. J Mol Biol 1997;266:831–846.
10. Duane S, Kennedy AD, Pendleton BJ, Roweth D. Hybrid Monte Carlo. Phys Lett B 1987;195:216–222.
11. Brass A, Pendleton BJ, Chen Y, Robson B. Hybrid Monte Carlo simulations theory and initial comparison with molecular dynamics. Biopolymers 1993;33:1307–1315.
12. Forrest BM, Suter UW. Hybrid Monte Carlo simulations of dense polymer systems. J Chem Phys 1994;101:2616–2629.
13. Clamp ME, Baker PG, Stirling CJ, Brass A. Hybrid Monte Carlo: an efficient algorithm for condensed matter simulation. J Comput Chem 1994;15:838–846.
14. Guarnieri F, Still WC. A rapidly convergent simulation method: mixed Monte Carlo/Stochastic dynamics. J Comput Chem 1994;15: 1302–1310.
15. Byrne D, Li J, Platt E, Robson B, Weiner P. Novel algorithms for searching conformational space. J Comput Aided Mol Des 1994;8: 67–82.
16. Leontidis E, Forrest BM, Widmann AH, Suter UW. Monte Carlo algorithms for the atomistic simulation of condensed polymer phases. J Chem Soc Faraday Trans 1995;91:2355–2368.
17. Novotny J, Bruccoleri R, Karplus M. An analysis of incorrectly folded protein models. Implications for structure predictions. J Mol Biol 1984;177:787–818.
18. Novotny J, Rashin AA, Bruccoleri RE. Criteria that discriminate between native proteins and incorrectly folded models. Proteins 1988;4:19–30.
19. Sippl MJ. Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structure in globular proteins. J Mol Biol 1990;213:859–883.
20. Sippl MJ. Boltzmann's principle, knowledge-based mean fields and protein folding: an approach to the computational determination of protein structures J Comput Aided Mol Des 1993;7:473–501.
21. Sippl MJ. Knowledge-based potentials for proteins. Curr Opin Struct Biol 1995;5:229–235.
22. Jernigan R, Bahar I. Structure-derived potentials and protein simulation. Curr Opin Struct Biol 1996;6:195–209.
23. Rooman MJ, Wodak SJ. Are database-derived potentials valid for scoring both forward and inverted protein folding? Protein Eng 1995;8:849–858.
24. Godzik A, Kolinski A, Skolnick J. Are proteins ideal mixtures of amino acids? Analysis of energy parameter sets. Protein Sci 1995;4:2107–2117.
25. Miyazawa S, Jernigan R. Residue-residue potentials with a favorable contact pair term and unfavorable high packing density term, for simulation and threading. J Mol Biol 1996;256:623–644.
26. Thomas PD, Dill KA. Statistical potentials extracted from protein structures: how accurate are they. J Mol Biol 1996;257:457–469.
27. Moult J. Comparison of database potentials and molecular mechanics force fields. Curr Opin Struct Biol 1997;7:194–199.
28. Vorobjev YN, Almagro JC, Hermans J. Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that used both dynamics simulations with an explicit solvent and an implicit solvent continuum model. Proteins 1998;32:399–413.
29. Zhang H. Fast and accurate prediction of protein local structures. Ph.D thesis, Peking Univsersity, 1996. p 1–25.
30. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. J Comput Chem 1983;4: 187–217.
31. Kolinski A, Skolnick J. Monte Carlo Simulations of Protein Folding. II. Application to Protein A, ROP Crambin. Proteins 1994;18:353–366.
32. Olszewski K, Kolinski A, Skolnick J. Folding simulations and computer redesign of Protein A three-helix bundle motifs. Proteins 1996;25:286–299.
33. Boczko EM, Brooks III, CL. First-principles calculation of the folding free energy and pathway for a three-helix bundle protein. Science 1995;269:393–396.
34. Holm L, Sander C. Evaluation of protein models by atomic solvation preference. J Mol Biol 1992;225:93–105.