Protein folding using fragment assembly and physical energy function

Seung-Yeon Kim
School of General Education, ChungJu National University, Chungju 380-702, Korea

Weontae Lee
Department of Biochemistry and HTSD-NMR & Application National Research Laboratory, College of Science, Yonsei University, Seoul 120-749, Korea

Julian Lee[a)]
Department of Bioinformatics and Life Science, Soongsil University, Seoul 156-743, Korea

(Received 24 May 2006; accepted 26 September 2006; published online 21 November 2006)

We perform a systematic study of the effects of sequence-independent backbone interactions and sequence-dependent side-chain interactions on protein folding using fragment assembly and physical energy function. Structures for ten proteins belonging to various structural classes are predicted only with Lennard-Jones interaction between backbone atoms. We find nativelike structures for β proteins, suggesting that for proteins in this class, the global tertiary structures can be determined mainly by sequence-independent backbone interactions. On the other hand, for α proteins, nonlocal hydrophobic side-chain interaction is also required to obtain nativelike structures. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2364500]

Understanding folding of a protein from its amino-acid sequence, especially the prediction of the native structure, is a longstanding challenge in theoretical biophysics. The information on the native structure of a protein is quite crucial in understanding its biological role. The most popular methods for protein structure prediction are knowledge-based methods such as comparative modeling and fold recognition.[1,2] In knowledge-based methods, there should exist a sequence with a known structure that is related to the query sequence. When homologous or weakly homologous sequences with known structures are not available, we turn to *ab initio* methods.[2–10] The *ab initio* structure prediction is based on the thermodynamic hypothesis.[11] Therefore, in the *ab initio* prediction method, also called the physics-based method, the native structure of a protein is predicted by obtaining a conformation that minimizes the free energy. Since the physics-based method is based on the fundamental principles of physics, the study of protein folding using this method provides us with a valuable insight into not only the native structure but also the folding mechanism.

Although some progress has been made in physics-based methods, the successful prediction based solely on the physical energy function still remains as a challenging and unsolved problem.[3] Therefore, recent popular trends in *ab initio* methods are to predict the protein structure by assembling fragments (local structures) collected from the protein data bank (PDB), where the global tertiary packing of fragments is determined by energy optimization. Fragment-assembly methods have shown the best performances in *ab initio* structure prediction.[2–8] Since the effect of local interactions are incorporated in fragments, one needs to include only nonlocal interactions in the energy function during fragment assembly. (Local and nonlocal interactions in this work mean interactions between residues near and far in sequence.) We can then perform a systematic study of various factors determining the protein folding: the effect of fragment selection method on correct local structure, and those of various nonlocal interactions, both sequence-independent backbone interactions and sequence-dependent side-chain interactions, on correct global tertiary packing. However, despite the remarkable success of the fragment-assembly method[2–8] on structure prediction, such a systematic analysis has seldom been performed. Moreover, various knowledge-based score terms have been used in earlier works on fragment assembly, where functional forms are obtained by fitting to the distributions of conformations in PDB, with little basis in physics. Not much physical insight can be obtained on the protein folding process from using such *ad hoc* score functions.

In this work, we conduct a systematic study of the effect of the nonlocal energy function on global tertiary structures. We use the minimal number of energy terms, whose functional forms are based mostly on physical considerations. First, we perform fragment assembly only with the Lennard-Jones interaction of the Chemistry at HARvard Molecular Mechanics (CHARMM) force field[12] for backbone atoms. The native structures are predicted for ten proteins in various structural classes, 1L2Y, 1F4I, 1BDD, 1PRB, 1E0L, 1BK2, 1M3B, 1E0G, 1P7E, and 1OQ3. Surprisingly, we find that for β proteins (1E0L, 1BK2, and 1M3B), the Lennard-Jones energy is enough to generate nativelike low-energy conformations. On the other hand, low-energy conformations of α proteins (1L2Y, 1F4I, 1BDD, and 1PRB) and α/β proteins (1E0G, 1P7E and 1OQ3) are less nativelike. The results indicate that for some β proteins (1E0L, 1BK2, 1M3B), the sequence effect plays a major role only in determining the local structures, and the global tertiary structures can be de-

[a)] Author to whom correspondence should be addressed. Electronic mail: jul@ssu.ac.kr

© 2006 American Institute of Physics

Downloaded 21 Nov 2006 to 203.253.28.138. Redistribution subject to AIP license or copyright, see http://jcp.aip.org/jcp/copyright.jsp

termined mainly by sequence-independent interactions, whereas for some proteins (1L2Y, 1F4I, 1BDD, 1PRB, 1E0G, 1P7E, and 1OQ3) containing helices, the nonlocal hydrophobic side-chain interaction is also required for correct tertiary packing of helices. By introducing a contact interaction between side chains that represents an implicit solvent effect, we can retrieve nativelike structures for most of the proteins, suggesting that it might be possible to perform fragment-assembly structure prediction with only a small number of energy terms. Also, one can understand the effects of sequence-dependent and sequence-independent interactions on the protein structure, thus gaining deeper insights into protein folding.

A fragment library for each residue is a set of 20 most probable conformations of the local neighborhood. The fragments are selected from a data set of nonredundant proteins, constructed by clustering the ASTRAL Structural Classification of Proteins (ASTRAL SCOP) set[13] so that no two proteins in the data set have more than 25% sequence identity with each other. The resulting data set consists of 4362 protein chains. The first stage in the fragment assembly is to find protein sequences homologous to the query protein using Position Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST),[14] from a *sequence* database, and to perform multiple alignment of these sequences. The mutation rate for each residue position is obtained, which is called the sequence profile. The sequence profile can be considered as containing evolutionary information that cannot be obtained from the raw sequence. Sequence profiles are precalculated also for proteins in the reference data set. For a given fragment of the query sequence, 20 fragments with similar sequence profiles are selected from the reference data set, using the fuzzy k-nearest-neighbor method.[15] The method is similar to that used in Rosetta,[4] but instead of fragments of size three and nine residues used in Rosetta, we used the size of five, which gave a marginally optimal performance for test predictions on a few proteins. We join fragments only when they share a residue with a common secondary structure and similar values of dihedral angles (difference of less than 15°), and since this residue is used as the junction point, the actual part of the fragment being used in the structure is between one and five residues long. Since the systematic study on the optimal method for fragment selection and assembly is out of the scope of the current work, and postponed to a future study, the method presented here is by no means optimal.

Various energy functions have been devised to be used in fragment-assembly methods. However, most of them have knowledge-based functional forms with little physical justification. Furthermore, possible redundancy between different interaction terms has not been eliminated in a rigorous manner. For example, an energy function consisting of ten components were used in Rosetta,[4] all of them derived by fitting to the distribution of conformations in PDB, rather than being motivated by physics. The energy function used in FRAGFOLD,[5] which contains five (six in the earliest version) components, also consists of similar kinds of knowledge-based potential terms. The energy function used in Solvent Induced Multibody FOrce fieLD (SimFold),[6,7] consisting of seven components, is more physically oriented, but it still contains many knowledge-based terms. Moreover, it contains local interaction terms, which seem to be redundant for the fragment-assembly method where local structures are collected from PDB. Despite excellent performance on protein structure prediction, it is difficult to obtain insights into physical principles underlying protein folding from energy functions in the literature, which contain comparatively large number of terms, most of them having knowledge-based functional forms.

On the other hand, we avoid excessive use of knowledge-based energy terms in this work, and the number of terms is kept to a minimum level in that it contains just one or two components. The energy function is

$$U = \sum_{i<j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) + w_c E_c, \tag{1}$$

where w_c is the weight parameter. The first term is the Lennard-Jones 6–12 van der Waals energy of the CHARMM force field,[12] where r_{ij} is the distance between the ith and jth backbone atoms. This interaction is present only between backbone atoms, since fragments are collected from protein structures with sequences different from the query protein, and hence the model protein has no side-chain atoms except β-carbon (C_β). Therefore this component is a sequence-independent backbone interaction. It should be noted that the Lennard-Jones interaction with CHARMM parameters has a much more firm physical foundation than van der Waals (vdW) type interactions used in earlier works,[6,7] where the functional forms of repulsive and attractive parts were determined from the distribution of conformations in PDB. In particular, the term called vdW interaction in Rosetta[4] contains only the repulsive part, and is not a vdW interaction in the true sense. The second term is a contact interaction between C_β atoms, which represents a hydrophobic interaction between the side chains. This term is included to implicitly incorporate the sequence-dependent interaction between the side chain and solvent. The functional form of the contact term is given by

$$E_c = \sum_{i<j} e_{a_i b_j} f\left(\frac{r_{ij} - r_0}{\Delta} \right), \tag{2}$$

where r_{ij} is the C_β–C_β distance between the ith and jth residues, a_i and b_j are their amino-acid types, $f(x)$ is a smoothed step function,[16] with r_0=6.5 Å and Δ=2.0 Å. The values of e_{ab}, representing relative strengths of hydrophobic interactions between the side chains, are adapted from Ref. 17. In Eq. (1), interaction within five residues are turned off in order to avoid intrafragment interaction.

The number of terms used in Eq. (1) is even less than that of energy function used by current authors in Ref. 8, where various *ad hoc* or redundant terms such as penalty on radius of gyration or hydrogen-bonding term were introduced. It should be noted that in Refs. 6 and 8 the hydrogen-bonding effect is incorporated by adding an explicit interaction term in addition to the vdW type attraction and repulsion terms. In this work, the hydrogen-bonding effect emerges as the result of the Lennard-Jones interaction of the CHARMM

Downloaded 21 Nov 2006 to 203.253.28.138. Redistribution subject to AIP license or copyright, see http://jcp.aip.org/jcp/copyright.jsp

force field, as can seen from the successful prediction of β proteins. On the other hand, adding the Coulomb interaction of the CHARMM force field gives much worse prediction results (data not shown). In fact, the nativelike structures are penalized due to highly repulsive energy values between strands, even when we use the native structure in PDB itself. It might be the case that the Coulomb interaction in the CHARMM force field is not optimized for the energy evaluation of the native structure without local relaxation, and for the energy calculation in the fragment-assembly methods where rigid fragments from PDB are used. By avoiding including redundant terms for the same interaction, and keeping the number of terms to a minimal level, we are able to study the effect of various interactions in a systematic way. In particular, we can easily investigate the relative importance of the sequence-independent backbone interaction and the sequence-dependent side-chain interaction for the global tertiary packing of a protein, by varying just one weight parameter w_c.

The structure predictions by the fragment-assembly method with the energy function in Eq. (1) have been performed for four α proteins (1L2Y, 1F4I, 1BDD, and 1PRB), three β proteins (1E0L, 1BK2, and 1M3B), and three α/β proteins (1E0G, 1P7E, and 1OQ3). The protein 1L2Y has an α helix and a 3_{10} helix, both 1F4Y and 1BDD contain three α helices, and 1PRB consists of three α helices and a 3_{10} helix. 1E0L is a three-stranded antiparallel β-sheet protein, 1BK2 consists of a five-stranded antiparallel β-sheet and a 3_{10} helix, and 1M3B is a five-stranded antiparallel β-sheet protein. The protein 1E0G consists of two α helices, a 3_{10} helix, and a two-stranded antiparallel β-sheet, 1P7E has an α helix and a four-stranded mixed β-sheet, and 1OQ3 consists of two α helices, a 3_{10} helix, a four-stranded antiparallel β-sheet. Conformational sampling is performed by the conformational space annealing (CSA) method.[18] Recently, the CSA method has been applied to the fragment-assembly method,[8] by defining a local minimum-energy conformation as the one whose energy is not minimized by fragment replacements. A detailed explanation on the sampling method can be found in Refs. 8 and 18.

Before the selection of the fragment, we removed those proteins from the reference data set of 4362 protein chains, which are homologous to the query protein, in order to remove any possible bias. We did it by performing a Basic Local Alignment Search Tool (BLAST) search of the query sequence against the data set, and eliminating any protein chain whose local alignments have a sequence identity of 70% or more of the query sequence length, except for the shortest protein 1L2Y, where the sequence homology was allowed up to 70%. The number of homologous chains in the data set after this filtering is shown in Table I, along with percentage sequence identities of the local alignments with respect to the query sequence length.

TABLE I. Proteins homologous to the query sequence in the data set of 4362 protein chains after filtering out proteins that have a homology of 70% (80% in the case of 1L2Y) or more with the query protein. The number of chains is shown for each of 10% ranges of percentage sequence identities. (See text for the definition.)

PDB ID	Residues	10–19	20–29	30–39	40–49	50–59	60–69	70–79
1L2Y	20	0	0	0	0	0	0	1
1F4I	40	0	2	1	0	0	0	0
1BDD	46	1	0	0	2	0	0	0
1PRB	53	2	1	0	0	1	0	0
1E0L	25	0	0	0	0	2	1	0
1BK2	57	5	8	11	2	0	0	0
1M3B	58	6	7	19	0	0	0	0
1E0G	41	2	4	0	0	0	0	0
1P7E	56	0	0	0	0	0	1	0
1OQ3	66	5	3	5	1	0	0	0

Tables II, III, and IV show the structure prediction results for the four α proteins, three β proteins, and three α/β proteins, respectively, where the parameter w_c has been changed from 0.0 to 0.6 in steps of 0.2. For each run, the number of local minimum-energy conformations is kept at 50. The α-carbon root-mean-square deviations (RMSDs) of the global minimum-energy conformations (GMECs), the minimum values of RMSD found among the final 50 low-energy conformations, and the correlations between the energy and RMSD are displayed in the tables.

As shown in Table II, for α proteins, RMSDs between the native structures and the GMECs are large without the contact interaction. When the contact term is added, RMSD values become smaller for most of them, and correlations between energy and RMSD also increase up to a certain point. The results indicate that the sequence-dependent solvation effect is important for global tertiary packing of helices. In the case of 1L2Y, the GMEC is non-native even after the introduction of the contact term, although there are nativelike conformations among the final 50 low-energy conformations. The performance of the prediction improves when we restore the interaction within five residues (num-

TABLE II. Prediction results for α proteins. w_c is the weight parameter for the contact interaction term. RMSD(GM) is the α-carbon RMSD between the native structure and the global minimum-energy conformation. RMSD (lowest) is the lowest value of RMSDs obtained from the 50 local minimum-energy conformations.

PDB ID	Residues	w_c	RMSD (GM)	RMSD (lowest)	Energy-RMSD correlation
1L2Y	20	0.0	5.37(4.14)[a]	2.36(0.93)	0.23(−0.28)
		0.2	5.01(0.95)	1.12(0.95)	0.46(0.47)
		0.4	5.05(1.15)	1.41(1.15)	0.53(0.49)
		0.6	5.04(5.28)	1.40(1.16)	0.38(0.28)
1F4I	40	0.0	2.61	1.76	0.45
		0.2	1.98	1.83	0.63
		0.4	1.99	1.72	0.74
		0.6	2.46	1.91	0.62
1BDD	46	0.0	3.50	2.83	−0.018
		0.2	8.37	2.72	−0.015
		0.4	5.10	2.90	0.13
		0.6	2.62	2.62	0.42
1PRB	53	0.0	4.93	3.15	0.50
		0.2	3.28	3.18	0.61
		0.4	3.02	3.02	0.71
		0.6	4.02	3.24	0.76

[a] The numbers in the parentheses are the results when the interaction within five residues are present (also see the text).

Downloaded 21 Nov 2006 to 203.253.28.138. Redistribution subject to AIP license or copyright, see http://jcp.aip.org/jcp/copyright.jsp

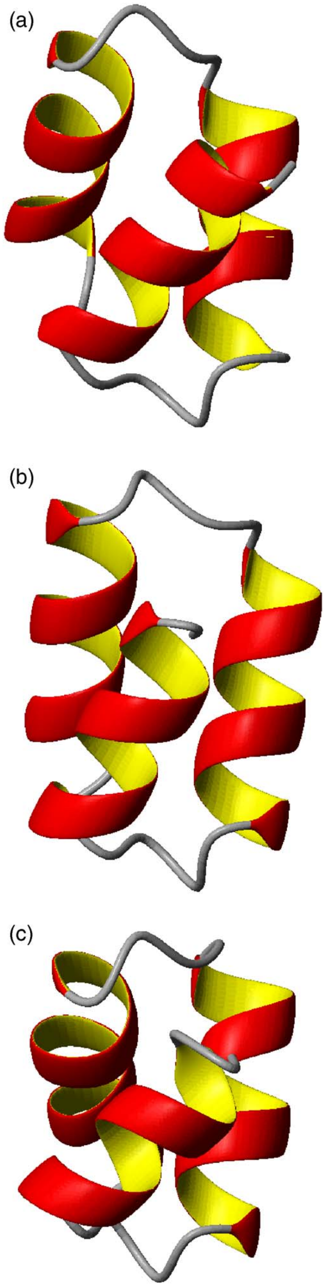

FIG. 1. (a) The native conformation and the global minimum-energy conformations at (b) w_c=0.0 (2.61 Å) and (c) w_c=0.2 (1.98 Å), for the protein 1F4I. The figures are prepared with the program MOLMOL (Ref. 19).

bers in parentheses in Table II). Since the actual size of the fragment being used may be less than five, removal of the interaction within five residues may be too stringent for an extremely short protein such as 1L2Y. Figure 1 shows GMECs of 1F4I for w_c=0.0 (b) and 0.2 (c), along with the native structure (a). At w_c=0.0 GMEC lacks nativelike packing, whereas at w_c=0.2 GMEC is nativelike.

As shown in Table III, for β proteins, GMECs obtained from fragment assembly are close to the native structures even without the contact term. It is amazing that nativelike low-energy conformations can be obtained only with vdW energy for β proteins. Therefore, for some β proteins, the sequence effect plays a role only in selecting appropriate fragments, and the global packing of fragments can be determined mainly by the sequence-independent interaction. Figure 2 shows GMECs of 1BK2 for w_c=0.0 (b) and 0.2 (c), along with the native structure (a). We see that GMEC is nativelike even for w_c=0.0.

TABLE III. Prediction results for β proteins.

PDB ID	Residues	w_c	RMSD (GM)	RMSD (lowest)	Energy-RMSD correlation
1E0L	25	0.0	1.51	1.51	0.42
		0.2	3.82	1.47	0.54
		0.4	5.02	2.72	0.39
		0.6	3.54	2.01	0.31
1BK2	57	0.0	2.35	2.20	0.63
		0.2	2.80	1.77	0.59
		0.4	2.35	2.32	0.40
		0.6	2.98	1.95	0.63
1M3B	58	0.0	2.73	2.09	0.59
		0.2	2.69	2.43	0.42
		0.4	2.65	2.31	0.56
		0.6	3.31	2.57	0.34

As shown in Table III, the prediction results for α/β are somewhat intermediate between the α and the β proteins. For 1E0G and 1P7E the performance of prediction becomes worse with the contact term, whereas for 1OQ3 it becomes better. Figure 3 shows GMECs of 1OQ3 for w_c=0.0 (b) and 0.2 (c), along with the native structure (a). We see that the orientation of helices is different from the native conformation for w_c=0.0 but becomes nativelike for w_c=0.2.

Figure 4 shows the energy landscapes for 1F4I (α protein) (a), 1BK2 (β protein) (b), and 1OQ3 (α/β protein) (c). As the contact term is added, the landscape for 1F4I and 1OQ3 is changed so that there is more correlation between energy and RMSD, whereas the correlation slightly decreases for 1BK2.

Our results indicate that for β proteins, the global tertiary structures can be determined mainly by sequence-independent interactions, whereas for α proteins, a nonlocal hydrophobic side-chain interaction is also required to obtain nativelike structures. The role of sequence-dependent and sequence-independent interactions were also investigated in Ref. 7. Excellent prediction results were obtained for α, β,

TABLE IV. Prediction results for α/β proteins.

PDB ID	Residues	w_c	RMSD (GM)	RMSD (lowest)	Energy-RMSD correlation
1E0G	41	0.0	3.94	2.89	0.12
		0.2	6.96	3.74	−0.28
		0.4	7.53	3.42	−0.52
		0.6	7.10	3.42	−0.31
1P7E	56	0.0	4.96	4.41	0.51
		0.2	6.38	5.24	0.53
		0.4	8.58	4.01	0.67
		0.6	5.52	3.35	0.68
1OQ3	66	0.0	7.43	1.97	0.36
		0.2	2.68	1.73	0.44
		0.4	3.74	2.01	0.57
		0.6	2.81	2.34	0.48

Downloaded 21 Nov 2006 to 203.253.28.138. Redistribution subject to AIP license or copyright, see http://jcp.aip.org/jcp/copyright.jsp

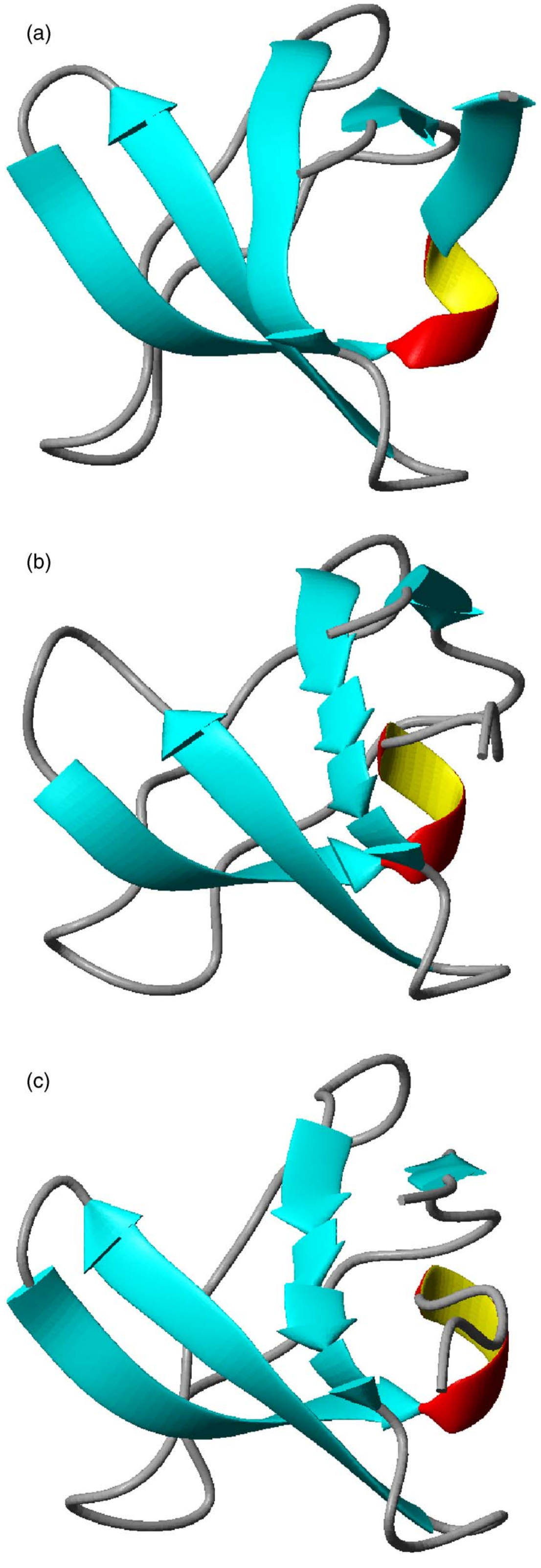

FIG. 2. (a) The native conformation and the global minimum-energy conformations at (b) w_c=0.0 (2.35 Å) and (c) w_c=0.2 (2.80 Å), for the protein 1BK2.

and α/β proteins in the chimera experiments, where artificial sequence-independent hydrophobic side-chain interactions were used in the prediction. This result implies that the sequence effect plays a role only in determining local structures, and the overall tertiary structure is determined mainly by the sequence-independent effect. The results in this work suggest a stronger statement, that for β proteins, protein folding may be performed without a nonlocal hydrophobic side-chain interaction altogether. It is interesting to note that, even in Ref. 7, the performance for the β protein is better than that for the α protein when sequence-independent nonlocal interactions are used.

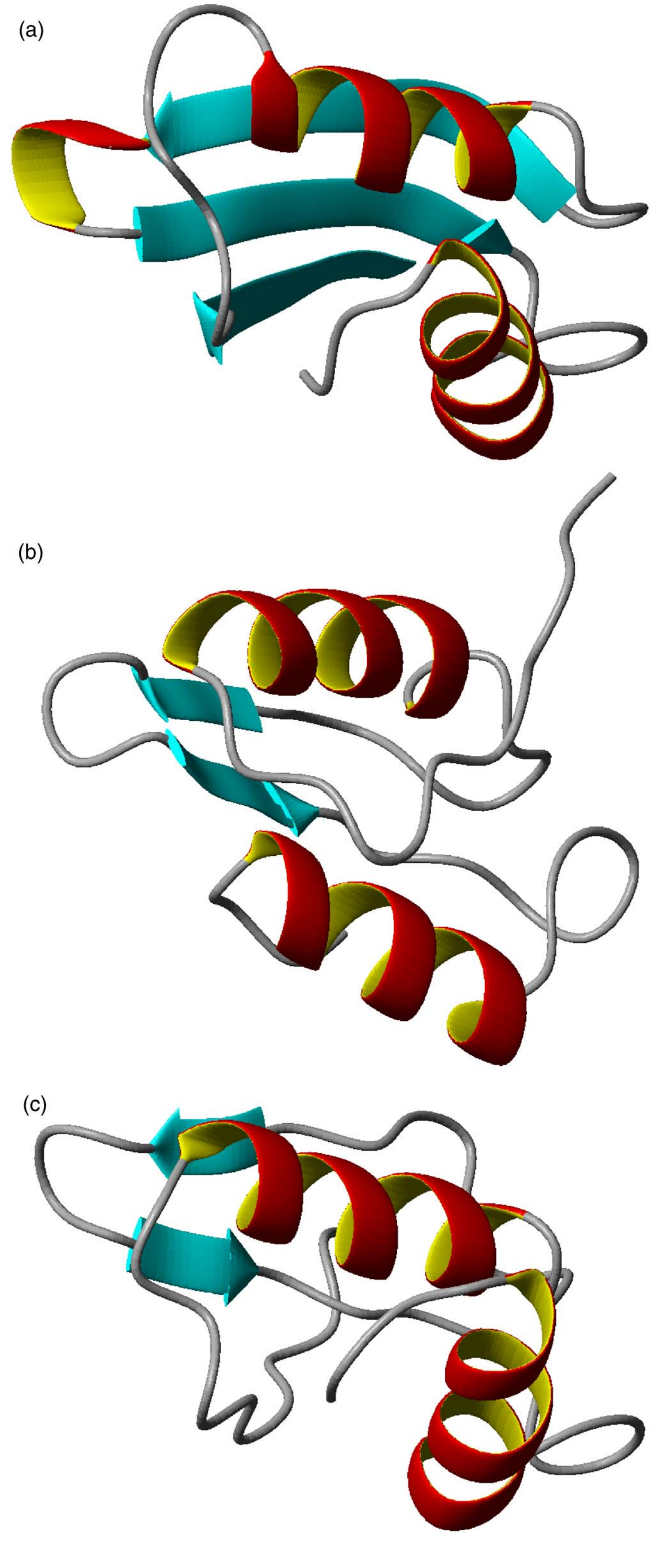

FIG. 3. (a) The native conformation and the global minimum-energy conformations at (b) w_c=0.0 (7.43 Å) and (c) w_c=0.2 (2.68 Å), for the protein 1OQ3.

Downloaded 21 Nov 2006 to 203.253.28.138. Redistribution subject to AIP license or copyright, see http://jcp.aip.org/jcp/copyright.jsp

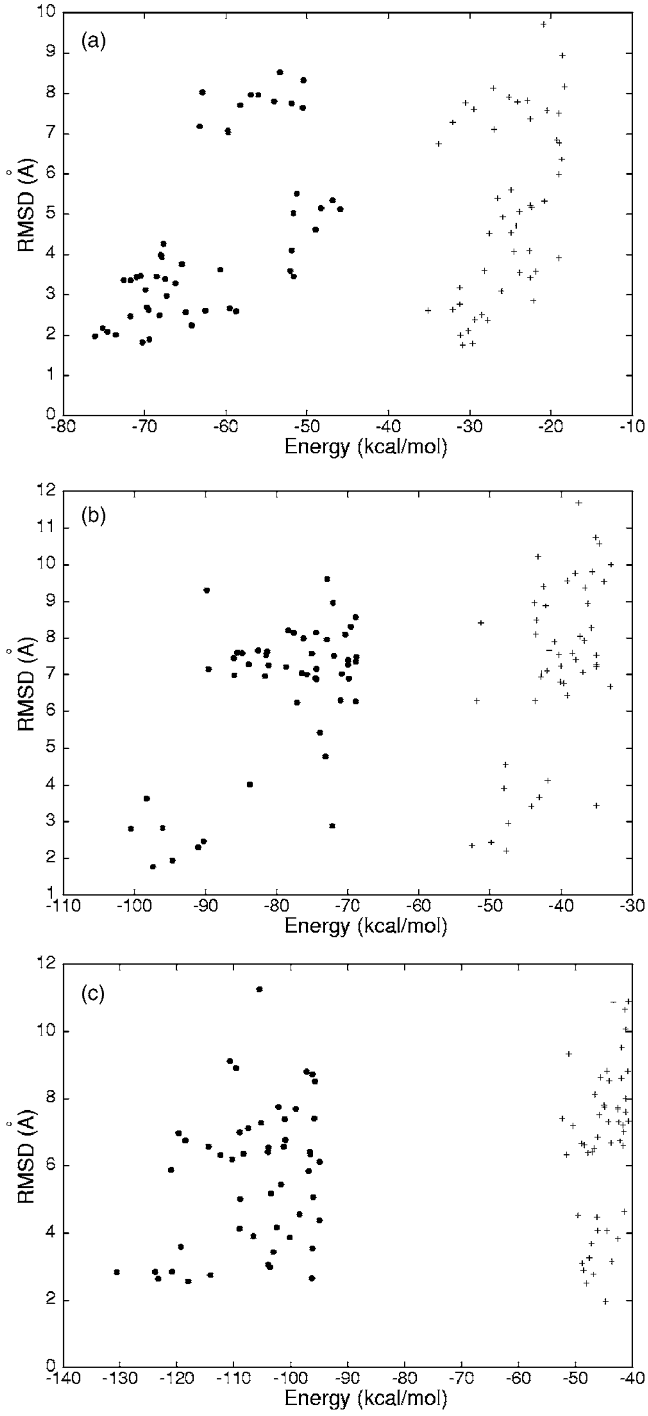

FIG. 4. Plots of energy and RMSD for (a) 1F4I (α protein) and (b) 1BK2 (β protein), and 1OQ3 (α/β protein), for w_c=0.0 (plus symbol) and 0.2 (filled circle). See Tables II–IV for the numerical values of the energy-RMSD correlation.

The major role of sequence-independent interaction in shaping up the free-energy landscape was also demonstrated in Ref. 9, in the context of a purely energy-based method of protein folding.

This work is a systematic study of the effects of sequence-independent backbone interactions and sequence-dependent side-chain interactions on protein folding using fragment assembly and physical energy function. It should be noted that although it would be ideal to find the functional form and the values of parameters for which a blind prediction can be performed for any protein sequence, finding the optimal energy parameter set separately for each of the structural classes would also be useful for developing methods for predicting the protein tertiary structure, when the secondary structure can be determined from experiments such as nuclear magnetic resonance.

This work was supported by the NRL program of MOST NRDP (M1–0203–00–0020). The authors thank Kyoungrim Lee for useful discussions.

[1] J. Moult, K. Fidelis, B. Rost, T.Hubbard, and A.Tramontano, Proteins **61**, 3 (2005).

[2] D. Baker and A. Sali, Science **294**, 93 (2001).

[3] A. M. Lesk, L. Lo Conte, and T. J. Hubbard, Proteins **45**, 98 (2001); P. Aloy, A. Stark, C. Hadley, and R. B. Russel, *ibid.* **53**, 436 (2003); J. J. Vincent, C.-H. Tai, B. K. Sathyanarayana, and B. Lee, *ibid.* **61**, 67 (2005).

[4] K. T. Simons, C. Kooperberg, E. Huang, and D. Baker, J. Mol. Biol. **268**, 209 (1997); C. A. Rohl, C. E. Strauss, K. M. Misura, and D. Baker, Methods Enzymol. **383**, 66 (2004).

[5] D. T. Jones, Proteins **45**, 127 (2001); **61**, 143 (2005).

[6] G. Chikenji, Y. Fujitsuka, and S. Takada, J. Chem. Phys. **119**, 6895 (2003); Y. Fujitsuka, G. Chikenji, and S. Takada, Proteins **62**, 381 (2006).

[7] G. Chikenji, Y. Fujitsuka, and S. Takada, Proc. Natl. Acad. Sci. U.S.A. **103**, 3141 (2006).

[8] J. Lee, S.-Y. Kim, K. Joo, I. Kim, and J. Lee, Proteins **56**, 704 (2004); J. Lee, S.-Y. Kim, and J. Lee, Biophys. Chem. **115**, 209 (2005); J. Korean Phys. Soc. **46**, 707 (2005).

[9] T. X. Hoang, L. Marsella, A. Trovato, F. Seno, J. R. Banavar, and A. Maritan, Proc. Natl. Acad. Sci. U.S.A. **101**, 7960 (2004); **103**, 6883 (2006).

[10] S. Oldziej, C. Czaplewski, A. Liwo *et al.*, Proc. Natl. Acad. Sci. U.S.A. **102**, 7547 (2005).

[11] C. B. Anfinsen, Science **181**, 223 (1973).

[12] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983); A. D. MacKerell, Jr., D. Bashford, M. Bellot *et al.*, J. Phys. Chem. B **102**, 3586 (1998).

[13] S. E. Brenner, P. Koehl, and M. Levitt, Nucleic Acids Res. **28**, 254 (2000).

[14] S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, Nucleic Acids Res. **25**, 3389 (1997).

[15] J. Sim, S.-Y. Kim, and J. Lee, Bioinformatics **21**, 2844 (2005).

[16] The function $f(x)$ is defined by

$$f(x)=\begin{cases}1\ (x\leq -1)\\ 1/2-15/16x+5/8x^3-3/16x^5\ (-1<x<1),\\ 0\ (x\geq 1),\end{cases}$$

where the coefficients were determined to make the function and the derivatives continuous.

[17] S. Miyazawa and R. L. Jernigan, Macromolecules **18**, 534 (1985); J. Mol. Biol. **256**, 623 (1996).

[18] J. Lee, I. H. Lee, and J. Lee, Phys. Rev. Lett. **91**, 080201 (2003) ;S.-Y. Kim, S. J. Lee, and J. Lee, J. Chem. Phys. **119**, 10274 (2003); Phys. Rev. E **72**, 011916 (2005).

[19] R. Koradi, M. Billeter, and K. Wuthrich, J. Mol. Graphics **14**, 51 (1996).

Downloaded 21 Nov 2006 to 203.253.28.138. Redistribution subject to AIP license or copyright, see http://jcp.aip.org/jcp/copyright.jsp