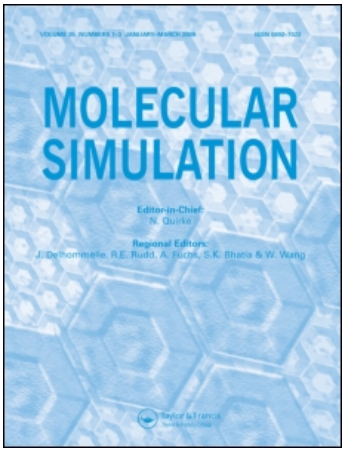

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Determination method of the balance of the secondary-structure-forming tendencies of force fields

Yoshitake Sakae[a]; Yuko Okamoto[a]
[a] Department of Physics, School of Science, Nagoya University, Aichi, Japan

First published on: 25 August 2009

To cite this Article Sakae, Yoshitake and Okamoto, Yuko(2010) 'Determination method of the balance of the secondary-structure-forming tendencies of force fields', Molecular Simulation, 36: 2, 159 — 165, First published on: 25 August 2009 (iFirst)
To link to this Article: DOI: 10.1080/08927020903131143
URL: http://dx.doi.org/10.1080/08927020903131143

Molecular Simulation
Vol. 36, No. 2, February 2010, 159–165

Taylor & Francis
Taylor & Francis Group

Determination method of the balance of the secondary-structure-forming tendencies of force fields

Yoshitake Sakae[1] and Yuko Okamoto*

Department of Physics, School of Science, Nagoya University, Nagoya, Aichi 464-8602, Japan

(*Received 6 April 2009; final version received 17 June 2009*)

We propose a new method of optimisation of backbone torsion-energy parameters in the force field for molecular simulations of protein systems. This method is based on the idea of balancing the secondary-structure-forming tendencies, namely, those of α-helix and β-sheet structures. We perform a minimisation of the backbone dihedral angle-based root-mean-square deviation of the helix and β structure regions in many protein structures. As an example, we optimised the backbone torsion-energy parameters of AMBER parm96 force field using 100 protein molecules from the Protein Data Bank. We then performed folding simulations of α-helical and β-hairpin peptides, using the optimised force field. The results imply that the new force-field parameters give structures more consistent with the experimental implications than the original AMBER parm96 force field.

Keywords: force field; protein; folding simulation; optimisation; secondary structure

1. Introduction

In the area of computational chemistry and biophysics, molecular simulations are often used within the framework of classical mechanics and quantum mechanics. Especially, the molecular simulations for proteins and peptides are performed using force fields, which are based on classical mechanics, defined by potential energy functions with force-field parameters. These force fields, such as AMBER [1–5], CHARMM [6,7], OPLS [8,9], GROMOS [10] and ECEPP [11], are widely used. Generally, the force-field parameters are determined based on experimental results for small molecules and theoretical results using quantum chemistry calculations of small peptides such as alanine dipeptide.

These force fields have been used for the protein structure prediction and the protein folding study. For some small proteins, such as a villin headpiece, protein A, Trp-cage and albumin binding domain, the folding simulations were performed, and the results showed good agreement with the experimental results [12–17].

Comparisons of six force fields by performing generalised-ensemble simulations [18] of two small peptides in explicit solvent [19,20] showed that the force fields have quite different secondary-structure-forming tendencies. For some force fields, their results and our simulation results [21–23] using the generalised Born/solvent-accessible surface area (GB/SA) model [24,25] give essentially the same secondary-structure-forming tendencies.

We believe that one reason for these discrepancies is due to the difference in backbone conformational preferences, such as α-helix structure and β-sheet structure, of force fields. For example, AMBER parm94 [1] and AMBER parm96 [2] have very different behaviours about the secondary-structure-forming tendencies, although these force fields differ only in the backbone torsion-energy terms for rotations of the backbone ϕ and ψ angles. Recently, new force-field parameters of the backbone torsion-energy term about ϕ and ψ angles have been developed, which are, for example, AMBER ff99SB [4], AMBER ff03 [5] and CHARMM 22/CMAP [7]. Additionally, it was proposed to set the backbone torsion-energy term simply to zero [26].

Therefore, we also focus on the backbone torsion-energy terms and propose a new optimisation method of the force-field parameters. This method optimises the balance between helix-structure-forming tendency and β-sheet-forming tendency using 100 protein molecules from the Protein Data Bank. As one of the optimisation methods of force-field parameters using tertiary structures of proteins, there is the method using *Z*-score that is the energy difference between average decoy structure and native structure in units of standard deviation [27,28]. However, this method needs many decoy structures, which are not known in general. On the other hand, our method uses only native structures. Namely, by the minimisation of the backbone torsion-angle root-mean-square (RMS) deviation of the helix and β-sheet structure regions

*Corresponding author. Email: okamoto@phys.nagoya-u.ac.jp

ISSN 0892-7022 print/ISSN 1029-0435 online

DOI: 10.1080/08927020903131143
http://www.informaworld.com

in native structures, we determined new force-field parameters.

In this paper, we first give the details of the new optimisation method of force-field parameters. We applied our method to AMBER parm96. We think that AMBER parm96 is a good force field to apply our optimisation method because this force field has the ability to change the secondary-structure-forming tendencies by the backbone torsion-energy term [29], and for this force field, successful results of folding simulations for some small proteins have recently been reported [30].

In Section 2, the details of our new optimisation method are given. In Section 3, the results of applications of the method to AMBER parm96 and those of folding simulations of two peptides are presented. Section 4 is devoted to conclusions.

2. Methods

2.1 Force-field parameters

The existing force fields for protein systems such as AMBER, CHARMM and OPLS use essentially the same functional forms for the potential energy E_{conf} except for minor differences. The conformational potential energy E_{conf} (in kcal/mol) can be written as, for instance,

$$E_{\text{conf}} = E_{\text{BL}} + E_{\text{BA}} + E_{\text{torsion}} + E_{\text{non-bond}}. \quad (1)$$

Here, E_{BL}, E_{BA}, E_{torsion} and $E_{\text{non-bond}}$ represent the bond-stretching term, the bond-bending term, the torsion-energy term and the non-bonded energy term, respectively. Each force field has similar but slightly different parameter values. For example, the torsion energy is usually given by

$$E_{\text{torsion}} = \sum_{\text{dihedral angle }\Phi} \sum_{n} \frac{V_n}{2}[1 + \cos(n\Phi - \gamma_n)], \quad (2)$$

where the first summation is taken over all dihedral angles Φ (both in the backbone and in the side chains), n is the number of waves, γ_n is the phase and V_n is the Fourier coefficient. Namely, the energy term E_{torsion} has n, γ_n and V_n as force-field parameters.

2.2 Optimisation method of force-field parameters

We now describe our new method for optimising the force-field parameters. We first select proteins from the Protein Data Bank. We try to choose proteins from folds (such as all α-helix, all β-sheet, α/β, etc.) and different homology classes as much as possible. If the force-field parameters are of ideal values, we expect that all the chosen native structures minimised by the ideal force field do not change. Namely, we believe that force-field parameters are better, if they have lower deviations obtained from minimisations of protein structures. Hence, we expect

$$\Phi\text{RMSD} = 0, \quad (3)$$

where

$$\Phi\text{RMSD} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\Phi_i^{\text{native}} - \Phi_i^{\text{min}})^2}. \quad (4)$$

Here, n is the total number of backbone dihedral angles (ϕ and ψ angles) in all molecules, Φ_i^{native} is the ith backbone dihedral angle of the native structures and Φ_i^{min} is the corresponding ith backbone dihedral angle of the minimised structures using the trial force-field parameters. In reality, $\Phi\text{RMSD} \neq 0$, because $\Phi\text{RMSD} \geq 0$, we expect that we can optimise the force-field parameters by minimising ΦRMSD with respect to these force-field parameters. In practice, we perform a simulation in the force-field parameter space for this minimisation.

However, our first aim is to determine the balance of secondary-structure-forming tendencies such as helix structure and β-sheet structure. Additionally, it is difficult to perform the minimisation of ΦRMSD in wider force-field parameter space until ΦRMSD is close to zero because of the computational cost. Therefore, we only focus on secondary-structure regions of helix structure and β-sheet structure in the amino-acid sequence. Namely, we only consider the backbone dihedral angles of residue in the native structure which are identified by the DSSP program [31] that they constitute one of α-helix, 3/10-helix, π-helix and β-sheet structures. We calculate two kinds of ΦRMSD for secondary structures, namely, $\Phi\text{RMSD}_{\text{helix}}$ and ΦRMSD_{β}. Here, $\Phi\text{RMSD}_{\text{helix}}$ stands for ΦRMSD of backbone dihedral angles of residues which have helix structures in the native structures, and ΦRMSD_{β} means that of only β-sheet structure in the native structures. Using these two ΦRMSDs, we want to optimise the torsion-energy parameters, which will have better balance of secondary-structure-forming tendencies. We propose the following combination:

$$\Phi\text{RMSD}_{\text{2ndly}} = \lambda\Phi\text{RMSD}_{\text{helix}} + \Phi\text{RMSD}_{\beta}, \quad (5)$$

where we have introduced a fixed scaling factor λ, $\Phi\text{RMSD}_{\text{helix}}$ and ΦRMSD_{β}.

Finally, by minimising $\Phi\text{RMSD}_{\text{2ndly}}$ with respect to the force-field parameters, we can obtain the optimised force-field parameters.

3. Results and discussion

3.1 Application of the optimisation method

We now present the results of the applications of our new optimisation method of force-field parameters.

Table 1. One hundred proteins used in the optimisation of force-field parameters.

Fold	PDB ID	Chain	PDB ID	Chain	PDB ID	Chain	PDB ID	Chain
All α	1DLW	A	1ERL	A	1UTG	A	2END	A
	1HBK	A	1TX4	A	1V54	E	1OR7	C
	1SK7	A	1POC	A	1G8Q	A	1NG6	A
	1DVO	A	1HFE	S	1J0P	A	1W53	A
	1Y02	A71–114	1IJY	A	1I2T	A	2LIS	A
	1G8E	A	1VKE	C	1FS1	A109–149	1S0P	A
	1S7Z	A	1AIL	A	1Q5Z	A	1UPT	H
	1Y9I	A						
All β	1CDC	A	1T2W	A	1GMU	C1	1AMM	A1
	1WJX	A	1NLQ	C	1BEH	A	1UA8	A
	1UXZ	A	1UB4	C	1OW1	A	1R75	A
	1PM4	A	1OU8	A	1V76	A	1UT7	B
	1OA8	D	1IFG	A				
α/β	1IO0	A	1U7P	A	1JKE	C	1MXI	A
	1LY1	A	1WKC	A	1IM5	A	1VC1	A
	1OGD	A	1T6T	2	1PYO	D	1MUG	A
	1J3A	A	1DZ3	A	1COZ	A	1D4O	A
α+β	1LNI	B	1PP0	B	1PZ4	A	1TU1	A
	1Q2Y	A	1SVY	A	1N9L	A	1LQV	B
	1A3A	A	1K2E	A	1TT8	A	1HUF	A
	1SXR	A	1CYO	A	1ID0	A	1UCD	A
	1F46	B	1KPF	A	1BYR	A	1Y60	D
	1SEI	A	1RL6	A	1WM3	A	1FTH	A
	1APY	B	1I7B	B	1LTS	C	1UGI	A
	1MWP	A	1PCF	A	1J98	A	1H6H	A
	1KAF	A	1JID	A	1JYO	A	1E87	A
	1MBY	A						

At first, we chose 100 PDB files with resolution 2.0 Å or better, with sequence similarity of amino acid 30.0% or lower, and with less than 200 residues (the average number of residues is 122.2) from PDB-REPRDB [32]. We selected the number of each fold (all α, all β, α/β and α+β) in 100 proteins based on the number of folds given by SCOP (version 1.73 at November 2007) [33]. Namely, we used 29 all α, 18 all β, 16 α/β and 37 (α+β) proteins (see Table 1).

The force field that we optimised is the AMBER parm96 version [2]. The backbone torsion-energy term $E_{\text{torsion}}(\Phi, \Psi)$ for this force field is given by

$$E_{\text{torsion}}(\Phi,\Psi) = \frac{V_1^{\phi}}{2}[1+\cos\phi] + \frac{V_2^{\phi}}{2}[1-\cos 2\phi] + \frac{V_1^{\psi}}{2}[1+\cos\psi] + \frac{V_2^{\psi}}{2}[1-\cos 2\psi], \quad (6)$$

where we have $V_1^{\phi} = 1.7$, $V_2^{\phi} = 0.6$, $V_1^{\psi} = 1.7$ and $V_2^{\psi} = 0.6$. Here, we have optimised only two parameters in the backbone torsion-energy term, namely, V_1^{ψ} and V_2^{ψ} for ψ angle. As described above, AMBER parm94 and AMBER parm96 have quite different secondary-structure-forming-tendencies, although these force fields differ only in the backbone torsion-energy terms for rotations of the ϕ and ψ angles. Moreover, we can easily imagine that force-field parameters V_1^{ψ} and V_2^{ψ} for ψ angle are important for the secondary-structure-forming-tendencies, because the energy surface in the Ramachandran space is quite sensitive to this energy term in the helix and β-sheet regions. Namely, if the torsion-energy term for the ψ angle changes, the stabilities of helix structure region and β-sheet region on the Ramachandran space change. Therefore, we considered some trial force-field parameters for V_1^{ψ} and V_2^{ψ}, which are given by the following equations:

$$V_1^{\text{trial}} = 1.7 \times 0.2i = 0.34i, \quad (7)$$

$$V_2^{\text{trial}} = 0.6 \times 0.2i = 0.12i. \quad (8)$$

Here, i is any real number. When i is 5, the force-field parameters V_1^{trial} and V_2^{trial} of ψ angle are equal to those of the original AMBER parm96. From our experience, if i has a small number ($i < 5$), the force field favours helix structure, and if i has a large number ($i > 5$), the force field favours β-sheet structure (see also Figures 2 and 3). We calculated $\Phi\text{RMSD}_{2\text{ndly}}$ values in Equation (5) about some trial force-field parameters obtained by changing i in Equations (7) and (8).

We performed the minimisation, which was terminated when the RMS potential energy gradients were less than 0.1 (kcal/mol/Å) by using TINKER program package [34]. For solvent effects, we used GB/SA solvent model in TINKER.

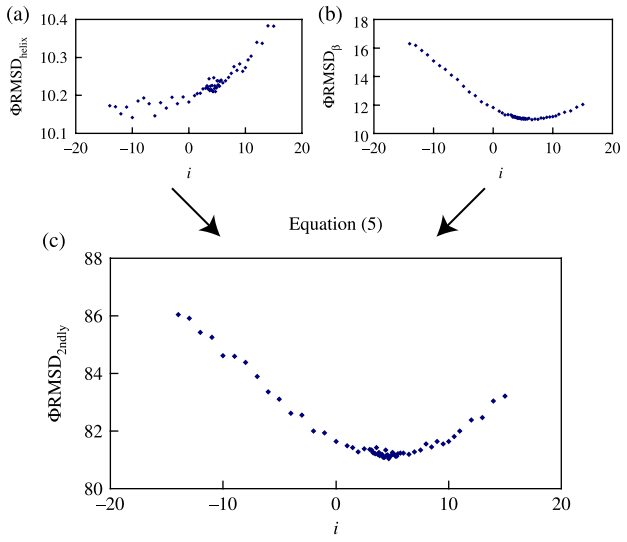

Figure 1. Distributions of (a) ΦRMSD$_{helix}$, (b) ΦRMSD$_{\beta}$ and (c) ΦRMSD$_{2ndly}$ obtained from the minimisation of 100 proteins using the trial force-field parameters V_1^{trial} and V_2^{trial} depending on the number i.

The results of ΦRMSD$_{helix}$ and ΦRMSD$_{\beta}$ are shown in Figure 1(a),(b), respectively. In these calculations, if the differences of the backbone dihedral angles between Φ_i^{native} and Φ_i^{min} in Equation (4) are more than 30°, they were ignored, assuming that the uncertainties in those angles are too large. We see that ΦRMSD$_{helix}$ decreases gradually with a decrease in i. If i decreases, the torsion energy of the helix structure region in the Ramachandran space also decreases. On the other hand, ΦRMSD$_{\beta}$ decreases gradually with an increase in i. If i increases, the torsion energy of the β structure region in the Ramachandran space decreases. Hence, this result is reasonable. However, ΦRMSD$_{\beta}$ reaches the global minimum, when i is 6.5. If i is larger than 6.5, ΦRMSD$_{\beta}$ increases gradually. This result implies that the ΦRMSD$_{\beta}$ does not correspond to the parameters V_1^{trial} and V_2^{trial} completely.

For ΦRMSD$_{helix}$ and ΦRMSD$_{\beta}$ in Figure 1(a),(b), we can see the difference clearly. The noteworthy point obtained from these results is that ΦRMSD can distinguish between helix structure and β structure.

Next, we combined ΦRMSD$_{helix}$ and ΦRMSD$_{\beta}$ by Equation (5). Here, in order to have roughly equal contributions from both terms, we can set the value of the scaling factor λ to be, for example, the coefficients of variations:

$$\lambda = \frac{\sigma_\beta / \mu_\beta}{\sigma_{helix} / \mu_{helix}}. \quad (9)$$

Here, μ_{helix} and μ_β are the averages and σ_{helix} and σ_β are the corresponding standard deviations for ΦRMSD$_{helix}$ and ΦRMSD$_{\beta}$. For the calculations, we have chosen a small number of i values in a range $i_{min} \le i \le i_{max}$. For $i_{min} = 0$ and $i_{max} = 10$, we obtained $\lambda = 6.857$, and this fixed value was used for all the calculations in the present work.

In Figure 1(c), the combined result is shown. The smallest ΦRMSD$_{2ndly}$ is obtained for the value $i = 4.7$, namely, the obtained force-field parameters are $V_1^{trial} = 1.598$ and $V_2^{trial} = 0.564$. These values are slightly smaller than those of the original AMBER parm96, which corresponds to $i = 5$. We can easily expect the new obtained force-field parameters to slightly favour helix structure more and β-sheet structure less than the original AMBER parm96.

3.2 Tests by folding simulations

In order to check the force-field parameters obtained by our optimisation method, we performed the folding simulations using two peptides, namely, C-peptide of ribonuclease A and the C-terminal fragment of the B1 domain of streptococcal protein G, which is sometimes referred to as G-peptide [35]. The C-peptide has 13 residues and its amino-acid sequence is Lys-Glu-Thr-Ala-Ala-Ala-Lys-Phe-Glu-Arg-Gln-His-Met. This peptide has been extensively studied by experiments and is known to form an α-helix structure [36,37]. Because the charges at peptide termini are known to affect helix stability [36,37], we blocked the termini by a neutral $COCH_3$— group and a neutral —NH_2 group. The G-peptide has 16 residues and its amino-acid sequence is Gly-Glu-Trp-Thr-Tyr-Asp-Asp-Ala-Thr-Lys-Thr-Phe-Thr-Val-Thr-Glu. The termini were kept as the usual zwitter ionic states, following the experimental conditions [35,38,39]. This peptide is known to form a β-hairpin structure by experiments [35,38,39].

For the folding simulations, we used replica-exchange molecular dynamics (REMD) [40]. REMD is one of the generalised-ensemble simulation methods, and has high conformational sampling efficiency by allowing configurations to heat up and cool down while maintaining proper Boltzmann distributions. We used the TINKER program package [34] modified by us for the folding simulations. The unit time step was set to 1.0 fs. Each simulation was carried out for 2 ns (hence, it consisted of 2,000,000 MD steps) with 16 replicas for 10 times. The temperature during MD simulations was controlled by Berendsen's method [41]. For each replica the temperature was used exponentially: 700, 662, 625, 591, 558, 528, 499, 471, 446, 421, 398, 376, 355, 336, 317 and 300 K. As for solvent effects, we used the GB/SA model [24,25] included in the TINKER program package [34]. These folding simulations were performed with different sets of randomly generated initial velocities.

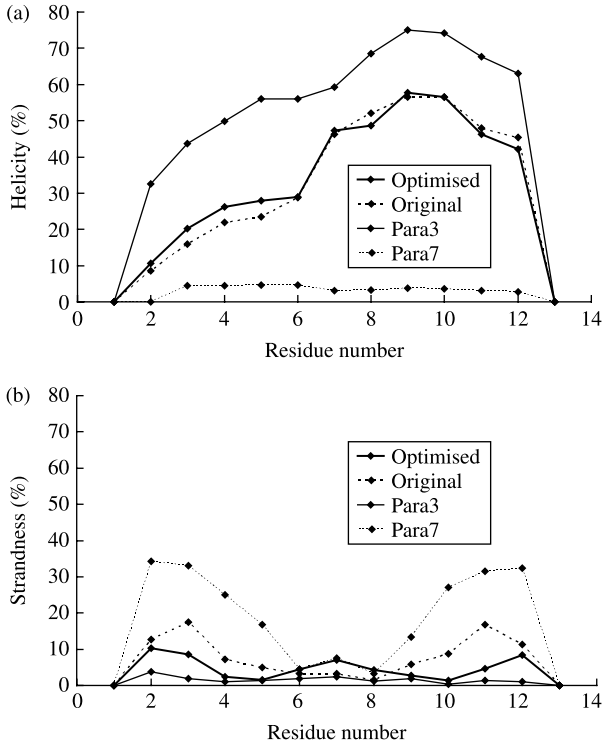

Figure 2. (a) Helicity and (b) strandness of C-peptide as functions of the residue number. These values are the averages of the 10 independent REMD [40] simulations at 300 K. Optimised, original, para3 and para7 stand for the optimised AMBER parm96 ($i = 4.7$), original AMBER parm96 ($i = 5.0$), trial force field para3 ($i = 3.0$) and trial force field para7 ($i = 7.0$), respectively.

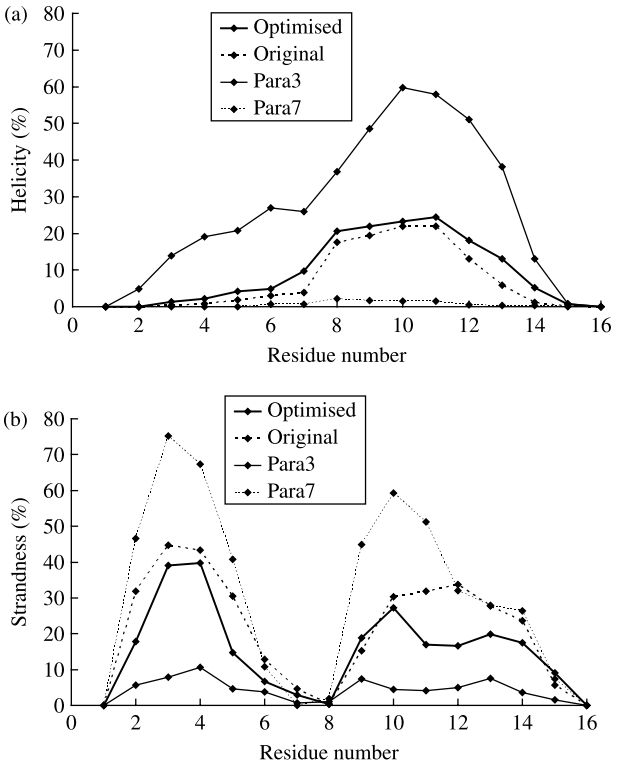

Figure 3. (a) Helicity and (b) strandness of G-peptide as functions of the residue number. These values are the averages of the 10 REMD [40] simulations at 300 K. Optimised, original, para3 and para7 stand for the optimised AMBER parm96 ($i = 4.7$), original AMBER parm96 ($i = 5.0$), trial force field para3 ($i = 3.0$) and trial force field para7 ($i = 7.0$), respectively.

In Figure 2, the helicity and strandness of C-peptide which were obtained with the original AMBER parm96 and its optimised force field are shown. These values are the averages of the 10 REMD simulations at 300 K. In comparison with the helicity of the original AMBER parm96, the helicity of the optimised force field is similar. However, the helicity of Thr3, Ala4 and Ala5 of the optimised force field slightly increases. In comparison with the strandness of the original AMBER parm96, the strandness of the optimised force field decreases except for those at Ala6, Lys7 and Phe8.

In Figure 3, the helicity and strandness of G-peptide at the original AMBER parm96 and its optimised force field are shown. In comparison with the helicity of the original AMBER parm96, the helicity of the optimised force field slightly increases, and in comparison with the strandness of the original AMBER parm96, the strandness of the optimised force field slightly decreases. For trial force fields of para3 and para7, the secondary-structure-forming-tendencies are similar to the case of C-peptide.

These results clearly show that the optimised force field favours helix structures and does not favour β structures in comparison with the original AMBER parm96. We can see that these secondary-structure-forming-tendencies of the optimised force field are better than those of the original AMBER parm96, because it is known that the AMBER parm96 slightly favours the β structure too much [19–23].

We also performed the folding simulations with two extreme cases of the trial force fields, namely, para3 ($i = 3.0$) and para7 ($i = 7.0$) (see Figures 2 and 3) for comparisons. The trial force field para3 favours helix structure strongly and does not favour β structure clearly. On the other hand, the trial force field para7 has the tendency that is quite reverse to para3. According to the results of $\Phi\text{RMSD}_{\text{helix}}$ and ΦRMSD_{β} in Figure 1(a),(b), $\Phi\text{RMSD}_{\text{helix}}$ decreases gradually with a decrease in i, and ΦRMSD_{β} reaches the global minimum, when i is 6.5. Namely, we can see that the values of $\Phi\text{RMSD}_{\text{helix}}$ and ΦRMSD_{β} are related to the stabilities of helix structure and β structure well.

4. Conclusions

In this paper, we proposed the new optimisation method of force-field parameters. This method can optimise

force-field parameters using only PDB structures, and determine the balance of the secondary-structure-forming tendencies, such as α-helix and β-sheet structures, for molecular simulations of protein systems. We applied this optimisation method to the AMBER parm96 using 100 protein molecules from the Protein Data Bank. We then performed folding simulations of α-helical and β-hairpin peptides. We found that the helicity of the optimised force field slightly increases, and the strandness of the optimised force field slightly decreases in comparison with those of the original AMBER parm96. The results imply that the optimised force-field parameters give structures more consistent with the experimental implications than the original AMBER parm96 force field.

We have shown that we can control the secondary-structure-forming tendencies by adjusting the backbone torsion-energy term in Ψ only for AMBER parm96. This finding greatly simplifies the refinement and optimisation of force-field parameter that will yield correct secondary-structure-forming tendencies in agreement with experiments.

Acknowledgements

The computations were performed on SuperNOVA of Miki Laboratory at Doshisha University and the computers at the Research Center for Computational Science, Institute for Molecular Science. This work was supported, in part, by the Grants-in-Aid for the Academic Frontier Project, 'Intelligent Information Science', for the Nano-Bio-IT Education Program, for Scientific Research on Innovative Areas ('Fluctuations and Biological Functions'), and for the Next Generation Super Computing Project, Nanoscience Program from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

Note

1. Email: sakae@tb.phys.nagoya-u.ac.jp.

References

[1] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, J. Kenneth, M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P.A. Kollman, *A second generation force field for the simulation of proteins, nucleic acids, and organic molecules*, J. Am. Chem. Soc. 117 (1995), pp. 5179–5197.

[2] P.A. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, and A. Pohorille, *The development/application of a 'minimalist' organic/biochemical molecular mechanic force field using a combination of* ab initio *calculations and experimental data*, in *Computer Simulations of Biological Systems*, W.F. van Gunsteren, ed., ESCOM, Dordrecht, 1997, pp. 83–96.

[3] J. Wang, P. Cieplak, and P.A. Kollman, *How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?*, J. Comput. Chem. 21 (2000), pp. 1049–1074.

[4] V. Hornak, A. Abel, R. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, *Comparison of multiple Amber force fields and development of improved protein backbone parameters*, Proteins 65 (2006), pp. 712–725.

[5] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, and T. Lee, *A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations*, J. Comput. Chem. 24 (2003), pp. 1999–2012.

[6] A.D. MacKerell, Jr, D. Bashford, M. Bellott, R.L. Dunbrack, Jr, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, *All-atom empirical potential for molecular modeling and dynamics studies of proteins*, J. Phys. Chem. B 102 (1998), pp. 3586–3616.

[7] A. MacKerell, Jr, M. Feig, and C. Brooks, III, *Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations*, J. Comput. Chem. 25 (2004), pp. 1400–1415.

[8] W.L. Jorgensen, D.S. Maxwell, and J. Tirado-Rives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids*, J. Am. Chem. Soc. 118 (1996), pp. 11225–11236.

[9] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, and W.L. Jorgensen, *Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides*, J. Phys. Chem. B 105 (2001), pp. 6474–6487.

[10] W.F. Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, and I.G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Vdf Hochschulverlag AG an der ETH Zürich, Zürich, 1996.

[11] G. Némethy, K.D. Gibson, K.A. Palmer, C.N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, and H. Scheraga, *Energy parameters in polypeptides. 10. Improved geometrical parameters and nonbonded interactions for use in the ECEPP/3 algorithm, with application to proline-containing peptides*, J. Phys. Chem. 96 (1992), pp. 6472–6484.

[12] Y. Duan and P.A. Kollman, *Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution*, Science 282 (1998), pp. 740–744.

[13] B. Zagrovic, C.D. Snow, M.R. Shirts, and V.S. Pande, *Simulation of folding of a small alpha-helical protein in atomistic detail using worldwide-distributed computing*, J. Mol. Biol. 323 (2002), pp. 927–937.

[14] J.A. Vila, D.R. Ripoll, and H.A. Scheraga, *Atomically detailed folding simulation of the B domain of staphylococcal protein A from random structures*, Proc. Natl Acad. Sci. USA 100 (2003), pp. 14812–14816.

[15] C. Simmerling, B. Strockbine, and A.E. Roitberg, *All-atom structure prediction and folding simulations of a stable protein*, J. Am. Chem. Soc. 124 (2002), pp. 11258–11259.

[16] J.W. Pitera and W. Swope, *Understanding folding and design: replica-exchange simulations of 'Trp-cage' miniproteins*, Proc. Natl Acad. Sci. USA 100 (2003), pp. 7587–7592.

[17] H. Lei and Y. Duan, Ab initio *folding of albumin binding domain from all-atom molecular dynamics simulation*, J. Phys. Chem. B 111 (2007), pp. 5458–5463.

[18] A. Mitsutake, Y. Sugita, and Y. Okamoto, *Generalized-ensemble algorithms for molecular simulations of biopolymers*, Biopolymers (Pept. Sci.) 60 (2001), pp. 96–123.

[19] T. Yoda, Y. Sugita, and Y. Okamoto, *Comparisons of force fields for proteins by generalized-ensemble simulations*, Chem. Phys. Lett. 386 (2004), pp. 460–467.

[20] T. Yoda, Y. Sugita, and Y. Okamoto, *Secondary-structure preferences of force fields for proteins evaluated by generalized-ensemble simulations*, Chem. Phys. 307 (2004), pp. 269–283.

[21] Y. Sakae and Y. Okamoto, *Optimization of protein force-field parameters with the Protein Data Bank*, Chem. Phys. Lett. 382 (2003), pp. 626–636.

[22] Y. Sakae and Y. Okamoto, *Protein force-field parameters optimized with the Protein Data Bank. I. Force-field optimizations*, J. Theor. Comput. Chem. 3 (2004), pp. 339–358.

[23] Y. Sakae and Y. Okamoto, *Protein force-field parameters optimized with the Protein Data Bank. II. Comparisons of force fields by folding simulations of short peptides*, J. Theor. Comput. Chem. 3 (2004), pp. 359–378.
[24] W.C. Still, A. Tempczyk, R.C. Hawley, and T. Hendrickson, *Semianalytical treatment of solvation for molecular mechanics and dynamics*, J. Am. Chem. Soc. 112 (1990), pp. 6127–6129.
[25] D. Qiu, P.S. Shenkin, F.P. Hollinger, and W.C. Still, *The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii*, J. Phys. Chem. A 101 (1990), pp. 3005–3014.
[26] A.E. Garcia and K.Y. Sanbonmatsu, *α-Helical stabilization by side chain shielding of backbone hydrogen bonds*, Proc. Natl Acad. USA 99 (2002), pp. 2782–2787.
[27] R. Goldstein, Z.A. Luthey-Schulten, and P.G. Wolynes, *Optimal protein-folding codes from spin-glass theory*, Proc. Natl Acad. Sci. USA 89 (1992), pp. 4918–4922.
[28] M. Hao and H.A. Scheraga, *Optimizing potential functions for protein folding*, J. Phys. Chem. 100 (1996), pp. 14540–14548.
[29] Y. Sakae and Y. Okamoto, *Secondary-structure design of proteins by a backbone torsion energy*, J. Phys. Soc. Jpn 75 (2006), 054802.
[30] M.S. Shell, R. Ritterson, and K.A. Dill, *A test on peptide stability of AMBER force fields with implicit solvation*, J. Phys. Chem. B 112 (2008), pp. 6878–6886.
[31] W. Kabsch and C. Sander, *Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features*, Biopolymers 22 (1983), pp. 2577–2637.
[32] T. Noguchi, K. Onizuka, Y. Akiyama, and M. Saito, *PDB-REPRDB: A database of representative protein chains in PDB (Protein Data Bank)*, in *Proceedings of the Fifth International Conference on Intelligent Systems for Molecular Biology*, AAAI Press, Menlo Park, CA, 1997, pp. 214–217.
[33] M. Levitt and C. Chothia, *Structural patterns in globular proteins*, Nature 261 (1976), pp. 552–558.
[34] TINKER program package, software available at http://dasher.wustl.edu/tinker/.
[35] S. Honda, N. Kobayashi, and E. Munekata, *Thermodynamics of a α-hairpin structure: evidence for cooperative formation of folding nucleus*, J. Mol. Biol. 295 (2000), pp. 269–278.
[36] K.R. Shoemaker, P.S. Kim, D.N. Brems, S. Marqusee, E.J. York, I.M. Chaiken, J.M. Stewart, and R.L. Baldwin, *Nature of the charged-group effect on the stability of the C-peptide helix*, Proc. Natl Acad. Sci. USA 82 (1985), pp. 2349–2353.
[37] J.J. Osterhout, Jr, R.L. Baldwin, E.J. York, J.M. Stewart, H.J. Dyson, and P.E. Wright, *1H NMR studies of the solution conformations of an analogue of the C-peptide of ribonuclease A*, Biochemistry 28 (1989), pp. 7059–7064.
[38] F.J. Blanco, G. Rivas, and L. Serrano, *A short linear peptide that folds into a native stable bold beta-hairpin in aqueous solution*, Nature Struct. Biol. 1 (1994), pp. 584–590.
[39] N. Kobayashi, S. Honda, H. Yoshii, H. Uedaira, and E. Munekata, *Complement assembly of two fragments of the streptococcal protein G B1 domain in aqueous solution*, FEBS Lett. 366 (1995), pp. 99–103.
[40] Y. Sugita and Y. Okamoto, *Replica-exchange molecular dynamics method for protein folding*, Chem. Phys. Lett. 314 (1999), pp. 141–151.
[41] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, and J.R. Haak, *Molecular dynamics with coupling to an external bath*, J. Chem. Phys. 81 (1984), pp. 3684–3690.