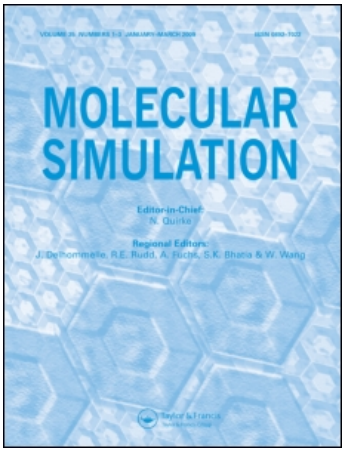

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Controlling the secondary-structure-forming tendencies of proteins by a backbone torsion-energy term

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

First published on: 27 August 2009

To cite this Article Sakae, Yoshitake and Okamoto, Yuko(2010) 'Controlling the secondary-structure-forming tendencies of proteins by a backbone torsion-energy term', Molecular Simulation, 36: 2, 138 — 158, First published on: 27 August 2009 (iFirst)
To link to this Article: DOI: 10.1080/08927020903124601
URL: http://dx.doi.org/10.1080/08927020903124601

Molecular Simulation
Vol. 36, No. 2, February 2010, 138–158

Taylor & Francis
Taylor & Francis Group

Controlling the secondary-structure-forming tendencies of proteins by a backbone torsion-energy term

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 16 June 2009*)

We examined a new backbone torsion-energy term proposed by us in the force field for protein systems. This torsion-energy term is represented by a double Fourier series in two variables, namely the backbone dihedral angles ϕ and ψ. It gives a natural representation of the torsion energy in the Ramachandran space in the sense that any two-dimensional energy surface periodic in both ϕ and ψ can be expanded by the double Fourier series. We can then easily control secondary-structure-forming tendencies by modifying the torsion-energy surface. For instance, we can increase or decrease the α-helix-forming-tendencies by lowering or raising the torsion-energy surface in the α-helix region and likewise increase or decrease the β-sheet-forming tendencies by lowering or raising the surface in the β-sheet region in the Ramachandran space. We applied this torsion-energy modification method to six force fields, AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L, and demonstrated that our modifications of the torsion-energy terms resulted in the expected changes of secondary-structure-forming tendencies by performing folding simulations of α-helical and β-hairpin peptides.

Keywords: force field; folding simulation; torsion energy; dihedral angle; secondary structure

1. Introduction

Molecular simulations for protein systems are often used within the framework of classical mechanics consisting of certain potential energy terms with force-field parameters. Some potential energy terms and force-field parameters have been proposed and developed. For example, AMBER [1–4], CHARMM [5], OPLS [6,7], GROMOS [8] and ECEPP [9] are used for many studies of molecular simulations. Usually, these force-field parameters are determined based on experimental results for small molecules and theoretical results using quantum chemistry calculations and molecular dynamics (MD) simulations. However, the simulations using them show different results because these force-field parameters have different values. Therefore, detailed comparisons of three versions of AMBER (parm94 [1], parm96 [2] and parm99 [3]), CHARMM [5], OPLS-AA/L [7] and GROMOS [8] were made by generalised-ensemble simulations [10] of two small peptides in an explicit solvent model [11,12]. The results of these comparisons suggested that these force fields show clearly different behaviours, especially, about secondary-structure-forming tendencies. The folding simulations of the two peptides in an implicit solvent model showed the same secondary-structure-forming tendencies [13–15].

These results imply that it is necessary to refine and improve the existing force-field parameters. Therefore, in order to yield the secondary-structure-forming tendencies that agree with experimental implications, some methods of force-field refinement have been proposed. In a force field, there are some energy terms, for example the bond-stretching term, the bond-bending term, the torsion-energy term and the non-bonded energy term. Among these energy terms, the torsion-energy term is the most problematic. For instance, the parm94, parm96 and parm99 versions of AMBER differ mainly in the backbone torsion-energy parameters. The main changes from OPLS-AA to OPLS-AA/L can also be found in the torsion-energy term [7]. The proposed methods of force-field refinement mainly concentrate on the torsion-energy terms. In some recent proposed force fields, new torsion-energy parameters or terms are proposed, such as AMBER parm99SB [16], AMBER parm03 [4] and CHARMM 22/CMAP [17,18]. These modifications of the torsion energy are usually based on quantum chemistry calculations [19–24]. It was also proposed to set the backbone torsion-energy term simply to zero [25].

We have also proposed a method of force-field refinement [13–15]. This method consists of minimising the sum of the square of the force acting on each atom in the proteins with the structures from the Protein Data Bank (PDB). We selected the partial-charge and backbone torsion-energy parameters for this optimisation and improved the three versions of AMBER (parm94, parm96 and parm99), CHARMM27 and OPLS-AA [13–15].

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

ISSN 0892-7022 print/ISSN 1029-0435 online

DOI: 10.1080/08927020903124601
http://www.informaworld.com

Moreover, we have proposed a new backbone torsion-energy term, which is represented by a double Fourier series in two variables, namely the backbone dihedral angles ϕ and ψ [26]. This expression gives a natural representation of the torsion energy in the Ramachandran space [27] in the sense that any two-dimensional energy surface periodic in both ϕ and ψ can be expanded by the double Fourier series. We can then easily control secondary-structure-forming tendencies by modifying the backbone torsion-energy surface. We have presented preliminary results for AMBER parm94 and AMBER parm96 elsewhere [26]. We remark that a double Fourier series for the backbone torsion energy was also considered in Ref. [17], but it was eventually replaced by a grid-based term in order to accurately reproduce the results of the quantum chemistry calculations. On the other hand, our emphasis is laid rather on the mathematical simplicity and beauty, and we further investigated the possibility of the double Fourier series.

In this article, we first give the details of the new backbone torsion-energy term and present the results of our applications to six existing force fields. Namely, we applied our approach to AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L and tested whether our modifications of the backbone torsion-energy term resulted in the expected change of secondary-structure-forming tendencies by performing folding simulations of α-helical and β-hairpin peptides.

In Section 2, the details of the double Fourier series as the backbone torsion-energy term are given. In Section 3, the results of applications of the double Fourier series to modifications of AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L force fields are presented. Section 4 is devoted to conclusions.

2. Methods

The 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 commonly used conformational potential energy E_{conf} is given by

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

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. 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. Separating the contributions $E(\phi, \psi)$ of the backbone dihedral angles ϕ and ψ from the rest of the torsion terms E_{rest}, we can write

$$E_{\text{torsion}} = E(\phi, \psi) + E_{\text{rest}}, \quad (3)$$

where we have

$$E(\phi, \psi) = \sum_{m} \frac{V_m}{2}[1 + \cos(m\phi - \gamma_m)] + \sum_{n} \frac{V_n}{2}[1 + \cos(n\psi - \gamma_n)]. \quad (4)$$

For example, the coefficients for the cases of six force fields are summarised in Table 1, and we can rewrite

Table 1. Torsion-energy parameters for the backbone dihedral angles ϕ and ψ for AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L in Equation (4).

	ϕ			ψ		
	m	$V_m/2$ (kcal/mol)	γ_m (radians)	n	$V_n/2$ (kcal/mol)	γ_n (radians)
parm94	2	0.2	π	1	0.75	π
				2	1.35	π
				4	0.4	π
parm96	1	0.85	0	1	0.85	0
	2	0.3	π	2	0.3	π
parm99	1	0.8	0	1	1.7	π
	2	0.85	π	2	2.0	π
CHARM	1	0.2	π	1	0.6	0
OPLS-AA	1	−1.1825	0	1	0.908	0
	2	0.456	π	2	0.611	π
	3	−0.425	0	3	0.7905	0
OPLS-AA/L	1	−0.298	0	1	0.3715	0
	2	0.1395	π	2	1.254	π
	3	−2.4565	0	3	−0.4025	0

$E(\phi, \psi)$ in Equation (4) as follows:

$$E_{\text{parm94}}(\phi, \psi) = 2.7 - 0.2\cos 2\phi - 0.75\cos\psi - 1.35\cos 2\psi - 0.4\cos 4\psi, \quad (5)$$

$$E_{\text{parm96}}(\phi, \psi) = 2.3 + 0.85\cos\phi - 0.3\cos 2\phi + 0.85\cos\psi - 0.3\cos 2\psi, \quad (6)$$

$$E_{\text{parm99}}(\phi, \psi) = 5.35 + 0.8\cos\phi - 0.85\cos 2\phi - 1.7\cos\psi - 2.0\cos 2\psi, \quad (7)$$

$$E_{\text{CHARMM}}(\phi, \psi) = 0.8 - 0.2\cos\phi + 0.6\cos\psi, \quad (8)$$

$$E_{\text{OPLS-AA}}(\phi, \psi) = 1.158 - 1.1825\cos\phi - 0.456\cos 2\phi - 0.425\cos 3\phi + 0.908\cos\psi - 0.611\cos 2\psi + 0.7905\cos 3\psi, \quad (9)$$

$$E_{\text{OPLS-AA/L}}(\phi, \psi) = -1.392 - 0.298\cos\phi - 0.1395\cos 2\phi - 2.4565\cos 3\phi + 0.3715\cos\psi - 1.254\cos 2\psi - 0.4025\cos 3\psi. \quad (10)$$

The backbone torsion-energy term $E(\phi, \psi)$ in Equation (4) is a sum of two one-dimensional Fourier series: one is for ϕ and the other for ψ. The two variables ϕ and ψ are decoupled, and no correlation of ϕ and ψ can be incorporated. On the other hand, any periodic function of ϕ and ψ with period 2π can be expanded by a double Fourier series. As a simple generalisation of $E(\phi, \psi)$, we therefore propose to express this backbone torsion energy by the following double Fourier series:

$$\mathcal{E}(\phi, \psi) = a + \sum_{m=1}^{\infty}(b_m\cos m\phi + c_m\sin m\phi) + \sum_{n=1}^{\infty}(d_n\cos n\psi + e_n\sin n\psi) + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}(f_{mn}\cos m\phi\cos n\psi + g_{mn}\cos m\phi\sin n\psi + h_{mn}\sin m\phi\cos n\psi + i_{mn}\sin m\phi\sin n\psi). \quad (11)$$

Here, m and n are the number of waves and a, b_m, c_m, d_n, e_n, f_{mn}, g_{mn}, h_{mn} and i_{mn} are the Fourier coefficients. This equation includes cross-terms in ϕ and ψ, while the original term in Equation (4) has no mixing of ϕ and ψ. Therefore, our new torsion-energy term can represent more complex energy surface than the conventional ones. The Fourier coefficients, by definition, are given by

$$c = \frac{1}{\alpha}\int_{-\pi}^{\pi}\mathrm{d}\phi\int_{-\pi}^{\pi}\mathrm{d}\psi\ \mathcal{E}(\phi, \psi)x(\phi, \psi) = \left(\frac{\pi}{180}\right)^2\frac{1}{\alpha}\int_{-180}^{180}\mathrm{d}\tilde{\phi}\int_{-180}^{180}\mathrm{d}\tilde{\psi}\ \mathcal{E}\left(\frac{\pi}{180}\tilde{\phi}, \frac{\pi}{180}\tilde{\psi}\right)x\left(\frac{\pi}{180}\tilde{\phi}, \frac{\pi}{180}\tilde{\psi}\right), \quad (12)$$

where α is the normalisation constant and $x(\phi, \psi)$ are the basis functions for the Fourier series. Table 2 summarises these coefficients and functions. Here, ϕ and ψ are given in radians, and $\tilde{\phi}$ and $\tilde{\psi}$ are in degrees ($\phi = (\pi/180)\tilde{\phi}, \psi = (\pi/180)\tilde{\psi}$). Hereafter, angular quantities without tilde and with tilde are in radians and in degrees, respectively.

Finally, $\mathcal{E}(\phi, \psi)$ in Equation (11) and E_{rest} in Equation (3) define our torsion-energy term in Equation (1) [instead of Equation (2)]:

$$E_{\text{torsion}} = \mathcal{E}(\phi, \psi) + E_{\text{rest}}. \quad (13)$$

The double Fourier series in Equation (11) is particularly useful, because it describes the backbone torsion-energy surface in the Ramachandran space. The Fourier series can express the torsion-energy surface $\mathcal{E}(\phi, \psi)$ that was obtained by any method including quantum chemistry calculations [7,19–24].

Moreover, one can refine the existing backbone torsion-energy term and control the secondary-structure-forming tendencies of the force fields. For example, α-helix is obtained for $(\tilde{\phi}, \tilde{\psi}) \approx (-57°, -47°)$, 3_{10}-helix for $(\tilde{\phi}, \tilde{\psi}) \approx (-49°, -26°)$, π-helix for $(\tilde{\phi}, \tilde{\psi}) \approx (-57°, -70°)$, parallel β-sheet for $(\tilde{\phi}, \tilde{\psi}) \approx (-119°, 113°)$, antiparallel β-sheet for $(\tilde{\phi}, \tilde{\psi}) \approx (-139°, 135°)$ and so on [27]. Hence, if the existing force field gives, say, too little α-helix-forming tendency compared to experimental results, one can lower the backbone torsion-energy surface near $(\tilde{\phi}, \tilde{\psi}) = (-57°, -47°)$ in order to enhance α-helix formations.

Table 2. Fourier coefficients c, normalisation constants α, and the basis functions $x(\phi,\psi)$ for the double Fourier series of the backbone torsion energy $\mathcal{E}(\phi,\psi)$ in Equations (11) and (12).

c	α	$x(\phi,\psi)$
a	$4\pi^2$	1
b_m	$2\pi^2$	$\cos m\phi$
c_m	$2\pi^2$	$\sin m\phi$
d_n	$2\pi^2$	$\cos n\psi$
e_n	$2\pi^2$	$\sin n\psi$
f_{mn}	π^2	$\cos m\phi\ \cos n\psi$
g_{mn}	π^2	$\cos m\phi\ \sin n\psi$
h_{mn}	π^2	$\sin m\phi\ \cos n\psi$
i_{mn}	π^2	$\sin m\phi\ \sin n\psi$

Table 3. Fourier coefficients in Equation (17) obtained from the numerical evaluations of the integrals in Equation (12). 'org94' stands for the original AMBER parm94 force field. 'mod94(α)' and 'mod94(β)' stand for AMBER parm94 force fields that were modified to enhance α-helix structures and β-sheet structures, respectively, by Equations (14) and (15). The bin size $\tilde{\epsilon}$ is the length of the sides of each unit square cell for the numerical integration in Equation (12).

Bin size $\tilde{\epsilon}$	1°			10°		
Coefficient	org94	mod94(α)	mod94(β)	org94	mod94(α)	mod94(β)
a	2.700000	2.308359	1.916719	2.700000	2.308370	1.916742
b_1	0.000000	−0.330937	0.781150	0.000000	−0.331053	0.781041
c_1	0.000000	0.509599	0.930938	0.000000	0.509517	0.930809
b_2	−0.200000	−0.101549	−0.115937	−0.200000	−0.101513	−0.115970
c_2	0.000000	0.221123	−0.476745	0.000000	0.221100	−0.476558
b_3	0.000000	−0.018073	0.031693	0.000000	−0.018084	0.031714
c_3	0.000000	−0.002862	−0.018298	0.000000	−0.003036	−0.018310
d_1	−0.750000	−1.164401	−0.052959	−0.750000	−1.164500	−0.052874
e_1	0.000000	0.444390	−0.995478	0.000000	0.444289	−0.995599
d_2	−1.350000	−1.333115	−1.184428	−1.350000	−1.333073	−1.184340
e_2	0.000000	0.241460	0.454905	0.000000	0.241451	0.455147
d_3	0.000000	−0.014220	0.035349	0.000000	−0.014143	0.035324
e_3	0.000000	−0.011515	0.009472	0.000000	−0.011671	0.009465
f_{11}	0.000000	−0.342789	−0.680493	0.000000	−0.343087	−0.680497
g_{11}	0.000000	0.367596	0.971845	0.000000	0.367697	0.971851
h_{11}	0.000000	0.527849	−0.810980	0.000000	0.527949	−0.810985
i_{11}	0.000000	−0.566049	1.158199	0.000000	−0.565751	1.158206
f_{21}	0.000000	0.090016	−0.064642	0.000000	0.090168	−0.064636
g_{21}	0.000000	−0.096530	0.092318	0.000000	−0.096472	0.092309
h_{21}	0.000000	0.202178	0.366601	0.000000	0.202421	0.366565
i_{21}	0.000000	−0.216810	−0.523561	0.000000	−0.216596	−0.523509
f_{12}	0.000000	0.012329	−0.142682	0.000000	0.012385	−0.142712
g_{12}	0.000000	0.176308	−0.392017	0.000000	0.176622	−0.392098
h_{12}	0.000000	−0.018984	−0.170042	0.000000	−0.019013	−0.170077
i_{12}	0.000000	−0.271490	−0.467187	0.000000	−0.271321	−0.467284
f_{22}	0.000000	−0.000586	−0.002453	−0.000001	−0.000585	−0.002451
g_{22}	0.000000	−0.008378	−0.006738	0.000000	−0.008397	−0.006733
h_{22}	0.000000	−0.001316	0.013909	0.000000	−0.001317	0.013897
i_{22}	0.000000	−0.018817	0.038215	0.000000	−0.018867	0.038183

We can thus write

$$\mathcal{E}(\phi, \psi) = E(\phi, \psi) - f(\phi, \psi), \tag{14}$$

where $E(\phi, \psi)$ is the existing backbone torsion-energy term that we want to refine and $f(\phi, \psi)$ is a function that has peaks around the corresponding regions where specific secondary structures are to be enhanced. There are many possible choices for $f(\phi, \psi)$. For instance, one can use the following function when one wants to lower the torsion-energy surface in a single region near $(\phi, \psi) = (\phi_0, \psi_0)$:

$$f(\phi, \psi) = \begin{cases} A \exp\left(\frac{B}{(\phi-\phi_0)^2+(\psi-\psi_0)^2-r_0}\right), \\ 0, \\ \quad \text{for } (\phi-\phi_0)^2 + (\psi-\psi_0)^2 < r_0^2, \\ \quad \text{otherwise}, \end{cases} \tag{15}$$

where A, B and r_0 are constants that we adjust for refinement. In this case, the energy surface is lowered by $f(\phi, \psi)$ in a circular region of radius r_0, which is centred at $(\phi, \psi) = (\phi_0, \psi_0)$. Note that we should also impose periodic boundary conditions on $f(\phi, \psi)$.

We then express $\mathcal{E}(\phi, \psi)$ in Equation (14) in terms of the double Fourier series in Equation (11), where the Fourier coefficients are obtained from Equation (12). Hence, we can fine-tune the backbone torsion-energy term by the above procedure so that it yields correct secondary-structure-forming tendencies.

3. Results and discussion

We now present the results of the applications of our backbone torsion energy in Equation (11). In our previous paper [26], we examined the following truncated Fourier series:

$$\begin{aligned} \mathcal{E}(\phi, \psi) = a &+ b_1 \cos\phi + c_1 \sin\phi + b_2 \cos 2\phi \\ &+ c_2 \sin 2\phi + d_1 \cos\psi + e_1 \sin\psi \\ &+ d_2 \cos 2\psi + e_2 \sin 2\psi \\ &+ f_{11} \cos\phi \cos\psi + g_{11} \cos\phi \sin\psi \\ &+ h_{11} \sin\phi \cos\psi + i_{11} \sin\phi \sin\psi. \end{aligned} \tag{16}$$

Table 4. Fourier coefficients in Equation (17) obtained from the numerical evaluations of the integrals in Equation (12). 'org96' stands for the original AMBER parm96 force field. 'mod96(α)' and 'mod96(β)' stand for AMBER parm96 force fields that were modified to enhance α-helix structures and β-sheet structures, respectively, by Equations (14) and (15). See also the caption of Table 3.

Bin size $\tilde{\epsilon}$	1°			10°		
Coefficient	org96	mod96(α)	mod96(β)	org96	mod96(α)	mod96(β)
a	2.300000	1.908359	1.908359	2.300000	1.908370	1.908371
b_1	0.850000	0.519063	1.240575	0.850000	0.518947	1.240521
c_1	0.000000	0.509599	0.465469	0.000000	0.509517	0.465404
b_2	−0.300000	−0.201549	−0.257968	−0.300000	−0.201513	−0.257985
c_2	0.000000	0.221123	−0.238372	0.000000	0.221100	−0.238279
b_3	0.000000	−0.018073	0.015847	0.000000	−0.018084	0.015857
c_3	0.000000	−0.002862	−0.009149	0.000000	−0.003036	−0.009155
d_1	0.850000	0.435599	1.198520	0.850000	0.435500	1.198563
e_1	0.000000	0.444390	−0.497739	0.000000	0.444289	−0.497800
d_2	−0.300000	−0.283115	−0.217214	−0.300000	−0.283073	−0.217170
e_2	0.000000	0.241460	0.227452	0.000000	0.241451	0.227573
d_3	0.000000	−0.014220	0.017675	0.000000	−0.014143	0.017662
e_3	0.000000	−0.011515	0.004736	0.000000	−0.011671	0.004733
f_{11}	0.000000	−0.342789	−0.340247	0.000000	−0.343087	−0.340249
g_{11}	0.000000	0.367596	0.485922	0.000000	0.367697	0.485925
h_{11}	0.000000	0.527849	−0.405490	0.000000	0.527949	−0.405492
i_{11}	0.000000	−0.566049	0.579100	0.000000	−0.565751	0.579103
f_{21}	0.000000	0.090016	−0.032321	0.000000	0.090168	−0.032318
g_{21}	0.000000	−0.096530	0.046159	0.000000	−0.096472	0.046154
h_{21}	0.000000	0.202178	0.183301	0.000000	0.202421	0.183283
i_{21}	0.000000	−0.216810	−0.261781	0.000000	−0.216596	−0.261755
f_{12}	0.000000	0.012329	−0.071341	0.000000	0.012385	−0.071356
g_{12}	0.000000	0.176308	−0.196008	0.000000	0.176622	−0.196049
h_{12}	0.000000	−0.018984	−0.085021	0.000000	−0.019013	−0.085039
i_{12}	0.000000	−0.271490	−0.233594	0.000000	−0.271321	−0.233642
f_{22}	0.000000	−0.000586	−0.001226	0.000000	−0.000585	−0.001226
g_{22}	0.000000	−0.008378	−0.003369	0.000000	−0.008397	−0.003366
h_{22}	0.000000	−0.001316	0.006955	0.000000	−0.001317	0.006949
i_{22}	0.000000	−0.018817	0.019108	0.000000	−0.018867	0.019091

This function has 13 Fourier coefficient parameters. We can easily modify the existing force fields so that specified secondary-structure-forming tendencies are enhanced for AMBER parm94 and AMBER parm96. In this article, we consider the following truncated Fourier series:

$$\begin{aligned}\mathcal{E}(\phi,\psi) =& a + b_1\cos\phi + c_1\sin\phi + b_2\cos 2\phi \\ &+ c_2\sin 2\phi + b_3\cos 3\phi + c_3\sin 3\phi \\ &+ d_1\cos\psi + e_1\sin\psi + d_2\cos 2\psi \\ &+ e_2\sin 2\psi + d_3\cos 3\psi + e_3\sin 3\psi \\ &+ f_{11}\cos\phi\cos\psi + g_{11}\cos\phi\sin\psi \\ &+ h_{11}\sin\phi\cos\psi + i_{11}\sin\phi\sin\psi \\ &+ f_{21}\cos 2\phi\cos\psi + g_{21}\cos 2\phi\sin\psi \\ &+ h_{21}\sin 2\phi\cos\psi + i_{21}\sin 2\phi\sin\psi \\ &+ f_{12}\cos\phi\cos 2\psi + g_{12}\cos\phi\sin 2\psi \\ &+ h_{12}\sin\phi\cos 2\psi + i_{12}\sin\phi\sin 2\psi \\ &+ f_{22}\cos 2\phi\cos 2\psi + g_{22}\cos 2\phi\sin 2\psi \\ &+ h_{22}\sin 2\phi\cos 2\psi + i_{22}\sin 2\phi\sin 2\psi. \end{aligned} \quad (17)$$

This function has 29 Fourier coefficient parameters. We will see below that this number of Fourier terms is sufficient for most of our purposes.

We first check how well the truncated Fourier series in Equation (17) can reproduce the six original backbone torsion-energy terms in Equations (5)–(10). Because these functions are already the sum of one-dimensional Fourier series and subsets of the double Fourier series in Equation (11), the Fourier coefficients in Equation (12) can be analytically calculated and agree with those in Equations (5)–(10) except for the last one (that for $\cos 4\psi$) in Equation (5). This term is missing in Equation (17). These cases thus give us good test of numerical integrations in Equation (12). The numerical integrations were evaluated as follows. We divided the Ramachandran space ($-180° < \tilde{\phi} < 180°, -180° < \tilde{\psi} < 180°$) into unit square cells of side length $\tilde{\epsilon}$ (in degrees). Hence, there are $(360/\tilde{\epsilon})^2$ unit cells altogether. The double integral on the right-hand side of Equation (12) was approximated by the sum of $[\mathcal{E}((\pi/180)\tilde{\phi},(\pi/180)\tilde{\psi})x((\pi/180)\tilde{\phi},(\pi/180)\tilde{\psi})] \times\tilde{\epsilon}^2$, where each $\mathcal{E}((\pi/180)\tilde{\phi},(\pi/180)\tilde{\psi})x((\pi/180)\tilde{\phi},(\pi/180)\tilde{\psi})$ was evaluated at one of the four corners of each unit cell. We tried two values of $\tilde{\epsilon}$ (1° and 10°).

Table 5. Fourier coefficients in Equation (17) obtained from the numerical evaluations of the integrals in Equation (12). 'org99' stands for the original AMBER parm99 force field. 'mod99(α)' and 'mod99(β)' stand for AMBER parm99 force fields that were modified to enhance α-helix structures and β-sheet structures, respectively, by Equations (14) and (15). See also the caption of Table 3.

Bin size $\tilde{\epsilon}$	1°			10°		
Coefficient	org99	mod99(α)	mod99(β)	org99	mod99(α)	mod99(β)
a	5.350000	4.958359	4.566718	5.350000	4.958370	4.566743
b_1	0.800000	0.469063	1.581150	0.800000	0.468947	1.581041
c_1	0.000000	0.509599	0.930938	0.000000	0.509517	0.930809
b_2	−0.850000	−0.751549	−0.765937	−0.850000	−0.751513	−0.765970
c_2	0.000000	0.221123	−0.476745	0.000000	0.221100	−0.476558
b_3	0.000000	−0.018073	0.031693	0.000000	−0.018084	0.031714
c_3	0.000000	−0.002862	−0.018298	0.000000	−0.003036	−0.018310
d_1	−1.700000	−2.114401	−1.002959	−1.700000	−2.114500	−1.002874
e_1	0.000000	0.444390	−0.995478	0.000000	0.444289	−0.995599
d_2	−2.000000	−1.983115	−1.834428	−2.000000	−1.983073	−1.834340
e_2	0.000000	0.241460	0.454905	0.000000	0.241451	0.455147
d_3	0.000000	−0.014220	0.035349	0.000000	−0.014143	0.035325
e_3	0.000000	−0.011515	0.009472	0.000000	−0.011671	0.009465
f_{11}	0.000000	−0.342789	−0.680493	0.000000	−0.343087	−0.680497
g_{11}	0.000000	0.367596	0.971845	0.000000	0.367697	0.971851
h_{11}	0.000000	0.527849	−0.810980	0.000000	0.527949	−0.810985
i_{11}	0.000000	−0.566049	1.158199	0.000000	−0.565751	1.158206
f_{21}	0.000000	0.090016	−0.064642	0.000000	0.090168	−0.064635
g_{21}	0.000000	−0.096530	0.092318	0.000000	−0.096472	0.092309
h_{21}	0.000000	0.202178	0.366601	0.000000	0.202421	0.366565
i_{21}	0.000000	−0.216810	−0.523561	0.000000	−0.216596	−0.523509
f_{12}	0.000000	0.012329	−0.142682	0.000000	0.012385	−0.142712
g_{12}	0.000000	0.176308	−0.392017	0.000000	0.176622	−0.392098
h_{12}	0.000000	−0.018984	−0.170042	0.000000	−0.019013	−0.170077
i_{12}	0.000000	−0.271490	−0.467187	0.000000	−0.271321	−0.467284
f_{22}	0.000000	−0.000586	−0.002453	0.000000	−0.000585	−0.002451
g_{22}	0.000000	−0.008378	−0.006738	0.000000	−0.008397	−0.006733
h_{22}	0.000000	−0.001316	0.013909	0.000000	−0.001317	0.013897
i_{22}	0.000000	−0.018817	0.038215	0.000000	−0.018867	0.038183

Both cases gave almost complete agreement of Fourier coefficients with the results of the analytical integrations (see Tables 3–8).

In Figure 1, we compare the six original backbone torsion-energy surfaces with those of the corresponding double Fourier series in Equation (17). Hereafter, the primed labels for figures such as (a′) indicate that the results are those of the double Fourier series. As can be seen from Figure 1, the backbone torsion-energy surfaces are in complete agreement for all force fields except for AMBER parm94, whereas we see a little difference for AMBER parm94 between Figure 1(a) and (a′). As discussed above, this slight difference for AMBER parm94 reflects the fact that the $\cos 4\psi$ term in Equation (5) is missing in the truncated double Fourier series in Equation (17).

We now consider the double Fourier series of non-trigonometric functions. The functions are those in Equations (14) and (15). We tried to fine-tune the six original force fields by subtracting $f(\phi, \psi)$ in Equation (15) from the original functions. The criterion for fine-tuning is, for instance, whether the refined force fields yield better agreement of the secondary-structure-forming tendencies with experimental implications than the original ones. For this, we need good experimental data. Because the purpose of the present article is to test whether or not we can control the secondary-structure-forming tendencies, we simply consider extreme cases where we try to modify the existing force fields so that desired secondary structures may be obtained regardless of the tendencies of the original force fields. Note that the six original force fields have quite different preferences for α-helix and β-sheet structures [11–15].

The function $f(\phi, \psi)$ in Equation (15) reduces the value of $E(\phi, \psi)$ in a circle of radius r_0 with the centre located at (ϕ_0, ψ_0). We used $\tilde{r}_0 = 100°$ and $\tilde{B} = 5000$ (degrees)^2. The coefficient A is calculated by Equation (15) from the other parameters $f(\tilde{\phi}_0, \tilde{\psi}_0)$, $\tilde{r}_0$ and $\tilde{B}$. Namely, we have

$$A = f(\tilde{\phi}_0, \tilde{\psi}_0)\exp\left(\frac{\tilde{B}}{\tilde{r}_0^2}\right). \tag{18}$$

We used $(\tilde{\phi}_0, \tilde{\psi}_0) = (-57°, -47°)$ and $(\tilde{\phi}_0, \tilde{\psi}_0) = (-130°, 125°)$ in order to enhance α-helix-forming

Table 6. Fourier coefficients in Equation (17) obtained from the numerical evaluations of the integrals in Equation (12). 'orgch' stands for the original CHARMM27 force field. 'modch(α)' and 'modch(β)' stand for CHARMM force fields that were modified to enhance α-helix structures and β-sheet structures, respectively, by Equations (14) and (15). See also the caption of Table 3.

Bin size $\tilde{\epsilon}$	1°			10°		
Coefficient	orgch	modch(α)	modch(β)	orgch	modch(α)	modch(β)
a	0.800000	0.408359	0.016719	0.800000	0.408370	0.016743
b_1	−0.200000	−0.530937	0.581150	−0.200000	−0.531053	0.581041
c_1	0.000000	0.509599	0.930938	0.000000	0.509517	0.930809
b_2	0.000000	0.098450	0.084063	0.000000	0.098487	0.084030
c_2	0.000000	0.221123	−0.476745	0.000000	0.221100	−0.476558
b_3	0.000000	−0.018073	0.031693	0.000000	−0.018084	0.031714
c_3	0.000000	−0.002862	−0.018298	0.000000	−0.003036	−0.018310
d_1	0.600000	0.185599	1.297041	0.600000	0.185500	1.297126
e_1	0.000000	0.444390	−0.995478	0.000000	0.444289	−0.995599
d_2	0.000000	0.016885	0.165572	0.000000	0.016927	0.165660
e_2	0.000000	0.241460	0.454905	0.000000	0.241451	0.455147
d_3	0.000000	−0.014220	0.035349	0.000000	−0.014143	0.035325
e_3	0.000000	−0.011515	0.009472	0.000000	−0.011671	0.009465
f_{11}	0.000000	−0.342789	−0.680493	0.000000	−0.343087	−0.680497
g_{11}	0.000000	0.367596	0.971845	0.000000	0.367697	0.971851
h_{11}	0.000000	0.527849	−0.810980	0.000000	0.527949	−0.810985
i_{11}	0.000000	−0.566049	1.158199	0.000000	−0.565751	1.158206
f_{21}	0.000000	0.090016	−0.064642	0.000000	0.090168	−0.064636
g_{21}	0.000000	−0.096530	0.092318	0.000000	−0.096472	0.092309
h_{21}	0.000000	0.202178	0.366601	0.000000	0.202421	0.366565
i_{21}	0.000000	−0.216810	−0.523561	0.000000	−0.216596	−0.523509
f_{12}	0.000000	0.012329	−0.142682	0.000001	0.012386	−0.142711
g_{12}	0.000000	0.176308	−0.392017	0.000000	0.176622	−0.392098
h_{12}	0.000000	−0.018984	−0.170042	0.000000	−0.019013	−0.170077
i_{12}	0.000000	−0.271490	−0.467187	0.000000	−0.271321	−0.467284
f_{22}	0.000000	−0.000586	−0.002453	0.000000	−0.000585	−0.002450
g_{22}	0.000000	−0.008378	−0.006738	0.000000	−0.008397	−0.006733
h_{22}	0.000000	−0.001316	0.013909	0.000000	−0.001317	0.013897
i_{22}	0.000000	−0.018817	0.038215	0.000000	−0.018867	0.038183

tendency and β-sheet-forming tendency, respectively. The central values $f(\tilde{\phi}_0, \tilde{\psi}_0)$ that we used were 3.0 and 6.0 kcal/mol for enhancing α-helix and β-sheet, respectively, in the case of AMBER parm94, AMBER parm99, CHARMM27 and OPLS-AA/L. They were both 3.0 kcal/mol in the case of AMBER parm96 and OPLS-AA.

We remark that the large value of $f(\tilde{\phi}_0, \tilde{\psi}_0)$, 6.0 kcal/mol, that was necessary to enhance β-sheet in the case of AMBER parm94, AMBER parm99, CHARMM27 and OPLS-AA/L reflects the fact that the original force field favours α-helix.

In Figure 2(a1)–(f1), we compare the six backbone torsion-energy surfaces modified according to Equation (14), which reduced the torsion energy in the α-helix region, with those of the corresponding double Fourier series in Equation (17). In Figure 2(a1)–(f1), α-helix is enhanced from the original AMBER parm94 (a1), AMBER parm96 (b1), AMBER parm99 (c1), CHARMM27 (d1), OPLS-AA (e1) and OPLS-AA/L (f1). In Figure 3(a1)–(f1), we show the case of the β-sheet region, and β-sheet is enhanced from the original AMBER parm94 (a1), AMBER parm96 (b1), AMBER parm99 (c1), CHARMM27 (d1), OPLS-AA (e1) and OPLS-AA/L (f1).

These modified backbone torsion-energy functions were expanded by the truncated double Fourier series in Equation (17) by evaluating the corresponding Fourier coefficients from Equation (12). For the numerical integration, we again tried two values of the bin size $\tilde{\epsilon}$ (1° and 10°). The obtained Fourier coefficients are summarised in Tables 3–8 in the case of AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L, respectively. For comparisons, the Fourier coefficients of the original AMBER force fields (before modifications) are also listed. We see that the two choices of the bin size $\tilde{\epsilon}$ gave essentially the same results (agreeing in about three digits).

In Figures 2(a1′)–(f1′) and 3(a1′)–(f1′), we show the backbone torsion-energy surfaces represented by the truncated double Fourier series. Comparing these with the original ones in Figures 2(a1)–(f1) and 3(a1)–(f1),

Table 7. Fourier coefficients in Equation (17) obtained from the numerical evaluations of the integrals in Equation (12). 'orgop' stands for the original OPLS-AA force field. 'modop(α)' and 'modop(β)' stand for OPLS-AA force fields that were modified to enhance α-helix structures and β-sheet structures, respectively, by Equations (14) and (15). See also the caption of Table 3.

Bin size $\tilde{\epsilon}$	1°			10°		
Coefficient	orgop	modop(α)	modop(β)	orgop	modop(α)	modop(β)
a	1.158000	0.766359	0.766359	1.158000	0.766370	0.766371
b_1	−1.182500	−1.513437	−0.791925	−1.182500	−1.513553	−0.791979
c_1	0.000000	0.509599	0.465469	0.000000	0.509517	0.465404
b_2	−0.456000	−0.357550	−0.413969	−0.456000	−0.357513	−0.413985
c_2	0.000000	0.221123	−0.238372	0.000000	0.221100	−0.238279
b_3	−0.425000	−0.443073	−0.409154	−0.425000	−0.443084	−0.409143
c_3	0.000000	−0.002862	−0.009149	0.000000	−0.003036	−0.009155
d_1	0.908000	0.493599	1.256520	0.908000	0.493500	1.256563
e_1	0.000000	0.444390	−0.497739	0.000000	0.444289	−0.497800
d_2	−0.611000	−0.594115	−0.528214	−0.610999	−0.594072	−0.528170
e_2	0.000000	0.241460	0.227452	0.000000	0.241451	0.227573
d_3	0.790500	0.776280	0.808175	0.790500	0.776357	0.808163
e_3	0.000000	−0.011515	0.004736	0.000000	−0.011671	0.004733
f_{11}	0.000000	−0.342789	−0.340247	0.000000	−0.343087	−0.340249
g_{11}	0.000000	0.367596	0.485922	0.000000	0.367697	0.485925
h_{11}	0.000000	0.527849	−0.405490	0.000000	0.527949	−0.405492
i_{11}	0.000000	−0.566049	0.579100	0.000000	−0.565751	0.579103
f_{21}	0.000000	0.090016	−0.032321	−0.000001	0.090167	−0.032318
g_{21}	0.000000	−0.096530	0.046159	0.000000	−0.096472	0.046154
h_{21}	0.000000	0.202178	0.183301	0.000000	0.202421	0.183283
i_{21}	0.000000	−0.216810	−0.261781	0.000000	−0.216596	−0.261755
f_{12}	0.000000	0.012329	−0.071341	0.000000	0.012385	−0.071356
g_{12}	0.000000	0.176308	−0.196008	0.000000	0.176622	−0.196049
h_{12}	0.000000	−0.018984	−0.085021	0.000000	−0.019013	−0.085039
i_{12}	0.000000	−0.271490	−0.233594	0.000000	−0.271321	−0.233642
f_{22}	0.000000	−0.000586	−0.001226	0.000000	−0.000585	−0.001226
g_{22}	0.000000	−0.008378	−0.003369	0.000000	−0.008397	−0.003366
h_{22}	0.000000	−0.001316	0.006955	0.000000	−0.001317	0.006949
i_{22}	0.000000	−0.018817	0.019108	0.000000	−0.018867	0.019091

we find that the overall features of the energy surfaces are well reproduced by the Fourier series. If more accuracy is desired, we can simply increase the number of Fourier terms in the expansion. As we will see below, the present accuracy of the Fourier series was sufficient for the purpose of controlling the secondary-structure-forming tendencies towards α-helix or β-sheet.

We examined the effects of the above modifications of the backbone torsion-energy terms in AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L (towards specific secondary structures) by performing the folding simulations of 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 [28]. 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 [29,30]. Because the charges at peptide termini are known to affect helix stability [29,30], 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 [28,31,32]. This peptide is known to form a β-hairpin structure by experiments [28,31,32].

Simulated annealing [33] MD simulations were performed for both peptides from fully extended initial conformations, where the 12 versions of the truncated double Fourier series [which were described in Tables 3–8 and in Figures 2(a1′)–(f1′) and 3(a1′)–(f1′)] were used for the backbone torsion-energy terms of AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L force fields. For comparisons, the simulations with the original force fields were also performed. The unit time step was set to 1.0 fs. Each simulation was carried out for 1 ns (hence, it consisted of 1,000,000 MD steps). The temperature during MD simulations was controlled by Berendsen's method [34].

Table 8. Fourier coefficients in Equation (17) obtained from the numerical evaluations of the integrals in Equation (12). 'orgopl' stands for the original OPLS-AA/L force field. 'modopl(α)' and 'modopl(β)' stand for OPLS-AA/L force fields that were modified to enhance α-helix structures and β-sheet structures, respectively, by Equations (14) and (15). See also the caption of Table 3.

Bin size $\tilde{\epsilon}$	1°			10°		
Coefficient	orgopl	modopl(α)	modopl(β)	orgopl	modopl(α)	modopl(β)
a	−1.392000	−1.783641	−2.175281	−1.392000	−1.783630	−2.175258
b_1	−0.298000	−0.628937	0.483150	−0.298001	−0.629054	0.483041
c_1	0.000000	0.509599	0.930938	0.000000	0.509517	0.930809
b_2	−0.139500	−0.041050	−0.055437	−0.139500	−0.041012	−0.055470
c_2	0.000000	0.221123	−0.476745	0.000000	0.221100	−0.476558
b_3	−2.456500	−2.474573	−2.424807	−2.456500	−2.474584	−2.424786
c_3	0.000000	−0.002862	−0.018298	0.000000	−0.003036	−0.018310
d_1	0.371500	−0.042901	1.068541	0.371500	−0.043000	1.068627
e_1	0.000000	0.444390	−0.995478	0.000000	0.444289	−0.995599
d_2	−1.254000	−1.237115	−1.088428	−1.254000	−1.237073	−1.088340
e_2	0.000000	0.241460	0.454905	0.000000	0.241451	0.455147
d_3	−0.402500	−0.416720	−0.367151	−0.402500	−0.416643	−0.367175
e_3	0.000000	−0.011515	0.009472	0.000000	−0.011671	0.009465
f_{11}	0.000000	−0.342789	−0.680493	0.000000	−0.343087	−0.680497
g_{11}	0.000000	0.367596	0.971845	0.000000	0.367697	0.971851
h_{11}	0.000000	0.527849	−0.810980	0.000000	0.527949	−0.810985
i_{11}	0.000000	−0.566049	1.158199	0.000000	−0.565751	1.158206
f_{21}	0.000000	0.090016	−0.064642	0.000000	0.090168	−0.064636
g_{21}	0.000000	−0.096530	0.092318	0.000000	−0.096472	0.092309
h_{21}	0.000000	0.202178	0.366601	0.000000	0.202421	0.366565
i_{21}	0.000000	−0.216810	−0.523561	0.000000	−0.216596	−0.523509
f_{12}	0.000000	0.012329	−0.142682	0.000000	0.012385	−0.142712
g_{12}	0.000000	0.176308	−0.392017	0.000000	0.176622	−0.392098
h_{12}	0.000000	−0.018984	−0.170042	0.000000	−0.019013	−0.170077
i_{12}	0.000000	−0.271490	−0.467187	0.000000	−0.271321	−0.467284
f_{22}	0.000000	−0.000586	−0.002453	0.000000	−0.000585	−0.002451
g_{22}	0.000000	−0.008378	−0.006738	0.000000	−0.008397	−0.006733
h_{22}	0.000000	−0.001316	0.013909	0.000000	−0.001317	0.013897
i_{22}	0.000000	−0.018817	0.038215	0.000000	−0.018867	0.038183

For each run, the temperature was decreased exponentially from 2000 to 250 K. We modified and used the program package TINKER version 4.1 (software available at http://dasher.wustl.edu/tinker/) for all the simulations. As for solvent effects, we used the Generalized Born/Surface Area (GB/SA) model [35,36] included in the TINKER program package. For both peptides, these folding simulations were repeated 60 times with different sets of randomly generated initial velocities.

In Table 9, the numbers of obtained conformations (final conformations) with α-helix structures and β-hairpin structures are listed for the case of C-peptide, which is known to form α-helix structures by experiments. We used DSSP [37] for the criterions of secondary-structure formations. We see that for the original AMBER parm94, 49 out of 60 conformations are α-helix structures, and 29 out of 60 conformations are 3_{10}-helix structures and that for the original AMBER parm99, 34 out of 60 conformations are α-helix structures and 32 out of 60 conformations are 3_{10}-helix structures. These results confirm that the original AMBER parm94 and AMBER parm99 favour helix structures more than other force fields. For the original AMBER parm96, 19 out of 60 conformations are α-helix structures, 11 out of 60 conformations are β-bridges, and 4 out of 60 conformations are extended strand structures. The number of β-structures of the original AMBER parm96 was the most among the six original force fields. AMBER parm96 favours β-hairpin structures (because C-peptide is a helix-forming peptide) [11,13]. For the original CHARMM27, 12 out of 60 conformations are α-helix structures, 15 out of 60 conformations are π-helix structures and 6 out of 60 conformations are β-bridge structures. Though we found fewer numbers of α-helix structures than in the case of the AMBER parm94, parm96 and parm99, more helix structures than β-structures were obtained. About half of the conformations are helical (12 for α-helix and 15 for π-helix), therefore the original CHARMM27 slightly favours helix structures compared to other force fields. For the original OPLS-AA, 11 out of 60 conformations are α-helix structures, 6 out of 60 conformations are 3_{10}-helix structures, 7 out of 60 conformations are β-bridge structures and 5 out of 60 conformations are extended strand structures. We cannot clearly tell whether the original OPLS favours

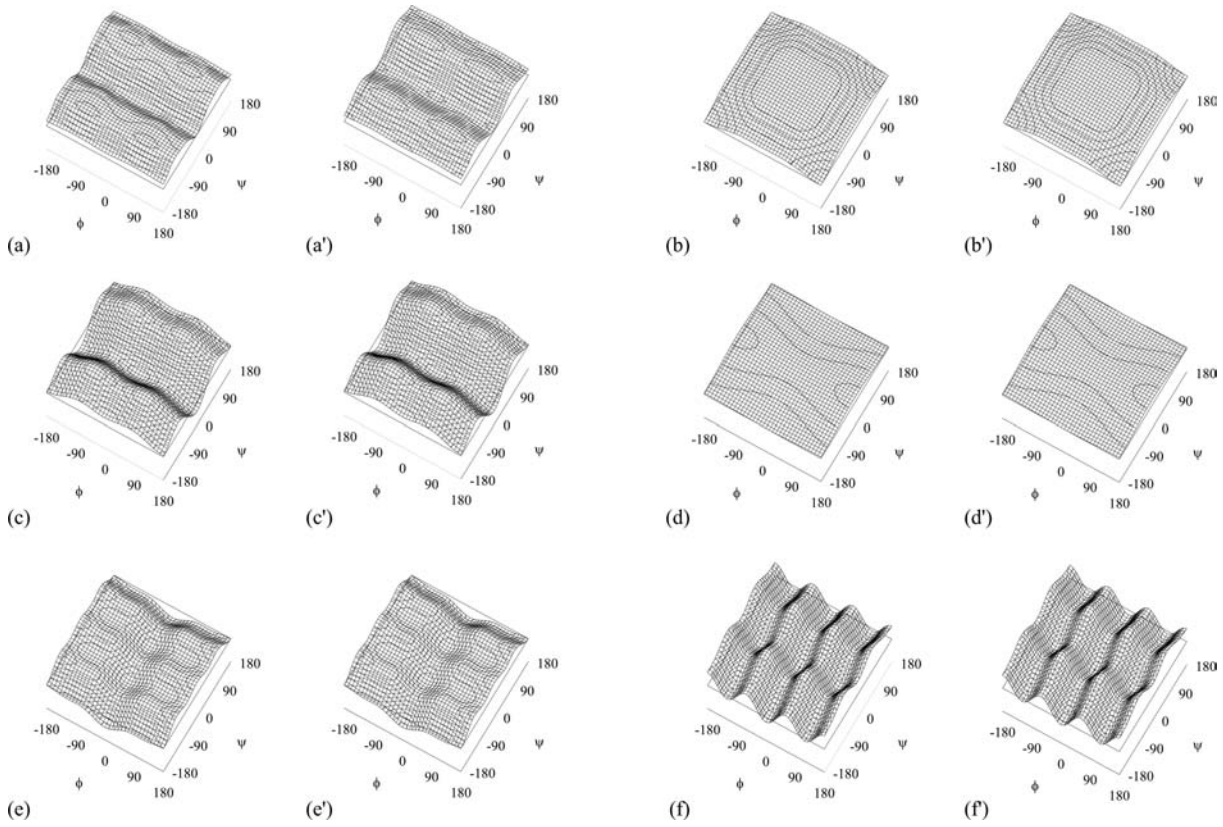

Figure 1. Backbone torsion-energy surfaces of six force fields. The backbone dihedral angles $\tilde{\phi}$ and $\tilde{\psi}$ in degrees (a) to (f) are those of the original AMBER parm94, the original AMBER parm96, the original AMBER parm99, the original CHARMM27, the original OPLS-AA and the original OPLS-AA/L, respectively. (a′) to (b′) are those of (a) to (b), respectively, that are expressed by the truncated double Fourier series in Equation (17). The contour lines are drawn every 0.5 kcal/mol.

helix structures or β-structures. For the original OPLS-AA/L, 15 out of 60 conformations are α-helix structures, 29 out of 60 conformations are 3_{10}-helix structures and 5 out of 60 conformations are β-bridge structures. Thus, the original OPLS-AA/L favours helix structures.

However, for all six force fields modified to enhance α-helix-forming tendency, almost all conformations are α-helix structures and there are no conformations with β-hairpin structures. For CHARMM27, the number of π-helix structures also increases. On the other hand, from about a quarter to about half the conformations exhibit β-hairpin structures for all six force fields modified to enhance β-sheet-forming tendency and no α-helix structures are found except for OPLS-AA. As for the OPLS-AA modified towards β-sheet, though the central values $f(\tilde{\phi}_0, \tilde{\psi}_0)$ that we used in Equation (18) were 3.0 kcal/mol, there are more number of β-structures than helix structures. Therefore, OPLS-AA modified towards β-sheet also favours β-structures.

In Table 10, the numbers of obtained conformations (final conformations) with α-helix structures and β-hairpin structures are listed for the case of G-peptide, which is known to form β-hairpin structures by experiments. The results are similar to those in Table 9. We see that for the original AMBER parm94, 47 out of 60 conformations are α-helix structures and 36 out of 60 conformations are 3_{10}-helix structures and that for the original AMBER parm99, 24 out of 60 conformations are α-helix structures and 46 out of 60 conformations are 3_{10}-helix structures. The original AMBER parm94 and AMBER parm99 strongly favour helix structures, and there are no β-structures observed. It is difficult for the original AMBER parm94 and AMBER parm99 to produce β structures. For the original CHARMM27 and OPLS-AA/L, there are no or few β structures. For the original AMBER parm96, 12 out of 60 conformations are α-helix structures, 2 out of 60 conformations are 3_{10}-helix structures, 12 out of 60 conformations are β-bridge structures and 7 out of 60 conformations are extended strand. For the original OPLS-AA, 8 out of 60 conformations are α-helix structures, 12 out of 60 conformations are β-bridge structures and 4 out of 60 conformations are extended strand. Hence, the original AMBER parm96 and OPLS-AA slightly favour β-structures.

In Figures 4–9, we show seven (out of 60) lowest-energy final conformations of C-peptide and G-peptide obtained by

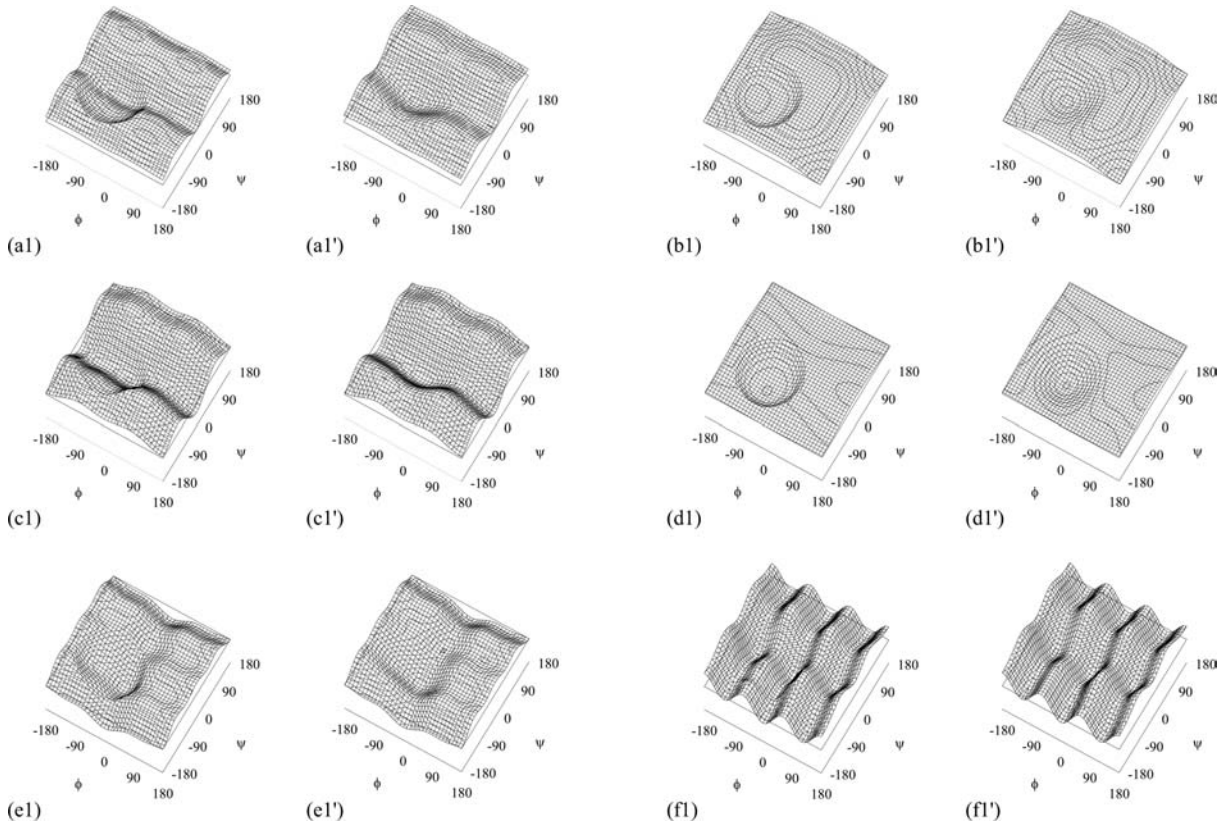

Figure 2. Backbone torsion-energy surfaces of six force fields that were modified by Equations (14), (15) and (17). (a1) to (f1) are those of AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L force fields that were modified to enhance α-helix structures, respectively. (a1′) to (f1′) are those of AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L force fields that were expanded by the truncated double Fourier series in Equation (17).

the simulated annealing MD simulations in the case of AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L, respectively. For these simulations, we used 18 force fields, namely the six original force fields, those six modified towards α-helix and those six modified towards β-sheet.

In Figure 4, we see that all conformations of the original AMBER parm94 (except for conformations 2 and 4 of G-peptide) and all conformations of its force field modified towards α-helix are α-helix structures (conformations 2 and 4 are 3_{10}-helix structures). The results show that the original AMBER parm94 favours α-helix structures, and, moreover, its force field modified towards α-helix favours α-helix structures more than the original force field in the sense that the obtained helices are more extended (and almost entirely helical). On the other hand, AMBER parm94 modified towards β-sheet favours β-structures strongly.

In Figure 5, conformations 2–5 are α-helix structures and conformations 3 and 5 are β-bridge structures for the original AMBER parm96 in the case of C-peptide, and conformation 1 is a helix structure, conformation 2 is a 3_{10}-helix structure, conformations 4 and 6 are β-bridge structures, and conformations 3 and 5 are extended β-strand structures for the original AMBER parm96 in the case of G-peptide. AMBER parm96 modified towards α-helix structures favours α-helix structures, and that modified towards β-sheet structures favours β-structures strongly, regardless of the secondary-structure-forming tendencies of the original AMBER parm96. In Figure 6, the results for the AMBER parm99 are almost the same as in the case of AMBER parm94. In Figure 7, conformations 2 and 4 are α-helix structures and conformations 1, 2 and 5 are π-helix structures for the original CHARMM27 in the case of C-peptide, and no conformations have secondary structures for the original CHARMM27 in the case of G-peptide. For CHARMM27 modified towards α-helix structures, all conformations of C-peptide and conformations 1, 3–6 of G-peptide have α-helix structures, and conformations 2, 3, 5 and 7 of C-peptide and conformations 1, 2, 4 and 7 of G-peptide have π-helix structures. These results show that CHARMM27 modified towards α-helix structures favours α-helix structures and π-helix structures. For CHARMM27 modified towards

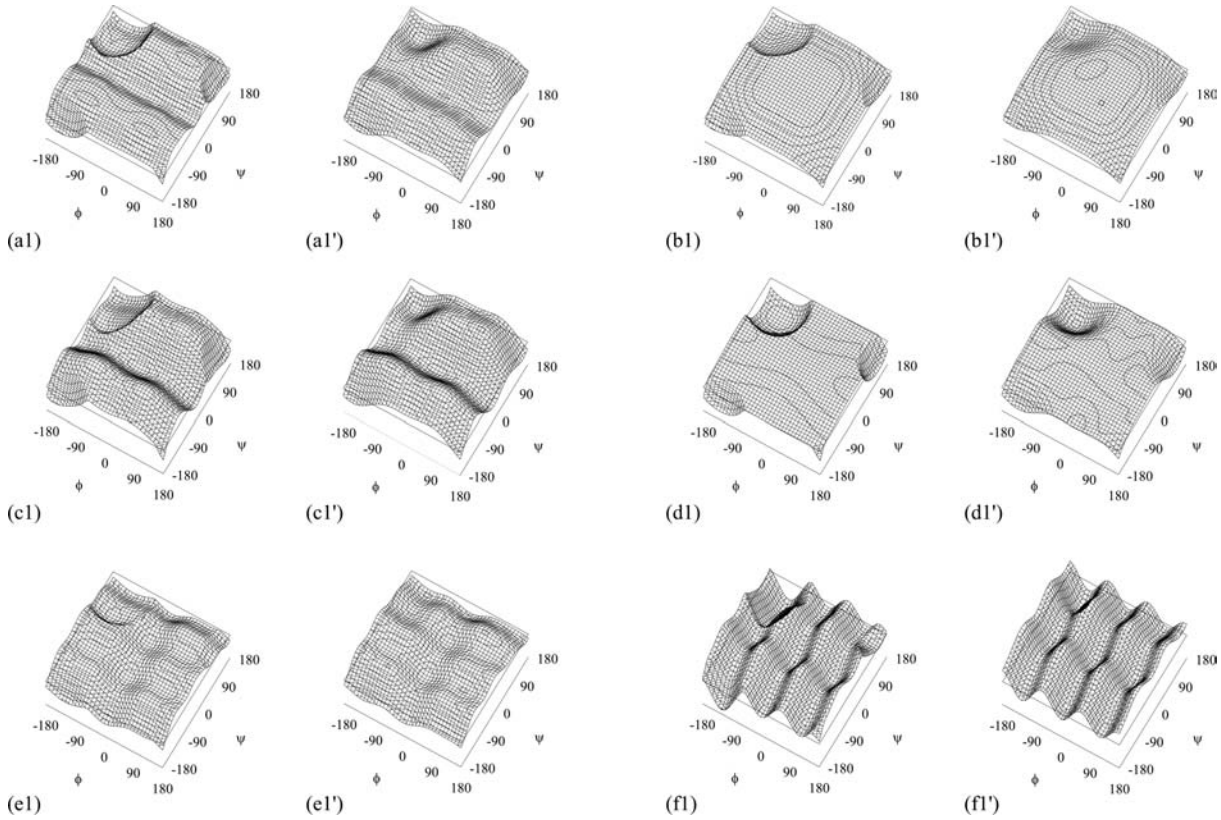

Figure 3. Backbone torsion-energy surfaces of six force fields that were modified by Equations (14), (15) and (17). (a1) to (f1) are those of AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L force fields that were modified to enhance β-sheet structures, respectively. (a1′) to (f1′) are those of AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L force fields that were expanded by the truncated double Fourier series in Equation (17).

Table 9. Number of final conformations with secondary structures obtained from the folding simulations of C-peptide. The total number of folding simulations in each case was 60. See also the captions of Tables 3–8. 'α-Helix', '3_{10}-helix', 'π-helix', 'β-bridges' and 'extended strand' stand for the number of conformations including the amino acids which were identified by DSSP.

	Helix			β-Structure	
	α-Helix	3_{10}-Helix	π-Helix	β-Bridge	Extended strand
org94	49/60	29/60	0/60	0/60	0/60
mod94(α)	59/60	1/60	0/60	0/60	0/60
mod94(β)	0/60	0/60	0/60	4/60	15/60
org96	19/60	2/60	1/60	11/60	4/60
mod96(α)	60/60	0/60	0/60	0/60	0/60
mod96(β)	0/60	0/60	0/60	9/60	26/60
org99	34/60	32/60	0/60	0/60	0/60
mod99(α)	56/60	15/60	0/60	0/60	0/60
mod99(β)	0/60	0/60	0/60	3/60	16/60
orgch	12/60	0/60	15/60	6/60	0/60
modch(α)	48/60	1/60	37/60	0/60	0/60
modch(β)	0/60	0/60	0/60	5/60	11/60
orgop	11/60	6/60	0/60	7/60	5/60
modop(α)	52/60	7/60	0/60	0/60	0/60
modop(β)	2/60	3/60	0/60	13/60	13/60
orgopl	15/60	29/60	0/60	5/60	0/60
modopl(α)	60/60	0/60	0/60	0/60	0/60
modopl(β)	0/60	0/60	0/60	9/60	26/60

Table 10. Number of final conformations with secondary structures obtained from the folding simulations of G-peptide. The total number of folding simulations in each case was 60. See also the captions of Tables 3–8. 'α-Helix' stands for the number of conformations including the amino acids which were identified to be 'H'(= α-helix) by DSSP. 'α-Helix', '3_{10}-helix', 'π-helix', 'β-bridges' and 'extended strand' stand for the number of conformations including the amino acids which were identified by DSSP.

	Helix			β-Structure	
	α-Helix	3_{10}-Helix	π-Helix	β-Bridge	Extended strand
org94	47/60	36/60	1/60	0/60	0/60
mod94(α)	60/60	4/60	0/60	0/60	0/60
mod94(β)	0/60	0/60	0/60	2/60	22/60
org96	12/60	2/60	0/60	12/60	7/60
mod96(α)	60/60	1/60	0/60	0/60	0/60
mod96(α)	0/60	0/60	0/60	3/60	16/60
org99	24/60	46/60	0/60	0/60	0/60
mod99(α)	58/60	29/60	0/60	0/60	0/60
mod99(β)	0/60	0/60	0/60	0/60	13/60
orgch	5/60	0/60	1/60	0/60	0/60
modch(α)	53/60	0/60	27/60	0/60	0/60
modch(β)	0/60	0/60	0/60	1/60	1/60
orgop	8/60	0/60	0/60	12/60	4/60
modop(α)	44/60	18/60	0/60	1/60	0/60
modop(β)	0/60	1/60	0/60	8/60	19/60
orgopl	7/60	20/60	1/60	2/60	1/60
modopl(α)	25/60	13/60	0/60	8/60	7/60
modopl(β)	0/60	0/60	0/60	13/60	16/60

β-sheet structures, conformations 1 and 3 have β-bridge structures and conformations 2, 4–7 have extended β-strand structures of C-peptide. However, only one conformation (no. 3) has an extended β-strand structure of G-peptide. We see that there are few β-structures in the case of CHARMM27 for G-peptide (see Table 10). For the original OPLS-AA (see Figure 8), conformations 2, 3, 5 and 7 are α-helix structures for C-peptide, and conformations 1, 3, 4 and 7 are β-bridge structures and conformations 2 and 6 are extended β-strand structures for G-peptide. From these results, we see that OPLS-AA favours α-helix for C-peptide and favours β-structures for G-peptide. These secondary-structure-forming tendencies agree with experimental implications. For OPLS-AA modified towards α-helix structures, all conformations (nos 1–7) have α-helix structures for both C-peptide and G-peptide. For OPLS-AA modified towards β-sheet structures, one conformation (no. 2) is a β-bridge structure and conformations 1, 4 and 5 are extended β-strand structures, and one conformation (no. 4) is an α-helix structure in the case of C-peptide and one conformation (no. 7) is a β-bridge structure and conformations 3, 4 and 6 are extended β-strand structures in the case of G-peptide. These results show that OPLS-AA modified towards α-helix and β-sheet structures favours α-helix structures and β-structures, respectively. However, these secondary-structure-forming tendencies are slightly weaker than those of AMBER parm94, AMBER parm96 and AMBER parm99 modified towards α-helix and β-sheet structures. In Figure 9, we see that the results for OPLS-AA/L are roughly the same as in the case of OPLS-AA.

Our modifications of the force fields resulted in the expected changes in the secondary-structure formations except for CHARMM27 modified towards β-sheet for G-peptide. For CHARMM27 modified towards β-sheet, only 1 out of 60 conformations was a β-bridge and one out of 60 conformations was an extended strand in spite of the central values $f(\tilde{\phi}_0, \tilde{\psi}_0)$ that we used in Equation (18) were 6.0 kcal/mol for enhancing β-sheet in the case of CHARMM27. However, in the case of C-peptide, the CHARMM27 modified towards β-sheet favours β-structures clearly. We believe that force-field parameters that are related to the side chains of amino acids cause this difference between C-peptide and G-peptide.

Therefore, regardless of the secondary-structure-forming tendencies of the original force fields, our modifications of the backbone torsion-energy term succeeded in enhancing the desired secondary structures.

4. Conclusions

In this article, we gave the details of our recently proposed backbone torsion-energy term that is represented by a double Fourier series in two variables, the backbone dihedral angles ϕ and ψ. This is a natural generalisation of the conventional torsion-energy terms. It is particularly useful in controlling secondary-structure-forming tendencies, because any function in the Ramachandran space can be expanded by this double Fourier series. We can

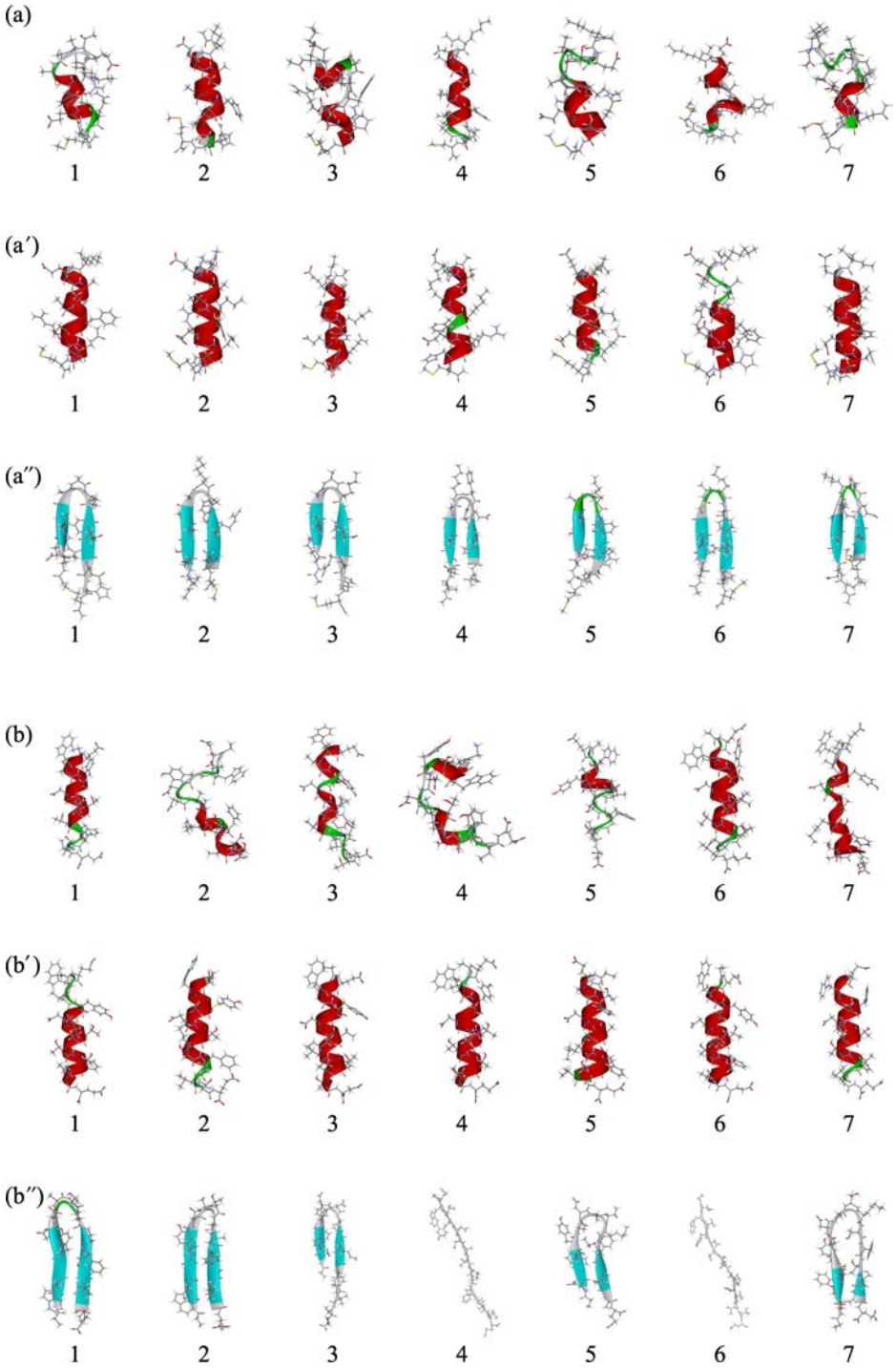

Figure 4. Seven lowest-energy final conformations of C-peptide (a)–(a″) and G-peptide (b)–(b″) obtained from six sets of 60 simulated annealing MD runs. (a) and (b) are the results of the original AMBER parm94. (a′) and (b′) are the results of AMBER parm94 of the truncated double Fourier series of six force fields that were modified to enhance α-helix structures. (a″) and (b″) are the results of AMBER parm94 of the truncated double Fourier series of six force fields that were modified to enhance β-sheet structures. The conformations are ordered in the increasing order of energy for each case. The figures were created with Accelrys DS Visualizer v1.5 (software available at www.accelrys.com/).

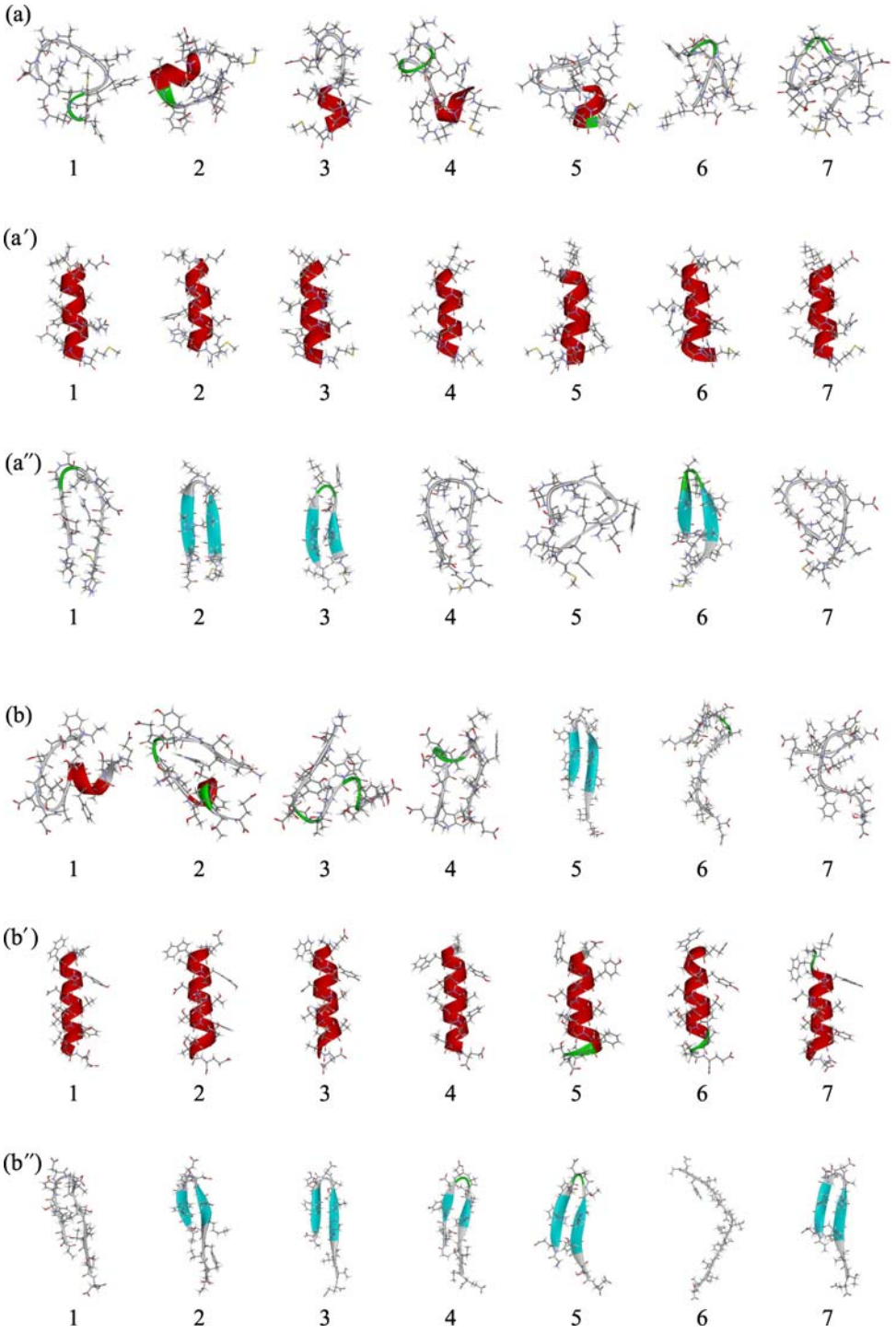

Figure 5. Seven lowest-energy final conformations of C-peptide (a)–(a″) and G-peptide (b)–(b″) obtained from six sets of 60 simulated annealing MD runs. (a) and (b) are the results of the original AMBER parm96. (a′) and (b′) are the results of AMBER parm96 of the truncated double Fourier series of six force fields that were modified to enhance α-helix structures. (a″) and (b″) are the results of AMBER parm96 of the truncated double Fourier series of six force fields that were modified to enhance β-sheet structures. The conformations are ordered in the increasing order of energy for each case. The figures were created with Accelrys DS Visualizer v1.5 (software available at www.accelrys.com/).

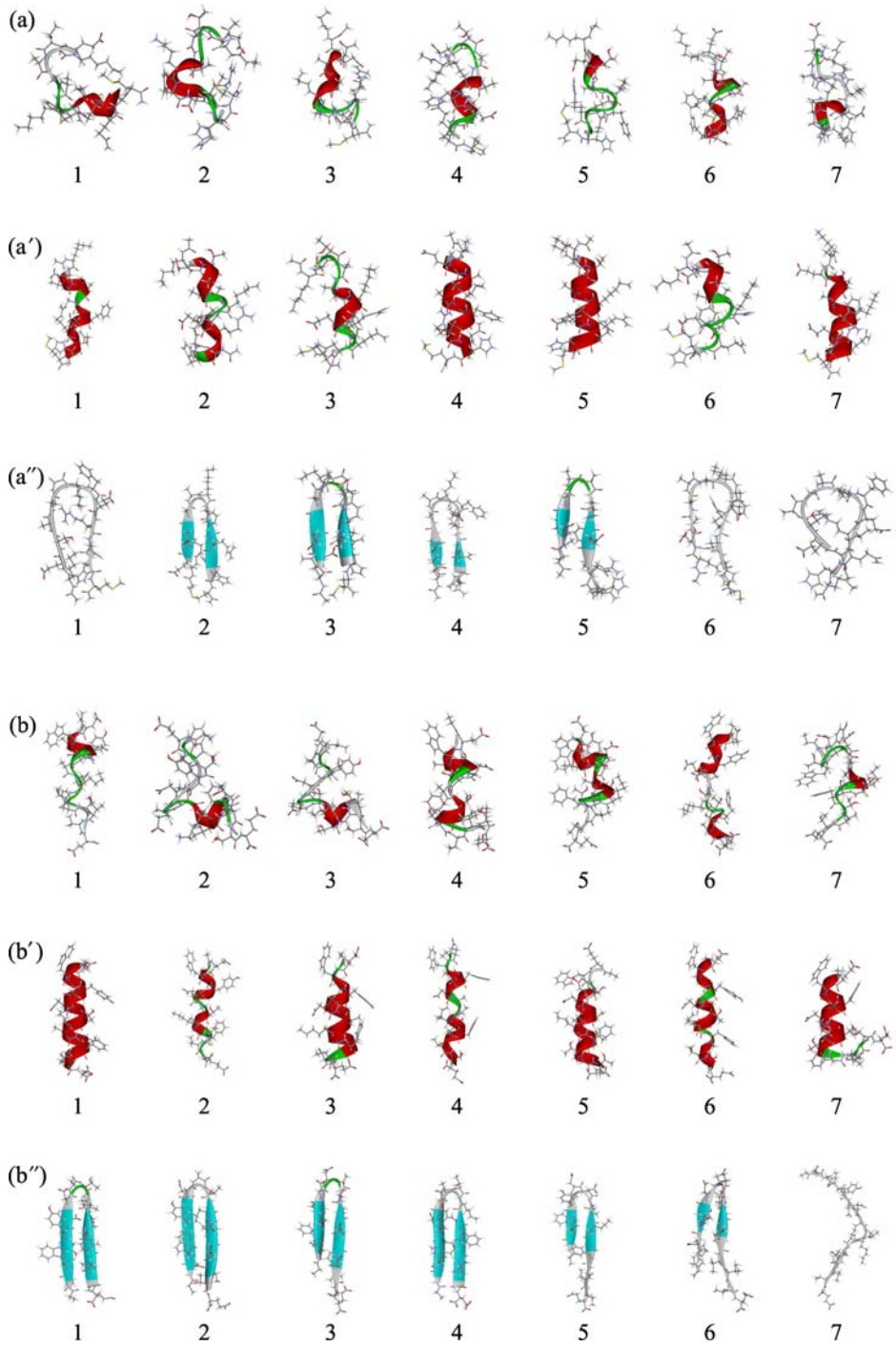

Figure 6. Seven lowest-energy final conformations of C-peptide (a)–(a″) and G-peptide (b)–(b″) obtained from six sets of 60 simulated annealing MD runs. (a) and (b) are the results of the original AMBER parm99. (a′) and (b′) are the results of AMBER parm99 of the truncated double Fourier series of six force fields that were modified to enhance α-helix structures. (a″) and (b″) are the results of AMBER parm99 of the truncated double Fourier series of six force fields that were modified to enhance β-sheet structures. The conformations are ordered in the increasing order of energy for each case. The figures were created with Accelrys DS Visualizer v1.5 (software available at www.accelrys.com/).

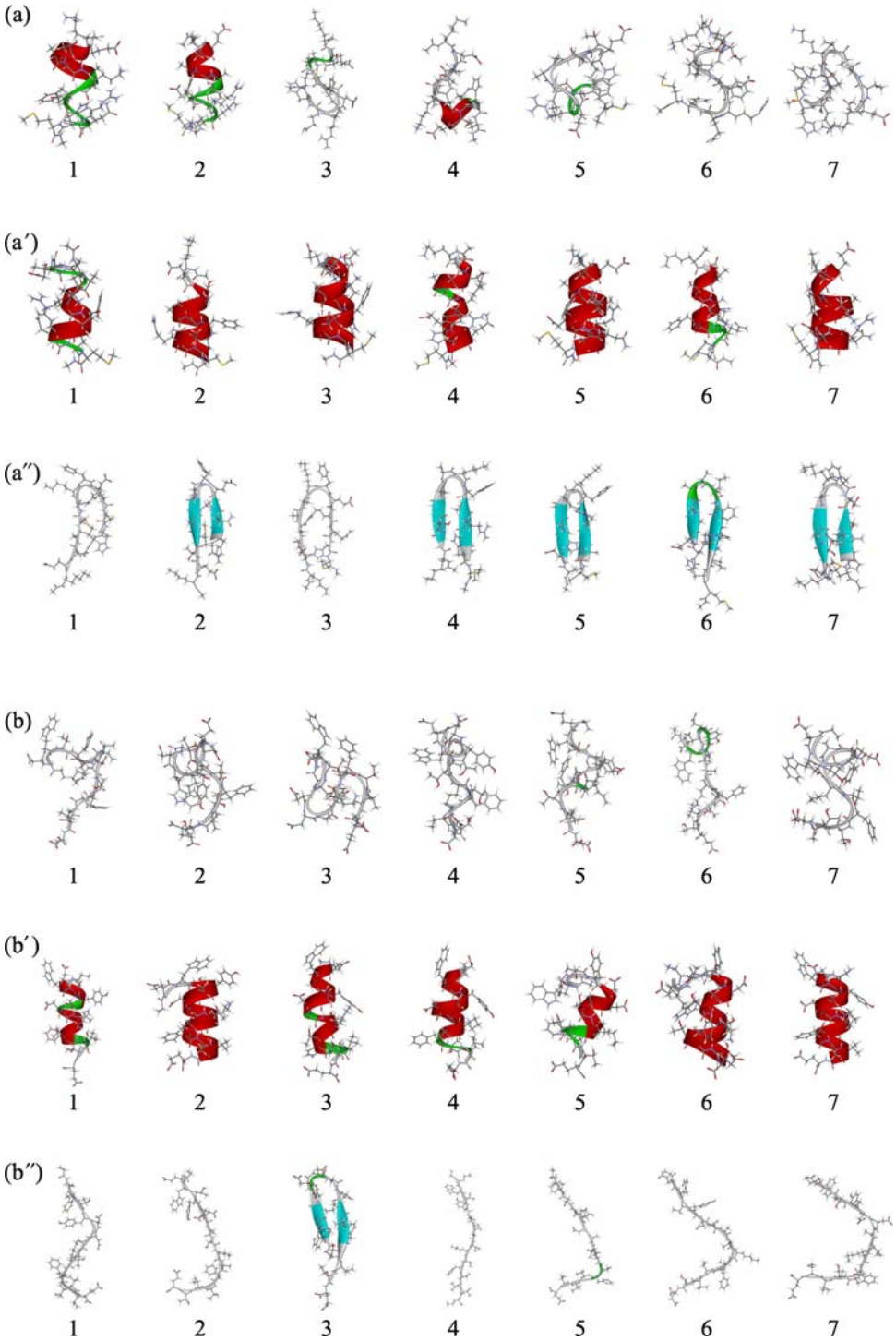

Figure 7. Seven lowest-energy final conformations of C-peptide (a)–(a″) and G-peptide (b)–(b″) obtained from six sets of 60 simulated annealing MD runs. (a) and (b) are the results of the original CHARMM27. (a′) and (b′) are the results of CHARMM27 of the truncated double Fourier series of six force fields that were modified to enhance α-helix structures. (a″) and (b″) are the results of CHARMM27 of the truncated double Fourier series of six force fields that were modified to enhance β-sheet structures. The conformations are ordered in the increasing order of energy for each case. The figures were created with Accelrys DS Visualizer v1.5 (software available at www.accelrys.com/).

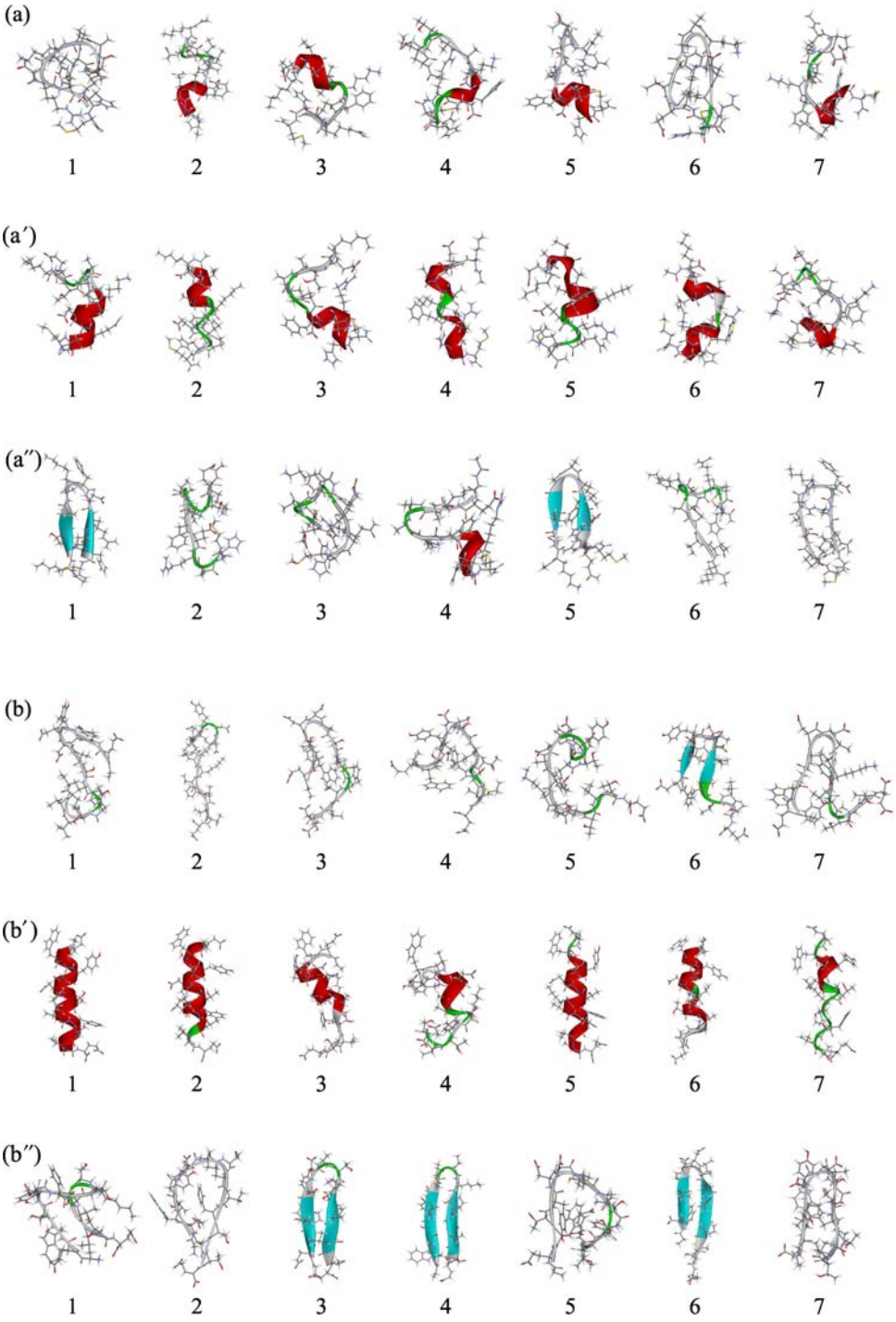

Figure 8. Seven lowest-energy final conformations of C-peptide (a)–(a″) and G-peptide (b)–(b″) obtained from six sets of 60 simulated annealing MD runs. (a) and (b) are the results of the original OPLS-AA. (a′) and (b′) are the results of OPLS-AA of the truncated double Fourier series of six force fields that were modified to enhance α-helix structures. (a″) and (b″) are the results of OPLS-AA of the truncated double Fourier series of six force fields that were modified to enhance β-sheet structures. The conformations are ordered in the increasing order of energy for each case. The figures were created with Accelrys DS Visualizer v1.5 (software available at www.accelrys.com/).

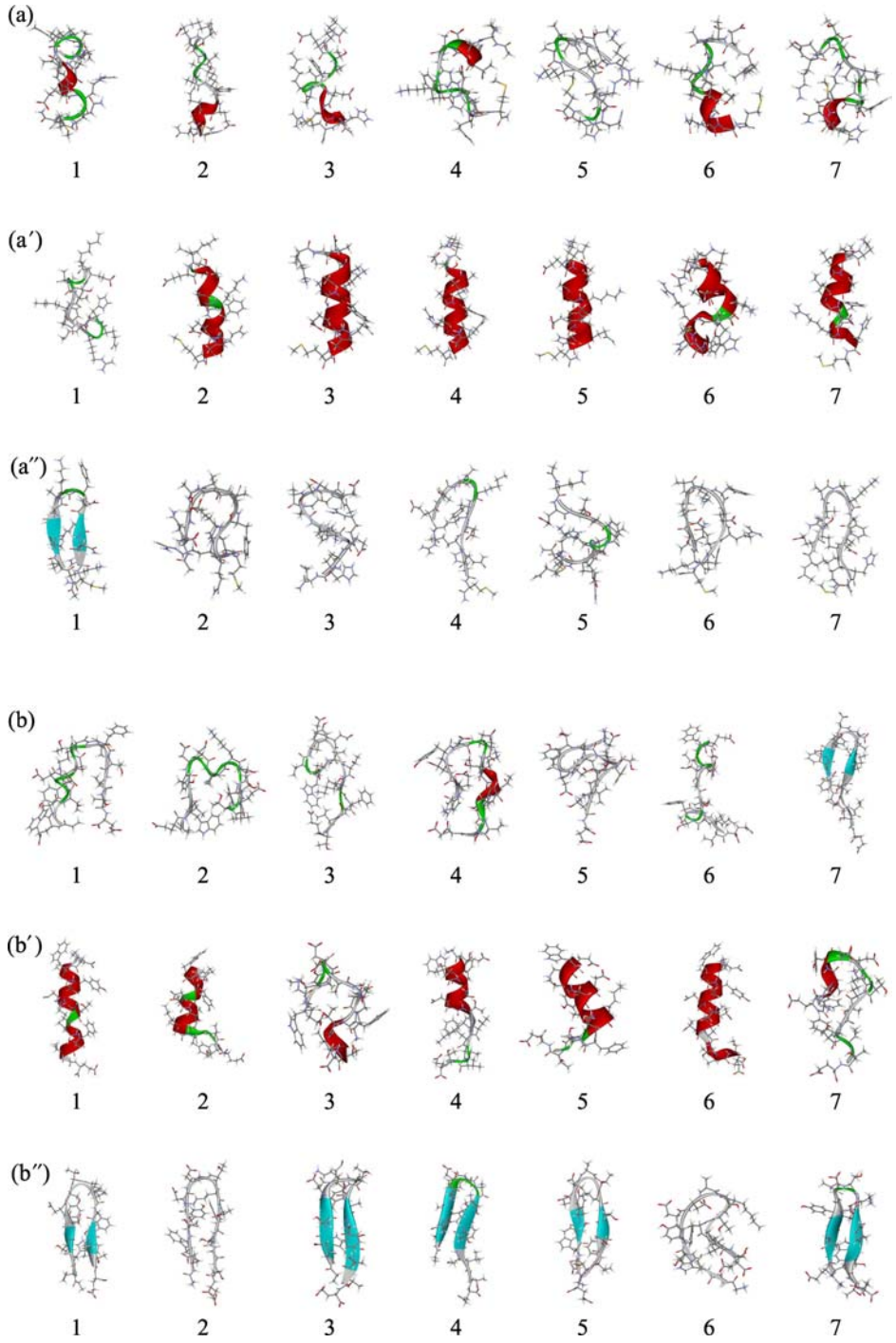

Figure 9. Seven lowest-energy final conformations of C-peptide (a)–(a″) and G-peptide (b)–(b″) obtained from six sets of 60 simulated annealing MD runs. (a) and (b) are the results of the original OPLS-AA/L. (a′) and (b′) are the results of OPLS-AA/L of the truncated double Fourier series of six force fields that were modified to enhance α-helix structures. (a″) and (b″) are the results of OPLS-AA/L of the truncated double Fourier series of six force fields that were modified to enhance β-sheet structures. The conformations are ordered in the increasing order of energy for each case. The figures were created with Accelrys DS Visualizer v1.5 (software available at www.accelrys.com/).

easily modify the existing force fields so that specified secondary-structure-forming tendencies are enhanced, by lowering the backbone torsion-energy surface in the corresponding regions of the Ramachandran space. We demonstrated this by taking the examples of AMBER parm94, AMBER parm96, AMBER parm99, CHARMM27, OPLS-AA and OPLS-AA/L force fields.

Besides the above 'manual' adjustment of the force fields, we can also apply our force-field refinement method [13–15] to this double Fourier series. Namely, we can determine the values of the Fourier coefficients so that the forces acting on the atoms with the coordinates of the PDB database become minimal. Work is in progress in this direction.

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, M. Kenneth, J. 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, *Computer Simulations of Biological Systems*, 3, Escom, The Netherlands, 1997, p. 83.

[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] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

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

[11] 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.

[12] 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.

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

[14] Y. Sakae and Y. Okamoto, *Protein force-field parameters optimized with the protein data bank. I. Force-field optimizations*, J. Theoret. Comput. Chem. 3 (2004), pp. 339–358.

[15] 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. Theoret. Comput. Chem. 3 (2004), pp. 359–378.

[16] V. Hornak, R. Abel, A. 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.

[17] 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. Comp. Chem. 25(11) (2004), pp. 1400–1415.

[18] A. MacKerell Jr., M. Feig, and D.N. Brems, *Improved treatment of the protein backbone in empirical force fields*, J. Am. Chem. Soc. 126 (2004), pp. 698–699.

[19] 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.

[20] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman, *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.

[21] M. Iwaoka and S. Tomoda, *The saap force field. a simple approach to a new all-atom protein force field by using single amino acid potential (saap) functions in various solvents*, J. Comput. Chem. 24 (2003), pp. 1192–1200.

[22] 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.

[23] A. MacKerell Jr., M. Feig, and C. Brooks III, *Improved treatment of the protein backbone in empirical force fields*, J. Am. Chem. Soc. 126 (2004), pp. 698–699.

[24] N. Kamiya, Y. Watanabe, S. Ono, and J. Higo, *Amber-based hybrid force field for conformational sampling of polypeptides*, Chem. Phys. Lett. 401 (2005), pp. 312–317.

[25] 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.

[26] Y. Sakae and Y. Okamoto, *Secondary-structure design of proteins by a backbone torsion energy*, J. Phys. Soc. Jpn. 75 (2006), 054802.
[27] G.N. Ramachandran and V. Sasisekharan, *Conformation of polypeptides and proteins*, Adv. Protein Chem. 23 (1968), pp. 283–438.
[28] S. Honda, N. Kobayashi, and E. Munekata, *Thermodynamics of a β^2-hairpin structure: Evidence for cooperative formation of folding nucleus*, J. Mol. Biol. 295 (2000), pp. 269–278.
[29] 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.
[30] J.J. Osterhout, Jr, R.L. Baldwin, E.J. York, J.M. Stewart, H.J. Dyson, and P.E. Wright, *1 h NMR studies of the solution conformations of an analogue of the C-peptide of ribonuclease a*, Biochemistry 28 (1989), pp. 7059–7064.
[31] F.J. Blanco, G. Rivas, and L. Serrano, *A short linear peptide that folds into a native stable bold beta-hairpin in aqueous solution*, Nat. Struct. Biol. 1 (1994), pp. 584–590.
[32] N. Kobayashi, S. Honda, H. Yoshii, H. Uedaira, and E. Munekata, *Complement assembly of two fragments of the streptococcal protein gb1 domain in aqueous solution*, FEBS Lett. 366 (1995), pp. 99–103.
[33] S. Kirkpatrick, C.D. Gelatt Jr., and M.P. Vecchi, *Optimization by simulated annealing*, Science 220 (1983), pp. 671–680.
[34] 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.
[35] 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.
[36] 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.
[37] W. Kabsch and C. Sander, *Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features*, Biopolymers 22 (1983), pp. 2577–2637.