# OPTIMIZATIONS OF PROTEIN FORCE-FIELD PARAMETERS WITH THE PROTEIN DATA BANK

A Thesis
Presented to the Department of Functional Molecular Science
School of Physical Sciences
The Graduate University for Advanced Studies
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Science

by

Yoshitake Sakae

September 2004

# Acknowledgements

My most heartfelt thanks go to my thesis advisor, Professor Yuko Okamoto, for his patient guidance and constant encouragement. I am greatly indebted to Drs. Hiroshi Noguchi, Hisashi Okumura, Takao Yoda, Takehiro Nagashima, Yuji Sugita, and Yukio Kawashima for useful discussions. I would like to thank Hironori Kokubo, Katsumi Murata, and Satoru Ito. I wish to express my gratitude to all the members of the Okamoto Group for discussions and encouragement. I am grateful to the members of the IMS theory groups for their generous support. I wish to thank the faculty and staff members of the Graduate University for Advanced Studies for their kindness. Finally, I am most thankful to my parents, Yasuo and Kazuyo Sakae, and my sister, Noriko Sakae, for their constant encouragement and support.

# Contents

# Chapter 1

# General Introduction

Native proteins have particular tertiary structures with specific biological functions. Tertiary structures of proteins are determined by their amino-acid sequence information. Each biological function of proteins is therefore closely related to their amino-acid sequences. However, the mechanism of protein folding has yet to be clarified. One of effective approaches to study the protein folding problem is a computational one. Various simulation techniques such as Monte Carlo (MC) and molecular dynamics (MD) methods have been developed. These methods often use a potential energy function, which is also referred to as a force field. The potential energy fucntion for the protein systems is very important ; if it is not accurate, one can not expect to obtain native conformations by molecular simulations. The goal of the present Thesis is address the problem of accuracy of existing force field and to develop an ideal potential energy function.

There are two major approaches for developing protein force field. One is a top-down approach, and the other a bottom-up approach. In the former, the database of known native protein tertiary structures is used (knowledge-based energy functions). It is often based on coarse-grained models of proteins. In the latter, a all-atom models of proteins a usually use and energy function parameters are determined to reproduce the physical data of small molecules.

One of the knowledge-based energy functions is based on statistical mechanics [1]. As a first approximation, only "contact" energy term, or spatially close interactions between pairs of amino acids, is used, and the medium and long-range interactions are neglected. The parameters for this interaction term are determined from the probability that each amino acid exists near another one. This method was further modified and refined method [2, 3]. There also exist other methods based on Bayes statistics [4, 5]. However, the formalisms for all these methods are essentially equivalent to one another.

In the bottom-up approach, examples of the energy terms are the bond-stretching, bond-bending, torsion-energy, Lennard-Jones, and Coulomb terms. The total number of parameters for these energy functions can be thousands. The energy parameters are determined by fitting the physical data of small molecules such as individual amino acids to those of experiments or quantum chemistry calculations. Well-known force fields (or potential energy functions) are AMBER [6]–[8], CHARMM [9], OPLS [10, 11], GROMOS

[12], and ECEPP [13]. These force fields are most widely used.

In order to compare the accuracy of energy parameters among force fields, we need a criterion for good force fields. One such measure is the Z-score [14, 15, 16]:

$$Z = \frac{E_N - < E >_{decoy}}{\sigma_{decoy}} , \tag{1.1}$$

where $E_N$ is the energy of the native structure, $< E >_{decoy}$ is the average energy of non-native structures (they are often referred to as decoys), and $\sigma_{decoy}$ is the corresponding standard deviation. Therefore, Z-score is the deviation of the energy of the native structure from the average energy of alternative conformations measured in units of standard deviation. One of popular optimization methods for force-field parameters is to find the parameters of the energy function that minimizes Z-score (maximizes the energy gap) for many proteins. This method, however, has a problem that the value of Z-score depends on the choice of the decoys and their number. The another method for force field optimization uses the spin glass theory [17], and the energy parameters are optimized by maximizing the ratio of the folding temperature to the glass transition temperature. An iterative algorithm that maximizes the free-energy difference between the native structure and the non-native states of the protein is also proposed [18]. However, the strategy of these methods is essentially the same as that of the above Z-score method.

In this Thesis, we propose a novel optimization method for the protein force-field parameters. Our method is different from most of previous works mainly in two points: We deal with all-atom force fields such as AMBER and we use only the Protein Data Bank (PDB) data without introducing decoys or Z-scores.

This Thesis is organized as follows. In Chapter 2, we describe our optimization method for force-field parameters and the folding simulation method. In Chapter 3, we present the results of the applications of our method to five force-field parameters sets. In Chapter 4, we give the results of the folding simulations of two peptides based on the new optimized force-field parameters. In Chapter 5, we present conclusions and future prospects.

# Bibliography


[1] S. Tanaka, H.A. Scheraga, Macromol. 9 (1976) 945.

[2] S. Miyazawa, R.L. Jernigan, Macromol. 18 (1985) 534.

[3] M.J. Sippl, J. Mol. Biol. 213 (1990) 859.

[4] K.T. Simons, C. Kooperberg, E. Huang, D. Baker, J. Mol. Biol. 268 (1997) 209.

[5] R. Samudrala, J. Moult, J. Mol. Biol. 275 (1998) 895.

[6] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, J. Am. Chem. Soc. 117 (1995) 5179.

[7] P.A. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, A. Pohorille, Computer Simulation of Biomolecular Systems, Eds. A. Wilkinson, P. Weiner and W.F. van Gunsteren 3 (1997) 83.

[8] J. Wang, P. Cieplak, P.A. Kollman, J. Comput. Chem. 21 (2000) 1049

[9] A.D. MacKerrell, 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. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, J. Phys. Chem. B 102 (1998) 3586.

[10] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, J. Am. Chem. Soc. 118 (1996) 11225.

[11] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, J. Phys. Chem. B 105 (2001) 6474.

[12] W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi (1996) Biomolecular simulation: the GROMOS96 manual and user guide. Zürich: Vdf Hochschulverlag AG an der ETH Zürich.

[13] M.J. Sippl, G. Némethy, H.A. Scheraga, J. Phys. Chem. 88 (1984) 6231, and references therein.

[14] L.A. Mirny, E.I. Shakhnovich, J. Mol. Biol. 264 (1996) 1164.

[15] S. Takada, Z.A. Luthey-Schulten, P.G. Wolynes, J. Chem. Phys. 110 (1999) 11616.

[16] Y. Xia, M. Levitt, J. Chem. Phys. 113 (2000) 9318.

[17] R.A. Goldstein, Z.A. Luthey-Schulten, P.G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 4918.

[18] M-H. Hao and H.A. Scheraga, J. Phys. Chem. 100 (1996) 14540.

# Chapter 2

# Methods

Yoshitake Sakae and Yuko Okamoto, "Optimization of protein force-field parameters with the Protein Data Bank," *Chemical Physics Letters* **382**, 626–636 (2003).

Yoshitake Sakae and Yuko Okamoto, "Protein force-field parameters optimized with the Protein Data Bank. I. Force-field optimizations," *Journal of Theoretical and Computational Chemistry* **3**, 339–358 (2004).

## 2.1 Optimization method of protein force-field parameters

The force fields for protein systems such as AMBER [1]–[3], CHARMM[4], and OPLS[5, 6] use essentially the same functional forms for the potential energy except for minor differences. The commonly used conformational potential energy $E_{\rm conf}$ is given by

$$E_{\rm conf} = E_{\rm BL} + E_{\rm BA} + E_{\rm torsion} + E_{\rm nonbond} \ , \tag{2.1}$$

where

$$E_{\rm BL} = \sum_{\text{bond length } \ell} K_\ell (\ell - \ell_{\rm eq})^2 \ , \tag{2.2}$$

$$E_{\rm BA} = \sum_{\text{bond angle } \theta} K_\theta (\theta - \theta_{\rm eq})^2 \ , \tag{2.3}$$

$$E_{\rm torsion} = \sum_{\text{dihedral angle } \Phi} \sum_n \frac{V_n}{2}[1 + \cos(n\Phi - \gamma_n)] \ , \tag{2.4}$$

$$E_{\rm nonbond} = \sum_{i<j} \left[ \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{332 q_i q_j}{\epsilon r_{ij}} \right] . \tag{2.5}$$

Here, $E_{\rm BL}$, $E_{\rm BA}$, and $E_{\rm torsion}$ represent the bond-stretching term, the bond-bending term, and the torsion-energy term, respectively. The bond-stretching and bond-bending energies are given by harmonic terms with the force constants, $K_\ell$ and $K_\theta$, and the equilibrium positions, $\ell_{\rm eq}$ and $\theta_{\rm eq}$. The torsion energy is, on the other hand, described by the Fourier series in Eq. (2.4), where the sum is taken over all dihedral angles $\Phi$, $n$ is the number of waves, $\gamma_n$ is the phase, and $V_n$ is the Fourier coefficient. The nonbonded energy in Eq. (2.5) is represented by the Lennard-Jones and Coulomb terms between pairs of atoms, $i$ and $j$, separated by the distance $r_{ij}$ (in Å). The parameters $A_{ij}$ and $B_{ij}$ in Eq. (2.5) are the coefficients for the Lennard-Jones term, $q_i$ (in units of electronic charges) is the partial charge of the $i$-th atom, and $\epsilon$ is the dielectric constant, where we usually set $\epsilon = 1$ (the value in vacuum). The factor 332 in the electrostatic term is a constant to express energy in units of kcal/mol. Hence, we have five classes of force-field parameters, namely, those in the bond-stretching term ($K_\ell$ and $\ell_{\rm eq}$), those in the bond-bending term ($K_\theta$ and $\theta_{\rm eq}$), those in the torsion term ($V_n$ and $\gamma_n$), those in the Lennard-Jones term ($A_{ij}$ and $B_{ij}$), and those in the electrostatic term ($q_i$).

Eq. (2.1) represents a standard set of the potential energy terms. As mentioned above, there are minor differences in the energy functions among different force fields. For instance, the Urey-Bradley term is used in CHARMM and OPLS, but not in AMBER. In our method, we try to optimize a certain set of parameters in the existing force fields without changing the functional forms. Therefore, if the original force field has non-standard terms, then the optimized one also has them.

Our optimization method for these force-field parameters is now described [7]. We first retrieve $N$ native structures (one structure per protein) from Protein Data Bank (PDB) [8]. We try to choose proteins from different folds (such as all $\alpha$-helix, all $\beta$-sheet, $\alpha/\beta$, etc.) and different homology classes as much as possible. If the force-field parameters are of ideal values, then all the chosen native structures are stable without any force acting on each atom in the molecules. Hence, we expect

$$F = 0 \ , \tag{2.6}$$

where

$$F = \sum_{m=1}^{N} \frac{1}{N_m} \sum_{i_m=1}^{N_m} \left| \vec{f}_{i_m} \right|^2 \ , \tag{2.7}$$

and

$$\vec{f}_{i_m} = -\frac{\partial E_{\rm tot}^{\{m\}}}{\partial \vec{x}_{i_m}} \ . \tag{2.8}$$

Here, $N_m$ is the total number of atoms in molecule $m$, $E_{\rm tot}^{\{m\}}$ is the total potential energy for molecule $m$, $\vec{x}_i$ is the Cartesian coordinate vector of atom $i$, and $\vec{f}_i$ is the force acting on atom $i$. In reality, $F \neq 0$, and because $F \geq 0$, we can optimize the force-field parameters by minimizing $F$ with respect to these parameters. In practice, we perform a simulation in the force-field parameter space for this minimization.

Proteins are usually in aqueous solution, and hence we also have to incorporate some kind of solvent effects. Because the more the total number of proteins ($N$) is, the better the force-field parameter optimizations are expected to be, we want to minimize our efforts in the calculations of the solvent effects. Here, we employ the generalized-Born surface area (GB/SA) terms for the solvent contributions [9, 10]. Hence, we use in Eq. (2.8) (we suppress the label $m$ for each molecule)

$$E_{\rm tot} = E_{\rm conf} + E_{\rm solv} \ , \tag{2.9}$$

where

$$E_{\rm solv} = E_{\rm GB} + E_{\rm SA} \ , \tag{2.10}$$

$$E_{\rm GB} = -166 \left(1 - \frac{1}{\epsilon_s}\right) \sum_{i,j} \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_{ij}^2 e^{-D_{ij}}}} \ , \tag{2.11}$$

$$E_{\rm SA} = \sum_k \sigma_k A_k \ . \tag{2.12}$$

Namely, in the GB/SA model, the total solvation free energy in Eq. (2.10) is given by the sum of a solute-solvent electrostatic polarization term, a solvent-solvent cavity term, and a solute-solvent van der Waals term. A solute-solvent electrostatic polarization term can be calculated by the generalized Born equation (2.11), where $\alpha_{ij} = \sqrt{\alpha_i \alpha_j}$, $\alpha_i$ is the so-called Born radius of atom $i$, $D_{ij} = r_{ij}^2/(2\alpha_{ij})^2$, and $\epsilon_s$ is the dielectric constant of bulk water (we take $\epsilon_s = 78.3$). A solvent-solvent cavity term and a solute-solvent van der Waals term can be approximated by the term in Eq. (2.12) that is proportional to the solvent accessible surface area. Here, $A_k$ is the total solvent-accessible surface area of atoms of type $k$ and $\sigma_k$ is an empirically determined proportionality constant [9, 10].

The flowchart of our method for the optimization of force-field parameters is shown in Fig. 2.1.

In Step 1 of the flowchart we try to obtain as many structures as possible from PDB. The number is limited by the computer power that we have available in our laboratory. We want to choose proteins with different sizes (numbers of amino acids), different folds, and different homology classes as much as possible. We also want to use only those with high experimental resolutions. Note that only atomic coordinates of proteins are extracted from PDB (and coordinates from other molecules such as crystal water are neglected).

If we use data from X-ray experiments, hydrogen atoms are missing, and thus in Step 2 we have to add hydrogen coordinates. Many protein simulation software packages provide with routines that add hydrogen atoms to the PDB coordinates, and one can use one of such routines.

We now have $N$ protein coordinates ready, but usually such "raw data" result in very high total potential energy and strong forces will be acting on some of the atoms in the molecules. This is because the hydrogen coordinates that we added as above are not based on experimental results and have rather large uncertainties. The coordinates of

heavy atoms from PDB also have experimental errors. We take the position that we leave the coordinates of heavy atoms as they are in PDB as much as possible, and adjust the hydrogen coordinates to reduce this mismatch. This is why we want to include as many PDB data as possible with high experimental resolutions (so that the effects of experimental errors in PDB may be minimal). We thus minimize the total potential energy $E_{\rm tot} = E_{\rm conf} + E_{\rm solv} + E_{\rm constr}$ with respect to the coordinates for each protein conformation, where $E_{\rm constr}$ is the constraint energy term that is imposed on the heavy atoms in PDB (it is referred to as the "predefined constraints" in Steps 3 and 5 in Fig. 2.1):

$$E_{\rm constr} = \sum_{\rm heavy\ atom} K_x(\vec{x} - \vec{x}_0)^2 \ . \tag{2.13}$$

Here, $K_x$ is the force constant of the restriction, and $\vec{x}_0$ are the original coordinate vectors of heavy atoms in PDB. Because we are searching for the nearest local-minimum states, usual minimization routines such as the conjugate-gradient method and Newton-Raphson method can be employed here. As one can see from Eq. (2.13), the coordinates of hydrogen atoms will be mainly adjusted, but unnatural heavy-atom coordinates will also be modified. We perform this minimization for all $N$ protein structures separately, and obtain $N$ refined structures.

Given $N$ sets of "ideal" reference coordinates in Step 3 of the flowchart, we now optimize the first set of force-field parameters in Step 4. In Eq. (2.1) we have five classes of force-field parameters as mentioned above. Namely, the force-field parameters are those in the bond-stretching term ($K_\ell$ and $\ell_{\rm eq}$), those in the bond-bending term ($K_\theta$ and $\theta_{\rm eq}$), those in the torsion term ($V_n$ and $\gamma_n$), those in the Lennard-Jones term ($A_{ij}$ and $B_{ij}$), and those in the electrostatic term ($q_i$). Because they are of very different nature, we believe that it is better to optimize these classes of force-field parameters separately (as in Steps 4, 6, and so on in Fig. 2.1). Note also that if we optimize all the parameters simultaneously, the null result (with all the parameter values equal to zero) is a solution to Eq. (2.6). This is the main reason why we optimize each class of parameters separatlely.

For each set of force-field parameters, the optimization is carried out by minimizing $F$ in Eq. (2.7) with respect to these parameters. Here, $E_{\rm tot}$ in Eq. (2.8) is given by Eq. (2.9). For this purpose usual minimization routines such as the conjugate-gradient

method are not adequate, because we need a global optimization. One should employ more powerful methods such as simulated annealing [11] and generalized-ensemble algorithms [12]. We perform this minimization simulation in the above parameter space to obtain the parameter values that give the global minimum of $F$.

These processes are repeated until the optimized force-field parameters converge.

## 2.2 Simulated annealing for folding simulations

We performed the folding simulations of two peptides, using simulated annealing [11] MD method in order to compare the optimized force fields with the original force fields. Simulated annealing is one of the most widely used global optimization methods. During a simulated annealing simulation, the temperature is lowered very slowly from a sufficiently high initial temperature to a "freezing" temperature. If the rate of temperature decrease is slow enough for the system to stay in thermodynamic equilibrium, then it is ensured that the system can avoid getting trapped in states of energy local minima and that the global minimum will be found.

In the present simulation the temperature was lowered exponentially in $N_S$ MD steps by setting the temperature to

$$T_n = T_I \gamma^{n-1} \ , \tag{2.14}$$

for the $n$-th MD step ($n = 1, 2, \cdots, N_S$). Here, $T_I$ is the initial temperature and $\gamma$ is given by

$$\gamma = \left( \frac{T_F}{T_I} \right)^{1/(N_S-1)} \ , \tag{2.15}$$

where $T_F$ is the final temperature. For a fixed value of $N_S$, these constants ($T_I$ and $T_F$) are free parameters and have to be tuned in such a way that the annealing process is optimized for the specific problem.

# Bibliography


[1] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P.A. Kollman, *J. Am. Chem. Soc.*, **117**, 5179 (1995).

[2] P.A. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, and A. Pohorille, in: A. Wilkinson, P. Weiner, and W.F. van Gunsteren, Elsevier, (Eds.), Computer Simulation of Biomolecular Systems, 3, Elsevier, 1997, p. 83.

[3] J. Wang, P. Cieplak, and P.A. Kollman, *J. Comput. Chem.*, **21**, 1049 (2000).

[4] A.D. MacKerrell, 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. Wiórkiewicz-Kuczera, and D. Yin, M. Karplus, *J. Phys. Chem. B*, **102**, 3586 (1998).

[5] W.L. Jorgensen, D.S. Maxwell, and J. Tirado-Rives, *J. Am. Chem. Soc.*, **118**, 11225 (1996).

[6] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, and W.L. Jorgensen, *J. Phys. Chem. B*, **105**, 6474 (2001).

[7] Y. Sakae and Y. Okamoto, *Chem. Phys. Lett.*, **382**, 626 (2003).

[8] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne, *Nucleic Acids Research*, **28**, 235 (2000).

[9] W.C. Still, A. Tempczyk, R.C. Hawley, and T. Hendrickson, *J. Am. Chem. Soc.*, **112**, 6127 (1990).

[10] D. Qiu, P.S. Shenkin, F.P. Hollinger, and W.C. Still, *J. Phys. Chem. A*, **101**, 3005 (1997).

[11] S. Kirkpatrick, C.D. Gelatt Jr., and M.P. Vecchi, *Science*, **220**, 671 (1983).

[12] A. Mitsutake, Y. Sugita, and Y. Okamoto, *Biopolymers (Pept. Sci.)*, **60**, 96 (2001).

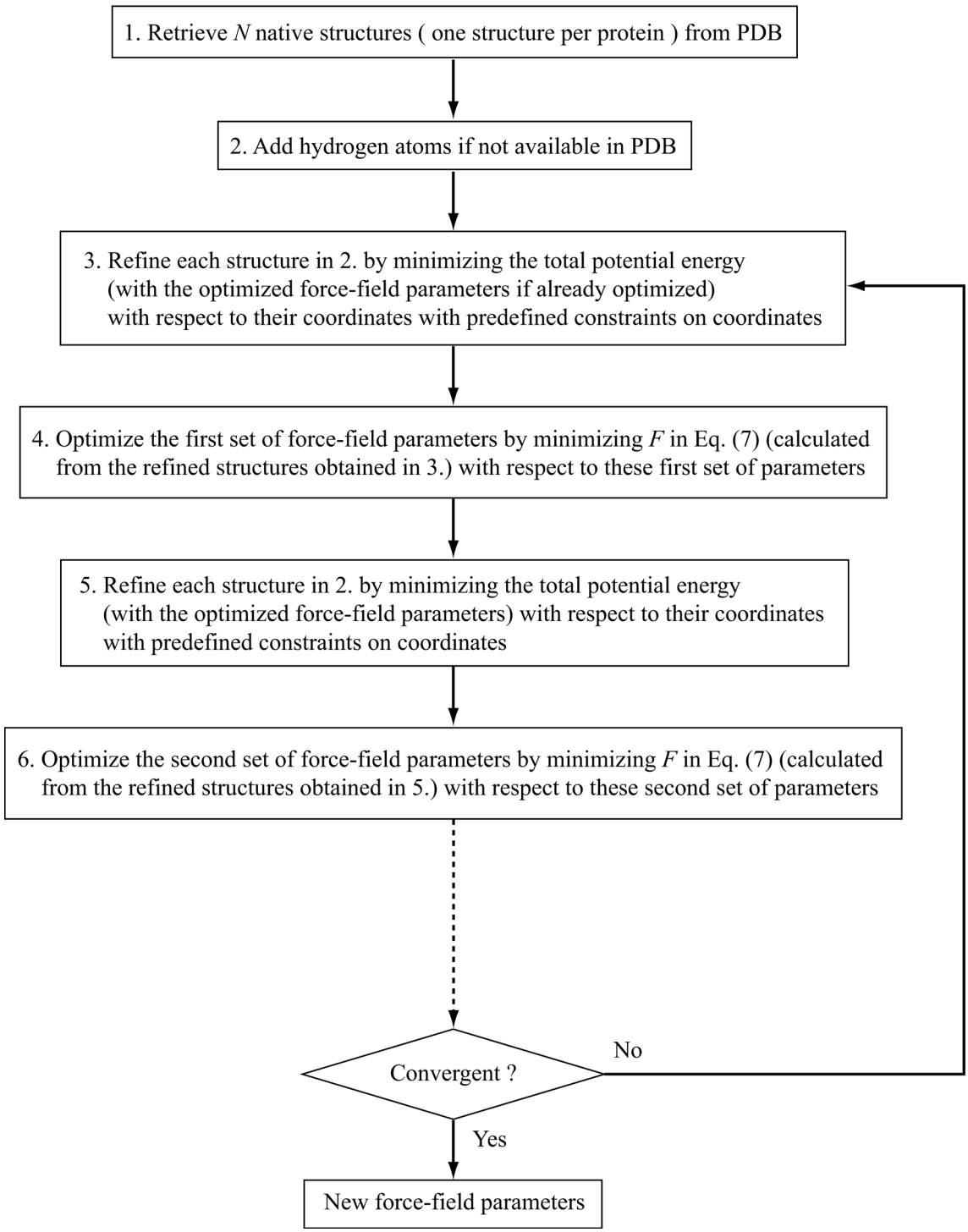


Figure 2.1: The flowchart of our method for the optimization of force-field parameters.

# Chapter 3

# Protein Force Field Optimizations

Yoshitake Sakae and Yuko Okamoto, "Optimization of protein force-field parameters with the Protein Data Bank," *Chemical Physics Letters* **382**, 626–636 (2003).

Yoshitake Sakae and Yuko Okamoto, "Protein force-field parameters optimized with the Protein Data Bank. I. Force-field optimizations," *Journal of Theoretical and Computational Chemistry* **3**, 339–358 (2004).

## 3.1 Introduction

Well-known force fields are AMBER [1]–[3], CHARMM [4], OPLS [5, 6], GROMOS [7], and ECEPP [8]. These force fields have been parameterized to fit the data of small molecules such as individual amino acids to those of experiments and quantum chemistry calculations. The force fields are thus optimal for simulations of these small molecules. However, it is not clear how accurate the force-field parameters are for simulating larger molecules such as proteins and nucleic acids. This is because the simulations of biomolecular systems have the problem that the simulations will get trapped in states of energy local minima, making it difficult to examine the validity of force-field parameters. It is therefore necessary to use powerful simulation algorithms such as generalized-ensemble algorithms (for a recent review, see Ref. [9]). In fact, detailed comparisons of three versions of AMBER (parm94 [1], parm96 [2], and parm99 [3]), CHARMM [4], OPLS-AA/L [6], and GROMOS [7] have recently been carried out by generalized-ensemble simulations of small peptides and it was found that none of these force-field parameters are perfect [10].

We therefore need to improve these existing force fields by some means. Our method [11, 12] utilizes the protein structures in the Protein Data Bank (PDB) [13]. In the previous Chapter we gave the details of the method and In this Chapter we present the results of our applications to five existing force-field parameters. Namely, we have optimized force-field parameters of AMBER parm94 [1], AMBER parm96 [2], AMBER parm99 [3], CHARMM version 22 [4], and OPLS-AA [5].

The validity of these optimized force fields should be tested by performing, for instance, folding simulations of small proteins. In the following Chapter, we will give the results of the folding simulations.

The present Chapter is organized as follows. In Section 2 the computational details are described. In Section 3 the results of the applications of our method to the five force-field parameter sets are presented.

## 3.2 Computational details

We can, in principle, optimize all the force-field parameters following the flowchart in Fig. 2.1. In the examples given below, however, we just optimize two classes of the force-field parameters for simplicity; namely, the partial charges and the backbone torsion-energy parameters. For the optimization of the partial charges ($q_i$), we impose a condition that the total charge of each amino acid remains constant, which is the usual assumption adopted by the force fields of Eq. (2.1) based on classical mechanics. As for the main chain torsion-energy parameters, we use the following functional form for each backbone dihedral angle $\phi$ and $\psi$ (see Eq. (2.4)):

$$E_{\Phi=\phi,\psi} = \frac{V_a}{2}\left[1+\cos(n_a\Phi-\gamma_a)\right]+\frac{V_b}{2}\left[1+\cos(n_b\Phi-\gamma_b)\right]+\frac{V_c}{2}\left[1+\cos(n_c\Phi-\gamma_c)\right] \ . \quad (3.1)$$

We optimize only the parameters ($V_a$, $V_b$, and $V_c$) and fix the number of waves ($n_a$, $n_b$, and $n_c$) and the phases ($\gamma_a$, $\gamma_b$, and $\gamma_c$) as in the original force field. This torsion-energy parameter optimization strongly depends on the values of the force constant $K_x$ of the constraint energy in Eq. (2.13): The larger the values of $K_x$ are, the larger those of $V_a$, $V_b$, and $V_c$ tend to be. In order to minimize such dependences, we impose the constraint that the total area enclosed by the curve of $|E_\Phi|$ (from $\Phi = -180^\circ$ to $180^\circ$) remains less than or equal to the original value during the optimization.

We believe that these two classes of parameters have the most uncertainty among all the force-field parameters. This is because partial charges are usually obtained by quantum chemistry calculations of an isolated amino acid in vacuum separately, which is a very different condition from that in amino acids of proteins in aqueous solution, and because the torsion-energy term is the most problematic (for instance, the parm94, parm96, and parm99 versions of AMBER differ mainly in torsion-energy parameters).

We modified and used the program package TINKER version 4.1 [28] for all the calculations in the present work.

## 3.3 Results and discussion

We now present the results of our force-field optimizations. In Step 1 of the flowchart in Fig. 2.1, we chose 100 PDB files ($N = 100$) from X-ray experiments with resolution 1.8 Å or better and with less than 200 residues (the average number of resiudes is 120.4) from PISCES [29]. Their PDB code names are 2LIS, 1EP0, 1TIF, 1EB6, 1C1L, 1CCW, 2PTH, 1I6W, 1DBF, 1KPF, 1LRI, 1AAP, 1C75, 1CC8, 1FK5, 1KQR, 1K1E, 1CZP, 1GP0, 1KOI, 1IQZ, 3EBX, 1I40, 1EJG, 1AMM, 1I07, 1GK8, 1GVP, 1M4I, 1EYV, 1E29, 1I2T, 1VCC, 1FM0, 1EXR, 1GUT, 1H4X, 1GBS, 1B0B, 119L, 1IFC, 1DLW, 1EAJ, 1GGZ, 1JR8, 1RB9, 1VAP, 1JZG, 1M55, 1EN2, 1C9O, 2ERL, 1EMV, 1F41, 1EW6, 2TNF, 1IFR, 1JSE, 1KAF, 1HZT, 1HQK, 1FXL, 1BKR, 1ID0, 1LQV, 1G2R, 1KR7, 1QTN, 1D4O, 1EAZ, 2CY3, 1UGI, 1IJV, 3VUB, 1BZP, 1JYR, 1DZK, 1QFT, 1UTG, 2CPG, 1I6W, 1C7K, 1I8O, 1LO7, 1LNI, 1EQO, 1NDD, 1HD2, 3PYP, 1FD3, 1DK8, 1WHI, 1FAZ, 4FGF, 2MHR, 1JB3, 2MCM, 1IGD, 1C5E, and 1JIG.

In Step 2 of the flowchart, we used the routine in the TINKER package to add hydrogen atoms to the PDB coordinates. The force fields that we optimized are the AMBER parm94 version [1], parm96 version [2], parm99 version [3], CHARMM version 22 [4], and OPLS-AA [5]. We have optimized only two sets of parameters. The first set is the partial-charge parameters ($q_i$ in Eqs. (2.5) and (2.11)). In order to simplify the constraint-imposing processes on the total charge, we did not optimize the charge of one of the hydrogen atoms (HN) in proline when it is located at tht N-terminus. In the original X-ray data, hydrogen coordinates are missing, and in the case of neutral histidine whether $N_\delta$ and $N_\epsilon$ are protonated or not is non-trivial to determine. Because we want to deal with as many as PDB data as possible, we treated all the histidine residues as positively charged histidine for simplicity. Among the five force fields, AMBER has the largest number of remaining partial-charge parameters (602). We thus optimized these 602 parameters for all the five force fields. The second set of parameters that we optimized is the backbone torsion-energy parameters ($V_a$, $V_b$, and $V_c$ in Eq. (3.1)) and there are six such parameters (three each for $\phi$ and $\psi$).

As explained in detail above, the coodinates of the 100 proteins molecules have been

prepared (Steps 1 and 2 of the flowchart in Fig. 2.1). The coordinate refinement in Step 3 of the flowchart was then carried out with the constraint in Eq. (2.13) on the heavy atoms. As for the force constant $K_x$ in Eq. (2.13), we have some freedom for the choice of the values. Our choice is: $K_x$ should be of the same order as $K_l$ in the bond-stretching term in Eq. (2.2). The force constant $K_l$ in AMBER varies from 166 kcal/mol/Å$^2$ to 656 kcal/mol/Å$^2$, and that in CHARMM varies from 173 kcal/mol/Å$^2$ to 650 kcal/mol/Å$^2$. Hence, in our first trial we set $K_x = 100$ kcal/mol/Å$^2$. After the refinements, we found good coincidence of heavy-atom coordinates and small deviations in the hydrogen coordinates, as expected (data not shown). For each force field (AMBER parm94, AMBER parm96, AMBER parm99, CHARMM version 22, and OPLS-AA), the average root-mean-square distances (RMSD) of all the 100 proteins were 0.029 Å, 0.030 Å, 0.029 Å, 0.029 Å, and 0.032 Å, respectively, for heavy atoms.

As also expected, the structure of the raw PDB data (that obtained in Step 2 of the flowchart) is not stable in the sense that the force acting on some atoms in the molecule will be large, while the refined structure (that obtained in Step 3 of the flowchart) will not have abnormally large forces acting. This is illustrated in Figs. 3.1(a1), 3.1(b1), $\cdots$, 3.1(e1) for the five force fields.

In Step 4 of the flowchart, we performed the optimization of the 602 partial-charge parameters by MC simulated annealing. Namely, we minimized $F$ in Eq. (2.7) by MC simulated annealing simulations of these parameters (the parameters were updated and the updates were accepted or rejected according to the Metropolis criterion). For this we introduced an effective "temperature" for the parameter space. The simulation run consisted of 50,000 MC sweeps with the temperature decreased exponentially from 20 to 0.01. The simulation was repeated 10 times with different initial random numbers. The time series of $F$ from these simulations are shown in Figs. 3.2(a)–3.2(e). We see that $F$ decreases quickly in the beginning until about 5,000 MC sweeps and then it decreases very slowly for all force fields; the total number of MC sweeps (50,000) seems sufficient. The optimized partial charges are taken from those that resulted in the lowest $F$ value.

Each term contributing to $F$ (i.e., each component of the force acting on the atoms) before and after the optimization is compared in Figs. 3.1(a2), 3.1(b2), $\cdots$, 3.1(e2). We

see that many terms decreased in magnitude as a whole.

In Tables 3.1–3.12, five examples (glycine, alanine, glutamic acid, lysine, and tyrosine) of the obtained partial charges together with the original force-field values are listed. We see from these tables that the values of the partial charges have not changed a lot. Although the sign of the partial charges remains the same for those with large magnitude, charges with small magnitude sometimes change their signs (see, for example, CA of glycine, CG of glutamic acid, CB of lysine, and CA, CB, HB, CG, CD, CE, CZ of tyrosine).

In Step 5 of the flowchart, the original coordinates obtained in Step 2 were again refined with the constraints in Eq. (2.13), but this time the optimized parameters from Step 4 were used. This time we used the value $K_x = 500$ kcal/mol/Å$^2$. For all force fields, the average RMSD of the 100 proteins is 0.012 Å, and the coordinates of heavy atoms have little changed.

In Step 6 of the flowchart, we carried out the optimization of the six torsion-energy parameters ($V_a$, $V_b$, and $V_c$ in Eq. (3.1) for both $\phi$ and $\psi$) by minimizing $F$ in Eq. (2.7) with MC simulated annealing simulations in this parameter space. The simulation run consisted of 10,000 MC sweeps with the temperature decreasing from 1,000 to 1.0. The simulation was repeated six times with different random numbers. We stopped after six trials because the convergence was very good. The optimized torsion-energy parameters are taken from those that resulted in the lowest $F$ value. The time series of $F$ from this simulation is shown in Figs. 3.3(a)–3.3(e). We see that $F$ decreases quickly in the beginning until about 3,000 MC sweeps and then the change of $F$ is very little; the total number of MC sweeps (10,000) seems sufficient. The obtained torsion-energy parameters are listed in Tables 3.64 and 3.65. We remark that the present choices of $n$ and $\gamma$ in Eq. (3.1) are slightly different from those used in our previous work [11]. Each term contributing to $F$ before and after the optimization is compared in Figs. 3.1(a3), 3.1(b3), $\cdots$, 3.1(e3). The term corresponding to the backbone atoms that are relevant to the torsion-energy parameters decreased, although not clear in the figures. Non-zero contributions are presumably to be reduced by optimization of other force-field parameters.

In the present work, we stopped our process in Step 6 of the flowchart and did not iterate the optimizations.

In order to examine how much the torsion-energy terms have changed after optimizations, we depict them in Fig. 3.4. Although the behaviors of the original force fields are quite different, those of the optimized force fields are rather similar. For example, the optimized torsion-energy curves for $\phi$ angles have two maximum peaks around $\phi \sim -60^\circ$ and $+60^\circ$ and a local minimum at $\phi = 0^\circ$, while those for $\psi$ angle have two peaks around $\psi \sim -100^\circ$ and $+100^\circ$ and a local minimum at $\psi = 0^\circ$ (the exceptions are those for CHARMM version 22 and OPLS-AA, which give the global maximum and a local maximum, respectively, at $\psi = 0^\circ$). These results suggest that our optimizations of the torsion-energy term yield a tendency for convergence towards a common function. Some remark is in order. The case for the optimized CHARMM is the most distinct from other optimized parameters in the sense that it gives the global maximum as $\psi = 0^\circ$ whereas that for other cases lie around $\psi \sim -100^\circ$ and $+100^\circ$.

In Fig. 3.5 the potential-energy surfaces of the alanine dipeptide (ACE-ALA-NME) are shown for the 10 force-field parameters: the original AMBER parm94, AMBER parm96, AMBER parm99, CHARMM version 22, OPLS-AA, and the corresponding optimized parameters. According to the *ab initio* quantum mechanical calculations, there exist three local-minimum states in the energy surface [1]. They are conformars $C_{7eq}$, $C_5$, and $C_{7ax}$, which correspond to $(\phi, \psi) \sim (-80^\circ, +80^\circ)$, $(-160^\circ, +160^\circ)$, and $(+75^\circ, -60^\circ)$, respectively ($C_{7eq}$ is the global-minimum state). We remark that these are the results of quantum chemistry calculations in vacuum, and so it is not clear how reliable the results are to represent the dipeptide in aqueous solution. The results of all five original force fields in Figs. 3.5(a1)–3.5(e1) seem to satisfy the above conditions. Namely, there are three local-minimum states at the locations of $C_{7eq}$, $C_5$, and $C_{7ax}$, and the global-minimum state is $C_{7eq}$. As for the results of the optimized force fields in Figs. 3.5(a2)–3.5(e2), those for CHARMM version 22 and OPLS-AA also satisfy the above conditions. Those of the optimized AMBER force fields are less consistent with the quantum mechanical calculations: $C_{7eq}$ is no longer the global-minimum state, but it is a local-minimum state. In particular, the optimized AMBER parm99 seems to be in the greatest disagreement in the sense that the $C_{7eq}$ state is almost disappering.

As we pointed out above, the present method for force-field optimizations depends

on the value of the force constant $K_x$ of the constraint term in Eq. (2.13). We have not examined in detail what the value of this constant should be. We just assumed that $K_x$ should be of the same order as $K_l$ in the bond-stretching term in Eq. (2.2), where the force constant $K_l$ varies, for instance, from 166 kcal/mol/Å$^2$ to 656 kcal/mol/Å$^2$ in AMBER. In the present work, we rather arbitrarily set $K_x = 100$ kcal/mol/Å$^2$ for the partial-charge optimizations and $K_x = 500$ kcal/mol/Å$^2$ for the torsion-energy optimizations. We now examine how much $K_x$ dependence we should expect in our method. For this purpose, we just repeat the optimizations of the torsion-energy term for AMBER parm99 and CHARMM version 22 with different values of $K_x$. As discussed above, the optimization results of these two force fields were more in variant than those of other force fields in two respects (see Fig. 3.5 for AMBER parm99 and Fig. 3.4 for CHARMM version 22). We tried to look for the values of $K_x$ where the optimized values of these two force fields approach closer to those of other force fields. We used $K_x = 200$ kcal/mol/Å$^2$ and $K_x = 100$ kcal/mol/Å$^2$ for AMBER parm99 and CHARMM version 22, respectively. The optimized torsion-energy parameters are listed in Tables 3.66 and 3.67, and these should be compared with the corresponding values in Tables 3.64 and 3.65, where $K_x = 500$ kcal/mol/Å$^2$ was used. There seems to be a tendency (especially for the CHARMM case) that the absolute values of $V_a$, $V_b$, and $V_c$ are reduced (because the magnitude of $K_x$ was reduced). In Fig. 3.6 we plot the torsion-energy functions for both cases, and these should be compared with the corresponding curves in Fig. 3.4. As for the optimized AMBER parm99, the curves did not change much (compare Figs. 3.4(c) and 3.6(a)). As for the optimized CHARMM version 22, on the other hand, the torsion-energy function for $\psi$ is improved in the sense that the value at $\psi = 0$ is no longer a distinct global maximum (compare Figs. 3.4(d2) and 3.6(b2)). The optimized AMBER parm99 is actually also improved with respect to the potential-energy surface. The results are shown in Fig. 3.7. While the local-minimum state $C_{7eq}$ is almost absent in Fig. 3.5(c2), we now observe a clear local-minimum state there (although it is not the global minimum). The results for CHARMM version 22 here are similar to those in Fig. 3.5(d2), and they are consistent with the quantum mechanical calculations.

Moreover, we used anoter approach for the optimization of torsion-energy term. This

approach does not use the constraint that the total area enclosed by the curve of $|E_\Phi|$ (from $\Phi = -180^\circ$ to $180^\circ$) remains less than or equal to the original value during the optimization and used the values of the number of waves ($n_a$, $n_b$, and $n_c$) and the phases ($\gamma_a$, $\gamma_b$, and $\gamma_c$) as in the OPLS-AA force field. This method was applied to the AMBER parm94 parameters. We used $K_x = 400$ kcal/mol/Å$^2$. The optimized torsion-energy parameters are listed in Tables 3.66 and 3.67. In Fig. 3.6(c) we plot the torsion-energy functions. These waveforms ($\phi$(c1) and $\psi$(c2)) resemble the optimized AMBER parm96 and the optimized AMBER parm99. In Fig. 3.7(c), the potential-energy surface is shown. The local-minimum state $C_{7eq}$ is almost absent.

# Bibliography


[1] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P.A. Kollman, *J. Am. Chem. Soc.*, **117**, 5179 (1995).

[2] P.A. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, and A. Pohorille, in: A. Wilkinson, P. Weiner, and W.F. van Gunsteren, Elsevier, (Eds.), Computer Simulation of Biomolecular Systems, 3, Elsevier, 1997, p. 83.

[3] J. Wang, P. Cieplak, and P.A. Kollman, *J. Comput. Chem.*, **21**, 1049 (2000).

[4] A.D. MacKerrell, 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. Wiórkiewicz-Kuczera, and D. Yin, M. Karplus, *J. Phys. Chem. B*, **102**, 3586 (1998).

[5] W.L. Jorgensen, D.S. Maxwell, and J. Tirado-Rives, *J. Am. Chem. Soc.*, **118**, 11225 (1996).

[6] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, and W.L. Jorgensen, *J. Phys. Chem. B*, **105**, 6474 (2001).

[7] W.F. van 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. Zürich: Vdf Hochschulverlag AG an der ETH Zürich," (1996).

[8] M.J. Sippl, G. Némethy, and H.A. Scheraga, *J. Phys. Chem.*, **88**, 6231 (1984), and references therein.

[9] A. Mitsutake, Y. Sugita, and Y. Okamoto, *Biopolymers (Pept. Sci.)*, **60**, 96 (2001).

[10] T. Yoda, Y. Sugita, and Y. Okamoto, *Chem. Phys. Lett.*, **386**, 460 (2004).

[11] Y. Sakae and Y. Okamoto, *Chem. Phys. Lett.*, **382**, 626 (2003).

[12] Y. Sakae and Y. Okamoto, *J. Theor. Comput. Chem.*, in press.

[13] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne, *Nucleic Acids Research*, **28**, 235 (2000).

[14] S. Tanaka and H.A. Scheraga, *Macromol.*, **9**, 945 (1976).

[15] S. Miyazawa and R.L. Jernigan, *Macromol.*, **18**, 534 (1985).

[16] M.J. Sippl, *J. Mol. Biol.*, **213**, 859 (1990).

[17] K.T. Simons, C. Kooperberg, E. Huang, and D. Baker, *J. Mol. Biol.*, **268**, 209 (1997).

[18] R. Samudrala and J. Moult, *J. Mol. Biol.*, **275**, 895 (1998).

[19] R.A. Goldstein, Z.A. Luthey-Schulten, and P.G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 4918 (1992).

[20] M-H. Hao and H.A. Scheraga, *J. Phys. Chem.*, **100**, 14540 (1996).

[21] L.A. Mirny and E.I. Shakhnovich, *J. Mol. Biol.*, **264**, 1164 (1996).

[22] S. Takada, Z.A. Luthey-Schulten, and P.G. Wolynes, *J. Chem. Phys.* **110**, 11616 (1999).

[23] Y. Xia and M. Levitt, *J. Chem. Phys.*, **113**, 9318 (2000).

[24] V.N. Maiorov and G.M. Crippen, *J. Mol. Biol.*, **227**, 876 (1992).

[25] W.C. Still, A. Tempczyk, R.C. Hawley, and T. Hendrickson, *J. Am. Chem. Soc.*, **112**, 6127 (1990).

[26] D. Qiu, P.S. Shenkin, F.P. Hollinger, and W.C. Still, *J. Phys. Chem. A*, **101**, 3005 (1997).

[27] S. Kirkpatrick, C.D. Gelatt Jr., and M.P. Vecchi, *Science*, **220**, 671 (1983).

[28] http://dasher.wustl.edu/tinker/.

[29] http://www.fccc.edu/research/labs/dunbrack/pisces/.

Table 3.1: Partial-charge parameters of glycine. AMB, CHA, and OPLS respectively stand for the original AMBER, CHARMM version 22, and OPLS-AA force fields. Opt(94), Opt(96), Opt(99), Opt(CH), and Opt(OP) are the optimized AMBER parm94, AMBER parm96, AMBER parm99, CHARMM version 22, and OPLS-AA, respectively.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3471 | −0.3614 | −0.3506 | −0.4700 | −0.4381 | −0.5000 | −0.5153 |
| CA | −0.0252 | 0.0175 | 0.0148 | 0.0166 | −0.0200 | 0.0185 | 0.0800 | 0.0909 |
| C | 0.5973 | 0.5526 | 0.5698 | 0.5577 | 0.5100 | 0.5309 | 0.5000 | 0.6459 |
| HN | 0.2719 | 0.2492 | 0.2509 | 0.2480 | 0.3100 | 0.3004 | 0.3000 | 0.2615 |
| O | −0.5679 | −0.5980 | −0.5977 | −0.5983 | −0.5100 | −0.5491 | −0.5000 | −0.5546 |
| HA | 0.0698 | 0.0629 | 0.0618 | 0.0633 | 0.0900 | 0.0687 | 0.0600 | 0.0358 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.2: Partial-charge parameters of alanine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3354 | −0.3483 | −0.3407 | −0.4700 | −0.3909 | −0.5000 | −0.5224 |
| CA | 0.0337 | 0.0545 | 0.0547 | 0.0511 | 0.0700 | 0.0427 | 0.1400 | 0.1301 |
| C | 0.5973 | 0.5141 | 0.5240 | 0.5235 | 0.5100 | 0.5215 | 0.5000 | 0.6687 |
| HN | 0.2719 | 0.2323 | 0.2346 | 0.2317 | 0.3100 | 0.2709 | 0.3000 | 0.2610 |
| O | −0.5679 | −0.5703 | −0.5599 | −0.5778 | −0.5100 | −0.5417 | −0.5000 | −0.5567 |
| HA | 0.0823 | 0.0901 | 0.0912 | 0.0900 | 0.0900 | 0.0741 | 0.0600 | 0.0786 |
| CB | −0.1825 | −0.0453 | −0.0470 | −0.0501 | −0.2700 | −0.2718 | −0.1800 | −0.0701 |
| HB | 0.0603 | 0.0200 | 0.0169 | 0.0241 | 0.0900 | 0.0984 | 0.0600 | 0.0036 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.3: Partial-charge parameters of valine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3486 | −0.3618 | −0.3553 | −0.4700 | −0.4286 | −0.5000 | −0.4963 |
| CA | −0.0875 | −0.0385 | −0.0429 | −0.0449 | 0.0700 | 0.0821 | 0.1400 | 0.2025 |
| C | 0.5973 | 0.5710 | 0.5932 | 0.5817 | 0.5100 | 0.5304 | 0.5000 | 0.6031 |
| HN | 0.2719 | 0.2617 | 0.2653 | 0.2614 | 0.3100 | 0.2973 | 0.3000 | 0.2315 |
| O | −0.5679 | −0.5701 | −0.5716 | −0.5765 | −0.5100 | −0.5321 | −0.5000 | −0.5396 |
| HA | 0.0969 | 0.1214 | 0.1207 | 0.1204 | 0.0900 | 0.0798 | 0.0600 | 0.0659 |
| CB | 0.2985 | 0.2752 | 0.2676 | 0.2673 | −0.0900 | −0.1376 | −0.0600 | −0.0591 |
| HB | −0.0297 | −0.0492 | −0.0487 | −0.0449 | 0.0900 | 0.1004 | 0.0600 | 0.0072 |
| CG1 | −0.3192 | −0.4065 | −0.3906 | −0.3915 | −0.2700 | −0.2832 | −0.1800 | 0.1116 |
| HG1 | 0.0791 | 0.1016 | 0.0935 | 0.0972 | 0.0900 | 0.1062 | 0.0600 | −0.0336 |
| CG2 | −0.3192 | −0.1785 | −0.1885 | −0.1828 | −0.2700 | −0.2824 | −0.1800 | −0.0707 |
| HG2 | 0.0791 | 0.0191 | 0.0256 | 0.0245 | 0.0900 | 0.0851 | 0.0600 | 0.0149 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.4: Partial-charge parameters of leucine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3374 | −0.3556 | −0.3447 | −0.4700 | −0.4123 | −0.5000 | −0.5173 |
| CA | −0.0518 | −0.0058 | −0.0106 | −0.0072 | 0.0700 | 0.0702 | 0.1400 | 0.1644 |
| C | 0.5973 | 0.5433 | 0.5598 | 0.5518 | 0.5100 | 0.5195 | 0.5000 | 0.6292 |
| HN | 0.2719 | 0.2461 | 0.2513 | 0.2451 | 0.3100 | 0.2886 | 0.3000 | 0.2582 |
| O | −0.5679 | −0.5778 | −0.5734 | −0.5843 | −0.5100 | −0.5335 | −0.5000 | −0.5392 |
| HA | 0.0922 | 0.1345 | 0.1346 | 0.1330 | 0.0900 | 0.0925 | 0.0600 | 0.1178 |
| CB | −0.1102 | −0.1031 | −0.1023 | −0.1043 | −0.1800 | −0.2347 | −0.1200 | −0.1771 |
| HB | 0.0457 | 0.0634 | 0.0616 | 0.0661 | 0.0900 | 0.1344 | 0.0600 | 0.0599 |
| CG | 0.3531 | 0.2996 | 0.2998 | 0.3019 | −0.0900 | −0.0626 | −0.0600 | 0.0053 |
| HG | −0.0361 | −0.0495 | −0.0487 | −0.0469 | 0.0900 | 0.0860 | 0.0600 | 0.0562 |
| CD1 | −0.4121 | −0.2813 | −0.2793 | −0.2866 | −0.2700 | −0.2677 | −0.1800 | −0.2648 |
| HD1 | 0.1000 | 0.0514 | 0.0504 | 0.0527 | 0.0900 | 0.0728 | 0.0600 | 0.0723 |
| CD2 | −0.4121 | −0.4673 | −0.4734 | −0.4655 | −0.2700 | −0.3401 | −0.1800 | −0.3670 |
| HD2 | 0.1000 | 0.1059 | 0.1078 | 0.1058 | 0.0900 | 0.1023 | 0.0600 | 0.0992 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.5: Partial-charge parameters of isoleucine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3507 | −0.3676 | −0.3588 | −0.4700 | −0.4268 | −0.5000 | −0.5004 |
| CA | −0.0597 | −0.0027 | −0.0057 | −0.0077 | 0.0700 | 0.0930 | 0.1400 | 0.2138 |
| C | 0.5973 | 0.5630 | 0.5830 | 0.5739 | 0.5100 | 0.5286 | 0.5000 | 0.6098 |
| HN | 0.2719 | 0.2462 | 0.2522 | 0.2460 | 0.3100 | 0.2917 | 0.3000 | 0.2261 |
| O | −0.5679 | −0.5723 | −0.5714 | −0.5773 | −0.5100 | −0.5420 | −0.5000 | −0.5633 |
| HA | 0.0869 | 0.1053 | 0.1051 | 0.1041 | 0.0900 | 0.0835 | 0.0600 | 0.0604 |
| CB | 0.1303 | 0.1874 | 0.1788 | 0.1844 | −0.0900 | −0.1291 | −0.0600 | −0.1041 |
| HB | 0.0187 | −0.0010 | −0.0011 | 0.0037 | 0.0900 | 0.1124 | 0.0600 | 0.0297 |
| CG1 | −0.0430 | −0.0716 | −0.0745 | −0.0748 | −0.1800 | −0.2641 | −0.1200 | 0.0498 |
| HG1 | 0.0236 | 0.0089 | 0.0139 | 0.0143 | 0.0900 | 0.1103 | 0.0600 | 0.0006 |
| CG2 | −0.3204 | −0.4567 | −0.4441 | −0.4548 | −0.2700 | −0.2406 | −0.1800 | 0.1440 |
| HG2 | 0.0882 | 0.1239 | 0.1168 | 0.1244 | 0.0900 | 0.0958 | 0.0600 | −0.0369 |
| CD | −0.0660 | 0.1304 | 0.1300 | 0.1365 | −0.2700 | −0.1898 | −0.1800 | −0.5018 |
| HD | 0.0186 | −0.0556 | −0.0543 | −0.0590 | 0.0900 | 0.0584 | 0.0600 | 0.1485 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.6: Partial-charge parameters of serine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3477 | −0.3591 | −0.3512 | −0.4700 | −0.4032 | −0.5000 | −0.5306 |
| CA | −0.0249 | 0.0146 | 0.0123 | 0.0121 | 0.0700 | 0.0294 | 0.1400 | 0.0995 |
| C | 0.5973 | 0.5441 | 0.5588 | 0.5517 | 0.5100 | 0.5685 | 0.5000 | 0.6559 |
| HN | 0.2719 | 0.2519 | 0.2526 | 0.2510 | 0.3100 | 0.3094 | 0.3000 | 0.2743 |
| O | −0.5679 | −0.5706 | −0.5675 | −0.5746 | −0.5100 | −0.5336 | −0.5000 | −0.5151 |
| HA | 0.0843 | 0.0992 | 0.0996 | 0.0988 | 0.0900 | 0.0813 | 0.0600 | 0.0780 |
| CB | 0.2117 | 0.2312 | 0.2252 | 0.2271 | 0.0500 | −0.0244 | 0.1450 | 0.1844 |
| HB | 0.0352 | 0.0221 | 0.0222 | 0.0243 | 0.0900 | 0.0980 | 0.0600 | 0.0274 |
| OG | −0.6546 | −0.6921 | −0.6908 | −0.6889 | −0.6600 | −0.6699 | −0.6830 | −0.7278 |
| HG | 0.4275 | 0.4252 | 0.4245 | 0.4254 | 0.4300 | 0.4465 | 0.4180 | 0.4266 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.7: Partial-charge parameters of threonine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3544 | −0.3653 | −0.3582 | −0.4700 | −0.4254 | −0.5000 | −0.5204 |
| CA | −0.0389 | 0.0115 | 0.0078 | 0.0069 | 0.0700 | 0.0768 | 0.1400 | 0.1277 |
| C | 0.5973 | 0.5408 | 0.5590 | 0.5491 | 0.5100 | 0.5336 | 0.5000 | 0.6162 |
| HN | 0.2719 | 0.2679 | 0.2691 | 0.2674 | 0.3100 | 0.3209 | 0.3000 | 0.2738 |
| O | −0.5679 | −0.5643 | −0.5646 | −0.5703 | −0.5100 | −0.5249 | −0.5000 | −0.5067 |
| HA | 0.1007 | 0.1026 | 0.1040 | 0.1018 | 0.0900 | 0.0701 | 0.0600 | 0.0837 |
| CB | 0.3654 | 0.4005 | 0.3907 | 0.3952 | 0.1400 | 0.0890 | 0.2050 | 0.2326 |
| HB | 0.0043 | 0.0043 | 0.0064 | 0.0070 | 0.0900 | 0.1028 | 0.0600 | 0.0437 |
| OG1 | −0.6761 | −0.7157 | −0.7108 | −0.7116 | −0.6600 | −0.6589 | −0.6830 | −0.7124 |
| HG1 | 0.4102 | 0.4025 | 0.4036 | 0.4028 | 0.4300 | 0.4502 | 0.4180 | 0.4195 |
| CG2 | −0.2438 | −0.1359 | −0.1353 | −0.1315 | −0.2700 | −0.1800 | −0.1800 | −0.0031 |
| HG2 | 0.0642 | 0.0134 | 0.0118 | 0.0138 | 0.0900 | 0.0486 | 0.0600 | −0.0182 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.8: Partial-charge parameters of cysteine (-SH). See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3709 | −0.3860 | −0.3745 | −0.4700 | −0.4270 | −0.5000 | −0.5385 |
| CA | 0.0213 | 0.0736 | 0.0719 | 0.0697 | 0.0700 | 0.0471 | 0.1400 | 0.1665 |
| C | 0.5973 | 0.5125 | 0.5251 | 0.5202 | 0.5100 | 0.5443 | 0.5000 | 0.6296 |
| HN | 0.2719 | 0.2508 | 0.2529 | 0.2485 | 0.3100 | 0.3016 | 0.3000 | 0.2591 |
| O | −0.5679 | −0.5443 | −0.5384 | −0.5488 | −0.5100 | −0.5130 | −0.5000 | −0.5133 |
| HA | 0.1124 | 0.1181 | 0.1187 | 0.1174 | 0.0900 | 0.0844 | 0.0600 | 0.0895 |
| CB | −0.1231 | −0.0302 | −0.0306 | −0.0359 | −0.1100 | −0.1527 | 0.0600 | 0.0886 |
| HB | 0.1112 | 0.0823 | 0.0813 | 0.0878 | 0.0900 | 0.1158 | 0.0600 | 0.0177 |
| SG | −0.3119 | −0.3141 | −0.3193 | −0.3124 | −0.2300 | −0.2810 | −0.4350 | −0.4017 |
| HG | 0.1933 | 0.1399 | 0.1431 | 0.1402 | 0.1600 | 0.1647 | 0.2550 | 0.1848 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.9: Partial-charge parameters of cysteine (-SS-). See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3450 | −0.3607 | −0.3513 | −0.4700 | −0.4311 | −0.5000 | −0.5186 |
| CA | 0.0429 | 0.0465 | 0.0409 | 0.0431 | 0.0700 | 0.0267 | 0.1400 | 0.1249 |
| C | 0.5973 | 0.5424 | 0.5600 | 0.5505 | 0.5100 | 0.5389 | 0.5000 | 0.6258 |
| HN | 0.2719 | 0.2546 | 0.2587 | 0.2546 | 0.3100 | 0.3118 | 0.3000 | 0.2555 |
| O | −0.5679 | −0.5666 | −0.5679 | −0.5726 | −0.5100 | −0.5279 | −0.5000 | −0.5155 |
| HA | 0.0766 | 0.0829 | 0.0844 | 0.0827 | 0.0900 | 0.0758 | 0.0600 | 0.0702 |
| CB | −0.0790 | −0.0044 | −0.0089 | −0.0094 | −0.1000 | −0.1751 | 0.0975 | 0.1189 |
| HB | 0.0910 | 0.0711 | 0.0723 | 0.0745 | 0.0900 | 0.1208 | 0.0600 | 0.0297 |
| SG | −0.1081 | −0.1526 | −0.1511 | −0.1466 | −0.0800 | −0.0607 | −0.2175 | −0.2206 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.10: Partial-charge parameters of proline. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.2548 | −0.2225 | −0.2387 | −0.2267 | −0.2900 | −0.2580 | −0.1400 | −0.2326 |
| CA | −0.0266 | 0.0008 | −0.0047 | 0.0024 | 0.0200 | −0.0194 | 0.0100 | 0.0287 |
| C | 0.5896 | 0.5767 | 0.5943 | 0.5784 | 0.5100 | 0.5854 | 0.5000 | 0.7098 |
| O | −0.5748 | −0.5701 | −0.5701 | −0.5700 | −0.5100 | −0.5623 | −0.5000 | −0.5597 |
| HA | 0.0641 | 0.0697 | 0.0700 | 0.0713 | 0.0900 | 0.0827 | 0.0600 | 0.0285 |
| CB | −0.0070 | −0.0167 | −0.0107 | −0.0173 | −0.1800 | −0.2029 | −0.1200 | −0.1056 |
| HB | 0.0253 | 0.0438 | 0.0431 | 0.0436 | 0.0900 | 0.1059 | 0.0600 | 0.0212 |
| CG | 0.0189 | −0.2199 | −0.2238 | −0.2231 | −0.1800 | −0.2775 | −0.1200 | 0.0695 |
| HG | 0.0213 | 0.0985 | 0.1000 | 0.1002 | 0.0900 | 0.1301 | 0.0600 | −0.0255 |
| CD | 0.0192 | 0.0482 | 0.0587 | 0.0436 | 0.0000 | −0.0548 | −0.0500 | −0.0524 |
| HD | 0.0391 | 0.0246 | 0.0194 | 0.0269 | 0.0900 | 0.1174 | 0.0600 | 0.0612 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.11: Partial-charge parameters of phenylalanine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3516 | −0.3667 | −0.3545 | −0.4700 | −0.4239 | −0.5000 | −0.5047 |
| CA | −0.0024 | 0.0417 | 0.0369 | 0.0405 | 0.0700 | 0.0526 | 0.1400 | 0.1678 |
| C | 0.5973 | 0.5259 | 0.5423 | 0.5343 | 0.5100 | 0.5304 | 0.5000 | 0.6237 |
| HN | 0.2719 | 0.2477 | 0.2511 | 0.2461 | 0.3100 | 0.2985 | 0.3000 | 0.2450 |
| O | −0.5679 | −0.5728 | −0.5666 | −0.5821 | −0.5100 | −0.5340 | −0.5000 | −0.5459 |
| HA | 0.0978 | 0.1159 | 0.1167 | 0.1130 | 0.0900 | 0.0788 | 0.0600 | 0.0813 |
| CB | −0.0343 | 0.0070 | 0.0043 | 0.0018 | −0.1800 | −0.1999 | −0.0050 | −0.0740 |
| HB | 0.0295 | 0.0270 | 0.0269 | 0.0273 | 0.0900 | 0.0982 | 0.0600 | 0.0270 |
| CG | 0.0118 | 0.1547 | 0.1512 | 0.1658 | 0.0000 | 0.1045 | −0.1150 | 0.2544 |
| CD | −0.1256 | −0.3557 | −0.3540 | −0.3529 | −0.1150 | −0.2126 | −0.1150 | −0.4278 |
| HD | 0.1330 | 0.1331 | 0.1330 | 0.1437 | 0.1150 | 0.0954 | 0.1150 | 0.1586 |
| CE | −0.1704 | 0.1587 | 0.1581 | 0.1348 | −0.1150 | 0.1134 | −0.1150 | 0.2149 |
| HE | 0.1430 | 0.0531 | 0.0518 | 0.0588 | 0.1150 | 0.0473 | 0.1150 | 0.0132 |
| CZ | −0.1072 | −0.2702 | −0.2719 | −0.2689 | −0.1150 | −0.3073 | −0.1150 | −0.3246 |
| HZ | 0.1297 | 0.0693 | 0.0711 | 0.0806 | 0.1150 | 0.1169 | 0.1150 | 0.1052 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.12: Partial-charge parameters of tyrosine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3515 | −0.3671 | −0.3608 | −0.4700 | −0.4350 | −0.5000 | −0.5111 |
| CA | −0.0014 | 0.0398 | 0.0350 | 0.0392 | 0.0700 | 0.0582 | 0.1400 | 0.1794 |
| C | 0.5973 | 0.5543 | 0.5708 | 0.5589 | 0.5100 | 0.5454 | 0.5000 | 0.6187 |
| HN | 0.2719 | 0.2446 | 0.2494 | 0.2439 | 0.3100 | 0.2932 | 0.3000 | 0.2357 |
| O | −0.5679 | −0.5803 | −0.5775 | −0.5870 | −0.5100 | −0.5412 | −0.5000 | −0.5413 |
| HA | 0.0876 | 0.0945 | 0.0962 | 0.1005 | 0.0900 | 0.0696 | 0.0600 | 0.0638 |
| CB | −0.0152 | 0.0666 | 0.0614 | 0.0396 | −0.1800 | −0.2011 | −0.0050 | −0.0243 |
| HB | 0.0295 | −0.0077 | −0.0067 | 0.0161 | 0.0900 | 0.0994 | 0.0600 | 0.0083 |
| CG | −0.0011 | −0.2058 | −0.2077 | −0.1631 | 0.0000 | −0.0780 | −0.1150 | 0.1976 |
| CD | −0.1906 | 0.0971 | 0.0971 | −0.0424 | −0.1150 | 0.0467 | −0.1150 | −0.3601 |
| HD | 0.1699 | 0.0655 | 0.0663 | 0.1277 | 0.1150 | 0.0490 | 0.1150 | 0.1947 |
| CE | −0.2341 | −0.3159 | −0.3146 | −0.2541 | −0.1150 | −0.1892 | −0.1150 | 0.0288 |
| HE | 0.1656 | 0.1383 | 0.1358 | 0.1541 | 0.1150 | 0.1531 | 0.1150 | 0.0940 |
| CZ | 0.3226 | 0.2900 | 0.2934 | 0.2488 | 0.1100 | 0.0656 | 0.1500 | −0.0805 |
| OH | −0.5579 | −0.5021 | −0.5055 | −0.5147 | −0.5400 | −0.5343 | −0.5850 | −0.5070 |
| HZ | 0.3992 | 0.3953 | 0.3958 | 0.3919 | 0.4300 | 0.4396 | 0.4350 | 0.4376 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.13: Partial-charge parameters of tryptophan. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3463 | −0.3612 | −0.3511 | −0.4700 | −0.4053 | −0.5000 | −0.5037 |
| CA | −0.0275 | 0.0226 | 0.0189 | 0.0172 | 0.0700 | 0.0603 | 0.1400 | 0.1430 |
| C | 0.5973 | 0.5247 | 0.5430 | 0.5334 | 0.5100 | 0.5323 | 0.5000 | 0.6137 |
| HN | 0.2719 | 0.2348 | 0.2377 | 0.2344 | 0.3100 | 0.2945 | 0.3000 | 0.2391 |
| O | −0.5679 | −0.5829 | −0.5811 | −0.5884 | −0.5100 | −0.5373 | −0.5000 | −0.5487 |
| HA | 0.1123 | 0.1014 | 0.1030 | 0.1013 | 0.0900 | 0.0824 | 0.0600 | 0.0520 |
| CB | −0.0050 | 0.1255 | 0.1178 | 0.1234 | −0.1800 | −0.2237 | −0.1200 | 0.0693 |
| HB | 0.0339 | −0.0046 | −0.0039 | −0.0005 | 0.0900 | 0.1138 | 0.0600 | −0.0265 |
| CG | −0.1415 | −0.2294 | −0.2236 | −0.2337 | −0.0300 | −0.0065 | 0.0750 | 0.0489 |
| CD1 | −0.1638 | 0.0281 | 0.0293 | 0.0243 | 0.0350 | −0.0105 | −0.1150 | −0.0986 |
| HD1 | 0.2062 | 0.1933 | 0.1912 | 0.1972 | 0.1150 | 0.1281 | 0.1150 | 0.1361 |
| CD2 | 0.1243 | 0.3478 | 0.3435 | 0.3616 | −0.0200 | −0.0180 | −0.0550 | 0.2884 |
| NE1 | −0.3418 | −0.3853 | −0.3883 | −0.3824 | −0.6100 | −0.6341 | −0.5700 | −0.5284 |
| HE1 | 0.3412 | 0.3340 | 0.3346 | 0.3317 | 0.3800 | 0.3842 | 0.4200 | 0.3865 |
| CE2 | 0.1380 | −0.0446 | −0.0385 | −0.0628 | 0.1300 | 0.1514 | 0.1300 | −0.1280 |
| CE3 | −0.2387 | −0.3892 | −0.3804 | −0.4000 | −0.1150 | −0.0061 | −0.1150 | −0.1744 |
| HE3 | 0.1700 | 0.1468 | 0.1455 | 0.1429 | 0.1150 | 0.0558 | 0.1150 | −0.0137 |
| CZ2 | −0.2601 | −0.3000 | −0.3044 | −0.2895 | −0.1150 | −0.0988 | −0.1150 | −0.1308 |
| HZ2 | 0.1572 | 0.1937 | 0.1970 | 0.1916 | 0.1150 | 0.0711 | 0.1150 | 0.1451 |
| CZ3 | −0.1972 | −0.0704 | −0.0801 | −0.0391 | −0.1150 | −0.3270 | −0.1150 | 0.1918 |
| HZ3 | 0.1447 | 0.1778 | 0.1805 | 0.1702 | 0.1150 | 0.1411 | 0.1150 | 0.0609 |
| CH2 | −0.1134 | −0.2815 | −0.2806 | −0.2857 | −0.1150 | 0.1219 | −0.1150 | −0.4572 |
| HH2 | 0.1417 | 0.2083 | 0.2040 | 0.2045 | 0.1150 | 0.0166 | 0.1150 | 0.2617 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.14: Partial-charge parameters of histidine (+). See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3479 | −0.2915 | −0.3099 | −0.2982 | −0.4700 | −0.4169 | −0.5000 | −0.5224 |
| CA | −0.1354 | −0.0650 | −0.0677 | −0.0664 | 0.0700 | 0.0615 | 0.1400 | 0.1764 |
| C | 0.7341 | 0.6530 | 0.6680 | 0.6614 | 0.5100 | 0.5396 | 0.5000 | 0.6420 |
| HN | 0.2747 | 0.2483 | 0.2538 | 0.2478 | 0.3100 | 0.2938 | 0.3000 | 0.2539 |
| O | −0.5894 | −0.5952 | −0.5908 | −0.5988 | −0.5100 | −0.5447 | −0.5000 | −0.5533 |
| HA | 0.1212 | 0.1338 | 0.1350 | 0.1359 | 0.0900 | 0.0663 | 0.0600 | 0.0629 |
| CB | −0.0414 | −0.0129 | −0.0131 | −0.0222 | −0.0500 | −0.0788 | −0.0050 | 0.0296 |
| HB | 0.0810 | 0.0801 | 0.0795 | 0.0861 | 0.0900 | 0.1024 | 0.0600 | 0.0339 |
| CG | −0.0012 | 0.0650 | 0.0595 | 0.0469 | 0.1900 | 0.2378 | 0.2150 | 0.3084 |
| ND1 | −0.1513 | −0.3095 | −0.3096 | −0.2882 | −0.5100 | −0.5012 | −0.5400 | −0.6897 |
| HD1 | 0.3866 | 0.4301 | 0.4304 | 0.4243 | 0.4400 | 0.4296 | 0.4600 | 0.4770 |
| CD2 | −0.1141 | −0.1171 | −0.1173 | −0.1057 | 0.1900 | 0.1636 | 0.2150 | 0.1161 |
| HD2 | 0.2317 | 0.2864 | 0.2862 | 0.2800 | 0.1300 | 0.1281 | 0.1150 | 0.1977 |
| CE1 | −0.0170 | −0.0092 | −0.0087 | −0.0204 | 0.3200 | 0.2829 | 0.3850 | 0.4910 |
| HE1 | 0.2681 | 0.2583 | 0.2589 | 0.2632 | 0.1800 | 0.1815 | 0.1150 | 0.1204 |
| NE2 | −0.1718 | −0.2472 | −0.2471 | −0.2438 | −0.5100 | −0.4948 | −0.5400 | −0.6662 |
| HE2 | 0.3911 | 0.4125 | 0.4134 | 0.4120 | 0.4400 | 0.4469 | 0.4600 | 0.4884 |
| Total | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

Table 3.15: Partial-charge parameters of aspartic acid. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.5163 | −0.4199 | −0.4377 | −0.4236 | −0.4700 | −0.4185 | −0.5000 | −0.4537 |
| CA | 0.0381 | 0.0743 | 0.0703 | 0.0738 | 0.0700 | 0.0775 | 0.1400 | 0.1396 |
| C | 0.5366 | 0.4768 | 0.4944 | 0.4842 | 0.5100 | 0.5129 | 0.5000 | 0.6620 |
| HN | 0.2936 | 0.2581 | 0.2620 | 0.2575 | 0.3100 | 0.2959 | 0.3000 | 0.2313 |
| O | −0.5819 | −0.6101 | −0.6060 | −0.6142 | −0.5100 | −0.5317 | −0.5000 | −0.5833 |
| HA | 0.0880 | 0.1013 | 0.0994 | 0.1011 | 0.0900 | 0.0710 | 0.0600 | 0.0666 |
| CB | −0.0303 | −0.0248 | −0.0267 | −0.0253 | −0.2800 | −0.2708 | −0.2200 | −0.2229 |
| HB | −0.0122 | −0.0096 | −0.0094 | −0.0087 | 0.0900 | 0.1035 | 0.0600 | 0.0041 |
| CG | 0.7994 | 0.7913 | 0.7944 | 0.7908 | 0.6200 | 0.5848 | 0.7000 | 0.8193 |
| OD | −0.8014 | −0.8139 | −0.8156 | −0.8135 | −0.7600 | −0.7640 | −0.8000 | −0.8336 |
| Total | -1.0000 | -1.0000 | -1.0000 | -1.0000 | -1.0000 | -1.0000 | -1.0000 | -1.0000 |

Table 3.16: Partial-charge parameters of asparagine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3251 | −0.3446 | −0.3322 | −0.4700 | −0.4107 | −0.5000 | −0.4827 |
| CA | 0.0143 | 0.0392 | 0.0358 | 0.0399 | 0.0700 | 0.0679 | 0.1400 | 0.1423 |
| C | 0.5973 | 0.5353 | 0.5522 | 0.5427 | 0.5100 | 0.5317 | 0.5000 | 0.6777 |
| HN | 0.2719 | 0.2384 | 0.2431 | 0.2378 | 0.3100 | 0.2951 | 0.3000 | 0.2460 |
| O | −0.5679 | −0.5827 | −0.5787 | −0.5886 | −0.5100 | −0.5384 | −0.5000 | −0.5811 |
| HA | 0.1048 | 0.1391 | 0.1389 | 0.1380 | 0.0900 | 0.0851 | 0.0600 | 0.0937 |
| CB | −0.2041 | −0.1941 | −0.1933 | −0.1927 | −0.1800 | −0.2159 | −0.1200 | −0.1446 |
| HB | 0.0797 | 0.0864 | 0.0862 | 0.0874 | 0.0900 | 0.1116 | 0.0600 | 0.0498 |
| CG | 0.7130 | 0.6446 | 0.6431 | 0.6451 | 0.5500 | 0.5535 | 0.5000 | 0.4988 |
| OD1 | −0.5931 | −0.5843 | −0.5852 | −0.5838 | −0.5500 | −0.5585 | −0.5000 | −0.4977 |
| ND2 | −0.9191 | −0.7908 | −0.7911 | −0.7906 | −0.6200 | −0.6644 | −0.7600 | −0.7458 |
| HD2 | 0.4196 | 0.3538 | 0.3537 | 0.3548 | 0.3100 | 0.3157 | 0.3800 | 0.3469 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.17: Partial-charge parameters of glutamic acid. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.5163 | −0.4248 | −0.4376 | −0.4302 | −0.4700 | −0.3961 | −0.5000 | −0.5401 |
| CA | 0.0397 | 0.0583 | 0.0553 | 0.0554 | 0.0700 | 0.0423 | 0.1400 | 0.1320 |
| C | 0.5366 | 0.4728 | 0.4873 | 0.4817 | 0.5100 | 0.5249 | 0.5000 | 0.6538 |
| HN | 0.2936 | 0.2595 | 0.2620 | 0.2590 | 0.3100 | 0.2845 | 0.3000 | 0.2626 |
| O | −0.5819 | −0.6181 | −0.6107 | −0.6248 | −0.5100 | −0.5603 | −0.5000 | −0.5777 |
| HA | 0.1105 | 0.1232 | 0.1232 | 0.1221 | 0.0900 | 0.0837 | 0.0600 | 0.0670 |
| CB | 0.0560 | 0.1226 | 0.1170 | 0.1217 | −0.1800 | −0.1634 | −0.1200 | −0.0517 |
| HB | −0.0173 | −0.0333 | −0.0334 | −0.0300 | 0.0900 | 0.0943 | 0.0600 | 0.0418 |
| CG | 0.0136 | −0.0678 | −0.0716 | −0.0659 | −0.2800 | −0.2870 | −0.2200 | −0.2185 |
| HG | −0.0425 | −0.0300 | −0.0297 | −0.0299 | 0.0900 | 0.1160 | 0.0600 | 0.0437 |
| CD | 0.8054 | 0.8293 | 0.8340 | 0.8292 | 0.6200 | 0.5465 | 0.7000 | 0.7320 |
| OE | −0.8188 | −0.8142 | −0.8163 | −0.8142 | −0.7600 | −0.7479 | −0.8000 | −0.8152 |
| Total | −1.0000 | −1.0000 | −1.0000 | −1.0000 | −1.0000 | −1.0000 | −1.0000 | −1.0000 |

Table 3.18: Partial-charge parameters of glutamine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3493 | −0.3657 | −0.3562 | −0.4700 | −0.4116 | −0.5000 | −0.5364 |
| CA | −0.0031 | 0.0328 | 0.0291 | 0.0285 | 0.0700 | 0.0403 | 0.1400 | 0.1459 |
| C | 0.5973 | 0.5314 | 0.5481 | 0.5402 | 0.5100 | 0.5451 | 0.5000 | 0.6336 |
| HN | 0.2719 | 0.2427 | 0.2466 | 0.2427 | 0.3100 | 0.2909 | 0.3000 | 0.2590 |
| O | −0.5679 | −0.5923 | −0.5870 | −0.5976 | −0.5100 | −0.5527 | −0.5000 | −0.5516 |
| HA | 0.0850 | 0.0961 | 0.0970 | 0.0946 | 0.0900 | 0.0810 | 0.0600 | 0.0681 |
| CB | −0.0036 | 0.0713 | 0.0661 | 0.0705 | −0.1800 | −0.1757 | −0.1200 | −0.0387 |
| HB | 0.0171 | 0.0080 | 0.0077 | 0.0093 | 0.0900 | 0.0994 | 0.0600 | 0.0340 |
| CG | −0.0645 | −0.1437 | −0.1445 | −0.1399 | −0.1800 | −0.2157 | −0.1200 | −0.1304 |
| HG | 0.0352 | 0.0518 | 0.0521 | 0.0524 | 0.0900 | 0.1176 | 0.0600 | 0.0590 |
| CD | 0.6951 | 0.6468 | 0.6487 | 0.6455 | 0.5500 | 0.5447 | 0.5000 | 0.4924 |
| OE1 | −0.6086 | −0.5871 | −0.5862 | −0.5849 | −0.5500 | −0.5621 | −0.5000 | −0.4788 |
| NE2 | −0.9407 | −0.7763 | −0.7796 | −0.7786 | −0.6200 | −0.6606 | −0.7600 | −0.7769 |
| HE2 | 0.4251 | 0.3540 | 0.3539 | 0.3559 | 0.3100 | 0.3212 | 0.3800 | 0.3639 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.19: Partial-charge parameters of methionine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.4157 | −0.3388 | −0.3517 | −0.3446 | −0.4700 | −0.3819 | −0.5000 | −0.5114 |
| CA | −0.0237 | 0.0207 | 0.0170 | 0.0172 | 0.0700 | 0.0687 | 0.1400 | 0.1557 |
| C | 0.5973 | 0.5499 | 0.5666 | 0.5604 | 0.5100 | 0.5316 | 0.5000 | 0.6635 |
| HN | 0.2719 | 0.2372 | 0.2409 | 0.2363 | 0.3100 | 0.2727 | 0.3000 | 0.2525 |
| O | −0.5679 | −0.5979 | −0.5921 | −0.6050 | −0.5100 | −0.5509 | −0.5000 | −0.5686 |
| HA | 0.0880 | 0.1145 | 0.1157 | 0.1136 | 0.0900 | 0.0861 | 0.0600 | 0.0921 |
| CB | 0.0342 | 0.0559 | 0.0527 | 0.0481 | −0.1800 | −0.2283 | −0.1200 | −0.1792 |
| HB | 0.0241 | 0.0312 | 0.0290 | 0.0351 | 0.0900 | 0.1217 | 0.0600 | 0.0541 |
| CG | 0.0018 | 0.0514 | 0.0488 | 0.0544 | −0.1400 | 0.0308 | 0.0975 | 0.2269 |
| HG | 0.0440 | 0.0244 | 0.0263 | 0.0266 | 0.0900 | 0.0531 | 0.0600 | 0.0101 |
| SD | −0.2737 | −0.2325 | −0.2380 | −0.2367 | −0.0900 | −0.2267 | −0.4350 | −0.3631 |
| CE | −0.0536 | −0.3055 | −0.3023 | −0.3004 | −0.2200 | −0.2193 | 0.0375 | −0.2421 |
| HE | 0.0684 | 0.1113 | 0.1106 | 0.1111 | 0.0900 | 0.0892 | 0.0600 | 0.1151 |
| Total | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

Table 3.20: Partial-charge parameters of lysine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3479 | −0.3044 | −0.3190 | −0.3110 | −0.4700 | −0.4059 | −0.5000 | −0.5274 |
| CA | −0.2400 | −0.1606 | −0.1610 | −0.1619 | 0.0700 | 0.0508 | 0.1400 | 0.1577 |
| C | 0.7341 | 0.6626 | 0.6776 | 0.6707 | 0.5100 | 0.5382 | 0.5000 | 0.6432 |
| HN | 0.2747 | 0.2498 | 0.2527 | 0.2495 | 0.3100 | 0.2854 | 0.3000 | 0.2587 |
| O | −0.5894 | −0.6064 | −0.6038 | −0.6111 | −0.5100 | −0.5397 | −0.5000 | −0.5419 |
| HA | 0.1426 | 0.1561 | 0.1564 | 0.1552 | 0.0900 | 0.0860 | 0.0600 | 0.0929 |
| CB | −0.0094 | 0.0348 | 0.0339 | 0.0306 | −0.1800 | −0.2348 | −0.1200 | −0.0988 |
| HB | 0.0362 | 0.0350 | 0.0345 | 0.0381 | 0.0900 | 0.1069 | 0.0600 | 0.0439 |
| CG | 0.0187 | 0.0325 | 0.0333 | 0.0255 | −0.1800 | −0.1185 | −0.1200 | −0.0713 |
| HG | 0.0103 | 0.0049 | 0.0030 | 0.0090 | 0.0900 | 0.0645 | 0.0600 | 0.0347 |
| CD | −0.0479 | −0.0496 | −0.0508 | −0.0494 | −0.1800 | −0.1354 | −0.1200 | −0.1761 |
| HD | 0.0621 | 0.0494 | 0.0501 | 0.0511 | 0.0900 | 0.0941 | 0.0600 | 0.0568 |
| CE | −0.0143 | −0.0746 | −0.0761 | −0.0742 | 0.2100 | 0.1543 | 0.1900 | 0.0432 |
| HE | 0.1135 | 0.1179 | 0.1180 | 0.1173 | 0.0500 | 0.0474 | 0.0600 | 0.1369 |
| NZ | −0.3854 | −0.3664 | −0.3654 | −0.3659 | −0.3000 | −0.2656 | −0.3000 | −0.3322 |
| HZ | 0.3400 | 0.3372 | 0.3370 | 0.3370 | 0.3300 | 0.3198 | 0.3300 | 0.3358 |
| Total | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

Table 3.21: Partial-charge parameters of arginine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3479 | −0.2942 | −0.3089 | −0.3013 | −0.4700 | −0.3950 | −0.5000 | −0.4992 |
| CA | −0.2637 | −0.1818 | −0.1818 | −0.1833 | 0.0700 | 0.0402 | 0.1400 | 0.1697 |
| C | 0.7341 | 0.6615 | 0.6757 | 0.6687 | 0.5100 | 0.5472 | 0.5000 | 0.6477 |
| HN | 0.2747 | 0.2447 | 0.2473 | 0.2442 | 0.3100 | 0.2742 | 0.3000 | 0.2411 |
| O | −0.5894 | −0.6182 | −0.6143 | −0.6218 | −0.5100 | −0.5543 | −0.5000 | −0.5488 |
| HA | 0.1560 | 0.1583 | 0.1586 | 0.1587 | 0.0900 | 0.0673 | 0.0600 | 0.0740 |
| CB | −0.0007 | 0.0594 | 0.0590 | 0.0554 | −0.1800 | −0.1602 | −0.1200 | −0.1092 |
| HB | 0.0327 | 0.0254 | 0.0241 | 0.0282 | 0.0900 | 0.0732 | 0.0600 | 0.0504 |
| CG | 0.0390 | −0.0022 | −0.0008 | −0.0047 | −0.1800 | −0.1134 | −0.0500 | 0.0429 |
| HG | 0.0285 | 0.0318 | 0.0306 | 0.0359 | 0.0900 | 0.0909 | 0.0600 | 0.0274 |
| CD | 0.0486 | 0.1216 | 0.1185 | 0.1167 | 0.2000 | 0.2972 | 0.1900 | 0.0772 |
| HD | 0.0687 | 0.0760 | 0.0766 | 0.0757 | 0.0900 | 0.0609 | 0.0600 | 0.1133 |
| NE | −0.5295 | −0.5465 | −0.5452 | −0.5510 | −0.7000 | −0.7342 | −0.7000 | −0.6452 |
| HE | 0.3456 | 0.3190 | 0.3189 | 0.3214 | 0.4400 | 0.4436 | 0.4400 | 0.4178 |
| CZ | 0.8076 | 0.8145 | 0.8126 | 0.8151 | 0.6400 | 0.6671 | 0.6400 | 0.5903 |
| NH | −0.8627 | −0.7923 | −0.7925 | −0.7904 | −0.8000 | −0.7966 | −0.8000 | −0.8199 |
| HH | 0.4478 | 0.3955 | 0.3957 | 0.3958 | 0.4600 | 0.4409 | 0.4600 | 0.4498 |
| Total | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

Table 3.22: Partial-charge parameters for N-termini of glycine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.2943 | 0.3629 | 0.3486 | 0.3594 | −0.3000 | −0.2681 | −0.3000 | −0.3153 |
| CA | −0.0100 | 0.0327 | 0.0300 | 0.0318 | 0.1300 | 0.1685 | 0.1900 | 0.2009 |
| C | 0.6163 | 0.5716 | 0.5888 | 0.5767 | 0.5100 | 0.5309 | 0.5000 | 0.6459 |
| HN | 0.1642 | 0.1566 | 0.1572 | 0.1562 | 0.3300 | 0.3267 | 0.3300 | 0.3172 |
| O | −0.5722 | −0.6024 | −0.6019 | −0.6026 | −0.5100 | −0.5490 | −0.5000 | −0.5546 |
| HA | 0.0895 | 0.0826 | 0.0815 | 0.0830 | 0.0900 | 0.0687 | 0.0600 | 0.0358 |

Table 3.23: Partial-charge parameters for N-termini of alanine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1414 | 0.2217 | 0.2088 | 0.2164 | −0.3000 | −0.2209 | −0.3000 | −0.3224 |
| CA | 0.1281 | 0.1489 | 0.1491 | 0.1455 | 0.2100 | 0.1827 | 0.2500 | 0.2401 |
| C | 0.6163 | 0.5332 | 0.5430 | 0.5426 | 0.5100 | 0.5215 | 0.5000 | 0.6687 |
| HN | 0.1997 | 0.1865 | 0.1873 | 0.1863 | 0.3300 | 0.3169 | 0.3300 | 0.3170 |
| O | −0.5722 | −0.5745 | −0.5641 | −0.5820 | −0.5100 | −0.5417 | −0.5000 | −0.5567 |
| HA | 0.0889 | 0.0967 | 0.0978 | 0.0966 | 0.1000 | 0.0841 | 0.0600 | 0.0786 |

Table 3.24: Partial-charge parameters for N-termini of valine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.0577 | 0.1248 | 0.1116 | 0.1181 | −0.3000 | −0.2587 | −0.3000 | −0.2963 |
| CA | 0.0023 | 0.0513 | 0.0469 | 0.0449 | 0.2100 | 0.2221 | 0.2500 | 0.3125 |
| C | 0.6163 | 0.5899 | 0.6122 | 0.6008 | 0.5100 | 0.5303 | 0.5000 | 0.6030 |
| HN | 0.2272 | 0.2238 | 0.2250 | 0.2237 | 0.3300 | 0.3258 | 0.3300 | 0.3072 |
| O | −0.5722 | −0.5744 | −0.5759 | −0.5807 | −0.5100 | −0.5322 | −0.5000 | −0.5396 |
| HA | 0.1093 | 0.1338 | 0.1331 | 0.1328 | 0.1000 | 0.0898 | 0.0600 | 0.0659 |

Table 3.25: Partial-charge parameters for N-termini of leucine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1010 | 0.1793 | 0.1611 | 0.1720 | −0.3000 | −0.2423 | −0.3000 | −0.3173 |
| CA | 0.0343 | 0.0803 | 0.0755 | 0.0789 | 0.2100 | 0.2102 | 0.2500 | 0.2744 |
| C | 0.6123 | 0.5583 | 0.5749 | 0.5667 | 0.5100 | 0.5195 | 0.5000 | 0.6293 |
| HN | 0.2148 | 0.2062 | 0.2079 | 0.2059 | 0.3300 | 0.3228 | 0.3300 | 0.3161 |
| O | −0.5713 | −0.5812 | −0.5768 | −0.5878 | −0.5100 | −0.5336 | −0.5000 | −0.5392 |
| HA | 0.1053 | 0.1476 | 0.1477 | 0.1461 | 0.1000 | 0.1025 | 0.0600 | 0.1178 |

Table 3.26: Partial-charge parameters for N-termini of isoleucine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.0311 | 0.0961 | 0.0792 | 0.0880 | −0.3000 | −0.2568 | −0.3000 | −0.3004 |
| CA | 0.0389 | 0.0959 | 0.0929 | 0.0909 | 0.2100 | 0.2330 | 0.2500 | 0.3238 |
| C | 0.6123 | 0.5780 | 0.5979 | 0.5888 | 0.5100 | 0.5286 | 0.5000 | 0.6098 |
| HN | 0.2329 | 0.2243 | 0.2263 | 0.2243 | 0.3300 | 0.3238 | 0.3300 | 0.3054 |
| O | −0.5713 | −0.5757 | −0.5749 | −0.5808 | −0.5100 | −0.5419 | −0.5000 | −0.5633 |
| HA | 0.1031 | 0.1215 | 0.1213 | 0.1203 | 0.1000 | 0.0935 | 0.0600 | 0.0604 |

Table 3.27: Partial-charge parameters for N-termini of serine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1849 | 0.2529 | 0.2415 | 0.2494 | −0.3000 | −0.2332 | −0.3000 | −0.3306 |
| CA | 0.0684 | 0.1079 | 0.1056 | 0.1054 | 0.2100 | 0.1694 | 0.2500 | 0.2095 |
| C | 0.6163 | 0.5631 | 0.5778 | 0.5707 | 0.5100 | 0.5686 | 0.5000 | 0.6559 |
| HN | 0.1898 | 0.1831 | 0.1834 | 0.1828 | 0.3300 | 0.3298 | 0.3300 | 0.3214 |
| O | −0.5722 | −0.5749 | −0.5717 | −0.5789 | −0.5100 | −0.5336 | −0.5000 | −0.5151 |
| HA | 0.0782 | 0.0931 | 0.0935 | 0.0927 | 0.1000 | 0.0913 | 0.0600 | 0.0780 |

Table 3.28: Partial-charge parameters for N-termini of threonine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1812 | 0.2425 | 0.2316 | 0.2387 | −0.3000 | −0.2554 | −0.3000 | −0.3204 |
| CA | 0.0332 | 0.0836 | 0.0799 | 0.0790 | 0.2100 | 0.2168 | 0.2500 | 0.2377 |
| C | 0.6163 | 0.5598 | 0.5780 | 0.5681 | 0.5100 | 0.5335 | 0.5000 | 0.6162 |
| HN | 0.1934 | 0.1921 | 0.1925 | 0.1919 | 0.3300 | 0.3336 | 0.3300 | 0.3213 |
| O | −0.5722 | −0.5686 | −0.5689 | −0.5746 | −0.5100 | −0.5250 | −0.5000 | −0.5067 |
| HA | 0.1087 | 0.1106 | 0.1120 | 0.1098 | 0.1000 | 0.0801 | 0.0600 | 0.0837 |

Table 3.29: Partial-charge parameters for N-termini of cysteine (-SH). See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1325 | 0.1773 | 0.1622 | 0.1737 | −0.3000 | −0.2570 | −0.3000 | −0.3385 |
| CA | 0.0978 | 0.1501 | 0.1484 | 0.1462 | 0.2100 | 0.1871 | 0.2500 | 0.2765 |
| C | 0.6123 | 0.5275 | 0.5401 | 0.5352 | 0.5100 | 0.5444 | 0.5000 | 0.6296 |
| HN | 0.2023 | 0.1953 | 0.1960 | 0.1945 | 0.3300 | 0.3272 | 0.3300 | 0.3164 |
| O | −0.5713 | −0.5476 | −0.5417 | −0.5522 | −0.5100 | −0.5130 | −0.5000 | −0.5133 |
| HA | 0.1411 | 0.1468 | 0.1474 | 0.1461 | 0.1000 | 0.0944 | 0.0600 | 0.0895 |

Table 3.30: Partial-charge parameters for N-termini of cysteine (-SS-). See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.2096 | 0.2803 | 0.2646 | 0.2740 | −0.3000 | −0.2611 | −0.3000 | −0.3186 |
| CA | 0.1138 | 0.1174 | 0.1118 | 0.1140 | 0.2100 | 0.1667 | 0.2500 | 0.2349 |
| C | 0.6163 | 0.5614 | 0.5790 | 0.5695 | 0.5100 | 0.5390 | 0.5000 | 0.6258 |
| HN | 0.1815 | 0.1757 | 0.1771 | 0.1757 | 0.3300 | 0.3306 | 0.3300 | 0.3152 |
| O | −0.5713 | −0.5700 | −0.5712 | −0.5759 | −0.5100 | −0.5279 | −0.5000 | −0.5155 |
| HA | 0.0922 | 0.0985 | 0.1000 | 0.0983 | 0.1000 | 0.0858 | 0.0600 | 0.0702 |

Table 3.31: Partial-charge parameters for N-termini of proline. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.2020 | −0.1697 | −0.1859 | −0.1739 | −0.0700 | −0.0380 | −0.2400 | −0.3326 |
| CA | 0.1589 | 0.1863 | 0.1808 | 0.1879 | 0.1600 | 0.1206 | 0.2500 | 0.2687 |
| C | 0.5260 | 0.5130 | 0.5308 | 0.5148 | 0.5100 | 0.5854 | 0.5000 | 0.7097 |
| HN | 0.3120 | 0.3120 | 0.3120 | 0.3120 | 0.2400 | 0.2400 | 0.3100 | 0.3100 |
| O | −0.5000 | −0.4944 | −0.4941 | −0.4928 | −0.5100 | −0.5173 | −0.5000 | −0.5315 |
| HA | 0.1000 | 0.1047 | 0.1047 | 0.1048 | 0.0900 | 0.0377 | 0.0600 | 0.0003 |
| CD | −0.0120 | 0.0170 | 0.0275 | 0.0124 | 0.1600 | 0.1052 | 0.1900 | 0.1876 |
| HD | 0.1000 | 0.0855 | 0.0803 | 0.0878 | 0.0900 | 0.1174 | 0.0600 | 0.0612 |

Table 3.32: Partial-charge parameters for N-termini of phenylalanine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1737 | 0.2378 | 0.2227 | 0.2349 | −0.3000 | −0.2540 | −0.3000 | −0.3047 |
| CA | 0.0859 | 0.1300 | 0.1252 | 0.1288 | 0.2100 | 0.1926 | 0.2500 | 0.2778 |
| C | 0.6123 | 0.5409 | 0.5574 | 0.5493 | 0.5100 | 0.5303 | 0.5000 | 0.6238 |
| HN | 0.1921 | 0.1840 | 0.1852 | 0.1835 | 0.3300 | 0.3262 | 0.3300 | 0.3117 |
| O | −0.5713 | −0.5762 | −0.5699 | −0.5856 | −0.5100 | −0.5341 | −0.5000 | −0.5458 |
| HA | 0.1041 | 0.1222 | 0.1230 | 0.1193 | 0.1000 | 0.0888 | 0.0600 | 0.0813 |

Table 3.33: Partial-charge parameters for N-termini of tyrosine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1940 | 0.2582 | 0.2426 | 0.2489 | −0.3000 | −0.2650 | −0.3000 | −0.3111 |
| CA | 0.0766 | 0.1178 | 0.1130 | 0.1172 | 0.2100 | 0.1982 | 0.2500 | 0.2894 |
| C | 0.6123 | 0.5693 | 0.5858 | 0.5739 | 0.5100 | 0.5455 | 0.5000 | 0.6188 |
| HN | 0.1873 | 0.1782 | 0.1798 | 0.1780 | 0.3300 | 0.3244 | 0.3300 | 0.3086 |
| O | −0.5713 | −0.5836 | −0.5808 | −0.5904 | −0.5100 | −0.5412 | −0.5000 | −0.5412 |
| HA | 0.0983 | 0.1052 | 0.1069 | 0.1112 | 0.1000 | 0.0796 | 0.0600 | 0.0638 |

Table 3.34: Partial-charge parameters for N-termini of tryptophan. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1913 | 0.2607 | 0.2458 | 0.2559 | −0.3000 | −0.2353 | −0.3000 | −0.3037 |
| CA | 0.0555 | 0.1056 | 0.1019 | 0.1002 | 0.2100 | 0.2003 | 0.2500 | 0.2530 |
| C | 0.6123 | 0.5397 | 0.5580 | 0.5484 | 0.5100 | 0.5323 | 0.5000 | 0.6138 |
| HN | 0.1888 | 0.1764 | 0.1774 | 0.1763 | 0.3300 | 0.3249 | 0.3300 | 0.3097 |
| O | −0.5713 | −0.5864 | −0.5846 | −0.5918 | −0.5100 | −0.5373 | −0.5000 | −0.5487 |
| HA | 0.1162 | 0.1053 | 0.1069 | 0.1052 | 0.1000 | 0.0924 | 0.0600 | 0.0520 |

Table 3.35: Partial-charge parameters for N-termini of histidine (+). See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.2560 | 0.3124 | 0.2940 | 0.3057 | −0.3000 | −0.2469 | −0.3000 | −0.3224 |
| CA | 0.0653 | 0.1357 | 0.1330 | 0.1343 | 0.2100 | 0.2015 | 0.2500 | 0.2864 |
| C | 0.7214 | 0.6404 | 0.6553 | 0.6487 | 0.5100 | 0.5396 | 0.5000 | 0.6420 |
| HN | 0.1704 | 0.1616 | 0.1634 | 0.1614 | 0.3300 | 0.3246 | 0.3300 | 0.3146 |
| O | −0.6013 | −0.6071 | −0.6027 | −0.6107 | −0.5100 | −0.5446 | −0.5000 | −0.5533 |
| HA | 0.1047 | 0.1173 | 0.1185 | 0.1194 | 0.1000 | 0.0763 | 0.0600 | 0.0629 |

Table 3.36: Partial-charge parameters for N-termini of aspartic acid. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.0782 | 0.1746 | 0.1568 | 0.1709 | −0.3000 | −0.2485 | −0.3000 | −0.2537 |
| CA | 0.0326 | 0.0688 | 0.0648 | 0.0683 | 0.2100 | 0.2175 | 0.2500 | 0.2496 |
| C | 0.5621 | 0.5023 | 0.5199 | 0.5097 | 0.5100 | 0.5129 | 0.5000 | 0.6620 |
| HN | 0.2200 | 0.2082 | 0.2095 | 0.2080 | 0.3300 | 0.3253 | 0.3300 | 0.3071 |
| O | −0.5889 | −0.6171 | −0.6130 | −0.6212 | −0.5100 | −0.5317 | −0.5000 | −0.5833 |
| HA | 0.1141 | 0.1274 | 0.1255 | 0.1272 | 0.1000 | 0.0810 | 0.0600 | 0.0666 |

Table 3.37: Partial-charge parameters for N-termini of asparagine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1801 | 0.2707 | 0.2512 | 0.2636 | −0.3000 | −0.2407 | −0.3000 | −0.2827 |
| CA | 0.0811 | 0.1060 | 0.1026 | 0.1067 | 0.2100 | 0.2079 | 0.2500 | 0.2523 |
| C | 0.6163 | 0.5543 | 0.5712 | 0.5617 | 0.5100 | 0.5317 | 0.5000 | 0.6776 |
| HN | 0.1921 | 0.1809 | 0.1825 | 0.1807 | 0.3300 | 0.3250 | 0.3300 | 0.3120 |
| O | −0.5722 | −0.5870 | −0.5830 | −0.5929 | −0.5100 | −0.5384 | −0.5000 | −0.5812 |
| HA | 0.1231 | 0.1574 | 0.1572 | 0.1563 | 0.1000 | 0.0951 | 0.0600 | 0.0937 |

Table 3.38: Partial-charge parameters for N-termini of glutamic acid. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.0017 | 0.0932 | 0.0804 | 0.0878 | -0.3000 | -0.2261 | -0.3000 | -0.3401 |
| CA | 0.0698 | 0.0884 | 0.0854 | 0.0855 | 0.2100 | 0.1823 | 0.2500 | 0.2420 |
| C | 0.5621 | 0.4983 | 0.5128 | 0.5072 | 0.5100 | 0.5249 | 0.5000 | 0.6538 |
| HN | 0.2391 | 0.2277 | 0.2286 | 0.2276 | 0.3300 | 0.3215 | 0.3300 | 0.3175 |
| O | -0.5889 | -0.6251 | -0.6177 | -0.6318 | -0.5100 | -0.5603 | -0.5000 | -0.5777 |
| HA | 0.1202 | 0.1329 | 0.1329 | 0.1318 | 0.1000 | 0.0937 | 0.0600 | 0.0670 |

Table 3.39: Partial-charge parameters for N-termini of glutamine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1493 | 0.2157 | 0.1993 | 0.2088 | −0.3000 | −0.2416 | −0.3000 | −0.3364 |
| CA | 0.0769 | 0.1128 | 0.1091 | 0.1085 | 0.2100 | 0.1803 | 0.2500 | 0.2559 |
| C | 0.6123 | 0.5464 | 0.5631 | 0.5552 | 0.5100 | 0.5450 | 0.5000 | 0.6335 |
| HN | 0.1996 | 0.1899 | 0.1912 | 0.1899 | 0.3300 | 0.3236 | 0.3300 | 0.3163 |
| O | −0.5713 | −0.5957 | −0.5904 | −0.6010 | −0.5100 | −0.5528 | −0.5000 | −0.5517 |
| HA | 0.1015 | 0.1126 | 0.1135 | 0.1111 | 0.1000 | 0.0910 | 0.0600 | 0.0681 |

Table 3.40: Partial-charge parameters for N-termini of methionine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1592 | 0.2361 | 0.2231 | 0.2303 | −0.3000 | −0.2120 | −0.3000 | −0.3114 |
| CA | 0.0429 | 0.0873 | 0.0836 | 0.0838 | 0.2100 | 0.2087 | 0.2500 | 0.2657 |
| C | 0.6123 | 0.5649 | 0.5815 | 0.5754 | 0.5100 | 0.5315 | 0.5000 | 0.6635 |
| HN | 0.1984 | 0.1868 | 0.1881 | 0.1865 | 0.3300 | 0.3176 | 0.3300 | 0.3142 |
| O | −0.5713 | −0.6013 | −0.5956 | −0.6084 | −0.5100 | −0.5510 | −0.5000 | −0.5686 |
| HA | 0.1116 | 0.1381 | 0.1393 | 0.1372 | 0.1000 | 0.0961 | 0.0600 | 0.0921 |

Table 3.41: Partial-charge parameters for N-termini of lysine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.0966 | 0.1401 | 0.1255 | 0.1335 | −0.3000 | −0.2359 | −0.3000 | −0.3274 |
| CA | −0.0101 | 0.0693 | 0.0689 | 0.0680 | 0.2100 | 0.1908 | 0.2500 | 0.2677 |
| C | 0.7214 | 0.6500 | 0.6648 | 0.6579 | 0.5100 | 0.5382 | 0.5000 | 0.6431 |
| HN | 0.2165 | 0.2082 | 0.2092 | 0.2081 | 0.3300 | 0.3218 | 0.3300 | 0.3162 |
| O | −0.6013 | −0.6182 | −0.6157 | −0.6230 | −0.5100 | −0.5397 | −0.5000 | −0.5419 |
| HA | 0.1180 | 0.1315 | 0.1318 | 0.1306 | 0.1000 | 0.0960 | 0.0600 | 0.0929 |

Table 3.42: Partial-charge parameters for N-termini of arginine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | 0.1305 | 0.1842 | 0.1695 | 0.1771 | −0.3000 | −0.2250 | −0.3000 | −0.2992 |
| CA | −0.0359 | 0.0460 | 0.0460 | 0.0445 | 0.2100 | 0.1802 | 0.2500 | 0.2797 |
| C | 0.7214 | 0.6489 | 0.6630 | 0.6559 | 0.5100 | 0.5472 | 0.5000 | 0.6477 |
| HN | 0.2083 | 0.1983 | 0.1992 | 0.1981 | 0.3300 | 0.3181 | 0.3300 | 0.3104 |
| O | −0.6013 | −0.6301 | −0.6262 | −0.6337 | −0.5100 | −0.5543 | −0.5000 | −0.5487 |
| HA | 0.1242 | 0.1265 | 0.1268 | 0.1269 | 0.1000 | 0.0773 | 0.0600 | 0.0740 |

Table 3.43: Partial-charge parameters for C-termini of glycine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3135 | −0.3278 | −0.3170 | −0.4700 | −0.4381 | −0.5000 | −0.5153 |
| CA | −0.2493 | −0.2066 | −0.2093 | −0.2075 | −0.0200 | 0.0185 | −0.0200 | −0.0091 |
| C | 0.7231 | 0.6784 | 0.6956 | 0.6835 | 0.3400 | 0.3609 | 0.7000 | 0.8459 |
| HN | 0.2681 | 0.2454 | 0.2471 | 0.2442 | 0.3100 | 0.3004 | 0.3000 | 0.2615 |
| OXT | −0.7855 | −0.8006 | −0.8003 | −0.8007 | −0.6700 | −0.6895 | −0.8000 | −0.8273 |
| HA | 0.1056 | 0.0987 | 0.0976 | 0.0991 | 0.0900 | 0.0687 | 0.0600 | 0.0358 |

Table 3.44: Partial-charge parameters for C-termini of alanine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3018 | −0.3147 | −0.3071 | −0.4700 | −0.3909 | −0.5000 | −0.5224 |
| CA | −0.1532 | −0.1324 | −0.1322 | −0.1358 | 0.0700 | 0.0427 | 0.0400 | 0.0301 |
| C | 0.7731 | 0.6900 | 0.6998 | 0.6994 | 0.3400 | 0.3515 | 0.7000 | 0.8687 |
| HN | 0.2681 | 0.2285 | 0.2308 | 0.2279 | 0.3100 | 0.2709 | 0.3000 | 0.2610 |
| OXT | −0.8055 | −0.8066 | −0.8015 | −0.8104 | −0.6700 | −0.6857 | −0.8000 | −0.8284 |
| HA | 0.1067 | 0.1145 | 0.1156 | 0.1144 | 0.0900 | 0.0741 | 0.0600 | 0.0786 |

Table 3.45: Partial-charge parameters for C-termini of valine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3150 | −0.3282 | −0.3217 | −0.4700 | −0.4287 | −0.5000 | −0.4963 |
| CA | −0.3352 | −0.2862 | −0.2906 | −0.2926 | 0.0700 | 0.0821 | 0.0400 | 0.1025 |
| C | 0.8350 | 0.8086 | 0.8309 | 0.8195 | 0.3400 | 0.3603 | 0.7000 | 0.8030 |
| HN | 0.2681 | 0.2579 | 0.2615 | 0.2576 | 0.3100 | 0.2973 | 0.3000 | 0.2315 |
| OXT | −0.8173 | −0.8184 | −0.8192 | −0.8216 | −0.6700 | −0.6810 | −0.8000 | −0.8198 |
| HA | 0.1438 | 0.1683 | 0.1676 | 0.1673 | 0.0900 | 0.0798 | 0.0600 | 0.0659 |

Table 3.46: Partial-charge parameters for C-termini of leucine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3038 | −0.3220 | −0.3111 | −0.4700 | −0.4123 | −0.5000 | −0.5173 |
| CA | −0.2874 | −0.2414 | −0.2462 | −0.2428 | 0.0700 | 0.0702 | 0.0400 | 0.0644 |
| C | 0.8326 | 0.7786 | 0.7952 | 0.7870 | 0.3400 | 0.3495 | 0.7000 | 0.8293 |
| HN | 0.2681 | 0.2423 | 0.2475 | 0.2413 | 0.3100 | 0.2886 | 0.3000 | 0.2582 |
| OXT | −0.8199 | −0.8248 | −0.8226 | −0.8281 | −0.6700 | −0.6819 | −0.8000 | −0.8196 |
| HA | 0.1346 | 0.1769 | 0.1770 | 0.1754 | 0.0900 | 0.0925 | 0.0600 | 0.1178 |

Table 3.47: Partial-charge parameters for C-termini of isoleucine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3171 | −0.3340 | −0.3252 | −0.4700 | −0.4268 | −0.5000 | −0.5004 |
| CA | −0.3070 | −0.2500 | −0.2530 | −0.2550 | 0.0700 | 0.0930 | 0.0400 | 0.1138 |
| C | 0.8343 | 0.8000 | 0.8199 | 0.8108 | 0.3400 | 0.3586 | 0.7000 | 0.8098 |
| HN | 0.2681 | 0.2424 | 0.2484 | 0.2422 | 0.3100 | 0.2917 | 0.3000 | 0.2261 |
| OXT | −0.8190 | −0.8212 | −0.8208 | −0.8238 | −0.6700 | −0.6860 | −0.8000 | −0.8316 |
| HA | 0.1375 | 0.1559 | 0.1557 | 0.1547 | 0.0900 | 0.0835 | 0.0600 | 0.0604 |

Table 3.48: Partial-charge parameters for C-termini of serine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3141 | −0.3255 | −0.3176 | −0.4700 | −0.4032 | −0.5000 | −0.5306 |
| CA | −0.2563 | −0.2168 | −0.2191 | −0.2193 | 0.0700 | 0.0294 | 0.0400 | −0.0005 |
| C | 0.8113 | 0.7581 | 0.7728 | 0.7657 | 0.3400 | 0.3986 | 0.7000 | 0.8559 |
| HN | 0.2681 | 0.2481 | 0.2488 | 0.2472 | 0.3100 | 0.3094 | 0.3000 | 0.2743 |
| OXT | −0.8132 | −0.8146 | −0.8130 | −0.8165 | −0.6700 | −0.6819 | −0.8000 | −0.8076 |
| HA | 0.1304 | 0.1453 | 0.1457 | 0.1449 | 0.0900 | 0.0813 | 0.0600 | 0.0780 |

Table 3.49: Partial-charge parameters for C-termini of threonine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3208 | −0.3317 | −0.3246 | −0.4700 | −0.4254 | −0.5000 | −0.5204 |
| CA | −0.2315 | −0.1811 | −0.1848 | −0.1857 | 0.0700 | 0.0768 | 0.0400 | 0.0277 |
| C | 0.7810 | 0.7245 | 0.7427 | 0.7328 | 0.3400 | 0.3635 | 0.7000 | 0.8162 |
| HN | 0.2681 | 0.2641 | 0.2653 | 0.2636 | 0.3100 | 0.3209 | 0.3000 | 0.2738 |
| OXT | −0.8044 | −0.8026 | −0.8028 | −0.8056 | −0.6700 | −0.6775 | −0.8000 | −0.8033 |
| HA | 0.1207 | 0.1226 | 0.1240 | 0.1218 | 0.0900 | 0.0701 | 0.0600 | 0.0837 |

Table 3.50: Partial-charge parameters for C-termini of cysteine (-SH). See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3373 | −0.3524 | −0.3409 | −0.4700 | −0.4270 | −0.5000 | −0.5385 |
| CA | −0.1598 | −0.1075 | −0.1092 | −0.1114 | 0.0700 | 0.0471 | 0.0400 | 0.0665 |
| C | 0.7497 | 0.6649 | 0.6775 | 0.6726 | 0.3400 | 0.3744 | 0.7000 | 0.8296 |
| HN | 0.2681 | 0.2470 | 0.2491 | 0.2447 | 0.3100 | 0.3016 | 0.3000 | 0.2591 |
| OXT | −0.7981 | −0.7863 | −0.7833 | −0.7886 | −0.6700 | −0.6714 | −0.8000 | −0.8066 |
| HA | 0.1396 | 0.1453 | 0.1459 | 0.1446 | 0.0900 | 0.0844 | 0.0600 | 0.0895 |

Table 3.51: Partial-charge parameters for C-termini of cysteine (-SS-). See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3114 | −0.3271 | −0.3177 | −0.4700 | −0.4311 | −0.5000 | −0.5186 |
| CA | −0.1283 | −0.1247 | −0.1303 | −0.1281 | 0.0700 | 0.0267 | 0.0400 | 0.0249 |
| C | 0.7618 | 0.7069 | 0.7245 | 0.7150 | 0.3400 | 0.3690 | 0.7000 | 0.8258 |
| HN | 0.2681 | 0.2508 | 0.2549 | 0.2508 | 0.3100 | 0.3118 | 0.3000 | 0.2555 |
| OXT | −0.8041 | −0.8034 | −0.8041 | −0.8064 | −0.6700 | −0.6790 | −0.8000 | −0.8078 |
| HA | 0.0938 | 0.1001 | 0.1016 | 0.0999 | 0.0900 | 0.0758 | 0.0600 | 0.0702 |

Table 3.52: Partial-charge parameters for C-termini of proline. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.2802 | −0.2479 | −0.2641 | −0.2521 | −0.2900 | −0.2580 | −0.1400 | −0.2326 |
| CA | −0.1236 | −0.0962 | −0.1017 | −0.0946 | 0.0200 | −0.0194 | −0.0900 | −0.0713 |
| C | 0.6631 | 0.6501 | 0.6679 | 0.6519 | 0.3400 | 0.4154 | 0.7000 | 0.9097 |
| OXT | −0.7697 | −0.7673 | −0.7674 | −0.7673 | −0.6700 | −0.6962 | −0.8000 | −0.8298 |
| HA | 0.0776 | 0.0832 | 0.0835 | 0.0848 | 0.0900 | 0.0827 | 0.0600 | 0.0285 |

Table 3.53: Partial-charge parameters for C-termini of phenylalanine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3180 | −0.3331 | −0.3209 | −0.4700 | −0.4240 | −0.5000 | −0.5047 |
| CA | −0.1756 | −0.1315 | −0.1363 | −0.1327 | 0.0700 | 0.0526 | 0.0400 | 0.0678 |
| C | 0.7660 | 0.6946 | 0.7111 | 0.7030 | 0.3400 | 0.3603 | 0.7000 | 0.8238 |
| HN | 0.2681 | 0.2439 | 0.2473 | 0.2423 | 0.3100 | 0.2985 | 0.3000 | 0.2450 |
| OXT | −0.8026 | −0.8050 | −0.8019 | −0.8097 | −0.6700 | −0.6820 | −0.8000 | −0.8229 |
| HA | 0.1098 | 0.1279 | 0.1287 | 0.1250 | 0.0900 | 0.0788 | 0.0600 | 0.0813 |

Table 3.54: Partial-charge parameters for C-termini of tyrosine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3179 | −0.3335 | −0.3272 | −0.4700 | −0.4350 | −0.5000 | −0.5111 |
| CA | −0.1911 | −0.1499 | −0.1547 | −0.1505 | 0.0700 | 0.0582 | 0.0400 | 0.0794 |
| C | 0.7817 | 0.7387 | 0.7552 | 0.7433 | 0.3400 | 0.3755 | 0.7000 | 0.8188 |
| HN | 0.2681 | 0.2408 | 0.2456 | 0.2401 | 0.3100 | 0.2932 | 0.3000 | 0.2357 |
| OXT | −0.8070 | −0.8132 | −0.8118 | −0.8166 | −0.6700 | −0.6856 | −0.8000 | −0.8206 |
| HA | 0.1092 | 0.1161 | 0.1178 | 0.1221 | 0.0900 | 0.0696 | 0.0600 | 0.0638 |

Table 3.55: Partial-charge parameters for C-termini of tryptophan. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3127 | −0.3276 | −0.3175 | −0.4700 | −0.4053 | −0.5000 | −0.5037 |
| CA | −0.2064 | −0.1563 | −0.1600 | −0.1617 | 0.0700 | 0.0603 | 0.0400 | 0.0430 |
| C | 0.7658 | 0.6932 | 0.7115 | 0.7019 | 0.3400 | 0.3623 | 0.7000 | 0.8138 |
| HN | 0.2681 | 0.2310 | 0.2339 | 0.2306 | 0.3100 | 0.2945 | 0.3000 | 0.2391 |
| OXT | −0.8011 | −0.8086 | −0.8077 | −0.8114 | −0.6700 | −0.6836 | −0.8000 | −0.8243 |
| HA | 0.1272 | 0.1163 | 0.1179 | 0.1162 | 0.0900 | 0.0824 | 0.0600 | 0.0520 |

Table 3.56: Partial-charge parameters for C-termini of histidine (+). See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3481 | −0.2917 | −0.3101 | −0.2984 | −0.4700 | −0.4169 | −0.5000 | −0.5224 |
| CA | −0.1503 | −0.0799 | −0.0826 | −0.0813 | 0.0700 | 0.0615 | 0.0400 | 0.0764 |
| C | 0.8032 | 0.7222 | 0.7371 | 0.7305 | 0.3400 | 0.3696 | 0.7000 | 0.8420 |
| HN | 0.2764 | 0.2500 | 0.2555 | 0.2495 | 0.3100 | 0.2938 | 0.3000 | 0.2539 |
| OXT | −0.8177 | −0.8206 | −0.8184 | −0.8224 | −0.6700 | −0.6873 | −0.8000 | −0.8267 |
| HA | 0.1115 | 0.1241 | 0.1253 | 0.1262 | 0.0900 | 0.0663 | 0.0600 | 0.0629 |

Table 3.57: Partial-charge parameters for C-termini of aspartic acid. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.5192 | −0.4228 | −0.4406 | −0.4265 | −0.4700 | −0.4185 | −0.5000 | −0.4537 |
| CA | −0.1810 | −0.1448 | −0.1488 | −0.1453 | 0.0700 | 0.0775 | 0.0400 | 0.0396 |
| C | 0.7256 | 0.6658 | 0.6834 | 0.6732 | 0.3400 | 0.3429 | 0.7000 | 0.8620 |
| HN | 0.3055 | 0.2700 | 0.2739 | 0.2694 | 0.3100 | 0.2959 | 0.3000 | 0.2313 |
| OXT | −0.7887 | −0.8028 | −0.8008 | −0.8048 | −0.6700 | −0.6809 | −0.8000 | −0.8416 |
| HA | 0.1046 | 0.1179 | 0.1160 | 0.1177 | 0.0900 | 0.0710 | 0.0600 | 0.0666 |

Table 3.58: Partial-charge parameters for C-termini of asparagine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.2915 | −0.3110 | −0.2986 | −0.4700 | −0.4107 | −0.5000 | −0.4827 |
| CA | −0.1927 | −0.1678 | −0.1712 | −0.1671 | 0.0700 | 0.0679 | 0.0400 | 0.0423 |
| C | 0.8050 | 0.7430 | 0.7599 | 0.7504 | 0.3400 | 0.3617 | 0.7000 | 0.8776 |
| HN | 0.2681 | 0.2346 | 0.2393 | 0.2340 | 0.3100 | 0.2951 | 0.3000 | 0.2460 |
| OXT | −0.8147 | −0.8221 | −0.8201 | −0.8251 | −0.6700 | −0.6843 | −0.8000 | −0.8406 |
| HA | 0.1358 | 0.1701 | 0.1699 | 0.1690 | 0.0900 | 0.0851 | 0.0600 | 0.0937 |

Table 3.59: Partial-charge parameters for C-termini of glutamic acid. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.5192 | −0.4277 | −0.4405 | −0.4331 | −0.4700 | −0.3961 | −0.5000 | −0.5401 |
| CA | −0.2000 | −0.1814 | −0.1844 | −0.1843 | 0.0700 | 0.0423 | 0.0400 | 0.0320 |
| C | 0.7420 | 0.6782 | 0.6927 | 0.6871 | 0.3400 | 0.3549 | 0.7000 | 0.8538 |
| HN | 0.3055 | 0.2714 | 0.2739 | 0.2709 | 0.3100 | 0.2845 | 0.3000 | 0.2626 |
| OXT | −0.7930 | −0.8111 | −0.8074 | −0.8144 | −0.6700 | −0.6951 | −0.8000 | −0.8389 |
| HA | 0.1399 | 0.1526 | 0.1526 | 0.1515 | 0.0900 | 0.0837 | 0.0600 | 0.0670 |

Table 3.60: Partial-charge parameters for C-termini of glutamine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3157 | −0.3321 | −0.3226 | −0.4700 | −0.4116 | −0.5000 | −0.5364 |
| CA | −0.2108 | −0.1749 | −0.1786 | −0.1792 | 0.0700 | 0.0403 | 0.0400 | 0.0459 |
| C | 0.7775 | 0.7116 | 0.7283 | 0.7204 | 0.3400 | 0.3750 | 0.7000 | 0.8335 |
| HN | 0.2681 | 0.2389 | 0.2428 | 0.2389 | 0.3100 | 0.2909 | 0.3000 | 0.2590 |
| OXT | −0.8042 | −0.8164 | −0.8138 | −0.8191 | −0.6700 | −0.6914 | −0.8000 | −0.8259 |
| HA | 0.1232 | 0.1343 | 0.1352 | 0.1328 | 0.0900 | 0.0810 | 0.0600 | 0.0681 |

Table 3.61: Partial-charge parameters for C-termini of methionine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3821 | −0.3052 | −0.3182 | −0.3110 | −0.4700 | −0.3820 | −0.5000 | −0.5114 |
| CA | −0.2441 | −0.1997 | −0.2034 | −0.2032 | 0.0700 | 0.0687 | 0.0400 | 0.0557 |
| C | 0.8013 | 0.7539 | 0.7705 | 0.7644 | 0.3400 | 0.3615 | 0.7000 | 0.8635 |
| HN | 0.2681 | 0.2334 | 0.2371 | 0.2325 | 0.3100 | 0.2727 | 0.3000 | 0.2525 |
| OXT | −0.8105 | −0.8255 | −0.8227 | −0.8290 | −0.6700 | −0.6905 | −0.8000 | −0.8343 |
| HA | 0.1277 | 0.1542 | 0.1554 | 0.1533 | 0.0900 | 0.0861 | 0.0600 | 0.0921 |

Table 3.62: Partial-charge parameters for C-termini of lysine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3481 | −0.3046 | −0.3192 | −0.3112 | −0.4700 | −0.4059 | −0.5000 | −0.5274 |
| CA | −0.2964 | −0.2170 | −0.2174 | −0.2183 | 0.0700 | 0.0508 | 0.0400 | 0.0577 |
| C | 0.8488 | 0.7774 | 0.7922 | 0.7853 | 0.3400 | 0.3682 | 0.7000 | 0.8431 |
| HN | 0.2764 | 0.2515 | 0.2544 | 0.2512 | 0.3100 | 0.2854 | 0.3000 | 0.2587 |
| OXT | −0.8252 | −0.8337 | −0.8324 | −0.8361 | −0.6700 | −0.6849 | −0.8000 | −0.8210 |
| HA | 0.1438 | 0.1573 | 0.1576 | 0.1564 | 0.0900 | 0.0860 | 0.0600 | 0.0929 |

Table 3.63: Partial-charge parameters for C-termini of arginine. See the caption in Table 3.1.

| Atom | AMB | Opt(94) | Opt(96) | Opt(99) | CHA | Opt(CH) | OPLS | Opt(OP) |
|---|---|---|---|---|---|---|---|---|
| N | −0.3481 | −0.2944 | −0.3091 | −0.3015 | −0.4700 | −0.3950 | −0.5000 | −0.4992 |
| CA | −0.3117 | −0.2298 | −0.2298 | −0.2313 | 0.0700 | 0.0402 | 0.0400 | 0.0697 |
| C | 0.8557 | 0.7832 | 0.7973 | 0.7902 | 0.3400 | 0.3772 | 0.7000 | 0.8477 |
| HN | 0.2764 | 0.2464 | 0.2490 | 0.2459 | 0.3100 | 0.2742 | 0.3000 | 0.2411 |
| OXT | −0.8266 | −0.8410 | −0.8390 | −0.8428 | −0.6700 | −0.6920 | −0.8000 | −0.8243 |
| HA | 0.1447 | 0.1470 | 0.1473 | 0.1474 | 0.0900 | 0.0673 | 0.0600 | 0.0740 |

Table 3.64: Torsion parameters of $\phi$ angle. Parm94, Parm96, Parm99, CHARMM, and OPLS are AMBER parm94, AMBER parm96, AMBER parm99, CHARMM version 22, and OPLS-AA force fields, respectively. "Optimized" stands for the corresponding optimized force field.

| Force field | $V_a$ | $n_a$ | $\gamma_a$ | $V_b$ | $n_b$ | $\gamma_b$ | $V_c$ | $n_c$ | $\gamma_c$ |
|---|---|---|---|---|---|---|---|---|---|
| Parm94 | 0.200 | 2 | 180.0 | — | — | — | — | — | — |
| Optimized | 0.191 | 1 | 0.0 | 0.146 | 2 | 180.0 | −0.223 | 3 | 0.0 |
| Parm96 | 0.850 | 1 | 0.0 | 0.300 | 2 | 180.0 | — | — | — |
| Optimized | 1.182 | 1 | 0.0 | 0.359 | 2 | 180.0 | −0.410 | 3 | 0.0 |
| Parm99 | 0.800 | 1 | 0.0 | 0.850 | 2 | 180.0 | — | — | — |
| Optimized | 1.380 | 1 | 0.0 | 0.599 | 2 | 180.0 | −0.330 | 3 | 0.0 |
| CHARMM | 0.200 | 1 | 180.0 | — | — | — | — | — | — |
| Optimized | −0.047 | 1 | 180.0 | 0.240 | 2 | 180.0 | −0.015 | 3 | 0.0 |
| OPLS | −2.365 | 1 | 0.0 | 0.912 | 2 | 180.0 | −0.850 | 3 | 0.0 |
| Optimized | 0.502 | 1 | 0.0 | 1.811 | 2 | 180.0 | −0.567 | 3 | 0.0 |

Table 3.65: Torsion parameters of $\psi$ angle. See the caption in Table 3.64.

| Force field | $V_a$ | $n_a$ | $\gamma_a$ | $V_b$ | $n_b$ | $\gamma_b$ | $V_c$ | $n_c$ | $\gamma_c$ |
|---|---|---|---|---|---|---|---|---|---|
| Parm94 | 0.750 | 1 | 180.0 | 1.350 | 2 | 180.0 | 0.400 | 4 | 180.0 |
| Optimized | −0.368 | 1 | 180.0 | 1.658 | 2 | 180.0 | 0.265 | 4 | 180.0 |
| Parm96 | 0.850 | 1 | 0.0 | 0.300 | 2 | 180.0 | — | — | — |
| Optimized | 0.039 | 1 | 0.0 | 1.011 | 2 | 180.0 | 0.104 | 3 | 0.0 |
| Parm99 | 1.700 | 1 | 180.0 | 2.000 | 2 | 180.0 | — | — | — |
| Optimized | 0.228 | 1 | 180.0 | 1.684 | 2 | 180.0 | −0.031 | 3 | 0.0 |
| CHARMM | 0.600 | 1 | 0.0 | — | — | — | — | — | — |
| Optimized | 0.321 | 1 | 0.0 | 0.028 | 2 | 180.0 | 0.251 | 3 | 0.0 |
| OPLS | 1.816 | 1 | 0.0 | 1.222 | 2 | 180.0 | 1.581 | 3 | 0.0 |
| Optimized | 0.880 | 1 | 0.0 | 1.479 | 2 | 180.0 | 0.952 | 3 | 0.0 |

Table 3.66: Torsion parameters of $\phi$ angle. Opt(99), Opt(CH) and Opt(94) are the optimized AMBER parm99 with $K_x = 200$ kcal/mol/Å$^2$, CHARMM version 22 with $K_x = 100$ kcal/mol/Å$^2$, and optimized AMBER parm94 with $K_x = 400$ kcal/mol/Å$^2$ with no constrain of torsion-energy parameters.

| Force field | $V_a$ | $n_a$ | $\gamma_a$ | $V_b$ | $n_b$ | $\gamma_b$ | $V_c$ | $n_c$ | $\gamma_c$ |
|---|---|---|---|---|---|---|---|---|---|
| Opt(99) ($K_x = 200$) | 1.258 | 1 | 0.0 | 0.611 | 2 | 180.0 | −0.220 | 3 | 0.0 |
| Opt(CH) ($K_x = 100$) | −0.025 | 1 | 180.0 | 0.204 | 2 | 180.0 | 0.014 | 3 | 0.0 |
| Opt(94) ($K_x = 400$) | 1.852 | 1 | 0.0 | 0.534 | 2 | 180.0 | −0.014 | 3 | 0.0 |

Table 3.67: Torsion parameters of $\psi$ angle. See the caption in Table 3.66.

| Force field | $V_a$ | $n_a$ | $\gamma_a$ | $V_b$ | $n_b$ | $\gamma_b$ | $V_c$ | $n_c$ | $\gamma_c$ |
|---|---|---|---|---|---|---|---|---|---|
| Opt(99) ($K_x = 200$) | 0.701 | 1 | 180.0 | 1.819 | 2 | 180.0 | −0.012 | 3 | 0.0 |
| Opt(CH) ($K_x = 100$) | 0.315 | 1 | 0.0 | 0.196 | 2 | 180.0 | 0.090 | 3 | 0.0 |
| Opt(94) ($K_x = 400$) | −0.185 | 1 | 0.0 | 1.323 | 2 | 180.0 | −0.014 | 3 | 0.0 |

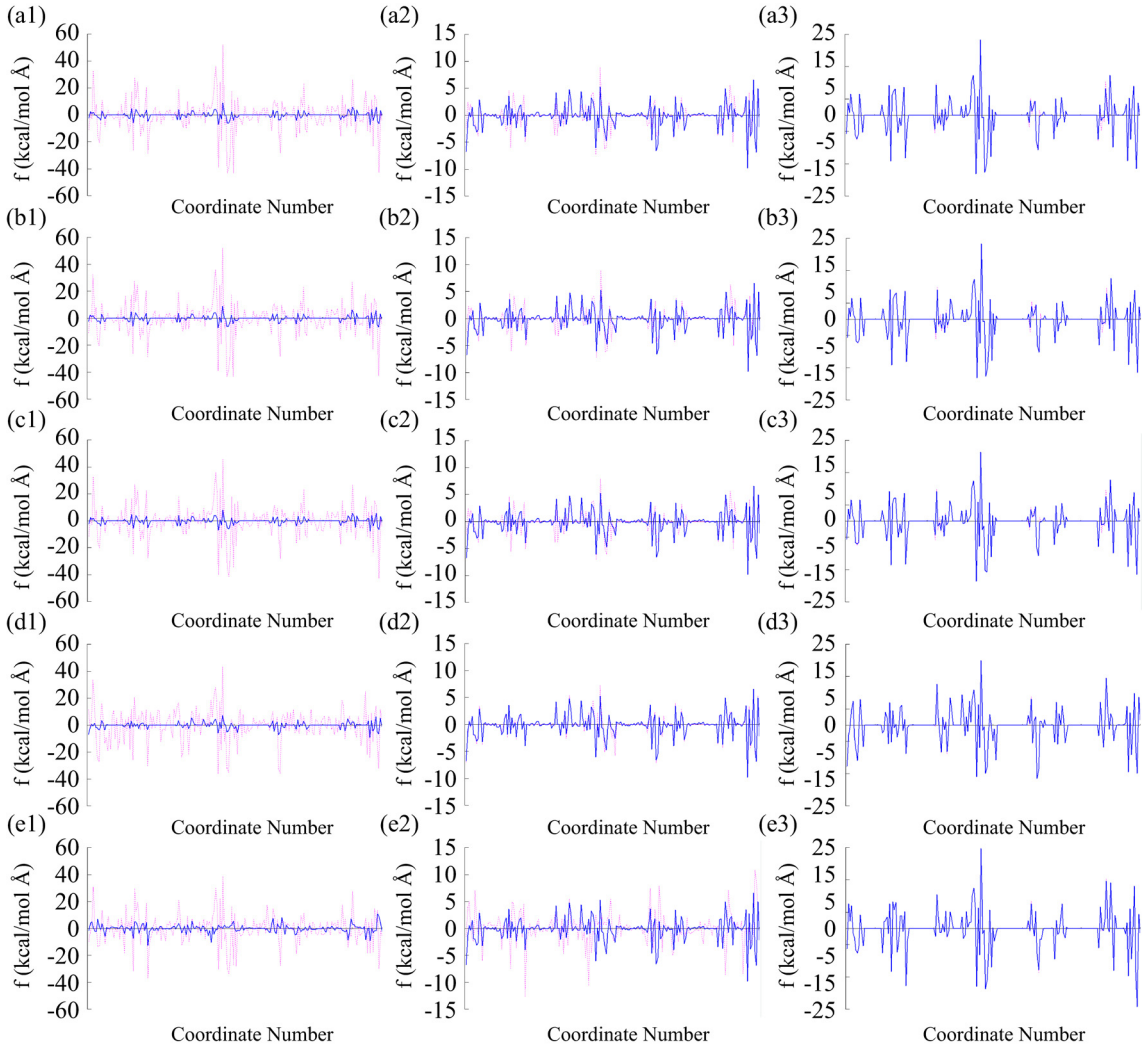


Figure 3.1: (a1)-(e1) Components of the force acting on the atoms in the original structure of 2LIS (red dotted curve) and its refined structure (blue solid curve). The abscissa is the components of the coordinates (i.e., $x_1$, $y_1$, $z_1$, $x_2$, $y_2$, $z_2$, $\cdots$) of the atoms, and the ordinate is the value of the corresponding components of the force acting on these atoms. (a2)-(e2) Components of the force acting on the atoms in the refined structure of 2LIS. The force fields are the original parameters (red dotted curve) and original parameters with the partial-charge parameters optimized (blue solid curve). (a3)-(e3) Components of the force acting on the atoms in the refined structure of 2LIS. The force field is the original parameters with the partial-charge parameters optimized (red dotted curve) and original parameters with the partial-charge and torsion-energy parameters optimized (blue solid curve). The force fields are the original AMBER parm94, AMBER parm96, AMBER parm99, CHARMM version 22, and OPLS-AA, respectively. Only 200 components are shown.

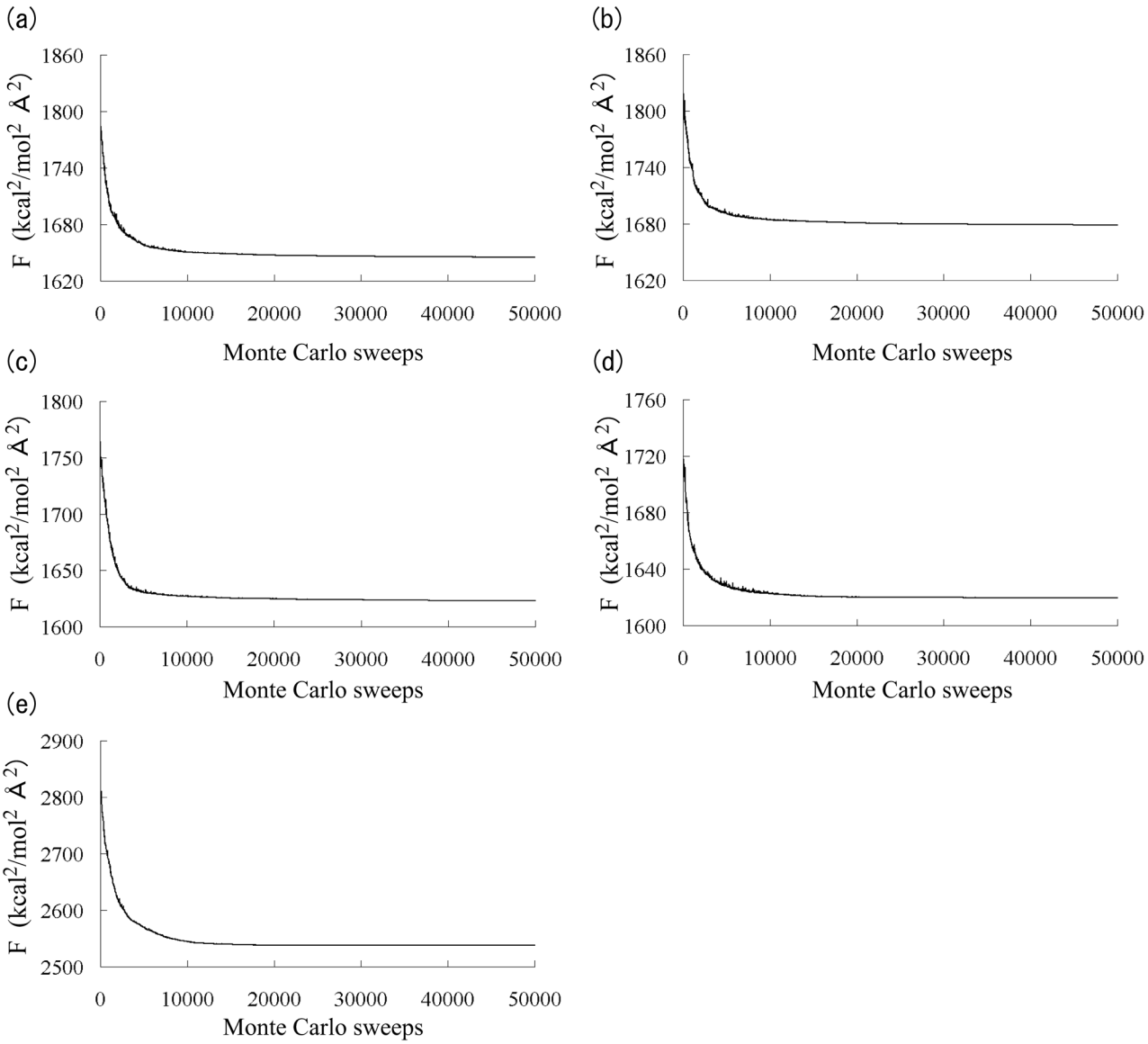


Figure 3.2: Time series of MC simulated annealing simulations in force-field parameter space of partial charges for AMBER parm94 (a), AMBER parm96 (b), AMBER parm99 (c), CHARMM version 22 (d), and OPLS-AA (e). The ordinate is the value of $F$ in Eq. (2.7).

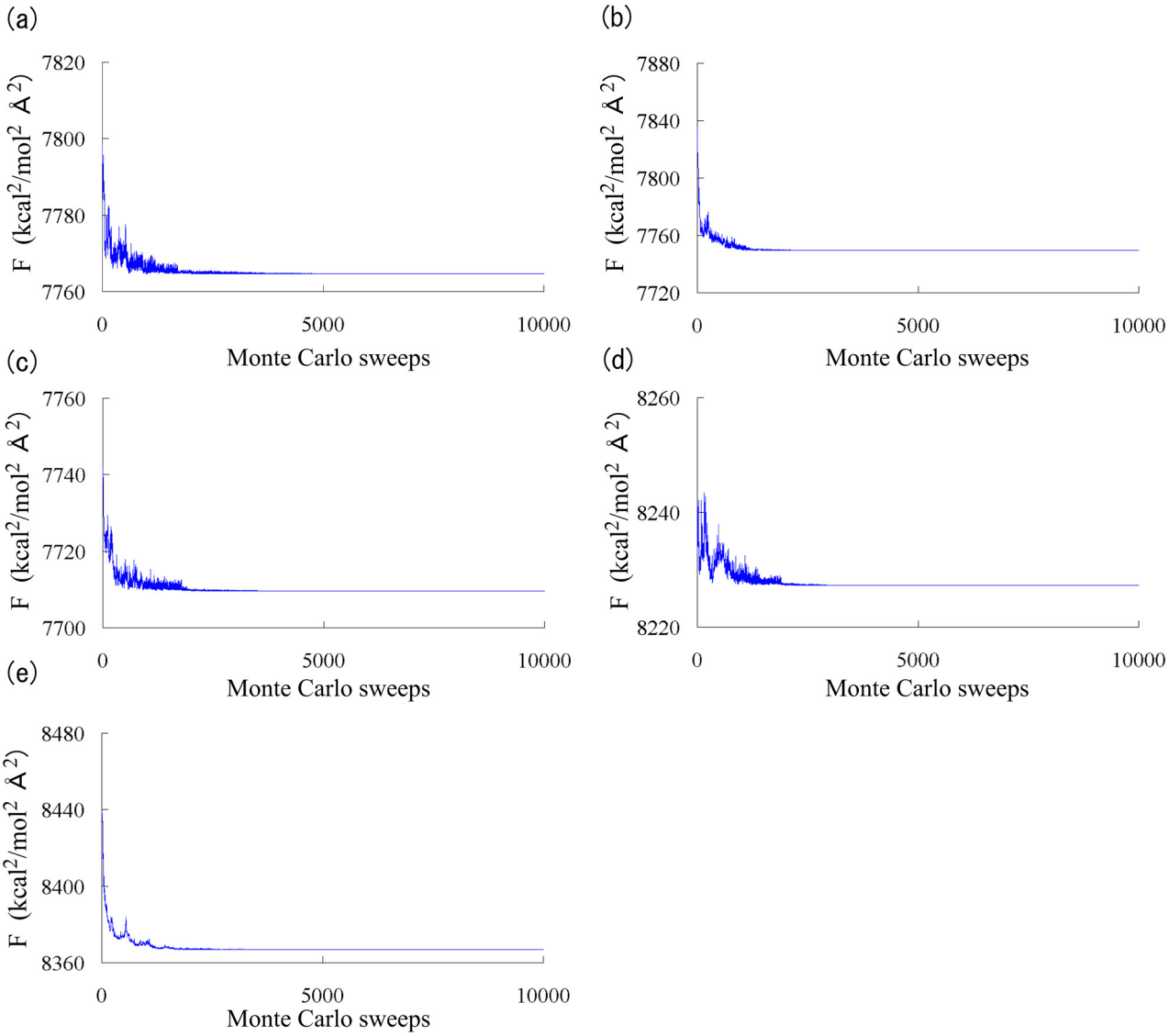


Figure 3.3: Time series of MC simulated annealing simulations in force-field parameter space of torsion-energy parameters of AMBER parm94 (a), AMBER parm96 (b), AMBER parm99 (c), CHARMM version 22 (d), and OPLS-AA (e). The ordinate is the value of $F$ in Eq. (2.7).

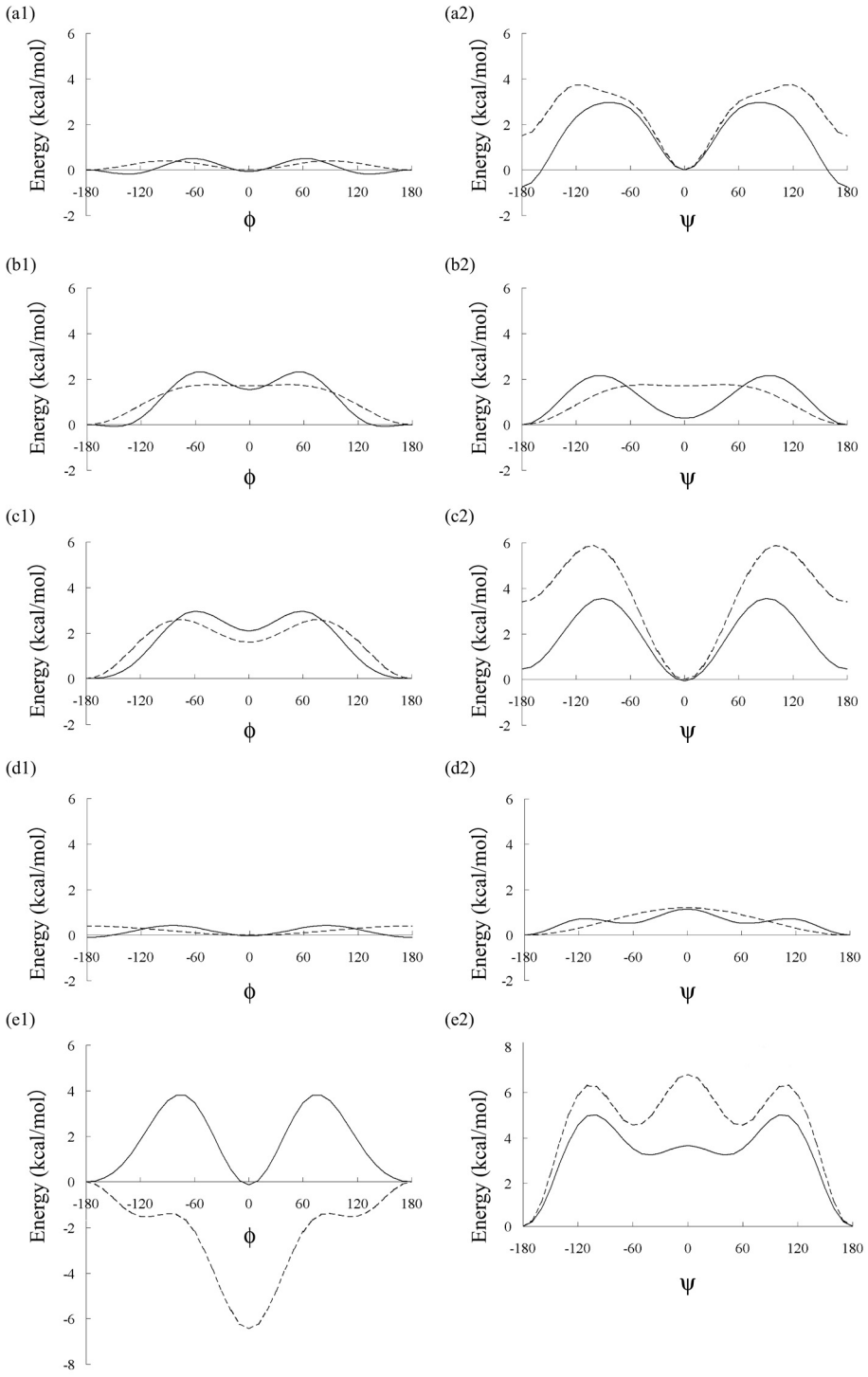


Figure 3.4: Backbone torsion-energy curves as functions of $\phi$ (in degrees) and $\psi$ (in degrees). The force fields are AMBER parm94 (a), AMBER parm96 (b), AMBER parm99 (c), CHARMM version 22 (d), and OPLS-AA (e). The results for the original force fields are represented by dotted curves, and those for the optimized force fields are by solid curves.

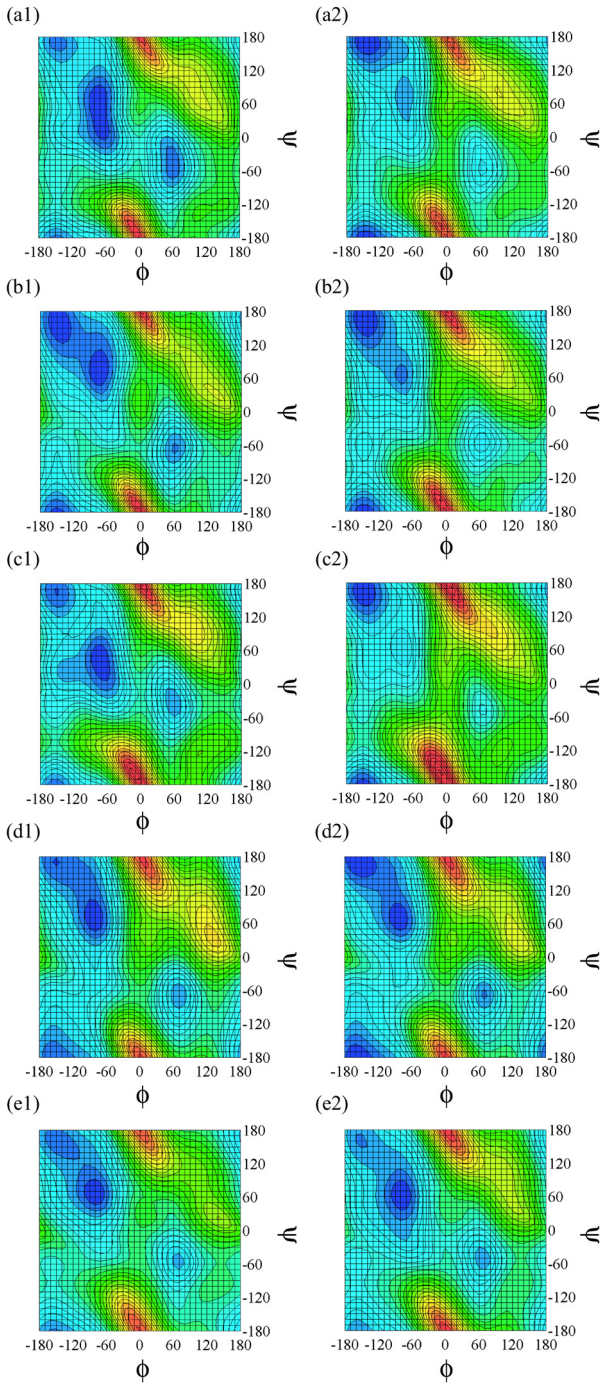


Figure 3.5: Potential-energy surfaces of alanine dipeptide. The force fields are the original AMBER parm94 (a1), AMBER parm96 (b1), AMBER parm99 (c1), CHARMM version 22 (d1), and OPLS-AA (e1), and the corresponding optimized parameters (a2)-(e2). The contour maps were evaluated every 10° of $\phi$ and $\psi$ angles and plotted every 1 kcal/mol, after minimizing the total potential energy in vacuum with the backbone structures fixed. The bluer the color is, the lower the potential energy surface is. As the potential-energy value increases, the color changes from blue to green, to yellow, and to red.

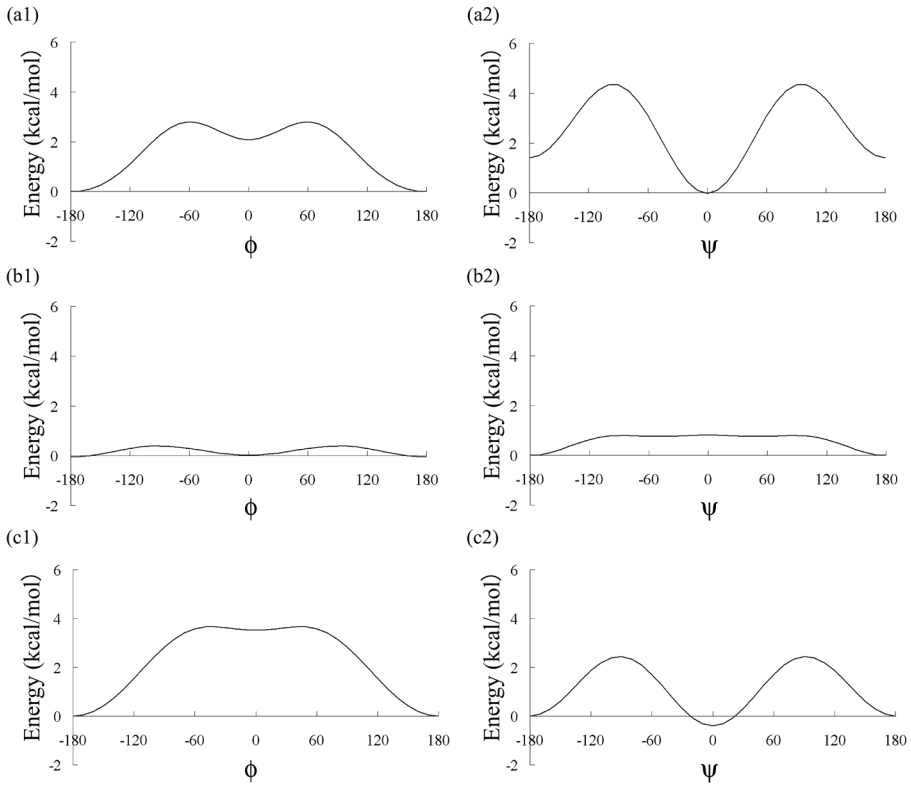


Figure 3.6: Backbone torsion-energy curves as functions of $\phi$ (in degrees) (a1) and (b1) and $\psi$ (in degrees) (a2) and (b2). (a) The optimized AMBER parm99 with $K_x = 200$ kcal/mol/Å$^2$. (b) The optimized CHARMM version 22 with $K_x = 100$ kcal/mol/Å$^2$.

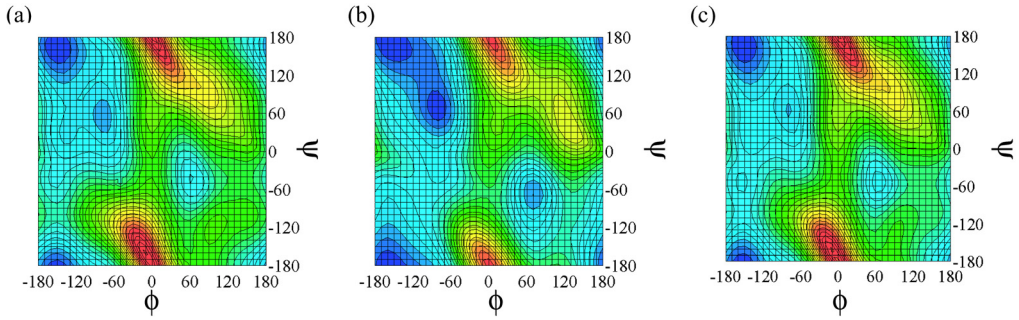


Figure 3.7: Potential-energy surfaces of alanine dipeptide for the optimized AMBER parm99 with $K_x = 200$ kcal/mol/Å$^2$ (a) and the optimized CHARMM version 22 with $K_x = 100$ kcal/mol/Å$^2$ (b). See the caption of Fig. 3.5.

# Chapter 4

# Comparisons of Force Fields by Folding Simulations of Short Peptides

Yoshitake Sakae and Yuko Okamoto, "Optimization of protein force-field parameters with the Protein Data Bank," *Chemical Physics Letters* **382**, 626–636 (2003).

Yoshitake Sakae and Yuko Okamoto, "Protein force-field parameters optimized with the Protein Data Bank. II. Comparisons of force fields by folding simulations of short peptides," *Journal of Theoretical and Computational Chemistry* **3**, 359–378 (2004).

## 4.1 Introduction

In Chapter 2, our optimization method [1] for the protein force-field parameters, which utilizes the protein structures in the Protein Data Bank (PDB) [2], was explained. In Chapter 3, we presented the results of our application of this method. We then applied this method to the optimizations of partial-charge and backbone torsion-energy parameters of five well-known force fields, namely, AMBER parm94 [3], AMBER parm96 [4], AMBER parm99 [5], CHARMM version 22 [6], and OPLS-AA [7]. In order to assess the validity of these optimized force fields, we have to perform folding simulations of small proteins. The force fields that yield the results which are in accord with experimental implications can be considered to be good ones, In this chapter, we present the results of the folding simulations of two small peptides with the above five original force fields and the corresponding optimized ones, and we try to judge whether any improvements are obtained.

The present Chapter is organized as follows. In Section 2 the computational details are described. In Section 3 the results of the simulations are given.

## 4.2 Computational details

Using the five original force fields as well as the optimized force fields that were obtained in Chapter 3, we tested their validity by applying them to the folding simulations of two peptides: 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 [14]. 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 $\alpha$-helix structure [15, 16]. Because the charges at peptide termini are known to affect helix stability [15, 16], we blocked the N-terminus and the C-terminus by a neutral $COCH_3$- group and a neutral -$NH_2$ group, respectively, except for CHARMM, where a neutral $COCH_3$- group and a neutral -$NHCH_3$ group were respectively used.

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 [17, 18, 14]. This peptide is known to form a $\beta$-hairpin structure by experiments [17, 18, 14].

Simulated annealing MD simulations of 1 ns were performed for both peptides from extended initial conformations. The unit time step was set to 1.0 fs (hence, each run consists of $N_S = 1,000,000$ MD steps). The temperature during MD simulations was controlled by Berendsen's method [19]. For each simulation the temperature was decreased exponentially from 2,000 K to 200 K (i.e., $T_I = 2,000$ K and $T_F = 200$ K) according to Eqs. (2.14) and (2.15). As for solvent effects, in order to carry out a quantitative test, we want to include accurate contributions such as explicit water molecules. Here, for simplicity, however, we used the generalized-Born/surface area (GB/SA) terms [20, 21] (see Eqs. (2.10)–(2.12)). For both peptides, these folding simulations were repeated 16 times with different initial conformations.

## 4.3 Results and discussion

We now present the results of the folding simulations (simulated annealing MD simulations of 1.0 ns) of two peptides, C-peptide and G-peptide. As for force fields, five well-known ones (namely, AMBER parm94 [3], AMBER parm96 [4], AMBER parm99 [5], CHARMM version 22 [6], and OPLS-AA [7]) and the corresponding optimized ones (see Tables 3.1–3.65) were used. In Figs. 4.1–4.10 the lowest-potential-energy conformations of C-peptide and G-peptide obtained from each of the 16 independent folding simulations are shown. They are numbered in increasing order of energy. In Figs. 4.11 and 4.12 we plot the number of conformations (out of 16 conformations in Figs. 4.1–4.10) with the secondary structures ($\alpha$-helix and $\beta$-hairpin) as functions of residue number in C-peptide and in G-peptide, respectively. These numbers represent the secondary-structure forming tendencies of each residue in the two peptides.

We now give the detailed discussions on these results. In Fig. 4.1(a) we see that all

conformations of C-peptide have $\alpha$-helix structures in the case of the original AMBER parm94. The root-mean-square deviations (RMSD; only $C_\alpha$ atoms are taken into account) from the corresponding part of the native X-ray structure of the entire ribonuclease A was the lowest for conformations Nos. 16 (2.37 Å), 12 (2.47 Å), and 5 (2.51 Å). For the optimized force field (see Fig. 4.1(b)), on the other hand, 12 out of 16 conformations are $\alpha$-helix structures, and $\beta$-hairpin structures also appeared in conformations Nos. 9, 14, and 16. The number of conformations with $\alpha$-helix structures increased by six compared with our previous results, and that with $\beta$-hairpin structures increased by one [1]. RMSD was the lowest for conformations Nos. 8 (1.06 Å), 4 (3.69 Å), and 12 (3.73 Å). Note that conformation No. 8 with the optimized parameters is in remarkable agreement with the native structure. The experiments imply only 30 % $\alpha$-helix formations around 0° C [15]. Hence, the results for the optimized AMBER parm94 are more consistent with the experiments than the original AMBER parm94, although the $\alpha$-helix-forming tendency is still very high.

In Fig. 4.2(a) we see that the results for G-peptide of the original AMBER parm94 are again all $\alpha$-helical. This is in contradiction with the experimental results [17, 18, 14], which imply the formation of a $\beta$-hairpin structure. RMSD from the corresponding part of the native X-ray structure of the entire protein G was the lowest for conformations Nos. 4 (5.92 Å), 9 (7.02 Å), and 6 (7.54 Å), which are very large. For the optimized AMBER parm94 (see Fig. 4.2(b)), two out of 16 conformations (Nos. 8 and 9) indeed yielded $\beta$-hairpin structures, and $\alpha$-helix structures also appeared in seven out of 16 conformations. The number of conformations with $\beta$-hairpin structures decreased by three as compared with our previous results, and that with $\alpha$-helix structures increased by five [1]. RMSD was the lowest for conformations Nos. 13 (3.01 Å), 8 (3.97 Å), and 16 (4.06 Å), which are better than those for the original AMBER parm94.

In Fig. 4.11(a1) and Fig. 4.12(a1) we find that the $\alpha$-helix-forming tendency of each residue for the original AMBER parm94 is about the same for C-peptide and G-peptide (almost entirely helical). The corresponding tendency for the optimized AMBER parm94 is less than that of the original AMBER parm94. The native C-peptide structure has $\alpha$-helix in residues from 4 to 12 [22], and the $\alpha$-helix-forming tendency for the optimized

AMBER parm94 is in accord with this experimental implication (see Fig. 4.11(a1)). As for G-peptide, the $\beta$-strand-forming tendency for the optimized AMBER parm94 is reasonable in the sense that residues 4–7 and 10–15 tend to form $\beta$-strands for the $\beta$-hairpin structure (see Fig. 4.11(a2)). The peak at residue 9 for the $\alpha$-helix-forming tendency for the optimized force field can be understood as the turn formations for the $\beta$-hairpin structure.

While the original AMBER parm94 favors $\alpha$-helix structures too much, the original AMBER parm96 tend to favor $\beta$-hairpin structures (see Figs. 4.3(a) and 4.4(a)). Six out of 16 conformations of C-peptide and five out of 16 conformations of G-peptide are all $\beta$-hairpin structures, and many of other conformations also have structures close to $\beta$-hairpin structures (but without the characteristic backbone hydrogen bonds). Only two conformations (Nos. 5 and 10) for C-peptide and one conformation (No. 11) for G-peptide are $\alpha$-helix structures. RMSD was the lowest for conformations Nos. 10 (3.37 Å), 5 (4.49 Å), and 12 (4.75 Å) for C-peptide and for conformations Nos. 4 (3.76 Å), 13 (3.76 Å), and 3 (4.29 Å) for G-peptide.

As for the optimized AMBER parm96, some improvement is achieved in the sense that more $\alpha$-helix structures are obtained (see Figs. 4.3(b) and 4.4(b)). Ten out of 16 conformations of C-peptide and seven out of 16 conformations of G-peptide now have $\alpha$-helix structures. Four out of 16 conformations of C-peptide and five out of 16 conformations of G-peptide have $\beta$-hairpin structures. RMSD was the lowest for conformations Nos. 10 (2.33 Å), 1 (2.92 Å), and 9 (3.50 Å) for C-peptide and for conformations Nos. 9 (3.11 Å), 1 (3.68 Å), and 4 (4.45 Å) for G-peptide, which are better than the original AMBER parm96 case.

The secondary-structure-forming tendencies as functions of residue number for the AMBER parm96 and optimized parameters are compared in Figs. 4.11(b) and 4.12(b). For C-peptide, which forms an $\alpha$-helix structure in experiments, we find an increase in $\alpha$-helix-forming tendency and a decrease in $\beta$-strand-forming tendency for the optimized parameters with respect to the original AMBER parm96. This is an improvement. For G-peptide, which forms a $\beta$-hairpin structure in experiments, however, we see almost no change in the $\beta$-strand-forming tendency, while there is an increase in $\alpha$-helix-forming

tendency. Hence, the optimized parameters seem to from $\alpha$-helix a little too much.

The original AMBER parm99 is similar to the original AMBER parm94 in that it favors $\alpha$-helix structures too much (compare Figs. 4.1(a) and 4.2(a) and Figs. 4.5(a) and 4.6(a)). The original AMBER parm94 is a little more $\alpha$-helix forming than the original AMBER parm99, because all the 32 conformations are $\alpha$-helix structures for the former (see Figs. 4.1(a) and 4.2(a)), while 30 out of 32 conformations are $\alpha$-helix structures for the latter (see Figs. 4.5(a) and 4.6(a)), and because more extended $\alpha$-helices are obtained for the former than for the latter. For C-peptide, RMSD was the lowest for conformations Nos. 7 (2.58 Å), 6 (2.67 Å), and 5 (2.81 Å) with the original AMBER parm99. As expected, RMSD for G-peptide is not as good; it is the lowest for conformations Nos. 12 (4.62 Å), 11 (5.08 Å), and 16 (5.32 Å).

As for the optimized AMBER parm99, an improvement is achieved because much less number of helical conformations were obtained (see Figs. 4.5(b) and 4.6(b)). Four out of 16 conformations of C-peptide are $\alpha$-helix structures, which is more consistent with the experimental results than with the original parameters, and $\beta$-hairpin structures also appeared in six out of 16 conformations. For G-peptide, on the other hand, three out of 16 conformations are $\beta$-hairpin structures, and no $\alpha$-helix structures were found. RMSD was the lowest for conformations Nos. 9 (3.68 Å), 6 (4.04 Å), and 15 (4.83 Å) for C-peptide, and for conformations Nos. 2 (2.22 Å), 3 (4.37 Å), and 4 (5.18 Å). Hence, the optimizedAMBER parm99 gives good agreement with the experimental results especially for G-peptide. However, it favors $\beta$-hairpin structures a little too much (see the results for C-peptide in Fig. 4.5(b) again).

The secondary-structure-forming tendencies as functions of residue number are shown in Figs. 4.11(c) and 4.12(c). We see that the conformations obtained with the original AMBER parm99 tend to favor $\alpha$-helix too much, and those with the optimized AMBER parm99 are more consistent with the experimental results, although the latter seems to prefer $\beta$-hairpin structures. For C-peptide, however, the results with the original AMBER parm99 are more consistent with the experimental results than the optimized AMBER parm99; especially $\alpha$-helix is favored in residues 4–12. For G-peptide, the results with the optimized AMBER parm99 are more consistent with the experimental results than the

original.

We have another optimized parameters for AMBER parm99 (see Tables 3.66 and 3.67). The results for this case are presented later below.

The case for the original CHARMM version 22 is less "extreme" than those for the three original AMBER parameters as far as the secondary-structure formations are concerned; there are less number os conformations with explicit secondary structures (compare Figs. 4.7(a) and 4.8(a) with Figs. 4.1(a), 4.2(a), 4.3(a), 4.4(a), 4.5(a), 4.6(a)). Only three out of 16 conformations for both C-peptide and G-peptide are $\alpha$-helix structures, whereas three out of 16 conformatios of C-peptide and one out of 16 conformations of G-peptide are $\beta$-hairpin structures (see Figs. 4.7(a) and 4.8(a)). RMSD was the lowest for conformations Nos. 2 (2.98 Å), 6 (3.45 Å), and 12 (3.70 Å) for C-peptide and for conformations Nos. 3 (5.48 Å), 2 (5.66 Å), and 10 (5.92 Å) for G-peptide. Hence, the original CHARMM version 22 better represent the $\alpha$-helix-forming peptide (C-peptide) than $\beta$-hairpin-forming peptide (G-peptide).

The optimized CHARMM version 22 is an improvement over the original parameters in the sense that we get more $\alpha$-helix structures for C-peptide and more $\beta$-hairpin structures for G-peptide (see Figs. 4.7(b) and 4.8(b)). The number of conformations with $\alpha$-helix structures increased by five for C-peptide, and those with $\beta$-hairpin structures increase by one for G-peptide with respect to the original CHARMM version 22. RMSD was the lowest for conformations Nos. 9 (2.28 Å), 1 (2.59 Å), and 6 (3.08 Å) for C-peptide and for conformations Nos. 13 (4.86 Å), 2 (5.32 Å), and 9 (6.00 Å) for G-peptide, and these are better than those of the original CHARMM version 22 above.

The secondary-structure-forming tendencies as functions of residue number for the CHARMM version 22 and optimized parameters are compared in Figs. 4.11(d) and 4.12(d). For C-peptide we find an increase in $\alpha$-helix-forming tendency and a decrease in $\beta$-strand-forming tendency for the optimized parameters with respect to the original CHARMM version 22, which is consistent with the experimental results. For G-peptide we see an increase in $\beta$-strand-forming tendency in the $\beta$-starnd regions and a concentration of $\alpha$-helix-forming tendency at the turn region for the optimized parameters compared to the original ones, while is an improvement. However, the optimized parameters seem to

from $\alpha$-helix a little too much.

We have another optimized parameters for CHARMM version 22 (see Tables 3.66 and 3.67). The results for this case are presented later below.

The last of the five force fields that we considered is OPLS-AA. The original OPLS-AA seems to favor $\beta$-hairpin structures, but this tendency is a little less than the original AMBER parm96 case (compare Figs. 4.9(a) and 4.10(a) with Figs. 4.3(a) and 4.4(a)). Especially, the lowest-energy conformation (conformation No. 1) among 16 is a $\beta$-hairpin structure for both C-peptide and G-peptide. However, the total number of conformations with $\alpha$-helix structures is more than that with $\beta$-hairpin structures (five versus three) and more extended $\alpha$-helix structures are found for C-peptide, while extended $\beta$-hairpin structures are obtained and $\alpha$-helices are short for G-peptide. RMSD was the lowest for conformations Nos. 6 (3.38 Å), 3 (3.73 Å), and 13 (4.12 Å) for C-peptide and for conformations Nos. 13 (3.82 Å), 2 (4.46 Å), and 5 (5.01 Å) for G-peptide. Hence, the original OPLS-AA is perhaps the best among the five original force fields that were considered in the present work, when we consider proteins with both $\alpha$-helix and $\beta$-sheet structures. When one considers proteins with only $\alpha$-helix structures, however, the original AMBER parm99 and CHARMM version 22 are better than others and when one considers proteins with only $\beta$-sheet structures, the original AMBER parm96 and OPLS-AA seem to be better than others.

We remark that quite similar results were obtained in a recent work [8], in which detailed comparisons of three versions of AMBER (parm94, parm96, and parm99), CHARMM version 22, OPLS-AA/L [9], and GROMOS [10] were carried out by generalized-ensemble simulations [11, 12] of the same two peptides in explicit water. The present work is based on an implicit solvent contributions, namely GB/SA, instead of explicit water. It is interesting to see that we still get good agreement on the simulation results.

The optimized OPLS-AA gives similar results to the original one (see Figs. 4.9(b) and 4.10(b)). It is improved (especially in G-peptide) in the sense that better RMSD values are obtained for the optimized parameters than the original ones. RMSD was the lowest for conformations Nos. 6 (3.13 Å), 1 (4.11 Å), and 16 (4.15 Å) for C-peptide and for conformations Nos. 7 (3.04 Å), 5 (3.51 Å), and 3 (4.33 Å) for G-peptide.

The secondary-structure-forming tendencies as functions of residue number for the OPLS-AA and optimized parameters are compared in Figs. 4.11(e) and 4.12(e). The above-mentioned improvement in G-peptide is now clearly seen in Fig. 4.12(e); we see an increase in $\beta$-strand-forming tendency in the $\beta$-starnd regions and an increase in $\alpha$-helix-forming tendency at the turn region for the optimized parameters compared to the original ones,

We now consider three additionally optimized force fields (AMBER parm99, CHARMM version 22, and AMBER parm94). These optimized parameters (backbone torision-energy parameters) are summarized in Tables 3.66 and 3.67. We want to compare the results of the folding simulations of the two peptides based on these new parameters with those of the previously optimized ones described above.

In Fig. 4.13 and Fig. 4.14, the lowest-potential-energy conformations that were obtained from each of the 16 independent folding simulations are shown for C-peptide and G-peptide, respectively, for the newly optimized AMBER parm99 (these should be compared with those in Figs. 4.5 and 4.6). The corresponding results for the newly optimized CHARMM version 22 are shown in Figs. 4.15 and 4.16 (these should be compared with those in Figs. 4.7 and 4.8).

In Figs. 4.5 and 4.6, we found that $\alpha$-helix structures are favored too much for the original AMBER parm99, while the previously optimized parameters favor $\beta$-hairpin structures a little too much. The new results lie in between these two cases (see Figs. 4.13 and 4.14). The number of $\alpha$-helix conformations (out of 16) is seven for C-peptide and five for G-peptide (these should be compared with 14 for C-peptide and 16 for G-peptide in the original AMBER parm99 and four for C-peptide and zero for G-peptide in the previously optimized AMBER parm99, respectively). Likewise, the number of $\beta$-hairpin conformations is four for both C-peptide and G-peptide (the corresponding numbers are zero for both C-peptide and G-peptide in the original AMBER parm99 and six for C-peptide and three for G-peptide in the previously optimized AMBER parm99). RMSD was the lowest for conformations Nos. 4 (2.72 Å), 1 (2.86 Å), and 3 (3.35 Å) for C-peptide and for conformations Nos. 2 (2.92 Å), 16 (4.96 Å), and 8 (5.07 Å) for G-peptide, which are reasonable.

The secondary-structure-forming tendencies as functions of residue number for the newly optimized AMBER parm99 are shown in Fig. 4.19(a) for C-peptide and Fig. 4.20(a) for G-peptide. These are in accord with the experimental implications except for the unexpected decrease in $\alpha$-helix propensity around the 7-th amino acid for C-peptide (see Fig. 4.19(a1)).

In Figs. 4.7(b) and 4.8(b), we found that $\alpha$-helix structures are favored a little too much for the previously optimized parameters. As for the newly optimized CHARMM version 22, this tendency is rectified (see Figs. 4.15 and 4.16). The number of $\alpha$-helix conformations (out of 16) is seven for C-peptide and zero for G-peptide (these should be compared with eight for C-peptide and four for G-peptide in the previously optimized CHARMM version 22). Likewise, the number of $\beta$-hairpin conformations is three for C-peptide and one for G-peptide (the corresponding numbers are two for both C-peptide and G-peptide in the previously optimized CHARMM version 22). RMSD was the lowest for conformations Nos. 15 (2.73 Å), 2 (2.75 Å), and 4 (3.01 Å) for C-peptide and for conformations Nos. 2 (3.37 Å), 6 (5.43 Å), and 14 (5.69 Å) for G-peptide, which are again reasonable.

The secondary-structure-forming tendencies as functions of residue number for the newly optimized CHARMM version 22 are shwn in Fig. 4.19(b) for C-peptide and Fig. 4.20(b) for G-peptide. These are consistent with the experimental results.

In Figs. 4.17 and 4.18, the simulation results using the optimized AMBER parm94 obtaind from the another optimization method with no constraint of the optimization of torsion-energy parameters are shown. As compared with the previously optimized parameters, the $\beta$-hairpin structures are favored. For C-peptide, there are more conformations of $\alpha$-helix than $\beta$-hairpin. For G-peptide, there are more conformations of $\beta$-hairpin than $\alpha$-helix. The secondary-structure-forming tendencies as functions of residue number for the newly optimized AMBER parm94 are shown in Fig. 4.19(c) for C-peptide and Fig. 4.20(c) for G-peptide. These are consistent with the experimental results of secondary-structure-forming tendencies of the two peptides.

# Bibliography


[1] Y. Sakae and Y. Okamoto, *Chem. Phys. Lett.* **382**, 626 (2003).

[2] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, Nucleic Acids Research 28 (2000) 235.

[3] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, J. Am. Chem. Soc. 117 (1995) 5179.

[4] P.A. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, and A. Pohorille, in: A. Wilkinson, P. Weiner, and W.F. van Gunsteren, Elsevier, (Eds.), Computer Simulation of Biomolecular Systems, 3, Elsevier, 1997, p. 83.

[5] J. Wang, P. Cieplak, P.A. Kollman, J. Comput. Chem. 21 (2000) 1049

[6] A.D. MacKerrell, 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. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, J. Phys. Chem. B 102 (1998) 3586.

[7] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, J. Am. Chem. Soc. 118 (1996) 11225.

[8] T. Yoda, Y. Sugita, Y. Okamoto, Chem. Phys. Lett. 386 (2004) 460.

[9] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, J. Phys. Chem. B 105 (2001) 6474.

[10] W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi (1996) Biomolecular simulation: the GROMOS96 manual and user guide. Zürich: Vdf Hochschulverlag AG an der ETH Zürich.

[11, 12]

[11] Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **329**, 261 (2000).

[12] A. Mitsutake, Y. Sugita, and Y. Okamoto, *J. Chem. Phys.* **118**, 6664 (2003); *ibid.* **118**, 6676 (2003).

[13] S. Kirkpatrick, C.D. Gelatt Jr., M.P. Vecchi, Science 220 (1983) 671.

[14] S. Honda, N. Kobayashi, E. Munekata, J. Mol. Biol. 295 (2000) 269.

[15] K.R. Shoemaker, P.S. Kim, D.N. Brems, S. Marqusee, E.J. York, I.M. Chaiken, J.M. Stewart, R.L. Baldwin, Proc. Natl. Acad. Sci. U.S.A. 82 (1985) 2349.

[16] J.J. Osterhout Jr., R.L. Baldwin, E.J. York, J.M. Stewart, H.J. Dyson, P.E. Wright, Biochemistry 28 (1989) 7059.

[17] F.J. Blanco, G. Rivas, L. Serrano, Nature Struct. Biol. 1 (1994) 584.

[18] N. Kobayashi, S. Honda, H. Yoshii, H. Uedaira, E. Munekata, FEBS Lett. 366 (1995) 99.

[19] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, J. Chem. Phys. 81 (1984) 3684.

[20] W.C. Still, A. Tempczyk, R.C. Hawley, T. Hendrickson, J. Am. Chem. Soc. 112 (1990) 6127.

[21] D. Qiu, P.S. Shenkin, F.P. Hollinger, W.C. Still, J. Phys. Chem. A 101 (1997) 3005.

[22] R. F. Tilton, J.C. Dewan, G.A. Petsko, Biochemistry 31 (1992) 2469.

[23] R.A. Sayle, E.J. Milner-White, Trends Biochem. Sci. 20 (1995) 374.

[24] W. Kabsch and C. Sander, Biopolymers 22 (1983) 2577.

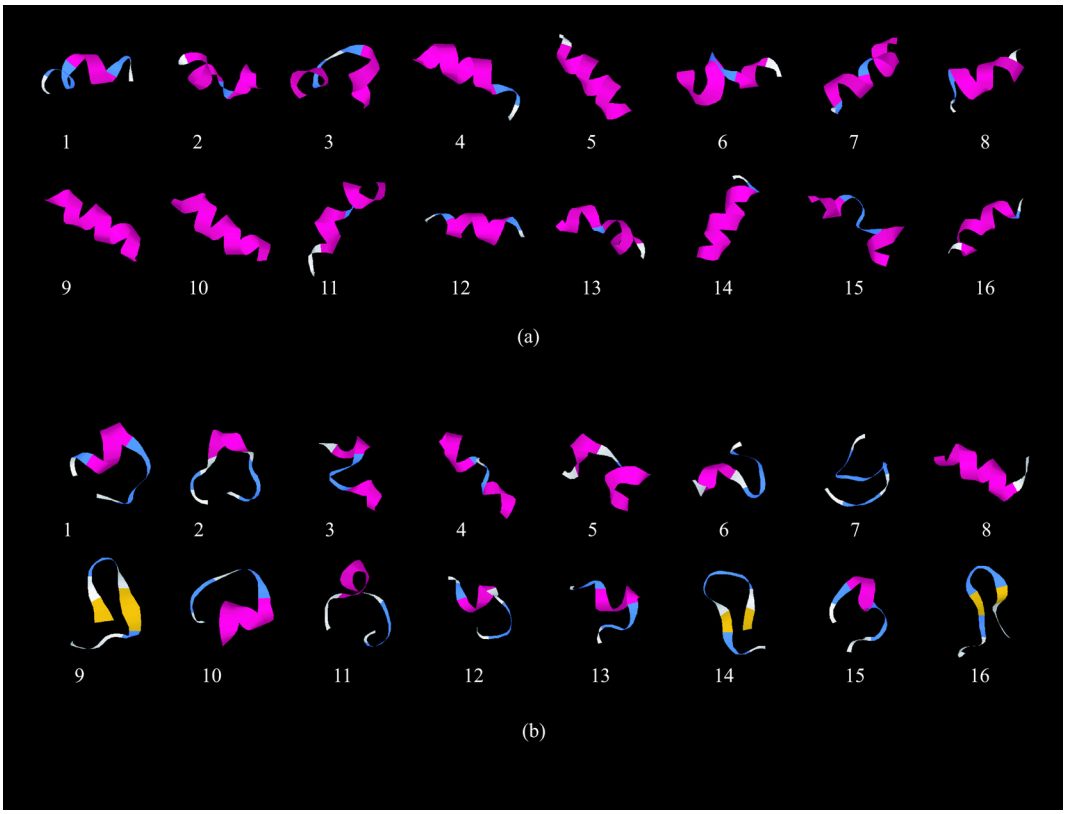


Figure 4.1: The lowest-energy conformations of C-peptide obtained by each of the simulated annealing MD simulations using the original AMBER parm94 (a) and the optimized force field (b). The conformations are ordered in the increasing order of energy. See Tables 3.1–3.65 for the values of the optimized parameters. The figures were created with RasMol [23].

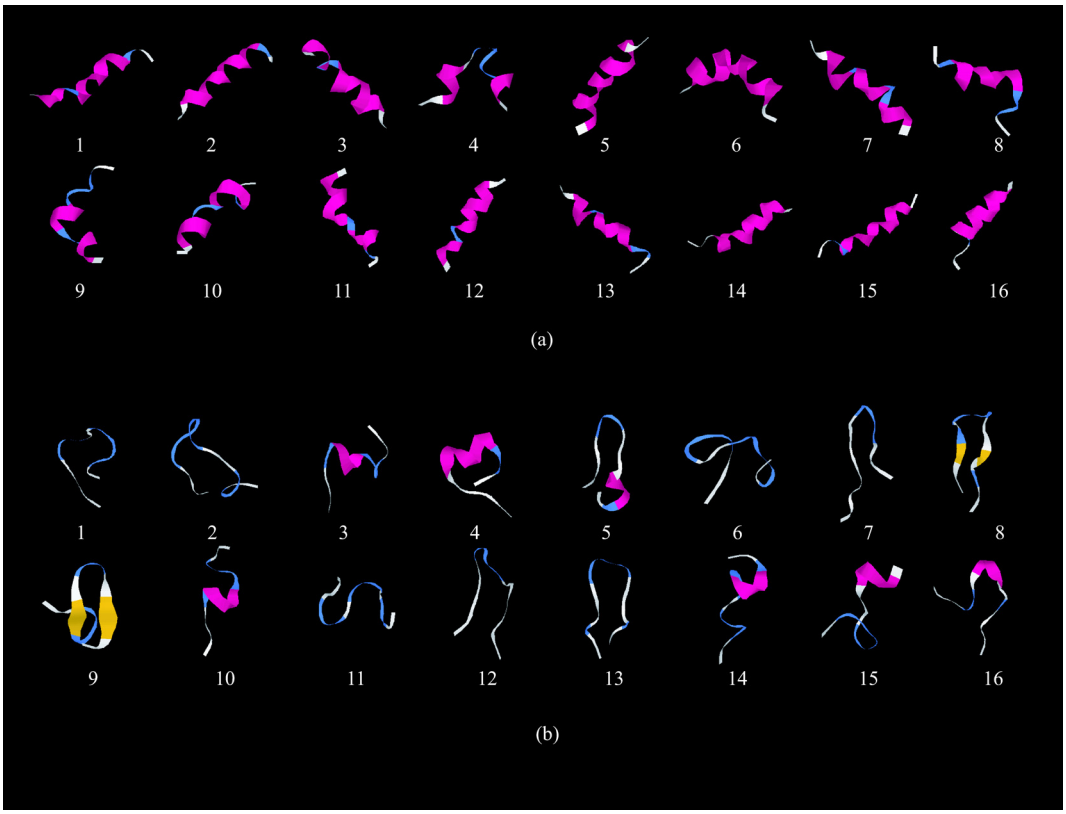


Figure 4.2: The lowest-energy conformations of G-peptide obtained by each of the simulated annealing MD simulations using the original AMBER parm94 (a) and the optimized force field (b). See the caption in Fig. 4.1.

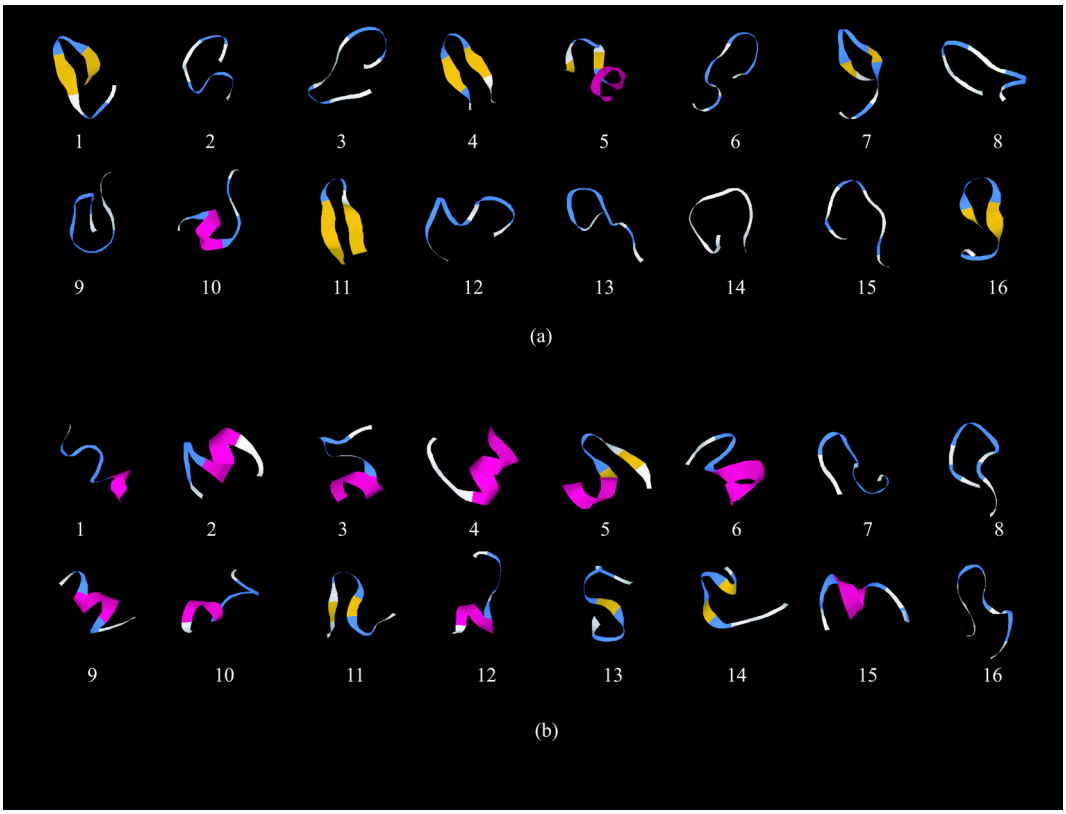


Figure 4.3: Figure 4.3. The lowest-energy conformations of C-peptide obtained by each of the simulated annealing MD simulations using the original AMBER parm96 (a) and the optimized force field (b). See the caption in Fig. 4.1.

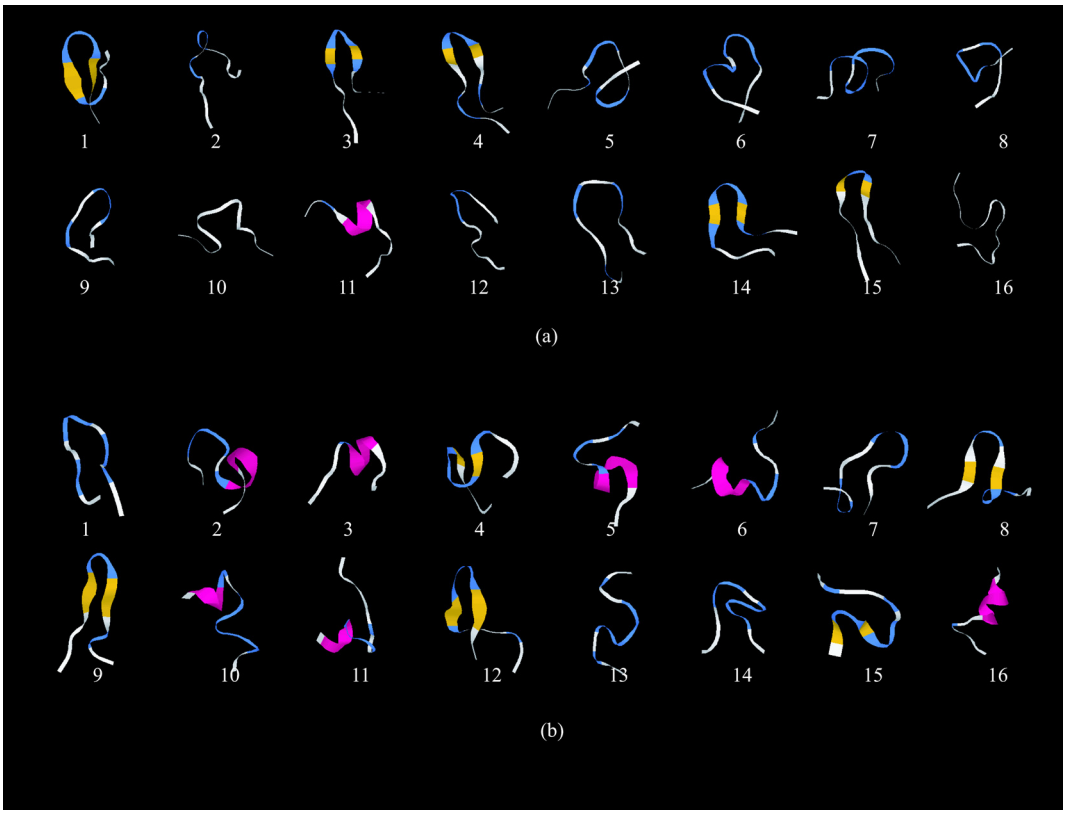


Figure 4.4: The lowest-energy conformations of G-peptide obtained by each of the simulated annealing MD simulations using the original AMBER parm96 (a) and the optimized force field (b). See the caption in Fig. 4.1.

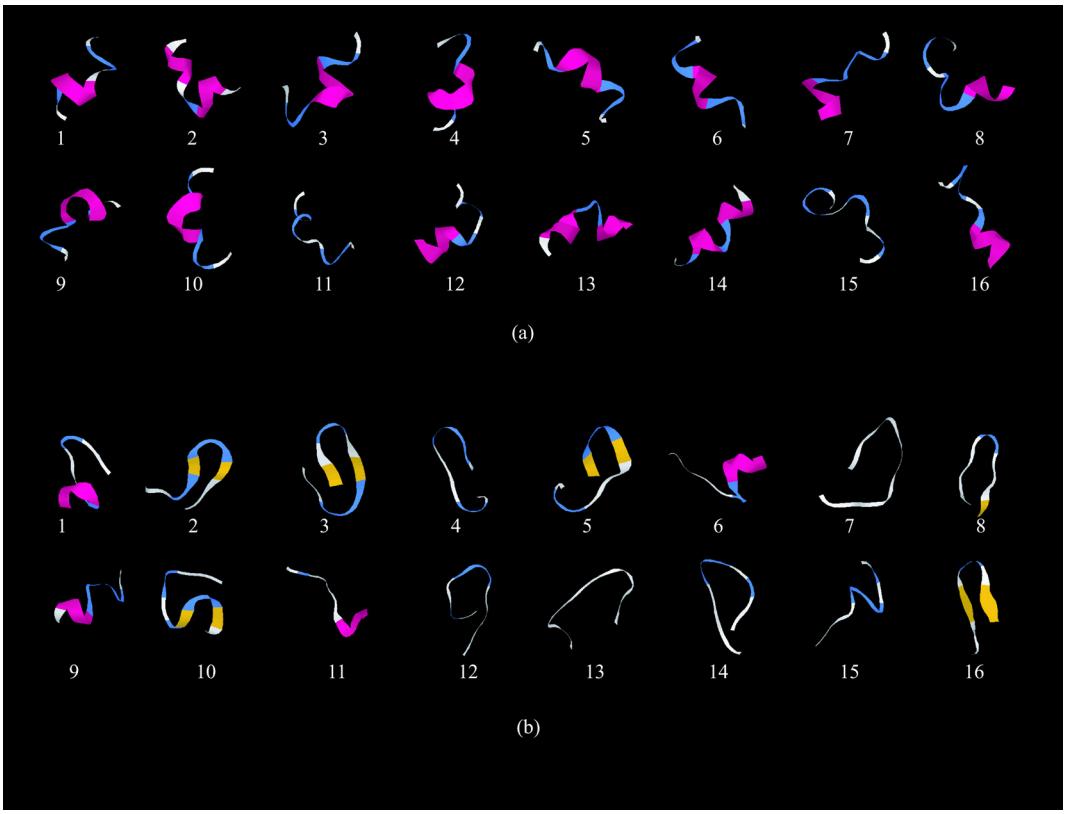


Figure 4.5: The lowest-energy conformations of C-peptide obtained by each of the simulated annealing MD simulations using the original AMBER parm99 (a) and the optimized force field (b). See the caption in Fig. 4.1.

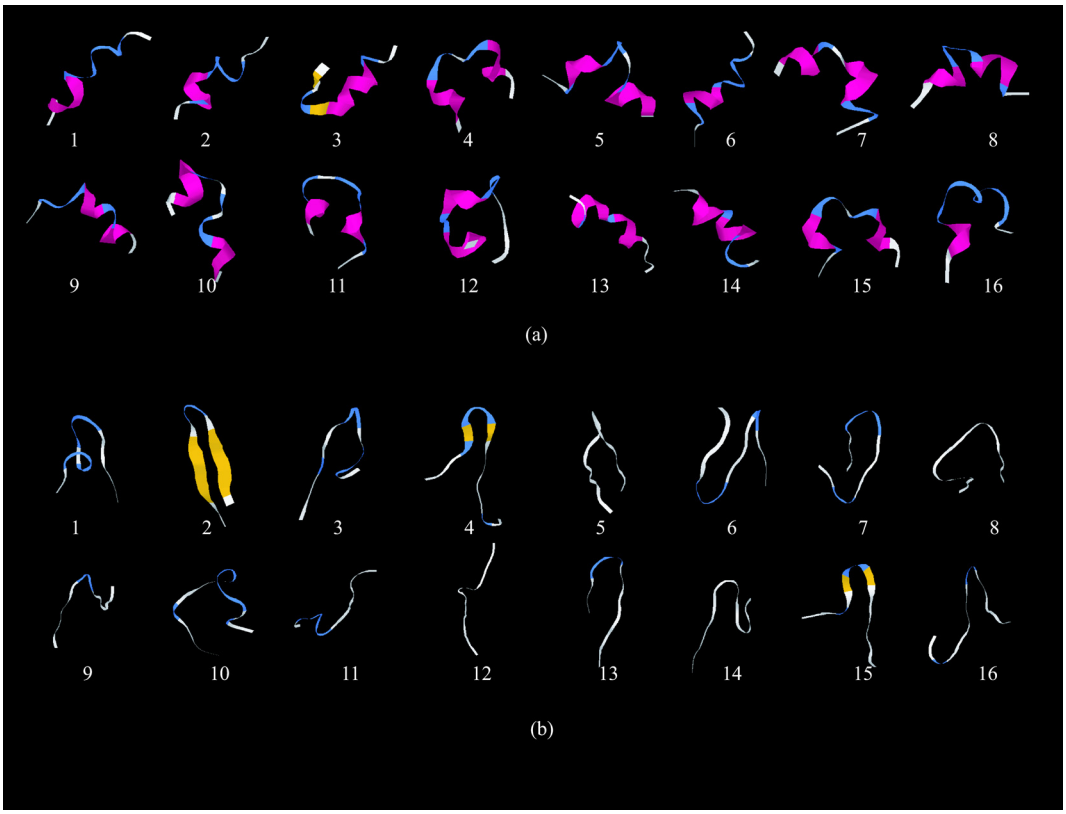


Figure 4.6: The lowest-energy conformations of G-peptide obtained by each of the simulated annealing MD simulations using the original AMBER parm99 (a) and the optimized force field (b). See the caption in Fig. 4.1.

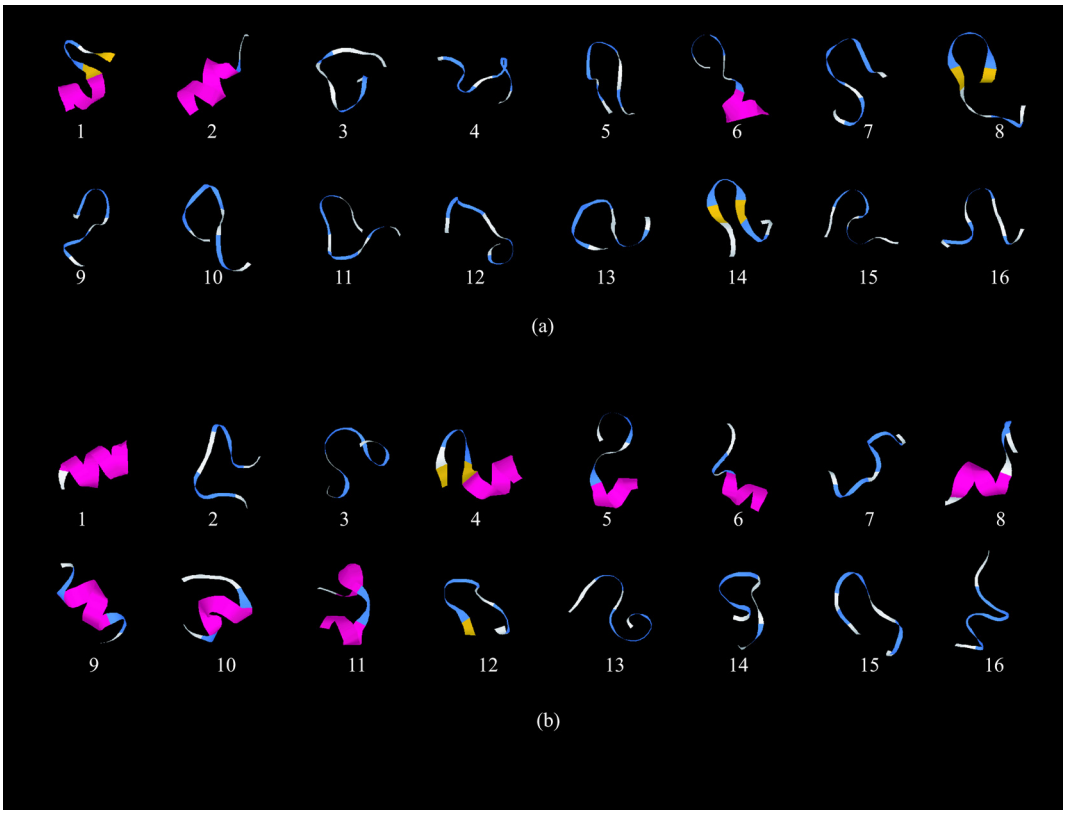


Figure 4.7: The lowest-energy conformations of C-peptide obtained by each of the simulated annealing MD simulations using the original CHARMM version 22 (a) and the optimized force field (b). See the caption in Fig. 4.1.

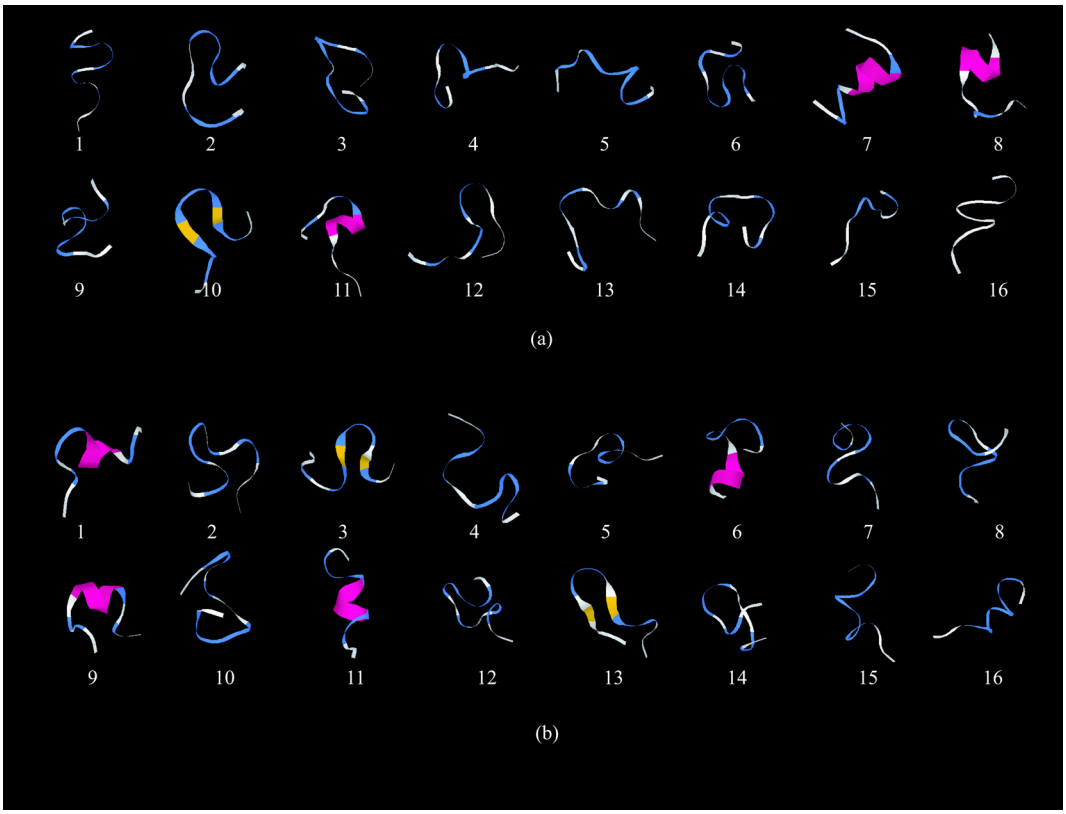


Figure 4.8: The lowest-energy conformations of G-peptide obtained by each of the simulated annealing MD simulations using the original CHARMM version 22 (a) and the optimized force field (b). See the caption in Fig. 4.1.

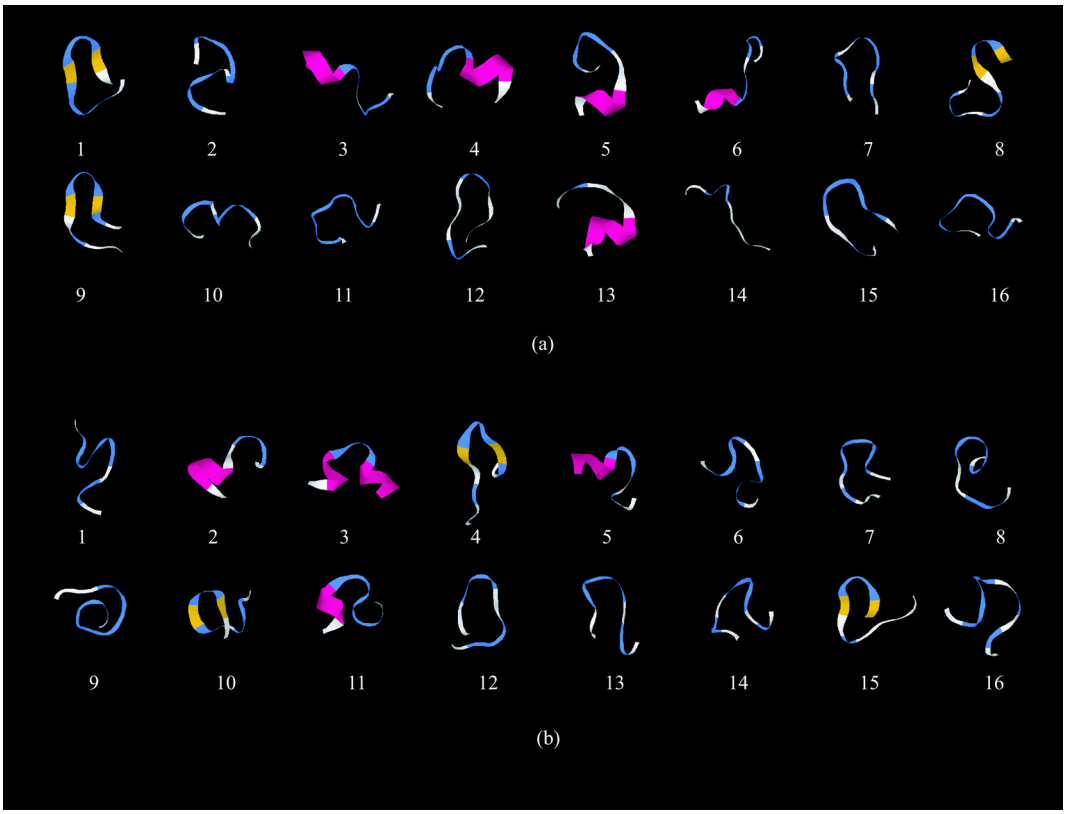


Figure 4.9: The lowest-energy conformations of C-peptide obtained by each of the simulated annealing MD simulations using the original OPLS-AA (a) and the optimized force field (b). See the caption in Fig. 4.1.

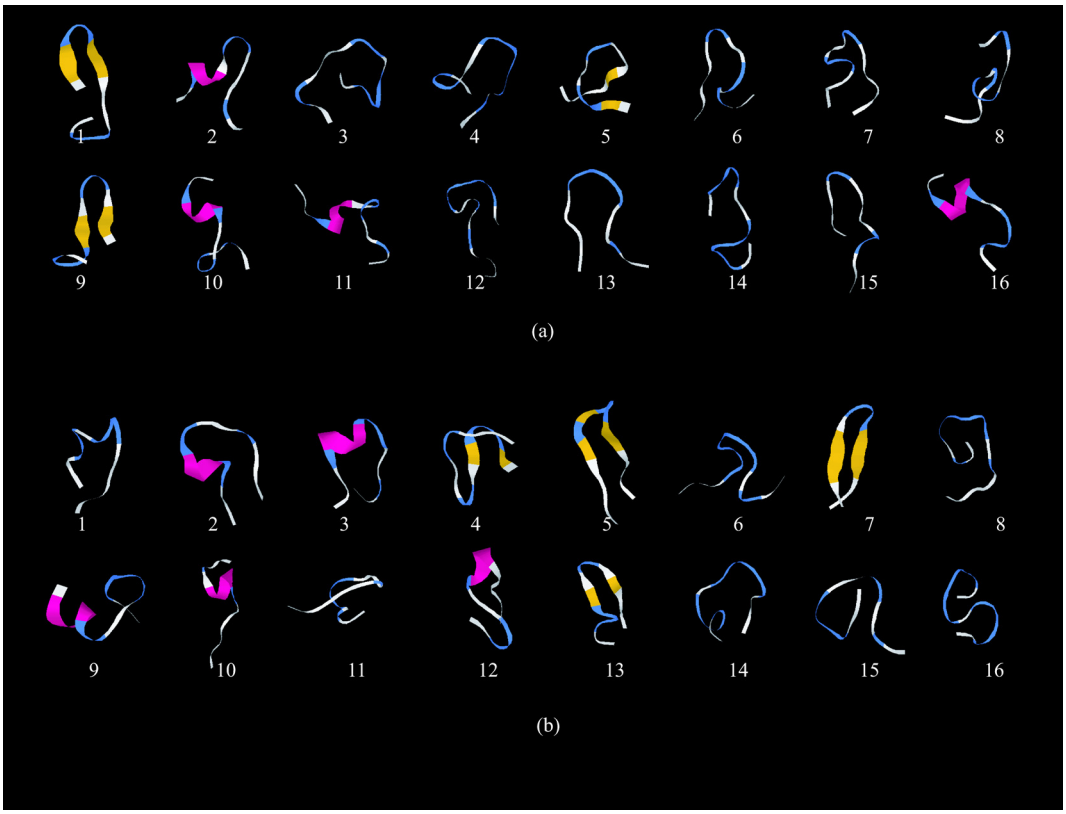


Figure 4.10: The lowest-energy conformations of G-peptide obtained by each of the simulated annealing MD simulations using the original OPLS-AA (a) and the optimized force field (b). See the caption in Fig. 4.1.

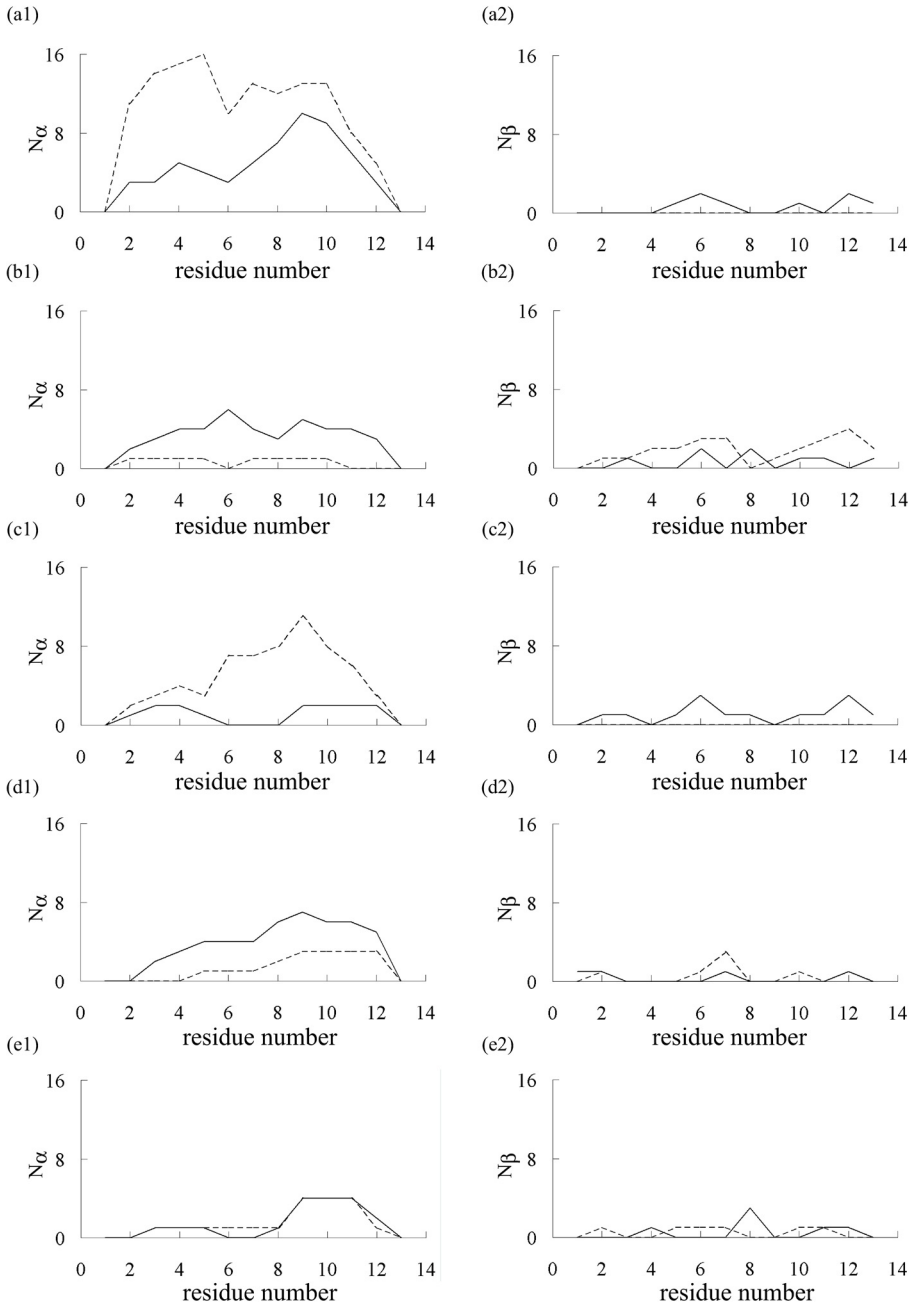


Figure 4.11: Number of conformations (out of 16 conformations in Figs. 4.1, 4.3, 4.5, 4.7, and 4.9) with the secondary structures as functions of residue number in C-peptide. $N_\alpha$ and $N_\beta$ stand for the numbers of conformations with $\alpha$-helix structures and $\beta$-hairpin structures, respectively. The criteria for the secondary-structure formations are those of DSSP [24]. (a), (b), (c), (d), and (e) correspond to the original force fields AMBER parm94, AMBER parm96, AMBER parm99, CHARMM version 22, and OPLS-AA (dotted curves), respectively, and the corresponding optimized force fields (solid curves).

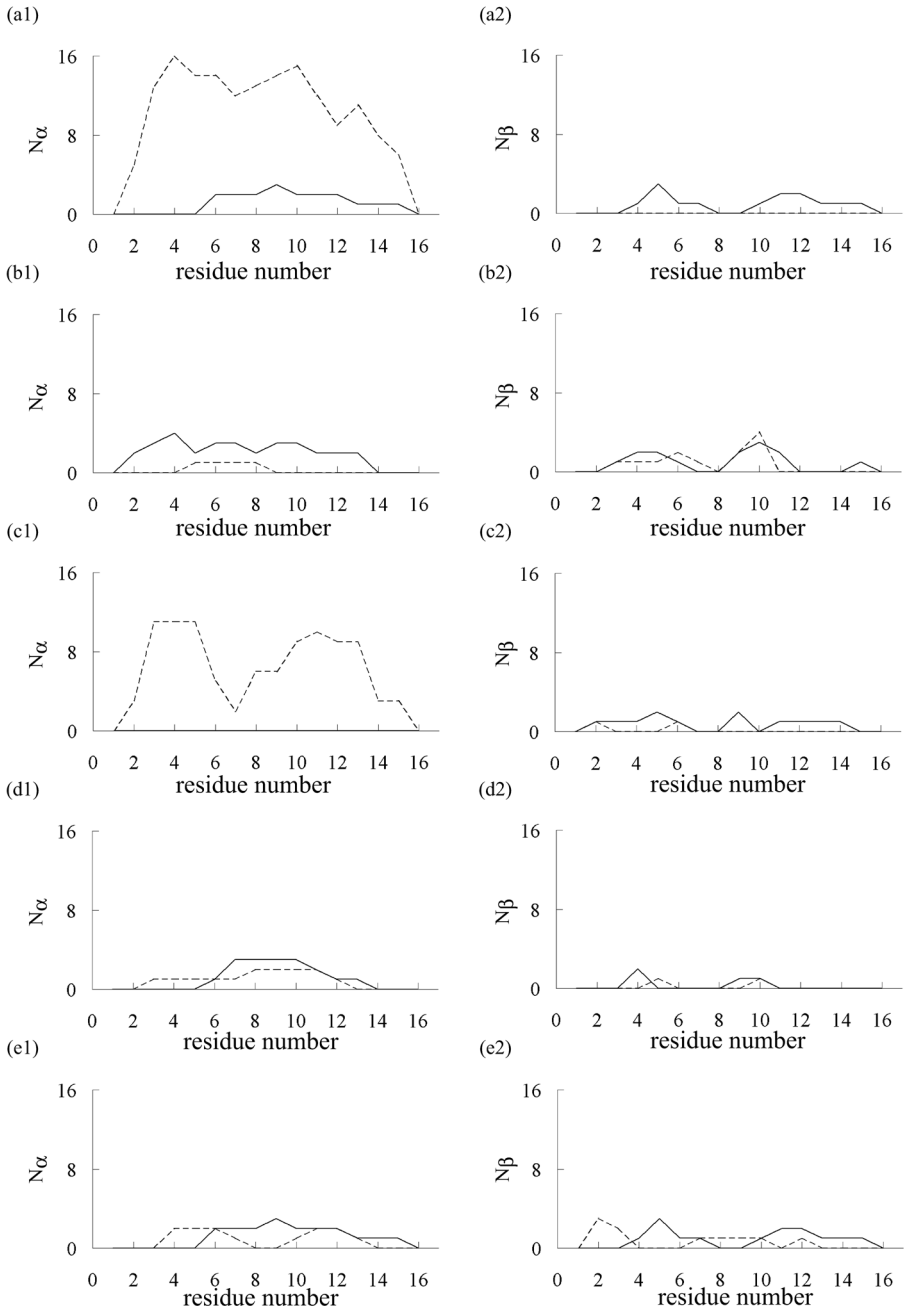


Figure 4.12: Number of conformations (out of 16 conformations in Figs. 4.2, 4.4, 4.6, 4.8, and 4.10) with the secondary structures as functions of residue number in G-peptide. See the caption in Fig. 4.11.

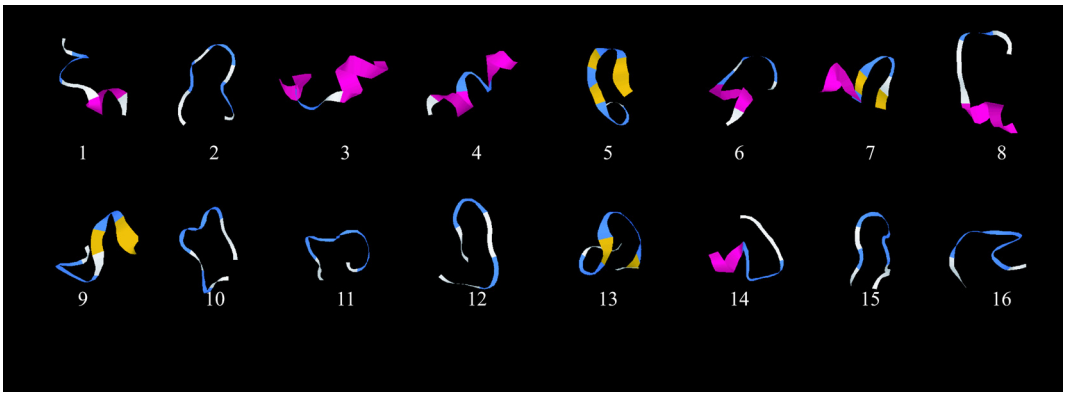


Figure 4.13: The lowest-energy conformations of C-peptide obtained by each of the simulated annealing MD simulations using another set of optimized parameters for AMBER parm99. The conformations are ordered in the increasing order of energy. See Tables 3.1–3.63, and Table 3.66 and 3.67 for the values of the optimized parameters. The figures were created with RasMol [23].

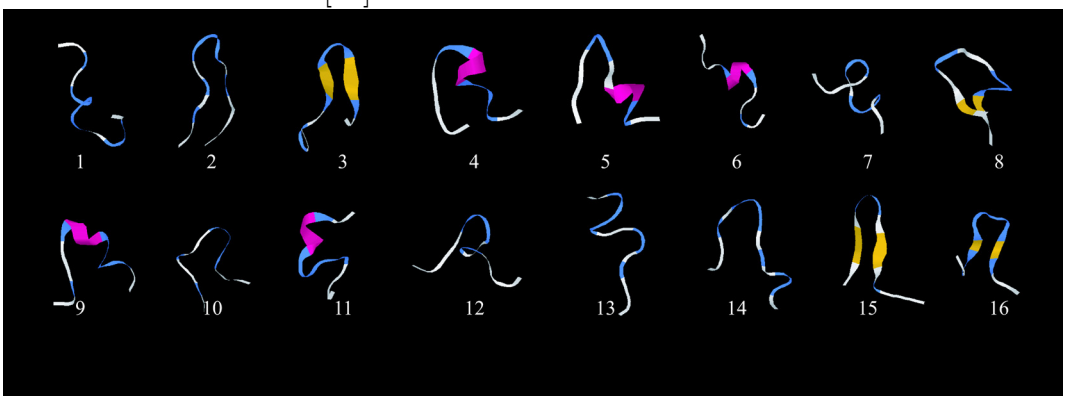


Figure 4.14: The lowest-energy conformations of G-peptide obtained by each of the simulated annealing MD simulations using another set of optimized parameters for AMBER parm99. See the caption in Fig. 4.13.

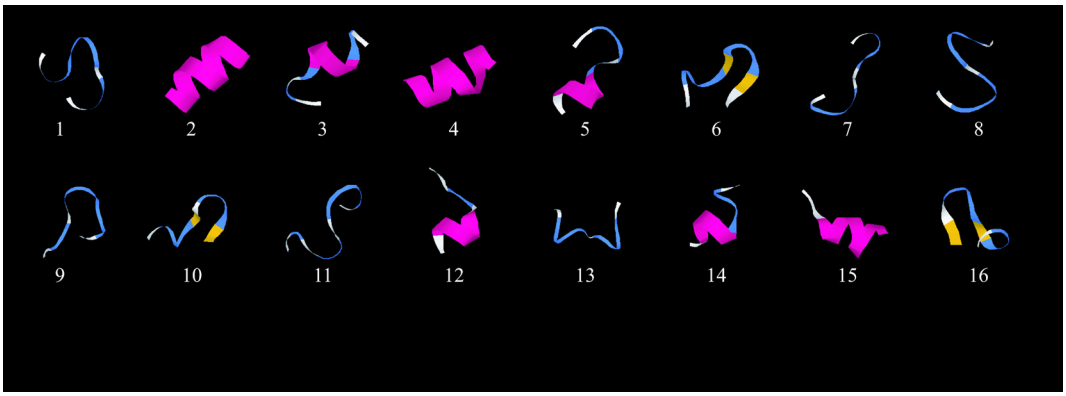


Figure 4.15: The lowest-energy conformations of C-peptide obtained by each of the simulated annealing MD simulations using another set of optimized parameters for CHARMM version 22. See the caption in Fig. 4.13.

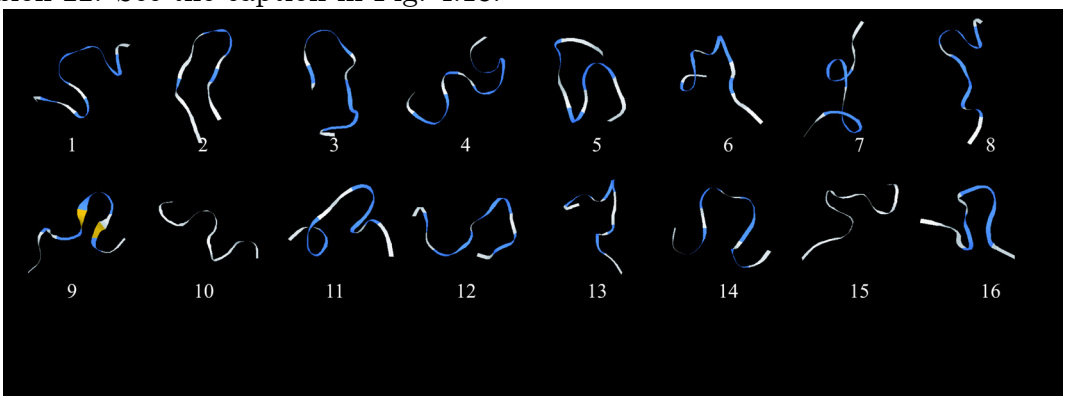


Figure 4.16: The lowest-energy conformations of G-peptide obtained by each of the simulated annealing MD simulations using another set of optimized parameters for CHARMM version 22. See the caption in Fig. 4.13.

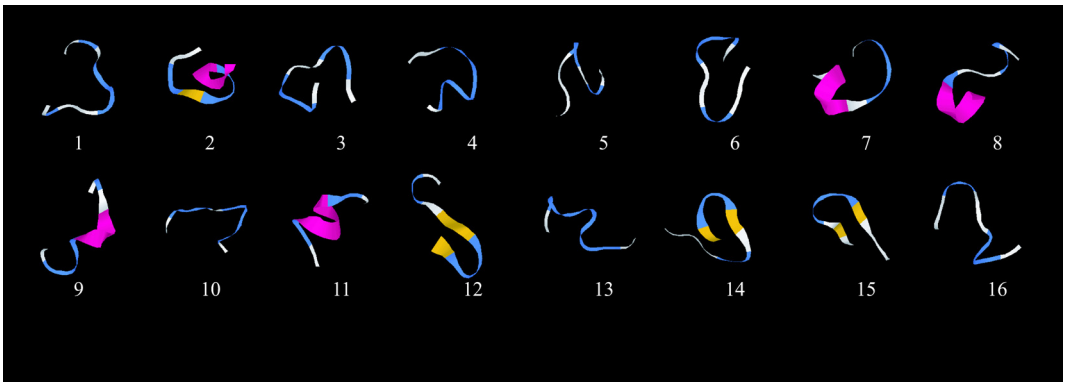


Figure 4.17: The lowest-energy conformations of C-peptide obtained by each of the simulated annealing MD simulations using another set of optimized parameters for AMBER parm94. See the caption in Fig. 4.13.

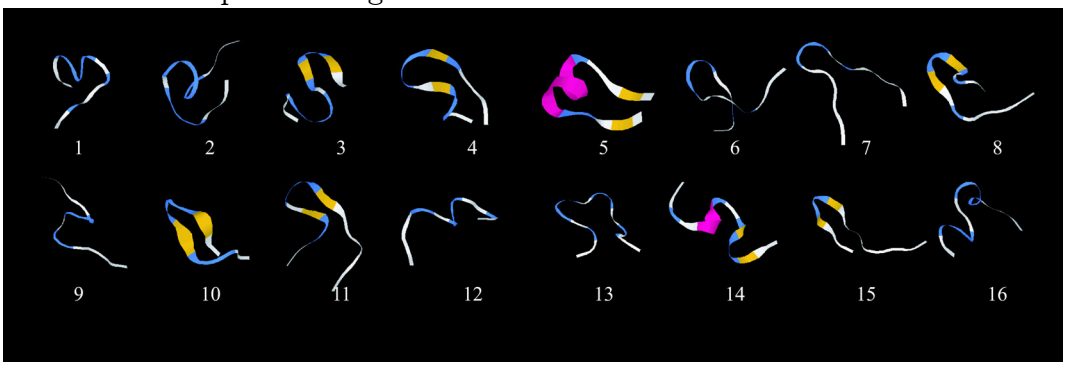


Figure 4.18: The lowest-energy conformations of G-peptide obtained by each of the simulated annealing MD simulations using another set of optimized parameters for AMBER parm94. See the caption in Fig. 4.13.

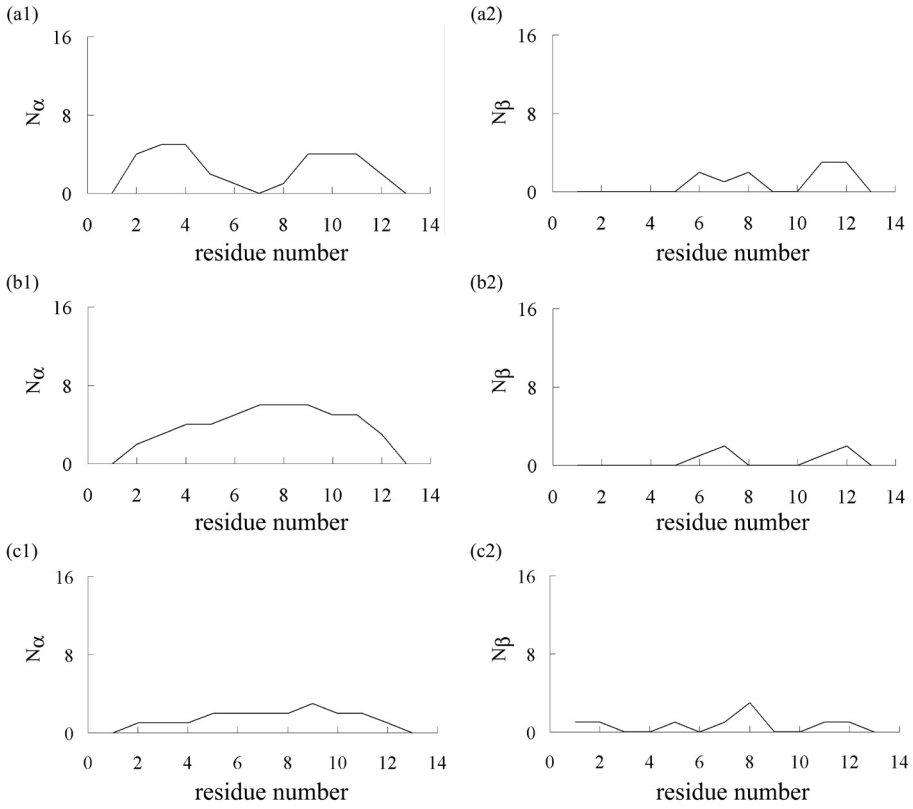


Figure 4.19: Number of conformations (out of 16 conformations in Figs. 4.13 and 4.15) with the secondary structures as functions of residue number in C-peptide. (a) and (b) correspond to second sets of optimized parameters for AMBER parm99, CHARMM version 22, and AMBER parm94, respectively. See the caption in Fig. 4.11.

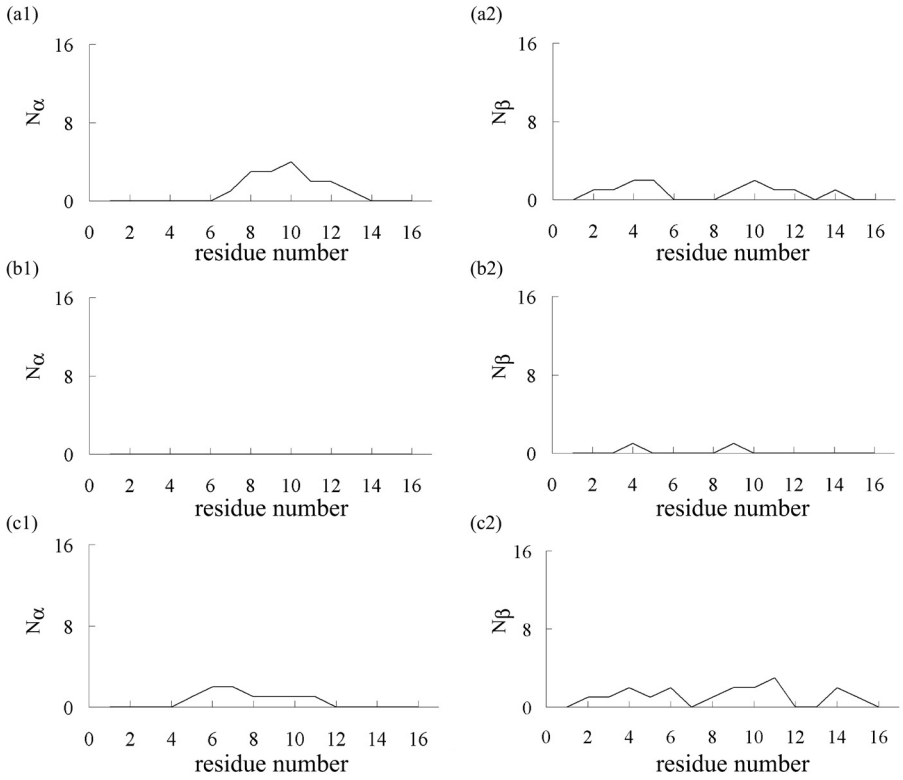


Figure 4.20: Number of conformations (out of 16 conformations in Figs. 4.14 and 4.16) with the secondary structures as functions of residue number in G-peptide. (a) and (b) correspond to second sets of optimized parameters for AMBER parm99, CHARMM version 22, and AMBER parm94, respectively. See the caption in Fig. 4.11.

# Chapter 5

# Conclusions

In this Thesis, we have proposed a new optimization method of the force-field parameters. This method uses many native structures from the Protein Data Bank, and therefore it is a knowledge-based method. The characteristic of the method lies in the fact that it uses the native structure only without introducing so-called decoys or misfolded structures. Althoguh our method is widely applicable to many force-field functions, we applied it to standard all-atom force fields. In Chapter 2, we gave the details of our method for optimizations of force-field parameters.

In Chapter 3, we applied this method to five existing force-field parameters, namely, AMBER parm94, AMBER parm96, AMBER parm99, CHARMM version 22, and OPLS-AA. The optimized parameters are the partial-charge parameters and the backbone torsion-energy parameters. For the partial-charge parameters, the optimized values of all the force fields did not change much from their respective original values. As for the backbone torsion-energy parameters, the optimized values changed a lot from their original ones. Although the values of the individual, optimized torsion-energy parameters are quite different, their overall shapes were found to be rather similar suggesting a tendency of convergence.

In Chapter 4, we found that all the optimized force-field parameters gave both $\alpha$-helical and $\beta$-hairpin structures more consistent with the experimental implications than the original force fields. It is indeed remarkable that just a small change in the backbone torsion-energy terms resulted in a significant variation in secondary-structure-forming tendencies of these peptides. This confirms the fact that the torsion-energy terms are the most problematic (for instance, the parm94, parm96, and parm99 versions of AM-

BER differ mainly in torsion-energy parameters). Our optimization method definitely improves the existing force fields and modifies different force fields towards a common function, which suggests that we are at least on the right track. However, we found that our optimization method depends on the value of force constant $K_x$ of the coordinate restrictions. We need a more systematic study on the $K_x$ dependence, and a work is in progress in this direction.