Assessment and enhancement of empirical force fields regarding the conformational equilibrium of carboxyl group in solution

Tetsuro Nagai[a]

[a]*Department of Physics, Graduate School of Science, Nagoya University, Nagoya, Aichi 464-8602, Japan*

Abstract

The equilibrium between the two conformers of carboxyl group is studied, whereby an improvement is proposed over the Amber force fields. The enhancement is based on the complete energetics, showing the qualitative inconsistency between the standard versions of the Amber force fields and higher-level calculations in solution, via the state-of-the-art techniques including the Hamiltonian replica-exchange molecular dynamics simulations, reweighting techniques, and 3D-RISM-KH calculations. Furthermore, the energetics is estimated by the 3D-RISM-KH calculations as regards other representative all-atom force fields, namely the CHARMM 36 and OPLS-AA/L force fields, showing noticeable difference in the contributions to the total energetics.

Keywords: Amber force field, carboxylic acid, glutamic acid, syn-/anti-conformational equilibrium, 3D-RISM theory, Hamiltonian replica-exchange molecular dynamics, solvation free energy, WHAM, MBAR, CHARMM 36, OPLS-AA/L

1. Introduction

Accurate empirical force fields are of great importance for the molecular dynamics (MD) and Monte Carlo simulations indispensable to study biological systems. Commonly used all-atom force fields for these systems include Amber [1–6], CHARMM [7], and OPLS-AA/L [8], often used together with the three site water model, such as TIP3P [9] and SPC [10]. Although these parameters were tested and revised extensively (see, e.g., Refs. [11–17]), some lack of reproducibility may persist. In the current work, we examine the balance between the syn- and anti-conformers (also known as Z and E conformers) in the carboxyl group of the glutamic acid with a main focus on Amber force fields. Here the syn- and anti-conformers correspond to the O=C−O−H dihedral angle of 0° and 180°, respectively. The balance has been studied with the acetic acid by QM/MM calculation; for example, the energy difference in gas phase was reported to be 5.7 kcal/mol (MP4SDTQ/6-311++G(d,p)) [18], 5.9 kcal/mol (MP3/6-311+G**) [19], 5.9 kcal/mol (AM1) [20], 5.3 kcal/mol (B3LYP/AUG-cc-pVDZ) [21], 6.9 kcal/mol [22] (RHF/DZP), and 6.5 kcal/mol (MP2/6-31++G**) [23], stabilizing the syn-form. On the other hand, in solution, the free energy difference is shown to drop to 0 kcal/mol (QM/MM with the energy representation method) [21], 1.1 kcal/mol (QM/MM) [20], and 1.7 kcal/mol (RISM-SCF/MCSCF) [22], yet the syn-form is more stable. Note that Ref. [21] by itself implies that its value is a slightly underestimated one. The syn-form thus should be more stable than the anti-form by ca. 6 kcal/mol in the gaseous phase and by ca. 1.5 kcal/mol in the aqueous solution. Nevertheless, the balance is not appreciated in most of the current Amber force fields in solution. In the article, therefore, the energetics is examined using the state-of-the-art techniques in the hope of a further development of the promising empirical force fields. This paper is organized as follows. In the Methods and Models section, the relevant force field parameters are briefly reviewed and the details of calculations are described. The section of Results and Discussion follows, giving the detailed

Email address: tnagai@nagoya-u.jp (Tetsuro Nagai)

energetics, whereby it is suggested that the value of V_1 of the O=C−O−H torsion should be 0.615 kcal/mol in some versions of Amber force fields such as Amber force field 03 (ff03) [3], 99SB (ff99SB) [4], and14SB (ff14SB) [6] for the shake of the balance in solution. Some of the other force fields including the CHARMM 36 and OPLS-AA/L force fields, are also analyzed using the 3D-RISM-KH theory [24–26]. The last section is devoted to Conclusions.

2. Methods and Models

2.1. Models

We briefly review the parameters of the dihedral potential energy related to the balance between the two conformers of the carboxyl group with a focus on the glutamic acid. In accordance with the common convention, the atoms in the carboxyl group of the glutamic acid are referred to as C_γ−C_δ($O_{\epsilon1}$)−$O_{\epsilon2}$−H. We review the parameters assigned to the C_γ−C_δ−$O_{\epsilon2}$−H and $O_{\epsilon1}$=C_δ−$O_{\epsilon2}$−H that are particularly relevant to the balance between the two conformers. When no confusion is expected, the subscripts are suppressed for the shake of clarity. Although there are many Amber force fields [1–6], the focus is put on Amber ff99SB [4], ff03 [3], and ff14SB [6]. Although becoming out of date, Amber ff99SB was studied because this must be one of the most widely spread Amber force fields. Amber ff03 employs the distinct partial charges, and Amber ff14SB is one of the newest variations.

In general, the dihedral interactions of the Amber force fields [2–6] are modeled by

$$\sum_n \frac{V_n}{2}\left(1-\cos(n\theta-\gamma)\right), \tag{1}$$

where θ and n are the dihedral angle and periodicity, respectively, and γ shifts the phase. As most of them such as Amber ff99SB [4] and ff03 [3] inherited some parameters from Amber force field 99 [2], the dihedral interactions related to C_δ−$O_{\epsilon2}$ are given by

$$V_{\mathrm{CCOH}} = \frac{4.60}{2}\left(1+\cos\left(2\theta-180\right)\right), \tag{2}$$

and

$$V_{\mathrm{OCOH}} = \frac{3.80}{2}\left(1+\cos\theta\right)+\frac{4.60}{2}\left(1+\cos\left(2\theta-180\right)\right). \tag{3}$$

For Amber ff14SB [6], the first term is replaced with

$$\begin{aligned} V_{\mathrm{CCOH}} = &\frac{0.226}{2}\left(1+\cos 4\theta\right)+\frac{0.958}{2}\left(1+\cos 3\theta\right)\\ &+\frac{5.412}{2}\left(1+\cos\left(2\theta-180\right)\right)+\frac{0.896}{2}\left(1+\cos\left(\theta-180\right)\right). \end{aligned} \tag{4}$$

Note that V_{OCOH} destabilizes the syn-form ($\theta = 0°$) by 3.80 kcal/mol compared to the anti-form ($\theta = 180°$). This treatment was applied based on the parameterization in the gas phase using the acetic acid [2].

It must be mentioned that, as Amber force field 14ipq (ff14ipq) [5] is a recently proposed version that has established a new family of the Amber force fields featuring the distinct and sophisticated fitting procedure over not only the nonbonded interactions but also the torsion interactions, this version has a very different trend from the other Amber force fields,

$$V_{\mathrm{CCOH}} = \frac{6.98568}{2}\left(1+\cos\left(2\theta-180\right)\right), \tag{5}$$

and

$$V_{\mathrm{OCOH}} = \frac{0.04146}{2}\left(1+\cos\theta\right)+\frac{3.34076}{2}\left(1+\cos\left(2\theta-180\right)\right). \tag{6}$$

Furthermore, we examine the counterparts of the CHARMM 36 [7] and OPLS-AA/L [8] force fields. As respects the former, the identical potential function is assigned to the C−C−O−H and O=C−O−H torsions,

$$V_{\mathrm{XCOH}} = \frac{4.10}{2}\left(1 + \cos\left(2\theta - 180\right)\right). \tag{7}$$

One the other hand, the latter has the potential functions given by

$$V_{\mathrm{OCOH}} = \frac{4.90}{2}\left(1 + \cos\left(2\theta - 180\right)\right) \tag{8}$$

and

$$V_{\mathrm{CCOH}} = \frac{3.00}{2}\left(1 + \cos\theta\right) + \frac{4.90}{2}\left(1 + \cos\left(2\theta - 180\right)\right). \tag{9}$$

Interestingly, this form lowers the potential energy of the syn-form, being opposite to the Amber force fields.

Another relevant interaction is the electrostatic potential energy between the carbonyl oxygen and acidic hydrogen atoms. The partial charges are therefore summarized in Table 1.

Table 1: Partial changes for the carbonyl oxygen and acidic hydrogen are summarized.

	ff99SB and ff14SB	ff03	ff14ipq	CHARMM36	OPLS-AA
$O_{\epsilon 1}$	−0.58380	−0.55903	−0.51143	−0.55	−0.44
H	0.46410	0.44766	0.40865	0.44	0.45

2.2. Methods

2.2.1. Hamiltonian replica-exchange molecular dynamics

First performed were the Hamiltonian replica-exchange MD (HREMD) simulations [27–30] with the Amber force fields. The barrier between the two conformers was reduced in 'hotter' replicas and thereby the conformational transitions happened during the simulations. In particular, the V_2 terms regarding the C−C−O−H and O=C−O−H torsions were divided by the factor up to 4, the effective temperature for these torsion energy terms corresponding to 1200 K in the spirit of the replica exchange with flexible tempering [31] or solute tempering [32, 33]. The obtained multiple trajectories were analyzed with the multiple-histogram reweighting techniques, namely WHAM [34, 35] and MBAR [36], and the thorough energetics was obtained for the original Amber 14SB force field. The 3D-RISM-KH [24–26, 37] calculations were performed as a part of analysis to obtain the solvation free energy for various water models including TIP3P [9], SPC [10], and TIP4P [9]. It is notable that the difference in the 3D-RISM-KH solvation free energy estimates between the different structures is reported to be pretty reasonable in contrast to the absolute values [38]. Finally, four values for V_1 of O=C−O−H were further investigated to extract an improved parameter for Amber ff99SB, ff03, and ff14SB via HREMD.

2.2.2. Details

The concrete conditions of the simulations are as follows. The blocked glutamic acid modeled by ff14SB was solvated in a cube with 760 TIP3P water molecules with the periodic boundary conditions. After equilibration with short NVT and NPT simulations, we further equilibrated the system with the NVT simulation over 1 ns and HREMD simulation of 1 ns per replica. The number of replicas was 6 and the scaling factors for V_2 were interpolated linearly from 1 (corresponding to the original force field) to 4 ('the hottest condition' effectively at 1200 K), yielding the large acceptance ratios of more than 80 %. The production run was executed over 1.6 μs per replica, namely in total 9.6 μs, to give the energetics over the whole angle range. The GROMACS 5 software package [39–42] was used, the time step was set to 2 fs, the LINCS was applied to the bond with hydrogen, the temperature was controlled at 300 K with the Nosé-Hoover thermostat [43, 44] to ensure the correct NVT ensemble, and the electrostatic potential was

treated with particle mesh Ewald (PME). The replica-exchange attempts were performed every 0.1 ps (50 MD steps). Additionally, four values for V_1 were studied with the similar procedure with the production run over 36 ns per replica. In the similar manner, Amber ff03 and ff99SB were studied; except that the number of replicas was set to 2 with the same range of the scaling factor for the shake of reduction in computation with Amber ff03, giving the fine acceptance ratio about 50 %. The blocked aspartic acid was also studied with ff14SB [see supplementary information (SI)]. The hydrogen bond analysis was performed based on the hydrogen–donor–accepter angle ($< 30°$) and donor–accepter distance (< 3.5 Å).

The average solvation free energy was obtained using the 3D-RISM-KH theory [24–26] as implemented in AmberTools [37, 45]. The average solvation free energy with TIP4P was estimated in an application of single-histogram reweighting techniques [46] to the coordinates obtained with the simulations with TIP3P. That is to say, the following formula was used

$$\langle \Delta G_{\mathrm{solv}} \rangle_{\mathrm{TIP4P}} = \frac{\left\langle \Delta G^{\mathrm{TIP4P}} \exp[-\beta \left(\Delta G^{\mathrm{TIP4P}} - \Delta G^{\mathrm{TIP3P}} \right)] \right\rangle}{\left\langle \exp[-\beta \left(\Delta G^{\mathrm{TIP4P}} - \Delta G^{\mathrm{TIP3P}} \right)] \right\rangle}, \tag{10}$$

where $\Delta G^{\mathrm{TIP3P}}$ and $\Delta G^{\mathrm{TIP4P}}$ denote the 3D-RISM solvation free energy estimates with respect to a given solute conformation with TIP3P and TIP4P, respectively. This is based on the natural assumption that the solute coordinates $\boldsymbol{r}^n$ are sampled by the MD simulations with the probability density of

$$p(\boldsymbol{r}^n) \propto \exp\left[-\beta \left(E_{\mathrm{solute}}(\boldsymbol{r}^n) + \Delta G^{\mathrm{TIP3P}}(\boldsymbol{r}^n) \right) \right], \tag{11}$$

where $E_{\mathrm{solute}}(\boldsymbol{r}^n)$ is the intra-solute interaction. As only the water model was modified, the intra-solute energy terms do not appear in Eq. 10. The identical procedure was also performed in regard to the SPC water model. Strictly speaking, the water models employed for the RISM calculations contain, as reported in Ref. [45] and implemented in Ambertools15 [37], the additional parameters for the hydrogen and dummy site. The syn- and anti-forms were defined by the configuration with the O=C−O−H torsion being $-90° < \theta < 90°$ and otherwise, respectively. AmberTools [37] and other programs [47, 48] were extensively used for handling the input and output.

2.2.3. Energetics by manually rotated structures

Furthermore, the 3D-RISM-KH [24–26, 37] calculations were performed over a series of manually prepared structures. These conformations were made by applying the rotation of the O=C−O−H torsion angle by the step of 1° to a conformation obtained with the above HREMD simulation with ff14SB. Confirming the consistency in the energetics between this series of structures and the HREMD simulation, the energetics was evaluated regarding other force fields, namely Amber ff14ipq, CHARMM 36 and OPLS-AA/L, in an application of the 3D-RISM-KH calculation to the conformations. It must be admitted that, as AmberTools was used in the 3D-RISM-KH calculations, the combination rule for the van der Waals interaction of OPLS-AA/L must have been mistaken. Nevertheless, the main contribution is the electrostatic interaction as shown below and some cancellation is anticipated in the relative values. Therefore, little effect is assumed. The other calculations concerning OPLS-AA/L were performed with GROMACS 5 free from this issue. Regarding the other force fields, the calculations were performed with Amber 14.

3. Results and Discussion

The energetics (Figure 1) detailed via HREMD in conjunction with reweighting demonstrates that the syn-form is unstable by ca. 2.0 kcal/mol compared to the anti-form in terms of the potential of mean force (PMF) [which is also referred to as 'free energy'] as well as the average potential energy, showing the discrepancy to the higher-level calculations mentioned in Introduction (see Fig. S1 in SI for an additional QM calculation). Two factors destabilize the syn-form. One is the dihedral potential energy that arises from V_1 of 3.80 kcal/mol for the O=C−O−H torsion. This V_1 value was introduced for an improved agreement between the empirical force field and higher-level calculations in the gas phase [2]. Presumably, the dihedral interaction was tweaked to compensate the strong intra-solute electrostatic interaction of 8.0 kcal/mol,

which should arise from the interaction between the carbonyl oxygen and acidic hydrogen. The other factor is the interaction between the solvent and solute, which is 5.9 kcal/mol. In total, the stabilization of 8.0 kcal/mol due to the electrostatic energy is defeated by the destabilization of 3.80 kcal/mol due to the dihedral interaction and 5.9 kcal/mol due to the interaction involving solvent, and therfore the syn-form is destabilized in the aqueous solution. This tendency persists even with other water models as per the 3D-RISM-KH solvation free energy analysis. While the difference in the average 3D-RISM-KH solvation free energy between the two conformations is estimated to be 7.30 ± 0.08 kcal/mol for TIP3P water model, it is estimated along with the single-histogram reweighting technique (Eq. 10) to be 6.85 ± 0.08 kcal/mol for TIP4P and 7.38 ± 0.08 kcal/mol for SPC. (For further reference, the correlation between the solvation free energy of TIP3P and other water models is given in Fig. S2.) Therefore only the marginal improvement is expected with TIP4P.

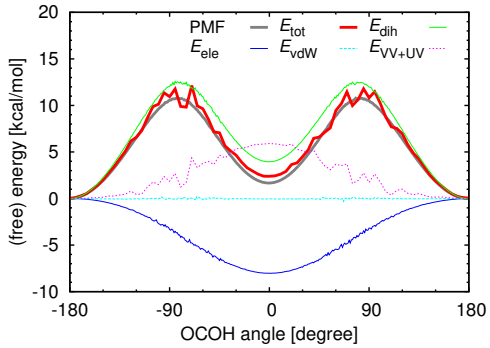

Figure 1: PMF and average potential energy components via the HREMD simulation in conjunction with the reweighting techniques are plotted as a function of the O=C−O−H dihedral angle. The thick gray and red lines represent the PMF and average total energy, respectively. The green solid, blue solid and cyan dash lines stand for the dihedral interaction, intra-solute electrostatic interaction, and intra-solute van der Waals interaction, respectively. The magenta dotted line marks the total interaction involving solvent due to both the electrostatic and van der Waals interactions. The electrostatic potential energy calculated in the reciprocal space by PME was simply suppressed from this decomposition. Hereafter, 'VV' and 'UV' stand for the solvent-solvent and solute-solvent interactions, respectively. Each term is shown relative to −180°. For the term involving the VV interaction, the moving average with five-degree window size is shown, whereas the other lines are based on the histogram with the bin width of 1°. The values for the PMF and average potential energy at 0° (corresponding to the syn-form) relative to −180° (to the anti-form) are 1.66 ± 0.01 kcal/mol and 2.4 ± 0.2 kcal/mol, respectively, indicating a small entropic effect of 0.7 ± 0.3 kcal/mol, where the errors were calculated through the standard errors of the 20 blocks into which the long HREMD simulation was split.

The breakdown of the electrostatic potential energy demonstrates that the anti-form is largely stabilized by the interaction between the solute and solvent at the sacrifice of the solvent-solvent interaction (Figure 2). A corresponding observation was found in the hydrogen bond analysis: the number of hydrogen bonds was 1.378 ± 0.002 (for the syn-form) and 1.701 ± 0.001 (for the anti-form) between $O_{\epsilon 1}$ and water; 0.5758 ± 0.0009 (for the syn-form) and 0.7936 ± 0.0002 (for the anti-form) between $O_{\epsilon 2}$ and water; and 1261.85 ± 0.02 (for the syn-form) and 1261.469 ± 0.004 (for the anti-form) between waters. In addition, Fig. 2 indicates that the main contribution to the energetics from solvation is the electrostatic rather than the van der Waals interactions.

Figure 3 shows the free energy difference (ΔF) between the two conformers as a function of the value for V_1 of the O=C−O−H torsion together with the digression line for Amber 14SB force field ($\Delta F = -2.14 + 1.04\,V_1$ [kcal/mol]), which also approximates ff03 and ff99SB well. The value of 0.615 kcal/mol for V_1 of O=C−O−H is recommended to get the stability of the syn-from by 1.5 kcal/mol in accordance with the previous QM/MM studies of the acetic acid in solution, as the intra-solute interaction between the side-chain and backbone has little to do with the balance in the glutamic acid (see Fig. S3 for a comparison with the aspartic acid). Nevertheless, excluding the V_1 term stabilizes the syn-form by 2.1 kcal/mol, which is still more or less in accord with the previous higher-level calculations. Whereas these modifications are adequate in solution, the stability of the syn-form would be overestimated in the gaseous phase. In other

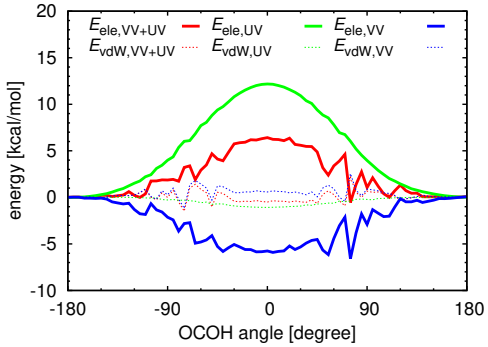

Figure 2: Components of the electrostatic and van der Waals energies are plotted against the O=C−O−H torsion angle. The three thick lines mark the electrostatic potential energy terms. The thick red line corresponds to the total electrostatic interaction involving solvent, the thick green line to the electrostatic solute-solvent interaction, and the thick blue line to the electrostatic solvent-solvent interaction. The three thin dashed lines are the counterparts in the van der Waals interaction. The contribution from reciprocal space was suppressed.

words, the qualitative consistency was obtained in both the situations by compromising on the quantitative agreement in the gas phase. One could modify the scaling factor for 1-4 interaction or partial charges to acquire an enhanced reproducibility in both of the environments. However, such a modification can demand the complete redesign of the Amber force field parameterization and was not included in the current work. As the reorganization energy is reported to interplay much in the seminal work [22], the polarizable force fields [49, 50] might be very useful for better transferability.

The present modification is particularly crucial when the peptide or proteins are studied at the low-pH conditions with the enhanced-sampling methods. That is, most conventional MD or MC simulations should hardly be affected, as the transitions would be prohibitively rare due to the barrier height between the two conformers as high as ca. 10 kcal/mol. Nevertheless, the defect of the force fields becomes obvious in virtue of the advanced technique for the enhanced sampling efficiency, and thereby unexpected transitions can distort the energetic analysis by ca. 7 kcal/mol through the solvation free energy, which is not always negligible.

Finally, the energetics was estimated in an application of 3D-RISM-KH theory to the manually prepared conformations, providing reasonable estimates in accord with the HREMD simulation with Amber ff14SB [see Figure 4(a)]. The syn-form is stabilized by 6.1 kcal/mol in gas and −1.1 kcal/mol in solution. Encouraged by the consistency, the energetics was further evaluated with regard to the Amber 14ipq, CHARMM 36, and OPLS-AA/L force fields, shown in Figure 4(b)–(d). These force fields demonstrate the syn-/anti-conformational equilibrium consistent with the higher-level calculations. As regards ff14ipq, the syn-form is more stable than the anti-form by 7.6 kcal/mol and 2.4 kcal/mol in the gaseous phase and in solution, respectively. In regard to the CHARMM 36 force field, the syn-form is more stable in the gas phase by 8.4 kcal/mol as well as in solution by 3.0 kcal/mol. With respect to the OPLS-AA/L force field, the intra-solute electrostatic potential energy does not stabilize the syn-form so much as with respect to the other force fields. This small stabilization is presumably ascribed to the smaller partial changes (see Table 1) and smaller value of the scaling factor for the 1-4 interactions. The multiplier is 0.5 in the OPLS-AA/L force field, while the Amber force fields set this to 0.83333 (= 1/1.2) and the CHARMM 36 force field to unity. It is noticeable that the electrostatic energy is almost perfectly cancelled by the 3D-RISM-KH solvation free energy as regards the OPLS-AA/L force field. Nevertheless, unlike the Amber force fields, the torsion interaction was designed to stabilize the syn-form, and overall the syn-form is more stabilized than the anti-form by 7.9 kcal/mol in gas and by 3.2 kcal/mol with RISM estimates, showing qualitative consistency with the higher-level calculations.

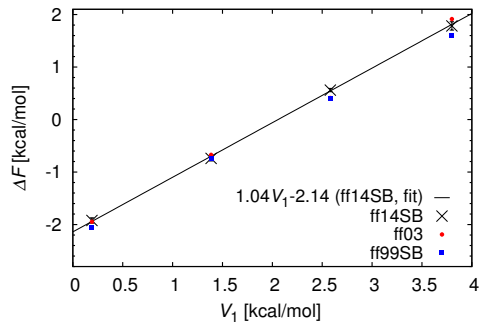

Figure 3: Free energy difference between the syn-form and anti-form ΔF is plotted as a function of V_1 of O=C−O−H together with the regression line for Amber ff14SB at 300 K. Black crosses, red circles, and blue squares mark ff14SB, ff03, and ff99SB, respectively. The black line is a linear regression obtained with ff14SB. The bars of black crosses express the errors obtained with splitting the 36-ns-per-replica HREMD simulations into 9 blocks. Regarding the other force fields, errors were at the similar level and suppressed for the shake of clarity. Data are collected from the replica coupling to the respective original force field with the modification concerned. Structures with the O=C−O−H torsion angle between −90° to 90° are defined as the syn-form and the other structures as the anti-form.

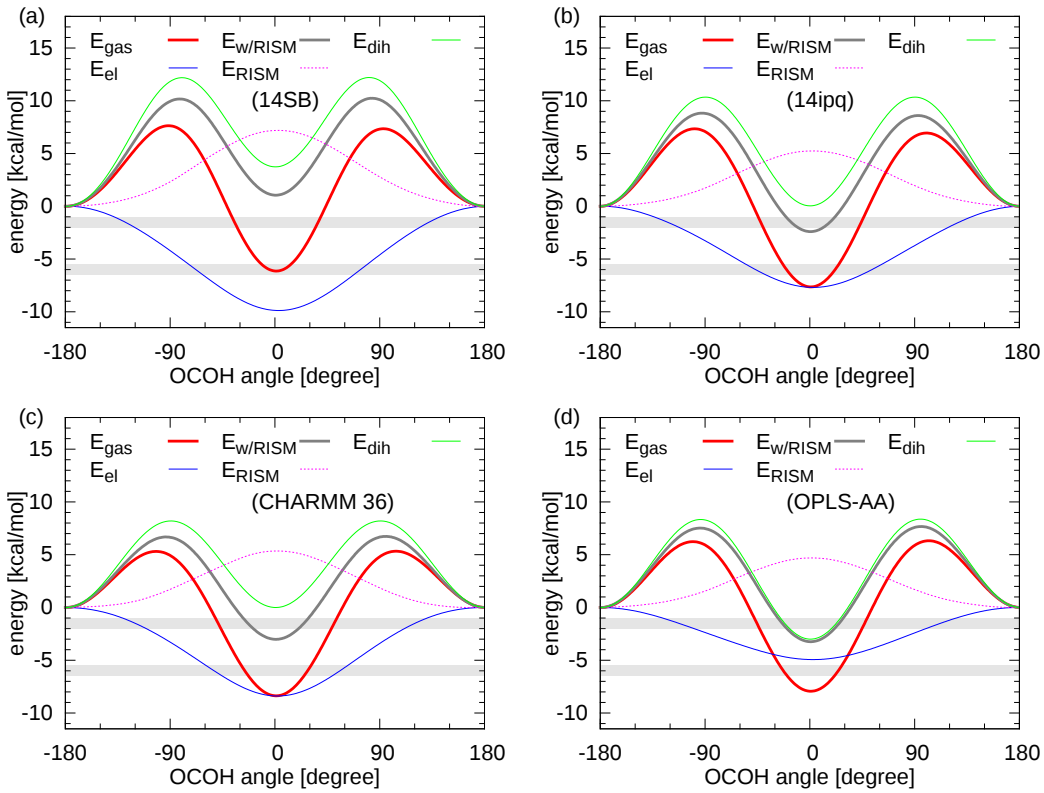

Figure 4: Intra-solute potential energy components and RISM solvation free energy estimates are plotted with respect to the O=C−O−H dihedral angle for Amber ff14SB. Each term is represented relative to −180°. The thick red and gray lines express the total intra-solute energy and the sum of the intra-solute energy and RISM solvation free energy, respectively. The green and blue solid lines correspond to the intra-solute dihedral interaction and electrostatic interaction, respectively, and the magenta dotted line to the RISM solvation free energy. The gray bands show the approximate implication due to the higher-level calculations described in Introduction, drawn between −2 to −1 kcal/mol and between −6.5 to −5.5 kcal/mol.

4. Conclusions

In this article, we examined the detailed energetics of the syn- and anti-conformers of glutamic acid via Hamiltonian replica-exchange molecular dynamics in conjunction with the reweighting techniques in addition to the 3D-RISM-KH estimates of the solvation free energy. In the analysis, the 3D-RISM-KH theory and reweighing technique were also combined. It is found that the common Amber force fields populate the anti-form more than the syn-form in solution, and therefore the balance is qualitatively inconsistent with the higher-level calculations. The anti-form is stabilized by the two factors: the dihedral interaction optimized in the gaseous phase and the solvation free energy. The unbalance persists even with other water models, such as TIP4P and SPC.

In the interest of the correct balance in solution, the value of 0.615 kcal/mol for V_1 was suggested for the O=C−O−H torsion interaction with respect to Amber ff14SB and related ones to read

$$V_{\mathrm{OCOH}} = \frac{0.615}{2}\left(1+\cos\theta\right) + \frac{4.60}{2}\left(1+\cos\left(2\theta-180\right)\right), \tag{12}$$

even though the syn-form is expected to be overstabilized in gas phase. Any of the scaling factor for 1-4 interaction and the partial changes was not modified, and thus the new parameter can be safely used with the other original parameters. The current modification is particularly important when peptides or proteins are studied with the enhanced sampling methods under the low-pH conditions. For example, the undesirable transitions to the anti-conformer will disturb the solvation free energy by ca. 7 kcal/mol, which may mislead the energetics qualitatively. Furthermore, the intra-solute energy and 3D-RISM-KH solvation free energy were analyzed over a series of conformations manually prepared. The simple procedure enabled the assessment regarding the Amber 14ipq, CHARMM 36, and OPLS-AA/L force fields. They are qualitatively fine and similar in the overall behavior, yet in particular the OPLS-AA/L force field reveals the different contribution to the total energetics. Further in-depth and accurate evaluation will be a future task of great interest.

Acknowledgement

The author thanks Drs. Straub and Panahi of Boston University, Drs. Takahashi and Sugita and Mr. Iwai of Ristumeikan University, and Drs. Okamoto and Fujimoto of Nagoya University for their comments and discussion. This work was supported in part by a Grant-in-Aid for Young Scientists (B) under Grant No. 26790083. Some of the computations were performed at the Research Center for Computational Science, Okazaki, Japan.

Appendix A. Supplementary material

Supplementary Information (SI) associated with this article can be found online at ********.

References

[1] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, Journal of the American Chemical Society 117 (1995) 5179–5197. doi:10.1021/ja00124a002.

[2] J. Wang, P. Cieplak, P. A. Kollman, How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?, Journal of Computational Chemistry 21 (2000) 1049–1074. doi:10.1002/1096-987X(200009)21:12<1049::AID-JCC3>3.0.CO;2-F.

[3] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, Journal of computational chemistry 24 (2003) 1999–2012. doi:10.1002/jcc.10349.

[4] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins 65 (2006) 712–25. doi:10.1002/prot.21123.

[5] D. S. Cerutti, W. C. Swope, J. E. Rice, D. A. Case, ff14ipq: A self-consistent force field for condensed-phase simulations of proteins, Journal of Chemical Theory and Computation 10 (2014) 4515–4534. doi:10.1021/ct500643c.

[6] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB, Journal of Chemical Theory and Computation 11 (2015) 3696–3713. doi:10.1021/acs.jctc.5b00255.

[7] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, A. D. MacKerell, Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ_1 and χ_2 dihedral angles, Journal of Chemical Theory and Computation 8 (2012) 3257–3273. doi:10.1021/ct300400x.

[8] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, The Journal of Physical Chemistry B 105 (2001) 6474–6487. doi:10.1021/jp003919d.

[9] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water, The Journal of Chemical Physics 79 (1983) 926. doi:10.1063/1.445869.

[10] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, in: B. Pullman (Ed.), Intermolecular Forces, Springer, 1981, pp. 331–342. doi:10.1007/978-94-015-7658-1_21.

[11] T. Yoda, Y. Sugita, Y. Okamoto, Secondary-structure preferences of force fields for proteins evaluated by generalized-ensemble simulations, Chemical Physics 307 (2004) 269–283. doi:10.1016/j.chemphys.2004.08.002.

[12] T. Yoda, Y. Sugita, Y. Okamoto, Comparisons of force fields for proteins by generalized-ensemble simulations, Chemical Physics Letters 386 (2004) 460–467. doi:10.1016/j.cplett.2004.01.078.

[13] A. D. Mackerell, M. Feig, C. L. Brooks, Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation, Journal of Computational Chemistry 25 (2004) 1400–1415. doi:10.1002/jcc.20065.

[14] I. Kurisaki, T. Takahashi, Assessment of dynamic properties of water around a monovalent ion: A classical molecular dynamics simulation study, Computational and Theoretical Chemistry 966 (2011) 26–30. doi:10.1016/j.comptc.2011.02.004.

[15] K. Takemura, A. Kitao, Water model tuning for improved reproduction of rotational diffusion and NMR spectral density, Journal of Physical Chemistry B 116 (2012) 6279–6287. doi:10.1021/jp301100g.

[16] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, D. E. Shaw, Systematic validation of protein force fields against experimental data, PLoS ONE 7 (2012) 1–6. doi:10.1371/journal.pone.0032131.

[17] S. Piana, J. L. Klepeis, D. E. Shaw, Assessing the accuracy of physical models used in protein-folding simulations: Quantitative evidence from long molecular dynamics simulations, Current Opinion in Structural Biology 24 (2014) 98–105. doi:10.1016/j.sbi.2013.12.006.

[18] M. T. Nguyen, D. Sengupta, G. Raspoet, L. G. Vanquickenborne, Theoretical study of the thermal decomposition of acetic acid: Decarboxylation versus dehydration, The Journal of Physical Chemistry 99 (1995) 11883–11888. doi:10.1021/j100031a015.

[19] K. B. Wiberg, K. E. Laidig, Barriers to rotation adjacent to double bonds. 3. The C–O barrier in formic acid, methyl formate, acetic acid, and methyl acetate. The origin of ester and amide "resonance", Journal of the American Chemical Society 109 (1987) 5935–5943. doi:10.1021/ja00254a006.

[20] J. Gao, J. Pavelites, Aqueous basicity of the carboxylate lone pairs and the carbon-oxygen barrier in acetic acid: A combined quantum and statistical mechanical study, Journal of the American Chemical Society (1992) 1912–1914.

[21] T. Hori, H. Takahashi, M. Nakano, T. Nitta, W. Yang, A QM/MM study combined with the theory of energy representation: Solvation free energies for anti/syn acetic acids in aqueous solution, Chemical Physics Letters 419 (2006) 240–244. doi:10.1016/j.cplett.2005.11.096.

[22] H. Sato, F. Hirata, The syn-/anti-conformational equilibrium of acetic acid in water studied by the RISM-SCF/MCSCF method, Journal of Molecular Structure: THEOCHEM 461-462 (1999) 113–120. doi:10.1016/S0166-1280(98)00436-9.

[23] D. Deerfield, L. Pedersen, Enol and deprotonated forms of acetic and malonic acid, Journal of Molecular Structure: THEOCHEM 368 (1996) 163–171. doi:10.1016/S0166-1280(96)90557-6.

[24] D. Beglov, B. Roux, An integral equation to describe the solvation of polar molecules in liquid water, Journal of Physical Chemistry B 101 (1997) 7821–7826. doi:10.1021/jp971083h.

[25] A. Kovalenko, F. Hirata, Three-dimensional density profiles of water in contact with a solute of arbitrary shape: A RISM approach, Chemical Physics Letters 290 (1998) 237–244. doi:10.1016/S0009-2614(98)00471-0.

[26] A. Kovalenko, F. Hirata, Self-consistent description of a metal–water interface by the Kohn–Sham density functional theory and the three-dimensional reference interaction site model, The Journal of Chemical Physics 110 (1999) 10095–10112. doi:10.1063/1.478883.

[27] K. Hukushima, K. Nemoto, Exchange Monte Carlo method and application to spin glass simulations, Journal of the Physics Society Japan 65 (1996) 1604–1608. doi:10.1143/JPSJ.65.1604.

[28] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chemical Physics Letters 314 (1999) 141–151. doi:10.1016/S0009-2614(99)01123-9.

[29] Y. Sugita, A. Kitao, Y. Okamoto, Multidimensional replica-exchange method for free-energy calculations, The Journal of Chemical Physics 113 (2000) 6042. doi:10.1063/1.1308516.

[30] H. Fukunishi, O. Watanabe, S. Takada, On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction, Journal of Chemical Physics 116 (2002) 9058–9067. doi:10.1063/1.1472510.

[31] S. L. C. Moors, S. Michielssens, A. Ceulemans, Improved replica exchange method for native-state protein sampling, Journal of Chemical Theory and Computation 7 (2011) 231–237. doi:10.1021/ct100493v.

[32] P. Liu, B. Kim, R. a. Friesner, B. J. Berne, Replica exchange with solute tempering: A method for sampling biological systems in explicit water, Proceedings of the National Academy of Sciences of the United States of America 102 (2005)

13749–13754. doi:10.1073/pnas.0506346102.

[33] T. Terakawa, T. Kameda, S. Takada, On easy implementation of a variant of the replica exchange with solute tempering in GROMACS, Journal of Computational Chemistry 32 (2011) 1228–1234. doi:10.1002/jcc.21703.

[34] A. Ferrenberg, R. Swendsen, Optimized Monte Carlo data analysis, Physical Review Letters 63 (1989) 1195–1198. doi:10.1103/PhysRevLett.63.1195.

[35] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, The weighted histogram analysis method for free-energy calculations on biomolecules. I. the method, Journal of Computational Chemistry 13 (1992) 1011–1021. doi:10.1002/jcc.540130812.

[36] M. R. Shirts, J. D. Chodera, Statistically optimal analysis of samples from multiple equilibrium states, The Journal of Chemical Physics 129 (2008) 124105. doi:10.1063/1.2978177.

[37] D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, P. A. Kollman, AMBER 15, University of California, San Francisco, CA, 2015.

[38] T. Sumi, A. Mitsutake, Y. Maruyama, A solvation-free-energy functional: A reference-modified density functional formulation, Journal of Computational Chemistry 36 (2015) 1359–1369. doi:10.1002/jcc.23942.

[39] H. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation, Computer Physics Communications 91 (1995) 43–56. doi:10.1016/0010-4655(95)00042-E.

[40] E. Lindahl, B. Hess, D. van der Spoel, GROMACS 3.0: A package for molecular simulation and trajectory analysis, Journal of Molecular Modeling 7 (2001) 306–317. doi:10.1007/s008940100045.

[41] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, GROMACS: fast, flexible, and free, Journal of Computational Chemistry 26 (2005) 1701–1718. doi:10.1002/jcc.20291.

[42] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29 (2013) 845–854.

[43] S. Nosé, A molecular dynamics method for simulations in the canonical ensemble, Molecular Physics 52 (1984) 255–268. doi:10.1080/00268978400101201.

[44] W. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Physical Review A 31 (1985) 1695–1697. doi:10.1103/PhysRevA.31.1695.

[45] T. Luchko, S. Gusarov, D. R. Roe, C. Simmerling, D. A. Case, J. Tuszynski, A. Kovalenko, Three-dimensional molecular theory of solvation coupled with molecular dynamics in Amber, Journal of Chemical Theory and Computation 6 (2010) 607–624. doi:10.1021/ct900460m.

[46] A. M. Ferrenberg, R. H. Swendsen, New Monte Carlo technique for studying phase transitions, Physical Review Letters 61 (1988) 2635–2638. doi:10.1103/PhysRevLett.61.2635.

[47] A. W. Sousa da Silva, W. F. Vranken, ACPYPE - AnteChamber PYthon Parser interfacE, BMC Research Notes 5 (2012) 367. doi:10.1186/1756-0500-5-367.

[48] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, MDAnalysis: A toolkit for the analysis of molecular dynamics simulations, Journal of Computational Chemistry 32 (2011) 2319–2327. doi:10.1002/jcc.21787.

[49] T. A. Halgren, W. Damm, Polarizable force fields, Current Opinion in Structural Biology 11 (2001) 236–42. doi:10.1016/S0959-440X(00)00196-2.

[50] J. W. Ponder, C. Wu, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. a. Distasio, M. Head-gordon, G. N. I. Clark, M. E. Johnson, T. Head-gordon, Current status of the AMOEBA polarizable force field, Journal of Physical Chemistry B 114 (2010) 2549–2564. doi:10.1021/jp910674d.