Peptide Free Energy Landscapes Calibrated by Molecular Orbital Calculations

S. ONO¹, M. KURODA², J. HIGO³, N. KAMIYA⁴, N. NAKAJIMA⁵ and H. NAKAMURA^{5,*}

¹Mitsubishi Pharma Corp., 1000 Kamoshida, Aoba-ku, Yokohama, Kanagawa 227-8502, Japan

²Tanabe Seiyaku Co., Ltd., 3-16-89 Kashima, Yodogawa-ku, Osaka 532-8505, Japan

³Tokyo University of Pharmacy and Life Science, 1432-1 Horinouchi, Hachioji, Tokyo 192-0392, Japan

⁴Department of Bioinformatics, Biomolecular Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan

⁵Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

(*Author for correspondence, e-mail: harukin@protein.osaka-u.ac.jp)

Abstract. Free energy landscapes of peptide conformations were calibrated by *ab initio* molecular orbital calculations, after enhanced conformational sampling using the multicanonical molecular dynamics simulations. Three different potentials of mean force for an isolated dipeptide were individually obtained using the conventional force fields, AMBER parm94, AMBER parm96, and CHARMM22. Each potential of mean force was calibrated based on the umbrella sampling algorithm from the adiabatic energy map that was calculated separately by the *ab initio* molecular orbital method. All the calibrated potentials of mean force coincided well. The calibration was applied to a peptide in explicit water, and the calibrated free energy landscapes did not depend on the force field used in conformational sampling, as far as the conformational space was well sampled.

Key words: force field, multicanonical molecular dynamics, peptide conformations, potential of mean force, quantum chemical calculation, structural ensemble

1. Introduction

As the post-genome sequencing era, the biochemical function of every protein is now one of the common issues that should be revealed through the protein-protein interaction. So far, the structure-function relationships have mainly been addressed mostly from the database analyses. However, even when the structural genomics may have covered the representative conformation for every protein family, deductive approaches other than the database analyses are required for providing precise structures, in particular, for modeling of low homologous local regions such as loops, of substantially flexible regions, and of structural docking of proteins and the ligands.

Enhanced conformational sampling provides a quite powerful tool for those deductive ways, so that the free energy landscapes are obtained and the stable

conformations are predicted in the sense of free energy [1–3]. In addition, for short peptide fragments, we have revealed the entire conformational ensembles, which agree well with experiments. Nakajima et al. [4] reproduced the β -turn conformational diversity depending on the amino acid residues, and Higo et al. [5] succeeded in obtaining the ensemble, the majority of which forms a β -hairpin conformation. Moreover, the conformational ensemble of a peptide fragment showing the random coil CD spectrum was also calculated, and almost all the conformations were the disordered structures, with rare possible conformations of α -helix and β -hairpins [6].

During the above conformation diversity search, we have noticed a substantial problem on the force field used in the simulations. In fact, the free energy landscapes of short peptides in explicit water were completely different, depending upon the force field AMBER parm94 and parm96 [7]. The difference between AMBER parm94 and CHARMM was also pointed out [4, 8]. Namely, AMBER parm94 [9] tends to produce helical and turn conformations, and AMBER parm96 [10] and CHARMM [11] tend to produce extended conformations in the ensemble at 300K.

Thus, one should select or develop an appropriate force field that gives reliable conformational ensemble. We propose another approach that the free energy landscape is calibrated from any force field, by correcting the contribution of the original potential function referring the adiabatic potential map given by the *ab initio* molecular orbital calculations [12]. The algorithm is a simple application of the umbrella sampling method [13], and applications of this approach are shown for several simple peptide systems.

2. Methods

The enhanced conformational sampling was performed by the multicanonical molecular dynamics (MD) method, which was precisely described elsewhere [4, 7]. The original idea for calibration of potential of mean force was given by Nakamura et al. [13]. A molecular conformational state is identified by a set of geometrical parameters $\{x\}$, and the difference is considered between the potential energy function $E_{MM}(x)$ calculated by a force field and the adiabatic energy value $E_{MO}(x)$ given by molecular orbital (MO) calculations.

$$\begin{aligned} E_{MM}(x) &= E_{MO}(x) + (E_{MM}(x) - E_{MO}(x)) \\ &= E_{MO}(x) + \Delta E(x) \end{aligned} \quad (1)$$

$$\Delta E(x) = E_{MM}(x) - E_{MO}(x) \quad (2)$$

Equations (1) and (2) mean that when $\Delta E(x)$ is assumed to be an umbrella potential for the system described by $E_{MO}(x)$, the conventional molecular mechanics calculation based on the force field $E_{MM}(x)$ equals to the umbrella sampling using $\Delta E(x)$ as the umbrella potential. Thus, after obtaining an appropriate canonical ensemble with the probability distribution $\rho_{MM}(x)$ by the conventional multicanonical MD,

the calibrated probability distribution $\rho_{MO}(x)$ and the corresponding potential of mean force (PMF), $W(x)$, are readily obtained.

$$\begin{aligned} W(x) &= -k_B T \ln \rho_{MO}(x) \\ &= -k_B T \ln \rho_{MM}(x) - \Delta E(x) \\ &\quad + k_B T \ln \langle \exp(\Delta E(x)/k_B T) \rangle_{MM} \end{aligned} \quad (3)$$

Here, $\langle \rangle_{MM}$ indicates the average for the canonical ensemble produced by the multicanonical MD using the force field $E_{MM}(x)$, and is independent of the parameter x . k_B is the Boltzmann constant. Thus, once $\Delta E(x)$ is given, the PMF corresponding to the free energy at the state x is correctly calibrated. In this paper, the model system is composed of Ala residues, and so (φ, ψ) is simply assumed as x approximating the entire conformational states.

The all-atom AMBER parm94 [9], AMBER parm96 [10], and CHARMM22 [11] force fields were used. The program PRESTO ver 3.0 [4, 14] was used for the multicanonical MD simulations.

The simulations of the dipeptide, acetyl-N methylated alanine (Ac-Ala-NMe), was carried out *in vacuo*, and the simulation of the peptide dimer blocked with N-terminal acetyl and C-terminal N-methyl groups (Ac-(Ala)₂-NMe), was also carried out in explicit water. The non-bonded interactions were not truncated *in vacuo*, and the non-bonded interactions longer than 12 Å were truncated in explicit water, where the solute was placed at the center of a sphere with the TIP3P water model [15]. The radii of the spheres were 12 Å. The numbers of water molecules were 219, and the CAP constraint was used. All the multicanonical MD simulations were performed with a constant temperature by the constraint method [16, 17], where the total kinetic energy was always constant. The time step was 0.5 fs for *in vacuo* simulations, and 1.0 fs for in explicit water. The SHAKE method [18] was used in explicit water to constrain every bond between a heavy atom and a hydrogen atom. The final multicanonical MD calculation was carried out for 5×10^7 steps for each force field.

The adiabatic energy map for E_{MO} was calculated for Ace-Ala-Nme by *ab initio* MO calculations at total 121 points, by choosing 11 points in φ , and 11 points in ψ . The HF/6-31G*//HF/6-31G* level of MO calculations were performed by Gaussian 98 [19]. Figure 1 shows the energy map showing E_{MO} . For the same optimized conformers, the classical potential energy for E_{MM} was calculated using the force field of AMBER parm94 [9], AMBER parm96 [10], and CHARMM22 [11] by minimizing the local peptide geometry with restraining the two parameters, φ and ψ . The average root-mean square deviation (r.m.s.d.) between the initial geometries, which were optimized by *ab initio* MO and by classical molecular mechanics, was less than 0.12 Å for each force field. Then, the difference energy map $\Delta E(\varphi, \psi)$ was constructed. In Figure 2, $E_{MM}(\varphi, \psi)$ and $\Delta E(\varphi, \psi)$ for the three force fields are shown.

The geometrical parameters of the peptide were monitored at every ten steps of the multicanonical MD simulations, and $\rho_{MM}(\varphi, \psi)$ was observed for each force

Figure 1. The adiabatic energy map of Ac-Ala-Nme computed by the HF/6-31G**/HF/3-21G level of MO calculations. The thick lines correspond to 2 kcal/mol above the energy minimum, and the contour intervals are all 2 kcal/mol.

field at 300 K. Following eq. (3), the probability distribution and the PMF were calibrated to give $\rho_{MO}(\phi, \psi)$ and the corresponding PMF from $\Delta E(\phi, \psi)$, which was computed from interpolating those values at the lattice points.

3. Results and Discussion

Figures 3(a)–3(c) show the original PMF maps for Ac-Ala-NMe *in vacuo* using the force field of AMBER parm94, AMBER parm 96, and CHARMM22. The corresponding calibrated PMFs are shown in Figures 3(d)–3(f), respectively.

As pointed out previously [4, 7, 8], the free energy of the Ala in Figure 3(a) is very low only around the α -helical conformations in the original PMF by AMBER parm94. In contrast, the extended conformations are much more stable in the original PMF by AMBER parm96 (Figure 3(b)) and CHARMM22 (Figure 3(c)). When those PMFs were calibrated, every free energy profile of the local backbone conformation of the Ala residue coincided well with each other, and both the extended and helical conformations appeared as the low free energy states (see Figures 3(d)–3(f)). Although CHARMM22 force field was developed by a different group with the different potential functions from AMBER force fields, all the PMFs show very similar features with each other.

Figure 2. The potential energy, $E_{MM}(\phi, \psi)$, of Ac-Ala-Nme for the force field of (a) AMBER parm94, (b) AMBER parm96, and (c) CHARMM22, and the energy difference, $E_{MM}(\phi, \psi) - E_{MO}(\phi, \psi) = \Delta E(\phi, \psi)$ for (d) AMBER parm94, (e) AMBER parm96, and (f) CHARMM22. The contour intervals are all 2 kcal/mol. For $E_{MM}(\phi, \psi)$, the thick lines correspond to 2 kcal/mol above each energy minimum. For $\Delta E(\phi, \psi)$, the solid thick lines correspond to 0 kcal/mol, and the solid and dashed thin lines correspond to positive and negative values, respectively.

Figure 3. The uncorrected original PMFs for the dipeptide Ac-Ala-NMe *in vacuo* using (a) AMBER parm94, (b) AMBER parm96, and (c) CHARMM22. The PMFs calibrated from the corresponding original ones using (d) AMBER parm94, (e) AMBER parm96, and (f) CHARMM22. The thick lines correspond to 2 kcal/mol above each PMF minimum, and the contour intervals are all 2 kcal/mol.

Figure 4. The original PMFs for Ac-(Ala)₂-NMe in explicit water of the first Ala using (a) AMBER parm94, (b) AMBER parm96, and (c) CHARMM22., and the calibrated potential of mean force using (d) AMBER parm94, (e) AMBER parm96, and (f) CHARMM22, respectively. Contour lines were drawn by the same way as in Figure 3.

Figure 5. Free energy profiles for Ac-(Ala)₂-NMe in explicit water along the distance, d , between the carbonyl oxygen of the N-terminal acetyl group, and the C-terminal amide proton at 300 K. (a) The original and (b) the calibrated PMFs. In both figures, solid lines are for the force field, AMBER parm94; broken lines, AMBER parm96; and dash-and-dotted lines, CHARMM22.

In explicit water, Figure 4 shows the original and calibrated PMFs for the first Ala residue of (Ac-(Ala)₂-NMe), using the three force fields, respectively. Again, the free energy of the Ala residue preferred the helical conformations in the original PMF by AMBER parm94, and, in contrast, the extended conformations tended to be assumed in the original PMFs by AMBER parm96 and CHARMM22. This tendency was also observed for the second Ala residue of the peptide in water [12]. The PMFs calibrated from the both AMBER parm94 and parm96 coincided well, however the PMF calibrated from CHARMM slightly deviated from the PMFs obtained from both two AMBER force fields around the region having positive φ and positive ψ . It is due to the fact that this conformational space was rarely sampled by CHARMM22 force field, and that it was difficult to calibrate it well.

Figure 5 shows the free energy landscapes for the peptide Ac-(Ala)₂-NMe in explicit water along the distance, d , between the carbonyl oxygen of the N-terminal acetyl group and the C-terminal amide proton at 300 K. The original PMFs in Figure 5(a) show very different features. AMBER parm94 prefers turn conformations with $d = 2$ Å, but the energy landscape with AMBER parm96 strongly suggests that this peptide prefers the extended conformations. CHARMM22 force field yields the free energy profile having the middle characteristics of the above two, but it lacks a stable turn conformation as previously reported by Tobias et al. [20]. In contrast, once they were calibrated as indicated in Figure 5(b), all the PMFs coincided well, although the PMF calibrated from CHARMM22 deviated slightly from those calibrated from AMBER parm94 and parm96.

In conclusion, we have indicated that the free energy landscapes of peptide molecules strongly depend on the force field used in the molecular simulation, and that the differences can be eliminated by the calibration procedures referring a reliable energy map that is provided from *ab initio* MO calculations. However, one should pay an attention on the fact that this calibration is not always available, in particular when the conformational sampling is not performed widely enough for the conformational space. In fact, the difference observed between the original and the calibrated PMFs from the both AMBER and the CHARMM22 force field was caused mainly because the region with positive φ and positive ψ has almost never been sampled during the simulation with CHARMM22 even with the enhanced sampling algorithm.

In *ab initio* molecular modeling of a short peptide [4, 5, 6], and of a flexible loop [1, 2], this calibration could give reliable estimate of the most stable and the meta-stable conformations. In addition, the future molecular docking study should have higher accuracy than the current one that is based on the conventional force field [3].

When we obtained the adiabatic energy map in this study, we optimized the peptide conformations keeping the φ and ψ values with the molecular orbital 6-31G*, and the energy was calculated with the 6-31G* orbital. The levels of the molecular orbital were not the highest one at present [21, 22], but we here provide the method to calibrate any force field. People who need a higher level orbital can

only compute more accurate energy map, $E_{MO}(\varphi, \psi)$, and calibrate the resulted PMFs following the current method.

This calibration is not only effective for analysis of the stable conformational states as described here, but also it is powerful for analysis of the transition states where the empirical force fields do not always give accurate energy values. In fact, the free energy landscape around the transition state conformations from *cis* to *trans* Pro residue due to the biased electron density was well calibrated by the current method [13, 23].

Acknowledgements

N.N. and H.N. were supported by a grant-in-aids for Scientific Research from the Japan Society for the Promotion of Science (12558083, 12680657, 12208029 and 12144206).

References

1. Shirai, H., Nakajima, N., Higo, J., Kidera, A. and Nakamura, H.: Conformational Sampling of CDR-H3 in Antibodies by Multicanonical Molecular Dynamics, *J. Mol. Biol.* **278** (1998), 481–496.
2. Kim, S.T., Shirai, N., Nakajima, N. and Nakamura, H.: Enhanced Conformational Diversity Search of CDR-H3 in Antibodies: Role of the First CDR-H3 Residue, *Proteins: Struct. Funct. Genet.* **37** (1999), 683–696.
3. Nakajima, N., Higo, J., Kidera, A. and Nakamura, H.: Flexible Docking of a Ligand Peptide to a Receptor Protein by Multicanonical Molecular Dynamics Simulation, *Chem. Phys. Lett.* **278** (1997), 297–301.
4. Nakajima, N., Higo, J., Kidera, A. and Nakamura, H.: Free energy Landscapes of Peptides by Enhanced Conformational Sampling, *J. Mol. Biol.* **296** (2000), 197–216.
5. Higo, J., Galzitskaya, O.V., Ono, S. and Nakamura, H.: Energy Landscape of a β -Hairpin Peptide in Explicit Water studied by Multicanonical Molecular Dynamics, *Chem. Phys. Lett.* **337** (2001), 169–175.
6. Higo, J., Ito, N., Kuroda, M., Ono, S., Nakajima, N. and Nakamura, H.: Energy Landscape of a Peptide Consisting of α -Helix, 3_{10} -Helix, β -Turn, β -Hairpin and Other Disordered Conformations, *Protein Sci.* **10** (2001), 1160–1171.
7. Ono, S., Nakajima, N., Higo, J. and Nakamura, H.: Peptide Free-Energy Profile is Strongly Dependent on the Force Field: Comparison of C96 and AMBER95, *J. Comput. Chem.* **21** (2000), 748–762.
8. Samuelson, S., Tobias, D.J., Martyna, G.J.: Modern Computational Methodology applied to the Simulation of Blocked Trialanine Peptide *in vacuo*, Water Clusters and Bulk Water, *J. Phys. Chem.* **101B** (1997), 7592–7603.
9. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.: A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids and Organic Molecules, *J. Am. Chem. Soc.* **117** (1995), 5179–5197.
10. Kollman, P., Dixon, R., Cornell, W., Fox, T., Chipot, C. and Pohorille, A.: The Development/Application of a ‘Minimalist’ Organic/Biochemical Molecular Mechanics Force Field using a Combination of *ab initio* Calculations and Experimental Data, In: W.F. van Gunsteren,

- P.K. Weiner and A.J. Wilkinson (eds.), *Computer Simulation of Biomolecular Systems* **3**, Kluwer/ESCOM, The Netherlands, 1997, pp. 83–96.
11. MacKerell, Jr., A.D., Bashford, D., Bellott, M., Dunbrack, Jr., R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kichnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, III, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. and Karplus, M.: All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins, *J. Phys. Chem.* **102B** (1998), 3586–3616.
 12. Ono, S., Kuroda, M., Higo, J., Nakajima, N. and Nakamura, H.: Calibration of Force Field Dependency in Free Energy Landscapes of Peptide Conformations by Quantum Chemical Calculations, *J. Comput. Chem.*, in press.
 13. Nakamura, H., Ono, S. and Higo, J.: A General *ab initio* Approach for Free Energy Landscapes of Biological Molecules around the Transition States, *Proc. Japan Acad.* **75B** (1999), 291–294.
 14. Morikami, K., Nakai, T., Kidera, A., Saito, M. and Nakamura, H.: PRESTO: A Vectorized Molecular Mechanics Program for Biopolymers, *Comput. Chem.* **16** (1992), 243–248.
 15. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L.: Comparison of Simple Potential Functions for Simulating Liquid Water, *J. Chem. Phys.* **79** (1983), 926–935.
 16. Hoover, W.G., Ladd, A.J.C. and Moran, B.: High-Strain-Rate Plastic Flow studied via Nonequilibrium Molecular Dynamics, *Phys. Rev. Lett.* **48** (1982), 1818–1820.
 17. Evans, D.J. and Morris, G.P.: The Isothermal/Isobaric Molecular Dynamics Ensemble, *Phys. Lett.* **98A** (1983), 433–436.
 18. Ryckaert, J.-P., Ciccotti, G. and Berendsen, H.J.C.: Numerical Integration of Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes, *J. Comput. Phys.* **23** (1977), 327–341.
 19. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Zakrzewski, V.G., Montgomery, J.A., Stratmann, R.E., J.C. Burant, J.C., Dapprich, S., Millam, J.M., Daniels, A.D., Kudin, K.N., Strain, M.C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Adamo, C.P.C., Clifford, S., Ochterski, J., Petersson, G.A., Ayala, P.Y., Cui, Q., Morokuma, K., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Cioslowski, J., Ortiz, J.V., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P.M.W., Johnson, B.G., Chen, W., Wong, M.W., Andres, J.L., Head-Gordon, M., Replogle, E.S. and Pople, J.A.: *Gaussian 98* (Revision A6), Gaussian, Inc., Pittsburgh PA, 1998.
 20. Tobias, D.J., Sneddon, S.F. and Brooks, C.L., III: Reverse Turns in Blocked Dipeptides are Intrinsically Unstable in Water, *J. Mol. Biol.* **216** (1990), 783–796.
 21. Beachy, M.D., Chasman, D., Murphy, R.B., Halgren, T.A. and Friesner, R.A.: Accurate *ab initio* Quantum Chemical Determination of the Relative Energies of Peptide Conformations and Assessment of Empirical Force Fields, *J. Am. Chem. Soc.* **119** (1997), 5908–5920.
 22. Kaminsky, G.A., Friesner, R.A., Tirado-Rives, J. and Jorgensen, W.L.: Evaluation of Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides, *J. Phys. Chem.* **105B** (2001), 6474–6487.
 23. Ono, S., Nakajima, N., Higo, J. and Nakamura, H.: The Multicanonical Weighted Histogram Analysis Method for the Free-Energy Landscape along Structural Transition Paths, *Chem. Phys. Lett.* **312** (1999), 247–254.