

A New Set of Atomic Radii for Accurate Estimation of Solvation Free Energy by Poisson–Boltzmann Solvent Model

Junya Yamagishi,^[a,b] Noriaki Okimoto,^{*,[b]} Gentaro Morimoto,^[b] and Makoto Tajji^[a,b]

The Poisson–Boltzmann implicit solvent (PB) is widely used to estimate the solvation free energies of biomolecules in molecular simulations. An optimized set of atomic radii (PB radii) is an important parameter for PB calculations, which determines the distribution of dielectric constants around the solute. We here present new PB radii for the AMBER protein force field to accurately reproduce the solvation free energies obtained from explicit solvent simulations. The presented PB radii were optimized using results from explicit solvent simulations of the large systems. In addition, we discriminated PB radii for N- and

C-terminal residues from those for nonterminal residues. The performances using our PB radii showed high accuracy for the estimation of solvation free energies at the level of the molecular fragment. The obtained PB radii are effective for the detailed analysis of the solvation effects of biomolecules. © 2014 The Authors. Journal of Computational Chemistry Published by Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23728

Introduction

Water, as a solvent, is crucial in the function of biomolecules; accordingly, it is of great importance to accurately account for solvation effects in the computational simulation of biomolecules. Explicit solvent models can provide a detailed description of the solvation effect^[1,2] but involve high computational costs because of the large amount of discrete water molecules in the system and the need for their conformational sampling. Implicit solvent models, which permit characterization of the mean effect of the surrounding solvents without including discrete water molecules in the system, have great advantages in terms of computational efficiency. For these reasons, implicit solvent models are widely used in molecular mechanics simulations as alternatives to explicit solvent models.^[3–7] Typical implicit solvent models decompose the solvation effect into nonpolar and polar components. The nonpolar component is often estimated from the molecular solvent-accessible surface area^[8,9] or volume.^[10–12] Continuum electrostatics have been extensively studied to estimate the polar component of the solvation effect.^[13–17]

The Poisson–Boltzmann implicit solvent (PB) is one of the major theories using continuum electrostatics.^[18–20] PB is often used to estimate the solvation free energy of the solute; PB estimates the polar component of the solvation effect by solving the Poisson–Boltzmann equation:

$$\nabla \cdot \varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -\rho(\mathbf{r}) - \sum_i q_i c_i \exp[-\beta q_i \phi(\mathbf{r})] \quad (1)$$

where $\varepsilon(\mathbf{r})$, $\phi(\mathbf{r})$, and $\rho(\mathbf{r})$ are the position-dependent dielectric constant, electrostatic potential, and charge density of the solute, respectively, q_i is the charge of the i -th ion type, c_i is the number density of the i -th ion type, and β is the inverse thermal energy. From eq. (1), PB energy is determined by the distribution

of dielectric constants that is often approximated by the dielectric boundary around the solute. The dielectric constants are switched between the high (solvent) and low values (solute) over the dielectric boundary according to the dielectric function.

In early studies on PB, the dielectric boundary was determined directly using van der Waals radii in the force field for molecular mechanics and molecular dynamics (MD) simulations; however, this method was found to be insufficient for the accurate estimation of the PB energy.^[21–24] Recent studies have proposed special collections of atomic radii for PB, called PB radii.^[23–25] Tan et al.^[25] presented PB radii for nucleic acids and proteins based on the atom types and partial charges in the AMBER protein force field.^[26,27] They optimized their PB radii by fitting the results from PB for amino acid and nucleic acid templates to those from explicit solvent simulations using TIP3P water.^[2] The PB radii obtained by Swanson et al.,^[24] based on the AMBER protein force field, were also optimized using the results from explicit solvent simulations using TIP3P water. Swanson et al.^[28] adopted the smoothing dielectric

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

[a] J. Yamagishi, M. Tajji

Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, 5-15 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan

[b] J. Yamagishi, N. Okimoto, G. Morimoto, M. Tajji

Laboratory for Computational Molecular Design, Quantitative Biology Center (QBiC), RIKEN, 1-6-5 Minatojima-Minatomachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan.

E-mail: okimoto@gsc.riken.jp

Contract grant sponsor: JSPS Fellows; Contract grant number: 236408; Contract grant sponsor: "Protein 3000 Project" of the Ministry of Education, Culture, Sports, Science, and Technology of Japan

© 2014 The Authors. Journal of Computational Chemistry Published by Wiley Periodicals, Inc.

function, which can provide accurate and numerically stable calculations of PB, instead of the traditional abrupt dielectric function.

Although the two recent PB methods have the same purpose of reproducing the results from explicit solvent simulations, their results for various proteins are significantly different (see Results and Discussion). We considered that this inconsistency might be attributed to the following points. The first reason is the difference in protocols of the explicit solvent simulations, in which MD-based approaches were used to calculate reference solvation free energies and forces of the template structures. The main difference in the two protocols was the boundary condition of the system. Tan et al.^[25] adopted the periodic boundary condition (PBC) with the particle mesh Ewald (PME) method,^[29] while Swanson et al.^[24] adopted the spherical boundary condition (SBC) with the spherical solvent boundary potential (SSBP).^[30] Even though SSBP was applied to reduce the artificial influence of water near the boundary, it was probable that the reference solvation free energies or forces calculated using SBC differed from those calculated using PBC. These errors in the reference solvation free energies or forces might contribute to the inconsistent results between the two PB methods.

The second reason for the inconsistency is the difference in the selection of training molecules used as template structures for the optimization of PB radii. This selection affects the transferability of PB radii between different sizes and conformations of proteins. Tan et al. used a few dipeptides and side-chain analogues of each amino acid.^[25] Their templates seemed to lack considerations for the protein secondary structure and the interactions between the backbone and side-chain atoms. In contrast, Swanson et al. included polyalanines in their consideration for the secondary structure of proteins and dipeptides of each amino acid in their training molecules.^[24] However, their PB radii were not discriminated among the N-, C-, and nonterminal residues. It could likely correlate with the inaccuracies for terminal residues, because the charge distributions of the terminal residues are different from those of nonterminal residues.

Considering the aspects described above, we propose new PB radii optimized for accurate estimations of the solvation free energies of proteins. To obtain the correct reference solvation free energies and forces from explicit solvent simulations, we apply a cut-off scheme to the calculation of the long-range interactions in the free energy calculation under SBC, which can remove the artificial influence from water near the boundary. In addition, we attempt to improve the accuracy for terminal residues by discriminating PB radii for N- and C-terminal residues from those for nonterminal residues. Details of the optimization of PB radii are given in the following section. The performances using our PB radii are compared with the two PB methods mentioned above in the Results and Discussion section.

Methods

Optimization of PB radii

Our PB radii were based on the partial charges in the AMBER protein force field (ff99SB).^[26] In previous studies,^[22–25] the

atoms in similar chemical environments had an identical PB radius; however, such an approximation can be one of the sources of inaccuracy in PB. In this study, only the four atoms of the peptide bond in nonterminal residues have common PB radii. The four atoms of the peptide bond were classified into four groups: positively charged, negatively charged, non-charged residues, and proline. The PB radii for each atom in the four groups were unified. For the other atoms, every atom had a distinct PB radius, although unified PB radii were assigned to only chemically equivalent atoms such as the hydrogen atoms in methyl groups.

The PB radii were optimized using a genetic algorithm (GA) to obtain good agreement between PB results and the reference solvation free energies and forces obtained from the explicit solvent simulations. The fitness score of GA is represented as follows:

$$SCORE = RMS(\Delta G_{\text{residue}}) + a \times RMS(\Delta G_{\text{atom}}) + b \times RMS(\Delta F^P) \quad (2)$$

where $RMS(\Delta G_{\text{residue}})$ and $RMS(\Delta G_{\text{atom}})$ are the root-mean-square error in the solvation free energy on a per-residue basis and on a per-atom basis as compared with the reference, respectively, and $RMS(\Delta F^P)$ is the root-mean-square error in the solvation force for each dimension of each atom as compared with the reference. We incorporated the atom-based and residue-based terms into the fitness score to reduce the dependence of the PB radius on amino acid sequences. Scale factors a and b were set to 0.5 and 1.0, respectively. The details of GA are as follows: the population size was 600. Five new individuals (children) were generated from each of 600 individuals (parents). Each parent and its children formed a family, and the best individual in each family became the new parent in the next generation. The rates for uniform-crossover and single-point mutation were 0.7 and 0.3, respectively.

Training molecules

The PB radii were optimized using training molecules consisting of multiple conformations of polyalanines and single amino acid-based molecules. Because the conformation of each structure was fixed during the optimization of PB radii, multiple conformations were necessary to enhance the transferability of PB radii between different conformations.

To account for the secondary structures of proteins, we first optimized the PB radii for the alanine residue using 14 polyalanine structures described in the previous study by Swanson et al.^[23] These structures were constructed using the specific regions of two X-ray crystallographic structures (Protein Data Bank (PDB)^[31] ID: 1AKI^[32] and 1EJG^[33]). Both the N- and C-terminal ends of the protein backbone structures were capped with *N*-acetyl (Ace) and *N*-methylamide (NMe) groups. All amino acids were mutated to alanines. All structures were energetically minimized in a box of TIP3P water^[2] using AMBER 12.^[34]

The remaining training molecules consist of the single amino acid-based molecules: nonterminal residues, and residues at N- and C-termini of the single amino acids (Supporting

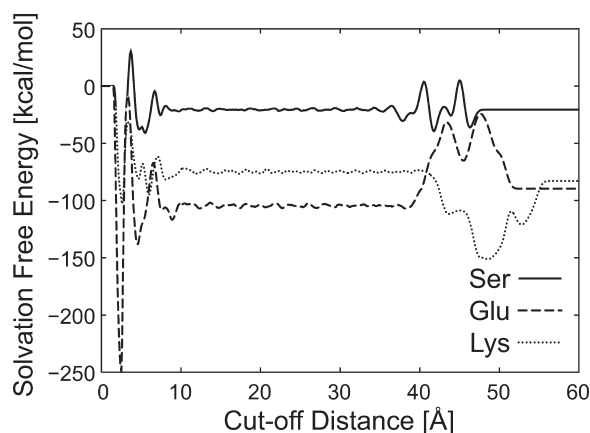


Figure 1. Solvation free energies of three dipeptides calculated by our protocol for explicit solvent simulations with respect to the cut-off distance used in the free energy calculations. Around 45 Å, the solvation free energies are influenced by SBC. The solvation free energies in the range of 20–30 Å (bulk water) and those at infinite distances are significantly different for Glu and Lys.

Information, Fig. S1). Nonterminal ends of each amino acid were capped with Ace and NMe groups. All conformations were generated using MD simulations and systematic conformational searches. After soaking the solutes in a box of TIP3P water, we performed 10 ns MD simulations at 300 K using AMBER 12.^[35] Additional systematic conformational searches were carried out with the generalized Born solvent^[36] using the Molecular Operating Environment (MOE) software package.^[37] More than three representative conformations of each residue were selected as training molecules from the MD trajectories and the results of conformational searches.

Test molecules

To measure the performances beyond the training molecules, we used 23 peptide structures of varying lengths. The 12 experimentally determined structures^[38–49] were obtained from PDB. The other 11 structures consisted of a wide variety of conformations of the mini-protein chignolin, and were generated by replica-exchange MD^[50] in our previous study^[51] from the PDB structure (ID: 1UAO^[52]). The protonation states of all structures were determined by the Protonate 3D module of MOE.^[37] All structures were energetically minimized in a box of TIP3P water.^[2] Detailed information of our test molecules is listed in Supporting Information, Table S1.

Explicit solvent simulations

We used the thermodynamic integration (TI) method^[53] to calculate the polar component of the solvation free energy of the solute using TIP3P water.^[2] The polar solvation free energy ΔG_{pol} is represented as:

$$\Delta G_{\text{pol}} = \Delta G_{\text{chg, wat}} - \Delta G_{\text{chg, vac}} \quad (3)$$

where $\Delta G_{\text{chg, wat}}$ and $\Delta G_{\text{chg, vac}}$ are the charging free energy of the solute in water and in vacuo, respectively. Each charging free energy was calculated using the TI method based on the

15 λ -point Gaussian quadrature. Initial structures were set up by locating the solute at the center of a sphere of TIP3P water. The radii of the solvent spheres for the training molecules and for the test molecules were 45 Å and 53 Å, respectively; they are sufficiently large for the calculation of the solvation free energy. Our free energy calculation consists of two separate steps. The first step is the conventional MD simulation for conformational sampling. At each λ point, we performed 500 ps MD for the equilibration and another 500 ps MD for the production. No cut-off scheme was employed for the conformational sampling by MD simulations. Snapshots were sampled every 20 fs in the production run, and a total of 25,000 snapshots were saved at every λ point. The Langevin thermostat^[54] was used to maintain the temperature of the system at 300 K. All solute atoms were harmonically restrained at their initial positions with a force constant of 50 kcal mol⁻¹ Å⁻¹. The second step is to perform the TI method using the MD trajectories obtained from the first step for the calculation of the solvation free energy. In this step, we applied a cut-off scheme for long-range interactions to remove the artificial influence from water molecules near the boundary. Because the free energy calculated using a given cut-off distance is largely fluctuated (Fig. 1), we adopted the average of the solvation free energies using multiple cut-off distances. The cut-off distances ranged from 22 to 30 Å and 28 to 36 Å for the training molecules and test molecules, respectively, with a step size of 0.1 Å. The dynamics of the water molecules in these ranges fully corresponded to those of bulk water molecules.

Polar solvation forces were also calculated using the explicit solvent simulations. As described by Wagoner,^[55] a polar component of the solvation force F^P is represented by:

$$F^P = \overline{F^{P+np}} - \overline{F^{np}} \quad (4)$$

where $\overline{F^{np}}$ is the averaged nonpolar force over an ensemble generated by MD with the noncharged solute, and $\overline{F^{P+np}}$ is the averaged total force (polar and nonpolar force) over an ensemble generated by MD with the fully charged solute. The simulation conditions of the two MD simulations were the same as those for the solvation free energy except for the handling of solute partial charges.

All explicit solvent simulations were carried out using AMBER 10^[35] modified for use on the special-purpose computer MD-GRAPE3.^[56,57]

Implicit solvent simulations

Our calculation conditions for PB were the same as those of Swanson et al.^[23,24] The nonlinear PB equation was solved using the Adaptive Poisson–Boltzmann Solver (APBS version 1.3).^[58] The PB grid spacing was 0.20 Å. The number of grid points in each dimension was determined as more than 10 Å larger than the size of the solute. The solute charges were distributed to the PB grid using a cubic B-spline discretization. The spline-smoothing dielectric function was used with a half-smoothing window of 0.3 Å.^[28] The dielectric constants inside and outside the solute were 1.0 and 78.4, respectively. The

Table 1. Solvation free energies calculated from explicit solvent simulations

Protocols ^[a]	Solvation free energy (kcal mol ⁻¹)		
	Ser	Glu	Lys
Tan et al. ^[b] (PBC with PME)	-20.18	-103.51	-74.18
Swanson et al. ^[c] (SBC with SSBP)	-20.40	-91.74	-84.95
Our protocol (SBC with cut-off)	-20.52	-104.58	-74.88
Our protocol (SBC without cut-off) ^[d]	-20.52	-89.64	-82.86

[a] The boundary conditions and the treatments of the long-range interactions used in the explicit solvent simulations are listed. [b] Ref. [25]. [c] Ref. [24]. [d] The difference between our two protocols was whether the cut-off scheme was applied to the TI method or not.

probe size of the water molecule was 1.4 Å. The bulk concentration of mobile ions was set at zero. To reduce dependence on the orientation of the solute to the PB grid, we prepared another orientation of each solute and averaged their results.

For comparison, we evaluated two other PB methods^[24,25] described in the Introduction section, which are commonly used with the AMBER force field.^[26] Tan's PB^[25] was carried out using an mm_pbsa.pl script in AMBER 12.^[35] The modifiable parameters in the script were set as follows: the dielectric constants inside and outside the solute were 1.0 and 80.0, respectively. The probe radius of the water molecule was set to 1.6 Å. The grid spacing was 0.333 Å. The calculation condition of Swanson's PB^[24] was the same as ours.

Results and Discussion

Optimization of PB radii

Prior to PB radius optimizations, we evaluated our protocol for explicit solvent simulations toward accurate estimation of reference solvation free energies. The solvation free energies of three dipeptides with different net charges were calculated using four protocols (Table 1). Our and Tan's protocol^[25] showed a similar tendency of the calculated solvation free energies, on the other hand, the solvation free energies calculated by Swanson's protocol^[23,24] were in disagreement with those calculated by our and Tan's one. These results were consistent with the previous study on the solvation free energies of ions.^[59] For a comparison purpose, we also calculated the solvation free energies using the same protocol for the MD simulation (for conformational sampling) as ours but without using the cut-off scheme in the TI calculations. These solvation free energies were different from those calculated by our protocol using the cut-off scheme (Table 1). Figure 1 clearly shows that these differences in the solvation free energies were caused by an artifact in the distribution of water molecules near the boundary. From the graph, applying the cut-off scheme was proven to be effective in removing the artifact introduced by SBC. Furthermore, Swanson's protocol and our protocol without the cut-off scheme showed similar tendencies in underestimating the solvation free energies for nega-

Table 2. Statistical performances on training molecules and test molecules for three PB methods

Methods	Mean and standard deviation of absolute errors ^[a] (kcal mol ⁻¹)	
	Training molecules	Test molecules
Tan et al. ^[b]	1.34 ± 1.79	21.24 ± 9.32
Swanson et al. ^[c]	8.72 ± 7.40	29.85 ± 13.04
Our PB	0.32 ± 0.31	2.43 ± 2.39

[a] Absolute errors denote the absolute difference between the total solvation free energies calculated by PB and the explicit solvent simulations. [b] Ref. [25]. [c] Ref. [24].

tively charged residues and overestimating those for positively charged residues, as compared with those calculated by the other two protocols (Table 1). These tendencies were caused by biased orientations of the oxygen and hydrogen atoms of water molecules near the boundary; that is, it is suggested that SSBP^[30] used in Swanson's protocol cannot fully remove the artifact introduced by SBC. From the above results, both our and Tan's protocols seem to provide accurate reference solvation free energies. However, the recent research reported that solvation free energies calculated using PBC with PME were dependent on the box size of PBC and the net charge of solute,^[60] which indicates that Tan's protocol is less reliable for the estimation of reference solvation free energies. Therefore, we chose our protocol from the point of view of the reliability of the protocol and our computational resource. Our results also suggested that the PB radii obtained by Swanson et al. were strongly influenced by errors in the reference solvation free energies.

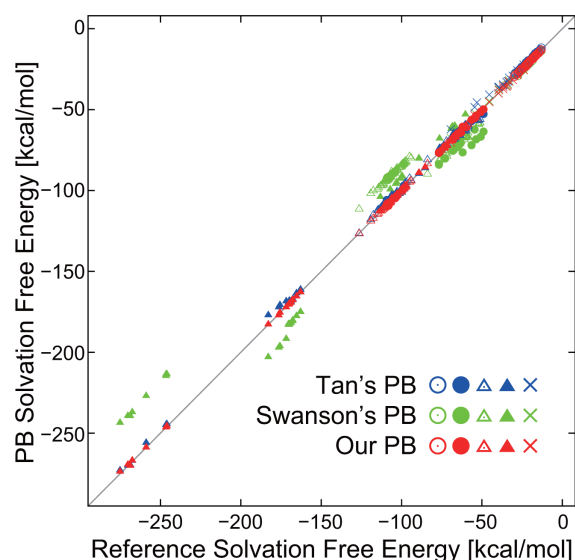


Figure 2. Correlation between the solvation free energies of training molecules calculated by PB and explicit solvent simulations. Circles are the non-terminal residues, whereas triangles are the N- or C-terminal residues. Open marks are the residues having noncharged side-chains, whereas filled marks are those having charged side-chains. The × labels correspond to polyalanines.

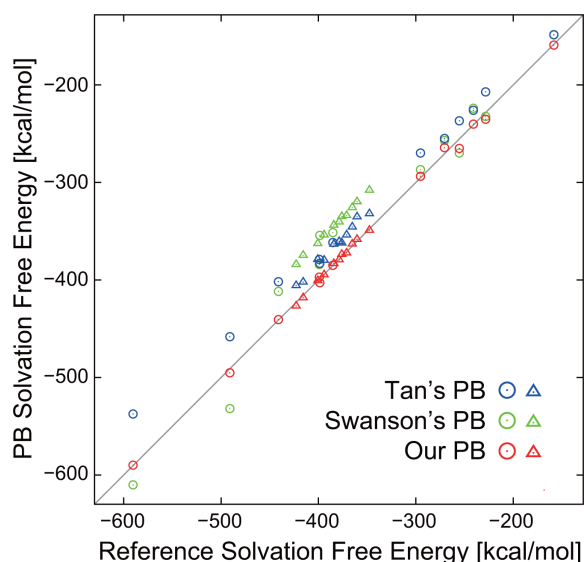


Figure 3. Correlation between the solvation free energies of test molecules calculated by PB and explicit solvent simulations. Circles are the 12 peptides from PDB, and triangles are the 11 conformations of chignolin.

Table 2 summarizes the statistical performances of our PB and other PB methods on the training molecules. Both the low mean and low standard deviation of absolute errors for our PB indicated that our optimization of PB radii was successfully accomplished. Note that the comparisons of statistical performances on simple training molecules between our PB and other PB methods are not so meaningful in themselves; however, it is useful to identify sources of errors in the solva-

tion free energies for large peptides. From Figure 2, we can see inaccurate solvation free energies calculated by Swanson's PB for charged molecules, as expected from the inaccuracy of the reference solvation free energies. Their inaccuracy for terminal residues suggested that PB radii for N- and C-terminal residues should be distinguished from those for nonterminal residues. In contrast, Tan's PB showed high accuracy for our training molecules; Tan's results suggested that the dependency of the solvation free energies to the box size of PBC and the net charge of the solute was less problematic for small training molecules.

Evaluation of solvation free energies for large peptides

The solvation free energies for molecules larger than the training molecules were further examined, because our PB radii were optimized using simple training molecules. The performances of the three PB methods on our test molecules are given in Table 2 and Figure 3. These results clearly identified our PB as the most accurate method among the three PB methods examined. In addition to the total solvation free energies, our PB showed good accuracy for the estimation of solvation free energies on a per-residue basis (Fig. 4 for nonterminal residues, and Supporting Information, Figs. S2–S3 for N- and C-terminal residues). In these graphs, Swanson's PB showed large errors in solvation free energies on a per-residue basis, especially for charged residues, as expected from the results for training molecules. Some errors in the total solvation free energy may be compensated by the underestimated and overestimated solvation free energies on each residue for Swanson's PB; this degree of compensations is dependent on

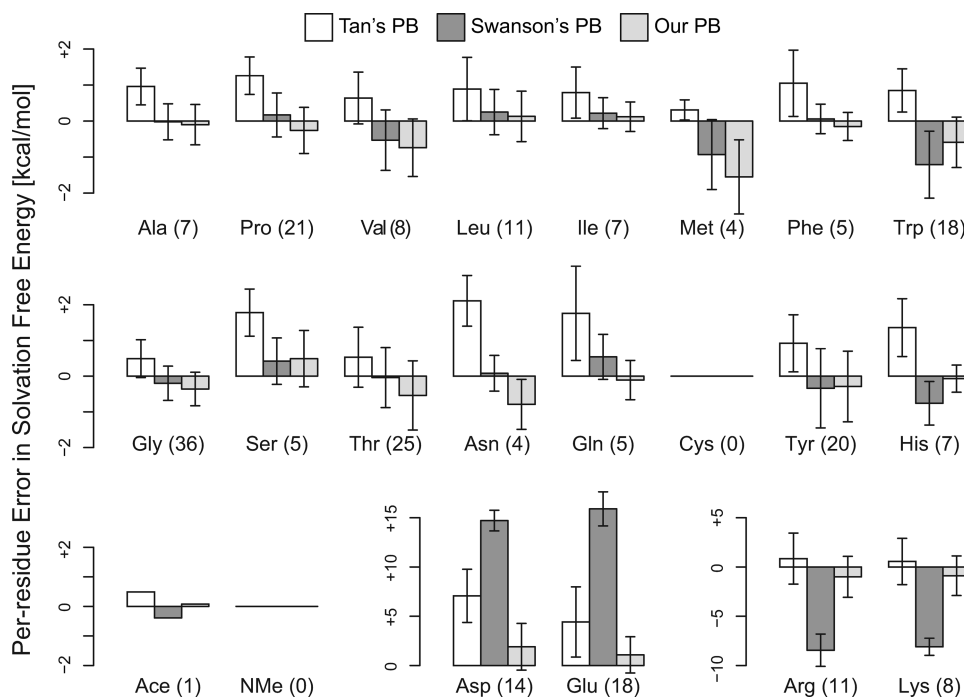


Figure 4. Per-residue errors in the solvation free energies on nonterminal residues in test molecules as compared with the reference solvation free energies. The total numbers of residues included in tested molecules are given in parentheses. The data for the N- and C-terminal residues are given in Supporting Information, Figures S2 and S3.

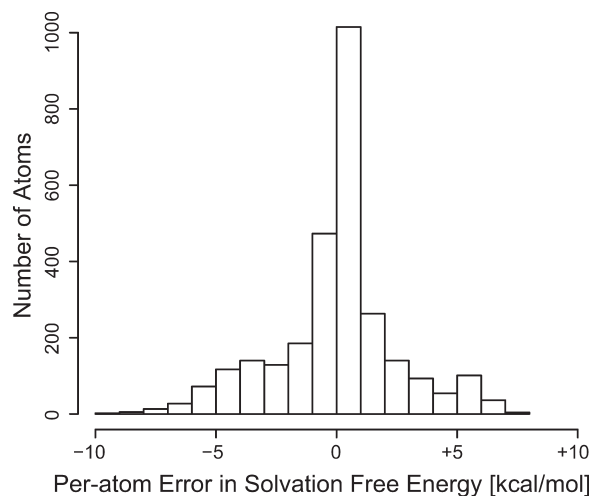


Figure 5. Distributions of per-atom errors in the solvation free energies in test molecules as compared with the reference solvation free energies.

the amino acid sequences of proteins. Thus, it is difficult when using Swanson's PB to compare the results from proteins with different amino acid sequences. Unlike the results for training molecules, Tan's PB tended to underestimate both the total and per-residue solvation free energies for the test molecules. This tendency may originate from the training molecules used in their optimization of PB radii; their PB radii were parameterized without consideration of the interactions between backbone and side-chain atoms. In addition, they lacked consideration for the secondary structures of proteins. Their insufficient treatment for the selection of training molecules would thus induce estimation inaccuracies for larger peptides.

It is difficult to demonstrate performances on molecules much larger than our test molecules due to the limitations of computational resources, in which a huge memory space is

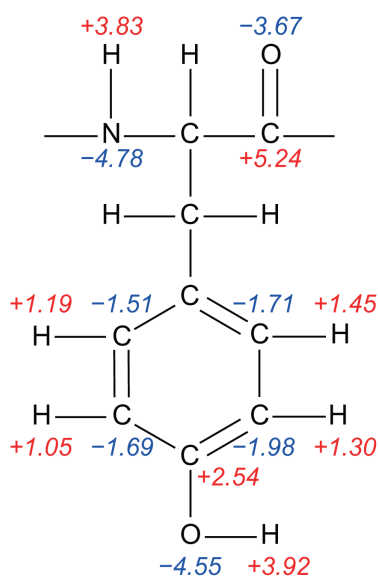


Figure 6. Example of compensations of per-atom errors in the solvation free energies within small molecular fragments. Per-atom errors for the tyrosine residue from PDB ID: 2DX4^[46] are illustrated. Small errors (under ± 1.0 kcal mol⁻¹) are omitted for clarity.

required for MD simulations of large isolated systems; however, the results suggested that our PB would show better performance than the other PB methods discussed.

Analysis at the atomic level

Further analysis of our PB results at the atomic level can provide more useful and interesting information. Figure 5 shows the distribution of errors in the solvation free energies on a per-atom basis, indicating relatively large errors on a per-atom basis as compared with those on a per-residue basis; large errors over ± 3.0 kcal mol⁻¹ were concentrated on atoms in polar or charged functional groups, while moderate errors ($\approx \pm 2.0$ kcal mol⁻¹) were also observed for atoms in nonpolar groups. Nevertheless, the errors on a per-residue basis for our PB were quite small. This can be attributed to compensations of errors within a small fragment or a functional group of the molecule. For example (Fig. 6), the error on the hydrogen atom in the hydroxyl group was almost compensated by that on the neighboring oxygen atom. Similarly, the errors on aromatic hydrogen atoms were compensated by those on the neighboring carbons. These data indicated that our PB can provide reliable results at the level of the small molecular fragment or the functional group.

Detailed analysis of PB radii

Finally, we address the comparison of our PB radii with those obtained by Swanson et al. Both radii were optimized using the same protocol to solve the PB equation.^[24] As expected from the solvation free energies in TIP3P water (Table 1), our and Swanson's PB radii for noncharged residues were almost similar each other. In contrast, large differences were found in PB radii for charged residues. More specifically, our PB radii for negatively charged residues became smaller, while those for positively charged residues became larger than Swanson's radii. We think that these differences mainly originated from the difference in the treatments of the boundary condition in the explicit solvent simulations.

Too small radii may be problematic because they cause small gaps in the solute interior; these gaps correspond to regions with high dielectric constants even if the water molecules do not occupy these spaces. The presence of these small gaps is known to induce errors in the solvation free energies and forces on atoms around the gap.^[24] Interestingly, such small PB radii are not found in Tan's radii, which were optimized using a protocol differing from ours to solve the PB equation.^[25] It is not clear yet why such differences in PB radii were obtained; future analyses are required to identify the reasons for this observation and to provide useful information for solving the problem.

Conclusions

In this study, we have presented new PB radii in order to improve the accuracy of PB for the estimation of solvation free energies of proteins. Based on previous studies, our PB radii were optimized using the smoothing dielectric function and

the AMBER protein force field. The radii were optimized via fitting the results from PB to those obtained by explicit solvent simulations using amino acid templates. The cut-off scheme was applied to the free energy calculations in explicit solvent simulations to remove the artificial influence of water near the spherical boundary; reliable reference solvation free energies and forces for the template structures could be thus obtained. In addition, we discriminated PB radii for N- and C-terminal residues from those for nonterminal residues. From performance evaluations on large peptides using the newly obtained PB radii, our PB showed a high estimation accuracy of solvation free energies not only at the molecular level, but also at the level of the small molecular fragment.

In light of the data presented here, our next challenge will involve the application of the above described PB to the design of novel peptide binders for the regulation of peptide-protein complex formation. The accurate estimation of solvation free energies in peptide design is highly essential, as proteins and peptides have far more polar or charged moieties than small organic molecules. From the results of our study, we believe that our PB is effective for peptide design.


Modifying a set of atomic radii will also be efficient in improving the accuracy of other implicit solvent models, such as the generalized Born implicit solvent (GB).^[14,17,61] Because GB is based on simple radii such as Bondi^[62] or its modifications,^[61] a detailed set of appropriate radii would greatly improve the estimation accuracy of solvation free energies and forces. In addition, our future study aims to improve on the drawback of too small or too large PB radii; the optimization of PB radii based on Tan's PB^[25] may be helpful for this problem since such radii are not found in Tan's PB radii. These studies are currently underway in our laboratories and will be presented in due course.

Acknowledgment

The calculations were performed using the RIKEN Integrated Cluster of Clusters (RICC) facility.

Keywords: Poisson-Boltzmann implicit solvent · implicit solvent · continuum electrostatics · solvation-free energy · molecular modeling

How to cite this article: J. Yamagishi, N. Okimoto, G. Morimoto, M. Tajji. *J. Comput. Chem.* **2014**, *35*, 2132–2139. DOI: 10.1002/jcc.23728

 Additional Supporting Information may be found in the online version of this article.

- [1] M. W. Mahoney, W. L. Jorgensen, *J. Chem. Phys.* **2000**, *112*, 8910.
- [2] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
- [3] J. Wang, P. Morin, W. Wang, P. A. Kollman, *J. Am. Chem. Soc.* **2001**, *123*, 5221.
- [4] C. S. Rapp, R. A. Friesner, *Proteins* **1999**, *35*, 173.
- [5] B. Kuhn, P. Gerber, T. Schulz-Gasch, M. Stahl, *J. Med. Chem.* **2005**, *48*, 4040.
- [6] D. C. Thompson, C. Humblet, D. Joseph-McCarthy, *J. Chem. Inf. Model* **2008**, *48*, 1081.
- [7] N. Okimoto, N. Futatsugi, H. Fujii, A. Suenaga, G. Morimoto, R. Yanai, Y. Ohno, T. Narumi, M. Tajji, *PLoS Comput. Biol.* **2009**, *5*, e1000528.
- [8] D. Eisenberg, A. D. McLachlan, *Nature* **1986**, *319*, 199.
- [9] K. Sharp, A. Nicholls, R. Fine, B. Honig, *Science* **1991**, *252*, 106.
- [10] R. M. Levy, L. Y. Zhang, E. Gallicchio, A. K. Felts, *J. Am. Chem. Soc.* **2003**, *125*, 9523.
- [11] C. Tan, Y.-H. Tan, R. Luo, *J. Phys. Chem. B* **2007**, *111*, 12,263.
- [12] A. Wallqvist, B. J. Berne, *J. Phys. Chem.* **1995**, *99*, 2893.
- [13] F. Floris, J. Tomasi, *J. Comput. Chem.* **1989**, *10*, 616.
- [14] G. D. Hawkins, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem.* **1996**, *100*, 19,824.
- [15] A. Klamt, G. Schuurmann, *J. Chem. Soc. Perkin 2* **1993**, 799.
- [16] A. W. Lange, J. M. Herbert, *J. Chem. Phys.* **2011**, *134*, 204110.
- [17] A. Onufriev, D. Bashford, D. A. Case, *J. Phys. Chem. B* **2000**, *104*, 3712.
- [18] B. Honig, A. Nicholls, *Science* **1995**, *268*, 1144.
- [19] I. Klapper, R. Hagstrom, R. Fine, K. Sharp, B. Honig, *Proteins* **1986**, *1*, 47.
- [20] J. Warwicker, H. C. Watson, *J. Mol. Biol.* **1982**, *157*, 671.
- [21] R. Zhou, B. J. Berne, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12,777.
- [22] D. Sitkoff, K. A. Sharp, B. Honig, *J. Phys. Chem.* **1994**, *98*, 1978.
- [23] J. M. J. Swanson, S. A. Adcock, J. A. McCammon, *J. Chem. Theory Comput.* **2005**, *1*, 484.
- [24] J. M. J. Swanson, J. A. Wagoner, N. A. Baker, J. A. McCammon, *J. Chem. Theory Comput.* **2007**, *3*, 170.
- [25] C. H. Tan, L. J. Yang, R. Luo, *J. Phys. Chem. B* **2006**, *110*, 18,680.
- [26] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins* **2006**, *65*, 712.
- [27] 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, *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- [28] W. Im, D. Beglov, B. Roux, *Comput. Phys. Commun.* **1998**, *111*, 59.
- [29] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10,089.
- [30] D. Beglov, B. Roux, *J. Chem. Phys.* **1994**, *100*, 9050.
- [31] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235.
- [32] P. J. Artymiuk, C. C. F. Blake, D. W. Rice, K. S. Wilson, *Acta Crystallogr. Sect. B* **1982**, *38*, 778.
- [33] C. Jelsch, M. M. Teeter, V. Lamzin, V. Pichon-Pesme, R. H. Blessing, C. Lecomte, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3171.
- [34] D. A. Case, T.A. Darden, I. T.E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossvary, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Q. C. Gohlke, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. A. Kollman, AMBER 12; University of California: San Francisco, **2012**.
- [35] D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossvary, F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, AMBER, 10-12; University of California: San Francisco, **2012**.
- [36] P. Labute, *J. Comput. Chem.* **2008**, *29*, 1693.
- [37] *Molecular Operating Environment (MOE), 2013.10*; Chemical Computing Group Inc.: Montreal, QC, Canada, **2013**.
- [38] I. Marcotte, F. Separovic, M. Auger, S. M. Gagné, *Biophys. J.* **2004**, *86*, 1587.
- [39] A. Fisher, P. B. Laub, B. S. Cooperman, *Nat. Struct. Biol.* **1995**, *2*, 951.
- [40] G. A. Spyroulias, P. Nikolakopoulou, A. Tzakos, I. P. Gerotheranassis, V. Magafa, E. Manessi-Zoupa, P. Cordopatis, *Eur. J. Biochem.* **2003**, *270*, 2163.
- [41] A. Phan-Chan-Du, M.-C. Petit, G. Guichard, J.-P. Briand, S. Muller, M. T. Cung, *Biochemistry* **2001**, *40*, 5720.
- [42] S. Bhattacharjya, P. N. Domadia, A. Bhunia, S. Malladi, S. A. David, *Biochemistry* **2007**, *46*, 5864.
- [43] M. Wang, L. Shan, J. Wang, *Biopolymers* **2006**, *83*, 268.

- [44] H. H. Jung, H. J. Yi, S. K. Lee, J. Y. Lee, H. J. Jung, S. T. Yang, Y.-J. Eu, S.-H. Im, J. I. Kim, *Biochemistry* **2007**, *46*, 14,987.
- [45] M. O. Steinmetz, I. Jelesarov, W. M. Matousek, S. Honnappa, W. Jahnke, J. H. Missimer, S. Frank, A. T. Alexandrescu, R. A. Kammerer, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7062.
- [46] M. Araki, A. Tamura, *Proteins* **2007**, *66*, 860.
- [47] J. W. Neidigh, R. M. Fesinmeyer, N. H. Andersen, *Nat. Struct. Biol.* **2002**, *9*, 425.
- [48] N. Page, N. Schall, J. M. Strub, M. Quinteret, O. Chaloin, M. Decossas, M. T. Cung, A. Van Dorsselaer, J. P. Briand, S. Muller, *PLoS One* **2009**, *4*, e5273.
- [49] R. Zahn, *J. Mol. Biol.* **2003**, *334*, 477.
- [50] Y. Sugita, Y. Okamoto, *Chem. Phys. Lett.* **1999**, *314*, 141.
- [51] H. X. Kondo, M. Taiji, *J. Chem. Phys.* **2013**, *138*, 244113.
- [52] S. Honda, K. Yamasaki, Y. Sawada, H. Morii, *Structure* **2004**, *12*, 1507.
- [53] T. P. Straatsma, J. A. McCammon, *J. Chem. Phys.* **1991**, *95*, 1175.
- [54] R. W. Pastor, B. R. Brooks, A. Szabo, *Mol. Phys.* **1988**, *65*, 1409.
- [55] J. Wagoner, N. A. Baker, *J. Comput. Chem.* **2004**, *25*, 1623.
- [56] T. Narumi, Y. Ohno, N. Okimoto, T. Koishi, A. Suenaga, N. Futatsugi, R. Yanai, R. Himeno, S. Fujikawa, M. Ikei, M. Taiji, In Proceedings of the 2006 ACM/IEEE conference on Supercomputing; ACM/IEEE: Tampa, FL, **2006** (CD-ROM).
- [57] M. Taiji, T. Narumi, Y. Ohno, N. Futatsugi, A. Suenaga, N. Takada, A. Konagaya, In Proceedings of the 2003 ACM/IEEE conference on Supercomputing; ACM/IEEE: Phoenix, AZ, **2003** (CD-ROM).
- [58] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, *Proc Natl Acad Sci U S A* **2001**, *98*, 10037-10041.
- [59] T. Darden, D. Pearlman, L. G. Pedersen, *J. Chem. Phys.* **1998**, *109*, 10,921.
- [60] G. J. Rocklin, D. L. Mobley, K. A. Dill, P. H. Hünenberger, *J. Chem. Phys.* **2013**, *139*, 184103.
- [61] A. Onufriev, D. Bashford, D. A. Case, *Proteins* **2004**, *55*, 383.
- [62] A. Bondi, *J. Phys. Chem.* **1964**, *68*, 441.

Received: 14 March 2014

Revised: 4 July 2014

Accepted: 20 August 2014

Published online on 15 September 2014