Showcasing research from the Group of Prof. Norio Yoshida and Prof. Haruyuki Nakano at Kyushu University, Japan

A computational scheme of $pK_a$ values based on the three-dimensional reference interaction site model self-consistent field theory coupled with the linear fitting correction scheme

This work proposes a scheme for quantitatively computing the $pK_a$ of hydrated molecules. It is based on 3D-RISM-SCF theory coupled with the linear fitting correction (LFC) scheme. In LFC/3D-RISM-SCF, $pK_a$ values are evaluated using the Gibbs energy difference between the protonated and unprotonated states and the parameters fitted to the experimental values of training set systems. The computed $pK_a$ values show quantitative agreement with the experimental data.

As featured in:
A computational scheme of pK\textsubscript{a} values based on the three-dimensional reference interaction site model self-consistent field theory coupled with the linear fitting correction scheme†

Ryo Fujiki,\textsuperscript{a} Yukako Kasai,\textsuperscript{a} Yuki Seno,\textsuperscript{a} Toru Matsui,\textsuperscript{b} Yasuteru Shigeta,\textsuperscript{c} Norio Yoshida\textsuperscript{d} \textsuperscript{**} and Haruyuki Nakano\textsuperscript{c} \textsuperscript{a}

A scheme for quantitatively computing the acid dissociation constant, pK\textsubscript{a} of hydrated molecules is proposed. It is based on the three-dimensional reference interaction site model self-consistent field (3D-RISM-SCF) theory coupled with the linear fitting correction (LFC) scheme. In LFC/3D-RISM-SCF, pK\textsubscript{a} values of target molecules are evaluated using the Gibbs energy difference between the protonated and unprotonated states calculated by 3D-RISM-SCF and the parameters fitted by the LFC scheme to the experimental values of training set systems. The pK\textsubscript{a} values computed by LFC/3D-RISM-SCF show quantitative agreement with the experimental data.

1. Introduction

Protonation and deprotonation are fundamental chemical reactions in solution and biological systems. Through these reactions, molecules change their charged states and interactions between surrounding environments. Such changes play an essential role in the solubility of molecules, higher-order structure formation of proteins, molecular recognition, and molecular transportation across membranes. An equilibrium constant of the reaction, an acid dissociation constant, K\textsubscript{a}, or its more commonly used logarithmic value, pK\textsubscript{a} = −\log_{10} K\textsubscript{a}, can usually be determined experimentally, by titration, for small molecules. The pK\textsubscript{a} values are strongly affected by the surrounding environment such as solvent and protein. Therefore, the pK\textsubscript{a} value or the protonation state of an amino acid residue often changes drastically.\textsuperscript{1–6} In practice, the protonation state of dissociative amino acid residues is measured by the neutron diffraction method or the nuclear magnetic resonance method.\textsuperscript{7–10} However, these methods have some disadvantages. The former requires a large crystal of the target protein; in the latter, it is difficult to specify the residue of the signal origin of proteins that have multiple dissociative residues. Therefore, because of the experimental difficulties, a theoretical method to compute pK\textsubscript{a} values has attracted considerable attention.

The pK\textsubscript{a} value is closely related to the Gibbs energy difference of the acid dissociation reaction, \Delta G, according to the relationship

\[
pK_a = \frac{\Delta G}{(\ln 10)RT}
\]  

(1)

where R and T are the gas constant and absolute temperature, respectively, and

\[
\Delta G = G(A^-) + G(H^+) - G(HA)
\]  

(2)

where \(G[X]\) denotes a Gibbs energy of species X. Here, HA and A\textsuperscript{-} represent the protonated and unprotonated states of an acid A. A straightforward way to calculate the Gibbs energy of a solvated molecule involves \textit{ab initio} molecular-dynamics-based methods.\textsuperscript{11–14} However, these methods require substantial computational costs, and it is therefore impractical to apply them to complex molecular systems. A more compact method to handle the Gibbs energy of solvated molecules is the hybrid quantum mechanics and molecular mechanics (QM/MM) method, which is commonly used for pK\textsubscript{a} evaluation.\textsuperscript{15–21} In this approach, only the reactive moiety is treated by the \textit{ab initio} molecular orbital (MO) or Kohn–Sham density functional theory (KS-DFT) and the remaining parts are treated by classical molecular mechanics. Even more compact methods in computational cost are hybrid methods with the implicit solvation models such as the polarizable continuum model (PCM), and the statistical mechanics integral equation theory of liquids, such as reference interaction site model (RISM), or three-dimensional RISM

---

\textsuperscript{a} Department of Chemistry, Graduate School of Science, Kyushu University, 744, Motooka, Nishi-ku, Fukuoka, 819-0395, Japan. E-mail: noriwo@chem.kyushu-univ.jp

\textsuperscript{b} Department of Chemistry, Graduate School of Pure and Applied Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8577, Japan

\textsuperscript{c} Center for Computational Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8577, Japan

\textsuperscript{d} Electronic supplementary information (ESI) available. See DOI: 10.1039/c8cp04354j
(3D-RISM) theory. These methods produce qualitatively good solvation free energies within a reasonable computational time. However, for the quantitative evaluation of pK\textsubscript{a}, computing the Gibbs energy of the proton, G(H\textsuperscript{+}), is problematic, because the number of water molecules surrounding the excess proton to be handled by quantum mechanics and the structure of the excess proton–water cluster are unclear. Many theoretical approaches assume that the excess proton exists as the hydroxium ion, H\textsubscript{3}O\textsuperscript{+}, and that such a simple treatment may cause an error in the pK\textsubscript{a} value. Therefore, empirical approaches are widely employed to evaluate the pK\textsubscript{a} value of amino acid residues. However, because such methods employ empirical parameters, there are reports that the methods sometimes produce unreliable values.

To achieve both low computational cost and high accuracy, Matsui et al. proposed a scheme based on the linear relationship between the pK\textsubscript{a} and the Gibbs energy difference between HA and \textit{A}\textsuperscript{−}. They suggested several types of methods with different approximation levels. In the present paper, we refer to the linear fitting correction (LFC) scheme as a generic term.

In the LFC scheme, pK\textsubscript{a} values of target molecules are evaluated using parameters fitted by the least squares method to the experimental values of training set systems. Using this scheme, the computation of the Gibbs energy of a dissociated proton in solution can be circumvented. This scheme has been successfully applied to the evaluation of the pK\textsubscript{a} values of amino acids. The results were in good agreement with the experimental observations. The original LFC scheme employs the PCM to take into account the solvent effect on the electronic structure. The PCM and related methods are widely used to reproduce the local molecular interactions, such as hydrogen bonding, and to define the dielectric constant of a heterogeneous environment, such as inside a protein.

In this paper, we propose a new scheme based on the LFC scheme employing the 3D-RISM as a solvent model. Hybrid methods of the 3D-RISM theory and the quantum chemical theory, such as KS-DFT and \textit{ab initio} MO, have been proposed by Kovalenko, Sato, and Hirata. These methods are referred to as KS-DFT/3D-RISM or three-dimensional reference interaction site model self-consistent field (3D-RISM-SCF). 3D-RISM-SCF has been applied to various chemical processes in a solution, including the pK\textsubscript{a} shift of drug molecules. The method allows us to treat a highly anisotropic solvent environment, such as inside a cavity and channel of a protein. Therefore, employing 3D-RISM-SCF, we expect to establish a method that is applicable to complex biological systems.

We first determined the parameters for the LFC scheme by a least square fitting based on the Gibbs energy of the training set molecules calculated by 3D-RISM-SCF and the corresponding experimental pK\textsubscript{a} value. We also examined the basis set dependency on the performance of the scheme. The scheme was applied to amino acids to assess the transferability of the fitted parameters.

2. Method

Linear fitting correction method with empirical parameters

The pK\textsubscript{a} value is related to the Gibbs energy difference of the acid dissociation reaction, \(\Delta G\), according to eqn (1). In the LFC scheme, eqn (1) is rewritten by introducing the scaling factor \(s\):

\[
pK_a = \frac{s(G(A^-) - G(HA))}{(\ln 10)RT} + \frac{sG(H^+)}{(\ln 10)RT} = k\Delta G_0 + C_0
\]

where

\[
k = \frac{s}{(\ln 10)RT}
\]

\[
\Delta G_0 = G(A^-) - G(HA)
\]

\[
C_0 = \frac{sG(H^+)}{(\ln 10)RT}
\]

The scaling factor \(s\) should be unity when the calculated Gibbs energy values are identical to exact values, and \(k = 0.733\) mol kcal\textsuperscript{−1} when \(s = 1\) at 298.15 K. The scaling factor \(s\) is an adjustable parameter, which corresponds to the activity coefficient of deprotonation reaction and corrects the systematic error of the computational method. The parameters \(k\) and \(C_0\) were determined by the least square fitting to minimize the errors of pK\textsubscript{a} values:

\[
e = \sum_i (pK_{a,i}^\text{exp} - (k\Delta G_{0,i} + C_0))^2
\]

where \(pK_{a,i}^\text{exp}\) is an experimental pK\textsubscript{a} value of molecule \(i\) and the summation over \(i\) is taken for all molecules in the training set that have the same dissociative chemical group and those pK\textsubscript{a} values are known. \(\Delta G_{0,i}\) is evaluated using \textit{ab initio} MO or KS-DFT with a solvation model such as the PCM. The parameters \(k\) and \(C_0\) are determined for each of the dissociative chemical groups, such as carboxyl, amine, alcohol, thiol, phenol, and imidazole.

3D-RISM-SCF theory

In the original LFC methods, the PCM is employed as a solvation model for \(\Delta G_0\) calculation. In the present study, we employed 3D-RISM-SCF instead of the PCM to take the solvation effect into account. As the details of the 3D-RISM-SCF method can be found in the literature, we only provide a brief explanation of the theory here.

The Gibbs energy of the solute molecule in the solvent at infinite dilution is defined as the sum of the solute electronic energy \((E_0)\), solvation free energy \((\Delta \mu)\), and the kinetic free energy \((G_{\text{kin}})\):

\[
G = E_0 + \Delta \mu + G_{\text{kin}}
\]

where \(E_0\) is given by

\[
E_0 = \langle \Psi | \hat{H}_0 | \Psi \rangle
\]

and where \(\hat{H}_0\) and \(\Psi\) denote the Hamiltonian of the isolated molecules and the electronic wave function of solute molecules. The kinetic free energy, \(G_{\text{kin}}\), includes the vibrational, rotational
and translational energies, which are obtained in a usual quantum mechanics manner after the normal mode analysis. In the present study, we ignore the kinetic term, $G_{\text{kin}}$, because the change in this term due to the deprotonation reaction is rather small and the adjustable parameter can absorb the error emerging from this approximation. However, it is noted that one cannot neglect this term in the case that significant geometry variations occur in the reaction. The solvation free energy is given by

$$\Delta \mu = k_{B}T \sum_{i} \rho_{i} \left\{ \frac{1}{2} h_{i}(r)^{2} \Theta(-h_{i}(r)) - c_{i}(r) - \frac{1}{2} h_{i}(r)c_{i}(r) \right\} dr$$

(8)

where $i$ runs over the solvent interaction sites. $\Theta$, $k_{B}$, $T$, and $\rho_{i}$ denote the Heaviside step function, the Boltzmann constant, the absolute temperature, and the number density of the solvent site $i$, respectively. $h_{i}(r)$ and $c_{i}(r)$ are total and direct correlation functions, obtained by solving the 3D-RISM equation coupled with the Kovalenko-Hirata closure.\(^{55}\)

$$h_{i}(r) = \sum_{j} c_{j}(r) + X_{j}(r)$$

(9)

$$h_{i}(r) = \begin{cases} \exp(d_{i}(r)) - 1 & \text{for } d_{i}(r) < 0 \\ -d_{i}(r) & \text{for } d_{i}(r) \geq 0 \end{cases}$$

(10a)

$$d_{i}(r) = \frac{1}{k_{B}T}u_{i}(r) + h_{i}(r) - c_{i}(r)$$

(10b)

where * denotes a convolution integral. $X_{j}(r)$ is a solvent susceptibility function, obtained by solving the RISM equation for pure solvent systems prior to 3D-RISM-KH calculation. $u_{i}(r)$ is an interaction potential function between a solute molecule and solvent molecules at position $r$. In the 3D-RISM-SCF framework, $u_{i}(r)$ is given by

$$u_{i}(r) = 4 \sum_{j} e_{ij} \left\{ \frac{\sigma_{i}}{r_{ij}} \right\}^{12} + \left\{ \frac{\sigma_{i}}{r_{ij}} \right\}^{6} + q_{i} \sum_{j} Z_{j} - q_{i} \int \frac{\psi(r')^{2}}{|r-r'|} dr'$$

(11)

where $e_{ij}$ and $\sigma_{ij}$ are the Lennard-Jones parameters (with usual meanings), and $q_{i}$ denotes the point electronic charge on the solvent site $i$. $Z_{j}$ is a nuclear charge of atom $j$.

### 3. Computational details

In the present study, the parameters for six types of chemical groups were determined, namely, alcohol, amine, imidazole, thiol, phenol, and carboxyl. Table S1 in the ESI\(^{\dagger}\) summarizes the training data sets for parameter fitting.

Prior to the Gibbs energy calculation, the structure optimization of protonated (HA) and unprotonated (A\(^{-}\)) states was performed at the B3LYP/6-31++G(d,p) level, in water, with the PCM, for all the training set molecules. For the Gibbs energy calculation, two different sizes of basis sets were employed, 6-31++G(d,p) and 6-31G, to examine the basis set dependency of the parameter fitting.

The parameters used in the 3D-RISM calculation were temperature of 298.15 K and density of solvent water of 1.0 g cm\(^{-3}\). The Lennard-Jones parameters for solute molecules were taken from the general Amber force field (Gaff) parameter set with antechamber software.\(^{56}\) The extended simple point charge model (SPC/E) parameter set for the geometrical and potential parameters for the solvent water was employed with modified hydrogen parameters ($\sigma = 1.0 \AA$ and $\varepsilon = 0.056$ kcal mol\(^{-1}\)).\(^{57,58}\) The grid spacing for the 3D grid was 0.5 Å and the number of grid points on each axis was 128.

All calculations were performed with a modified version of the GAMESS program package, for which the 3D-RISM-SCF program has been implemented.\(^{59-62}\)

### 4. Results and discussion

Table 1 summarizes the fitted parameters. The computed and experimental $pK_{a}$ values are compared in Fig. 1, which also depicts the $pK_{a}$ values computed without using the LFC scheme, as well as the LFC values. For the $pK_{a}$ computation without the LFC, we assumed the acid dissociation reaction

$$HA + H_{2}O = A^{-} + H_{3}O^{+}$$

(12)

and the associated $pK_{a}$ formula

$$pK_{a} = \frac{G(A^{-}) + G(H_{3}O^{+}) - G(HA) - G(H_{2}O)}{(\ln 10)RT}$$

(13)

where the Gibbs energy of each molecule is calculated by 3D-RISM-SCF. Hereafter, we refer to this treatment as a direct 3D-RISM-SCF scheme.

Fig. 1 shows that the accuracy of the computed $pK_{a}$ value is drastically improved for all the chemical groups by introducing the LFC scheme to 3D-RISM-SCF; we refer to this as the LFC/3D-RISM-SCF scheme. LFC/3D-RISM-SCF shows an excellent score for both the root mean square error (RMSE), 0.709, and the correlation factor, $s = 0.978$. Although the direct 3D-RISM-SCF scheme shows good correlation with the experimental value, $r = 0.912$, the absolute $pK_{a}$ values are significantly overestimated (the RMSE of direct 3D-RISM-SCF is 18.43), indicating that the Gibbs energy of the reaction is overestimated by direct 3D-RISM-SCF. The overestimation of the Gibbs energy of the reaction is suppressed by the scaling factor $s$, which shows the range 0.43 to 0.67. In addition, the contribution of the Gibbs energy of the proton, $G(H^{+})$, is also well parameterized by $C_{0}$.

<table>
<thead>
<tr>
<th>Group</th>
<th>$ka$</th>
<th>$C_{0}$</th>
<th>$s$</th>
<th>RMSE</th>
<th>$r$</th>
<th>$G(H^{+})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol</td>
<td>0.443</td>
<td>−112.460</td>
<td>0.604</td>
<td>1.175</td>
<td>0.698</td>
<td>−254.0</td>
</tr>
<tr>
<td>Amine</td>
<td>0.396</td>
<td>−98.600</td>
<td>0.540</td>
<td>0.469</td>
<td>0.973</td>
<td>−248.9</td>
</tr>
<tr>
<td>Imidazole</td>
<td>0.338</td>
<td>−84.960</td>
<td>0.462</td>
<td>0.629</td>
<td>0.927</td>
<td>−251.1</td>
</tr>
<tr>
<td>Thiol</td>
<td>0.490</td>
<td>−123.403</td>
<td>0.668</td>
<td>0.821</td>
<td>0.750</td>
<td>−252.0</td>
</tr>
<tr>
<td>Phenol</td>
<td>0.317</td>
<td>−78.788</td>
<td>0.432</td>
<td>0.423</td>
<td>0.931</td>
<td>−248.5</td>
</tr>
<tr>
<td>Carboxyl</td>
<td>0.319</td>
<td>−81.352</td>
<td>0.435</td>
<td>0.661</td>
<td>0.832</td>
<td>−253.3</td>
</tr>
<tr>
<td>Total</td>
<td>0.709</td>
<td>0.978</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Unit of $k$ is mol kcal\(^{-1}\). $^b$ Unit of $G(H^{+})$ is kcal mol\(^{-1}\).
Table 2 summarizes the $G(H^+) = C_0/k$ values. The $G(H^+)$ values take in the range $-255$ to $-248$ kcal mol$^{-1}$ is shown. These results are comparable with those obtained in the previous LFC approach by Matsui et al., which ranged from $-268$ to $-246$ kcal mol$^{-1}$, and in the experimental and other theoretical approaches, which ranged from $-264$ to $-259$ kcal mol$^{-1}$.

In Table 2 and Fig. 3, the parameter $k$ for 6-31G takes a value of 0.23–0.42, while for 6-31+G(d,p) the range in values is 0.3–0.48. Here, the 6-31G basis set, both the RMSE and correlation values are slightly worse than with the 6-31+G(d,p) basis set; however, the accuracy of the results determined using LFC/3D-RISM-SCF is acceptable. The direct 3D-RISM-SCF results for the thiol group highlight interesting behavior; thiol shows considerably shifted values (compared with other groups). When using a small basis set, the description of the electronic structure is inadequate, which may be the cause of the shift. Surprisingly, such an irregular behavior of a specific chemical group can be compensated by the parameters in the LFC scheme. This result clearly indicates that LFC/3D-RISM-SCF allows us to use the computationally cheaper basis set, thereby providing a significant advantage when the scheme is applied to large molecular systems such as biomolecules.

To assess the solvent model dependencies on the effective-$K_a$ values by the LFC and direct schemes are compared with the experimental $K_a$ values for the dissociative amino acids by LFC/3D-RISM-SCF were examined. Table 3 and Fig. 4 respectively compare the experimental and computed $K_a$ values of several amino acid side chains. Here, we examined an aspartic acid (Asp), glutamic acid (Glu), cysteine (Cys), histidine (His), lysine (Lys), and tyrosine (Tyr). The computed $K_a$ values by LFC/3D-RISM-SCF show quantitative agreement with the experimental data; in contrast, direct 3D-RISM-SCF shows serious deviation (the RMSE of LFC/3D-RISM-SCF is 0.39, that of direct 3D-RISM-SCF is 18.7). These results indicate that the parameters created by LFC/3D-RISM-SCF have good transferability and that it can be used for the $K_a$ prediction of proteins.

Although the correlation of the direct PCM values with the experimental values is relatively low, 0.80, and the total RMSE is very high, 27.9, the computed $K_a$ values by the LFC/PCM scheme show high accuracy and good correlation (the total RMSE and correlation factor are 0.72 and 0.98) as noted by
Matsui et al. (Fig. 5a and b). LFC/3D-RISM-SCF shows slightly better values in the RMSE and correlation factor than LFC/PCM. In the case of the application of the fitted parameters to the amino acids, LFC/PCM shows excellent transferability (Fig. 5c and Table S3 in the ESI†). The RMSE for the amino acids by LFC/PCM is 1.03, and that by LFC/3D-RISM-SCF is 0.39. This result indicates that LFC/3D-RISM-SCF has better transferability of the LFC scheme to biomolecules than the PCM.

Table 3 Computed and experimental pKₐ values of amino acids

| Amino Acid | Chemical Group | LFC/3D-RISM-SCF | Direct 3D-RISM-SCF | Experimental
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>Carboxyl</td>
<td>3.92</td>
<td>22.86</td>
<td>3.86</td>
</tr>
<tr>
<td>Cys</td>
<td>Thiol</td>
<td>9.14</td>
<td>25.07</td>
<td>8.33</td>
</tr>
<tr>
<td>Glu</td>
<td>Carboxyl</td>
<td>3.94</td>
<td>22.92</td>
<td>4.25</td>
</tr>
<tr>
<td>His(D/E)</td>
<td>Imidazole</td>
<td>6.31/6.27</td>
<td>22.92/24.39</td>
<td>6.04</td>
</tr>
<tr>
<td>Lys</td>
<td>Amine</td>
<td>10.69</td>
<td>24.29</td>
<td>10.53</td>
</tr>
<tr>
<td>Tyr</td>
<td>Phenol</td>
<td>9.64</td>
<td>28.95</td>
<td>10.07</td>
</tr>
</tbody>
</table>

a Taken from ref. 63. b D and E denote the positions where protonation occurs, the delta and epsilon nitrogens, respectively.

Fig. 2 Comparison of the computed pKₐ values with the experimental values. Values for each of the chemical groups are presented in separate panels: (a) alcohol, (b) carboxyl, (c) phenol, (d) amine, (e) imidazole, and (f) thiol. The filled squares and circles denote the pKₐ values determined by direct 3D-RISM-SCF and LFC/3D-RISM-SCF, respectively.

Fig. 3 Comparison of the computed pKₐ values using the 6-31G basis set with the experimental values, using (a) LFC/3D-RISM-SCF and (b) direct 3D-RISM-SCF. The references for the experimental values are given in Table S1 in ESI†.

Fig. 4 Comparison of the computed pKₐ values for amino acids with values determined experimentally. The filled squares and circles denote the computed pKₐ values by direct 3D-RISM-SCF and LFC/3D-RISM-SCF, respectively.
5. Summary

We have proposed a scheme for computing \( pK_a \) values based on 3D-RISM-SCF with the LFC scheme. According to this scheme, the \( pK_a \) value is computed by utilizing the linear relationship between the \( pK_a \) value and the Gibbs energy difference between the protonated and unprotonated states of target molecules. The parameters were determined by the least square fitting for the experimental values of a training set for each chemical group. The parameters introduced here correspond to the Gibbs energy of the excess proton and the scaling factor. The error of the computed \( pK_a \) values arising from the treatment of the excess proton in water and the computational condition such as basis sets for electronic structure calculations are well absorbed by the parameters. It is suggested that, with this scheme, the computationally cheap basis set can be used for \( pK_a \) calculations. The parameters were applied to the amino acid molecules which were not included in the training set, and a good performance was found. Furthermore, LFC/3D-RISM-SCF shows better performance than the LFC/PCM scheme, especially in terms of the transferability of the parameters.

These features may allow us to use this scheme for the prediction of \( pK_a \) values of amino acids in biological systems. In order to apply the LFC/3D-RISM-SCF scheme to amino acids in proteins, a method taking account of environment other than water, such as surrounding residue and ions, which are not currently considered, is necessary. Previously, we proposed the use of advanced methods of 3D-RISM-SCF, in combination with quantum chemical methods, applicable to the biomolecular systems, which we referred to as the quantum mechanics/molecular mechanics/RISM (QM/MM/RISM) and the fragment molecular orbital/3D-RISM (FMO/3D-RISM) methods.\(^{61, 62}\) The combinational use of the scheme proposed here, and QM/MM/RISM or FMO/3D-RISM, may be a powerful tool to tackle the problems related to the protonation and deprotonation of dissociated amino acid residues in biological systems. Such studies with the LFC/3D-RISM-SCF are in progress.

Matsui et al. investigated the redox potentials of several half reactions, metal complexes, and physiologically active molecules using the LFC scheme with the PCM.\(^{64-67}\) As the present LFC/3D-RISM-SCF scheme outperforms the LFC/PCM scheme, it is expected that extension to the redox potentials can improve the accuracy of the estimation. These issues will be addressed in future work.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by Grants-in-Aid [16H00842, 16K05519, 18K05036] from MEXT, Japan. NY also acknowledges support of TOYOTA Riken Scholar from the Toyota Physical and Chemical Research Institute. The authors are grateful to Prof. Yoshihiro Watanabe (Kyushu University) for helpful discussions.

References