

# Analytical energy gradient for reference interaction site model self-consistent field explicitly including spatial electron density distribution

Daisuke Yokogawa, Hirofumi Sato,<sup>a)</sup> and Shigeyoshi Sakaki

*Department of Molecular Engineering, Graduate School of Engineering, Kyoto University, Nishikyo-ku, Kyoto 615-8510, Japan*

(Received 18 August 2009; accepted 29 October 2009; published online 3 December 2009)

Analytical energy gradient formula was derived for reference interaction site model self-consistent field explicitly including spatial electron density distribution (RISM-SCF-SEDD). RISM-SCF-SEDD is a combination method of *ab initio* electronic structure theory and statistical mechanics for molecular liquids [D. Yokogawa *et al.*, *J. Chem. Phys.* **126**, 244504 (2007)]. As shown previously, RISM-SCF-SEDD is numerically stable and has expanded the applicability of the solvation theory. The energy gradient is an indispensable tool to compute molecular geometry and its implementation further extends the capability of RISM-SCF-SEDD. The present method was applied to chemical systems in aqueous solution; hydration structure and geometry of phosphate anion  $\text{PO}_4^{3-}$  and tautomerization between 2-pyridone and 2-hydroxypyridine. Compared to available experimental data, the present method correctly reproduced the geometries and relative energies of solvated molecules with microscopic solvent distribution. It is clearly shown that highly sophisticated quantum chemical calculation such as coupled cluster with single and double and perturbative triple excitations coupled with solvation effect is a powerful tool to accurately evaluate molecular properties. © 2009 American Institute of Physics.

[doi:[10.1063/1.3265856](https://doi.org/10.1063/1.3265856)]

## I. INTRODUCTION

Molecular geometry is a fundamental information to understand chemical phenomena and extensive studies are performed over a long period of time. For molecule in isolated state, several methods become mature and precise determination of geometrical parameters is realizable, especially for small polyatomic molecule. *Ab initio* electronic structure computation is now routinely available due to the development of energy gradient technique for various electronic structure theories and highly accurate geometrical parameters—the error is sometimes smaller than experimental one—are available. Contrastingly, the situation is different for molecule in solution phase. Most widely employed experimental approaches to obtain geometrical data are x-ray and neutron scatterings, which allow us to estimate bond lengths and bond angles of solvated molecule. However, in general, it is still experimentally difficult to obtain geometrical parameters in high accuracy.

Nowadays, development of *ab initio* quantum chemical methods involving solvation effect has been strenuously proceeding. Polarizable continuum model<sup>1</sup> (PCM) is one of the most popular methods, in which solvent is treated as dielectric continuum surrounding the solute molecule.<sup>2</sup> The energy gradient technique has been developed in 1994 (Ref. 2) and now it is routinely executable in various program packages. The method is very useful but it has been well recognized that the evaluation of hydrogen bonding is rather poor.<sup>3</sup> To make up for the defect, a few solvent molecules are often considered explicitly within the system. Quantum mechani-

cal and molecular mechanics (QM/MM) method utilizing molecular simulation method such as molecular dynamics naturally takes into account the effect of microsolvation. In this method, solute molecule is computed with quantum chemical calculation under the potential produced by the surrounding solvent molecules that are dealt with classical mechanics. Statistically averaged energy gradient is applied to obtain equilibrium geometry in solution.<sup>4,5</sup> In principle, accurate geometrical parameters are expected to be obtained if adequate ensemble of solvent configuration is feasible.

Reference interaction site model self-consistent field (RISM-SCF) is another method,<sup>6–8</sup> where solvent is treated with statistical mechanics and the electronic structure is evaluated under the electric fields generated by microscopic solvation structure, i.e., distribution function. Thanks to the analytic, algebraic nature of RISM theory, statistical ensemble of solvent configuration is perfectly computed with reasonable cost. The analytical energy gradient formula has been proposed in the early stage,<sup>8</sup> and successfully applied to a large number of systems.<sup>9,10</sup> However, the calculation of the original RISM-SCF often diverges when the system contains buried atoms. Recently, we proposed a new generation of RISM-SCF to overcome the weak point.<sup>11</sup> In the new method, spatial electron density distribution (SEDD) was explicitly introduced in RISM-SCF scheme (RISM-SCF-SEDD). Not only the description of physical nature of chemical system becomes accurate but also the numerical stability of the computation is significantly improved. In this article, we derived an analytical gradient formula based on RISM-SCF-SEDD. To obtain accurate geometry in solution phase, the RISM-SCF-SEDD calculation was performed

<sup>a)</sup>Electronic mail: [hirofumi@moleng.kyoto-u.ac.jp](mailto:hirofumi@moleng.kyoto-u.ac.jp).

with density functional theory (DFT) and coupled cluster expansion using single and double excitations with perturbation treatment of triples [CCSD(T)].

This article is organized as follows. In Sec. II, the equations employed in this work were derived. The detail of the calculation was described in Sec. III. The present method was applied to two demonstrative examples; hydration structure and geometry of phosphate anion  $\text{PO}_4^{3-}$  in aqueous phase and the tautomerization between 2-pyridone (PY) and 2-hydroxypyridine (HP) in aqueous phase. The results were compared with available experimental data in Sec. IV.

## II. THEORY

Before deriving the equation of energy gradient, we briefly summarize key equations employed in this work. Solvation structure around a solute molecule is calculated by RISM equation with hypernetted-chain (HNC) closure,

$$h_{\alpha s}^{UV} = \sum_{\gamma t} \omega_{\alpha \gamma}^U * c_{\gamma t}^{UV} * \chi_{ts}^{VV}, \quad (1)$$

$$\chi_{ts}^{VV} = \omega_{ts}^V + \rho_t^V h_{ts}^{VV}, \quad (2)$$

$$h_{\alpha s}^{UV} = \exp[-\beta \phi_{\alpha s}^{UV} + h_{\alpha s}^{UV} - c_{\alpha s}^{UV}] - 1, \quad (3)$$

where  $\beta = 1/k_B T$  and  $k_B$  is Boltzmann's constant.  $h_{\alpha s}^{UV}$  and  $c_{\alpha s}^{UV}$  are total and direct correlation functions between solute ( $U$ ) site  $\alpha$  and solvent ( $V$ ) site  $s$ ,  $h_{st}^{VV}$  is the total correlation function for solvent, and  $\{\phi_{\alpha s}^{UV}\}$  is the solute-solvent interaction potential.  $\{\omega_{\alpha \gamma}^U\}$  and  $\{\omega_{st}^V\}$  are intramolecular correlation functions, which define solute and solvent molecular geometries, respectively.  $\rho_t^V$  is the number density of the solvent site  $t$ .

The solute-solvent interaction potential,  $\{\phi_{\alpha s}^{UV}\}$ , is the sum of Coulombic and Lennard-Jones (LJ) terms,

$$\phi_{\alpha s}^{UV}(r) = \phi_{\alpha s}^{CL}(r) + 4\epsilon_{\alpha s} \left\{ \left( \frac{\sigma_{\alpha s}}{r} \right)^{12} - \left( \frac{\sigma_{\alpha s}}{r} \right)^6 \right\}. \quad (4)$$

The Coulombic term  $\phi_{\alpha s}^{CL}(r)$  is usually written as classical expression, namely,  $q_\alpha q_s / r$ . However, in RISM-SCF-SEDD, it is calculated directly from electron density  $\rho_e(\mathbf{r})$  of the solute molecule.<sup>11</sup>  $\rho_e(\mathbf{r})$  is expanded with the spherical auxiliary basis sets (ABSs),  $\{f_i\}$ , on each solute atomic site,

$$\rho_e(\mathbf{r}) = \sum_i d_i f_i(\mathbf{r}), \quad (5)$$

with the coefficient  $\mathbf{d}$  and ABS,  $\phi_{\alpha s}^{CL}(r)$  is thus expressed as follows:

$$\phi_{\alpha s}^{CL}(r) = -q_s \sum_{i \in \alpha} d_i \int \frac{f_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{q_s Z_\alpha}{r}, \quad (6)$$

where  $q_s$  is a partial charge on a solvent site  $s$ ,  $Z_\alpha$  is the nuclear charge of a solute atomic site  $\alpha$ , and  $\mathbf{R}_\alpha$  is the position vector of atomic site  $\alpha$ . The expansion coefficient  $\{d_i\}$  is evaluated by the following equation:

$$\mathbf{d} = \mathbf{X}^{-1} \text{tr}(\mathbf{P}\mathbf{Y}) - \frac{\mathbf{Z}'\mathbf{X}^{-1} \text{tr}(\mathbf{P}\mathbf{Y}) - N_e \mathbf{X}^{-1}\mathbf{Z}}{\mathbf{Z}'\mathbf{X}^{-1}\mathbf{Z}}, \quad (7)$$

where  $\mathbf{P}$  is usual ‘‘P-matrix’’ related to the electron density and  $N_e$  is the number of electron in solute molecule.  $\mathbf{X}$ ,  $\mathbf{Y}$ , and  $\mathbf{Z}$  are matrices which are defined with the given ABSs, atomic orbitals, and geometry of solute molecule, as follows:<sup>11</sup>

$$X_{ij} = \int \int f_i(\mathbf{r}_1) |\mathbf{r}_1 - \mathbf{r}_2| f_j(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2, \quad (8)$$

$$Y_{\mu\nu,i} = \int \int \chi_\mu(\mathbf{r}_1) \chi_\nu(\mathbf{r}_1) |\mathbf{r}_1 - \mathbf{r}_2| f_i(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2, \quad (9)$$

$$Z_i = \int f_i(\mathbf{r}) d\mathbf{r}, \quad (10)$$

where  $\chi$  is the basis function employed in molecular orbital calculation.

In the RISM-HNC framework, excess chemical potential ( $\Delta\mu$ ) given by

$$\begin{aligned} \Delta\mu &= -\frac{1}{\beta} \sum_{\alpha s} \rho_s^V \int d\mathbf{r} \left[ c_{\alpha s}^{UV}(r) - \frac{1}{2} (h_{\alpha s}^{UV}(r))^2 + \frac{1}{2} h_{\alpha s}^{UV}(r) c_{\alpha s}^{UV}(r) \right] \\ &= -\frac{1}{\beta} \sum_{\alpha s} \rho_s^V \int d\mathbf{r} \left[ e^{-\beta \phi_{\alpha s}^{UV}(r) + t_{\alpha s}^{UV}(r)} - 1 - t_{\alpha s}^{UV}(r) \right. \\ &\quad \left. + \frac{1}{2} (h_{\alpha s}^{UV}(r))^2 - h_{\alpha s}^{UV}(r) t_{\alpha s}^{UV}(r) \right] \\ &\quad + \frac{1}{(2\pi)^3 \beta} \int d\mathbf{k} \left[ \sum_{\alpha s} \rho_s^V \hat{h}_{\alpha s}^{UV}(k) \hat{c}_{\alpha s}^{UV}(k) \right. \\ &\quad \left. - \frac{1}{2} \sum_{\alpha \gamma st} \rho_s^V \hat{c}_{\alpha s}^{UV} \hat{c}_{\gamma t}^{UV} \hat{\chi}_{st}^{VV} \hat{\omega}_{\alpha \gamma}^U \right], \end{aligned} \quad (11)$$

where  $t_{\alpha s}^{UV} = h_{\alpha s}^{UV} - c_{\alpha s}^{UV}$  and hat ( $\hat{\phantom{x}}$ ) means the function in reciprocal space. RISM equation and HNC closure are derived from Eq. (11) by the well-known variational stationarity.<sup>12</sup> In analogy with the RISM-HNC as well as with the original RISM-SCF method,<sup>8</sup>  $\mathcal{A}$  corresponding to the Helmholtz free energy of the system is defined as follows:

$$\mathcal{A} = E_{\text{solute}} + \Delta\mu, \quad \text{where } E_{\text{solute}} = \langle \Psi | H | \Psi \rangle. \quad (12)$$

$H$  is the Hamiltonian for a solute molecule in the standard quantum chemical method and  $\Psi$  is the wave function of the solute molecule in solvent. Thanks to the variational stationarity of RISM-HNC theory, the first derivative of  $\mathcal{A}$  with respect to the nuclear coordinate of the solute molecule  $R_\alpha$  is simply given by (see Appendix for the detail)

$$\begin{aligned} \frac{\partial \mathcal{A}}{\partial R_a} &= \frac{\partial E_{\text{solute}}}{\partial R_a} + \frac{\partial \Delta \mu}{\partial R_a} \\ &= \frac{\partial E_{\text{solute}}}{\partial R_a} - \mathbf{V}^t \frac{\partial \mathbf{d}}{\partial R_a} \\ &\quad - \frac{1}{2(2\pi)^3 \beta} \sum_{\alpha\gamma st} \rho_s \int d\mathbf{k} \hat{c}_{\alpha s}^{UV}(k) \hat{c}_{\gamma t}^{UV}(k) \frac{\partial \omega_{\alpha\gamma}^U(k)}{\partial R_a} \hat{\chi}_{st}^{VV}(k). \end{aligned} \quad (13)$$

The first and third terms in the last equation are the same as those obtained by previous work.<sup>8</sup> Column vector  $\mathbf{V}$  in the second term represents the electrostatic potential on ABSs induced by averaged solvent distribution, whose  $i$ th element ( $V_i$ ) is given by

$$V_i = \sum_s \rho_s^V q_s \iint \frac{f_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} h_{\alpha s}^{UV}(|\mathbf{r} - \mathbf{r}'|) d\mathbf{r} d\mathbf{r}' \quad (i \in \alpha). \quad (14)$$

The derivative of  $\mathbf{d}$  with respect to  $R_a$  in the second term is

$$\begin{aligned} \frac{\partial \mathbf{d}}{\partial R_a} &= - \left[ \mathbf{X}^{-1} - \frac{\mathbf{X}^{-1} \mathbf{Z}}{\mathbf{Z}^t \mathbf{X}^{-1} \mathbf{Z}} \mathbf{Z}^t \mathbf{X}^{-1} \right] \left[ \frac{\partial \mathbf{X}}{\partial R_a} \mathbf{d} - \text{tr} \left( \mathbf{P} \frac{\partial \mathbf{Y}}{\partial R_a} \right) \right] \\ &\quad + \frac{\mathbf{X}^{-1} \mathbf{Z}}{\mathbf{Z}^t \mathbf{X}^{-1} \mathbf{Z}} \text{tr} \left( \mathbf{P} \frac{\partial \mathbf{S}}{\partial R_a} \right), \end{aligned} \quad (15)$$

where  $\mathbf{S}$  is the usual overlap matrix. In our program code, the matrices,  $\mathbf{X}$ ,  $\mathbf{Y}$ , and  $\mathbf{Z}$ , and the first derivative of  $\mathbf{X}$  and  $\mathbf{Y}$  are calculated by the Obara–Saika recursion relation.<sup>13,14</sup>

### III. COMPUTATIONAL DETAILS

All calculations were performed using the GAMESS program package,<sup>15</sup> where RISM-SCF-SEDD and the energy gradient technique were implemented by us. Spherical convolution was computed with Ohura's fast Fourier transform program.<sup>16</sup> For the acceleration of the RISM-SCF convergence, M-DIIS method<sup>17</sup> was employed. Dielectric correction for RISM was not used and the standard geometry optimization algorithm implemented in GAMESS was employed.

The present calculation was performed with DFT using the B3LYP functional (VWN 5) and CCSD(T).<sup>18</sup> Hydration effect was evaluated by RISM-SCF-SEDD and for phosphate molecule, C-PCM (Ref. 19) calculation was also performed for comparison. The basis set for  $\text{PO}_4^{3-}$  was cc-pV(T+d)Z for P atom and cc-pVTZ for O atom. For all atoms in PY and HP, aug-cc-pVDZ basis set was used. ABS was constructed from the s-type primitive atomic orbital basis of these basis sets [please see Ref. 8 in our previous work (Ref. 11) for the details]. The diffuse functions were not included in the ABS construction for PY/HP.

The LJ parameters for each atom are summarized in Table I and potentials between atoms were evaluated with the Lorentz–Berthelot mixing rule [ $\sigma_{ij} = (\sigma_i + \sigma_j)/2$  and  $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ ]. For solvent water molecule, the SPC model was employed<sup>20</sup> with the LJ parameters of the hydrogen sites ( $\sigma = 1.0$  Å and  $\epsilon = 0.056$  kcal/mol). All calculations were performed at the temperature of 298.15 K and at the number of density of 0.033 426 molecule/Å<sup>3</sup> ( $= 1$  g cm<sup>-3</sup>).

TABLE I. LJ parameters.

	$\sigma$ (Å)	$\epsilon$ (kcal/mol)
		$\text{PO}_4^{3-a}$
P	3.742	0.200
O	2.960	0.210
		PY/HP <sup>b</sup>
H	1.782	0.020
C	3.750	0.110
N	3.250	0.170
O	2.960	0.210

<sup>a</sup>Reference 28.

<sup>b</sup>References 24, 29, and 30.

### IV. RESULTS AND DISCUSSION

In this work, two chemical systems were studied; hydration structure and geometry of phosphate anion  $\text{PO}_4^{3-}$  in aqueous phase and tautomerization between PY and HP in aqueous phase. The former system was recently studied with neutron scattering experiments, and hydration structure as well as intramolecular bond lengths were precisely reported.<sup>21</sup> The latter system has been widely studied with theoretical and experimental methods.

#### A. Phosphate anion $\text{PO}_4^{3-}$

The geometry optimizations of  $\text{PO}_4^{3-}$  were performed with DFT (B3LYP) calculation under the  $T_d$  symmetry. In aqueous solution, both of RISM-SCF-SEDD and C-PCM were employed. The geometrical parameters optimized with gas phase (geom I), RISM-SCF-SEDD (geom II), and with C-PCM (geom III) are shown in Table II together with experimental data. In gas phase, P–O and O–O distances were 1.582 and 2.583 Å, respectively. These distances became significantly shorter in aqueous solution; the optimized distances of P–O and O–O with RISM-SCF-SEDD were 1.539 and 2.514 Å, respectively, which were very close to the experimental data (1.53 Å for P–O and 2.49 Å for O–O). Computation with CCSD(T) method gives essentially the same results.<sup>22</sup> Although the optimized distances by C-PCM also became shorter than those of geom I, they were somewhat longer than experimental data.

The computed radial distribution functions (RDFs) [ $g(r)$ ] of water oxygen ( $\text{O}_w$ ) and water hydrogen ( $\text{H}_w$ ) around solute P and O atoms are shown in Fig. 1, together with the each coordination numbers [ $N_i(r)$ ],

TABLE II. Geometrical parameters of  $\text{PO}_4^{3-}$  (Å) in gas phase and in aqueous solution.

	Gas phase	Aqueous solution		
	geom I	RISM-SCF-SEDD geom II	C-PCM geom III	Expt. <sup>a</sup>
P–O	1.582	1.539	1.556	1.53
O–O	2.583	2.514	2.541	2.49

<sup>a</sup>Reference 21.

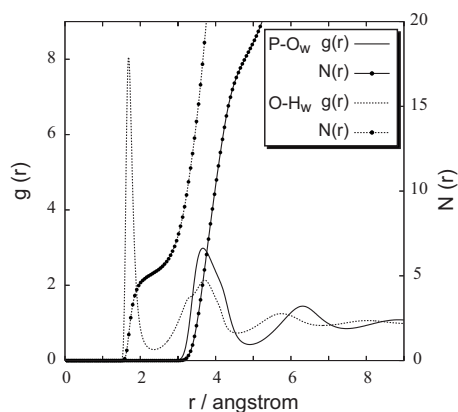


FIG. 1. Radial distribution functions [ $g(r)$ ] and coordination numbers [ $N(r)$ ] of water oxygen ( $O_w$ ) and water hydrogen ( $H_w$ ) around solute P and O atoms.

$$N_t(r) = 4\pi\rho_t^V \int_0^r x^2 g_{Xt}(x) dx \quad (X = P \text{ or } O), \quad (16)$$

where  $\rho_t^V$  is the number density of water. A high and narrow peak was seen at 1.69 Å in  $g_{O_{H_w}}(r)$ , indicating a strong hydrogen bonding between O sites of  $PO_4^{3-}$  and  $H_w$ . The P– $O_w$  peak at 3.67 Å was broad compared to O– $H_w$  peak. This is because  $O_w$  does not interact directly with  $PO_4^{3-}$  and fluctuates around  $H_w$ . The peak position of RDF and coordination number obtained by RISM-SCF-SEDD is compared to the experimental data<sup>21</sup> in Table III. The computed distances show a good agreement, especially the P– $O_w$  distance is almost identical to the experimental data. The coordination numbers corresponding to the first shell are somewhat larger, but reasonably accord with those obtained by the neutron scattering experiments.

## B. PY/HP

In this molecule, tautomeric isomerization occurs between amide form (PY) and imidic acid forms (HP). Three conformers, PY,  $HP_{anti}$ , and  $HP_{syn}$  are considered in terms of the position of the transferring H atom. Geometries of these conformers were optimized under  $C_s$  symmetry and shown in Fig. 2. The obtained geometry reasonably agrees with the experimental data determined by microwave spectroscopy.<sup>23</sup> Root mean square deviation in the bond length between the DFT (B3LYP) calculations and experimentally obtained data in gas phase was 0.015 Å for PY and 0.010 Å for  $HP_{syn}$ . The

TABLE III. The RDF peak position and the coordination number.

	This work	Expt. <sup>a</sup>
P– $O_w$ distance (Å)	3.67	3.7
O– $H_w$ distance (Å)	1.69	1.9
$N$ (first shell of P– $O_w$ )	18.4	$15 \pm 3$
$N$ (first shell of O– $H_w$ )	5.2	3

<sup>a</sup>Reference 21.

deviation is satisfactorily small, suggesting that calculation is reliable for the geometry optimization. On the other hand, much highly sophisticated quantum chemical method is necessary to adequately evaluate the relative energy. The energy obtained by DFT (B3LYP) shows that PY is slightly stable, contradicting the experimental data. By performing CCSD(T) calculation the relative stability between PY and HP was properly reproduced (Table IV).

The geometries in aqueous phase calculated by RISM-SCF-SEDD are also shown in Fig. 2. The solvation does not affect the geometry of  $HP_{syn}$  and  $HP_{anti}$  so much. The maximum change in distances from gas to aqueous phase was only 0.004 Å for  $HP_{syn}$  and 0.007 Å for  $HP_{anti}$ . On the other hand, the geometry of PY in aqueous phase is quite different from that in gas phase. In particular, the bond lengths of C2–N1 and C3–C2 decrease by 0.029 and 0.022 Å, respectively, and that of O7–C2 increases by 0.033 Å. This large change may be understood with resonance structures in PY. As shown in Scheme 1, PY has zwitterionic resonance structures in addition to amide form. In gas phase, amide form is stable compared to zwitterionic form, while the zwitterionic resonance structures are much stabilized in aqueous phase by solvent water. In aqueous phase, the double bond character becomes greater in C2–N1 and C3–C2 bonds while single bond character becomes greater in O7–C2 bond.

The relative free energy in aqueous phase calculated by the present method is shown in Table IV, where those computed by the original RISM-SCF (HF level)<sup>24</sup> and by QM/MM (Ref. 25) are also shown. In the case of CCSD(T) calculation in aqueous phase, the energies were also evaluated with a different type of LJ parameter, general amber force field (GAFF).<sup>26</sup> A qualitative trend, namely, PY became more stable compared to HP in aqueous solution, is reproduced by all methods, but there are quantitative differences among the methods. DFT level calculation with RISM-SCF-SEDD and the original RISM-SCF overestimate the stability

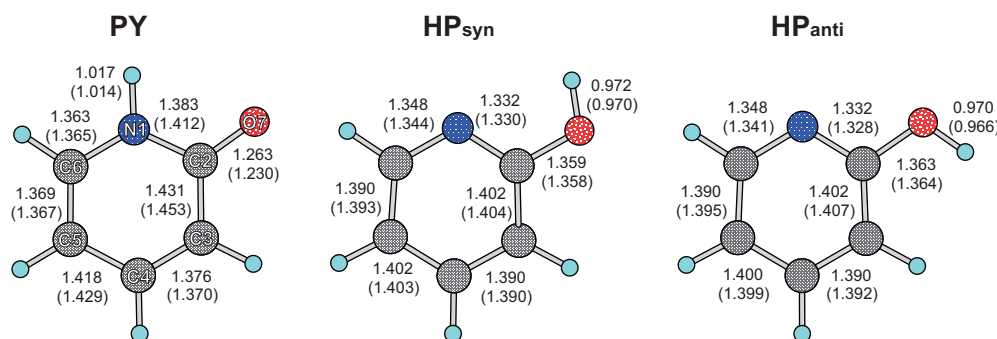


FIG. 2. Optimized bond lengths in PY,  $HP_{syn}$ , and  $HP_{anti}$  (Å) in aqueous phase. The values in gas phase are in parenthesis.

TABLE IV. Computed and experimentally obtained relative energies (kcal/mol). The energies in aqueous phase calculated by CCSD(T) with GAFF parameter are shown in parentheses.

	Gas phase			Aqueous solution		
	PY	HP <sub>anti</sub>	HP <sub>syn</sub>	PY	HP <sub>anti</sub>	HP <sub>syn</sub>
DFT <sup>a</sup>	0.0	5.3	0.4	0.0	8.5	9.2
CCSD(T) <sup>a</sup>	0.0	3.7	-1.1	0.0	6.1 (6.0)	6.9 (7.0)
HF <sup>b</sup>	0.0	...	-0.7	0.0	...	9.7
QM/MM (AM1) <sup>c</sup>	0.0	...	-0.4	0.0	...	5.7
Expt. <sup>d</sup>	0.0	...	-0.8	0.0	4.0	

<sup>a</sup>This work.<sup>b</sup>Reference 24.<sup>c</sup>Reference 25.<sup>d</sup>References 23, 31, and 32.

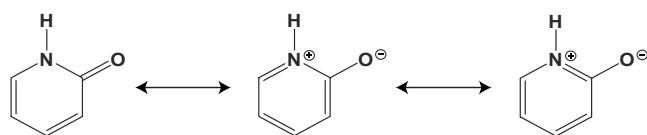
of PY. The agreement is certainly improved with the present RISM-SCF-SEDD coupled with CCSD(T) calculation. In the calculation, the sensitivity of LJ parameters was significantly small.

Figure 3 displays the radial distribution functions of hydrogen ( $H_w$ ) of solvent water around N and O atoms in the three isomers obtained by RISM-SCF-SEDD method. In HP<sub>syn</sub> and HP<sub>anti</sub>, the peak around 2 Å corresponds to the hydrogen bonding.  $H_w$  coordinates to both N and O atoms, which are exposed to the solvent. In the case of PY, N atom is already occupied by the transferring H atom and  $H_w$  cannot interact, leading to the disappearance of the peak around 2 Å seen in HP. At the same time, O atom in PY shows alkoxide character and its charge becomes more negative compared to that in HP. Consequently, O strongly attracts  $H_w$ , which makes RDF of O- $H_w$  higher. This is an important factor that PY is strongly stabilized in aqueous phase.

## V. CONCLUSIONS

We derived an analytical energy gradient formula based on RISM-SCF-SEDD. The present method was applied to two demonstrative systems in aqueous phase, phosphate anion  $PO_4^{3-}$ , and tautomerization between PY and HP.

The computational procedure of RISM-SCF-SEDD is numerically stable compared to the original RISM-SCF. The new method has expanded the applicability of *ab initio* electronic structure calculation coupled with solvation effect and the development of energy gradient technique reported here further extends the capability of the method. It should be emphasized that both of the original RISM-SCF and RISM-SCF-SEDD methods are equivalent to *ab initio* QM/MM method, in which the electronic structure of the solute molecule is determined under the influence of solvent molecules. Thanks to the analytical treatment based on statistical mechanics, the computational time to evaluate the statistical ensemble is negligibly short. This feature enables us to combine highly sophisticated electronic structure theory, e.g.,



SCHEME 1. Resonance structures in PY.

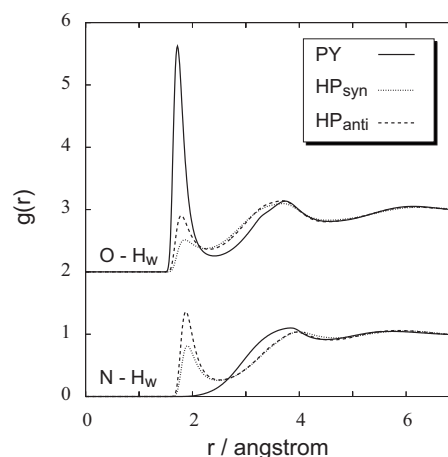
CCSD(T), DFT. We believe that the present method will become the powerful tools to analyzing interesting chemical systems, such as chemical reactions in solution, with reasonable computational cost.

## ACKNOWLEDGMENTS

This work was financially supported in part by the Grant-in Aid for Scientific Research on Priority Areas “Water and Biomolecules” (Grant No. 430-18031019), “Molecular Theory for Real Systems” (Grant No. 461), “Ionic Liquids” (Grant No. 452-20031014), “Molecular Science of Fluctuations” (Grant No. 2006-21107511), and by the Grant-in Aid for General Research (Grant No. 19350010). D.Y. thanks the Grant-in Aid for JSPS Fellows. All of them were supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Japan.

## APPENDIX: DERIVATION OF EQUATION (13)

After some mathematical manipulation, the derivative of excess chemical potential  $\Delta\mu$  with respect to nuclear coordinate  $R_\alpha$  is obtained as follows:

FIG. 3. Radial distribution functions of water hydrogen ( $H_w$ ) around N and O atoms in PY, HP<sub>syn</sub>, and HP<sub>anti</sub>.

$$\begin{aligned}
\frac{\partial \Delta \mu}{\partial R_\alpha} = & -\frac{1}{\beta} \sum_{\alpha, s} \rho_s^V \int d\mathbf{r} \left[ \left( e^{-\beta \phi_{as}^{UV}(r) + t_{as}^{UV}(r)} - 1 - h_{as}^{UV}(r) \right) \right. \\
& \times \frac{\partial t_{as}^{UV}(r)}{\partial R_\alpha} + \left( -t_{as}^{UV}(r) + h_{as}^{UV}(r) - c_{as}^{UV}(r) \right) \frac{\partial h_{as}^{UV}(r)}{\partial R_\alpha} \\
& + \left( -h_{as}^{UV}(r) + \sum_{\gamma t} \omega_{\alpha\gamma}^U c_{\gamma s}^{UV} \chi_{st}^{VV}(r) \right) \frac{\partial c_{as}^{UV}(r)}{\partial R_\alpha} \\
& + \left. \left( -\beta e^{-\beta \phi_{as}^{UV}(r) + t_{as}^{UV}(r)} \frac{\partial \phi_{as}^{UV}}{\partial R_\alpha} \right. \right. \\
& \left. \left. + \left( \frac{1}{2} \sum_{\gamma t} \frac{\partial \omega_{\alpha\gamma}^U}{\partial R_\alpha} c_{as}^{UV} c_{\gamma t}^{UV} \chi_{st}^{VV}(r) \right) \right) \right]. \quad (\text{A1})
\end{aligned}$$

The terms with the derivatives of  $t^{UV}$ ,  $h^{UV}$ , and  $c^{UV}$  are vanished using RISM and HNC equations. Because the distance  $r$  between solute and solvent atoms in the solute-solvent interaction potential  $\{\phi_{as}^{UV}\}$  does not depend on the nuclear coordinates  $R_\alpha$  of solute atoms, the derivative of  $\phi_{as}^{UV}$  in Eq. (A1) is given by

$$\frac{\partial \phi_{as}^{UV}}{\partial R_\alpha} = -q_s \sum_{i \in \alpha} \frac{\partial d_i}{\partial R_\alpha} \int \frac{f_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \quad (\text{A2})$$

By substituting Eq. (A2) into Eq. (A1), Eq. (13) is obtained.

<sup>1</sup>J. Tomasi, B. Mennucci, and R. Cammi, *Chem. Rev. (Washington, D.C.)* **105**, 2999 (2005).

<sup>2</sup>R. Cammi and J. Tomasi, *J. Chem. Phys.* **100**, 7495 (1994); **101**, 3888 (1994); M. Cossi, B. Mennucci, and R. Cammi, *J. Comput. Chem.* **17**, 57 (1996).

<sup>3</sup>P. Bandyopadhyay, M. S. Gordon, B. Mennucci, and J. Tomasi, *J. Chem. Phys.* **116**, 5023 (2002).

<sup>4</sup>N. Okuyama-Yoshida, K. Kataoka, M. Nagaoka, and T. Yamabe, *J. Chem. Phys.* **113**, 3519 (2000).

<sup>5</sup>M. E. Martín, A. M. Losa, I. Fdez Galván, and M. A. Aguilar, *J. Chem. Phys.* **121**, 3710 (2004).

<sup>6</sup>*Molecular Theory of Solvation*, edited by F. Hirata (Kluwer, Dordrecht, 2003).

<sup>7</sup>S. Ten-no, F. Hirata, and S. Kato, *J. Chem. Phys.* **100**, 7443 (1994).

<sup>8</sup>H. Sato, F. Hirata, and S. Kato, *J. Chem. Phys.* **105**, 1546 (1996).

<sup>9</sup>S. Hayaki, K. Kido, D. Yokogawa, H. Sato, and S. Sakaki, *J. Phys. Chem. B* **113**, 8227 (2009).

<sup>10</sup>K. Iida, D. Yokogawa, A. Ikeda, H. Sato, and S. Sakaki, *Phys. Chem. Chem. Phys.* **11**, 8556 (2009); K. Iida, D. Yokogawa, H. Sato, and S. Sakaki, *Chem. Phys. Lett.* **443**, 264 (2007).

<sup>11</sup>D. Yokogawa, H. Sato, and S. Sakaki, *J. Chem. Phys.* **126**, 244504 (2007).

<sup>12</sup>S. J. Singer and D. Chandler, *Mol. Phys.* **55**, 621 (1985).

<sup>13</sup>S. Obara and A. Saika, *J. Chem. Phys.* **84**, 3963 (1986).

<sup>14</sup>R. Ahlrichs, *Phys. Chem. Chem. Phys.* **8**, 3072 (2006).

<sup>15</sup>M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comput. Chem.* **14**, 1347 (1993).

<sup>16</sup>Useful libraries for FFT calculation can be download from the following web page: <http://www.kurims.kyoto-u.ac.jp/~ooura/>.

<sup>17</sup>A. Kovalenko, S. Ten-no, and F. Hirata, *J. Comput. Chem.* **20**, 928 (1999).

<sup>18</sup>CCSD(T) calculation in this work was performed using the subroutine implemented in GAMESS (Ref. 27).

<sup>19</sup>V. Barone and M. Cossi, *J. Phys. Chem. A* **102**, 1995 (1998).

<sup>20</sup>H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981).

<sup>21</sup>P. E. Mason, J. M. Cruickshank, G. W. Neilson, and P. Buchanan, *Phys. Chem. Chem. Phys.* **5**, 4686 (2003).

<sup>22</sup>To check the accuracy of the present calculation level, the optimized distances were also computed by RISM-SCF-SEDD with CCSD(T). Because analytical gradient of CCSD(T) was not available in the current version of GAMESS package, the optimized geometry was determined numerically by evaluating free energy surface. The obtained distances were 1.536 Å for P-O and 2.507 Å for O-O, which are very close to those obtained by DFT (B3LYP) calculation.

<sup>23</sup>L. D. Hatherley, R. D. Brown, P. D. Godfrey, and A. P. Pierlot, *J. Phys. Chem.* **97**, 46 (1993).

<sup>24</sup>H. Sato, F. Hirata, and S. Sakaki, *J. Phys. Chem. A* **108**, 2097 (2004).

<sup>25</sup>J. Gao and L. Shao, *J. Phys. Chem.* **98**, 13772 (1994).

<sup>26</sup>J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, *J. Comput. Chem.* **25**, 1157 (2004).

<sup>27</sup>P. Piecuch, S. A. Kucharski, K. Kowalski, and M. Musial, *Comput. Phys. Commun.* **149**, 71 (2002).

<sup>28</sup>W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.* **117**, 5179 (1995).

<sup>29</sup>W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* **110**, 1657 (1988).

<sup>30</sup>S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, Jr., and P. Weiner, *J. Am. Chem. Soc.* **106**, 765 (1984).

<sup>31</sup>P. Beak, *Acc. Chem. Res.* **10**, 186 (1977).

<sup>32</sup>J. Frank and A. R. Katritzky, *J. Chem. Soc., Perkin Trans. 2* **1976**, 1428.