A highly parallelizable integral equation theory for three dimensional solvent distribution function: Application to biomolecules

Daisuke Yokogawa,[1] Hirofumi Sato,[1,a)] Takashi Imai,[2] and Shigeyoshi Sakaki[1]

[1]*Department of Molecular Engineering, Graduate School of Engineering, Kyoto University, Nishikyo-ku, Kyoto 615-8510, Japan*
[2]*Computational Science Research Program, RIKEN, Wako, Saitama 351-0198, Japan*

Three dimensional (3D) hydration structure is informative to clarify the functions of hydrated waters around a protein. We develop a new approach to calculate 3D solvation structure with reasonable computational cost. In the present method, the total solvation structure is obtained using conventional one dimensional reference interaction site model (RISM) followed by integrating the 3D fragment data, which are evaluated around each atom (site) of solute. Thanks to this strategy, time-consuming 3D fast Fourier transformation, which is required in 3D-RISM theory, can be avoided and high-parallel performance is achieved. The method is applied to small molecular systems for comparison with 3D-RISM. The obtained results by the present method and by 3D-RISM show good agreement. The hydration structures for a large protein computed by the present method are also consistent with those obtained by x-ray crystallography. © *2009 American Institute of Physics*. [DOI: 10.1063/1.3077209]

I. INTRODUCTION

Most of biochemical reactions proceed in aqueous phase, where the water molecules locate inside and around biomolecules. Almost all of the functions of proteins can proceed only in this environment. For example, bacteriorodopsin, which is a light-driven proton pump, prepares hydrogen bonding network between bound waters and residues inside. Protons are considered to be transferred across the membrane through the network.[1,2] To elucidate the functions of these waters, the information of their distribution is indispensable.

Experimental studies to obtain the hydration structure have been performed using several powerful techniques such as x-ray diffraction at low temperature and the scattering with very high-power neutron source. Thanks to these studies, high resolution data, which are accurate enough to discuss the hydration structure around protein, have become available. However, such solvation information is still very limited. Furthermore, almost all of them are in crystal, which are believed to be essentially different from those in aqueous environment.

Theoretical approach to investigate the hydration structure has been also performed. The most popular treatment is molecular simulation, e.g., molecular dynamics. Although the method has been applied to various types of solvation systems, very long simulation is necessary to obtain the hydration structure inside and around a protein with high quality. Recently a theoretical determination of hydration structure around a protein was done by Imai *et al.*[3] based on statistical mechanics. They employed three dimensional reference interaction site model[4,5] (3D-RISM) for hen egg-white lysozyme and showed that 3D-RISM is capable of computing the solvation structure not only around a protein

but also inside appropriately. Yoshida *et al.*[6] showed that 3D-RISM can also correctly reproduce the selective ion binding with human lysozyme.

In 3D-RISM calculation, the hydration structure is calculated on 3D lattice grids. Applicable size of protein is determined by the number of grid points. In principle, it would be possible to compute finer hydration structure around larger proteins of any size as increasing the number of grid points. However, to obtain such results, greater memory size and very long calculation are required.

In this work, a new method for computing 3D solvation structure was developed. Since the calculations are performed with a combination of logarithmic grid and real solid harmonics around each solute atom (site), the number of grid points can be adjusted more flexibly compared to 3D lattice grid employed in 3D-RISM. Moreover, the present method does not require time-consuming 3D fast Fourier transformation (3D FFT), which is indispensable for 3D-RISM theory. These features allow the present approach to readily achieve a high-parallel performance and the computational time can be dramatically reduced. It is well recognized that the efficient parallelization of 3D FFT is generally difficult[7] and avoiding FFT is an effective technique for solvation structure computations.[8,9]

The organization of this paper is as follows. In Secs. II and III, the formalism of the present method and the computational details are respectively explained. In Sec. IV, the following three topics are described. First, the parallel performance of the present method is demonstrated. Second, the results calculated by the present method are compared with those by 3D-RISM in terms of hydration structure and partial molar volume (PMV). Finally, the hydration structure around

---

a)Electronic mail: hirofumi@moleng.kyoto-u.ac.jp.

a protein called Fv fragment is discussed. The computed results are compared to the high-quality results obtained by x-ray crystallography.

II. METHOD

A. Ornstein–Zernike-type equation

In the present theory, 3D total and direct correlation functions, H_α and C_α, are expressed as a sum of reference and residual correlation functions, as follows:

$$H_\alpha(\mathbf{r}) = H_\alpha^{\text{ref}}(\mathbf{r}) + \Delta H_\alpha(\mathbf{r}), \tag{1}$$

$$C_\alpha(\mathbf{r}) = C_\alpha^{\text{ref}}(\mathbf{r}) + \Delta C_\alpha(\mathbf{r}). \tag{2}$$

Here, $\mathbf{r}$ is a position vector in laboratory frame.[10] The reference correlation functions, H_α^{ref} and C_α^{ref} are defined with one dimensional (1D) correlation functions,

$$H_\alpha^{\text{ref}}(\mathbf{r}) \equiv \sum_{\beta\gamma} c_{\beta\gamma} * [\omega_{\gamma\alpha}^V + \rho h_{\gamma\alpha}^V](r_\beta), \tag{3}$$

$$C_\alpha^{\text{ref}}(\mathbf{r}) \equiv \sum_{\beta} c_{\beta\alpha}(r_\beta), \tag{4}$$

where $r_\beta = |\mathbf{r} - \mathbf{R}_\beta|$, $\mathbf{R}_\beta$ is the position of solute atomic site β, and ω^V is intramolecular correlation function of solvent. h^V is the total correlation function of solvent. The Ornstein–Zernike-type equation for the residual correlation functions, ΔH_α and ΔC_α, may be given by[11]

$$\Delta H_\alpha(\mathbf{r}) = \sum_{\beta} \Delta C_\beta * [\omega_{\beta\alpha}^V + \rho h_{\beta\alpha}^V](\mathbf{r}). \tag{5}$$

In the present method, the residual functions are expressed by the components on each solute atomic site,

$$\Delta H_\alpha(\mathbf{r}) = \sum_{\beta} w_\beta(\mathbf{r}) \Delta H_\alpha^{(\beta)}(\mathbf{r}_\beta), \tag{6}$$

$$\Delta C_\alpha(\mathbf{r}) = \sum_{\beta} w_\beta(\mathbf{r}) \Delta C_\alpha^{(\beta)}(\mathbf{r}_\beta), \tag{7}$$

where $w_\beta(\mathbf{r})$ is the weight function for solute atomic site β at the position $\mathbf{r}$. $\Delta H_\alpha^{(\beta)}$ and $\Delta C_\alpha^{(\beta)}$ are further expanded with real solid harmonics S_{lm} around atom center β, as follows:[11]

$$\Delta H_\alpha^{(\beta)}(\mathbf{r}_\beta) = \sum_{lm} \Delta H_{lm,\alpha}^{(\beta)}(r_\beta) S_{lm}(\hat{\mathbf{r}}_\beta), \tag{8}$$

$$\Delta C_\alpha^{(\beta)}(\mathbf{r}_\beta) = \sum_{lm} \Delta C_{lm,\alpha}^{(\beta)}(r_\beta) S_{lm}(\hat{\mathbf{r}}_\beta), \tag{9}$$

where $\hat{\mathbf{r}}_\beta$ is the unit vector directed along $\mathbf{r}_\beta$. A component of residual total correlation function is evaluated by the following equation:

$$\Delta H_{lm,\alpha}^{(\beta)}(r_\beta) = \sum_{\delta} \Delta C_{lm,\delta}^{(\beta)} * [\omega_{\delta\alpha}^V + \rho h_{\delta\alpha}^V](r_\beta). \tag{10}$$

B. Closure

In 1D-RISM and 3D-RISM frameworks, there are many closure equations. Kovalenko and Hirata (KH) proposed the following closures for 3D-RISM:[12]

$$H_\alpha(\mathbf{r}) = \begin{cases} \exp(\chi_\alpha(\mathbf{r})) - 1 & \text{for } \chi_\alpha(\mathbf{r}) \le 0, \\ \chi_\alpha(\mathbf{r}) & \text{for } \chi_\alpha(\mathbf{r}) > 0, \end{cases}$$

$$\chi_\alpha(\mathbf{r}) = -\beta u_\alpha(\mathbf{r}) + H_\alpha(\mathbf{r}) - C_\alpha(\mathbf{r}), \tag{11}$$

where $\beta = 1/k_B T$, k_B is Boltzmann's constant, and $u_\alpha(\mathbf{r})$ is the intermolecular potential function between solute and solvent site α. To solve Eqs. (8)–(10), we elaborated the following KH-type closure:

$$\Delta H_\alpha^{(\gamma)}(\mathbf{r}_\gamma)$$
$$= \begin{cases} \exp(\chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma)) - 1 - H_\alpha^{\text{ref}}(\mathbf{r}_\gamma) & \text{for } \chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma) \le 0, \\ \chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma) - H_\alpha^{\text{ref}}(\mathbf{r}_\gamma) & \text{for } \chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma) > 0, \end{cases}$$

$$\chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma) = -\beta u_\alpha(\mathbf{r}_\gamma) + \{H_\alpha^{\text{ref}}(\mathbf{r}_\gamma) + \Delta H_\alpha^{(\gamma)}(\mathbf{r}_\gamma)\}$$
$$- \{C_\alpha^{\text{ref}}(\mathbf{r}_\gamma) + \Delta C_\alpha^{(\gamma)}(\mathbf{r}_\gamma)\}, \tag{12}$$

where $u_\alpha(\mathbf{r}_\gamma)$ is the intermolecular potential function between solute and solvent site α, which is defined around solute site γ. The difference between $u_\alpha(\mathbf{r})$ in Eq. (11) and $u_\alpha(\mathbf{r}_\gamma)$ in Eq. (12) is only the origins of the vectors $\mathbf{r}$ and $\mathbf{r}_\gamma$ (see the Appendix for the details).

The procedure of the present method is summarized as follows. The reference correlation functions, H_α^{ref} and C_α^{ref}, are calculated by Eqs. (3) and (4) with the correlation functions obtained by 1D-RISM (Refs. 13–15) for the whole system (step 1). The residual correlation functions $\Delta H_\alpha^{(\beta)}(\mathbf{r}_\beta)$, $\Delta C_\alpha^{(\beta)}(\mathbf{r}_\beta)$ are then calculated on solute site β with Eqs. (8)–(10) and (12) (step 2). In the final step, the solvation structure $G_\alpha(\mathbf{r})$ $(=H_\alpha(\mathbf{r})+1)$ for the solvent site α is obtained by Eqs. (1), (3), and (6) (step 3). In this algorithm, step 2, which is most time demanding, is highly parallelized because of the application of the weight function $w_\alpha(\mathbf{r})$ for decomposition of correlation functions into one-site components and replacement of the 3D KH closure by the closure [Eq. (12)] for the one-site components.

III. COMPUTATIONAL DETAILS

In this work, logarithmic grid[16] for radial part and the Lebedev grid[17] for angular part were employed. The general importance of the logarithmic grid is to account for a molecular ordering that varies rapidly at short distances and slowly at long distances with the minimum numerical work.[18,19] With the grid set, the convolution in Eq. (10) was performed by spherical Bessel transformation.[16] The number of grid points in this work is 512 for radial part and 194 for angular part on each atom. For the weight function $w_\beta(\mathbf{r})$, Becke's function was employed.[20] Again, we would like to emphasize that 3D lattice grids and 3D FFT are not necessary in the present method. The calculation of the present method was parallelized with MPICH2.[21]

The geometries of the proteins were taken from the Protein Data Bank (PDB) data and those of amino acids from *Klotho* (biochemical compounds declarative database).[22] The potential functions for the amino acids and the proteins were united-atom model of optimized potential for liquid simulations (OPLS).[23] For solvent water molecule, simple point charge (SPC)-like model was employed[24] with a correction
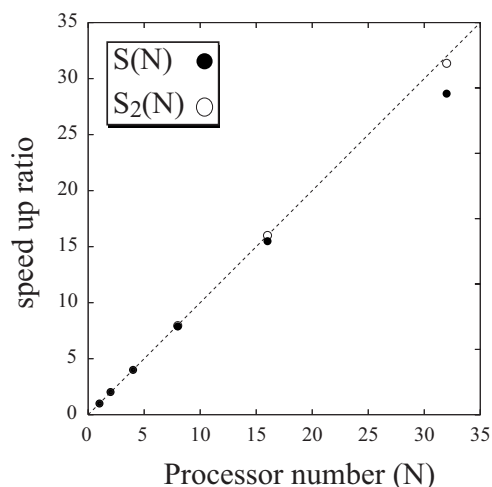
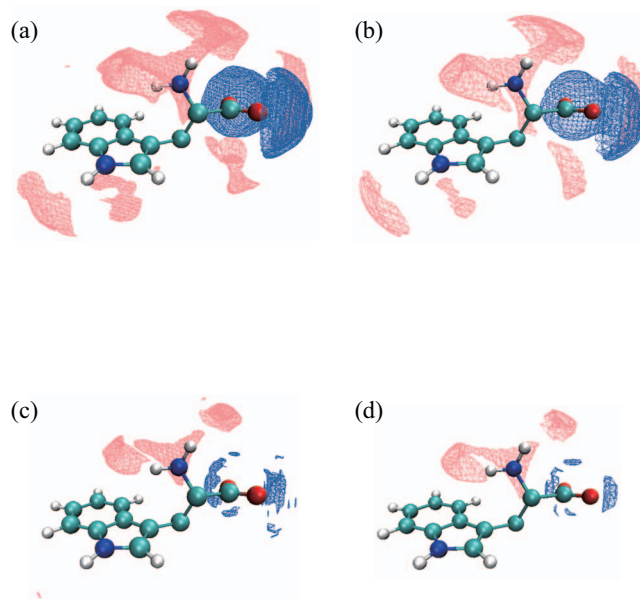FIG. 1. Speed up ratio as a function of the number of used processors (N).

FIG. 2. (Color) 3D hydration structures of oxygen site (pink) and hydrogen site (blue) around tryptophan calculated by the present method [(a) and (c)] and by 3D-RISM [(b) and (d)]. The isodensity surface of the solvation structure is 3.5 for (a) and (b) and 4.5 for (c) and (d).

concerning the Lennard-Jones parameters of the hydrogen sites ($\sigma=1.0$ Å, $\epsilon=0.056$ kcal/mol). The maximum number of angular momentum l in Eqs. (8) and (9) was 10 in this work. To visualize 3D solvation structure, VMD software was used.[25]

IV. RESULTS AND DISCUSSION

A. Computational performance of parallelization

The computational performance of the parallelization was evaluated with a speed up ratio. The benchmark calculation was performed using a small protein, Chignolin.[26]

Speed up ratio $S(N)$ using N processors is defined as

$$S(N) = \frac{\text{execution time with 1 processor}}{\text{execution time with } N \text{ processors}}. \quad (13)$$

If a program is completely parallelized, $S(N)=N$ (dashed line in Fig. 1), corresponding to linear speed up. Although the evaluated speed up ratio $S(N)$ is slightly smaller than N (filled circle in Fig. 1), the present method shows good performance. Since the execution time for the calculation is essentially a sum of the times for steps 1 and 2, the ratio of step 2 $S_2(N)$ was separately shown in Fig. 1, indicating the degradation of the computational performance mainly comes from step 1.

In step 1, the 1D correlation functions for all solute sites and solvent sites are evaluated at the same time. These functions among solute sites are integrated through the spherical convolution integral in 1D-RISM equation. Because transfer of very large data among each processor is required, the performance of this step is degraded. On the other hand, step 2 does not include the convolution integral between the solute sites, allowing to calculate the functions independently. Note that step 2 is dominative in the total time and the contribution from steps 1 is significantly small. As a consequence, virtually linear speed up is achieved.

B. Comparison of the present method with 3D-RISM

To evaluate the accuracy of the present method, comparison with 3D-RISM was performed on hydration structure $G_\alpha(\mathbf{r})$ ($\alpha=$O or H) and partial molar volume (PMV).

The hydration structures around tryptophan in zwitterionic form evaluated by the present method and by 3D-RISM were shown in Fig. 2. Because tryptophan is a molecule including both of hydrophobic group (benzene ring) and hydrophilic groups (NH_3^+, CO_2^-, and NH in pyrrole), it is a good example for the comparison. When the threshold of the isodensity is 3.5 [Figs. 2(a) and 2(b)], the broad distributions of water hydrogen (blue region) and of water oxygen (pink region) appear around hydrophilic groups. The shape of the distributions evaluated by the present method [Fig. 2(a)] shows good agreement with that obtained by 3D-RISM [Fig. 2(b)]. As increasing the threshold, strongly binding hydration waters selectively appear. In Figs. 2(c) and 2(d), the hydration structures with the threshold of 4.5 were shown. The distributions of water oxygen and hydrogen sites, which strongly bind with NH_3^+ and CO_2^-, can be seen in both figures. Therefore, it is concluded that the present method gives essentially the same distribution of hydration structure computed by 3D-RISM.[27] Preliminary estimated coordination number of solvent water around carboxylic oxygen site (O···H) is 3.6 for the present method and 3.2 for 3D-RISM. The numbers around ammonium nitrogen site (N···O) are 5.4 (the present method) and 5.1 (3D-RISM). The coordination numbers show reasonable agreement, clearly indicating again that the solvation structures evaluated by the two methods are essentially the same. It should be mentioned that the shape of solvation structures calculated by the present method did not change with higher quality grid (the number of grid points on each atom: 1024 for radial part and 230 for angular part) and the real solid harmonics up to $l=11$.
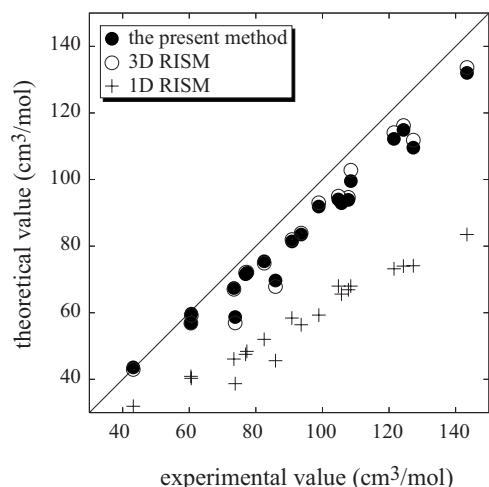
FIG. 3. PMVs of 20 amino acids calculated by 1D-RISM, the present method, and 3D-RISM in comparison with the corresponding experimental data (Refs. 30–33).

PMV of the solute molecule in molecular solvent, $\bar{V}_u^0$, is expressed in terms of the 3D solute-solvent direct correlation functions by the relation[28,29]

$$\bar{V}_u^0 = k_B T \chi_T^0 \left(1 - \rho \sum_\beta \int C_\beta(\mathbf{r}) d\mathbf{r}_\beta \right), \qquad (14)$$

where χ_T^0 is the isothermal compressibility of the pure solvent. By inserting Eq. (2) into Eq. (14), $\bar{V}_u^0$ is expressed with a sum of the reference part and the residual part, as follows:

$$\bar{V}_u^0 = k_B T \chi_T^0 \left(1 - \rho \sum_{\alpha\beta} \int c_{\alpha\beta}(\mathbf{r}_\gamma) d\mathbf{r}_\gamma \right)$$
$$- k_B T \chi_T^0 \rho \sum_\gamma \int \Delta C_\gamma(\mathbf{r}_\gamma) d\mathbf{r}_\gamma = \bar{V}_u^{\text{ref}} + \Delta \bar{V}_u. \qquad (15)$$

In Fig. 3, the PMVs of 20 amino acids in zwitterionic form calculated by 1D-RISM, the present method, and 3D-RISM were plotted in comparison with the corresponding experimental data.[30–33] If the theoretical values reproduce the corresponding experimental data exactly, the data are plotted on the solid line. As shown previously,[29] the PMVs calculated by 1D-RISM are much smaller than experimental data. On the other hand, the present method considerably improves the agreement with the experimental values. The accuracy is almost the same as that of 3D-RISM.

C. Hydration structure around a protein

The present method was applied to a protein called Fv fragment, which is a part of an anti-dansyl antibody. The protein was thoroughly studied with x-ray crystallography at low temperature and high-quality crystal structures were obtained in different states.[34,35] Because the hydration structures around them were also discussed in their works, it is a good system to compare the calculated hydration structure with experimental one. In this work, the hydration structures around the Fv fragment in unliganded state[34] and around the Fv fragment in complex state with ϵ-dansyl-L-lysine (DNS-lys) (Ref. 35) were evaluated. The geometries of the unliganded state and the complex state were taken from PDB data (PDBID: 2dlf.pdb and 1wz1.pdb). The molecular weights of these systems are 26 254 and 26 737 Da, respectively. All the water molecules were removed from the data before the computation. This calculation was done with 32 cores of PRIMEQUEST in Institute for Molecular Science (Okazaki Japan). The required time and memory per one core were 8.6 h and 0.9 GBytes for the Fv fragment in complex state.[27]

The hydration structures of water oxygen $G(\mathbf{r})$ around the unliganded state with different threshold of isodensity surfaces were shown in Fig. 4 together with the hydration waters determined by x-ray crystallography [Fig. 4(c)]. When the threshold is 2.5 [Fig. 4(a)], broad hydration structure is observed. The broadening of the distribution reflects the fluctuation of solvent waters at room temperature. The hydration structure with threshold 4.0 is considerably localized [Fig. 4(b)]. The distribution shows the area where waters bind strongly with the protein, corresponding to peaks in $G(\mathbf{r})$. These positions are in reasonable agreement with the experimental data, which were obtained at low temperature.

The Fv fragment has a binding pocket for DNS-lys and the structure around the site is much different between the unliganded state and the complex state. In Fig. 5, the calculated hydration structures around the binding site for these states were shown together with the experimentally determined waters (white spheres).[36] In the unliganded state, the binding pocket is surrounded by TYR96, HIS98, TYR99, and TYR102. Around the residues, there are many waters observed experimentally and the distributions calculated by the present method [green mesh area in Fig. 5(a)] agree well with the positions of them. In the case of the complex state, the residues flip away and the pocket opens. Along the flip, hydrophilic groups such as N and O are pushed away and hydrophobic groups such as benzene rings of TYR96 and TYR99 appear. The distribution of waters is seen around the
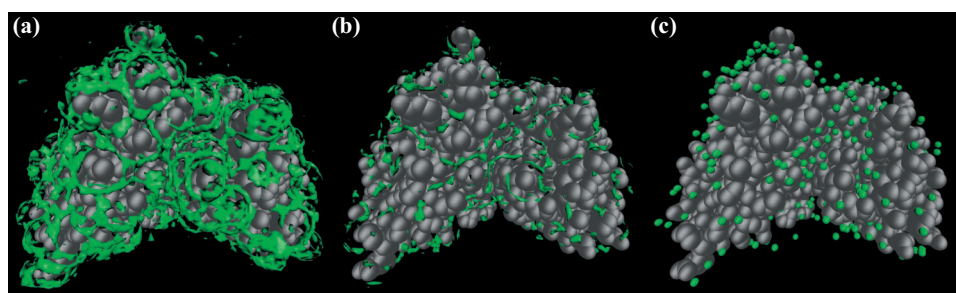

FIG. 4. (Color) 3D hydration structures of oxygen site; calculated by the present method (green area) [(a) and (b)] and determined by x-ray crystallography (green spheres) (c). The isodensity surfaces of these structures are 2.5 (a) and 4.0 (b), respectively. For sake of viewability, all of the residues are shown with black spheres.
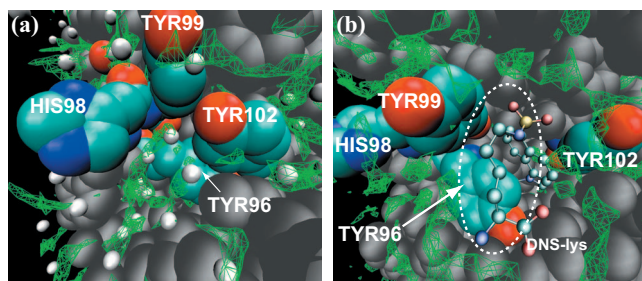
FIG. 5. (Color) Hydration structure of oxygen site (green area) around the binding pocket of DNS-lys for the unliganded state (a) and the complex state (b). The isodensity surface is 3.5. The residues except HIS98, TYR96, TYR99, and TYR102 are shown with black spheres. The red, blue, and cyan spheres correspond to oxygen, nitrogen, and carbon atoms, respectively. The water molecules determined by x-ray crystallography are shown with white spheres.

binding pocket and there is a hydrophobic space where DNS-lys binds [dotted circle in Fig. 5(b)], which is consistent with the previous work.[35]

In Fig. 6, the bottom of the binding pocket in the complex state was focused to see water oxygen (green) and water hydrogen (white). The positively charged water hydrogen sites are around N and O site of DNS-lys and TYR102, respectively, and the negatively charged water oxygen site is around H site of TYR102. These distributions indicate three hydrogen bondings between DNS-lys and TYR102 (dotted line in Fig. 6). These distributions correspond to the water experimentally observed by x-ray diffraction technique at low temperature.[35]

V. CONCLUSION

In this article, a new method to compute 3D solvation distribution function was proposed. The solvation structure is calculated with three steps; calculation of the reference correlation functions (1D functions, step 1), residual correlation functions (3D functions, step 2), and the building up of 3D solvation structure from the results obtained by the previous steps (step 3). Because the most time demanding step (step 2) is highly parallelized, the computational time can be re-
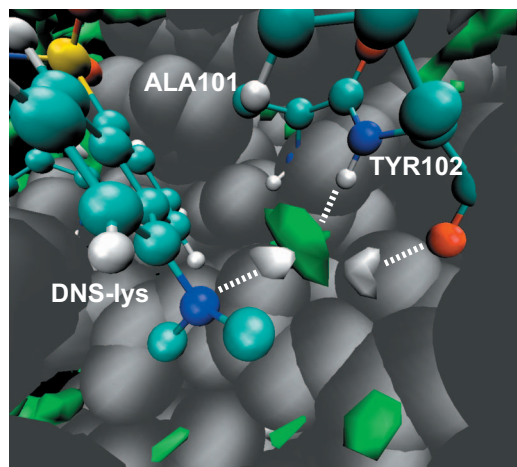

FIG. 6. (Color) Hydration structure of oxygen site (green area) and hydrogen site (white area) in the pocket of DNS-lys, where the isodensity is 3.0. The residues except ALA 101 and TYR102 are shown with black spheres.

duced dramatically. One of the most crucial and important difference between 3D-RISM and the present method is whether 3D FFT is required or not. The present method gave the almost the same results for the hydration structure of tryptophan and PMVs of amino acids evaluated by 3D-RISM.

The method was applied to the hydration structure around the binding pocket between Fv fragment and DNS-lys. By drawing the 3D solvation structure with small and large thresholds of isodensity, the fluctuation of waters and tightly binding waters can be respectively investigated. The calculated hydration structure shows good agreement with the experimentally observed results. By evaluating not only water oxygen sites but also water hydrogen sites, which are difficult to be determined by x-ray crystallography, the direction of the hydrogen bonding was clearly shown.

APPENDIX: Derivation of KH type [Eq. (12)]

Equation (11) is rewritten by using Eqs. (1), (2), (6), and (7),

$$\sum_\epsilon w_\epsilon(\mathbf{r})\Delta H_\alpha^{(\epsilon)}(\mathbf{r}_\epsilon)$$

$$= \begin{cases} \exp(\chi_\alpha(\mathbf{r})) - 1 - H_\alpha^{\text{ref}}(\mathbf{r}) & \text{for } \chi_\alpha(\mathbf{r}) \le 0, \\ \chi_\alpha(\mathbf{r}) - H_\alpha^{\text{ref}}(\mathbf{r}) & \text{for } \chi_\alpha(\mathbf{r}) > 0, \end{cases}$$

$$\chi_\alpha(\mathbf{r}) = -\beta u_\alpha(\mathbf{r}) + \left\{ H_\alpha^{\text{ref}}(\mathbf{r}) + \sum_\epsilon w_\epsilon(\mathbf{r})\Delta H_\alpha^{(\epsilon)}(\mathbf{r}_\epsilon) \right\}$$
$$- \left\{ C_\alpha^{\text{ref}}(\mathbf{r}) + \sum_\epsilon w_\epsilon(\mathbf{r})\Delta C_\alpha^{(\epsilon)}(\mathbf{r}_\epsilon) \right\}. \quad (A1)$$

Here we introduce an assumption that all the sites are independently determined, namely, $w_\epsilon(\mathbf{r})$ in the equation is replaced by $\delta_{\epsilon\gamma}$. Then Eq. 16 is given by

$$\Delta H_\alpha^{(\gamma)}(\mathbf{r}_\gamma)$$

$$= \begin{cases} \exp(\chi_\alpha^{(\gamma)}(\mathbf{r})) - 1 - H_\alpha^{\text{ref}}(\mathbf{r}) & \text{for } \chi_\alpha^{(\gamma)}(\mathbf{r}) \le 0, \\ \chi_\alpha^{(\gamma)}(\mathbf{r}) - H_\alpha^{\text{ref}}(\mathbf{r}) & \text{for } \chi_\alpha^{(\gamma)}(\mathbf{r}) > 0, \end{cases}$$

$$(A2)$$

$$\chi_\alpha^{(\gamma)}(\mathbf{r}) = - \beta u_\alpha(\mathbf{r}) + \{H_\alpha^{\text{ref}}(\mathbf{r}) + \Delta H_\alpha^{(\gamma)}(\mathbf{r}_\gamma)\}$$
$$- \{C_\alpha^{\text{ref}}(\mathbf{r}) + \Delta C_\alpha^{(\gamma)}(\mathbf{r}_\gamma)\}.$$

By setting the origin of the position vector $\mathbf{r}$ on the atomic site γ, the closure [Eq. (12)] is obtained. In fact, Eq. 17 is valid only around the atomic site γ since the weight function $w_\gamma(\mathbf{r})$ is virtually 1 in the vicinity of atomic site γ and gradually becomes zero far from the site. This treatment does not lead to any serious problems because the total and direct correlation functions are constructed essentially only from the contributions localized around each atomic site [Eqs. (1), (2), (6), and (7)].

[1] H. Luecke, B. Schobert, H.-T. Richter, J.-P. Cartailler, and J. K. Lanyi, J. Mol. Biol. **291**, 899 (1999).
[2] H. Kandori, Biochim. Biophys. Acta **1460**, 177 (2000).
[3] T. Imai, R. Hiraoka, A. Kovalenko, and F. Hirata, J. Am. Chem. Soc. **127**, 15334 (2005).
[4] A. Kovalenko and F. Hirata, Chem. Phys. Lett. **290**, 237 (1998).
[5] D. Beglov and B. Roux, J. Phys. Chem. B **101**, 7821 (1997).
[6] N. Yoshida, S. Phongphanphanee, and F. Hirata, J. Phys. Chem. B **111**, 4588 (2007).
[7] R. H. Bisseling, *Parallel Scientific Computation: A Structured Approach Using BSP and MPI* (Oxford University Press, Oxford, 2004).
[8] M. P. Sears and L. J. D. Frink, J. Comput. Phys. **190**, 184 (2003).
[9] G. N. Chuev and M. V. Fedorov, J. Comput. Chem. **25**, 1369 (2004).
[10] The origin of the position vector $\mathbf{r}$ can be set on the arbitrarily point at this moment. However, as discussed later the origin is reset to each atomic site.
[11] D. Yokogawa, H. Sato, and S. Sakaki, Chem. Phys. Lett. **432**, 595 (2006).
[12] A. Kovalenko and F. Hirata, J. Chem. Phys. **110**, 10095 (1999).
[13] D. Chandler and H. C. Andersen, J. Chem. Phys. **57**, 1930 (1972).
[14] F. Hirata and P. J. Rossky, Chem. Phys. Lett. **83**, 329 (1981).
[15] *Molecular Theory of Solvation*, edited by F. Hirata (Kluwer, Dordrecht, 2003).
[16] J. D. Talman, J. Comput. Phys. **29**, 35 (1978).
[17] V. I. Lebedev and D. N. Laikov, Dokl. Math. **59**, 477 (1999).
[18] J. M. Caillol, D. Levesque, and J. J. Weis, Mol. Phys. **44**, 733 (1981).
[19] P. H. Fries, W. Kunz, P. Calmettes, and P. Turq, J. Chem. Phys. **101**, 554 (1994).
[20] A. D. Becke, J. Chem. Phys. **88**, 2547 (1988).
[21] MPICH2, high-performance and widely portable implementation of MPI, http://www.mcs.anl.gov/research/projects/mpich2/index.php.
[22] B. Cannon, S. Sadekar, D. Taylor, B. H. Dunford-Shore, W. Sulaman, B. Feng, F. Fabrizio, J. Holcomb, W. Wise, and T. Kazic, http://www.biocheminfo.org/klotho/, 1994 for more information on biochemical compounds declarative database.
[23] W. L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc. **110**, 1657 (1988).
[24] 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).
[25] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).
[26] S. Honda, K. Yamasaki, Y. Sawada, and H. Morii, Structure (London) **12**, 1507 (2004).
[27] Since the computational time and required memories highly depend on the architecture of computer and the used library code (e.g., FFT), direct comparison between the present method and the original 3D-RISM is very difficult. In the present system, the method took less than one hour with 230 Mbytes memory per 1 core, performed by Core2 2 cores, 2.93 GHz (grids are 512 and 194, l is up to 10). The original 3D-RISM was performed by ITANIUM2, 1.4 GHz (256^3 grids) and took a few hours. It should be emphasized that both of them are much faster than standard molecular simulations.
[28] T. Imai, M. Kinoshita, and F. Hirata, J. Chem. Phys. **112**, 9469 (2000).
[29] Y. Harano, T. Imai, A. Kovalenko, M. Kinoshita, and F. Hirata, J. Chem. Phys. **114**, 9506 (2001).
[30] M. Kikuchi, M. Sakurai, and K. Nitta, J. Chem. Eng. Data **40**, 935 (1995).
[31] A. A. Zamyatnin, Annu. Rev. Biophys. Bioeng. **13**, 145 (1984).
[32] Y. Yasuda, N. Tochio, M. Sakurai, and K. Nitta, J. Chem. Eng. Data **43**, 205 (1998).
[33] M. Mizuguchi, M. Sakurai, and K. Nitta, J. Solution Chem. **26**, 579 (1997).
[34] M. Nakasako, H. Takahashi, N. Shimba, I. Shimada, and Y. Arata, J. Mol. Biol. **291**, 117 (1999).
[35] M. Nakasako, T. Oka, M. Mashumo, H. Takahashi, I. Shimada, Y. Yamaguchi, K. Kato, and Y. Arata, J. Mol. Biol. **351**, 627 (2005).
[36] Because the experimental data for the position of hydrated waters in unliganded state are not available in pdb data (1wz1.pdb), white spheres are missing in Fig. 5(b).
[37] A. Kovalenko and F. Hirata, J. Chem. Phys. **112**, 10391 (2000).