Supporting Information

Application of reference-modified density functional theory: Temperature and pressure dependences of solvation free energy

Tomonari Sumi[*1,2], **Yutaka Maruyama**[3], **Ayori Mitsutake**[4], **Kenji Mochizuki**[1], **and Kenichiro Koga**[1,2]

[1]*Research Institute for Interdisciplinary Science, Okayama University, 3-1-1 Tsushima-Naka, Kita-ku, Okayama 700-8530, Japan*

[2]*Department of Chemistry, Faculty of Science, Okayama University, 3-1-1 Tsushima-Naka, Kita-ku, Okayama 700-8530, Japan*

[3]*Co-Design Team, FLAGSHIP 2020 Project, RIKEN Advanced Institute for Computational Science, 7-1-26, Minatojima-minami-machi, Kobe 650-0047, Japan*

[4]*Department of Physics, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223–8522, Japan*

[*]E-mail: sumi@okayama-u.ac.jp

The RMDFT functional derived by applying the EDA to the reference HS system

In this study, we applied the effective-density approximation (EDA)[1] to the excess intrinsic free-energy functional for the reference HS system $F_{HS}^{ex}[n]$ in the RMDFT. It is relatively easy to perform the RMDFT calculation of the SFE if the EDA functional is employed as $F_{HS}^{ex}[n]$; it is because, the EDA functional is much simpler than the one obtained by, for example, the fundamental measure theory (FMT)[2] or its modified versions.[3,4] We believe that the values of SFE obtained by the RMDFT would not significantly be affected by the difference in the details of the free energy functional that is applied to the reference HS fluid, but rather depend on the HS diameter of the reference system.

Equation (14) is rewritten as follows if one uses the EDA functional for $F_{HS}^{ex}[n]$:

$$\begin{aligned}
\Delta G_{solv} = &-\frac{1}{\beta}\sum_{a=1}^{P}\int d\mathbf{r}_1^a \left[n_a\left(\mathbf{r}_1^a\middle|\left\{U_\lambda^{PR}\right\}\right) - n_0\right] + \int d\mathbf{r}_1^0 \Big[n_0\left(\mathbf{r}_1^0\middle|\left\{U_\lambda^{PR}\right\}\right) f_{HS}\left(n_0^{eff}\left(\mathbf{r}_1^0\middle|n_0\right)\right) - n_0 f_{HS}(n_0)\Big] \\
&- \int d\mathbf{r}_1^0 \Big[n_0\left(\mathbf{r}_1^0\middle|\left\{U_\lambda^{PR}\right\}\right) \Big\{f_{HS}\left(n_0^{eff}\left(\mathbf{r}_1^0\middle|n_0\right)\right) \\
&+ \int d\mathbf{r}_2^0\, W_{00}\left(\left|\mathbf{r}_1^0 - \mathbf{r}_2^0\right|\right) n_0\left(\mathbf{r}_2^0\middle|\left\{U_\lambda^{PR}\right\}\right) f_{HS}'\left(n_0^{eff}\left(\mathbf{r}_2^0\middle|n_0\right)\right)\Big\} \\
&- n_0 \Big\{f_{HS}(n_0) + n_0 f_{HS}'(n_0)\int d\mathbf{r}_2^0\, W_{00}\left(\left|\mathbf{r}_1^0 - \mathbf{r}_2^0\right|\right)\Big\}\Big] \\
&+ \frac{n_0}{\beta}\sum_{a=1}^{P}\sum_{b=1}^{P}\int d\mathbf{r}_1^a d\mathbf{r}_1^b\, C_{ab}^{ex}\left(\left|\mathbf{r}_1^a - \mathbf{r}_2^b\right|\right)\left[n_b\left(\mathbf{r}_1^b\middle|\left\{U_\lambda^{PR}\right\}\right) - n_0\right] \\
&+ \frac{1}{2\beta}\sum_{a=1}^{P}\sum_{b=1}^{P}\int d\mathbf{r}_1^a d\mathbf{r}_1^b\, C_{ab}^{ex}\left(\left|\mathbf{r}_1^a - \mathbf{r}_2^b\right|\right)\left[n_a\left(\mathbf{r}_1^a\middle|\left\{U_\lambda^{PR}\right\}\right) - n_0\right]\left[n_b\left(\mathbf{r}_1^b\middle|\left\{U_\lambda^{PR}\right\}\right) - n_0\right]
\end{aligned} \tag{A1}$$

Here, $f_{HS}'(n)$ is the first derivative of $f_{HS}(n)$, which is the excess free energy of the HS system per particle. A highly accurate expression for $f_{HS}(n)$, obtained from the Carnahan–Starling (CS) equation of state,[5,6] is available:

$$f_{HS}(n) \equiv \frac{A_{HS}^{ex}}{N} = \frac{1}{\beta}\frac{\eta(4-3\eta)}{(1-\eta)^2}, \tag{A2}$$

$$\eta = \pi n_0 (d_{HS})^3/6, \tag{A3}$$

where d_{HS} is the diameter of the reference HS fluid. $n_0^{eff}\left(\mathbf{r}_1^0\middle|n_0\right)$ in equation (A1) is defined by the EDA excess intrinsic free-energy functional for the reference HS system:

$$F_{HS}^{ex}[n_O] = \int d\mathbf{r}_1^0\, n_0\left(\mathbf{r}_1^0\middle|\left\{U_\lambda^{PR}\right\}\right) f_{HS}\left(n_0^{eff}\left(\mathbf{r}_1^0\middle|n_0\right)\right), \tag{A4}$$

where $n_0^{eff}(\mathbf{r}_1^0|n_0)$ is the effective density, which is assumed to be a functional of $n_0(\mathbf{r}|\{U_\lambda^{PR}\})$.

$n_0^{eff}(\mathbf{r}|n_0)$ is approximated by the first-order density functional Taylor series expansion:

$$n_0^{eff}(\mathbf{r}|n_0) = n_0 + \int d\mathbf{r}_1^0\, W_{00}(|\mathbf{r}-\mathbf{r}_1^0|)[n_0(\mathbf{r}_1^0|\{U_\lambda^{PR}\}) - n_0]. \tag{A5}$$

This expansion coefficient $W_{00}(r)$ appears in equation (A1) and is related to the second-order direct correlation function $C_{OO}^{HS}(r)$ for the reference HS fluid via

$$\widehat{W}_{00}(k) = \left[-2\beta f'_{HS}(n_0) + \sqrt{\left(2\beta f'_{HS}(n_0)\right)^2 - 4n_0\beta f''_{HS}(n_0)\hat{C}_{OO}^{HS}(k)}\right]/2n_0\beta f''_{HS}(n_0), \tag{A6}$$

where $\widehat{W}_{00}(k)$ and $\hat{C}_{OO}^{HS}(k)$ are the Fourier transforms of $W_{OO}(r)$ and $C_{OO}^{HS}(r)$, respectively, and $f''_{HS}(n)$ is the second derivative of $f_{HS}(n)$. The density functional differentiation of $F_{HS}^{ex}[n_O]$ appearing in equation (A1) is obtained from equation (A4) as follows:

$$\frac{\delta F_{HS}^{ex}[n_O]}{\delta n_O(\mathbf{r}|\{U_\lambda^{PR}\})} = f_{HS}\left(n_0^{eff}(\mathbf{r}|n_0)\right) + \int d\mathbf{r}_1^0\, W_{00}(|\mathbf{r}-\mathbf{r}_1^0|)n_0(\mathbf{r}_1^0|\{U_\lambda^{PR}\})f'_{HS}\left(n_0^{eff}(\mathbf{r}|n_0)\right). \tag{A7}$$

In this study, we calculated the site-site direct correlation functions $\{C_{\alpha\beta}(r)\}$ for bulk water and the site-density distribution functions of the solvent around the solute molecule $\{n_\alpha(\mathbf{r}|\{U_\lambda^{PR}\})\}$ using the 1D-RISM-KH and 3D-RISM-KH integral equations [7,8], respectively. Before we calculate the SFE using equation (A1) with the sets of $\{C_{\alpha\beta}(r)\}$ and $\{n_\alpha(\mathbf{r}|\{U_\lambda^{PR}\})\}$, it is necessary to determine $W_{00}(r)$ or $C_{OO}^{HS}(r)$ by solving the Ornstein-Zernike (OZ) integral equation,[5] in which the following EDA equation combined with the Percus' relation [5,9] is used as the closure:

$$H^{HS}(r) = n(\mathbf{r}|U_{PR}^{HS})/n_0 - 1, \tag{A8}$$

$$n(r|U_{PR}^{HS}) = n_0 exp(-\beta U^{IG}(r)), \tag{A9}$$

$$
\begin{aligned}
U^{IG}(r) &= U_{PR}^{HS}(r) + \frac{\delta F_{HS}^{ex}[n]}{\delta n(\mathbf{r}|\{U_{PR}^{HS}\})} - \mu^{exHS} \\
&= U_{PR}^{HS}(r) + f_{HS}\left(n_0^{eff}(r|n_0)\right) - f_{HS}(n_0) \\
&= \int d\mathbf{r_1}\, W(|\mathbf{r}-\mathbf{r_1}|)\left[n(\mathbf{r_1}|U_{PR}^{HS})f'_{HS}\left(n_0^{eff}(\mathbf{r_1}|n)\right) - n_0 f'_{HS}(n_0)\right], \qquad \text{(A10)}
\end{aligned}
$$

$$
\widehat{H}^{HS}(k) = \hat{C}^{HS}(k) + n_0\hat{C}^{HS}(k)\widehat{H}^{HS}(k). \qquad \text{(A11)}
$$

Here, $U_{PR}^{HS}(r)$ in the Percus' relation in equation (A8) is equal to the interaction potential between HS particles of the reference system.

3D-RISM calculation

We need the site-density distribution functions $n_\alpha\left(\mathbf{r}\middle|\left\{U_\lambda^{PR}\right\}\right)$ to calculate the SFE using eq. (A1). In the present study, we employed the three-dimensional reference-interaction-site-model (3D-RISM) theory [7,8] to calculate $n_\alpha\left(\mathbf{r}\middle|\left\{U_\lambda^{PR}\right\}\right)$.

The 3D-RISM theory is an integral equation theory that we can use to obtain the solvent distribution functions from the intermolecular potential functions and the thermodynamic conditions (i.e., temperature and density). It produces the distribution functions of the solvent molecules around the solute molecule. The calculations involved in this method have two steps. The first step is to calculate the pair correlation functions in the aqueous solution on the basis of the solvent-solvent 1D-RISM theory,[10,11] which represents the microscopic structure of the distribution of the water molecules. In the second step, a solute molecule is immersed into the solvent and the 3D distribution functions of the solvent molecule are calculated using the 3D-RISM integral equations. For an infinitely dilute solution, the solute-solvent 3D-RISM integral equations can be written as

$$\hat{h}_\gamma(\mathbf{k}) = \sum_{a=1}^{p} \hat{c}_a(\mathbf{k}) * \hat{\sigma}_{a\gamma}(k) + n_0 \sum_{a=1}^{p} \hat{c}_a(\mathbf{k}) * \hat{H}_{a\gamma}(k) \qquad \text{(A12)}$$

Here, $\hat{h}_\gamma(\mathbf{k})$ and $\hat{c}_\gamma(\mathbf{k})$ are the Fourier transforms of $h_\gamma(\mathbf{r})$ and $c_\gamma(\mathbf{r})$, which are the solute-solvent total and direct correlation functions between the solute molecule and solvent site γ. It is noted that eq. (A12) contains the solvent-solvent total correlation function $H_{\alpha\gamma}(r)$, which is defined by eq. (13) in the main text. Thus, one needs to solve eq. (13) before solving eq. (A12). We can calculate $n_\alpha\left(\mathbf{r}\middle|\left\{U_\lambda^{PR}\right\}\right)$ from $h_\gamma(r)$ using the Percus' relation, $n_\alpha\left(\mathbf{r}\middle|\left\{U_\lambda^{PR}\right\}\right) = n_0(h_\alpha(\mathbf{r})+1)$.[12,13]

Equations (13) in the text and (A12) both have two unknown functions, namely, the total and direct correlation functions. Therefore, an additional equation is required to solve them. We employed partially linearized HNC (PLHNC) equation,[7] which is called Kovalenko-Hirata (KH) equations, [8] as the closure:

$$H_{\alpha\gamma}(r) = \begin{cases} \exp(\chi_{\alpha\gamma}(r)) & \chi_{\alpha\gamma}(r) < 1 \\ \chi_{\alpha\gamma}(r) - 1 & \chi_{\alpha\gamma}(r) \geq 1 \end{cases} \qquad \text{(A13)}$$
$$\chi_{\alpha\gamma}(r) = -\beta v_{\alpha\gamma}(r) + H_{\alpha\gamma}(r) - C_{\alpha\gamma}(r)$$

and

$$h_\gamma(\mathbf{r}) = \begin{cases} \exp(\chi_\gamma(\mathbf{r})) & \chi_\gamma(\mathbf{r}) < 1 \\ \chi_\gamma(\mathbf{r}) - 1 & \chi_\gamma(\mathbf{r}) \geq 1 \end{cases} \qquad \text{(A14)}$$
$$\chi_\gamma(\mathbf{r}) = -\beta U_\gamma^{PR}\left(\mathbf{r}\middle|\left\{\mathbf{r}_1^\lambda\right\}\right) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})$$

where $v_{\alpha\gamma}(r)$ is the intermolecular interaction potential between sites α and γ on the solvent molecule, and $U_{\gamma}^{PR}\left(\mathbf{r}\middle|\left\{\mathbf{r}_1^{\lambda}\right\}\right)$ is given by the external field caused by the solute molecule having M interaction sites,

$$U_{\alpha}^{PR}\left(\mathbf{r}\middle|\left\{\mathbf{r}_1^{\lambda}\right\}\right)=\sum_{b=1}^{M}u_{\alpha b}\left(\left|\mathbf{r}-\mathbf{r}_1^{\lambda}\right|\right), \tag{A15}$$

where $u_{\alpha b}(r)$ is the intermolecular interaction potential between site α of the solvent molecule and site *b* of the solute molecule. We used eqs. (13) and (A13) for the first step and eqs. (A12) and (A14) for the second step. To solve the 1D-RISM and 3D-RISM integral equations, we respectively introduce the following difference functions:

$$T_{\alpha\gamma}(r)=H_{\alpha\gamma}(r)-C_{\alpha\gamma}(r) \tag{A16}$$

and

$$t_{\gamma}(\mathbf{r})=h_{\gamma}(\mathbf{r})-c_{\gamma}(\mathbf{r}) \tag{A17}$$

The numerical procedure for the first/second step can be briefly summarized as follows.

1. Calculate $u_{\alpha\gamma}(r)/U_{\gamma}^{PR}\left(\mathbf{r}\middle|\left\{\mathbf{r}_1^{\lambda}\right\}\right)$ at each 1D/3D grid point.
2. Initialize $T_{\alpha\gamma}(r)/t_{\gamma}(\mathbf{r})$ to zero.
3. Calculate $H_{\alpha\gamma}(r)/h_{\gamma}(\mathbf{r})$ using Eq. A13/A14.
4. Calculate $C_{\alpha\gamma}(r)/c_{\gamma}(\mathbf{r})$ from $C_{\alpha\gamma}(r)=H_{\alpha\gamma}(r)-T_{\alpha\gamma}(r)/c_{\gamma}(\mathbf{r})=h_{\gamma}(\mathbf{r})-t_{\gamma}(\mathbf{r})$.
5. Transform $C_{\alpha\gamma}(r)$ / $c_{\gamma}(\mathbf{r})$ to $\hat{C}_{\alpha\gamma}(k)$ / $\hat{c}_{\gamma}(\mathbf{k})$ using the 1D/3D fast Fourier transform (1D-FFT/3D-FFT).

6. Calculate $\hat{H}_{\alpha\gamma}(k) / \hat{h}_{\gamma}(\mathbf{k})$ using eq. 13/A12.

7. Invert $\hat{H}_{\alpha\gamma}(k) / \hat{h}_{\gamma}(\mathbf{k})$ to $H_{\alpha\gamma}(r) / h_{\gamma}(\mathbf{r})$ using the 1D-FFT/3D-FFT.

8. Calculate $T_{\alpha\gamma}(r) / t_{\gamma}(\mathbf{r})$ using Eq. A16/A17.

9. Calculate the new $T_{\alpha\gamma}(r) / t_{\gamma}(\mathbf{r})$ from $T_{\alpha\gamma}(r) / t_{\gamma}(\mathbf{r})$ and its history using an acceleration method.

10. Repeat steps 3–9 until the input and output functions become identical within the convergence tolerance.

We had proposed a few acceleration methods for solving these integral equations. In this study, a modified version of the Anderson method was employed.[14]

HS-diameter dependence of the SFEs for methane, propane, and isobutane

We investigated how the grid spacing used in the EDA calculation for the reference HS fluid affects the SFE value obtained by the RMDFT. Figure S1 shows HS-diameter dependences of the SFE for methane, propane, and isobutane calculated by the RMDFT using the grid spacing 0.01 and 0.00125 Å in the EDA calculation of the reference HS fluid. In Fig. S1, the SFE values calculated by using the fine grid spacing of 0.00125 Å seem to be shifted by 0.05 Å toward the smaller value of d_{HS} compared with those with the grid spacing of 0.01 Å. The optimal HS diameter by the fine grid spacing 0.00125 Å is determined as 2.875 Å that might correspond to the middle point between the discrete points of 2.87 and 2.88 Å provided by the grid spacing of 0.01 Å. Note that the radial distribution function $g(r)$ obtained by the EDA calculation with the grid spacing of 0.01 Å discontinuously changes from zero to nonzero between at the distances of 2.87 and 2.88 Å. These observations suggest that the SFE obtained by the RMDFT based on the EDA solution calculated by using the grid spacing 0.01 Å with the HS diameter of 2.88 Å approximately corresponds to the SFE obtained by the RMDFT based on the EDA solution calculated by using the HS diameter of 2.875 Å (i.e., (2.87+2.88)/2 Å) with a sufficiently fine grid spacing.

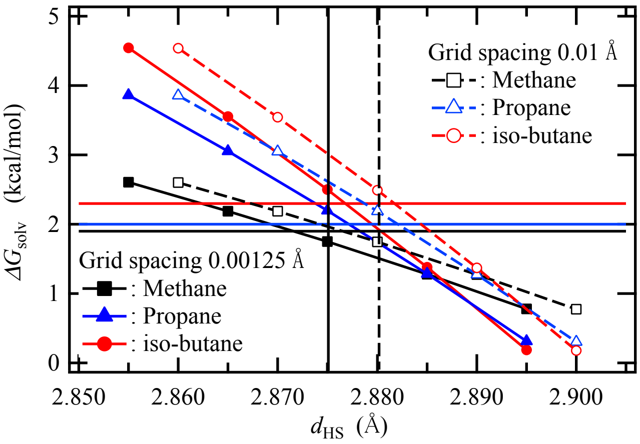

Figure S1. HS-diameter dependence of the SFE values for methane, propane, and isobutane obtained by the RMDFT based on the EDA solution calculated with the grid spacing 0.01 and 0.00125 Å. The horizontal lines with black, blue, and red indicate the experimental values of SFE for methane, propane, and isobutane, respectively. The perpendicular solid and dashed lines at d_{HS} of 2.875 and 2.88 Å indicate the optimal HS diameters determined with the use of the grid spacing 0.00125 and 0.01 Å, respectively.

Schematic illustration of the free energy surface of chignolin based on the experimental data

On the basis of the experimental data for temperature dependence of the molar fraction of the native state, f_N, we can estimate the difference of the free energy between the native and denatured states, $\Delta F/k_\mathrm{B}T=\left(F_D-F_N\right)/k_\mathrm{B}T$, where F_N and F_D are the free energy for the native and denatured states, respectively, using $\Delta F/k_\mathrm{B}T=-\ln\left[\left(1-f_N\right)/f_N\right]$. From Fig. 5 G in Ref. [15], the values of f_N are obtained as 0.6 at 298K and as 0.1 at 373K, as a result, we estimate $\Delta F/k_\mathrm{B}T$ to be about 0.4 at 298 K and –2.2 at 373 K. Based on these values for $\Delta F/k_\mathrm{B}T$, we can illustrate a schematic free energy surface as a function of the radius of gyration R_g of chignolin, $F\left(R_g\right)/k_BT$. In this figure, the values of $F\left(R_g\right)/k_BT$ for the native state at both the temperatures are assumed to be zero.

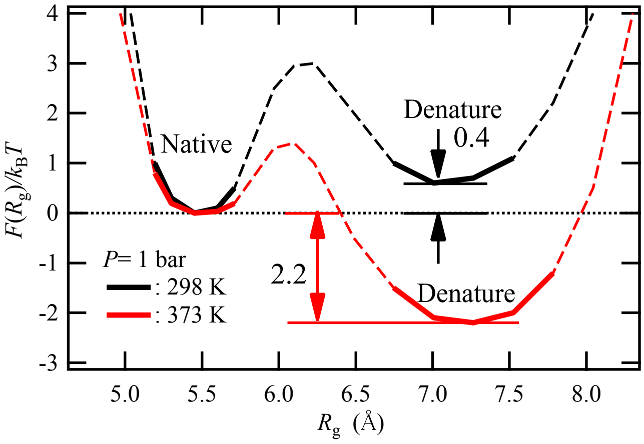

Figure S2. Schematic free energy surfaces of chignolin $F\left(R_g\right)/k_\mathrm{B}T$ as a function of the radius of gyration R_g at temperatures of 298 and 373 K. These are illustrated based on the experimental data on the difference in the free energy between for the native (N) state and for the denatured (D) state, i.e. $\Delta F/k_\mathrm{B}T=\left(F_D-F_N\right)/k_\mathrm{B}T$, that is, 0.4 at 298 K and -2.2 at 373 K [15]. The reason why most parts of the free energy surfaces are shown as broken lines is having no experimental data except for the difference in the free energy between at the native state and at the denatured state.

References

[1] T. Sumi, H. Sekino, *J. Phys. Soc. Jpn.* **2008**, *77*, 034605.

[2] Y. Rosenfeld, *Phys. Rev. Lett.* **1989**, *63*, 980–983.

[3] R. Roth, R. Evans, A. Lang, G. Kahl, *J. Phys.: Condens. Matter* **2002**, *14*, 12063–12078.

[4] Y.-X. Yu, J. Wu, *J Chem Phys* **2002**, *117*, 10156.

[5] J.-P. Hansen, I. R. McDonald, *Theory of Simple Liquids; 2nd Ed.*, Academic Press, London, **1986**.

[6] N. F. Carnahan, K. E. Starling, *The Journal of Chemical Physics* **1969**, *51*, 635–636.

[7] A. Kovalenko, F. Hirata, *The Journal of Chemical Physics* **1999**, *110*, 10095–10112.

[8] F. Hirata, Ed., *Molecular Theory of Solvation*, Kluwer Academic Publishers, Dordrecht, **2003**.

[9] J. K. Percus, *Phys. Rev. Lett.* **1962**, *8*, 462–463.

[10] F. Hirata, P. J. Rossky, *Chemical Physics Letters* **1981**, *83*, 329–334.

[11] F. Hirata, *J Chem Phys* **1982**, *77*, 509.

[12] T. Sumi, T. Imai, F. Hirata, *J Chem Phys* **2001**, *115*, 6653.

[13] T. Sumi, H. Sekino, *J Chem Phys* **2006**, *125*, 34509.

[14] Y. Maruyama, F. Hirata, *J. Chem. Theory Comput.* **2012**, *8*, 3015–3021.

[15] S. Honda, K. Yamasaki, Y. Sawada, H. Morii, *Structure* **2004**, *12*, 1507–1518.