PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2016, **18**, 32406

Statistical thermodynamics of aromatic–aromatic interactions in aqueous solution

Tomohiko Hayashi and Masahiro Kinoshita*

Received 31st August 2016,
Accepted 7th November 2016

DOI: 10.1039/c6cp06000e

www.rsc.org/pccp

To elucidate the interactions between aromatic rings, which are believed to play essential roles in a variety of biological processes, we analyze the water-mediated interactions between toluene molecules along face-to-face stacked (FF) and point-to-face T-shaped (TS) paths using a statistical-mechanical theory of liquids combined with a molecular model for water. The theory enables us to decompose each interaction into physically insightful components, revealing detailed hydration effects. The dimers (*i.e.*, molecules in contact with each other) formed in the FF and TS paths, which are referred to as "FF stacking" and "TS contact", respectively, share almost the same stability in vacuum. In water, however, the stability of the FF stacking increases whereas that of the TS contact decreases. By the energetic hydration effect, for the FF stacking, more than half of the London dispersion attractive interaction is cancelled out and the electrostatic repulsive interaction is significantly screened. Importantly, a large gain of water entropy occurs. For the TS contact, the London dispersion interaction is almost completely cancelled out and the electrostatic component of the water-mediated interaction becomes repulsive. It is accompanied by a water-entropy gain. The water-entropy effect is crucially important for the participation of aromatic side chains in the close packing of a protein as well as FF stacked arrangements of aromatic rings in the case of nucleotide base interactions. The term "π–π stacking" is inappropriate for the stacking in aqueous solution, because it sounds as if the London dispersion interaction was the only contributor to it as in vacuum.

1 Introduction

Planar moieties are frequently found in a monomeric or polymeric biomolecule. Good examples are the aromatic side chain (phenylalanine, tyrosine, or tryptophan) in a protein, an aromatic nucleobase in a polynucleotide such as DNA or RNA, and an aromatic steroid ring system in a steroid hormone. Generally, an aromatic ring possesses one or more circular π-bonds through which the electrons are delocalized above and below the ring (*i.e.*, aromaticity).[1,2] Structural analyses of the biomolecules or their complexes showed that a variety of biological self-assembly and molecular-recognition processes (*e.g.*, protein folding, double strand formation of DNA, and hormone–receptor binding) accompany the contact and/or stacking of the aromatic rings. It is widely recognized that the London dispersion attractive interaction between aromatic rings, which is usually referred to as "π–π interaction" in the biological research community, plays a dominant role in the binding properties of nucleic acids and the structural stability of proteins.[3,4]

It has been known since a long time ago that the interaction between nonpolar, polar, or ionic solutes in water is substantially different from that in vacuum.[5–9] The interaction in water is represented in terms of the potential of mean force (PMF): the water-mediated interaction defined as the sum of direct and water-induced interactions (the direct interaction is the interaction in vacuum). However, the water-induced interaction is largely dependent on hydration properties of the solutes and quite variable. Despite a number of theoretical and computer-simulation studies being performed, there are still a lot of unresolved issues and controversial aspects to overcome. This is because the water-induced interaction is simultaneously influenced by multiple physicochemical factors. The interaction between aromatic rings mentioned above is no exception in the sense that their hydration properties play an imperative role.

Jorgensen and Severance[10] studied orientationally averaged benzene–benzene interactions in liquid benzene and in water using a Monte Carlo (MC) simulation. Afterwards, the face-to-face stacking and the T-shaped contact of aromatic rings have attracted much attention as two fundamental patterns.[3] Face-to-face stacked arrangements of aromatic rings are generally observed in nucleotide base interactions, whereas T-shaped conformations are more favored in the native structures of proteins.[3] Linse[11] calculated the face-to-face stacked and point-to-face T-shaped interactions (see Fig. 1) between benzene molecules in vacuum and in water (the molecules are immersed in water at infinite dilution) using a molecular dynamics (MD) simulation. In the calculation, the angle (θ) between normals to the rings is fixed and the distance

Institute of Advanced Energy, Kyoto University, Gokasho, Uji, Kyoto, 611-0011, Japan. E-mail: kinoshit@iae.kyoto-u.ac.jp; Fax: +81 774 38 4695; Tel: +81 774 383503

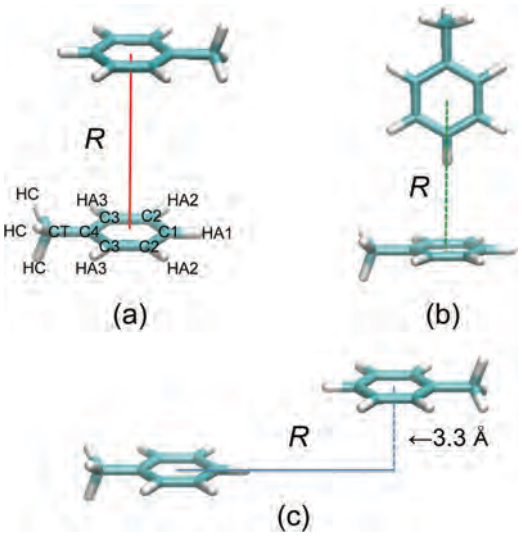

Fig. 1 Paths along which two toluene molecules approach each other. (a) Face-to-face stacked (FF) path. (b) Point-to-face T-shaped (TS) path. (c) Sliding path. In (a) and (b), R is the distance between ring centroids. In (c), R is defined as shown in the figure. The atom names and corresponding nonbonded potential parameters used are listed in Table 1.

(R) between ring centroids is gradually decreased toward the dimer formation ("dimer" represents molecules in contact with each other). The cases of $\theta = 0^\circ$ and $\theta = 90^\circ$, respectively, correspond to the face-to-face stacked and point-to-face T-shaped interactions. A significant conclusion is that the T-shaped contact is more stable than the face-to-face stacking as the dimer. (The dimers formed by FF stacking and TS contact are simply referred to as "FF stacking" and "TS contact", respectively.)

Chipot *et al.*[12] showed that toluene–toluene interactions are substantially different from benzene–benzene interactions not only in vacuum but also in water. An MD simulation was employed. As they emphasized, toluene is more suitable as a model of an aromatic ring or side chain found in the biomolecules. Compared to the interaction in vacuum comprising Lennard-Jones (LJ) and electrostatic terms, that in water (*i.e.*, PMF) is characterized by the following: it is oscillatory with a main periodicity of d_S (d_S is the molecular diameter of water = 2.8 Å); a repulsive peak appears at a surface separation of $h \sim 0.6d_S$; and it possesses a second minimum at $h \sim d_S$. An important result for toluene–toluene interactions is that the face-to-face stacking is as stable as the T-shaped contact in vacuum, but the former is considerably more stable than the latter in water. Chelli *et al.*[13] compared the PMFs (functions of θ and R) between toluene and *p*-cresol molecules in CCl_4, MeOH, and water by means of an MD simulation. They found that, unlike in vacuum and in the other two solvents, there is a global free-energy minimum at $(\theta, R) = (\theta^*, R^*)$ ($0^\circ < \theta^* < 10^\circ$ and 3.5 Å $< R^* <$ 4.0 Å) which approximately corresponds to the face-to-face stacking. However, the physical origins of the PMF in water reported by Chipot *et al.*[12] and Chelli *et al.*[13] have not been clarified yet. We note that benzene and toluene can be categorized as rather hydrophobic solutes.

A statistical-mechanical theory, which enables us to decompose the PMF into physically insightful components, is much more suited to the elucidation of the hydration effects than the MC and MD simulations. In this study, we analyze the toluene–toluene PMF in water by employing the three-dimensional reference interaction site model (3D-RISM) theory.[14,15] It has been applied to important problems in biophysics and biochemistry such as the hydration of peptides and proteins,[16,17] receptor–ligand binding,[18,19] association of proteins,[20,21] and discrimination of the relative propensities of proteins to aggregate.[22] A nice review on the theory and its applications has been given in a recent article.[23] We consider three different paths including the face-to-face stacked and point-to-face T-shaped paths. It is shown that the hydration effects on the PMF reported by Chipot *et al.*[12] can well be reproduced by the theory. The water-induced interaction included in the PMF is calculated and decomposed into energetic and entropic components. The energetic component is further decomposed into van der Waals (vdW) and electrostatic components: the two components are added to the direct vdW and electrostatic interactions, respectively. As in the MD simulation, the London dispersion interaction between toluene molecules is expressed by the direct vdW interaction, the attractive part of the LJ one.

Some of our important results can be summarized as follows. The toluene–toluene interactions are substantially influenced by water both energetically and entropically. The London dispersion interaction is made significantly weaker by the energetic hydration effect. In the entropic component of the PMF, regions where the change is negative and positive, respectively, appear alternately. The physical origin of this behavior is discussed in terms of the two principal geometric measures, the excluded volume and the water-accessible surface area of the pair of toluene molecules. A negative value of the entropic component implies that the water crowding in the bulk is more serious than in the case where toluene molecules are infinitely separated. The face-to-face stacking is highly stabilized in water: the stabilizing free energy is $\sim -6k_BT$ (k_B is the Boltzmann constant and T is the absolute temperature), and a water-entropy gain makes a significantly large contribution to this stabilization. The water-entropy gain originates from an increase in the total volume available for the translational displacement of water molecules in the system. We find a path along which the orientations of toluene molecules are fixed and the stacking can be reached with essentially no free-energy barrier. The path is distinct from the face-to-face stacked one. An essential point is that the ways of stabilizing the stacking are quite different from those in vacuum: a large water-entropy gain is crucial in water. The term, "π–π stacking", is frequently used for the face-to-face stacking of aromatic rings in biological systems. However, this terming is misleading because it is associated solely with the London dispersion interaction. We also discuss the relevance of the results obtained to intermolecular interactions involving aromatic rings in biological processes.

2 Theoretical method

2.1 Model

We adopt an all-atom model for a toluene molecule. A pair of molecules or an isolated molecule is immersed in water.

The nonbonded potential parameters for toluene are taken from those used in the MD simulation by Chipot *et al.*[12] and listed in Table 1. The LJ parameters are those of the general Amber force field (GAFF).[24] The London dispersion interaction between toluene molecules is expressed as the attractive, vdW part of the LJ interaction. The bond lengths between atoms used in this study are slightly different from those in the MD simulation (the maximum difference is only ~1%) for the following reason. In optimizing the structure of a toluene molecule using a quantum chemical calculation, the Hartree–Fock approximation used by Chipot *et al.* is replaced by the second-order Møller–Plesset approximation taking account of the many-body electron correlation.

For the water model, TIP3P[25] was adopted in the MD simulation. When a statistical-mechanical theory like the 3D-RISM theory is employed, LJ potential parameters must newly be assigned to the hydrogen sites to obtain converged solutions of the basic equations. Thus, it is not possible to compare the MD and 3D-RISM results using completely the same water model. On the other hand, the water model referred to as "cSPC/E"[26] was recommended as the one best suited to the 3D-RISM theory: it is a modified version of the extended single point charge model (SPC/E)[27] with assignment of LJ potential parameters to the hydrogen sites. We believe that the thermodynamic quantities of hydration calculated are not significantly dependent on the water model as long as popular models such as TIP3P,[25] TIP4P,[28] SPC/E,[27] and related versions are employed (this was actually verified by an MD simulation for LJ solutes[29]).

2.2 Three paths for calculating toluene–toluene interactions

Fig. 1 illustrates three different paths considered for calculating the interactions between toluene molecules with fixed orientations. In the figure, paths (a) and (b) correspond to the face-to-face stacked and point-to-face T-shaped interactions, respectively. The angle (θ) between normals to the aromatic rings is fixed and the distance (R) between ring centroids is gradually decreased toward the dimer formation. In the additional path, path (c) (this can be referred to as the "sliding path"), R is differently defined as explained in the figure. We note that $\theta = 0^\circ$ in paths (a) and (c) and $\theta = 90^\circ$ in path (b).

Table 1 Nonbonded potential parameters used in this study (also see Fig. 1) (ecu denotes "elementary charge unit")

Molecule	Atom name/ atom type	Charge (ecu)	Lennard-Jones parameters σ (Å)	ε (kcal mol^{-1})
Toluene	C1/ca	−0.189	3.816	0.0860
	HA1/ha	0.151	2.918	0.0150
	C2/ca	−0.128	3.816	0.0860
	HA2/ha	0.147	2.918	0.0150
	C3/ca	−0.279	3.816	0.0860
	HA3/ha	0.158	2.918	0.0150
	C4/ca	0.353	3.816	0.0860
	CT/c3	−0.574	3.816	0.1094
	HC/hc	0.154	2.974	0.0157
cSPC/E water	OW/OW	−0.8476	3.1658	0.15530
	HW/OH	0.4238	1.1658	0.01553

R considered in paths (a) and (b) is in the range from 8.0 to 3.0 Å with a decrement of 0.1 Å. R considered in path (c) is in the range from 8.0 to 0 Å with the same decrement. The face-to-face stacking in water occurs at $R \sim 3.3$ Å in path (a) and at $R = 0$ Å in path (c). The T-shaped contact in water occurs at $R \sim 4.9$ Å in path (b).

2.3 Free-energy function

For a solute molecule immersed in water at infinite dilution, we define the free-energy function G_{water} as

$$G_{\mathrm{water}} = E_{\mathrm{C}} - TS_{\mathrm{C}} + \mu_{\mathrm{H}} \quad (1)$$

where E_{C}, S_{C}, and μ_{H} are the conformational (solute intramolecular) energy, conformational entropy, and hydration free energy (HFE) of the solute, respectively, and T is the absolute temperature. The quantity μ_{H} is given by

$$\mu_{\mathrm{H}} = \varepsilon_{\mathrm{VH}} - TS_{\mathrm{VH}} \quad (2)$$

where $\varepsilon_{\mathrm{VH}}$ and S_{VH} are the hydration energy and entropy, respectively, and the subscript "VH" denotes hydration under isochoric conditions. The quantity $\varepsilon_{\mathrm{VH}}$ comprises the solute–water interaction energy generated upon solute insertion and the energy change due to the structural reorganization of water especially near the solute. S_{VH} denotes the change in water entropy upon solute insertion. We note that μ_{H} is independent of the solute insertion conditions, isobaric or isochoric. Substituting eqn (2) into eqn (1) yields

$$G_{\mathrm{water}} = E_{\mathrm{C}} - TS_{\mathrm{C}} + \varepsilon_{\mathrm{VH}} - TS_{\mathrm{VH}}. \quad (3)$$

G_{water} is independent of the solute insertion conditions.

E_{C}, which is calculated using a molecular mechanical potential, can be decomposed into the bonded and nonbonded components as

$$E_{\mathrm{C}} = E_{\mathrm{B}} + E_{\mathrm{vdW}} + E_{\mathrm{ES}} \quad (4)$$

where E_{B} is the bonded energy comprising the bond-stretching, angle-bending, and torsional terms, E_{vdW} is the vdW interaction energy, and E_{ES} is the electrostatic interaction energy. In a strict sense, E_{vdW} arises from the LJ interaction potential. However, the repulsive part is a minor contributor except in the region where the core repulsion is dominant and the corresponding configuration of toluene molecules is hardly accessible: It is referred to as "vdW interaction energy". (The nonbonded potential parameters are listed in Table 1.) We decompose $\varepsilon_{\mathrm{VH}}$ as

$$\varepsilon_{\mathrm{VH}} = \varepsilon_{\mathrm{VH,vdW}} + \varepsilon_{\mathrm{VH,ES}} \quad (5)$$

where $\varepsilon_{\mathrm{VH,vdW}}$ and $\varepsilon_{\mathrm{VH,ES}}$ are the vdW and electrostatic contributions to $\varepsilon_{\mathrm{VH}}$, respectively. The decomposition of $\varepsilon_{\mathrm{VH}}$ is performed as follows: first, we calculate the hydration energy of a hypothetical solute molecule whose partial charges are all switched to zero, $\varepsilon_{\mathrm{VH,vdW}}$; second, we obtain $\varepsilon_{\mathrm{VH,ES}}$ from $\varepsilon_{\mathrm{VH,ES}} = \varepsilon_{\mathrm{VH}} - \varepsilon_{\mathrm{VH,vdW}}$.

Substituting eqn (4) and eqn (5) into eqn (3) yields

$$G_{\mathrm{water}} = E_{\mathrm{B}} + (E_{\mathrm{vdW}} + \varepsilon_{\mathrm{VH,vdW}}) + (E_{\mathrm{ES}} + \varepsilon_{\mathrm{VH,ES}}) - TS_{\mathrm{C}} - TS_{\mathrm{VH}}. \quad (6)$$

On the other hand, the free-energy function for a solute molecule in vacuum G_{vacuum} is expressed as

$$G_{vacuum} = E_B + E_{vdW} + E_{ES} - TS_C. \quad (7)$$

2.4 Calculation of hydration free energy, entropy, and energy

The solute, a pair of toluene molecules or an isolated toluene molecule, is inserted into the model water under isochoric conditions at infinite dilution. S_{VH} and ε_{VH} are calculated using the 3D-RISM theory.[14,15,23] In this theory, a solvent molecule is represented by atomic sites referred to as ''interaction sites'' (*i.e.*, a water molecule has three sites: an oxygen and two hydrogens). A water–water site–site correlation function is dependent only on the distance between centers of the two interaction sites. In the present study, cSPC/E[26] is employed as the water model (see Table 1). The water–water site–site correlation functions are first calculated using the dielectrically consistent RISM (DRISM) theory.[30,31] The correlation function between the solute and each interaction site of water is then calculated using the 3D-RISM theory.

T = 298.15 K, ε_S = 78.4, and $\rho_S d_S^3$ = 0.7317 (d_S = 2.8 Å) form part of the input data. Here, ε_S, ρ_S, and d_S denote the dielectric constant, number density, and molecular diameter of water. The basic equations of the 3D-RISM theory are numerically solved on a 3D cubic grid. The grid spacing (Δx, Δy, and Δz) is set at 0.5 Å, and the minimum distance between the solute and each edge of the solvent box is set at 10 Å. It has been confirmed that the spacing is sufficiently small and the box size ($N_x\Delta x$, $N_y\Delta y$, $N_z\Delta z$) is large enough for the result obtained to be identical within convergence tolerance. We use the Singer–Chandler formula[32] for calculating μ_H. S_{VH} is evaluated through the numerical differentiation of μ_H with respect to T as

$$S_{VH} = -(\partial\mu_H/\partial T)_V \sim -\{\mu_H(T+\Delta T) - \mu_H(T-\Delta T)\}/(2\Delta T),\ \Delta T = 5\ \text{K} \quad (8)$$

where the subscript ''V'' denotes the differentiation under isochoric conditions. The quantity ε_{VH} is obtained from $\varepsilon_{VH} = \mu_H + TS_{VH}$.

We employ the hypernetted-chain (HNC) approximation[33] in the closure equations. The Kovalenko–Hirata (K–H) closure[15] has widely been employed instead of the HNC closure, especially in the 3D-RISM theory. This is because the HNC closure often gives rise to divergence in the numerical solution explained above unless the partial charges of the solute atoms are sufficiently small. However, we have recently suggested that the HNC closure be adopted when it gives convergence in the numerical solution. For example, the results obtained from the HNC closure are significantly better than those obtained from the K–H closure in the calculation of thermodynamic quantities of hydration (*e.g.*, hydration free energy, energy, and entropy) of a nonpolar solute.[34]

2.5 Calculation of potential of mean force

Let us consider the paths along which two toluene molecules approach in vacuum and in water (see Fig. 1). The toluene–toluene interaction $\Phi(R)$ is defined as

$$\Phi(R) = G(\text{toluene–toluene}, R) - \{G(\text{toluene}) + G(\text{toluene})\} \quad (9)$$

where G(toluene–toluene, R) and G(toluene) are the free-energy function of a pair of toluene molecules at R on a path and that of an isolated toluene molecule, respectively. $G = G_{water}$ in water and $G = G_{vacuum}$ in vacuum.

The PMF in water $\Phi_{water}(R)$ is expressed as

$$\Phi_{water}(R) = \Delta\{E_{vdW}(R) + \varepsilon_{VH,vdW}(R)\} + \Delta\{E_{ES}(R) + \varepsilon_{VH,ES}(R)\} - T\Delta S_{VH}(R), \quad (10)$$

where ''Δ'' signifies the value for a pair of toluene molecules at R on a path relative to that at $R \to \infty$. $\Delta\{E_{vdW}(R) + \varepsilon_{VH,vdW}(R)\}$, $\Delta\{E_{ES}(R) + \varepsilon_{VH,ES}(R)\}$, and $-T\Delta S_{VH}(R)$ are the vdW-energetic, electrostatic-energetic, and water-entropic contributions to $\Phi_{water}(R)$, respectively. For all R, $\Delta E_B = 0$ because the toluene molecules are not covalently bonded and $-T\Delta S_C = 0$ because the conformations of toluene molecules remain unchanged. The interaction in vacuum $\Phi_{vacuum}(R)$ is given by

$$\Phi_{vacuum}(R) = \Delta E_{vdW}(R) + \Delta E_{ES}(R). \quad (11)$$

We note that $\Phi_{vacuum}(R)$ is simply the interaction potential (*i.e.*, direct interaction) between two toluene molecules with fixed orientations.

We are concerned primarily with the signs and magnitudes of $\Phi_{vacuum}(R)$, $\Delta E_{vdW}(R)$, and $\Delta E_{ES}(R)$ in vacuum and those of $\Phi_{water}(R)$, $\Delta\{E_{vdW}(R) + \varepsilon_{VH,vdW}(R)\}$, $\Delta\{E_{ES}(R) + \varepsilon_{VH,ES}(R)\}$, and $-T\Delta S_{VH}(R)$ in water.

2.6 Reliability of the 3D-RISM results

The 3D-RISM theory is capable of handling a solute with a polyatomic structure immersed in water based on a molecular model without mathematical complexity. However, the reliability of the calculation results is dependent on the solute properties and the subjects discussed. In what follows, we argue that the results presented in this study should sufficiently be reliable.

The theory is not good at elucidating the temperature and pressure dependences of the solute hydration.[34,35] For example, it cannot reproduce the weakening of the hydrophobic effect at low temperatures manifested by, for example, cold denaturation of a protein. The theory predicts for a protein that the so-called swelling structure (*i.e.*, pressure-denatured structure) is more stable than the native structure even at normal pressure. However, it gives successful results in significantly many other cases.[14–23] We summarize some of the important remarks for the RISM and related theories[14,15,30,31] as follows.

(1) When the HNC closure gives converged solutions of the basic equations of the theory, it should be used instead of the K–H closure to obtain better values of the thermodynamic quantities of hydration.[34]

(2) Still, the theory tends to give too high a value of the HFE of a solute. However, the difference between two values of the HFE, which are calculated for two different solute structures or for a solute structure in pure water and that in salt solution, becomes much more reliable due to the cancellation of errors. For example, the theory can reproduce the effect of salt addition

on the solubility of a noble gas in water[36] and the change in the HFE upon protein folding.[17]

(3) As the solute hydrophobicity decreases, the calculation results become more reliable.[34] When solute–water electrostatic and vdW interactions are present, as they become stronger, the theory gives more reliable results.

In this study, we do not consider the temperature and pressure dependences of the solute hydration. We employ the HNC closure. Only the difference between two values of the HFE, those for a solute molecule and a pair of solute molecules, is discussed. The hydrophobicity of benzene or toluene is rather low (much lower than that of alkane). The solubility (in mol L^{-1}) of benzene into water is about 15, 30, and 200 times higher than the solubilities of methane, cyclohexane, and hexane, respectively.[37] The solubility of toluene into water is as high as $\sim$1/3.5 of that of benzene.[37] Thus, toluene as well as benzene are not highly hydrophobic. Taken together, the 3D-RISM results presented in this study should sufficiently be reliable (see Section 3.2 in which qualitative and quantitative comparisons are made between the 3D-RISM and MD results).

3 Results and discussion

3.1 Aromatic–aromatic interaction in vacuum

Unless otherwise specified, our discussion is focused on the face-to-face stacked (FF) and point-to-face T-shaped (TS) paths. Fig. 2(a) shows the curves of $\Phi_{vacuum}(R)$ along the FF and TS paths. The two curves are almost indistinguishable from those reported by Chipot *et al.*,[12] respectively. $\Phi_{vacuum}(R)$ becomes progressively lower as R decreases (except in the region where the core repulsion in the direct interaction is dominant) and has the global minimum value of -2.53 kcal mol^{-1} at $R = R_{VM} \sim 3.6$ Å in the FF path and that of -2.39 kcal mol^{-1} at $R = R_{VM} \sim 5.1$ Å in the TS path (the subscripts "V" and "M" denote "vacuum" and "minimum"), respectively.

The presence of π-electrons may result in considerable polarization of aromatic substances accompanied by a many-body effect. This effect cannot be taken into consideration by the simplified force-field method employed in the MD and 3D-RISM calculations. Chipot *et al.*[12] compared the results obtained from the AMBER force-field method and from a high-quality *ab initio* method at the second-order Møller–Plesset level of approximation combined with the 6-31+G(2d,p) basis set. The latter method takes account of the many-body effect. They found the following: in both of the FF and TS paths, $\Phi_{vacuum}(R_{VM})$ from the simplified force-field method is higher by $\sim$0.8 kcal mol^{-1} than that from the *ab initio* method; and the difference between the values of $\Phi_{vacuum}(R_{VM})$ in the two paths, which is the most important quantity, remains unchanged by the omission of the many-body effect.

The vdW ($\Delta E_{vdW}(R)$) and electrostatic ($\Delta E_{ES}(R)$) contributions to $\Phi_{vacuum}(R)$ are shown in Fig. 2(b) and (c), respectively. We note that "vdW" represents the "London dispersion" in our model. $\Delta E_{vdW}(R_{VM})$ (the value of $\Delta E_{vdW}(R)$ at $R = R_{VM}$) is -4.50 kcal mol^{-1} in the FF path and -1.73 kcal mol^{-1} in the TS path. The former is much lower than the latter because the number of contacting atoms in the FF dimer is larger than that in the TS dimer. $\Delta E_{ES}(R)$ is positive for all R and becomes higher as R decreases in the FF path, whereas it is negative for all R and becomes lower as R decreases in the TS path. The values of $\Delta E_{ES}(R_{VM})$ are 1.97 and -0.66 kcal mol^{-1} for the FF and TS dimers, respectively. Though contacts of like-charged atoms (C and C as well as H and H) occur in the FF dimer, those of unlike-charged atoms (C and H) occur in the TS dimer. The values of $\Phi_{vacuum}(R_{VM})$ in the FF and TS dimers are almost the same, but the vdW and electrostatic contributions are significantly different.

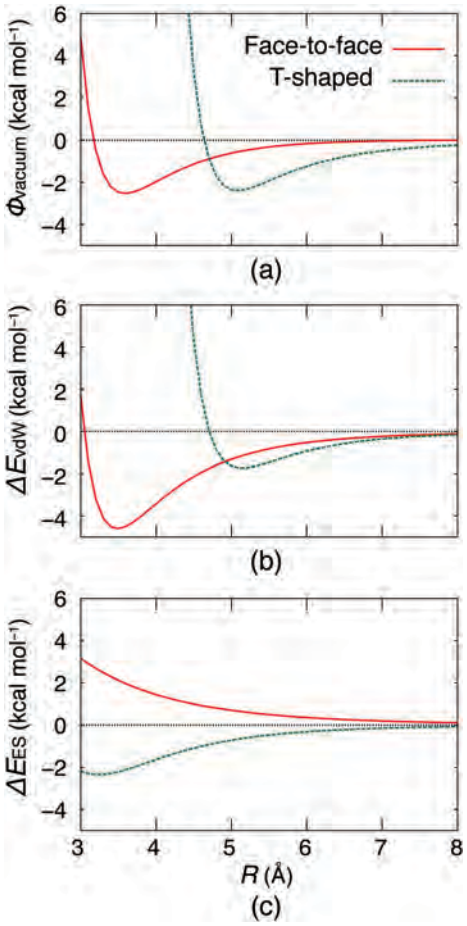

Fig. 2 (a) Interaction between toluene molecules in vacuum Φ_{vacuum}. (b) van der Waals contribution to Φ_{vacuum}, ΔE_{vdW}. (c) Electrostatic contribution to Φ_{vacuum}, ΔE_{ES}. $\Phi_{vacuum} = \Delta E_{vdW} + \Delta E_{ES}$. The red and green curves are for the FF and TS paths illustrated in Fig. 1(a) and (b), respectively.

3.2 Water-mediated aromatic–aromatic interaction: potential of mean force (PMF) in water

Fig. 3 shows the curves of $\Phi_{water}(R)$ (*i.e.*, PMFs) along the FF and TS paths. Unlike $\Phi_{vacuum}(R)$, $\Phi_{water}(R)$ is oscillatory with a main periodicity of d_S (d_S is the molecular diameter of water, 2.8 Å) and possesses local minima and maxima. The global minimum is referred to as the "close-contact minimum (CCM)". The CCM occurs at $R = R_{CCM} \sim 3.3$ Å in the FF dimer and $R = R_{CCM} \sim 4.9$ Å in the TS dimer ($R_{CCM} < R_{VM}$). In the FF dimer, $\Phi_{water}(R_{CCM}) = -3.71$ kcal mol^{-1} that is lower than

$\Phi_{vacuum}(R_{VM}) = -2.53$ kcal mol^{-1}. In the TS dimer, $\Phi_{water}(R_{CCM}) = -0.66$ kcal mol^{-1} that is higher than $\Phi_{vacuum}(R_{VM}) = -2.39$ kcal mol^{-1}. These changes in the FF and TS interactions are caused by the hydration effects. The PMFs along the two paths are qualitatively similar to those calculated by Chipot *et al.*[12] in their MD simulation.

We quantitatively compare the local minima and maxima of $\Phi_{water}(R)$ from this study with those from the MD simulation mentioned above (see Table 2). The MD results for the CCM are as follows: $R_{CCM} \sim 3.5$ Å and $\Phi_{water}(R_{CCM}) = -3.41$ kcal mol^{-1} in the FF dimer and $R_{CCM} \sim 5.0$ Å and $\Phi_{water}(R_{CCM}) = -2.29$ kcal mol^{-1} in the TS dimer. Our results are in quantitatively good accord with the MD results, with the exception that $\Phi_{water}(R_{CCM})$ in the TS dimer in ours is somewhat higher. The second minimum is referred to as the "water-separated minimum (WSM)". The presence of WSM was first pointed out by Geiger *et al.*[8] for the PMF between LJ particles calculated by an MD simulation with explicit water molecules. We note that $R_{WSM} \sim R_{CCM} + d_S$ and $\Phi_{water}(R_{WSM}) > \Phi_{water}(R_{CCM})$. The values of R_{WSM} for the FF and TS paths from this study are ~6.1 and ~7.6 Å, respectively, and those from the MD simulation are ~6.7 and ~8.1 Å, respectively. The values of $\Phi_{water}(R_{WSM})$ for the FF and TS paths from this study are −0.72 and 0.34 kcal mol^{-1}, respectively, and those from the MD simulation are −0.80 and −0.85 kcal mol^{-1}, respectively. Again, only $\Phi_{water}(R_{WSM})$ in the TS path is calculated to be somewhat higher in this study. A repulsive peak (RP) appears at $R = R_{RP}$. Our results are as follows: "$R_{RP} \sim 5.0$ and ~6.1 Å" and "$\Phi_{water}(R_{RP})$ = 4.04 and 1.71 kcal mol^{-1}" in the FF and TS paths, respectively. According to the MD results, "$R_{RP} \sim 5.5$ and ~6.7 Å" and "$\Phi_{water}(R_{RP})$ = 0.46 and 0.66 kcal mol^{-1}" in the FF and TS paths, respectively. $\Phi_{water}(R_{RP})$ in the FF path from this study seems to be significantly higher.

The quantitative disagreement described above is attributable not only to the approximations employed in the 3D-RISM theory[34] but also to the strategy of the MD simulation. The so-called thermodynamic integration is adopted in the MD simulation,[12] and the results are not completely free from inaccuracy arising from insufficiently large numbers of water molecules explicitly incorporated and of system configurations taken for the ensemble average. Moreover, the PMF is calculated for not eqn (9) but $\Phi(R) = G$(toluene–toluene, R) − G(toluene–toluene, $R = R^+$) where R^+ is 8 Å for the FF path and 10.5 Å for the TS path. With $R^+ = 8$ Å, for example, the surface separation between two toluene molecules is ~5 Å which is only ~1.8 times larger than the molecular diameter of water. In the 3D-RISM results, the PMF is considerably longer ranged. It is ~0.9 kcal mol^{-1} at $R \sim 8.5$ Å in the FF path and ~0.5 kcal mol^{-1} at $R \sim 11$ Å in the TS path. Setting the PMF at zero for $R \geq R^+$ in the MD simulation can be accompanied by nontrivial discrepancy from the correct PMF. Nevertheless, it is important to note that the agreement between the 3D-RISM and MD results for the CCM is fairly good especially in the case of FF stacking. Our principal objective of this study is to analyze the hydration effects on the aromatic–aromatic contact, and the arguments presented in Sections 3.3 through 3.6 are not likely to be altered by the uncertainty of the 3D-RISM and MD results.

Table 2 Comparison between 3D-RISM and MD results. The MD values are extracted from Fig. 3(b) in ref. 12 using a software. $\Delta(E_{vdW} + \varepsilon_{vdW})(R_{CCM})$, for example, denotes $\Delta\{E_{vdW}(R_{CCM}) + \varepsilon_{vdW}(R_{CCM})\}$, the value of $\Delta\{E_{vdW}(R) + \varepsilon_{vdW}(R)\}$ at $R = R_{CCM}$. In vacuum, $R_{VM} = 3.6$, $\Delta\Phi_{vacuum}(R_{VM}) = -2.53$, $\Delta E_{vdW}(R_{VM}) = -4.50$, and $\Delta E_{ES}(R_{VM}) = 1.97$ for "face-to-face" and $R_{VM} = 5.1$, $\Delta\Phi_{vacuum}(R_{VM}) = -2.39$, $\Delta E_{vdW}(R_{VM}) = -1.73$, and $\Delta E_{ES}(R_{VM}) = -0.66$ for "T-shaped". Except for R_X (X = VM, CCM, RP, or WSM) given in Å, the values are given in kcal mol^{-1}

	Face-to-face		T-shaped	
Quantities	3D-RISM	MD	3D-RISM	MD
R_{CCM}	3.3	3.5	4.9	5.0
$\Delta\Phi_{water}(R_{CCM})$	−3.71	−3.41	−0.66	−2.29
$\Delta(E_{vdW} + \varepsilon_{vdW})(R_{CCM})$	−1.89		−0.02	
$\Delta(E_{ES} + \varepsilon_{ES})(R_{CCM})$	0.43		0.59	
$-T\Delta S_{VH}(R_{CCM})$	−2.25		−1.23	
R_{RP}	5.0	5.5	6.1	6.7
$\Delta\Phi_{water}(R_{RP})$	4.04	0.46	1.71	0.66
$\Delta(E_{vdW} + \varepsilon_{VH,vdW})(R_{RP})$	1.28		0.61	
$\Delta(E_{ES} + \varepsilon_{VH,ES})(R_{RP})$	1.15		0.91	
$-T\Delta S_{VH}(R_{RP})$	1.61		0.19	
R_{WSM}	6.1	6.7	7.6	8.1
$\Delta\Phi_{water}(R_{WSM})$	−0.72	−0.80	0.34	−0.85
$\Delta(E_{vdW} + \varepsilon_{VH,vdW})(R_{WSM})$	−0.01		0.24	
$\Delta(E_{ES} + \varepsilon_{VH,ES})(R_{WSM})$	−0.34		0.28	
$-T\Delta S_{VH}(R_{WSM})$	−0.36		−0.18	

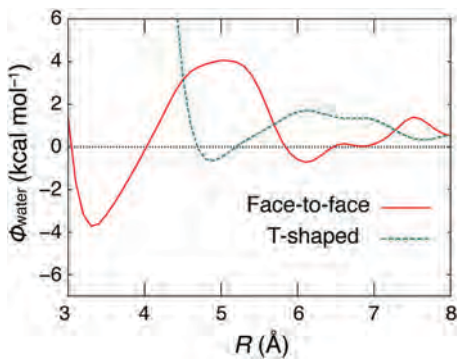

Fig. 3 Potential of mean force between toluene molecules in water Φ_{water}. The red and green curves are for the FF and TS paths illustrated in Fig. 1(a) and (b), respectively.

3.3 Energetic and entropic components of PMF

Fig. 4 shows $\Delta\{E_{vdW}(R) + \varepsilon_{VH,vdW}(R)\}$, $\Delta\{E_{ES}(R) + \varepsilon_{VH,ES}(R)\}$, and $-T\Delta S_{VH}(R)$ along the FF and TS paths. $\Delta\{E_{vdW}(R) + \varepsilon_{VH,vdW}(R)\}$ and $\Delta\{E_{ES}(R) + \varepsilon_{VH,ES}(R)\}$ are substantially different from $\Delta E_{vdW}(R)$ and $\Delta E_{ES}(R)$, respectively, and $-T\Delta S_{VH}(R)$ makes a significantly large contribution to $\Phi_{water}(R)$, which indicates that the hydration effects are quite large. The values of $\Delta\{E_{vdW}(R) + \varepsilon_{VH,vdW}(R)\}$, $\Delta\{E_{ES}(R) + \varepsilon_{VH,ES}(R)\}$, and $-T\Delta S_{VH}(R)$ at $R = R_{CCM}$, R_{RP}, and R_{WSM} are given in Table 2. The term which is the most responsible for the repulsive peak is $-T\Delta S_{VH}$ for the FF path and $\Delta(E_{ES} + \varepsilon_{VH,ES})$ for the TS path.

Fig. 5 compares the components of direct interaction and energetic hydration, $\Delta E_{vdW}(R)$ and $\Delta\varepsilon_{VH,vdW}(R)$ or $\Delta E_{ES}(R)$ and

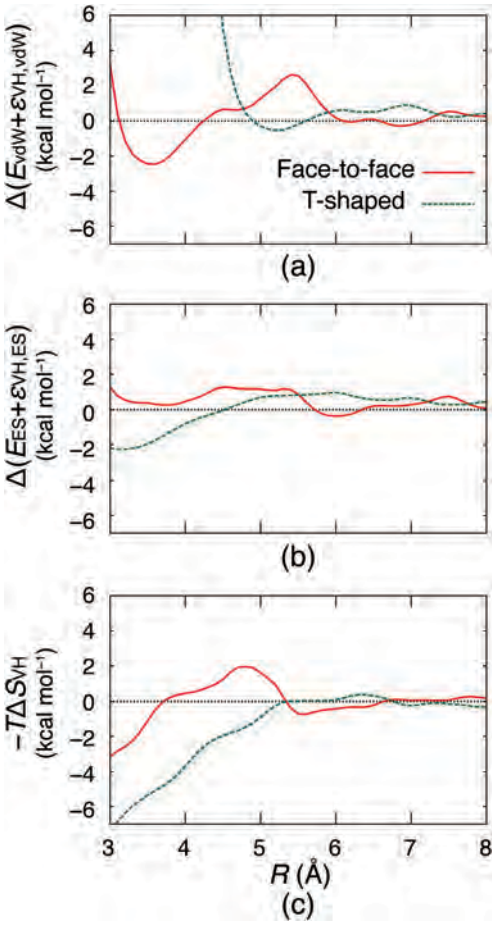

Fig. 4 Decomposition of the potential of mean force between toluene molecules in water Φ_{water} into van der Waals energetic ((a): $\Delta(E_{vdW} + \varepsilon_{VH,vdW})$), electrostatic energetic ((b): $\Delta(E_{ES} + \varepsilon_{VH,ES})$), and entropic ((c): $-T\Delta S_{VH}$) components. $\Phi_{water} = \Delta(E_{vdW} + \varepsilon_{VH,vdW}) + \Delta(E_{ES} + \varepsilon_{VH,ES}) - T\Delta S_{VH}$. The red and green curves are for the FF and TS paths illustrated in Fig. 1(a) and (b), respectively.

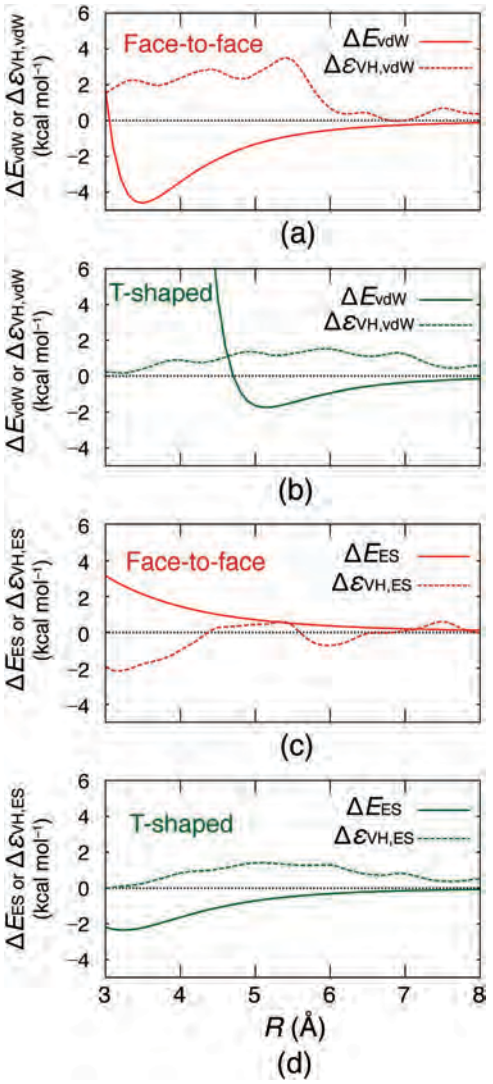

Fig. 5 Comparison between the contributions from direct interaction (ΔE_{vdW} or ΔE_{ES}) and energetic hydration ($\Delta\varepsilon_{VH,vdW}$ or $\Delta\varepsilon_{VH,ES}$). (a) van der Waals contributions (ΔE_{vdW} and $\Delta\varepsilon_{VH,vdW}$) for the FF path. (b) van der Waals contributions (ΔE_{vdW} and $\Delta\varepsilon_{VH,vdW}$) for the TS path. (c) Electrostatic contributions (ΔE_{ES} and $\Delta\varepsilon_{VH,ES}$) for the FF path. (d) Electrostatic contributions (ΔE_{ES} and $\Delta\varepsilon_{VH,ES}$) for the TS path. The red and green curves are for the FF and TS paths illustrated in Fig. 1(a) and (b), respectively.

$\Delta\varepsilon_{VH,ES}(R)$, along the FF and TS paths. $\Delta E_{vdW}(R)$ and $\Delta\varepsilon_{VH,vdW}(R)$ are compensating except in the region where the core repulsion in the direct interaction dominates. This is primarily because a gain of the toluene–toluene vdW interaction is accompanied by a loss of the toluene–water vdW interaction. $\Delta E_{ES}(R)$ and $\Delta\varepsilon_{VH,ES}(R)$ are also compensating. In general, when solute atoms with like partial charges come closer to each other, their direct interaction becomes more repulsive, causing larger, positive ΔE_{ES}. However, the electric field near the solute atoms increases to a significant extent. Consequently, hydrogen atoms (with positive partial charges) or oxygen atoms (with negative partial charges) of water molecules are more strongly attracted to the solute atoms and the stabilization by the solute–water electrostatic attractive interaction becomes stronger, leading to larger, negative $\Delta\varepsilon_{VH,ES}$. When solute atoms with unlike partial charges come closer to each other, their direct interaction becomes more attractive, leading to larger, negative ΔE_{ES}. However, the electric field near the solute atoms decreases to a significant extent. As a result, hydrogen atoms or oxygen atoms of water molecules are less strongly attracted to the solute atoms and the stabilization by the solute–water electrostatic attractive interaction becomes weaker, giving rise to larger, positive $\Delta\varepsilon_{VH,ES}$.

3.4 Global minimum values in energetic and entropic components of PMF

We are concerned with the FF dimers in vacuum and in water. $\Delta E_{vdW}(R_{VM}) = -4.50$ kcal mol^{-1} but $\Delta\{E_{vdW}(R_{CCM}) + \varepsilon_{VH,vdW}(R_{CCM})\} = -1.89$ kcal mol^{-1}: the vdW-energetic contribution inducing the contact is largely weakened in water because the contact is accompanied by the loss of toluene–water vdW interaction. This factor (factor 1) makes the FF dimer less stabilized in water. $\Delta E_{ES}(R_{VM}) = 1.97$ kcal mol^{-1} but $\Delta\{E_{ES}(R_{CCM}) + \varepsilon_{VH,ES}(R_{CCM})\} = 0.43$ kcal mol^{-1}, since $\Delta\varepsilon_{VH,ES}(R)$ makes compensation for $\Delta E_{ES}(R)$ at $R = R_{CCM}$ as explained above.

This factor (factor 2) makes the FF dimer less destabilized in water. Factor 1 is larger than factor 2, but $-T\Delta S_{VH}(R_{CCM}) = -2.25$ kcal mol^{-1} ($\sim -3.8k_BT$) originating from a gain of water entropy comes into play, with the result that the FF dimer is more stabilized in water: $\Phi_{water}(R_{CCM}) < \Phi_{vacuum}(R_{VM})$.

As for the TS dimers in vacuum and in water, $\Delta E_{vdW}(R_{VM}) = -1.73$ kcal mol^{-1} but $\Delta\{E_{vdW}(R_{CCM}) + \varepsilon_{VH,vdW}(R_{CCM})\} = -0.02$ kcal mol^{-1}: an almost complete cancellation occurs. Thus, factor 1 makes the TS dimer significantly less stabilized in water. $\Delta E_{ES}(R_{VM}) = -0.66$ kcal mol^{-1} but $\Delta\{E_{ES}(R_{CCM}) + \varepsilon_{VH,ES}(R_{CCM})\} = 0.59$ kcal mol^{-1}, since $\Delta\varepsilon_{VH,ES}(R)$ works against $\Delta E_{ES}(R)$ at $R = R_{CCM}$ as explained above. This factor, which can also be categorized as factor 2, makes the TS dimer rather destabilized in water. $-T\Delta S_{VH}(R_{CCM}) = -1.23$ kcal mol^{-1} comes into play in water but is not large enough to suppress the destabilization effects by factors 1 and 2. Consequently, the TS dimer is less stabilized in water: $\Phi_{water}(R_{CCM}) > \Phi_{vacuum}(R_{VM})$.

3.5 Physical origin of the behavior of entropic components of PMF

$-T\Delta S_{VH}(R)$ exhibits oscillatory behavior (see Fig. 4). That is, when the water-entropy change is considered as a function of R, regions where the change is negative and positive, respectively, appear alternately. In particular, as R decreases, a region where a water-entropy loss occurs is encountered before the two toluene molecules contact each other. The loss is quite large in the FF path. In this section, we provide a physical interpretation of this behavior.

First, we discuss the hydration entropy of a solute. Upon solute insertion into water, the translational and rotational freedom of water molecules is more restricted: the resultant water-entropy loss comprises the translational and rotational contributions. We showed that the translational contribution is much larger than the rotational one.[38] Hereafter, the discussion is focused on the translational contribution.

The hydration entropy S_{VH} can be expressed as the linear combination of four terms which depend on the excluded volume (EV) V_{ex}, water-accessible surface area (WASA) A_{WAS}, and integrated mean and Gaussian curvatures of water-accessible surface (X_{WAS} and Y_{WAS}), respectively:[39,40]

$$S_{VH} = C_1V_{ex} + C_2A_{WAS} + C_3X_{WAS} + C_4Y_{WAS}. \quad (12)$$

We showed for a solute that at ambient temperature and pressure the EV and WASA terms are dominant, $C_1 < 0$, and $C_2 > 0$.[34,35]

The reason for $C_1 < 0$ can readily be understood. The presence of a solute generates a space which the centers of water molecules cannot enter. The volume of the space is the EV. Due to the solute presence, the total volume available to the translational displacement of water molecules in the system is reduced by the EV, giving rise to a water-entropy loss. If the EV decreases, the loss becomes smaller and the water entropy increases ($C_1 < 0$).

What is the reason for $C_2 > 0$? We note that the presence of a water molecule also generates an EV for the other water molecules, and all of the water molecules in the system are entropically correlated. This entropic correlation in the bulk is referred to as "water crowding".[35,41] When some water molecules come very close to the solute (see Fig. 6) and a layer within which the water density is higher than in the bulk is formed near the solute, the translational displacement of these water molecules (especially the water molecules in contact with the solute) is more restricted (effect 1): They undergo an entropic loss. This loss is approximately in proportion to the number of water molecules in the vicinity of the solute, and this number is also approximately in proportion to the WASA. However, the EVs generated by the solute and by these water molecules overlap. This overlap leads to an increase in the total volume available to the translational displacement of the other water molecules (*i.e.*, water molecules that are not in the vicinity of the solute), which is followed by the reduction of their crowding (effect 2): an entropic gain is conferred on the other water molecules. This gain is approximately in proportion to the net overlapped volume and therefore to the number of water molecules in the vicinity of the solute and the WASA. The density structure of water near the solute in the equilibrium state is determined by the competition of effects 1 and 2. We have shown that in the equilibrium state effect 2 is larger than effect 1 with the result of a gain of water entropy upon the formation of the density structure.[35,41] It follows that smaller WASA is entropically less favorable in this sense. If the WASA decreases, the water entropy becomes lower ($C_2 > 0$).

Here, we give the following two remarks.[35,41] First, the energetic component of water can be argued by looking primarily at the water near the solute, but this is not the case for the entropic component. Second, it is true that the water near the solute is entropically unstable due to the solute-induced water structure and makes a negative contribution to S_{VH}, but this contribution is not large enough to make C_2 negative. The concept of water crowding in the bulk is crucially important. The solute hydrophobicity is attributed primarily to an increase in the water crowding upon solute insertion. For example, the cold and pressure denaturating of a protein can be elucidated only by introducing this concept.[35,41] The view concerning the water-entropy effect given by Graziano and coworkers[42,43] (they used a theoretical method which is different from ours: the scaled particle theory) is in line with ours in respect that the translational entropy of water in the system is emphasized and the iceberg structure[44] of water near a nonpolar solute is shown to be irrelevant. The significance of the iceberg structure for the

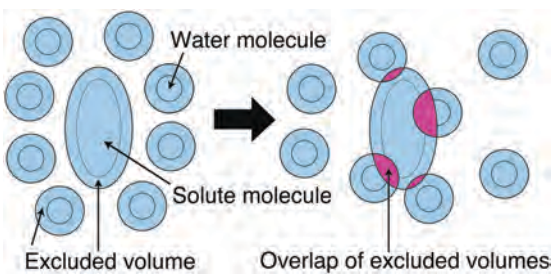

Fig. 6 Overlap of excluded volumes generated by a solute and by some water molecules in the vicinity of the solute.

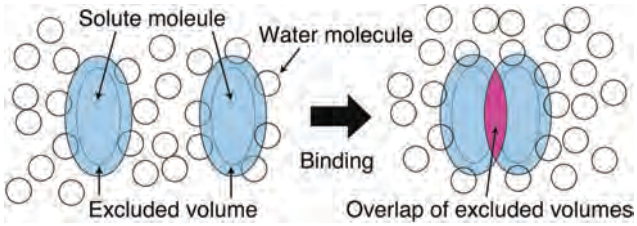

Fig. 7 Overlap of excluded volumes generated by two solutes.

phenomena in aqueous environments was also questioned by Ben-Naim.[9]

The behavior of $-T\Delta S_{VH}(R)$ ($\Delta S_{VH} \sim C_1\Delta V_{ex} + C_2\Delta A_{WAS}$) in Fig. 4 can be understood by the competition between the EV and WASA effects. When two solute molecules come closer to each other, both the EV and the WASA generated by them decrease (a cartoon is illustrated in Fig. 7 for the solute contact). Fig. 8 shows the EV change $\Delta V_{ex}(R)$ and the WASA change $\Delta A_{WAS}(R)$. It is observed that $d^2\Delta V_{ex}(R)/dR^2 < 0$ in both of the FF and TS paths but $d^2\Delta A_{WAS}(R)/dR^2 > 0$ for the FF path and $d^2\Delta A_{WAS}(R)/dR^2 \sim 0$ for the TS path. For larger R, the WASA effect is larger and the net change in water entropy is negative ($-T\Delta S_{VH} > 0$). For smaller R, the EV effect is larger and the net change in water entropy is positive ($-T\Delta S_{VH} < 0$). Negative $\Delta S_{VH}(R)$ implies that the water crowding in the bulk is more serious relative to that in the case of $R \to \infty$.

3.6 Another path toward face-to-face stacking of aromatic rings

There are a number of paths other than the FF path for reaching the FF dimer. The sliding path illustrated in Fig. 1(c) is an example. Fig. 9 shows $\Phi_{water}(R)$, $\Delta\{E_{vdW}(R) + \varepsilon_{VH,vdW}(R)\}$, $\Delta\{E_{ES}(R) + \varepsilon_{VH,ES}(R)\}$, and $-T\Delta S_{VH}(R)$ along the sliding path. This figure should be compared with Fig. 3 and 4 for the FF path. $\Phi_{water}(R)$ for the sliding path features much less oscillatory behavior and a remarkably lower free-energy barrier for reaching the FF dimer. If the 3D-RISM theory always predicts too high a barrier, the PMF along the sliding path actually possesses essentially no barrier. Thus, toluene molecules can readily reach the FF stacking even with their orientations fixed. The presence of a repulsive peak in the PMF along the FF path is not essential.

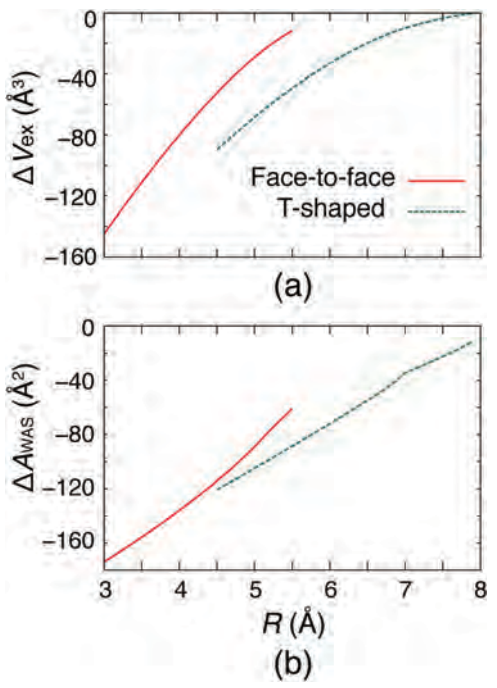

Fig. 8 Changes in the excluded volume ΔV_{ex} (a) and in the water accessible surface area ΔA_{WAS} (b) accompanied by the approach of two toluene molecules. The red and green curves are for the FF and TS paths illustrated in Fig. 1(a) and (b), respectively.

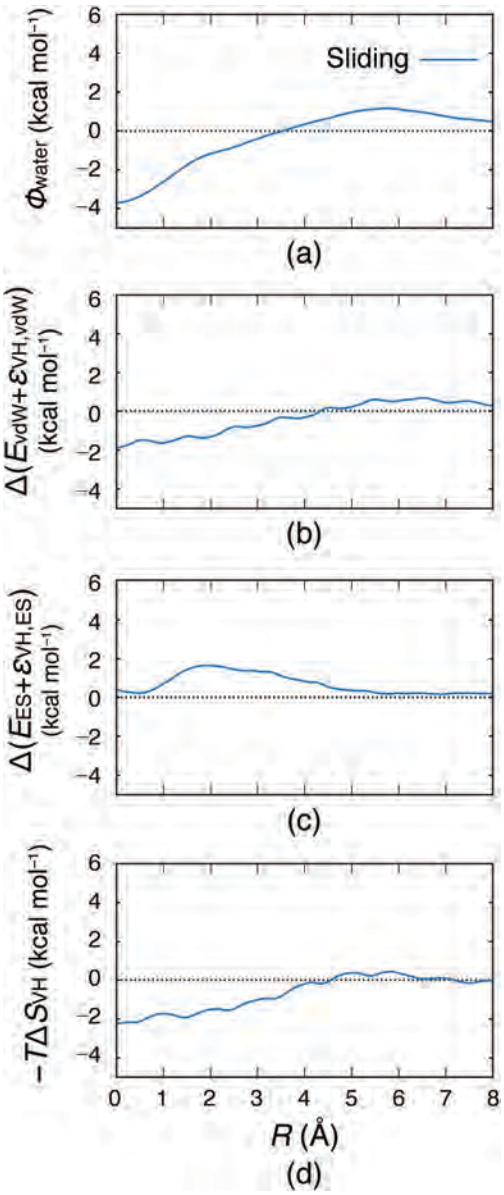

Fig. 9 (a) Potential of mean force between toluene molecules in water Φ_{water}. Decomposition of the potential of mean force between toluene molecules in water Φ_{water} into van der Waals energetic ((b): $\Delta(E_{vdW} + \varepsilon_{VH,vdW})$), electrostatic energetic ((c): $\Delta(E_{ES} + \varepsilon_{VH,ES})$), and entropic ((d): $-T\Delta S_{VH}$) components. $\Phi_{water} = \Delta(E_{vdW} + \varepsilon_{VH,vdW}) + \Delta(E_{ES} + \varepsilon_{VH,ES}) - T\Delta S_{VH}$. The sliding path illustrated in Fig. 1(c) is considered.

3.7 Relevance to intramolecular and intermolecular interactions involving aromatic rings in biological processes

Here, we comment on the aromatic–aromatic stacking interaction in the biomolecular system. As argued above, the FF dimer is more stable than the TS dimer in aqueous solution. It is no wonder that FF stacked arrangements of aromatic rings are commonly observed in the case of nucleotide base interactions.[45] As a typical example, in the binding of Musashi1 (an RNA-binding protein) to mRNA of Numb, adenine is sandwiched by two phenylalanines and guanine is stacked on tryptophan[46] (see Fig. 10(a)).

In proteins, however, the FF stacked arrangements are not frequently observed.[3] The reason for this could be the following: The orientational constraints of dihedral angles of the main chain also limit the aromatic–aromatic contact; and in general, aromatic rings far separated from each other are not allowed to reach the FF stacking. Nevertheless, a protein exhibits high structural stability when the aromatic side chains with relatively larger sizes efficiently participate in close packing of the backbone and side chains of the protein. The close packing is quite advantageous in terms of the water entropy. For instance, CLN025, a peptide with only 10 residues, folds into a specific structure, thus possessing exceptionally high structural stability.[48] We showed that this feature is ascribed to a large gain of water entropy upon folding.[49] As illustrated in Fig. 10(b), its folded structure is characterized by two aromatic side chains stacking on flat portions formed by the backbone. This type of stacking, which is not relevant to the π–π interaction, leads to a large decrease in the total excluded volume followed by a corresponding water-entropy gain (see Fig. 7). An arrangement of aromatic side chains which is similar to the TS contact is also observed. The water-entropy gain resulting from the TS contact is $\sim 2.1k_B$ which is more than half of that brought by the FF contact (the FF contact leads to a water-entropy gain of $\sim 3.8k_B$; see Section 3.4 and Table 2). Therefore, the TS contact often occurs in a protein.[3]

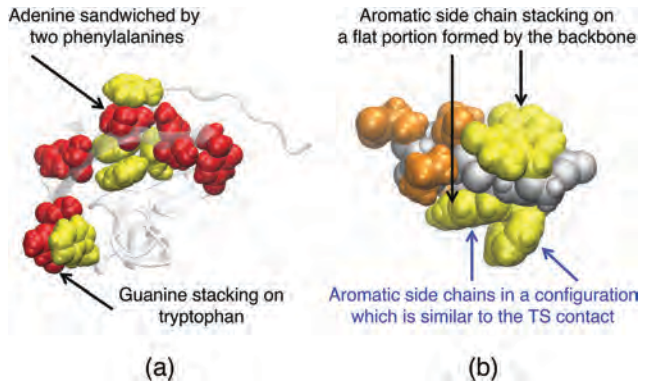

Fig. 10 (a) Binding interface between Musashi1 (colored in yellow), an RNA-binding protein, and mRNA of Numb (colored in red). Space-filling models are employed except for the backbones. Adenine is sandwiched by two phenylalanines and guanine is stacked on tryptophan. (b) Space-filling models of the native structure of CLN025. The backbone, aromatic side chains, and the other side chains are colored in gray, yellow, and orange, respectively. The aromatic side chains indicated by the black arrows stack on flat portions formed by the backbone. The aromatic side chains indicated by the blue arrows are in a configuration which is similar to the TS contact. This figure was drawn by VMD 1.9.1.[47]

It is possible that the biological system utilizes the sliding path for reactions involving the aromatic–aromatic contact. A good example is the intercalative binding of a drug possessing an aromatic moiety[50] or an aromatic side chain of a protein into a DNA double helix.[51] In the complexes formed, the aromatic units are sandwiched by two nucleotide bases. Through the sliding path, the intercalation of the aromatic units into a DNA helix can readily be achieved, realizing the gene regulation by the drug–DNA or protein–DNA bindings.

4 Conclusion

We have analyzed the potential of mean force (PMF) between two toluene molecules in water using the three-dimensional interaction site model (3D-RISM) theory.[14,15,23] The contribution from hydration to the PMF is calculated and decomposed into energetic and entropic components. The energetic component is further decomposed into van der Waals (vdW) and electrostatic terms. These decompositions, which cannot readily be made by a molecular dynamics simulation, have revealed the physical origins of differences between the interactions in vacuum and in water.

In vacuum, the London dispersion (*i.e.*, vdW) attractive interaction strongly induces a contact of toluene molecules. It is much stronger for the face-to-face (FF) stacking than for the T-shaped (TS) contact. However, the electrostatic interaction is repulsive for the former, whereas it is attractive for the latter. As a consequence, the FF stacking and the TS contact share almost the same stability.

In water, on the other hand, the FF stacking is considerably more stable than the TS contact. An important point is that the energetic component of water-induced interaction (*i.e.*, energetic hydration effect) works for opposing the direct interaction (*i.e.*, interaction in vacuum). For the FF stacking, more than half of the London dispersion interaction is cancelled out by the energetic hydration effect. However, the electrostatic repulsive interaction is significantly screened, and a large gain of water entropy occurs. It turns out that the FF stacking is more stabilized by the hydration effects. For the TS contact, a water-entropy gain occurs though it is smaller than in the case of FF stacking. However, the energetic hydration effect cancels out the London dispersion interaction almost completely. At the same time, the electrostatic component of the water-mediated interaction becomes repulsive. The TS contact is less stabilized by the hydration effects. The water-entropy gain, which originates from an increase in the total volume available to the translational displacement of water molecules in the system, plays an essential role in the increased stabilization of the FF stacking in aqueous solution. An experimental study showed that the association of benzene molecules in aqueous solution is entropically driven, manifesting the importance of the water-entropy gain.[52]

There is a general trend that only the direct interaction between aromatic rings is emphasized for biomolecules immersed in aqueous solution. Even in studies using MD simulations with

explicit water, the interaction energy between aromatic rings is often calculated using only their interaction potential: It is regarded that water affects only their relative configuration (*i.e.*, the distance between their centers and the orientations) and the interaction energy is calculated without accounting for the biomolecule–water and water–water interaction energies.[53] The term ''π–π stacking'' is inappropriate for the stacking in aqueous solution, because it sounds as if the hydration effects were minor and the stacking was driven dominantly by the London dispersion interaction.

We have considered the path illustrated in Fig. 1(c) (sliding path) in addition to the FF and TS paths. In comparison to the PMF along the FF path, the one along the sliding path is characterized by significantly less oscillatory behavior and much lower free-energy barrier for toluene molecules to reach the FF stacking. Thus, toluene molecules can readily reach the FF stacking even with their orientations fixed.

In the case of nucleotide base interactions, FF stacked arrangements of aromatic rings are commonly observed.[45,46] They make essential contributions to the biomolecular structural stability. In proteins, the FF stacking cannot necessarily be allowed due to the orientational constraints of dihedral angles of the main chain. However, the participation of aromatic side chains with relatively larger sizes in the close packing of the protein becomes crucially important through the water-entropy effect.[49] For example, the following structural units are often observed in proteins: the stacking of an aromatic side chain on a flat portion formed by the backbone and an arrangement of aromatic side chains which is similar to the TS contact. Further, the intercalative binding of an aromatic unit into a DNA helix[50,51] can readily be achieved through the sliding path, by which the gene transcription is regulated through the drug–DNA or protein–DNA bindings.

Fig. 4 shows that $-T\Delta S_{\mathrm{VH}}(R)$ is oscillatory. In particular, as R decreases, a region where $-T\Delta S_{\mathrm{VH}}(R)$ is significantly repulsive (*i.e.*, a large water-entropy loss occurs) is encountered before the contact of two toluene molecules is reached. This can be interpreted by the competition of the effects of excluded volume (EV) and water-accessible surface area (WASA). The decreases in EV and WASA lead to water-entropy gain and loss, respectively. In the region where the loss is larger than the gain, for example, $\Delta S_{\mathrm{VH}}(R)$ is negative. Negative $\Delta S_{\mathrm{VH}}(R)$ implies that the water crowding in the bulk is more serious relative to that in the case of $R \rightarrow \infty$ (see Section 3.5). For many of the host–guest systems such as the complexation of aromatic solutes in apolar cyclophane cavities, the experimental measurements showed that the process is enthalpically driven because it is accompanied by a loss of entropy and a decrease in enthalpy.[3,4] When the WASA effect is larger than the EV one even after the complexation is accomplished, a decrease in water entropy occurs. This is a possible reason for the entropic loss. How can the enthalpy decrease be explicated? We have investigated the dependence of water structure near an apolar, convex surface on the curvature radius.[54,55] As the radius increases, it becomes more difficult to maintain hydrogen bonds near the surface. It is probable that near an apolar, concave surface, a number of hydrogen bonds are unavoidably broken. Presumably, the water inside a cyclophane cavity is energetically unstable due to the break of hydrogen bonds, and the release of such water to the bulk probably leads to a large decrease in energy. This can be the primary reason for the enthalpy decrease. We intend to analyze the host–guest systems in the near future.

References

1 A. W. Hofmann, *Proc. R. Soc. London*, 1856, **8**, 1–3.
2 E. Hückel, *Z. Phys.*, 1931, **70**, 204–286.
3 E. A. Meyer, R. K. Castellano and F. Diederich, *Angew. Chem., Int. Ed.*, 2003, **42**, 1210–1250.
4 E. Persch, O. Dumele and F. Diederich, *Angew. Chem., Int. Ed.*, 2015, **54**, 3290–3327.
5 R. W. Gurney, *Ionic processes in solution*, McGraw Hill, New York, 1953.
6 H. L. Friedmann and C. V. Krishnan, in *Aqueous solutions of simple electrolytes*, ed. F. Franks, Water: A comprehensive treatise, Plenum, New York, 1973, ch. 1, vol. 3, pp. 1–118.
7 D. G. Duff and C. H. Giles, in *Aqueous Solutions of Amphiphiles and Macromolecules*, ed. F. Franks, Water: A comprehensive treatise, Plenum, New York, 1975, ch. 3, vol. 4, pp. 169–207.
8 A. Geiger, A. Rahaman and F. H. Stillinger, *J. Chem. Phys.*, 1979, **70**, 263–276.
9 A. Ben-Naim, *Hydrophobic interactions*, Plenum, New York, 1980.
10 W. L. Jorgensen and D. L. Severance, *J. Am. Chem. Soc.*, 1990, **112**, 4768–4774.
11 P. Linse, *J. Am. Chem. Soc.*, 1992, **114**, 4366–4373.
12 C. Chipot, R. Jaffe, B. Maigret, D. A. Pearlman and P. A. Kollman, *J. Am. Chem. Soc.*, 1996, **118**, 11217–11224.
13 R. Chelli, F. L. Gervasio, P. Procacci and V. Schettino, *J. Am. Chem. Soc.*, 2002, **124**, 6133–6143.
14 D. Beglov and B. Roux, *J. Chem. Phys.*, 1995, **103**, 360–364.
15 A. Kovalenko and F. Hirata, *J. Chem. Phys.*, 1999, **110**, 10095–10112.
16 T. Imai, Y. Harano, M. Kinoshita, A. Kovalenko and F. Hirata, *J. Chem. Phys.*, 2006, **125**, 024911.
17 T. Imai, Y. Harano, M. Kinoshita, A. Kovalenko and F. Hirata, *J. Chem. Phys.*, 2007, **126**, 225102.
18 T. Imai, K. Oda, A. Kovalenko, F. Hirata and A. Kidera, *J. Am. Chem. Soc.*, 2009, **131**, 12430–12440.
19 T. Imai, N. Miyashita, Y. Sugita, A. Kovalenko, F. Hirata and A. Kidera, *J. Phys. Chem. B*, 2011, **115**, 8288–8295.
20 S.-H. Chong, M. Park and S. Ham, *J. Chem. Theory Comput.*, 2012, **8**, 724–734.
21 S.-H. Chong and S. Ham, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 7636–7641.
22 S.-H. Chong and S. Ham, *Angew. Chem., Int. Ed.*, 2014, **53**, 3961–3964.
23 E. L. Ratkova, D. S. Palmer and M. V. Fedorov, *Chem. Rev.*, 2015, **115**, 6312–6356.
24 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

25 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
26 T. Luchko, S. Gusarov, D. R. Roe, C. Simmerling, D. A. Case, J. Tuszynski and A. Kovalenko, *J. Chem. Theory Comput.*, 2010, **6**, 607–624.
27 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.
28 W. L. Jorgensen and J. D. Madura, *Mol. Phys.*, 1985, **56**, 1381–1392.
29 D. Paschek, *J. Chem. Phys.*, 2004, **120**, 6674–6690.
30 J. S. Perkyns and B. M. Pettitt, *Chem. Phys. Lett.*, 1992, **190**, 626–630.
31 J. S. Perkyns and B. M. Pettitt, *J. Chem. Phys.*, 1992, **97**, 7656–7666.
32 S. J. Singer and D. Chandler, *Mol. Phys.*, 1985, **55**, 621–625.
33 J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, Academic, London, 3rd edn, 2006.
34 T. Hayashi, H. Oshima, Y. Harano and M. Kinoshita, *J. Phys.: Condens. Matter*, 2016, **28**, 344003.
35 H. Oshima and M. Kinoshita, *J. Chem. Phys.*, 2015, **142**, 145103.
36 M. Kinoshita and F. Hirata, *J. Chem. Phys.*, 1997, **106**, 5202–5215.
37 PubChem Open Chemistry Database; https://pubchem.ncbi.nlm.nih.gov/compound/CID. Methane: CID = 297, hexane: CID = 8058, cyclohexane: CID = 8078, benzene: CID = 241, toluene: CID = 1140.
38 M. Kinoshita, *J. Chem. Phys.*, 2008, **128**, 024507.
39 P.-M. König, R. Roth and K. R. Mecke, *Phys. Rev. Lett.*, 2004, **93**, 160601.
40 R. Roth, Y. Harano and M. Kinoshita, *Phys. Rev. Lett.*, 2006, **97**, 078101.
41 M. Kinoshita, *Biophys. Rev.*, 2013, **5**, 283–293.
42 G. Graziano, *Phys. Chem. Chem. Phys.*, 2014, **16**, 21755–21767.
43 A. Pica and G. Graziano, *Phys. Chem. Chem. Phys.*, 2015, **17**, 27750–27757.
44 H. S. Frank and M. W. Evans, *J. Chem. Phys.*, 1945, **13**, 507–532.
45 T. Hermann and D. J. Patel, *Science*, 2000, **287**, 820–825.
46 T. Ohyama, T. Nagata, K. Tsuda, N. Kobayashi, T. Imai, H. Okano, T. Yamazaki and M. Katahira, *Nucleic Acids Res.*, 2012, **40**, 3218–3231.
47 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
48 S. Honda, T. Akiba, Y. S. Kato, Y. Sawada, M. Sekijima, M. Ishimura, A. Ooishi, H. Watanabe, T. Odahara and K. Harata, *J. Am. Chem. Soc.*, 2008, **130**, 15327–15331.
49 S. Yasuda, T. Hayashi and M. Kinoshita, *J. Chem. Phys.*, 2014, **141**, 105103.
50 C.-C. Tsai, S. C. Jain and H. M. Sobell, *Proc. Natl. Acad. Sci. U. S. A.*, 1975, **72**, 628–632.
51 M. H. Werner, A. M. Gronenborn and G. M. Clore, *Science*, 1996, **271**, 778–784.
52 E. E. Tucker, E. H. Lane and S. D. Christian, *J. Solution Chem.*, 1981, **10**, 1–20.
53 M. P. D. Hatfield, R. F. Murphy and S. Lovas, *J. Phys. Chem. B*, 2010, **114**, 3028–3037.
54 M. Kinoshita, *J. Mol. Liq.*, 2005, **119**, 47–54.
55 M. Kinoshita, N. Matubayasi, Y. Harano and M. Nakahara, *J. Chem. Phys.*, 2006, **124**, 024512.