Intermolecular potential for benzoic acid–water based on the test-particle model and statistical mechanical simulations of benzoic acid in aqueous solutions

Kritsana Sagarik [a,*], Bernd M. Rode [b]

[a] *School of Chemistry, Institute of Science, Suranaree University of Technology, 111 University Avenue, Nakhon Ratchasima 30000, Thailand*

[b] *Department of Theoretical Chemistry, Institute for General, Inorganic and Theoretical Chemistry, University of Innsbruck, Innrain 52a, A6020, Innsbruck, Austria*

Received 5 May 2000

Abstract

Structures and interaction energies of benzoic acid–water (BA–H_2O) 1:1, 1:2, 2:1 and 2:2 complexes were investigated using intermolecular potentials derived from the test-particle model (T-model). The absolute and some lowest-lying minimum energy geometries of the 1:1 and 1:2 complexes were examined using ab initio calculations with the Hartree–Fock (HF) and the second order Møller–Plesset (MP2) perturbation theories. The T-model, HF and MP2 calculations revealed that cyclic arrangements of hydrogen bonds (H-bonds) between the COOH group and H_2O represent the absolute minimum energy geometries of the 1:1 and 1:2 complexes. The results on the 2:1 and 2:2 complexes showed that the cyclic H-bonds in the dimers could be opened to allow insertion of water molecules. Based on the T-model potentials, aqueous solutions of BA and $(BA)_2$ were investigated by conducting a series of molecular dynamic (MD) simulations. It was found that, except at the solute–solute H-bonds, the hydration structures of the cyclic H-bond planar (CHP) and side-on type (SOT) dimers are not substantially different from a single BA. The atom–atom pair correlation functions ($g(R)$) derived from MD simulations suggested that, in very dilute aqueous solution, the cyclic H-bonds in the CHP and SOT dimers are not stable and can be disrupted by the solute–solvent H-bond interaction and thermal energy fluctuation. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Ab initio; T-model; Benzoic acid; Dimer; Aqueous solution

1. Introduction

The parallel stacking arrangement between π-electrons of aromatic rings and aromatic side chains of proteins has been observed frequently in protein crystal structures [1]. Much work has been carried out to investigate hydrogen bonding (H-bonding) between aromatic side chains and amino groups in protein crystal structures [2,3]. Although the force that drives aromatic–aromatic interactions is not yet clear, the π–π interaction has been pointed out in many cases to be responsible for the interlayer or stacking of the phenyl rings [4]. The π-electrons in aromatic rings were suggested to

act as H-bond acceptors towards H-bond donors [5].

H-bonding between organic acids and aromatic van der Waals cluster formation in dilute aqueous solutions have also been a subject of great interest in recent years, both from experimental and theoretical points of view. In particular, the influence of hydration on H-bonding and aromatic interaction between solute molecules in aqueous solutions remains controversial. The term ''hydrophobic interaction'' has been invented and frequently used to explain the stability of such solute aggregates in aqueous solutions. The behavior of acetic acid (AA) [6], ethanol [7] and aromatic clusters [8] in dilute aqueous solutions represent good examples.

Benzoic acid (BA) is an aromatic carboxylic acid from which various types of intermolecular interactions could be studied. When two or more BA molecules form clusters, either among themselves or with other molecules, H-bonds, as well as X–H···π [9–12] and π–π types of interaction [13,14], could be expected. Most crystal structures of BA and its derivatives are characterized by stacking of phenyl rings diagonally in one direction and cyclic H-bonded pairs in the direction nearly normal to the planes of the phenyl rings [15–17]. In the crystalline state, reactions between these substances with gaseous molecules were found to be anisotropic and the reactivity could be correlated with their crystal structure [16]. The concerted double-hydrogen transfer in BA dimer in their crystalline environment has continuously been the subject of interest from both experimental and theoretical points of view [18–22]. Having a phenyl ring and a carboxylic functional group in the same molecule, BA has been chosen as a model system to investigate the dissolution mechanism of organic acid into aqueous solutions [23]. This is because the information on the structures of water in the vicinities of the phenyl ring and the carboxylic group could serve as a model to understand the hydration pattern of salicylic acid and its derivatives, from which many biological active molecules could be constructed. According to our literature survey, only limited number of theoretical investigations on BA and its derivatives in aqueous solutions have been presented [24,25]. The results on aqueous solution of 2- as well as 4-

hydroxybenzoic acids reported by Nagy et al. [24] seem to represent the most extensive theoretical study on hydration structures and energetic of BA building block.

In our previous work, H-bonds [26–29], as well as π–π interactions [30–32], in clusters and in liquids have been investigated successfully based on the test-particle model (T-model) potentials. Since the T-model includes the dispersion energy contributions in an approximate way, the molecules considered could be in a wide range. They can be ranging from small H-bonded systems, such as $CHClF_2$ [33], NH_3 [26], HCN [28], $HCONH_2$ [27] and NH_2OH [29], to aromatic compounds, such as benzene [32], pyridine [30], phenol [31] and the base pairs of DNA, adenine–thymine (A–T), guanine and cytosine (G–C) [28]. For small molecules, the T-model potentials gave the results comparable with ab initio calculations at the MP2 level [29]. Whereas for aromatic compounds, the results on benzene [32], pyridine [30] and phenol [31] showed that the T-model can give reliable results with much less computational efforts, compared to supermolecule ab initio calculations that include the effects of the electron correlation. The advantages and disadvantages of the T-model over ab initio supermolecular approach have been discussed in Refs. [30,31], and will not be repeated here in detail.

In the present report, the T-model potential was derived for BA and applied in the investigation of equilibrium structures and interaction energies of the BA–H_2O 1:1 and 1:2 complexes. In order to check the reliability of the T-model potential, the results of the 1:1 and 1:2 complexes were compared with those obtained from ab initio calculations at the Hartree–Fock (HF) and the second-order Møller–Plesset (MP2) perturbation theories. As a primary step in the study of the stability of BA dimers in aqueous solution ($[(BA)_2]_{aq}$), the BA–H_2O 2:1 and 2:2 complexes were investigated using the T-model potential. The T-model potential was further applied in the investigation of the hydration structures of $[BA]_{aq}$ and $[(BA)_2]_{aq}$ using molecular dynamics (MD) simulations. A series of MD simulations were performed to get insight on the structures and stability of BA dimers in aqueous solution. The MD results were discussed

and compared with previous theoretical and experimental results on the same and similar systems.

2. The test-particle model

The derivation of the T-model has been presented in detail elsewhere [34–36]. Here, only some important aspects of the T-model will be briefly summarized. In the present work, the BA geometry was optimized using ab initio calculations at the MP2/6-31G(d,p) level of theory. The optimized geometry of BA was used in the construction of the T-model potential. It was kept constant throughout the calculations. The Cartesian coordinates of atoms of BA are given in Table 1, together with an atom numbering system. Within the framework of the T-model, the interaction energy ($\Delta E^{\text{T-model}}$) of an interacting system A...B is written as a sum of the first-order interaction energy (ΔE^1_{SCF}) and a higher-order term (ΔE^r).

$$\Delta E^{\text{T-model}} = \Delta E^1_{\text{SCF}} + \Delta E^r. \tag{1}$$

ΔE^1_{SCF} accounts for the exchange repulsion and electrostatic energy contributions. It is computed from ab initio wave functions and takes the following form:

$$\Delta E^1_{\text{SCF}} = \sum_{i \in A} \sum_{j \in B} \left[\exp\left[\frac{-R_{ij} + \sigma_i + \sigma_j}{\rho_i + \rho_j} \right] + \frac{q_i q_j}{R_{ij}} \right], \tag{2}$$

where i and j label the sites of molecules. σ_i, ρ_i and q_i are site parameters. R_{ij} are the site–site distances. The exponential term in Eq. (2) represents the exchange repulsion energy contributions, which are related to the size and shape of the molecules A and B. Thus, it is relatively easy to fine-tune the repulsion energy contributions to obtain the best intermolecular distance between the AB system. The exponential parameters in Eq. (2) are derived from the first-order SCF perturbation theory [34]. The point charges (q_i and q_j) are obtained from the requirement that a point-charge model reproduces the electrostatic potentials of molecules of interest. In the previous study [29–31], it was shown that the potential derived

Table 1

Optimized geometry of BA, obtained from ab initio calculations with MP2/6-31G(d,p) and Cartesian coordinates of BA and H_2O (values in atomic unit)
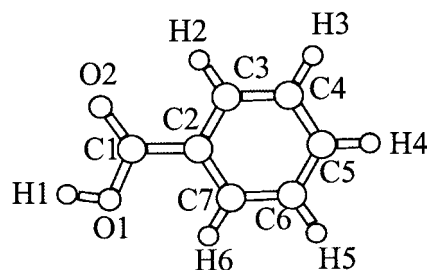
Atom	X	Y	Z
BA			
O1	1.835270	0.000000	0.000000
O2	0.936451	4.180855	0.000000
C1	2.501011	2.484408	0.000000
C2	5.282599	2.858911	0.000000
C3	6.186526	5.344863	0.000000
C4	8.780373	5.790621	0.000000
C5	10.472869	3.764263	0.000000
C6	9.563473	1.286594	0.000000
C7	6.970289	0.820848	0.000000
H1	0.000000	0.000000	0.000000
H2	4.840495	6.883395	0.000000
H3	9.482609	7.711856	0.000000
H4	12.488716	4.113946	0.000000
H5	10.875142	−0.283149	0.000000
H6	6.253941	−1.091959	0.000000
H_2O			
O	0.000000	0.000000	0.000000
H	−1.110670	1.434450	0.000000
H	−1.110670	−1.434450	0.000000
D[a]	−0.500000	0.000000	0.000000

[a] D represents dummy charge.

(PD) charges [37], as well as the charges determined by a fit of the electrostatic potentials at points selected according to the CHelpG scheme [38], are both applicable. In the present study, MP2/6-311G(2d,2p) calculations were employed to generate density matrices, and q_i and q_j of BA were computed based on the CHelpG scheme. About 10,000 electrostatic energies were used in the determination of the point charges. The dipole moment derived from the CHelpG charges was 1.81 D, compared with the experimental value of 1.78 D [39]. All the T-model parameters in Eq. (2) are listed in Table 2.

Table 2

Parameters for the T-model potential of BA and H_2O (atom numbering system in Table 1)

Atom	σ_i	ρ_i	q_i
BA			
O1	1.138607	0.236810	−0.548532
O2	1.129559	0.245188	−0.465275
C1	0.529477	0.396220	0.519581
C2	1.342103	0.214636	0.100950
C3	1.120595	0.307058	−0.180925
C4	1.245962	0.261224	−0.017924
C5	0.929315	0.367166	−0.124305
C6	1.329701	0.234160	−0.038697
C7	1.028291	0.337961	−0.164809
H1	−0.243465	0.288071	0.409564
H2	0.029207	0.270760	0.118136
H3	0.081551	0.269920	0.083979
H4	0.179658	0.228821	0.097361
H5	0.048919	0.283433	0.086936
H6	0.029352	0.265465	0.123960
H_2O[a]			
O	1.284091	0.200370	−0.451660
H	−0.318644	0.331849	0.514110
D[b]			−0.576560

[a] Values taken from Ref. [28].

[b] D = a dummy charge on the C_2 axis of H_2O, 0.26 Å from oxygen and in the opposite direction of the lone pair. The solute–solute and solute–solvent repulsion energies were scaled with a factor of 0.7.

The higher-order energy term, ΔE^r in Eq. (1), represents the dispersion and polarization contributions of the T-model potential. ΔE^r can be determined based on both theoretical and experimental data. ΔE^r takes the following form:

$$\Delta E^r = -\sum_{i \in A} \sum_{j \in B} C_{ij}^6 F_{ij}(R_{ij}) R_{ij}^{-6}, \tag{3}$$

where

$$F_{ij}(R_{ij}) = \exp\left[-\left(1.28\, R_{ij}^0/R_{ij} - 1\right)^2\right],$$
$$R_{ij} < 1.28\, R_{ij}^0, = 1, \text{ elsewhere} \tag{4}$$

and

$$C_{ij}^6 = C_6 \frac{3}{2} \frac{\alpha_i \alpha_j}{(\alpha_i/N_i)^{1/2} + (\alpha_j/N_j)^{1/2}}, \tag{5}$$

R_{ij}^0 in Eq. (4) is the sum of the van der Waals radii of the interacting atoms. Eq. (5) is the Slater–

Kirkwood relation, in which α_i and N_i denote the atomic polarizability and the number of valence electrons of the corresponding atoms, respectively. $F_{ij}(R_{ij})$ is a damping function, introduced to correct the behavior of R_{ij}^{-6} at short R_{ij} distance. Only C_6 in Eq. (5) has to be determined for a particular pair of molecules. The determination of ΔE^r could be regarded as a calibration of the incomplete potential to the properties related to intermolecular interaction energies. Second virial coefficient $(B(T))$, dimerization energy and potential energy of liquid have been shown in our previous studies to be applicable.

A four-site model, similar to the MCY potential [40], was used in the derivation of the T-model parameters of water [28]. They are included in Table 2. For the interaction between BA and water, there is no appropriate experimental data to determine C_6. In our experience, the value of C_6 could be ranging from 0.8 to 1.4, and variation of C_6 within this range does not lead to a substantial change in the potential energy surface. Therefore, in the present study, C_6 was chosen to be 0.8. The value is the same as those of phenol–H_2O [31] and benzene–H_2O complexes [32].

3. BA–H_2O complexes

Although BA cluster in the gas phase is not the primary objective of the present study, remarks should be made on the equilibrium structures and interaction energies of BA dimers derived from the T-model potential [41]. The T-model potential predicted a cyclic H-bond planar (CHP) structure to be the absolute minimum energy geometry, with the interaction energy of −47.00 kJ mol^{-1}. This is in line with experimental and theoretical results in the gas phase [25,42,43]. The molecular mechanical program AMBER, with the charges of BA obtained from a fit to the electrostatic potential around the molecule, predicted a similar dimer to possess the interaction energy of only −33.44 kJ mol^{-1} [25]. The AMBER result also showed that a sandwich structure is as stable as the CHP structure. The authors discussed extensively on the possibility that this type of weak interaction may exist in solutions. Several local minimum energy geome-

tries of the dimer were suggested based on the T-model potential. The so-called "side-on type" (SOT) dimer was predicted to be the second lowest minimum energy geometry, with the interaction energy of -31.06 kJ mol^{-1}. Our results on the BA dimers are in line with the AA dimers [6], in which the CHP and SOT dimers were predicted to be the absolute and local minimum energy geometries, with the interaction energies of -51.46 and -22.03 kJ mol^{-1}, respectively. There was no direct evidence from the T-model results showing the existence of the sandwich structure, as well as the π–π type interaction [44]. This finding might be an indication that the AMBER force fields with the fit charges employed in Ref. [25] are inadequate to describe accurately the interaction in BA clusters, in which the interaction between aromatic rings, though weak, is not negligible, and must be taken into account properly.

The T-model potential described in the previous section was applied in the calculations of the equilibrium structures and interaction energies of BA–H$_2$O 1:1, 1:2, 2:1 and 2:2 complexes. Starting from randomly generated positions and orientations of water and BA, an energy minimization routine was employed in searching for minimum energy geometries of the H-bonded complexes. The absolute and two low-lying energy geometries of each BA–H$_2$O complex are shown in Figs. 1–4, respectively.

In order to examine the T-model results, ab initio calculations were performed on the BA–H$_2$O 1:1 and 1:2 complexes. Since the BA–H$_2$O complexes are considered to be large, care must be exercised in choosing appropriate level of theory, as well as basis sets used in ab initio calculations. We were aware of the problems of the basis set superposition error (BSSE) arising in ab initio standard gradient optimization of H-bonded systems, addressed in detail by Hobza and Havlas [45,46], who pointed out that the potential energy surface (PES) derived from standard gradient optimization and the counterpoise (CP) corrected PES, for which the CP method is applied in each gradient cycle, are not the same. The differences depend on both the level of theory and the basis set used. It was also pointed out that the differences are larger at the electron correlated level.
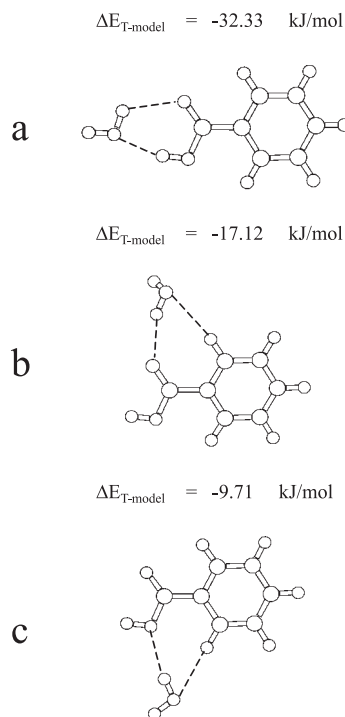
$\Delta E_{\text{T-model}} = -32.33$ kJ/mol

a

$\Delta E_{\text{T-model}} = -17.12$ kJ/mol

b

$\Delta E_{\text{T-model}} = -9.71$ kJ/mol

c

Fig. 1. The absolute and two local minimum energy geometries of BA–H$_2$O 1:1 complexes computed from the T-model potential.

Hobza and Havlas also showed the evidence for the necessity of using the CP-corrected gradient optimization [45]. Since ab initio calculations on the CP corrected PES are very computer intensive especially for large systems and the main objective of the present ab initio gradient optimizations was only to sample check some minimum energy geometries predicted by the T-model potential, single-point CP correction was applied on the BA–H$_2$O optimized geometry. In response to the remarks made by Hobza and Havlas [45] and to ensure that the stability orders of the 1:1 and 1:2 complexes derived from ab initio calculations are correct, a simple statistical analysis was made on the BSSE. The mean was used to assess the quality of the basis set and the standard deviation to measure the anisotropy of the BSSE.

In the present study, all the 1:1 complexes shown in Fig. 1 were employed as starting configurations in the MP2/6-311G(d,p) gradient optimizations.
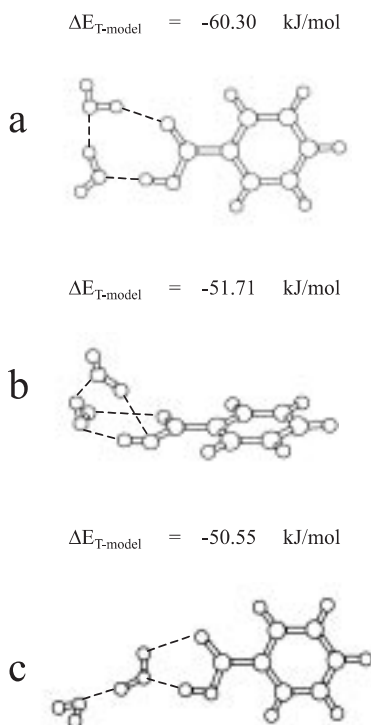
$\Delta E_{\text{T-model}}$ =   -60.30   kJ/mol

a

$\Delta E_{\text{T-model}}$ =   -51.71   kJ/mol

b

$\Delta E_{\text{T-model}}$ =   -50.55   kJ/mol

c

Fig. 2. The absolute and two local minimum energy geometries of BA–H_2O 1:2 complexes computed from the T-model potential.

$\Delta E_{\text{T-model}}$ =   -66.87   kJ/mol

a

$\Delta E_{\text{T-model}}$ =   -66.77   kJ/mol

b

$\Delta E_{\text{T-model}}$ =   -64.20   kJ/mol

c

Fig. 3. The absolute and two local minimum energy geometries of BA–H_2O 2:1 complexes computed from the T-model potential.

Six intermolecular geometrical parameters in the Z-matrix of internal coordinate [47] were identified for each 1:1 complex. They were optimized using the Berny optimization routine [48] built in the GAUSSIAN 94 package [49]. The interaction energies of the optimized geometries were corrected using conventional CP correction of the BSSE. A remark should be made on the 6-311G(d,p) basis sets employed in the present ab initio gradient optimizations. In Ref. [29], we showed that, although the structures of the H-bonded complexes were predicted reasonably well by MP2/6-311G(d,p), the BSSE were still large and sensitive to the structure of the H-bonded complex. The BSSE were significantly reduced when single-point MP2/6-311++G(d,p) calculations were applied on the optimized H-bonded structures. Therefore, in the present case, additional single-point ab initio calculations were made at MP2/6-311++G(d,p) level to minimize the BSSE.
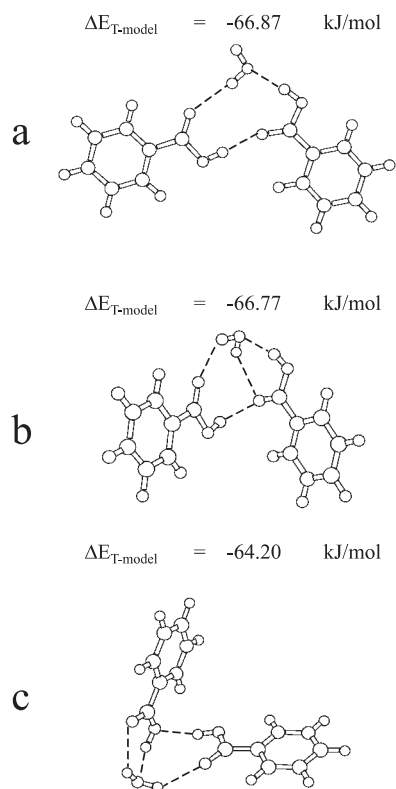
The same procedure was followed for the BA–H_2O 1:2 complexes. In this case, the ab initio gradient optimizations were made at the MP2/6-311G level to maintain reasonable computer time. Since the BSSE was expected to be more pronounced when the number of the interacting molecules increased, two sets of single-point ab initio calculations were made on the MP2/6-311G optimized geometries namely, at the MP2/6-311++G(d,p) and MP2/6-311++G(2d,2p) levels. The energy results of the 1:1 and 1:2 complexes are listed in Tables 3 and 4, respectively.

The T-model potential predicted a CHP structure to be the absolute minimum energy geometry of the 1:1 complex, structure **a** in Fig. 1. Structure **a** consists of two H-bonds, in which water acts as proton donor towards C=O and as proton acceptor towards O–H of BA. The interaction energy
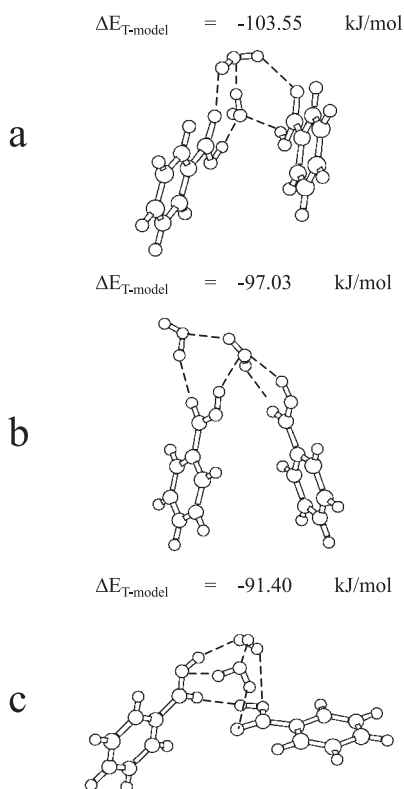
$\Delta E_{T\text{-model}}$ = -103.55   kJ/mol

a

$\Delta E_{T\text{-model}}$ = -97.03   kJ/mol

b

$\Delta E_{T\text{-model}}$ = -91.40   kJ/mol

c

Fig. 4. The absolute and two local minimum energy geometries of BA–H_2O 2:2 complexes computed from the T-model potential.

in this case is -32.33 kJ mol^{-1}. The MP2/6-311G(d,p) gradient optimizations predicted virtually the same results, with slightly shorter H-bond distances, as expected. The CP corrected interaction energy of structure **a** is -33.16 kJ mol^{-1}, compared with -31.81 kJ mol^{-1} from MP2/6-311++G(d,p)//MP2/6-311G(d,p). The C–H groups adjacent to the COOH group of BA were predicted to be involved in the H-bond formation with water, structures **b** and **c** in Fig. 1. Structure **b** is 15.21 kJ mol^{-1} less stable than structure **a**. It is stabilized by C=O···H–O and C–H···O H-bonds. Structure **c** possesses the interaction energy of only -9.71 kJ mol^{-1}. It is obvious from the analysis of the interaction energies in Table 3 that the T-model and ab initio calculations at both HF and MP2 levels gave the same stability order for the BA–H_2O 1:1 complexes. This is partly because the interaction energies of structures **a**–**c** are well separated at all levels of theory. The standard deviations in Table 3 showed that the BSSE for MP2/6-311++G(d,p) are smaller and more isotropic compared to MP2/6-311G(d,p). These findings confirm the stability order and ensure that it will remain the same even if the size of the basis set is further increased.

For the BA–H_2O 1:2 complexes, the COOH group is still the most preferential binding site for water. The absolute minimum energy geometry of the 1:2 complex predicted by the T-model was structure **a** in Fig. 2. Structure **a** consists of three planar H-bonds, with the interaction energy of -60.30 kJ mol^{-1}. Structures **b** and **c** have a unit resembling the absolute minimum energy geometry of the 1:1 complex. The T-model suggested that structures **b** and **c** possess comparable interaction energy, -51.71 and -50.55 kJ mol^{-1}, respectively. Structure **b** is stabilized by four cyclic H-bonds, whereas structure **c** by three H-bonds. Ab initio gradient optimizations at the MP2/6-311G level predicted the same absolute minimum energy geometry, with slightly shorter H-bond distances. The analysis of the ab initio interaction energies in Table 4 revealed that, for both HF/6-311G and MP2/6-311G calculations, structure **c** is more stable than structure **b**. This is partly due to the BSSE. The basis set extension and CP correction considerably improved the interaction energies of structures **b** and **c**. At the MP2/6-311++G(d,p)//MP2/6-311G level, the stability of structures **b** and **c** are comparable. Whereas, at the MP2/6-311++G(2d,2p)//MP2/6-311G level, structure **b** is slightly more stable than structure **c**. This is in good agreement with the T-model results. At the HF level, however, the increase of the size of the basis sets did not reverse the stability order of structures **b** and **c**. In order to further investigate this finding, a separate set of ab initio gradient optimizations at the HF/6-311++G(2d,2p) level was performed, starting from structure **a**–**c**. The global minimum energy geometry was the same as structures **a**. However, starting from structure **b**, the Berny optimization routine yielded structure **c**. This implies that structure **b** is stabilized to a large extent by the dispersion energy and structure **b** need not be a local minimum on the HF energy surface.

Table 3
Interaction energies and BSSE of BA–H$_2$O 1:1 complexes derived from ab initio calculations (energy in kJ mol^{-1})

Structure	a	b	c	Mean	SD
SCF-A	−39.03	−25.70	−10.31		
SCFCP-A	−31.43	−15.44	−5.38		
BSSE$_{SCF-A}$	7.60	10.26	4.93	7.60	2.67
MP2-A	−47.75	−34.06	−18.48		
MP2CP-A	−33.16	−15.19	−8.99		
BSSE$_{MP2-A}$	14.59	18.87	9.49	14.32	4.70
SCF-B	−32.90	−17.73	−7.12		
SCFCP-B	−30.00	−15.16	−5.32		
BSSE$_{SCF-B}$	2.90	2.57	1.80	2.42	0.56
MP2-B	−39.74	−23.33	−15.21		
MP2CP-B	−31.81	−15.86	−9.72		
BSSE$_{MP2-B}$	7.93	7.47	5.49	6.96	1.30

SCF-X: SCF calculations using X basis set, SCFCP-X: SCF-X with BSSE correlation, BSSE$_{SCF-X}$: SCFCP-X–SCF-X, MP2-X: MP2 calculation using X basis set, MP2CP-X: MP2-X with BSSE correction and BSSE$_{MP2-X}$: MP2CP-X–MP2-X (X = A,B). A: ab initio calculations with MP2/6-311G(d,p) and B: ab initio calculations with MP2/6-311++G(d,p)//MP2/6-311G(d,p).

Table 4
Interaction energies and BSSE of BA–H$_2$O 1:2 complexes derived from ab initio calculations (energy in kJ mol^{-1})

Structure	a	b	c	Mean	SD
SCF-A	−119.00	−87.63	−95.11		
SCFCP-A	−96.71	−66.49	−76.92		
BSSE$_{SCF-A}$	22.29	21.14	18.19	20.54	2.11
MP2-A	−128.61	−103.10	−104.92		
MP2CP-A	−87.40	−64.13	−71.24		
BSSE$_{MP2-A}$	41.21	38.97	33.68	37.95	3.87
SCF-B	−69.33	−46.52	−51.92		
SCFCP-B	−60.89	−39.15	−44.41		
BSSE$_{SCF-B}$	8.44	7.37	7.51	7.77	0.58
MP2-B	−90.14	−69.39	−69.91		
MP2CP-B	−66.04	−48.33	−49.56		
BSSE$_{MP2-B}$	24.10	21.06	20.35	23.17	1.99
SCF-C	−60.82	−40.36	−43.67		
SCFCP-C	−55.89	−35.66	−39.85		
BSSE$_{SCF-C}$	4.93	4.70	3.82	4.48	0.59
MP2-C	−84.74	−65.14	−63.45		
MP2CP-C	−69.68	−52.04	−51.62		
BSSE$_{MP2-C}$	15.06	13.10	11.83	13.33	1.63

SCF-X: SCF calculations using X basis set, SCFCP-X: SCF-X with BSSE correlation, BSSE$_{SCF-X}$: SCFCP-X–SCF-X, MP2-X: MP2 calculation using X basis set, MP2CP-X: MP2-X with BSSE correction, and BSSE$_{MP2-X}$: MP2CP-X–MP2-X (X = A,B,C). A: ab initio calculations with MP2/6-311G, B: ab initio calculations with MP2/6-311++G(d,p)//MP2/6-311G and C: ab initio calculations with MP2/6-311++G(2d,2p)//MP2/6-311G.

For the BA–H$_2$O 2:1 complexes, the planar structure, in which water acts simultaneously as proton donor and acceptor bridging the C=O and O–H groups of both BA molecules, was found to be the absolute minimum energy geometry, struc-

ture **a** in Fig. 3. The interaction energy of structure **a** is −66.87 kJ mol^{-1}. Several herringbone structures, similar to those found in the phenol [31] and pyridine dimers [30], were local minimum energy geometries. Structures **b** and **c** are slightly less

stable than structure **a**. They possess a unit resembling the absolute minimum energy geometry of the BA–H_2O 1:1 complex. Structure **a** and **b** can be constructed by breaking one H-bond of the CHP dimer and inserting a water molecule between the C=O and O–H groups, whereas structure **c** can be obtained by inserting a water molecule in the SOT dimer. A herringbone structure appeared to be the absolute minimum energy geometry of the BA–H_2O 2:2 complex, structure **a** in Fig. 4. Structure **a** consists of a unit resembling the absolute minimum energy geometry of the water dimer, with the interaction energy of -103.55 kJ mol^{-1}. Including two water molecules in the BA dimer seems to increase the flexibility of the H-bonded cluster, and hence the possibility to form the π–π interaction. A similar 2:2 complex was suggested by the AMBER force fields [25] to be a stable H-bonded cluster. Based on the interaction energies of 1:1, 1:2, 1:3, as well as 2:1 and 2:2 complexes, the authors concluded that the CHP dimer is not well hydrated.

It should be noted at this point that the structural and energetic results on the BA dimer and BA–H_2O complexes suggest only the possibility of breaking the cyclic H-bonds by a limited number of water molecules. Attempt to explain the situation in aqueous solutions using only this structural and energetic information seems not appropriate since the H-bond interactions in water are rather strong. Reliable information on the stability of BA dimer in aqueous solution should be obtained from statistical mechanical simulations, in which solute–solute, solute–solvent and solvent–solvent interactions, as well as the effects of thermal energy fluctuation and dynamics of the system are all taken into account.

4. Molecular dynamic simulations of [BA]$_{aq}$ and [(BA)$_2$]$_{aq}$

BA is a weak organic acid with a relatively low solubility in water. At 298 K, the acid dissociation constant, K_a, and the solubility of BA in water were reported to be 6.32×10^{-5} mol dm^{-3} [50] and 0.0275 mol dm^{-3} [51], respectively. Hence, we were

interested only in unionized BA in aqueous solutions.

The T-model potential tested in the previous sections was applied in MD simulations of [BA]$_{aq}$ and [(BA)$_2$]$_{aq}$. In order to investigate the hydration structures of [BA]$_{aq}$ and [(BA)$_2$]$_{aq}$, a series of NVE–MD simulations were carried out at 298 K. In each MD run, a single BA molecule or a BA dimer was put in a cubic box filled with 300 water molecules. The simulation box was subject to periodic boundary conditions. The density of the aqueous solutions was maintained at 1.0 g ml^{-1}. The cut-off radius was half of the box length. The long range Coulomb interactions were computed using the Ewald summation. The time step applied in solving the equations of motions was 0.0005 ps. Ten thousand MD steps were devoted to the equilibration and 20,000 steps to property calculations. The hydration structures of aqueous solutions were derived based on selected atom–atom pair correlation functions $(g(R))$ and the corresponding running coordination numbers, as well as oxygen (PDO) and hydrogen probability distribution (PDH) maps [52]. The calculations of the PDO and PDH maps were aiming at visualization of the average three-dimensional arrangement of water molecules in the vicinity of solute. In the calculation of PDO and PDH maps, a solute molecule was identified as a reference and was not allowed to move in the course of MD simulations. The molecular plane of the fixed solute was assumed to coincide with the XY plane of the simulation box (with $Z = 0$ Å). The volumes above and below the molecular plane of solute were divided into layers with the thickness of 1 Å, e.g. from $Z = -0.5$ to 0.5, 0 to 1, 1 to 2, 2 to 3, 3 to 4 and 4 to 5 Å. In each layer, the PDO and PDH maps were computed at the 61×61 grid intersections, by following the trajectories of oxygen and hydrogen atoms of water in the course of MD simulations. The PDO and PDH maps were represented by contour lines. For simplicity, the minimum and maximum values of the contour lines as well as the contour interval were chosen to be the same for all PDO and PDH maps.

Since the main aims of the present MD simulations were to investigate the structures and stability of BA dimers, as well as hydration structures

of $[BA]_{aq}$ and $[(BA)_2]_{aq}$, attempt was not made to compute dynamic properties.

4.1. Hydration structure of $[BA]_{aq}$

The first set of MD simulations (MD-1) involved a single BA solvated by 300 water molecules. The PDO and PDH maps derived from MD-1 are shown in Fig. 5. It is clearly seen from the PDO maps that both polar and non-polar functional groups of BA are hydrated. The water molecule labeled with A in Fig. 5 is obviously the most localized one at the O–H group, especially

within the layer closest to the molecular plane of the solute (Z between −0.5 and 0.5 Å). The O–H group is shared by a water molecule at F. The water molecule at F acts as proton donor towards the O–H group. For the layer with $Z = 1$–2 Å, the PDO map shows more than one water molecules in the vicinity of the C=O group. They are labeled with B in Fig. 5. The water molecules at E, F, G, A, B, C and D are parts of the H-bond networks around BA. On the PDO map, for the layer with $Z = 3$–4 Å, water molecules are seen to hydrate above the aromatic ring. This is in line with the results of $[phenol]_{aq}$ [31], $[benzene]_{aq}$ [32] and
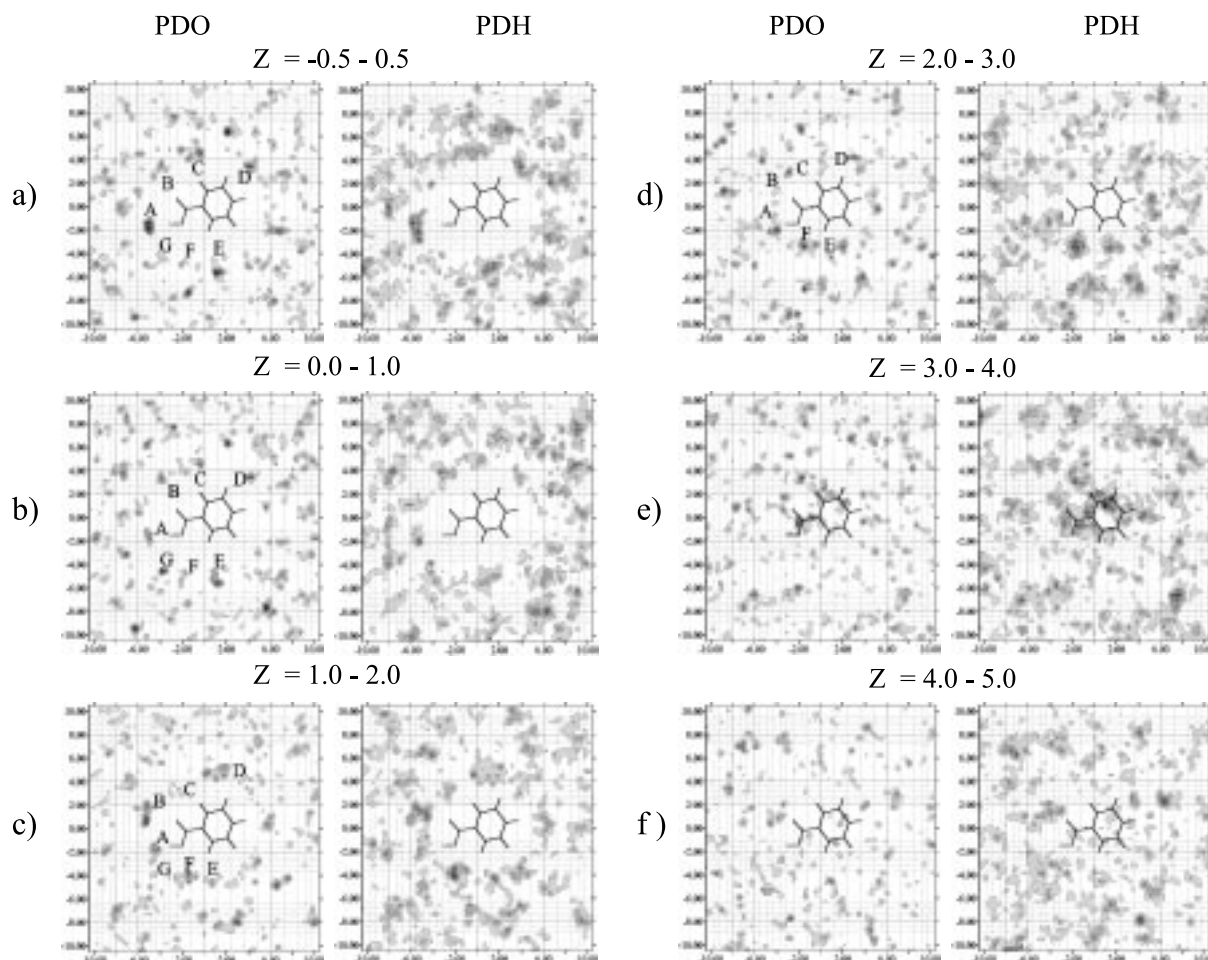
Fig. 5. PDO and PDH maps of $[BA]_{aq}$ at 298 K derived from MD-1 (X-, Y- and Z-axis in Å).

[hydroxybenzoic acid]$_{aq}$ [24], and suggests further that the hydration structure above the aromatic ring does not change substantially by substituents such as the COOH and OH groups.

Some $g(R)$ derived from MD-1 are shown in Fig. 6. The analysis of $g(R_{Ow-O1})$, $g(R_{Ow-O2})$ and $g(R_{Ow-H1})$ quantifies the above results on the hydration structure of [BA]$_{aq}$. For $g(R_{Ow-O1})$, the main peak is quite structured and located at $R = 2.96$ Å. The integration of $g(R_{Ow-O1})$ to the position of the main peak resulted in more than one (1.47) water molecule in closest contact with the O–H group of BA. These correspond to the water molecule at A and partly at F on the PDO map. The main peak of $g(R_{Ow-O2})$ is broader, with the maximum at $R = 3.26$ Å. The integration of $g(R_{Ow-O2})$ to the position of the maximum yielded

about three (2.56) water molecules closest to the C=O group. They can be attributed to the water molecules at B and C on the PDO maps. Therefore, on average, three to four water molecules are expected to be in close contact with the COOH group of BA. The integration of $g(R_{Ow-O1})$ and $g(R_{Ow-O2})$ to the first minimum suggested about six (5.80) and seven (7.02) water molecules in the first hydration shells of the O–H and C=O groups, respectively. Since the main peak of $g(R_{Ow-H1})$ is associated with the H-bond between the oxygen atom of water and the H–O group of BA, it could be used to measure the H-bond donating ability of BA in aqueous solution. In this case, $g(R_{Ow-H1})$ is quite structured and seen at $R = 2.01$ Å. The integration of $g(R_{Ow-H1})$ to the first minimum resulted in about one (1.32) water molecule localized
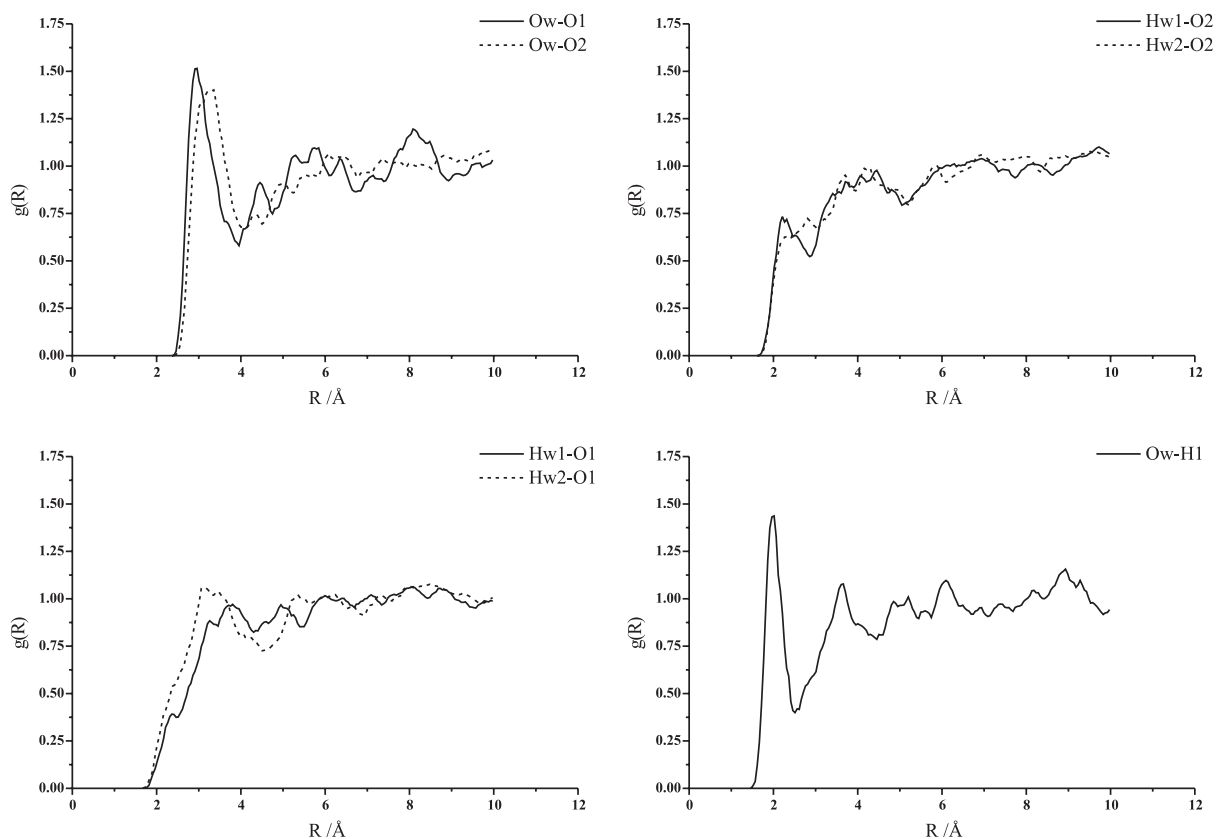

Fig. 6. Selected $g(R)$ of [BA]$_{aq}$ at 298 K derived from MD-1. Atom numbering system is in Table 1. Ow and Hw are the oxygen and hydrogen atoms of water, respectively.

in the vicinity of the H1 atom, which is in good agreement with the analysis of $g(R_{\text{Ow–O1}})$.

Aqueous solutions of 2- and 4-hydroxybenzoic acids were investigated using Monte Carlo (MC) simulations [24], in which the 12-6-1 OPLS and TIP4P intermolecular potential functions were employed. From the analysis of $g(R)$ and snapshots of the aqueous solutions with a solute molecule, the authors concluded that the hydration structures of the carboxylic and phenolic OH groups of 4-hydroxybenzoic acid are independent. The conclusion is supported by experimental studies on H-bonding in dilute solutions [53,54], in which the solvation numbers of some aromatic solutes were shown to be additive, e.g. the solvation number of hydroquinone in methanol is twice of that of phenol. Thus based on the solvation numbers of 4-chlorobenzoic acid, 4-methylbenzoic acid and 2-nitrotoluene, the solvation number of a single BA in methanol is estimated to be three. This is in reasonable agreement with three to four in the case of dilute aqueous solution.

The above findings allow a direct comparison of the hydration structure at the carboxylic group of 4-hydroxybenzoic acids [24] with that of BA in the present study. It was reported in Ref. [24] that water molecules are more stable when localized at the O–H group, compared to the C=O group. Based on a H-bond energy analysis, one strong linear H-bond and a partial H-bond, in which O1 acts as proton acceptor, were predicted at the O–H group and only one water molecule was expected to hydrate directly at the C=O group. The former corresponds to the water molecules at A and F and the latter to a water molecule at B in present study.

Hydration structure of $[\text{AA}]_{\text{aq}}$ was examined using MC simulations with hybrid AM1/TIP3P and OPLS functions in Ref. [55]. The hydration structure at the COOH group, inferred from MC simulations, is in general in good accord with the present results. The AM1/TIP3P model, however, predicted a rather week H-bond donating ability of AA, leading to the conclusion that the C=O group is more favorable for the H-bond interaction with water.

The average structure of H-bond network in aqueous solution was not discussed in Refs. [24,55]. This is probably due to the fact that the

information obtained from $g(R)$ and the snapshots are not enough to provide insight on the average orientation of water molecules in the vicinity of the solute. The PDO and PDH maps are, in our opinion, more appropriate for this purpose.

4.2. Hydration structures of $[(BA)_2]_{aq}$

In order to model hydration structures of $[(BA)_2]_{aq}$ and get insight on the stability of the CHP and SOT dimers in aqueous solutions, three sets of MD simulations (MD-2, MD-3 and MD-4) were carried out. The CHP dimer was chosen since it is the smallest H-bonded pair on the solid surface and in the crystal structures [15–17]. The investigation on $[\text{SOT}]_{\text{aq}}$ was aiming mainly at comparison with $[\text{AA}]_{\text{aq}}$ [6].

In MD-2 and MD-2a, the CHP and SOT dimers were treated as supermolecules, in which both BA in the dimers were not allowed to move in the course of MD simulations. Since the interfacial resistance to the dissolution of the CHP dimer is negligible [56], the results obtained from MD-2 could represent the earliest stage of the dissolution process at the solid surface.

The PDO and PDH maps derived from MD-2 are given in Fig. 7. Some selected $g(R)$ are illustrated in Fig. 8. Since the C=O and O–H groups of the CHP dimer are nearly completely H-bonded, the hydration structure of $[\text{CHP}]_{\text{aq}}$ was expected to be less pronounced, compared to both $[\text{BA}]_{\text{aq}}$ and $[\text{SOT}]_{\text{aq}}$. It is seen from the PDO maps in Fig. 7 that water molecules form H-bond networks above the cyclic H-bonds of the CHP dimer, as well as in the molecular planes. The H-bond networks span from C to G to E to F, clearly seen from the PDO map within the layer of $Z = 2$–3 Å. In Fig. 8, the structures of both $g(R_{\text{Ow–O1}})$ and $g(R_{\text{Ow–O2}})$ are quite different from $[\text{BA}]_{\text{aq}}$. In the case of $[\text{CHP}]_{\text{aq}}$, the main peak of $g(R_{\text{Ow–O1}})$ is considerably less structured and shifted to $R = 3.26$ Å. The integration of $g(R_{\text{Ow–O1}})$ to the position of the main peak gave about two (1.70) water molecules in closest contact with O1, compared to about one (1.47) in the case of $[\text{BA}]_{\text{aq}}$. In going from $[\text{BA}]_{\text{aq}}$ to $[\text{CHP}]_{\text{aq}}$, the C=O group becomes obviously less hydrated. The position of the main peak of $g(R_{\text{Ow–O2}})$ remains the same but the peak

PDO   PDH   PDO   PDH

Z = -0.5 - 0.5   Z = 2.0 - 3.0

a)

d)

Z = 0.0 - 1.0   Z = 3.0 - 4.0

b)

e)
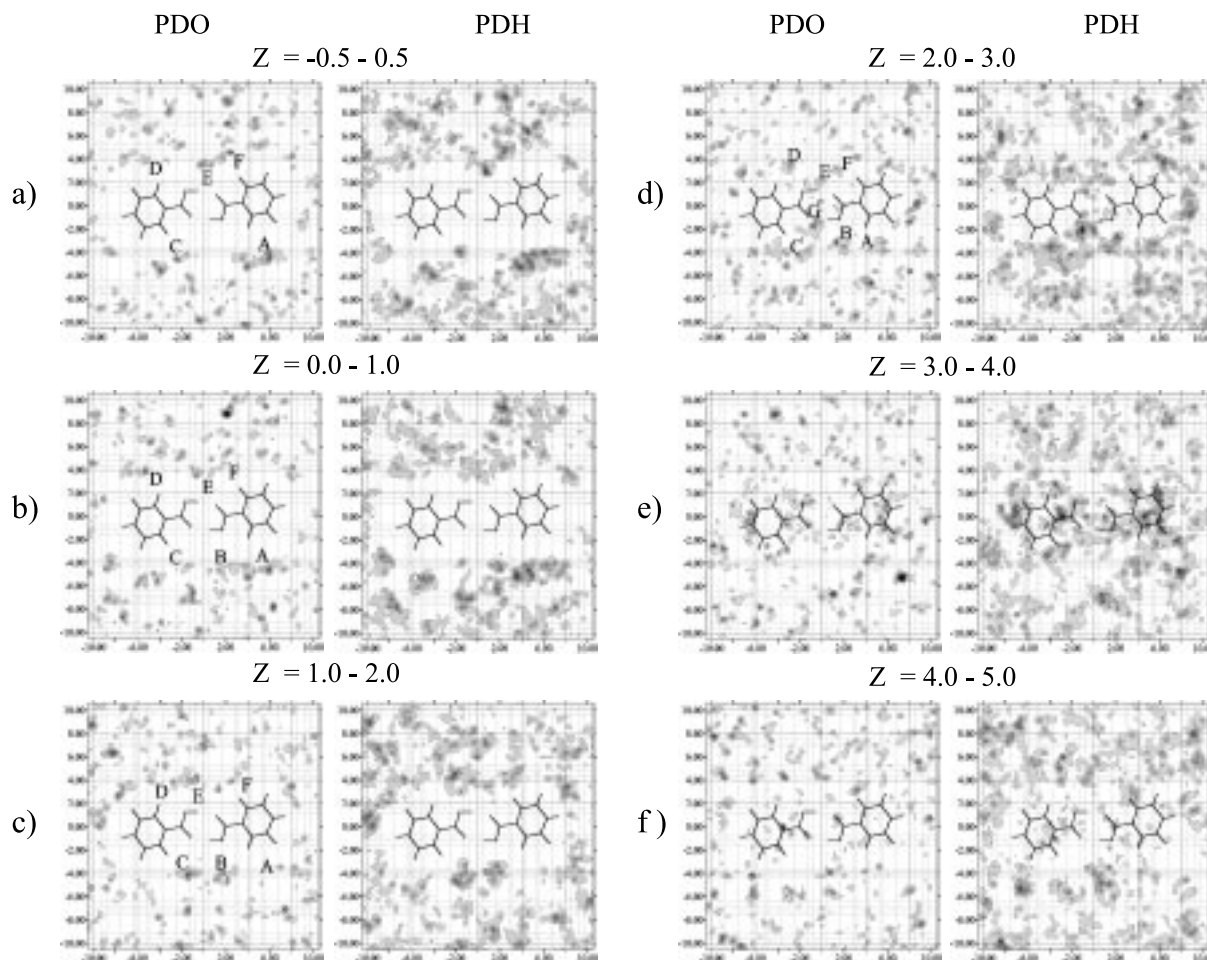
Z = 1.0 - 2.0   Z = 4.0 - 5.0

c)

f)

Fig. 7. PDO and PDH maps of $[CHP]_{aq}$ at 298 K derived from MD-2 (X-, Y- and Z-axis in Å).

height decreases substantially. The integration of $g(R_{Ow–O2})$ to the position of the main peak indicated that there is only about one (1.37) water molecule in closest contact with O2, compared to about three in $[BA]_{aq}$. For $g(R_{Ow–H1})$, the position of the main peak shifted to longer distance, compared to $[BA]_{aq}$. The new peak position at $R = 3.36$ Å is too long to match the H-bond distance and implies that the H-bond donating ability of BA is considerably reduced upon the CHP dimer formation.

The PDO and PDH maps derived from MD-2a are given in Fig. 9. Some selected $g(R)$ are illustrated in Fig. 10. In general, there is no significant

difference in the hydration structures of $[CHP]_{aq}$ and $[SOT]_{aq}$, except in the area near the COOH groups. For $[SOT]_{aq}$, the PDO maps showed that both C=O and O–H groups can be partly hydrated. Water molecules obviously prefer to hydrate the SOT dimer in the areas labeled with A, E and F in Fig. 9, especially in the layer with $Z = 2–3$ Å. The water molecule at A acts as both proton acceptor and donor towards the O–H and C=O groups on different BA molecule. Water molecule at E also H-bonds simultaneously with the C–H and O–H groups of the SOT dimer. The analysis of $g(R_{Ow–O1})$ and $g(R_{Ow–O2})$ in Fig. 10 showed that there are more water molecules (1.99) in close
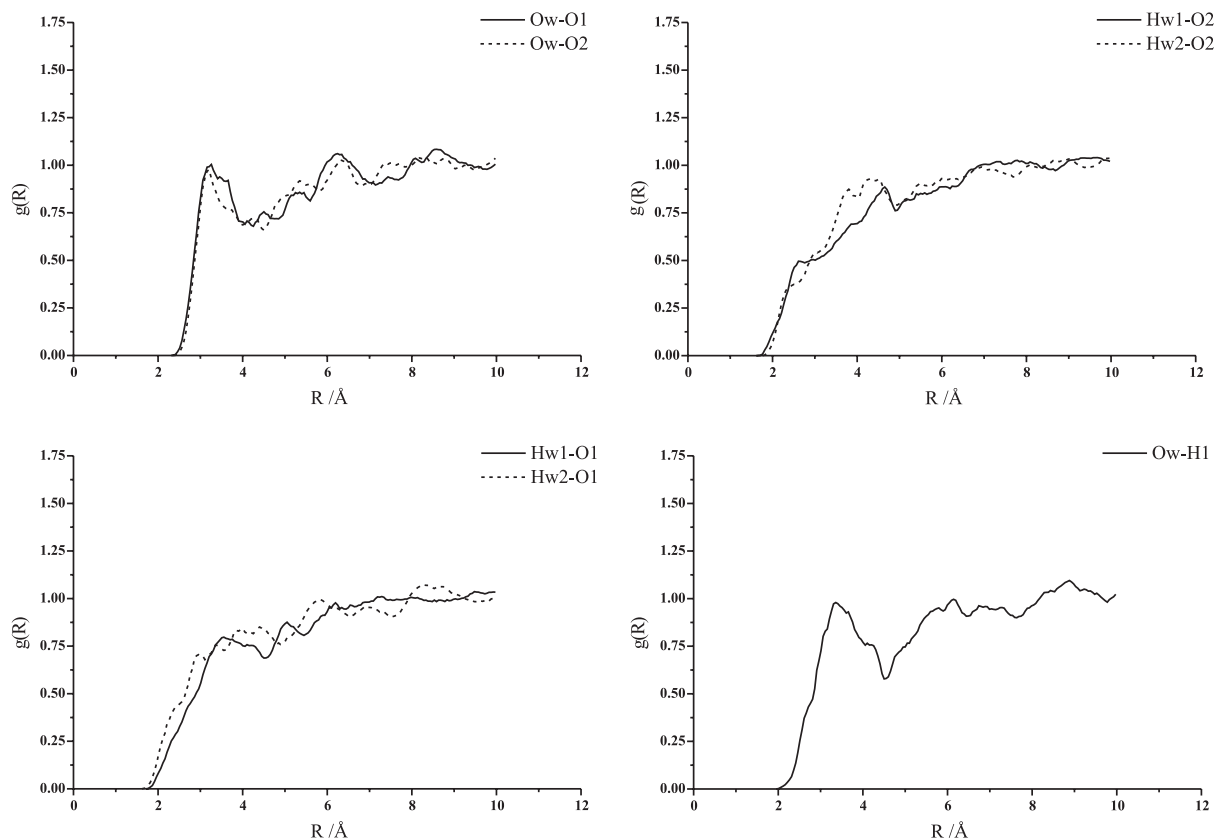
Fig. 8. Selected $g(R)$ of $[CHP]_{aq}$ at 298 K derived from MD-2. Atom numbering system is in Table 1. Ow and Hw are the oxygen and hydrogen atoms of water, respectively.

contact with O1 and less (1.79) with O2, compared to $[BA]_{aq}$. The main peak of $g(R_{Ow-O2})$ shifted slightly towards shorter distance, with the peak height larger than in MD-2. This indicates an increase of the H-bond accepting ability of the SOT dimer, compared to the CHP dimer. The main peak position at $R = 3.11$ Å could be attributed to the water molecule at A in Fig. 9. $g(R_{Ow-H1})$ showed that the H-bond donating ability of BA is considerably reduced upon the SOT dimer formation.

For both CHP and SOT dimers, water molecules were also found to hydrate above the aromatic rings in the layer with $Z = 3–4$ Å. The above discussion on the hydration structures leads to the conclusion that both C=O and O–H groups in the supermolecule CHP and SOT dimers are not very

active in the H-bond formation with water. However, the SOT dimer seems to be a better proton acceptor towards water molecules, compared to the CHP dimer.

In enzyme catalysis, the most important steps have long been known to involve H-bond formation and disruption. In such catalytic reactions, active sites of amino acid residues of enzymes could be considered as being in the stationary phase and substrates in the mobile phase. Also in separation chemistry, apart from the steric effects, the diffusivity of permeating compounds, as well as van der Waals and H-bond interactions between permeating compounds and polymer membranes, are of primary importance, especially in the determination of the retention time and membrane selectivity [57,58]. It was pointed out in Ref. [59]
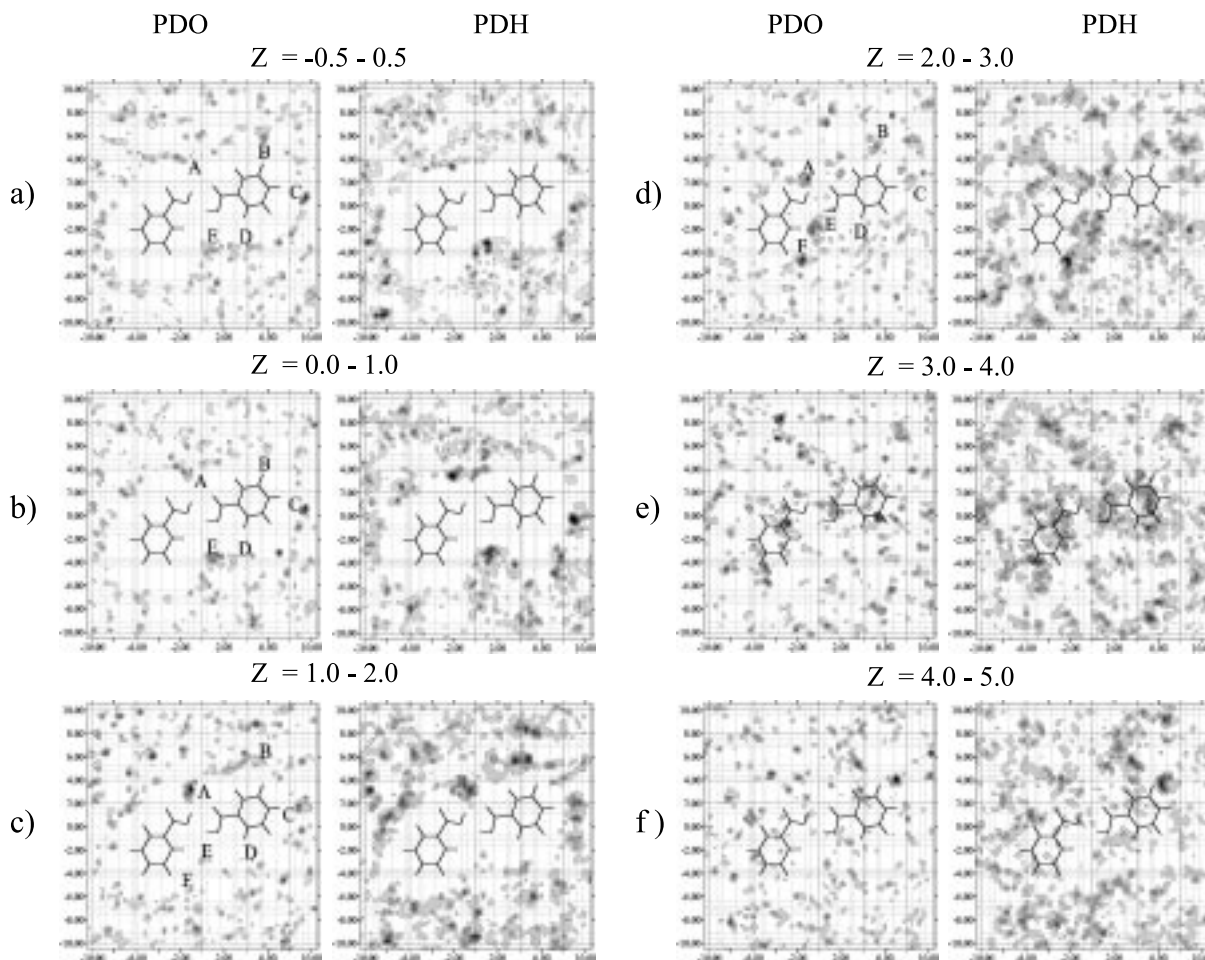
PDO   PDH   PDO   PDH
Z = -0.5 - 0.5   Z = 2.0 - 3.0

a)

Z = 0.0 - 1.0   Z = 3.0 - 4.0

b)

Z = 1.0 - 2.0   Z = 4.0 - 5.0

c)

Fig. 9. PDO and PDH maps of [SOT]$_{aq}$ at 298 K derived from MD-2a (X-, Y- and Z-axis in Å).

that BA can be efficiently removed from aqueous solution by polymer membrane only if the majority of BA present as neutral molecules.

In order to investigate the stability of the H-bond donor and acceptor interactions in mobile and stationary phases, one BA molecule in the CHP and SOT dimers was allowed to move in the course of MD-3 and MD-3a simulations, respectively. The equilibrium configurations of [CHP]$_{aq}$ and [SOT]$_{aq}$, obtained from MD-2 and MD-2a, were employed as starting configurations in MD-3 and MD-3a, respectively. Since the main aims of MD-3 and MD-3a were to investigate the stability and structures of the CHP and SOT dimers in

aqueous solutions, the discussion was made based only on selected $g(R)$.

Selected $g(R)$ derived from MD-3 are shown in Fig. 11. The characteristic structures of $g(R_{Ow-O1})$ and $g(R_{Ow-O2})$ derived from MD-3 are similar to those from MD-2. The ones resulted from MD-3 are slightly more structured. However, the main peaks of $g(R_{Ow-H1})$ are quite different. The structures and positions of the main peak of $g(R_{Ow-H1})$ obtained from MD-3 are more similar to those of MD-1, with a smaller peak height. This suggested the possibility to have an open or bent CHP dimer in aqueous solution. For $g(R_{O1-O2})$, a well-defined peak and a shoulder are seen at $R = 2.91$ and
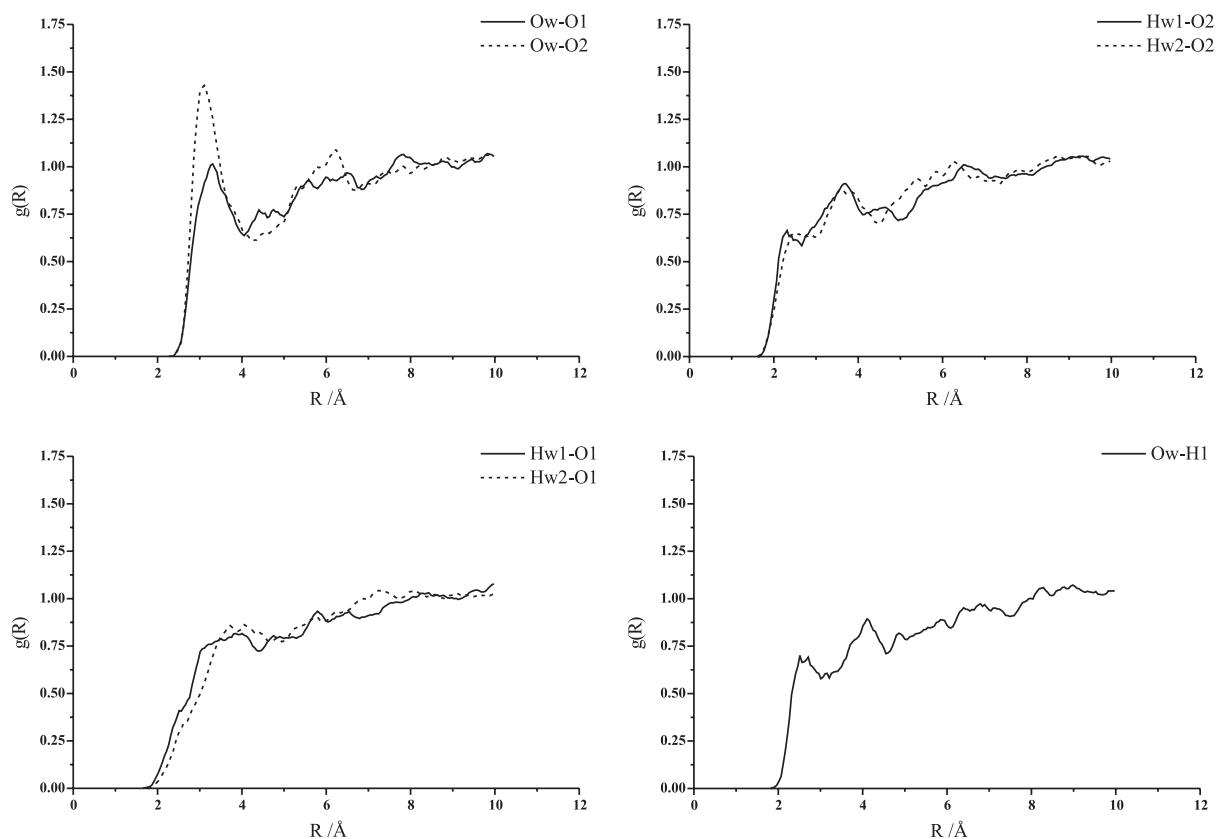
Fig. 10. Selected $g(R)$ of $[SOT]_{aq}$ at 298 K derived from MD-2a. Atom numbering system is in Table 1. Ow and Hw are the oxygen and hydrogen atoms of water, respectively.

3.41 Å, respectively. The main peak suggested that, on average, the CHP dimer is able to maintain its structure in aqueous solution. However, the presence of the shoulder of $g(R_{O1-O2})$ and the similarity between $g(R_{Ow-H1})$ obtained from MD-3 and MD-1 made believe that one of the C=O···H–O cyclic H-bonds in the CHP dimer could be bent or disrupted in the course of MD-3 simulations. Additional information on the structure of the CHP dimer can be acquired from $g(R_{O1-O1})$ and $g(R_{O2-O2})$ in Fig. 11. The main peak of $g(R_{O1-O1})$ splits into two peaks, one at $R = 4.01$ Å and the other at $R = 4.40$ Å. The former matches the O1···O1 non-H-bonded distance in the CHP dimer, whereas the latter represents the situation in which one of the cyclic C=O···H–O H-bonds is bent or disrupted. Since the main peak

of $g(R_{O2-O2})$ does not split, the bending could be attributed to the movement of the O–H group, away from the C=O group. Two snapshots displayed in Fig. 12 compare the CHP and bent CHP dimers.

Selected $g(R)$ derived from MD-3a are shown in Fig. 13. Similar results were found when a BA molecule in the SOT dimer was allowed to move in MD-3a. The characteristic structures of $g(R_{Ow-O1})$ and $g(R_{Ow-O2})$ derived from MD-3a and MD-2a are slightly different. The main peak of $g(R_{Ow-O1})$ obtained from MD-3a splits into two peaks, with the maximum at $R = 2.91$ and 3.26 Å, respectively. The former suggested an opening of the O···H–O H-bond in the SOT dimer for hydration. This is in accord with the structure and peak position of $g(R_{Ow-H1})$ which indicates that the H-bond do-
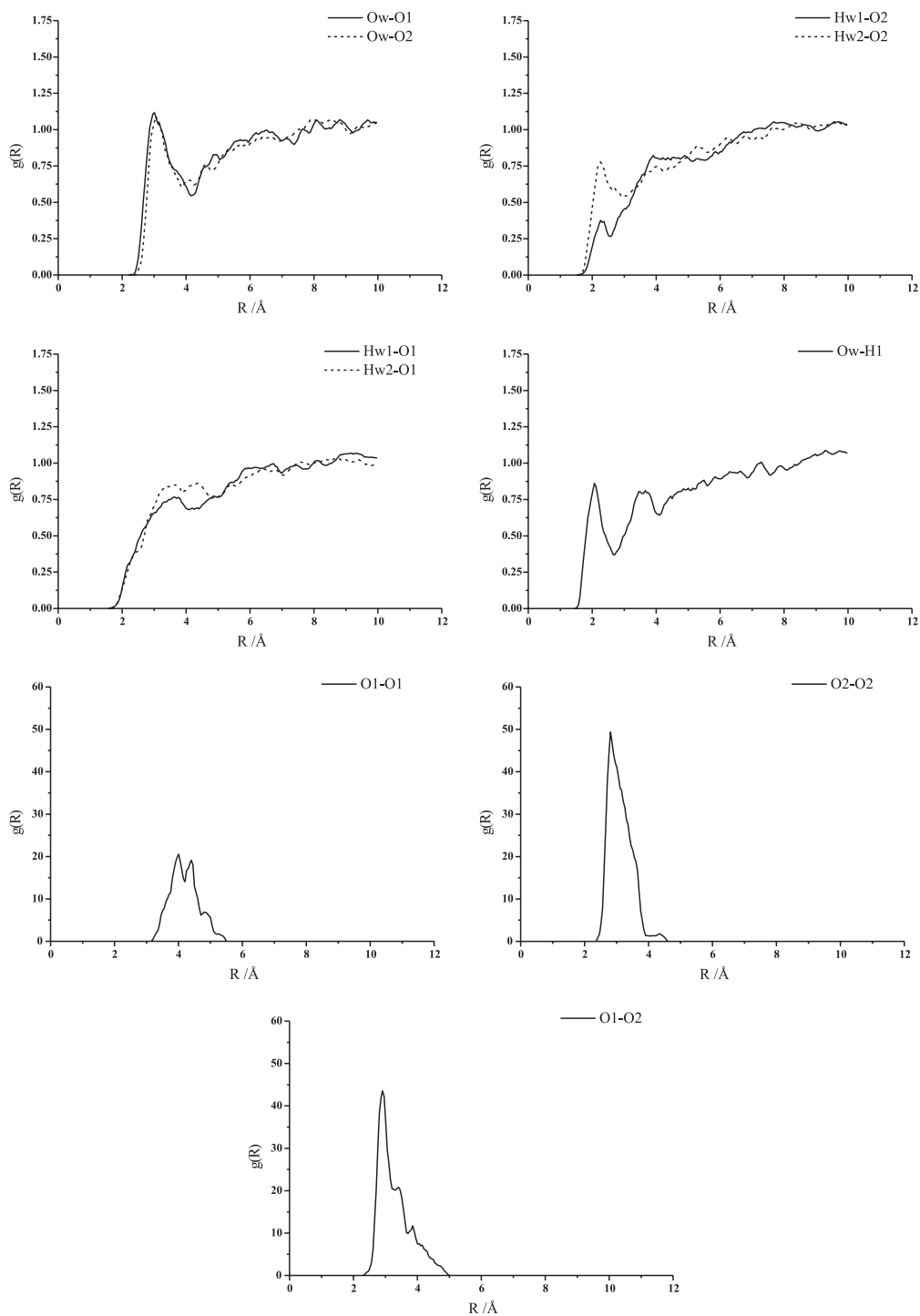
Fig. 11. Selected $g(R)$ of $[CHP]_{aq}$ at 298 K derived from MD-3. Atom numbering system is in Table 1. Ow and Hw are the oxygen and hydrogen atoms of water, respectively.

Fig. 12. Snapshots of the CHP dimers in [CHP]$_{aq}$ derived from MD-3. The initial and final configurations of the CHP and bent CHP dimers are shown without water molecules.

nating ability of the SOT dimer increases when a BA molecule is allowed to move. The main peak of $g(R_{O1-O2})$ also splits into two peaks, with the maximum at $R = 2.71$ and 3.01 Å, respectively. The position of the main peak of $g(R_{O1-O2})$ confirms that the O–H$\cdots$O=C H-bond remains associated while the SOT dimer is bending. Fig. 14 compares the SOT and bent SOT dimers.

Both BA molecules in the dimers were allowed to freely move in MD-4 and MD-4a. The results of MD-4 and MD-4a could suggest in detail how the cyclic H-bonds in the dimers can be destabilized or disrupted by solute–solvent H-bond interaction and thermal energy fluctuation. Since the structures and positions of the main peaks of $g(R_{Ow-O1})$, $g(R_{Ow-O2})$ and $g(R_{Ow-H1})$ obtained from MD-4 and MD-4a are in general not much different from [BA]$_{aq}$, they are not shown here. The results on $g(R_{O1-O2})$ indicated that, when both BA molecules in the CHP and SOT dimers were free to move, the cyclic H-bonds in the dimers were opened and finally disrupted in the course of MD simulations.

4.3. Stability of [BA]$_{aq}$ and [(BA)$_2$]$_{aq}$

The average potential energies ($\langle E_{aq}^{pot} \rangle$) as well as the translational ($\langle E_{aq}^{tr-BA} \rangle$) and rotational kinetic energies ($\langle E_{aq}^{rot-BA} \rangle$) of BA in aqueous solutions obtained from MD simulations are summarized in Table 5. The reported energy values were the results of the average over the number of water molecules and MD steps. In the present study, the stability of hydrated species is discussed using the average potential energy of pure water ($\langle E_w^{pot} \rangle$) as a reference. $\langle E_w^{pot} \rangle$ obtained from MD simulations of 300 water molecules at 298 K with T-model potential is -30.48 kJ mol^{-1}. The energy values in Table 5 comply with the above

discussion on the hydration structures of [BA]$_{aq}$ and [(BA)$_2$]$_{aq}$. A single BA is energetically the most well hydrated solute, with $\langle E_{aq}^{pot} \rangle = -30.76$ kJ mol^{-1}. Our MD simulations on [benzene]$_{aq}$, with the T-model potential and the same number of water molecules, yielded $\langle E_{aq}^{pot} \rangle = -30.71$ kJ mol^{-1} [32]. The energy results of MD-2 and MD-2a also confirmed that the supermolecule CHP and SOT dimers are not well hydrated. The stability of [(BA)$_2$]$_{aq}$ is slightly increased when a BA molecule in the CHP and SOT dimers was allowed to move in MD-3 and MD-3a. In this case, [SOT]$_{aq}$ is slightly more stable than [CHP]$_{aq}$. For [CHP]$_{aq}$, $\langle E_{aq}^{tr-BA} \rangle$ is larger than $\langle E_{aq}^{rot-BA} \rangle$, while the situation is opposite in the case of [SOT]$_{aq}$. When both BA molecules in the dimers were allowed to move in MD-4 and MD-4a, due to the thermal energy fluctuation at room temperature, the total kinetic energy of the solutes became large enough to break the cyclic H-bonds of both CHP and SOT dimers.

As mentioned earlier that some information on [AA]$_{aq}$ has been accumulated and the behavior of organic acid clusters in aqueous solutions might be extrapolated from this prototype system. Although the CHP dimer of AA possesses higher stability in the gas phase the results of ab initio calculations and Raman spectroscopic study [6] revealed that only the SOT dimer with bent H-bonds is predominant in aqueous solution. No direct evidence from the low-frequency Raman spectra showed the existence of stable acid–water pairs in [AA]$_{aq}$. The Raman spectra strongly suggested the presence of two microphases with homogeneous associated molecules: water cluster phase and AA cluster phase. Another low-frequency Raman experiment suggested the existence of an open AA dimer in [AA]$_{aq}$ [60]. In aqueous
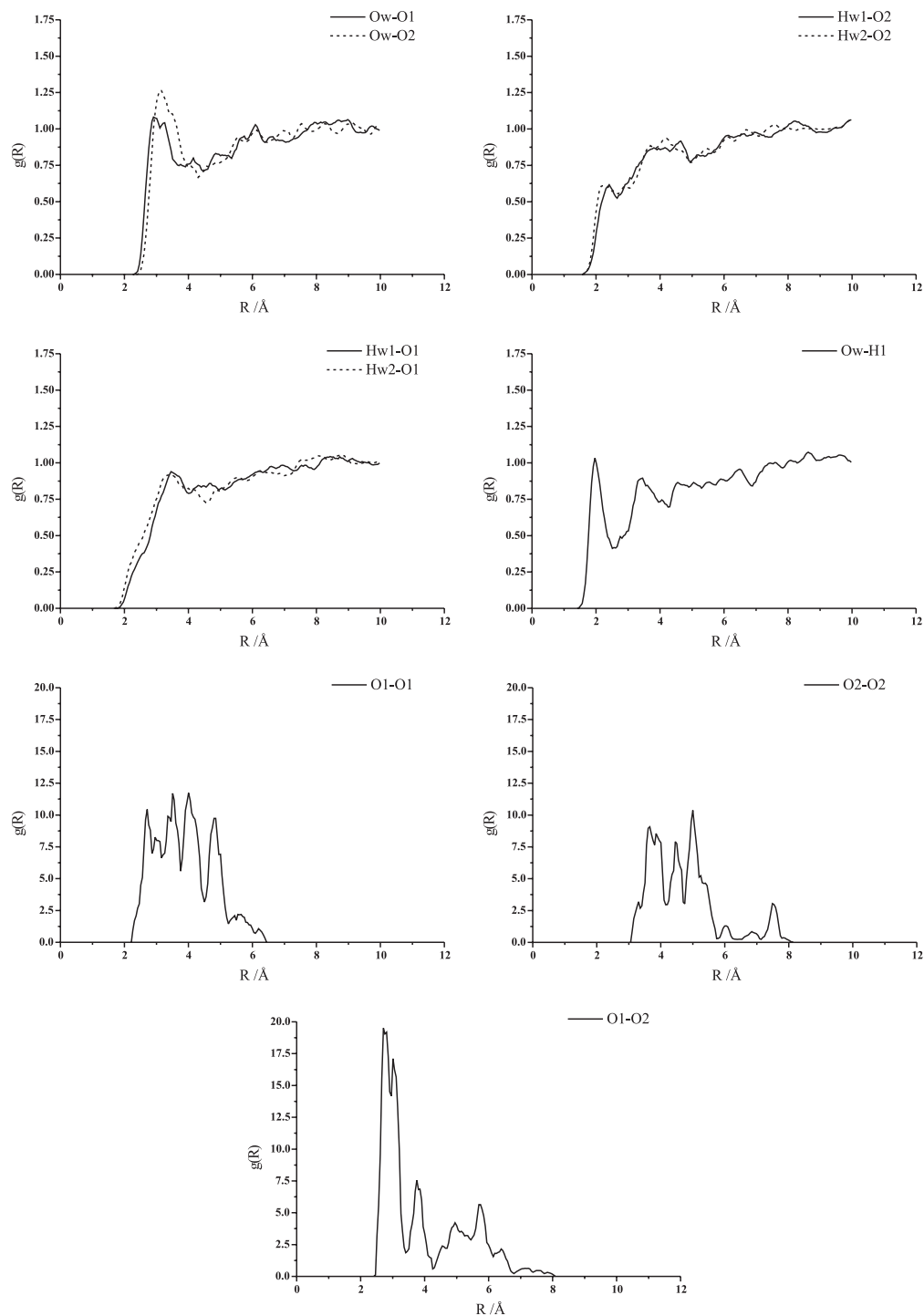
Fig. 13. Selected $g(R)$ of $[SOT]_{aq}$ at 298 K derived from MD-3a. Atom numbering system is in Table 1. Ow and Hw are the oxygen and hydrogen atoms of water, respectively.
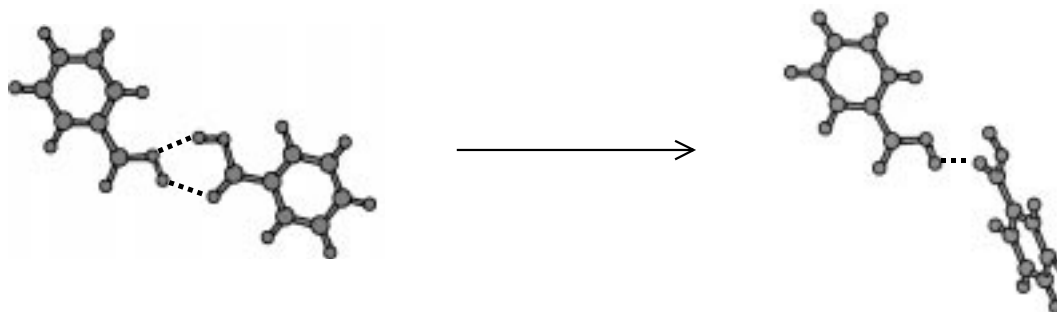
Fig. 14. Snapshots of SOT dimers in $[SOT]_{aq}$, derived from MD-3a. The initial and final configurations of the SOT and bent SOT dimers are shown without water molecules.

Table 5
The average potential energies ($\langle E_{aq}^{pot} \rangle$) of $[BA]_{aq}$ and $[(BA)_2]_{aq}$ at 298 K, derived from MD simulations. $\langle E_{aq}^{tr-BA} \rangle$ and $\langle E_{aq}^{rot-BA} \rangle$ are translational and rotational kinetic energies of BA, respectively[a]

	$\langle E_{aq}^{pot} \rangle$ (kJ mol^{-1})	$\langle E_{aq}^{tr-BA} \rangle$ (kJ mol^{-1})	$\langle E_{aq}^{rot-BA} \rangle$ (kJ mol^{-1})
MD-1	−30.76	–	–
MD-2	−30.41	–	–
MD-2a	−30.50	–	–
MD-3	−30.71	4.21	3.64
MD-3a	−30.75	3.42	4.29
MD-4	−30.45	4.11	4.06
		3.74	3.66
MD-4a	−30.63	2.44	3.63
		3.43	3.70

[a] The definitions of MD-*x* appeared in the text.

solution, formic acid (FA) behaves in the same way. Since the H-bond interaction in $(FA)_2$ is slightly stronger than that in $(AA)_2$, [61], the degrees of association of the dimer in aqueous solution could be higher. For $[FA]_{aq}$, significant and strong cyclic H-bond interactions between FA molecules were inferred from the results of Raman spectral analysis [62], in which the resonance energy transfer and non-coincident effects were applied to estimate the strength of the molecular associations.

It has been reported that a small fraction of benzene can be dissolved in aqueous solution [63,64]. The hydration structure of $[(benzene)_2]_{aq}$ was pointed out to be similar to $[benzene]_{aq}$ [64]. Our MD results on $[(benzene)_2]_{aq}$ and $[(benzene)_3]_{aq}$ revealed that two or three benzene

molecules do not form stable clusters in aqueous solutions [32]. The situation in $[(benzene)_2]_{aq}$ is similar to $[(BA)_2]_{aq}$, but different from $[(AA)_2]_{aq}$ and $[(FA)_2]_{aq}$. Apparently, the discrepancy is partly due to the presence of the aromatic ring, which results in a higher molecular mass of the solute, as well as larger area for solute–solvent interaction, compared to $[(AA)_2]_{aq}$ and $[(FA)_2]_{aq}$. The two effects lead naturally to a preferential hydration of a single solute molecule. An experimental investigation on BA in dilute aqueous solution suggested that the concentration of $(BA)_2$ is less than 3% of the concentration of the unionized BA [65,66].

The effects of molecular association on the diffusivities of various pseudoplanar solute at trace concentration in methanol and acetone were experimentally studied in Ref. [53]. In this study, the relative strength of H-bonding between solute and solute, as well as solute and solvent, was determined through diffusion measurements. The number of solvent molecules associated with aromatic solutes and the degrees of association were also determined from the diffusivities. The authors pointed out that all the aromatic acids considered behave more as H-bond donor than acceptor and the motions of the solute–solvent H-bonded complexes can be described in terms of the diffusivities of non-associated solutes and the H-bond acidity/basicity of the solutes. For aromatic solutes containing one polar functional group, the so-called excess reciprocal diffusivities (ΔD_{12}^{-1}) were required in the calculations of the diffusivities of the solute–solvent H-bonded complexes. ΔD_{12}^{-1} is a

function of the interaction energy of solute–solvent 1:1 complex obtainable from ab initio calculations. Apparently, no information on the solute–solute H-bond interaction was required in this approach [53]. This could mean that the effects of solute–solute H-bond interaction are negligible or the dimers do not exist in an appreciable amount in very dilute methanol and acetone solutions. The latter seems to be more conceivable and, therefore, in line with the present results.

Since in the gas phase $(BA)_2$ is much more associated than $(benzene)_2$, one could expect to find $(BA)_2$ in benzene solution. The self-association of BA and substituted BA in dilute benzene solutions was found indeed in an isopiestic experiment at 298.15 K [67,68]. The dimerization constant for the unsubstituted BA was reported to be the highest one. In benzene solutions, the mean association numbers of BA inferred from the dimerization constants were suggested to be between 1.6 and 1.8, depending on the concentration.

Hydration of polar organic molecules in organic solvents has been a subject of interest in the area of molecular association in solutions [65–73]. The partition of BA between the benzene and water phases has received much attention, especially in the experimental points of view [65,66,69–73]. It appeared that the increase in the distribution of BA in the benzene phase is directly related to the concentration of BA [70]. This experimental finding was explained as a consequence of the association of BA molecules in benzene [70], whereas many experimental investigations confirmed the existence of BA–H_2O complexes in the organic phase [65]. In Ref. [65], an experimental study on the hydration of BA in benzene solution, using coulometric titration and near IR spectrophotometric method, revealed that the concentrations of non-H-bonded water and H-bonded water, as well as the concentration of the hydrated BA monomers, increase as a function of the total water and BA concentrations in the benzene phase. With the assumption that the hydrated BA dimers can be ignored, the authors concluded that a single BA is hydrated by one water molecule in benzene solution. This conclusion is opposite to the results reported in Refs. [71–73], in which the excess of water solubility was explained by the tendency of

water to form H-bonds among themselves and two water molecules were suggested to H-bond directly to a single BA as well as the dimer in benzene solution [71,72].

Although the information obtained from the present investigation and from $[benzene]_{aq}$ [32] cannot provide a complete explanation on the experimental finding [70], comments could be made on the BA–H_2O–benzene systems. It is quite obvious from the minimum energy geometries and the interaction energies obtained from the T-model potentials that $(BA)_2$ and $(benzene)_2$ can be separated by one or two water molecules in the gas phase. However, the two monomers can still be linked together by water molecules, through the O–H$\cdots\pi$ and O–H$\cdots$O H-bonds in the case of benzene and BA, respectively. Our preliminary theoretical study revealed that the T-shaped structure, in which the O–H group of BA acts as proton donor towards the π clouds of benzene, represents the absolute minimum energy geometry of the benzene–BA 1:1 complex. The interaction energy for this H-bond complex is -16.97 kJ mol^{-1}, compared with -9.26 and -11.87 kJ mol^{-1} for $(benzene)_2$ and benzene–H_2O 1:1 complex, respectively. Therefore, from the energetic viewpoint the degrees of molecular association in BA–H_2O–benzene system decrease in the sequence:

$(BA)_2 >$ BA–H_2O 1:1 complex $> (H_2O)_2 >$ benzene–BA 1:1 complex $>$ benzene–H_2O 1:1 complex $> (benzene)_2$.
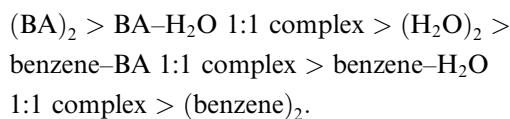
Although the H-bond interaction in $(BA)_2$ predominates in this sequence, it was shown in the present work that $(BA)_2$ cannot maintain their structures in aqueous solution. This makes one believe that both weak and strong intermolecular interactions contribute more or less to the behavior of molecules in the BA–H_2O–benzene system and they must be taken into consideration in both theoretical [25] and experimental model [70]. It should also be pointed out that the thermal energy fluctuation must be taken into account if a reasonable explanation on the behavior of associated molecules in solutions is the major objective [70]. Progress is being made in our laboratory to apply

the T-model potentials and MD simulations in the study of the BA–H$_2$O–benzene system.

5. Conclusion

Intermolecular potential to describe the interaction in BA clusters and aqueous solution was constructed using the T-model. The T-model potential was applied in the calculations of equilibrium structures and interaction energies of BA–H$_2$O 1:1, 1:2, 2:1 and 2:2 complexes. In order to check the performance of the computed T-model potential, representative results of the 1:1 and 1:2 complexes were examined using ab initio calculations, at both HF and MP2 levels of accuracy. It was found that the structural and energetic results obtained from the T-model potential are in good agreement with ab initio calculations. Cyclic H-bond arrangements were suggested in the present study to be the most preferential binding features for all BA–H$_2$O complexes. There was no direct evidence showing the existence of the O–H⋯π interaction in BA–H$_2$O complexes. The results on the 2:1 and 2:2 complexes suggested the possibility to break cyclic H-bonds in the dimers by BA–H$_2$O H-bond interaction. The possibility to observe π–π interaction increased when going from 2:1 to 2:2 complexes, in which several herringbone structures were found to be minimum energy geometries. The present results also suggested the possibility that the AMBER force fields with the fit charges are inadequate to describe accurately the interaction in BA clusters, in which the interaction between the aromatic rings must be properly taken into account.

The hydration structure of [BA]$_{aq}$ was investigated in the present study by conducting a series of MD simulations. The PDO and PDH maps as well as g(R), derived from MD simulations, revealed that the OH group represents the most preferential hydration site of BA and three to four water molecules H-bond directly to the COOH group. Water molecules were found to hydrate at the distance of 3–4 Å above the aromatic ring, which is in good agreement with the results of [benzene]$_{aq}$, [phenol]$_{aq}$ and [hydroxybenzoic acid]$_{aq}$. This finding leads to the conclusion that the hydration

structure above the aromatic ring does not change substantially by substituents such as the COOH and OH group.

In general, the hydration structures of the supermolecule CHP and SOT dimers are not much different, except at the cyclic H-bonds. In both cases, hydration structures could be established both at the polar and non-polar parts of the BA dimers. Both C=O and O–H groups in the supermolecule CHP dimer cannot be fully active in the H-bond formation with water. MD results revealed that the supermolecule SOT dimer becomes a slightly better proton acceptor towards water molecules, compared to BA and the supermolecule CHP dimer.

In order to model the interactions between H-bond donor and acceptor in mobile and stationary phases, as well as the possibility to break cyclic H-bonds in the CHP and SOT dimers by solute–solvent interaction, additional MD simulations were performed. One BA molecule in the CHP and SOT dimers was allowed to move in the course of MD simulations. It was found that one H-bond in both CHP and SOT dimers was disrupted by the solute–solvent H-bond interaction and both dimer structures were transformed into bent structures. Due to the thermal energy fluctuation and solute–solvent H-bond interaction, the cyclic H-bonds were completely disrupted when two BA molecules in the dimers were allowed to move. These results prove that, in aqueous solution, (BA)$_2$ are not stable and they could remain associated only when a BA molecule is put in a stationary phase. The results seem to be different when (BA)$_2$ are put in non-polar solvent such as benzene. In this case, the self-association between BA molecules was evident in an isopiestic experiment. The preferential solute–solvent H-bond interaction in [BA]$_{aq}$ suggested in the present work is compatible with the experimental results on solvation of aromatic acids in dilute methanol and acetone solutions. In aqueous solution, one may conclude that the solute–solute H-bond interactions between small organic acids, e.g. (AA)$_2$ and (FA)$_2$, are stronger than those in aromatic acid dimers. It was also iterated in the present report that all weak and strong intermolecular interactions, as well as the effects of thermal energy fluctuation, must be

taken into consideration if a reasonable explanation on the behavior of molecular associations in solutions is the major objective.

It should be noted that the MD results reported in the present paper were based on a pair-wise additive scheme, in which the many-body effects are neglected. It has been well recognized that the cooperative effects tend to increase due to an increase in the cluster size. Although in aqueous solutions the BA dimers are not strongly associated, it is reasonable to believe that, due to the cooperative and hydrophobic effects, the H-bond association between the BA solutes could be more pronounced if the number of molecules in the aggregate increase. Therefore, the presence of microphases of BA clusters in aqueous solution, as in the case of $[AA]_{aq}$ and ethanol, is likely possible.

Acknowledgements

References

[1] M.M. Flocco, S.L. Mowbray, J. Mol. Biol. 235 (1994) 709.
[2] S.K. Burley, G.A. Petsko, FEBS Lett. 203 (1986) 139.
[3] S.K. Burley, G.A. Petsko, Adv. Protein Chem. 39 (1988) 125.
[4] S. Sun, E.R. Bernstein, J. Phys. Chem. 100 (1996) 13348.
[5] S. Suzuki, P.G. Green, R.E. Bumgarner, S. Dasgupta, W.A. Goddard III, G.A. Blake, Science 257 (1992) 942.
[6] N. Nishi, T. Nakabayashi, K. Kosugi, J. Phys. Chem. A 103 (1999) 10858.
[7] N. Nishi, S. Takahashi, M. Matsumoto, A. Tanaka, K. Muraya, T. Takamuku, T. Yamaguchi, J. Phys. Chem. 99 (1995) 462.
[8] H.R. Mulla, A. Cammers-Goodwin, J. Am. Chem. Soc. 122 (2000) 738.
[9] U. Samanta, P. Chakrabarti, J. Chandrasekhar, J. Phys. Chem. 102 (1998) 8964.
[10] H. Park, S. Lea, Chem. Phys. Lett. 301 (1999) 487.
[11] P. Hobza, H.L. Selzle, E.W. Schlag, J. Phys. Chem. 100 (1996) 18790.
[12] P. Hobza, H.L. Selzle, E.W. Schlag, J. Phys. Chem. 97 (1993) 3937.
[13] V. Lukes, M. Breza, S. Biskupic, Theor. Chem. Acc. 101 (1999) 319.
[14] W.L. Jorgensen, D.L. Severance, J. Am. Chem. Soc. 112 (1990) 4768.
[15] R.S. Miller, D.Y. Curtin, I.C. Paul, J. Am. Chem. Soc. 96 (1974) 6329.
[16] R.S. Miller, D.Y. Curtin, I.C. Paul, J. Am. Chem. Soc. 96 (1974) 6340.
[17] L. Leiserowitz, Acta. Cryst. B 32 (1976) 775.
[18] C.C. Wilson, N. Shankland, A.J. Florence, Chem. Phys. Lett. 253 (1996) 103.
[19] C. Scheurer, P. Saalfrank, J. Chem. Phys. 104 (1996) 2869.
[20] H.R. Zelsmann, Z. Mielke, Chem. Phys. Lett. 186 (1991) 501.
[21] A. Stoeckli, B.H. Meier, R. Kreis, R. Meyer, R.R. Ernst, J. Chem. Phys. 93 (1990) 1502.
[22] S. Nagaoka, N. Hirota, T. Matsushita, K. Nishimoto, Chem. Phys. Lett. 92 (1982) 498.
[23] R.G. Compton, M.S. Harding, M.R. Pluck, J.H. Atherton, C.M. Brennan, J. Phys. Chem. 97 (1993) 10416.
[24] P.I. Nagy, W.J. Dunn III, G. Alagona, C. Ghio, J. Phys. Chem. 97 (1993) 4628.
[25] P. Cieplak, W. Pawlowski, E. Wieckowska, Z. Phys. Chem. 177 (1992) 63.
[26] K.P. Sagarik, R. Ahlrichs, S. Brode, Mol. Phys. 57 (1986) 1247.
[27] K.P. Sagarik, R. Ahlrichs, J. Chem. Phys. 86 (1987) 5117.
[28] K.P. Sagarik, V. Pongpituk, S. Chiyapongs, P. Sisot, Chem. Phys. 156 (1991) 439.
[29] K.P. Sagarik, J. Mol. Struct. (Theochem) 465 (1999) 141.
[30] K.P. Sagarik, E. Spohr, Chem. Phys. 199 (1995) 73.
[31] K.P. Sagarik, P. Asawakun, Chem. Phys. 219 (1997) 173.
[32] K.P. Sagarik, P. Sisot, in preparation.
[33] K.P. Sagarik, R. Ahlrichs, Chem. Phys. Lett. 131 (1986) 74.
[34] H.J. Boehm, R. Ahlrichs, J. Chem. Phys. 77 (1982) 2028.
[35] H.J. Boehm, C. Meissner, R. Ahlrichs, Mol. Phys. 53 (1984) 651.
[36] H.J. Boehm, R. Ahlrichs, Mol. Phys. 55 (1985) 1159.
[37] D.E. Williams, R.R. Weller, J. Am. Chem. Soc. 105 (1983) 4143.
[38] C.M. Breneman, K.B. Wiberg, J. Comput. Chem. 11 (1990) 361.
[39] C.S. Brooks, M.E. Hobbs, J. Am. Chem. Soc. 62 (1940) 2851.
[40] O. Matsuoka, E. Clementi, M. Yoshimine, J. Chem. Phys. 64 (1976) 1351.

[41] K.P. Sagarik, P. Asawakun, in preparation.

[42] M. Neumann, D.F. Brougham, C.J. McGloin, M.R. Johnson, A.J. Horsewill, H.P. Trommsdorf, J. Chem. Phys. 109 (1998) 7300.

[43] G. Meijer, M.S. de Vries, H.E. Hunziker, H.R. Wendt, J. Chem. Phys. 92 (1990) 7625.

[44] K. Coutinho, S. Canuto, M.C. Zerner, Int. J. Quant. Chem. 65 (1997) 885.

[45] P. Hobza, Z. Havlas, Theor. Chem. Acc. 99 (1998) 372.

[46] P. Hobza, Z. Havlas, Chem. Phys. Lett. 303 (1999) 447.

[47] J.B. Foresman, A. Frisch, Exploring Chemistry with Electronic Structure Methods, GAUSSIAN, Pittsburgh, PA, 1996.

[48] H.B. Schlegel, J. Comput. Chem. 3 (1982) 214.

[49] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T. Keith, G.A. Petersson, J.A. Montgomery, K. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. Defrees, J. Baker, J.P. Stewart, M. Head-Gordon, C. Gonzalez, J.A. Pople, GAUSSIAN, Pittsburgh, PA, 1995.

[50] F.G. Brockman, M. Kilpatric, J. Am. Chem. Soc. 56 (1934) 1483.

[51] M. Eisenberg, P. Chang, G.W. Tobias, C.R. Wilke, AIChE J. 1 (1955) 558.

[52] E. Clementi, Computational Aspects for Large Chemical Systems, Lecture Notes in Chemistry 2, Springer, Berlin, 1980.

[53] J.G. Lu, R. Kong, T.C. Chan, J. Chem. Phys. 110 (1999) 3003.

[54] T.C. Chan, N.L. Ma, N. Chen, J. Chem. Phys. 107 (1997) 1890.

[55] J. Gao, J. Phys. Chem. 96 (1992) 6432.

[56] R.A. Noulty, D.G. Leaist, J. Chem. Engng. Data 32 (1987) 418.

[57] G. Laufenberg, S. Hausmanns, B. Kunz, J. Membr. Sci. 110 (1996) 59.

[58] M. Waksmundzka-Hajnos, J. Chromatogr. B 717 (1998) 93.

[59] K. Rzeszutek, A. Chow, Talanta 47 (1998) 697.

[60] K. Kosugi, T. Nakabayashi, N. Nishi, Chem. Phys. Lett. 291 (1998) 253.

[61] B.S. Jursic, J. Mol. Struct. Theochem. 417 (1997) 89.

[62] R.J. Bartholomew, D.E. Irish, J. Raman Spectrosc. 30 (1999) 325.

[63] C. Chipot, P.A. Kollman, D.A. Pearlman, J. Comput. Chem. 17 (1996) 1112.

[64] P. Linse, J. Am. Chem. Soc. 114 (1992) 4366.

[65] E. Wieckowska, W. Pawlowski, Z. Phys. Chem. 166 (1990) 167.

[66] W. Pawlowski, E. Wieckowska, Z. Phys. Chem. 168 (1990) 205.

[67] R. Barela, G. Liwski, H. Szatylowicz, Fluid Phase Equilibria 105 (1995) 119.

[68] R. Barela, H. Buchowski, H. Szatylowicz, Fluid Phase Equilibria 92 (1994) 303.

[69] H. Yamada, K. Yajima, H. Wada, G. Nakagawa, Talanta 42 (1995) 789.

[70] W. Nernst, Z. Phys. Chem. 8 (1891) 110.

[71] R. van Duyne, S.A. Taylor, S.D. Christian, H.E. Affsprung, J. Phys. Chem. 71 (1976) 3427.

[72] L. Barcza, A. Buvari, M. Vajda, Z. Phys. Chem. 102 (1976) 35.

[73] Y. Hasegawa, T. Unno, E.R. Choppin, Inorg. Nucl. Chem. 43 (1981) 2154.

[74] K. Refson, MOLDY User's Manual, Department of Earth Science, Oxford University, 1999.