

Supplementary Material (ESI) for Chemical Communications

Parallel alignment of water and aryl rings – crystallographic and theoretical evidence for the interaction

Bojana D. Ostojić^a, Goran V. Janjić^a, and Snežana D. Zarić^{*,b}

^a*IHTM, University of Belgrade, Njegoševa 12, P. O. Box 473, 11001 Belgrade, Serbia*

^b*Department of Chemistry, University of Belgrade, Studentski trg 16, P.O. Box 158, 11001 Belgrade, Serbia*

*To whom the correspondence should be addressed e-mail: szaric@chem.bg.ac.yu

Contents:

I Database screening and the results of the CSD search

II *Ab initio* calculations

III References

I Database screening and the results of the CSD search

The statistical study is based on the crystal structures archived in the Cambridge Structural Database (November 2007 release, version 5.29).¹ The crystal structures involving non coordinated water molecule and C₆-aryl group were screened for intermolecular contacts. The CSD search program ConQuest 1.10² was used to retrieve structures satisfying the following criteria: a) the crystallographic R factor \downarrow 10% b) the error-free coordinates according to the criteria used in the CSD c) the H-atom positions were normalized using the CSD default X-H bond lengths d) no polymer structure e) no powder structures f) not disordered structures g) 3D coordinates determined. The number of structures bearing a water molecule and a six-membered aromatic moiety (Ar) was 18902. The geometric parameters used and the atom labelling scheme are displayed in Fig. S1.

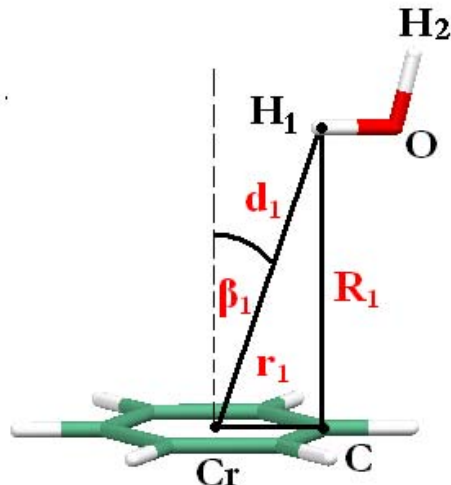


Figure S1. The geometric parameters and atom labelling used for the description of parallel-alignment interaction. C_r - the centroid of the C₆-aromatic ring; d₁, d₂, and d₀ – the distances between H₁, H₂, and O, respectively, and the centroid of the ring; R₁, R₂, and R₀ - the normal distances from the ring plane to the H₁, H₂, and O atoms, respectively; the offsets r₁, r₂, and r₀ - the distances from the ring centroid to the projection of the H₁, H₂, and O atom position, respectively, on the plane of the ring; the angles β_1 , β_2 , and β_0 are measured between the C_r-H₁, C_r-H₂, and C_r-O lines, respectively, and the normal to the ring plane.

In the initial step of the analysis we searched for all contacts found in crystal structures characterized by the distance d_1 less than 4.5 Å ($d_1 < 4.5$ Å). In further step of the analysis intermolecular contacts were retrieved with the distance R_1 less than 4.0 Å and the offset r_1 less than 2.0 Å. The hydrogen atom of the water molecule that is in that area and that is closer to the center of the ring comparing to the second hydrogen atom of the water molecule is denoted by H_1 . The number of entries with short intermolecular contacts was 1598.

In the next step of the analysis the structures with the plane–plane angle between water molecule and aryl-fragments of less than 10° (approximately parallel molecular planes) were isolated (type A set). In the subsequent search, the type B set comprised of structures containing the O- H_1 bond approximately parallel to the aryl ring plane (the angle between the O- H_1 vector and the normal to the ring plane was in the range 80° to 90°) was obtained. Finally, the structures of the type C characterized by the O- H_2 bond in a parallel alignment with the ring plane (the angle between the O- H_2 vector and normal to the ring plane was in the range 80° to 90°) were isolated. In the final step of data gathering the structures passing the above criteria were individually inspected on the graphics display. On the data obtained from the CSD area corrections³ were also performed.

A survey of crystal structures from the CSD that contain water molecule and aryl groups revealed that the total number of observed parallel contacts was 720. The number of intermolecular contacts satisfying the criteria for the type A was 79. 391 contacts were retrieved in which the O- H_1 bond was parallel to the ring plane (type B set), while the number of contacts that involved O- H_2 bond parallel to the ring plane (type C set) was 250.

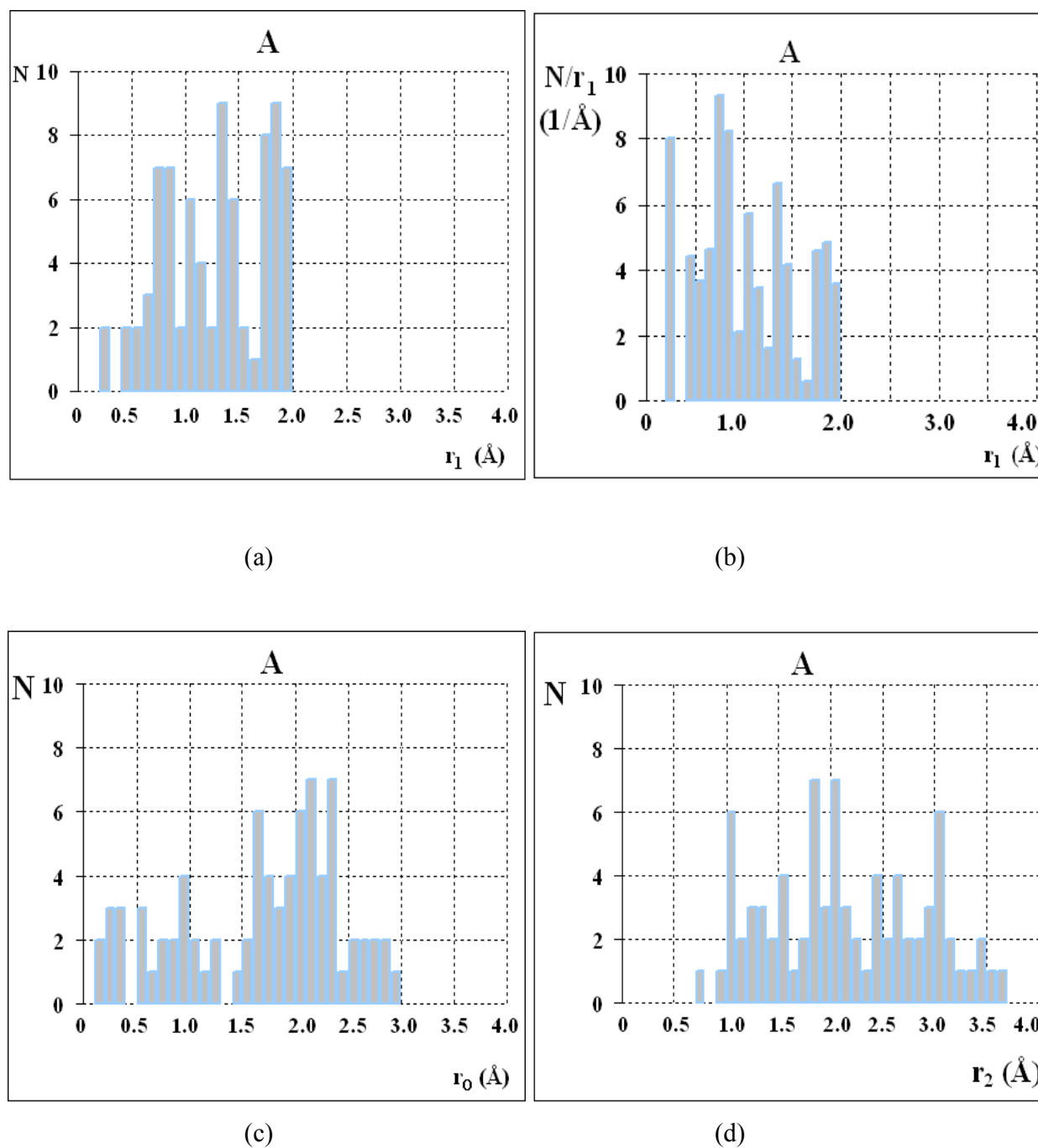


Figure S2. The frequency distributions of offsets r_1 (a), area-corrected offsets r_1 (b), offsets r_0 (c), and offsets r_2 (d) in retrieved structures of the type A. (A bar centered at, e.g., 1.0 Å represents contacts ≥ 1.0 Å and < 1.1 Å.)

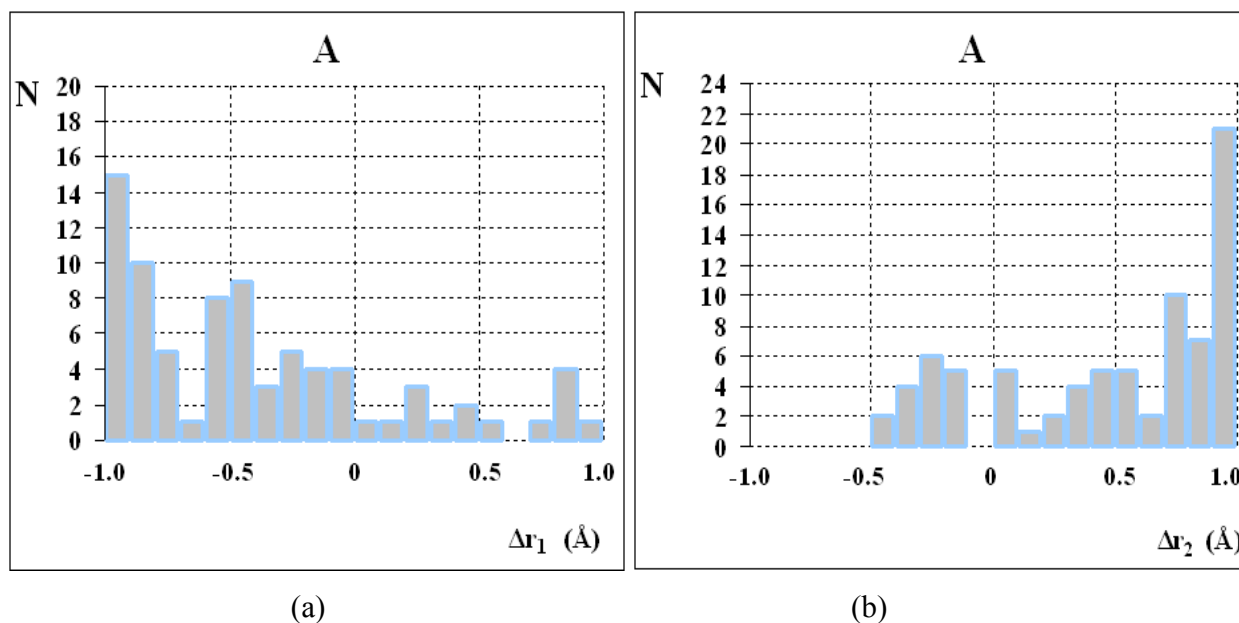


Figure S3. The histograms of the parameters $\Delta r_1 = r_1 - r_0$ (a) and $\Delta r_2 = r_2 - r_0$ (b) in retrieved crystal structures of the type A.

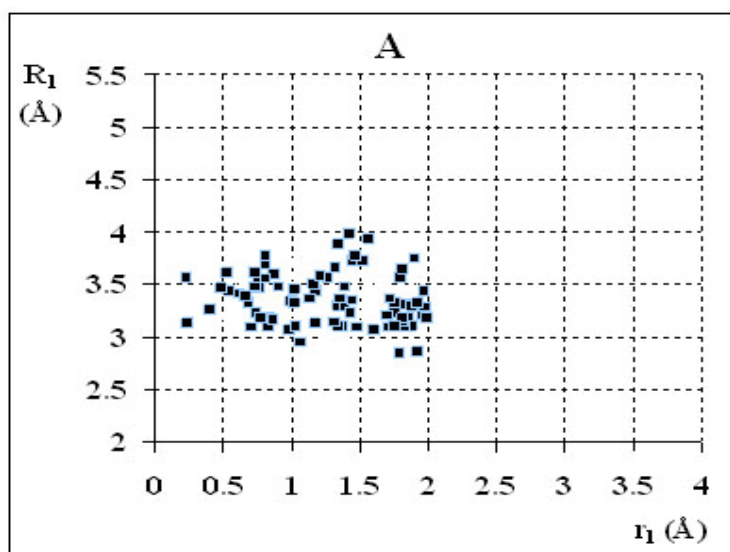


Figure S4. The scattergram for the correlation of the normal distance R_1 with the offset r_1 in intermolecular contacts of the type A.

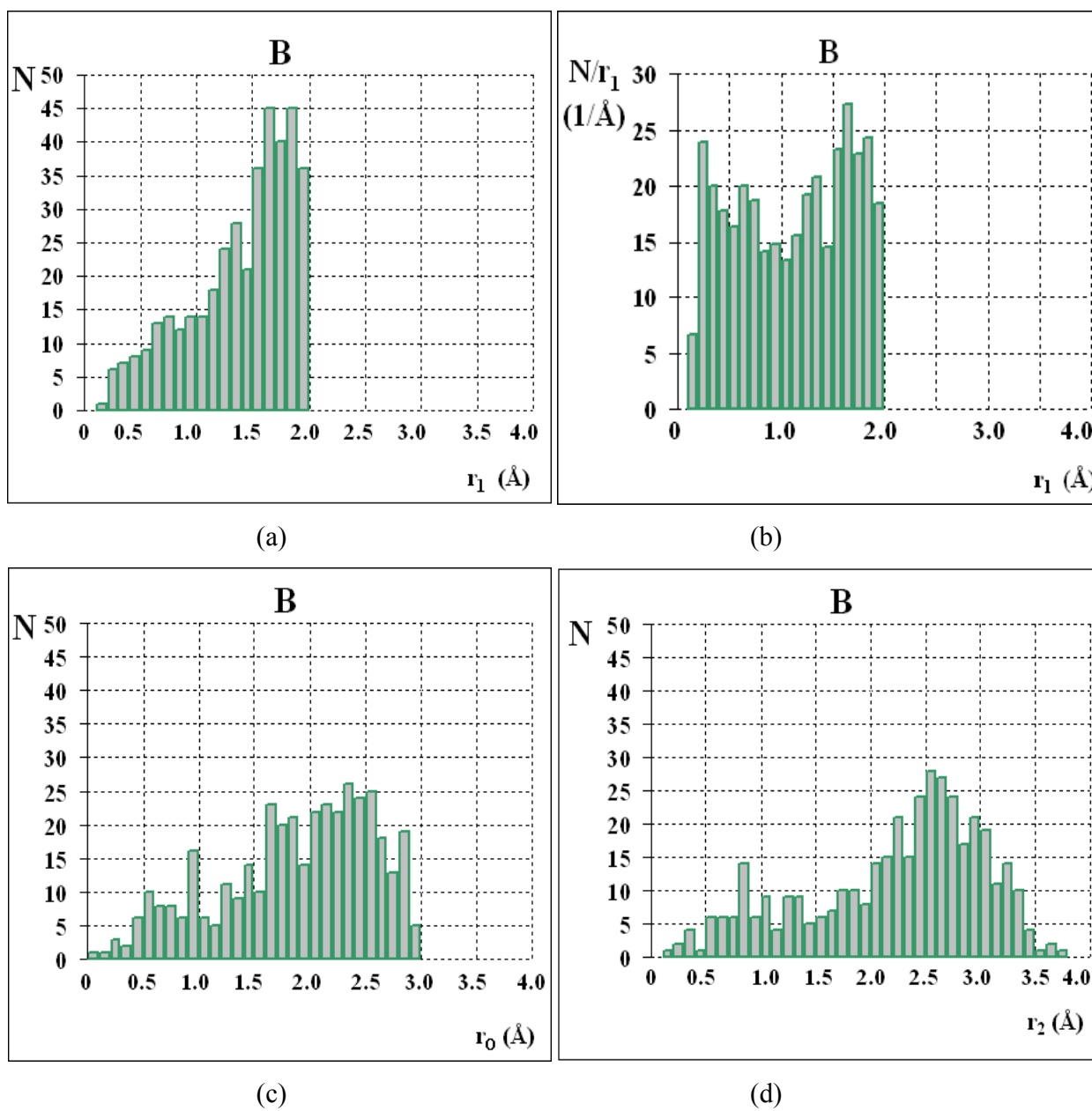


Figure S5. The distributions of offsets r_1 (a), area-corrected offsets r_1 (b), offsets r_0 (c), and offsets r_2 (d) in intermolecular contacts of the type B.

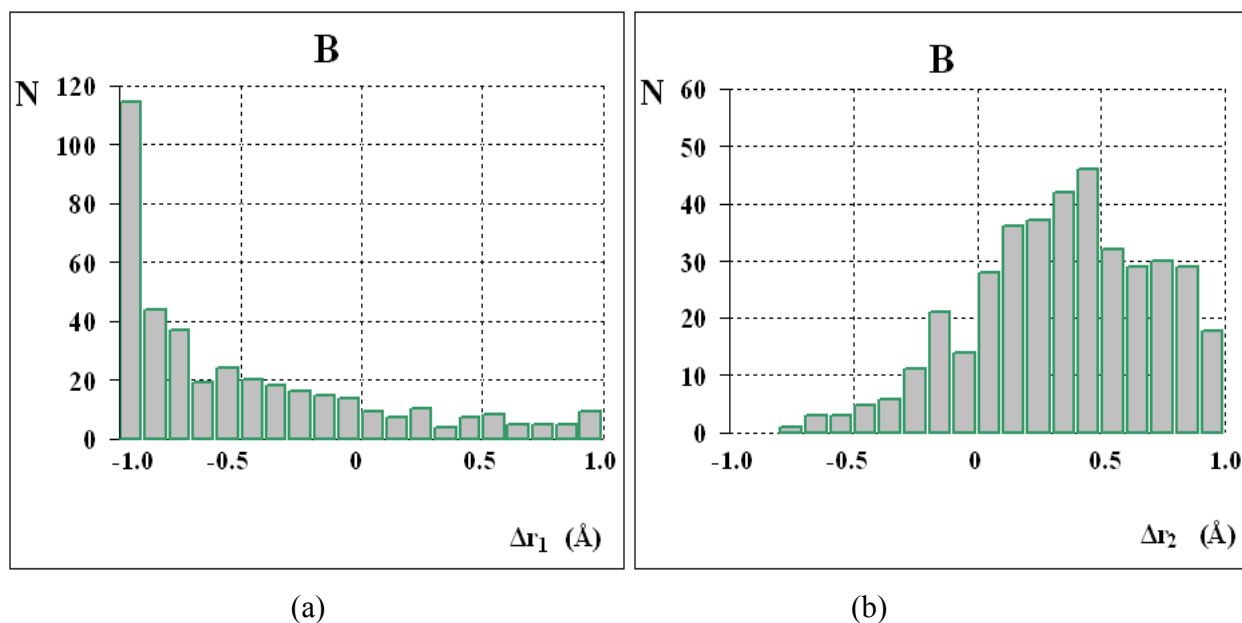


Figure S6. The histograms of the parameters $\Delta r_1 = r_1 - r_0$ (a) and $\Delta r_2 = r_2 - r_0$ (b) in retrieved crystal structures of the type B.

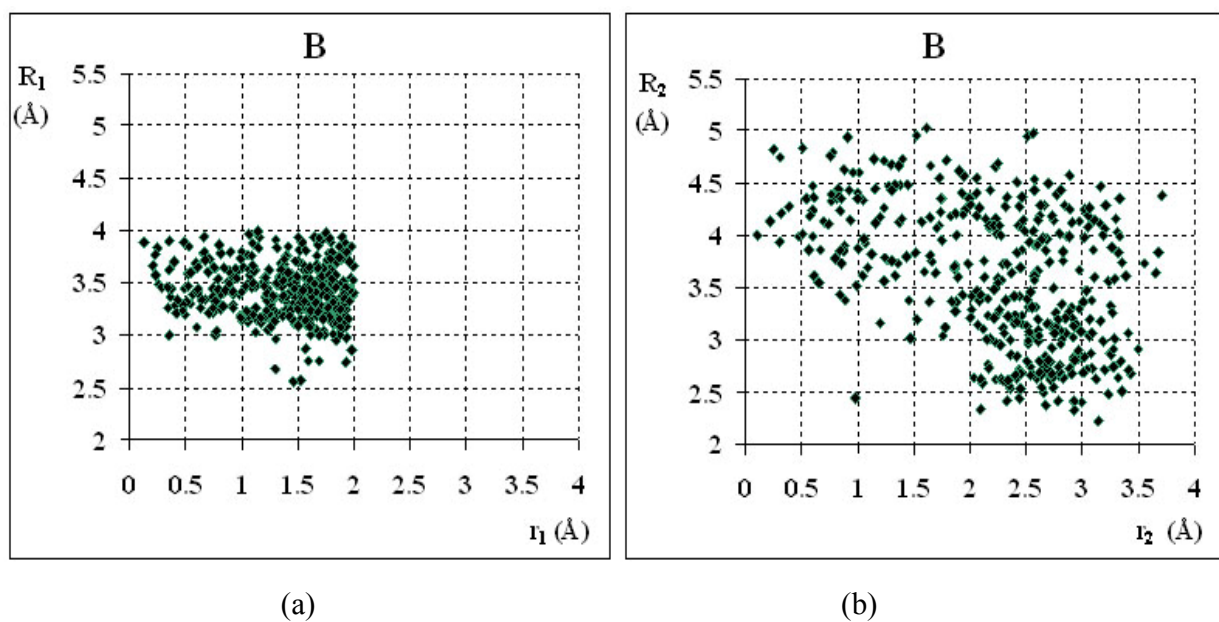


Figure S7. The scattergrams for the correlation of the normal distance R_1 with the offset r_1 (a) and of the normal distance R_2 with the offset r_2 (b) in intermolecular contacts of the type B.

The results of the additional CSD search that involved only selected structures of the type A

Another survey of crystal structures from the CSD included additional criterion to those seven mentioned above. We carefully selected only structures with water bound hydrogen atoms that were located from the difference Fourier maps in the X-ray structure determination. The new type A set involved structures with hydrogen positions fixed on their parent oxygen atom as well as those structures which hydrogen atom positions were refined with isotropic displacement parameters. In the latter case the number of structures is rather small (only 11). After duplicate REFCODES were removed the number of entries with short intermolecular contacts was 44.

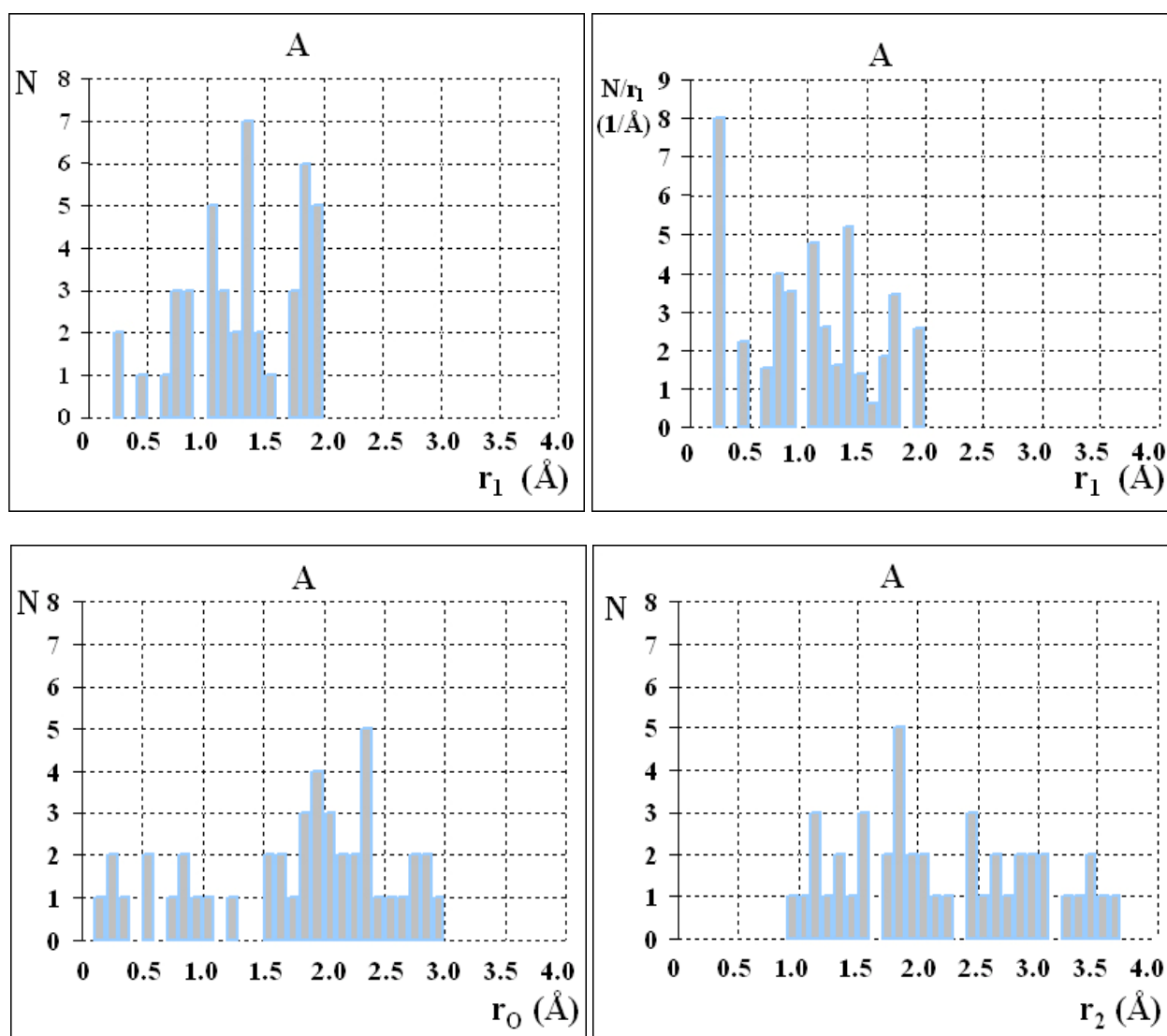


Figure S8. The frequency distributions of offsets r_1 (a), area-corrected offsets r_1 (b), offsets r_0 (c), and offsets r_2 (d) in retrieved structures of the new type A set. (A bar centered at, e.g., 1.0 Å represents contacts ≥ 1.0 Å and < 1.1 Å.)

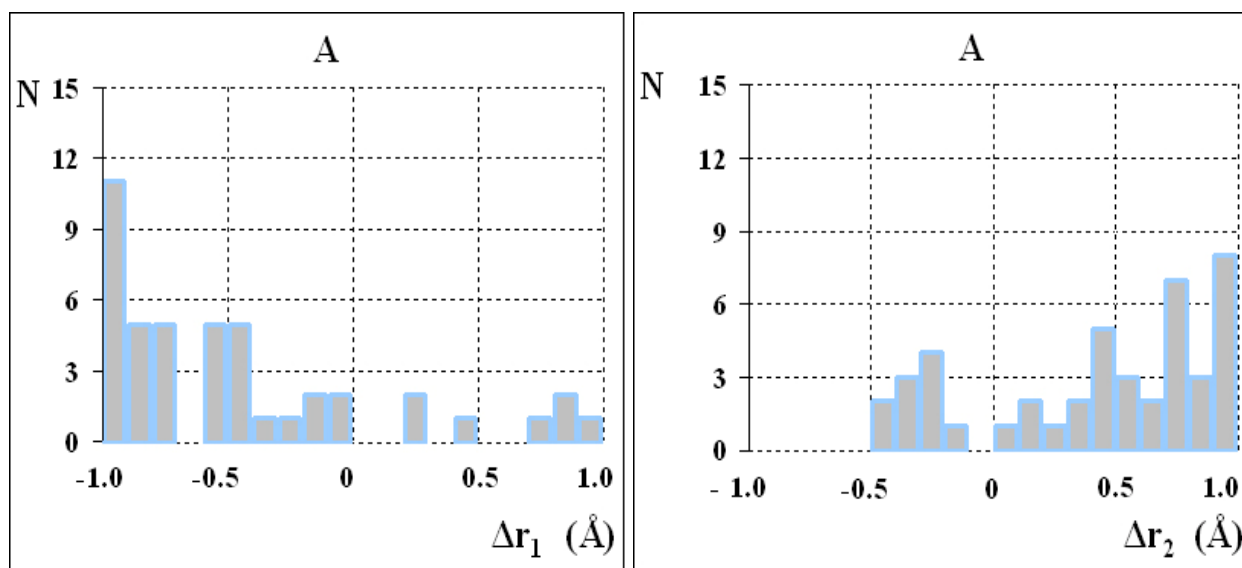


Figure S9. The histograms of the parameters $\Delta r_1 = r_1 - r_0$ (a) and $\Delta r_2 = r_2 - r_0$ (b) in retrieved crystal structures of the new type A set.

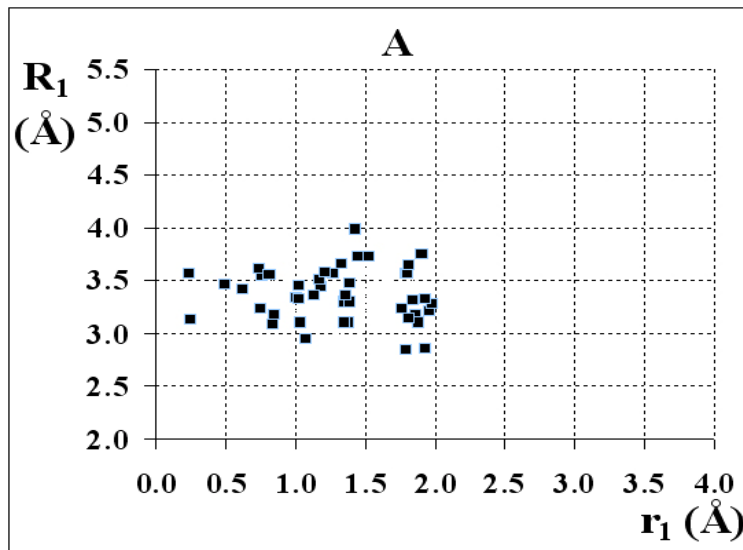


Figure S10. The scattergram for the correlation of the normal distance R_1 with the offset r_1 in intermolecular contacts of the new type A set.

The comparison of the corresponding figures obtained from data for two searches of crystal structures from the CSD (described above) that belong to the type A shows similar patterns. It indicates that the quality

of the structures does not influence our conclusions about the trends in parallel-alignment conformations in crystal structures.

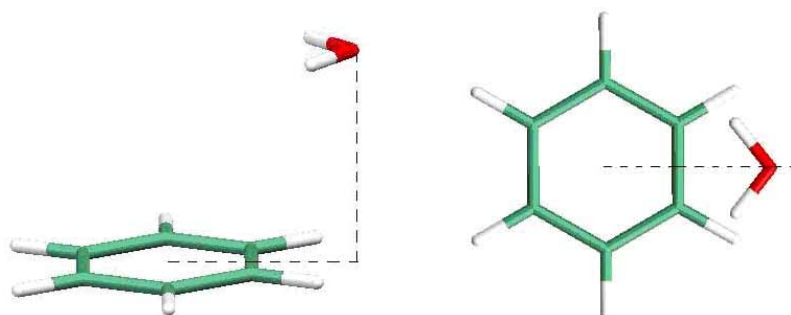
II *Ab initio* calculations

The interactions were modeled by the water–benzene dimers (Fig. S14). The geometries of isolated molecules, the benzene and the water molecule, were optimized by the gradient optimization using the cc-pVTZ basis set⁴ and Møller-Plesset second-order perturbation method (MP2)⁵. They were used for the calculation of interaction energies. The vertical distance between the benzene molecular plane and the OH bond (or both OH bonds) of the water molecule (denoted by R) was systematically varied while the monomer geometries were kept rigid. These optimum intermonomer distances were computed at the MP2/cc-pVTZ level. Only the structure that corresponds to the absolute minimum of the water-benzene dimer (C_{\min}) was fully optimized at the MP2/aug-cc-pVDZ level.

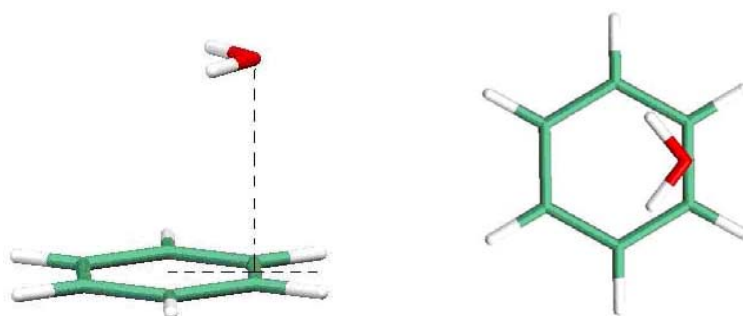
The supermolecular approach was employed to investigate the interaction energy. Interaction energy was determined as a difference of the dimer energy and the sum of monomer energies calculated in the dimer basis set, i.e., the basis set superposition error (BSSE) was properly eliminated.⁶ The intermolecular interaction potentials of complexes were calculated at the MP2 level using the Dunning's correlation consistent basis sets cc-pVDZ, cc-pVTZ, cc-pVQZ, aug-cc-pVDZ and aug-cc-pVTZ. The largest basis set used with the complex was cc-pVQZ which entailed 625 basis functions. In addition we carried out the coupled cluster calculations using single and double excitations with noniterative triple excitations (CCSD(T))⁷ to include electron correlation beyond MP2. Gaussian 98 series of programs were employed for the calculations.⁸

The MP2 interaction energy at the basis set limit, $E_{\text{MP2}(\text{limit})}$, was estimated by the methods of Helgaker et al.,⁹ Truhlar,¹⁰ and Feller.¹¹ In the method of Helgaker et al. the calculated MP2 interaction energies with the Dunning's correlated basis sets were fitted to the form $E_m = E_\infty + a/m^3$ where m is 2 for aug-cc-pVDZ and 3 for aug-cc-pVTZ. In the method of Truhlar the calculated SCF and MP2 correlation interaction energies for the cc-pVDZ and cc-pVTZ basis sets were fitted using the power-law expressions $E_n^{\text{HF}} = E_\infty^{\text{HF}} + An^{-\alpha}$ and $E_n^{\text{corrMP2}} = E_\infty^{\text{corrMP2}} + Bn^{-\beta}$, respectively, where n is 2 for the cc-pVDZ basis and 3 for the cc-pVTZ basis; α and β are parameters fitted by the author. In the method proposed by Feller the calculated MP2 interaction energies were fitted to the form $E_p = E_\infty + b/\exp(-cp)$ where p is 2 for cc-pVDZ, 3 for cc-pVTZ, and 4 for cc-pVQZ. The CCSD(T) interaction energy at the basis set limit $E_{\text{CCSD(T)}(\text{limit})}$ was estimated as the sum of

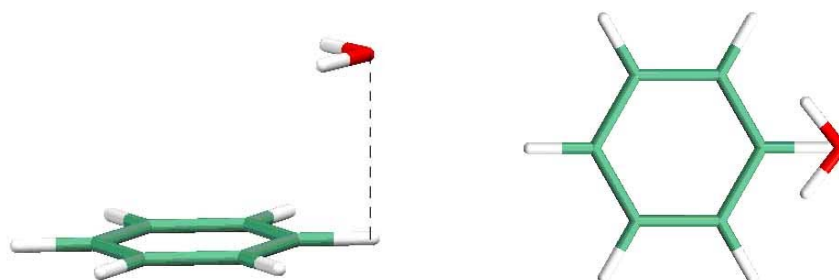
$E_{\text{MP2}(\text{limit})}$ and CCSD(T) correction ($\Delta\text{CCSD(T)}$). The $\Delta\text{CCSD(T)}$ term is the difference between the CCSD(T) and MP2 interaction energies using medium size basis set.¹² The calculated interaction energies and normal distances for the benzene-water complexes are summarized in Table S1. The intermolecular interaction potentials of the complexes A₁ and B₁ calculated at the MP2 and CCSD(T) levels using different basis sets are shown in Figs. S12 and S13.



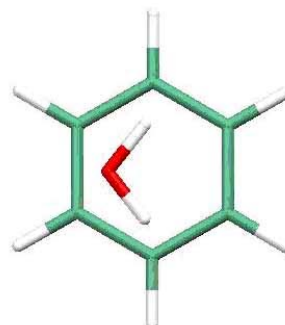
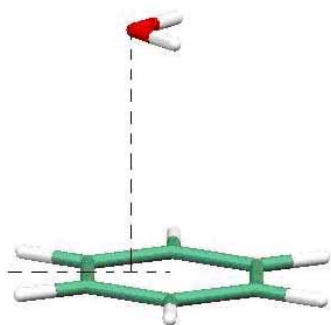
A₁ dimer $E_{\text{CCSD(T)}(\text{limit})} = -1.45$ kcal/mol $R = 3.1$ Å



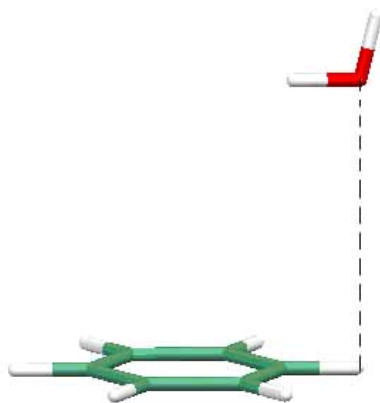
A₂ dimer $E_{\text{CCSD(T)}(\text{limit})} = -0.90$ kcal/mol $R = 3.6$ Å



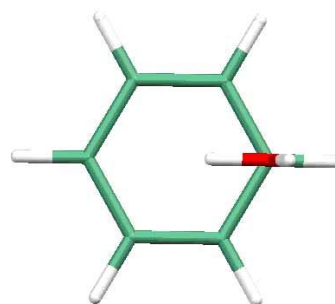
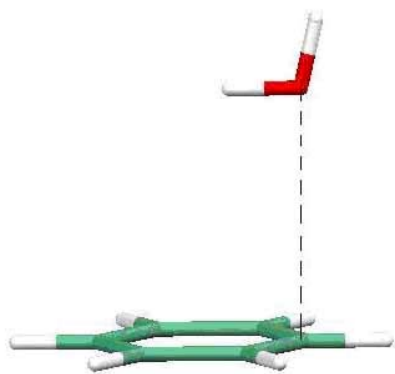
A₃ dimer $E_{\text{CCSD(T)}(\text{limit})} = -1.32$ kcal/mol $R = 3.3$ Å



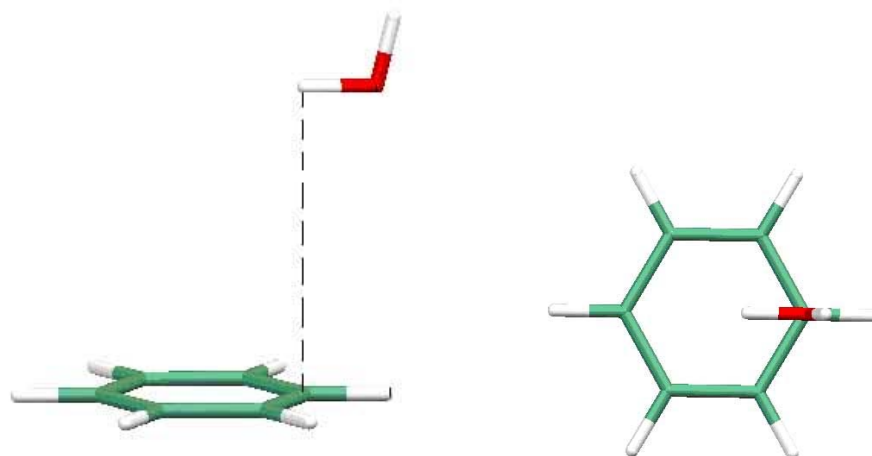
A₄ dimer $E_{\text{CCSD(T)}(\text{limit})} = -0.74$ kcal/mol $R = 3.6$ Å



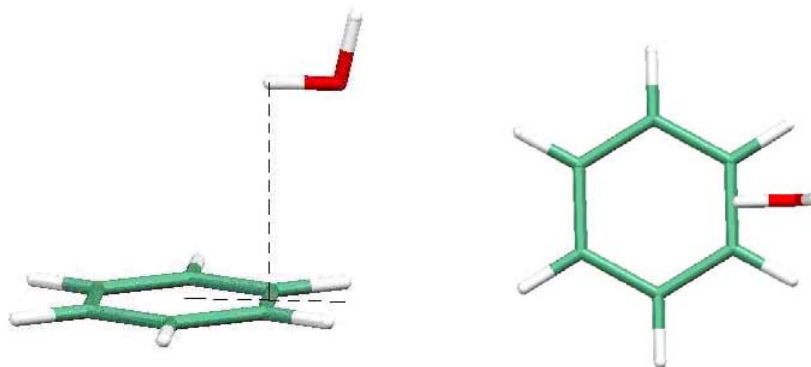
B₁ dimer $E_{\text{CCSD(T)}(\text{limit})} = -1.60$ kcal/mol $R = 3.0$ Å



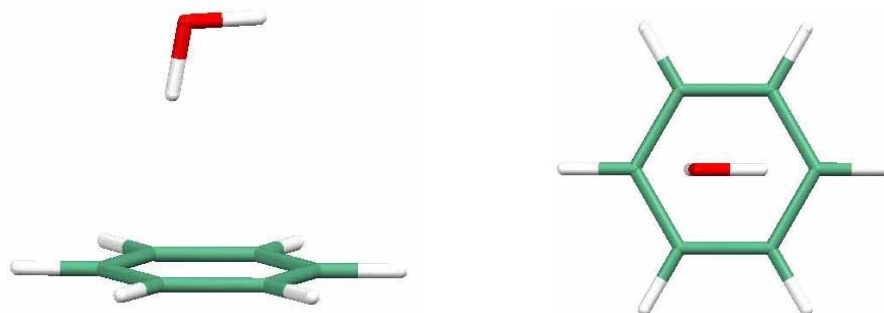
B₂ dimer $E_{\text{CCSD(T)}(\text{limit})} = -0.78$ kcal/mol $R = 3.4$ Å



B₃ dimer $E_{\text{CCSD(T)}(\text{limit})} = -1.52$ kcal/mol $R = 3.0$ Å



B₄ dimer $E_{\text{CCSD(T)}(\text{limit})} = -1.38$ kcal/mol $R = 3.1$ Å



C_{min} dimer $E_{\text{CCSD(T)}(\text{limit})} = -3.29$ kcal/mol $R = 3.3$ Å

Figure S11. The geometries of the water-benzene complexes considered in this work (side-on view (left); overhead view (right)). The values of R are optimal distances between the benzene-ring plane and the parallel OH bond (type B) or whole water molecule (type A).

Table S1. Calculated interaction energies of the benzene-water complexes A₁, A₂, A₃, A₄, B₁, B₂, B₃, B₄, C_{min}, and C^a

Method	N ^b	A ₁	A ₂	A ₃	A ₄	B ₁	B ₂	B ₂	B ₄	C _{min} ^c	C ^d
HF/cc-pVDZ	138	-0.28	0.40	-0.29	0.82	-0.04	0.84	0.17	0.26	-0.34	-0.29
HF/cc-pVTZ	322	-0.24	0.39	-0.25	0.76	-0.04	0.76	0.17	0.24	-0.38	-0.33
HF/cc-pVQZ	625	-0.23	0.37	-0.24	0.72	-0.04	0.73	0.17	0.24	-0.38	-0.33
HF/aug-cc-pVDZ	233	-0.29	0.35	-0.29	0.72	-0.12	0.70	0.09	0.17	-0.35	-0.30
HF/aug-cc-pVTZ	506	-0.23	0.35	-0.23	0.68	-0.04	0.71	0.17	0.24	-0.35	-0.30
MP2/cc-pVDZ	138	-0.70	-0.15	-0.68	0.15	-0.79	0.02	-0.66	-0.57	-1.72	-1.65
MP2/cc-pVTZ	322	-1.15	-0.57	-1.06	-0.34	-1.35	-0.50	-1.26	-1.14	-2.90	-2.81
MP2/cc-pVQZ	625	-1.34	-0.77	-1.23	-0.58	-1.60	-0.73	-1.52	-1.38	-3.34	-3.23
MP2/aug-cc-pVDZ	233	-1.31	-0.77	-1.21	-0.59	-1.59	-0.78	-1.50	-1.39	-2.90	-2.80
MP2/aug-cc-pVTZ	506	-1.43	-0.90	-1.31	-0.75	-1.71	-0.89	-1.64	-1.50	-3.42	-3.31
CCSD(T)/cc-pVDZ	138	-0.62	-0.07	-0.59	0.22	-0.63	0.13	-0.49	-0.40	-1.29	-1.24
CCSD(T)/aug-cc-pVDZ	233	-1.27	-0.70	-1.17	-0.51	-1.43	-0.62	-1.33	-1.21	-2.54	-2.46
E _{MP2(limit)} ^e		-1.49	-0.96	-1.36	-0.82	-1.76	-0.93	-1.70	-1.55	-3.65	-3.53
Δ _{CCSD(T)} ^f		0.04	0.07	0.04	0.08	0.16	0.15	0.18	0.17	0.36	0.33
E _{CCSD(T)(limit)} ^g		-1.45	-0.90	-1.32	-0.74	-1.60	-0.78	-1.52	-1.38	-3.29	-3.19
E _{MP2(limit)} ^h		-1.47	-0.86	-1.34	-0.65	-1.74	-0.83	-1.67	-1.53	-3.71	-3.59
Δ _{CCSD(T)} ⁱ		0.09	0.08	0.08	0.07	0.17	0.11	0.18	0.16	0.43	0.41
E _{CCSD(T)(limit)} ^g		-1.39	-0.78	-1.26	-0.58	-1.58	-0.72	-1.49	-1.36	-3.28	-3.18
E _{MP2(limit)} ^j		-1.49	-0.96	-1.36	-0.83	-1.79	-0.91	-1.73	-1.56	-3.61	-3.49
Δ _{CCSD(T)} ⁱ		0.09	0.08	0.08	0.07	0.17	0.11	0.18	0.16	0.43	0.41
E _{CCSD(T)(limit)} ^g		-1.40	-0.88	-1.28	-0.76	-1.62	-0.80	-1.55	-1.39	-3.18	-3.08

Table continued on the next page.

Table S1 continuation

Method	N ^b	A ₁	A ₂	A ₃	A ₄	B ₁	B ₂	B ₂	B ₄	C _{min} ^c	C ^d
R ^k		3.1	3.6	3.3	3.6	3.0	3.4	3.0	3.1	3.3	3.3
D ₀ ^f - E _{CCSD(T)(limit)} - ΔZPE										2.29 ^l	
D ₀ (exp.)										2.44±0.09 ^m 2.25±0.28 ⁿ 1.63–2.78 ^o	

^a Energies in kcal/mol. BSSE corrected interaction energies. The geometries of the complexes are shown in Fig. S11. The geometries of monomers (except the C_{min} and C dimers) were optimized at the MP2/cc-pVTZ level.

^b Number of contracted Gaussian basis function used for the calculation of the complex.

^c Single point calculations on the MP2/aug-cc-pVDZ fully optimized structure (Table S2). The BSSE was corrected by the counterpoise method.

^d Single point calculations on the monomer optimized distances and angles (MP2/cc-pVTZ); the distance C_r-O and the angle C_r-O-H₁ were retained from the MP2/aug-cc-pVDZ fully optimized structure C_{min}. The BSSE was corrected by the counterpoise method.

^e The estimated MP2 level interaction energy at the basis set limit by the method of Helgaker et al.⁹ employing the aug-cc-pVDZ and aug-cc-pVTZ basis sets.

^f CCSD(T) correction term. Difference between the E_{CCSD(T)} and E_{MP2} obtained with the aug-cc-pVDZ basis set.

^g The estimated CCSD(T) level interaction energy at the basis set limit according to the equation E_{CCSD(T)(limit)} = E_{MP2(limit)} + Δ_{CCSD(T)}.

^h The estimated MP2 level interaction energy at the basis set limit. Truhlar's method¹⁰ employing the cc-pVDZ and cc-pVTZ basis sets is used for the estimation.

ⁱ CCSD(T) correction term. Difference between the E_{CCSD(T)} and E_{MP2} obtained with the cc-pVDZ basis set.

^j The estimated MP2 level interaction energy at the basis set limit. Feller's method¹¹ employing the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets is used for the estimation.

^k Intermolecular distances in Å. See Fig. S11.

^l This work.

^m Experimental D₀ value from Courty et al.¹³

ⁿ Experimental D₀ value from Cheng et al.¹⁴

^o Experimental D₀ value from Gotch and Zwier¹⁵

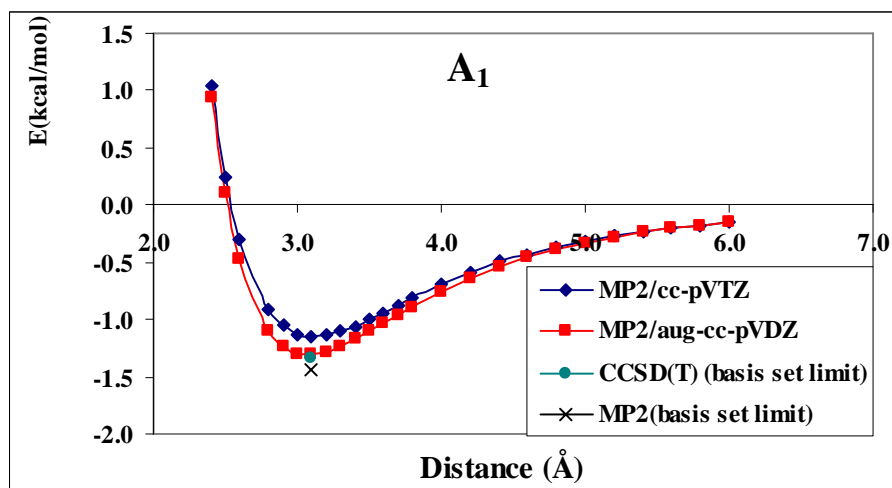


Figure S12. The MP2 interaction potentials of the benzene-water complex A_1 calculated with several basis sets. (The distance on the abscissa axis is the intermonomer distance (R), see Fig. S11)

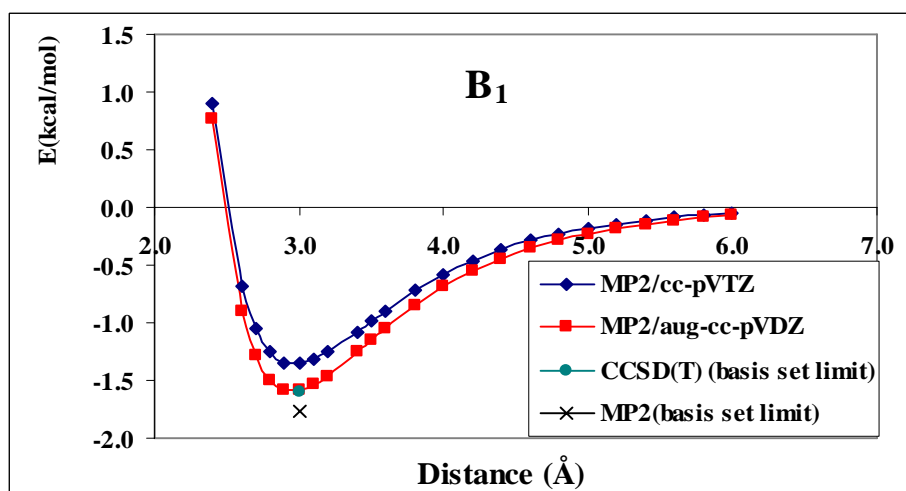


Figure S13. The MP2 interaction potentials of the benzene-water complex B_1 calculated with several basis sets. (The distance on the abscissa axis is the intermonomer distance (R), see Fig. S11)

Table S2. The geometry and the energy of the dimer C_{\min}

Dimer C_{\min}

Standard orientation:

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms)		
			X	Y	Z
1	6	0	0.614619	1.461373	0.000000
2	6	0	0.614135	0.757519	1.220058
3	6	0	0.614135	-0.650810	1.220535
4	6	0	0.614361	-1.355455	0.000000
5	6	0	0.614135	-0.650810	-1.220535
6	6	0	0.614135	0.757519	-1.220058
7	1	0	0.614275	2.555568	0.000000
8	1	0	0.609023	1.304134	2.167618
9	1	0	0.605059	-1.198077	2.167504
10	1	0	0.608840	-2.449358	0.000000
11	1	0	0.605059	-1.198077	-2.167504
12	1	0	0.609023	1.304134	-2.167618
13	8	0	-2.665522	-0.304238	0.000000
14	1	0	-2.731382	0.659315	0.000000
15	1	0	-1.708849	-0.459746	0.000000

EUMP2= -0.30780883384489E+03 (aug-cc-pVDZ)

III References

1. F.H. Allen, *Acta Cryst., Sect. B*, 2002, **58**, 380.
2. I.J. Bruno, J.C. Cole, P.R. Edgington, M. Kessler, C.F. Macrae, P. McCabe, J. Pearson, R. Taylor, *Acta Cryst., Sect. B.*, 2002, **58**, 389.
3. Z. Ciunik, G.R. Desiraju, *Chem. Commun.*, 2001, 703.
4. R.A. Kendall, T.H. Dunning, R.J. Harrison, *J.Chem.Phys.*, 1992, **96**, 6796.
5. (a) C. Møller, M.S. Plesset, *Phys. Rev.* 1934, **46**, 618; b) M.Head-Gordon, J.A. Pople, M.J. Frish, *Chem. Phys. Lett.*, 1988, **87**, 153.
6. S.F. Boys, F. Bernardi, *Mol. Phys.*, 1970, **19**, 553.
7. J. A. Pople, M. Head-Gordon, K. Raghavachari, *J. Chem. Phys.*, 1987, **87**, 5968.
8. M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Jr. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, J.L. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian 98, Revision A.6, Gaussian Inc., Pittsburgh, PA, 1998.
9. T. Helgaker, W. Klopper, H. Koch, J. Noga, *J. Chem. Phys.*, 1997, **106**, 9639.
10. D.G. Truhlar, *Chem.Phys.Lett.*, 1998, **294**, 45.
11. D. Feller, *J. Chem. Phys.*, 1992, **96**, 6104.
12. S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, *J.Am.Chem.Soc.*, 2002, **124**, 104.
13. A. Courty, M. Mons, I. Dimicoli, F. Puzzi, M.-P. Gageot, V. Brenner, P.d. Pujo, P. Millie, *J.Phys.Chem. A*, 1998, **102**, 6590.
14. B.-M. Cheng, J.R. Grover, E.A. Walters, *Chem.Phys.Lett.*, 1995, **232**, 364.
15. A.J. Gotch, T.S. Zwier, *J.Chem.Phys.*, 1992, **96**, 3388.