

July 23, 2005  
Final Author Version

## **Infinite-Basis Calculations of Binding Energies for the Hydrogen Bonded and Stacked Tetramers of Formic Acid and Formamide and Their Use for Validation of Hybrid DFT and Ab Initio Methods**

Yan Zhao and Donald G. Truhlar  
*Department of Chemistry and Supercomputing Institute,  
University of Minnesota, Minneapolis, MN 55455-0431*

### ***Abstract.***

Benchmark stabilization energies for planar H-bonded and stacked structures of formic acid tetramers and formamide tetramers were determined as the sum of the infinite basis set limit of MP2 energies and a CCSD(T) correction term evaluated with the 6-31G(0.25) basis set. The infinite basis (IB) set limit of MP2 energies was determined by two-point extrapolation using the aug-cc-pVXZ basis sets for X = D and T and separate extrapolation of the Hartree-Fock and correlation energies with new IB parameters for augmented basis sets determined here. Final stabilization energies (kcal/mol) for the tetramer studied are in the range of 4.6~6.7 kcal/mol and they were used as reference data to test thirteen density functionals. Among the tested DFT methods, PWB6K gives the best performance with an average error equal to only 30% of the average binding energy. In contrast the popular B3LYP functional has an average error of 85%. We recommend the PWB6K method for exploring the potential energy surfaces of organic complexes and clusters and supramolecular assemblies.

## 1. Introduction

Formic acid is an important constituent of clouds and fog<sup>1,2</sup> and is widely used in manufacturing. Formamide is the smallest organic compound containing the amide group, and it has been studied as a model compound for protein residues.<sup>3</sup> Studies of clusters of formic acid and formamide are of fundamental importance because they give insights into solvation, molecular recognition, crystal packing, protein folding, formation of aerosols, and supramolecular chemistry. Recently Roy and Thakkar used the popular B3LYP density functional to investigate the potential energy surfaces of tetramers<sup>4</sup> and pentamers<sup>5</sup> of formic acid, and they found 75 local minima for the tetramer and 205 local minima for the pentamer. Unfortunately, as pointed out by Wang<sup>6</sup> recently, popular density functional theory (DFT) methods, such as B3LYP, are not capable of quantitatively describing the dispersion-like interactions that play an important role in the stacking interaction of these clusters. The binding energies of the two stacked stationary points of the tetramer that were located in Roy and Thakkar's study were respectively 1.8 and 2.0 kcal/mol higher than their global minimum. Roy and Thakkar discussed the stacked structures and concluded that the CH $\cdots$ O forces in the planar clusters are stronger than the forces holding the dimers together in the stacked tetramers. In contrast, Wang<sup>6</sup> employed second-order Møller-Plesset perturbation theory (MP2)<sup>7</sup> method, and he found that a stacked tetramer is more stable than the planar hydrogen bonded ones. However, the largest basis set that Wang used is 6-311++G(3df,3pd) and is far from the complete basis set (CBS) limit. One goal of this paper is to obtain benchmark results for the interaction energies of the hydrogen bonded and stacked formic acid tetramer and formamide tetramer by employing separate extrapolation of Hartree-Fock and correlation energies,<sup>8</sup> which is called<sup>9</sup> the IB method, to obtain the infinite-basis MP2 limits and then add a correction for CCSD(T) effects. In particular, we employ the aug-cc-pVXZ basis sets with X = D and T for the MP2 part, and we employ the 6-31G\*(0.25) basis set to estimate the CCSD(T) correction term.

Because the previous parametrization of the IB<sup>9</sup> method was carried out for unaugmented basis sets, as a preliminary step in the present study, we determined new extrapolation parameters ( $\alpha$  and  $\beta$ ) for augmented basis sets.

For clusters the size of formic acid tetramer, it is too computationally demanding to use MP2 (even with a double zeta basis set) or any other reliable wave function theory (WFT) to investigate the whole PES of these clusters, because there are many stationary points on the PES. The second goal of the present study is to validate a number of low-cost DFT methods, and to determine if there are DFT methods that can describe hydrogen bonding and stacking interactions sufficiently well for practical studies of these small organic clusters.

The new parametrization of the IB method is presented in section 2. The computational details and DFT methods are described in section 3, and results and discussion are in section 4. Section 5 presents the concluding remarks.

## 2. Parametrization

To parametrize an IB method for the augmented basis sets, we use the HF and MP2 CBS data tabulated by Fast *et al.*<sup>9</sup> for eighteen molecules that have only hydrogen and first-row elements. These molecules are C<sub>2</sub>, CF, CH, CN, HCN, CO, HCO, FH, H<sub>2</sub>, H<sub>2</sub>CO, H<sub>2</sub>O, HCCH, N<sub>2</sub>, NH, NH<sub>2</sub>, NO, O<sub>2</sub>, and OH. The Hartree-Fock (HF) energies are extrapolated by

$$E^{HF}(n) = E_{\infty}^{HF} + A^{HF} n^{-\alpha} \quad (1)$$

and the MP2 correlation energies are extrapolated by

$$E^{cor}(n) = E_{\infty}^{cor} + A^{cor} n^{-\beta} \quad (2)$$

where  $n$  represent the highest angular momentum in an augmented correlation-consistent basis set;  $n = 2$  for the aug-cc-pVDZ basis, and  $n = 3$  for the aug-cc-pVTZ basis. The parameters  $\alpha$  and  $\beta$  are determined by fitting  $\alpha$  to the HF CBS energies and  $\beta$  to the MP2 CBS correlation energies. The value determined for  $\alpha$  is 4.93 and that for  $\beta$  is 2.13.

The new coefficients have been added to both of our computer programs<sup>10,11</sup> for carrying out multilevel calculations.

### 3. Computational Methods

All calculations were carried out using a locally modified *Gaussian03*<sup>12</sup> program. The geometries of all tetramers and dimers are optimized at the MP2/6-31+G(d,p) level of theory. We found the same two local minima for the stacked formic acid tetramer that Roy and Thakkar<sup>4</sup> found; their F456 and F448 correspond to (HCOOH)<sub>4</sub>- $\pi$ -1 and (HCOOH)<sub>4</sub>- $\pi$ -2 in this paper (Wang<sup>6</sup> reported only one stacked structure). We also located a stacked minimum for the formamide tetramer.

To extrapolate to the MP2 CBS limit, we used the above mentioned extrapolation scheme<sup>8</sup> with the aug-cc-pVDZ and aug-cc-pVTZ basis sets and the newly determined  $\alpha$  and  $\beta$ . The counterpoise correction<sup>13,14</sup> for basis set superposition error (BSSE) was not used in determining  $\alpha$  and  $\beta$ , but all WFT and DFT complexation energies were computed both with and without counterpoise corrections. All correlated WFT calculations have a frozen core on C, N, and O. The CBS extrapolation was applied to all calculated WFT energies of formic acid and formamide complexes (even for BSSE).

Hobza and coworkers have shown that difference between CCSD(T) and MP2 interaction energies ( $\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}}$ ) has a small basis set dependence,<sup>15</sup> so the CBS CCSD(T) interaction energy can be approximated as:

$$\Delta E^{\text{CCSD(T)}}_{\text{CBS}} = \Delta E^{\text{MP2}}_{\text{CBS}} + (\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}})_{\text{small basis}} \quad (3)$$

We use the 6-31G\*(0.25) basis set for the ( $\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}}$ ) term. The 6-31G\*(0.25) basis set<sup>16,17</sup> is the 6-31G\* basis with the standard d polarization functions (with exponent of 0.8) replaced by more diffuse ones (exponent of 0.25) in order to improve the description of the dispersion interaction. Hobza and Spomer<sup>17</sup> have shown that the 6-31G\*(0.25) basis set already yields a satisfactory value of the ( $\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}}$ ) difference for the hydrogen bonding and stacking interactions of nucleic base pairs.

In the present study, the dissociation energy  $D_e$  for the tetramers is defined as the equilibrium dissociation energy of the tetramer dissociated into two dimers at their equilibrium geometries.

To test the DFT methods for these tetramers, we used Eq. (3) to calculate benchmark reference data, as shown in Table 1, at MP2/6-31+G(d,p) geometries, which are given in the supporting information for both tetramers and dimers. The reference data was calculated by adding the MP2 IB energies (without counterpoise corrections) and the counterpoise corrected  $\Delta(\text{CCSD(T)}-\text{MP2})$  values. The reason that we use the MP2 IB energies without counterpoise corrections is because the extrapolation parameters  $\alpha$  and  $\beta$  are determined without counterpoise corrections.

We tested a number of hybrid DFT methods. In particular, we assessed B3LYP,<sup>18</sup> B97-1,<sup>19</sup> B97-2,<sup>20</sup> B98,<sup>21</sup> MPW1K,<sup>22</sup> MPW3LYP,<sup>23</sup> PBE1PBE,<sup>24-26</sup> MPW1B95,<sup>23</sup> MPWB1K,<sup>23</sup> PW6B95,<sup>27</sup> PWB6K,<sup>27</sup> TPSS1KCIS,<sup>28</sup> TPSSh,<sup>29</sup> and X3LYP.<sup>30</sup> Note that these functionals are in several cases built upon earlier functionals that should be credited for their role in the evolution of the final functionals, in particular the B88,<sup>31</sup> LYP,<sup>32</sup> mPW,<sup>33</sup> B95,<sup>34</sup> and PW91<sup>35</sup> functionals. (These papers in turn credit even earlier work.)

We used the 6-311+G(2df,2p)<sup>36 37</sup> basis set for DFT calculations.

## 4. Results and Discussion

Figure 1 shows the structures of the tetramers investigated in the present study.

### 4.1. Benchmark calculations

Table 1 gives the benchmark results for the interaction energies of the studied tetramers. Table 1 shows that Hartree-Fock (HF) theory gives reasonable descriptions for the hydrogen bonded tetramers, but fails seriously for the stacked ones. For the two stacked formic acid tetramers, HF theory behaves differently. HF theory gives more repulsive interaction energy for  $(\text{HCOOH})_{4-\pi-1}$  than for  $(\text{HCOOH})_{4-\pi-2}$ . This means that correlation contributes more stabilization to  $(\text{HCOOH})_{4-\pi-1}$  than to  $(\text{HCOOH})_{4-\pi-2}$ . The

$(\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}})$  corrections contribute 0.2~0.6 kcal/mol to the final dissociation energies in all five cases. There are only very small differences between the counterpoise corrected and the uncorrected value for the  $(\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}})$ .

Table 1 also shows that the stacked tetramers are more stable than the planar hydrogen bonded ones, and this confirms the results of Wang.<sup>6</sup> Even though these are gas-phase results, the substantial difference, 1.43 kcal/mol, between the dissociation energies of  $(\text{HCONH}_2)_4\text{-}\pi$  and  $(\text{HCONH}_2)_4\text{-HB}$  implies that stacking interactions may play an important role in protein folding.

The mean value of the five reference complexation energies in Table 1 is 5.5 kcal/mol, and in the next section we look for DFT methods whose mean error is much less than this.

#### 4.2. Assessment of DFT and ab initio methods

Table 2 gives the results obtained by both DFT and WFT methods. We also tabulated two error quantities in Table 2. One is mean unsigned error (MUE), and the other one is MMUE which is defined as:

$$\text{MMUE} = [\text{MUE}(\text{cp}) + \text{MUE}(\text{nocp})]/2 \quad (4)$$

where cp denotes counterpoise correction for BSSE, and nocp means without counterpoise correction. As indicated in section 2, all DFT calculations are carried out at the DFT/6-311+G(2df,2p)//MP2/6-31+G(d,p) level. The 6-311+G(2df,2p) basis set is a practical compromise between high accuracy and low cost.<sup>38</sup> The performance of various basis sets depends on the property examined (e.g., proton affinity, covalent bond strength, or various types of noncovalent complexation strengths) and the level of WFT (e.g., Hartree-Fock or MP2) or the functional of DFT. There is no systematic study available that tells us how close 6-311+G(2df,2p) is to the basis set limit for DFT calculations of  $\pi$  stacking complexation strength, and such a study would be beyond the scope of the present article. However, general experience with basis sets indicates that a correlation-balanced,<sup>39</sup> multiply polarized valence-triple-zeta basis set with polarization functions

two units of angular momentum beyond the normal valence shell and diffuse functions on heavy atoms (like 6-311+G(2df,2p)) is a good general-purpose basis set and is adequate for testing whether given functionals give realistic descriptions of various classes of interaction energies. We adopt this as a working hypothesis. The calculations presented here are one step toward confirming that working hypothesis, and we make a few additional remarks about the basis set at the end of this section.

Among the tested DFT methods, PWB6K gives the best performance for calculating complexation energies of these hydrogen bonded and stacked tetramers. Its MMUE is 1.69 kcal/mol, which is 30% of the mean complexation energy of 5.5 kcal/mol. Other DFT methods with relatively good performance are MPWB1K, PW6B95, and MPW1B95 with MMUE 2.7-3.1 kcal/mol. All four of these best performing DFT methods include kinetic energy density in the functional. Among the DFT methods without kinetic energy density, B97-1 gives the best performance, followed by MPW3LYP. The most popular DFT method, B3LYP, gives a MMUE of 4.70 kcal/mol, which translates to a percentage error of 85%, and this result shows that B3LYP is not suitable for exploration of the PES of organic clusters.

For the WFT methods in Table 2, HF gives errors similar to B3LYP. It is not surprising that the MP2/aug-cc-pVTZ method gives the overall best results, since it is a much more expensive method than the PWB6K/6-311+G(2df,2p) method. Unfortunately it is prohibitive to use MP2/aug-cc-pVTZ to explore the whole PES of the clusters the size of formic acid tetramers.

For all DFT methods, with the basis set 6-311+G(2df,2p), the BSSE is less for the hydrogen bonded tetramer (0.2 ~ 0.4 kcal/mol) than for the stacked ones (0.7 ~ 1.0 kcal/mol). From the standpoint of BSSE, 6-311+G(2df,2p) does not reach the basis set limit for these DFT calculations for the stacked complexes, but the BSSE is low enough that it does not dominate the error in the calculations. Note that MP2 calculations, even with the aug-cc-pVTZ basis set (which is much larger than 6-311+G(2df,2p)), have a

BSSE of  $\sim 0.7$  kcal/mol for the hydrogen bonded tetramers and a BSSE of  $\sim 1.4$  kcal/mol for the stacked ones. This is consistent with our general experience that, for a given basis set, DFT calculations have smaller BSSE than correlated WFT calculations.

## 5. Concluding remarks

Planar H-bonded and stacked structures of formic acid tetramers and formamide tetramers were optimized at the MP2 level using the 6-31+G(d,p) basis set. Stabilization energies for all structures were determined as the sum of the infinite basis set limit of MP2 energies and a  $(\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}})$  correction term evaluated with the 6-31G\*(0.25) basis set. The infinite basis set limit of the MP2 energies was determined by two-point extrapolation using the aug-cc-pVXZ basis sets for X = D and T and the economical IB extrapolation procedure presented previously, with new parameters determined here for augmented basis sets. The present stabilization energies for H-bonded and stacked formic acid and formamide tetramers are the most accurate and reliable values available, and they were used as reference data to test a number of DFT methods. Among the tested DFT methods, PWB6K performs very well for calculating the stabilization energies for both hydrogen bonded and stacked tetramers, with a mean unsigned error only 36% as large as the mean unsigned error of the popular B3LYP method. This example confirms the ability of the new PWB6K density functional to treat a wide range of chemical interactions, and we recommend the PWB6K method for exploring the potential energy surfaces of clusters of small organic compounds.

**Supporting Information Available:** The MP2/6-31+G(d,p) geometries for all tetramers and dimers are available via the Internet at <http://www.rsc.org/>.

## Acknowledgment

This work was supported in part by the U.S. Department of Energy, Office of Basic Energy Sciences.



- (1) Graedel, T. E.; Weschler, C. J. *Rev. Geophys. Space Phys.* **1981**, *19*, 505.
- (2) Aloisio, S.; Hintze, P. E.; Vaida, V. *J. Phys. Chem. A* **2002**, *106*, 363.
- (3) Torii, H.; Tasumi, M. *In. J. Quantum Chem.* **1998**, *70*, 241.
- (4) Roy, A. K.; Thakkar, A. J. *Chem. Phys. Lett.* **2004**, *393*, 347.
- (5) Roy, A. K.; Thakkar, A. J. *Chem. Phys.* **2005**, *312*, 119.
- (6) Wang, W. *Chem. Phys. Lett.* **2005**, *402*, 54.
- (7) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
- (8) Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *294*, 45.
- (9) Fast, P. L.; Sánchez, M. L.; Truhlar, D. G. *J. Chem. Phys.* **1999**, *111*, 2921.
- (10) Zhao, Y.; Rodgers, J. M.; Lynch, B. J.; Fast, P. L.; Chuang, Y.-Y.; Pu, J.; Truhlar, D. G. *MULTILEVEL-version 4.1*; University of Minnesota: Minneapolis, 2004.
- (11) Zhao, Y.; Truhlar, D. G. *MLGAUSS-version 1.1*; University of Minnesota: Minneapolis, 2004.
- (12) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr., T. V.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision C.01; Gaussian, Inc.: Pittsburgh PA, 2003.
- (13) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
- (14) Schwenke, D. W.; Truhlar, D. G. *J. Chem. Phys.* **1985**, *82*, 2418.
- (15) Jureka, P.; Hobza, P. C. P. L., 365, 89. *Chem. Phys. Lett.* **2002**, *365*, 89.
- (16) Hobza, P.; Sponer, J. *Chem. Rev.* **1999**, *99*, 3247.
- (17) Hobza, P.; Sponer, J. *J. Am. Chem. Soc.* **2002**, *124*, 11802.
- (18) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
- (19) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. *J. Chem. Phys.* **1998**, *109*, 6264.
- (20) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. *J. Chem. Phys.* **2001**, *115*, 9233.
- (21) Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624.
- (22) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811.
- (23) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908.
- (24) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett* **1996**, *77*, 3865.
- (25) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029.
- (26) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.
- (27) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656.

- (28) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43.
- (29) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (30) Xu, X.; Goddard, W. A. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 2673.
- (31) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- (32) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (33) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664.
- (34) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.
- (35) Perdew, J. P. In *Electronic Structure of Solids '91*; Ziesche, P., Eschig, H., Eds.; Akademie Verlag: Berlin, 1991; p 11.
- (36) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
- (37) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.
- (38) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384.
- (39) Almlöf, J.; Taylor, P. R. *J. Chem. Phys.* **1987**, *7*, 4070.

Table 1. Benchmark results (in kcal/mol) for the binding energies<sup>a</sup> of the tetramers.

Complexes	HF/IB	$\Delta$ MP2/IB <sup>b</sup>	$\Delta$ (CCSD(T)-MP2)	CCSD(T) CBS <sup>c</sup>
nocp				
(HCOOH) <sub>4</sub> -HB	2.37	1.73	0.39	4.50
(HCOOH) <sub>4</sub> - $\pi$ -1	-1.43	6.38	0.24	5.19
(HCOOH) <sub>4</sub> - $\pi$ -2	0.02	5.33	0.58	5.93
(HCONH <sub>2</sub> ) <sub>4</sub> -HB	2.68	2.08	0.45	5.21
(HCONH <sub>2</sub> ) <sub>4</sub> - $\pi$	0.06	6.08	0.53	6.68
cp				
(HCOOH) <sub>4</sub> -HB	2.30	1.45	0.48	4.22
(HCOOH) <sub>4</sub> - $\pi$ -1	-1.64	6.08	0.17	4.61
(HCOOH) <sub>4</sub> - $\pi$ -2	-0.22	5.05	0.56	5.39
(HCONH <sub>2</sub> ) <sub>4</sub> -HB	2.60	1.75	0.52	4.87
(HCONH <sub>2</sub> ) <sub>4</sub> - $\pi$	-0.14	5.82	0.57	6.24
reference data				
(HCOOH) <sub>4</sub> -HB	2.37	1.73	0.48	4.59
(HCOOH) <sub>4</sub> - $\pi$ -1	-1.43	6.38	0.17	5.13
(HCOOH) <sub>4</sub> - $\pi$ -2	0.02	5.33	0.56	5.91
(HCONH <sub>2</sub> ) <sub>4</sub> -HB	2.68	2.08	0.52	5.28
(HCONH <sub>2</sub> ) <sub>4</sub> - $\pi$	0.06	6.08	0.57	6.71

<sup>a</sup> The dissociation energy  $D_e$  for the tetramers is defined in the present study as the equilibrium dissociation energy of the tetramer dissociated into two dimers.

<sup>b</sup> The extrapolated ( $\Delta E_{\text{MP2}} - \Delta E_{\text{HF}}$ ) results.

<sup>c</sup> This column gives  $D_e$  estimated using Eq. (3).

<sup>d</sup> The reference data was calculated by adding the MP2 IB energy without counterpoise corrections and the counterpoise corrected  $\Delta$ (CCSD(T)-MP2) value.

Table 2. Binding energies and mean errors (in kcal/mol) for DFT and WFT methods.

Method	(HCOOH) <sub>4</sub> -HB		(HCOOH) <sub>4</sub> - $\pi$ -1		(HCOOH) <sub>4</sub> - $\pi$ -2		(HCONH <sub>2</sub> ) <sub>4</sub> -HB		(HCONH <sub>2</sub> ) <sub>4</sub> - $\pi$		MUE		MMUE
	nocp	cp	nocp	cp	nocp	cp	nocp	cp	nocp	cp	nocp	cp	
DFT													
B3LYP	2.62	2.32	-0.45	-1.40	0.40	-0.52	2.96	2.73	0.14	-0.52	4.39	5.00	4.70
B97-1	3.40	3.09	1.52	0.53	2.36	1.39	3.79	3.55	2.40	1.73	2.83	3.47	3.15
B98	3.13	2.81	0.90	-0.14	1.75	0.74	3.50	3.25	1.69	0.99	3.33	3.99	3.66
B97-2	2.13	1.79	-0.97	-2.09	-0.04	-1.14	2.44	2.19	-0.32	-1.08	4.88	5.59	5.23
MPW1B95	3.13	2.81	1.68	0.63	2.66	1.64	3.41	3.16	2.74	2.03	2.80	3.47	3.14
MPW1K	3.06	2.69	0.47	-0.74	1.57	0.40	3.45	3.17	1.42	0.59	3.53	4.30	3.92
MPW3LYP	3.47	3.16	1.27	0.26	2.12	1.13	3.81	3.56	1.92	1.22	3.01	3.66	3.33
MPWB1K	3.42	3.09	2.12	1.05	3.19	2.15	3.72	3.46	3.32	2.60	2.37	3.05	2.71
PBE1PBE	3.23	2.90	1.10	0.03	2.03	0.98	3.64	3.38	1.94	1.20	3.14	3.83	3.48
PW6B95	3.36	3.05	2.11	1.12	3.15	2.16	3.63	3.39	3.28	2.58	2.42	3.06	2.74
PWB6K	4.04	3.72	3.36	2.34	4.43	3.43	4.35	4.09	4.64	3.95	1.36	2.02	1.69
TPSSh	2.57	2.22	-0.18	-1.36	0.64	-0.51	2.48	1.99	-0.09	-1.61	4.44	5.38	4.91
TPSS1KCIS	2.67	2.34	0.00	-1.12	0.81	-0.28	2.87	2.41	0.53	-0.84	4.15	5.02	4.59
X3LYP	3.13	2.84	0.58	-0.38	1.44	0.50	3.49	3.25	1.23	0.56	3.55	4.17	3.86
WFT													
HF/aug-cc-pVDZ	2.90	2.42	0.38	-0.72	1.02	-0.17	3.17	2.78	0.72	-0.17	3.89	4.70	4.29

HF/aug-cc-pVTZ	2.45	2.31	-0.51	-0.83	0.15	-0.21	2.75	2.62	0.15	-0.14	4.53	4.78	4.65
MP2/aug-cc-pVDZ	4.63	3.26	6.63	3.57	7.21	3.77	5.17	3.80	7.78	4.45	0.80	1.76	1.28
MP2/aug-cc-pVTZ	4.18	3.51	5.28	3.98	5.84	4.37	4.79	4.07	6.64	5.17	0.24	1.31	0.77
MP2/6-31G*(0.25)	6.96	3.13	11.35	2.60	10.56	2.76	7.49	3.73	11.25	2.91	4.00	2.50	3.25
MP2/6-31G*(0.25)	6.96	3.13	11.35	2.60	10.56	2.76	7.49	3.73	11.25	2.91	4.00	2.50	3.25
MP3/6-31G*(0.25)	6.93	3.41	10.33	2.92	9.76	2.30	7.55	4.10	10.44	2.48	3.48	2.48	2.98
MP4(SDQ)/6-31G*(0.25)	7.06	3.34	10.80	3.12	10.22	2.62	7.61	3.96	10.71	2.66	3.76	2.39	3.07
CCSD(T)/6-31G*(0.25)	7.36	3.61	11.68	2.92	11.14	3.32	7.94	4.26	11.78	3.47	4.46	2.01	3.23

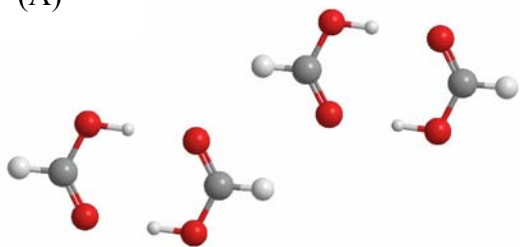
<sup>a</sup> MP2/6-31+G(d,p) geometries are used for all calculations, and the 6-311+G(2df,2p) basis set is used for all DFT calculation. In column headings, nocp denotes calculations without counterpoise correction, and cp denotes calculations with counterpoise correction. The reference data for the tetramers are taken from Table 1.

<sup>b</sup> MMUE=1/2[MUE(cp)+MUE(nocp)], and MUE denotes mean unsigned error (same as mean absolute deviation, MAD).

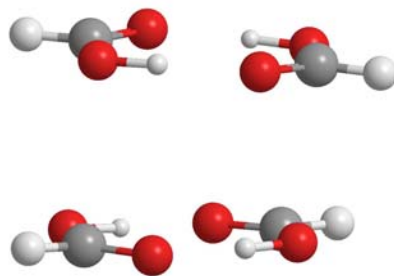
**Figure caption**

Figure 1. Geometries of the tetramers studied. (A)  $(\text{HCOOH})_4\text{-HB}$ , (B)  $(\text{HCOOH})_4\text{-}\pi\text{-1}$ , (C)  $(\text{HCOOH})_4\text{-}\pi\text{-2}$ , (D)  $(\text{HCONH}_2)_4\text{-HB}$ , (E)  $(\text{HCONH}_2)_4\text{-}\pi$ .

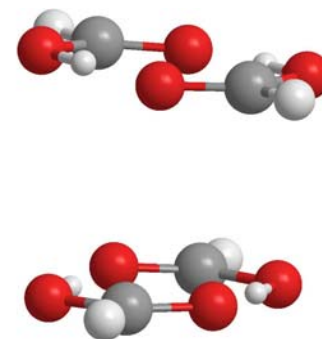
(A)



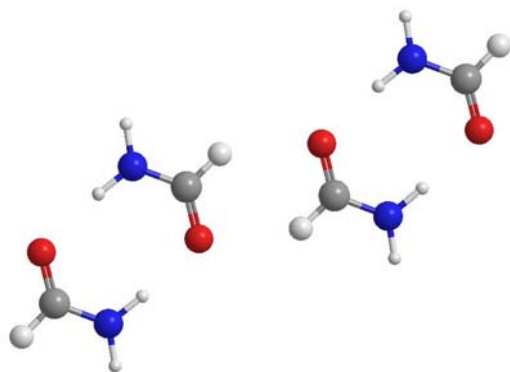
(B)



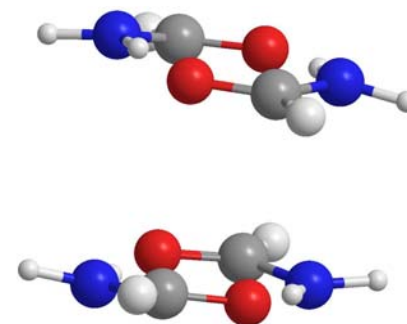
(C)



(D)



(E)



## Table of Contents (TOC) graphic

