An ab Initio Benchmark Study of Hydrogen Bonded Formamide Dimers

Jann A. Frey and Samuel Leutwyler*

Departement für Chemie und Biochemie, Universität Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

Received: July 25, 2006; In Final Form: September 7, 2006

The five singly and doubly hydrogen bonded dimers of formamide are calculated at the correlated level by using resolution of identity Møller−Plesset second-order perturbation theory (RIMP2) and the coupled cluster with singles, doubles, and perturbative triples [CCSD(T)] method. All structures are optimized with the Dunning aug-cc-pVTZ and aug-cc-pVQZ basis sets. The binding energies are extrapolated to the complete basis set (CBS) limit by using the aug-cc-pVXZ (X = D, T, Q) basis set series. The effect of extending the basis set to aug-cc-pV5Z on the geometries and binding energies is studied for the centrosymmetric doubly N−H···O bonded dimer FA1 and the doubly C−H···O bonded dimer FA5. The MP2 CBS limits range from −5.19 kcal/mol for FA5 to −14.80 kcal/mol for the FA1 dimer. The ΔCCSD(T) corrections to the MP2 CBS limit binding energies calculated with the 6-31+G(d,p), aug-cc-pVDZ, and aug-cc-pVTZ basis sets are mutually consistent to within ≤0.03 kcal/mol. The ΔCCSD(T) correction increases the binding energy of the C−H···O bonded FA5 dimer by 0.4 kcal/mol or ∼9% over the distance range ±0.5 Å relative to the potential minimum. This implies that the ubiquitous long-range C−H···O interactions in proteins are stronger than hitherto calculated.

I. Introduction

Amide groups give rise to omnipresent hydrogen bond donor/acceptor motifs in peptides and proteins and play a crucial role in defining their three-dimensional structures.[1−4] The accurate description of the N−H···O and C−H···O hydrogen bond geometries and interaction energies are important for understanding the biological functionality. *cis*-Amide groups constitute or contribute to the hydrogen-bonding sites of thymine, uracil, cytosine, and guanine, and the amide N−H···O═C hydrogen bonds are centrally involved in the structure and function of DNA and RNA. Since such biological molecules are still too large to be treated by highly accurate ab initio quantum chemical methods, highly accurate calculations on hydrogen-bonded amides are important for understanding the local characteristics of H-bonding interactions.

Formamide is the smallest amide and several ab initio quantum chemical studies have treated the cyclic formamide dimer, denoted FA1 in Figure 1.[5−8] Calculations have also been performed on the *N*-methylformamide dimer,[6] $(NMA)_2$, and recently on the formamide tetramer.[9] Vargas et al. have located five different minimum-energy structures of $(FA)_2$ using Møller−Plesset second-order perturbation theory (MP2).[6] These structures, which they have denoted FA1 to FA5 in order of decreasing stability,[6] are shown in Figure 1. FA1 has two N−H···O═C and FA5 two C−H···O═C hydrogen bonds; both dimers are centrosymmetric. FA2 is bound by an N−H···O═C and a C−H···O═C hydrogen bond and FA3 by a short N−H···O═C and a long C−H···N hydrogen bond, and FA4 is stabilized by a single N−H···O═C hydrogen bond. Vargas et al. optimized the structures with the aug-cc-pVTZ basis set and extrapolated the MP2 energies to the basis set limit, using the Dunning aug-cc-pVXZ series of basis sets.[10] However, they correlated all electrons, which is not appropriate for these basis sets.[11,12]

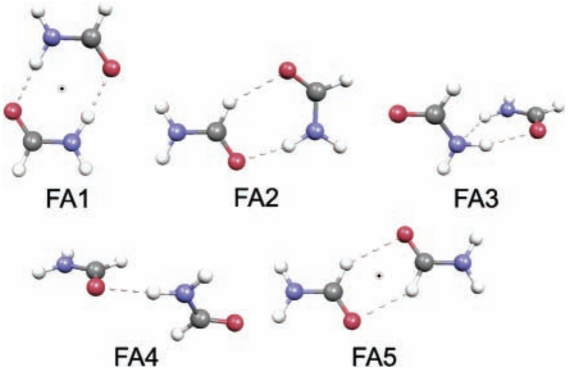

Figure 1. The five hydrogen bonded formamide dimers (nomenclature according to ref 6).

Among the high-level correlated ab initio approaches, the CCSD(T) (coupled-cluster with singles, doubles, and noniterative triple excitations) method has proven to give excellent results for structures and binding energies of molecules[13−15] and dimers.[16] However, the steep dependence of the CCSD(T) method on the number of basis functions ($\sim N^4_{\mathrm{func}}$) and occupied orbitals ($\sim N^3_{\mathrm{occ}}$) precludes the use of large basis sets for all but the smallest systems. The total correlation energy is known to converge only slowly with increasing basis set size. However, it has been noted that the difference of the MP2 and CCSD(T) stabilization energies, also known as the ΔCCSD(T) correction term, $\Delta\mathrm{CCSD(T)} = D_{\mathrm{e,CCSD(T)}} - D_{\mathrm{e,MP2}}$, becomes nearly constant for basis sets of quite moderate size.[7,16−19] This suggests that the CCSD(T) interaction energy at the basis set limit, $D^{\infty}_{\mathrm{e,CCSD(T)}}$ can be accurately estimated from the MP2 interaction energy at the complete basis set (CBS) limit $D^{\infty}_{\mathrm{e,MP2}}$, combined with the ΔCCSD(T) contribution calculated with a smaller basis set.[7,16−19]

Here, we calculate the interaction energy contributions to the FA1 to FA5 forms of the formamide dimer using the aug-cc-pVXZ (aVXZ) basis sets up to quintuple-ζ ($X = 2, 3, 4, 5$).

---

* Address correspondence to this author. E-mail: leutwyler@iac.unibe.ch.

10.1021/jp064730q CCC: $33.50

TABLE 1: Selected Experimental and Theoretical Structural Parameters of Formamide and of the Centrosymmetric Formamide Dimers FA1 and FA5[a]

	formamide, exptl		aug-cc-pVTZ			aug-cc-pVQZ		
	microwave[b]	el. diffraction[c]	formamide	FA1	FA5	formamide	FA1	FA5
r(C−N)	1.352(12)	1.367(4)	1.358	1.340	1.355	1.355	1.338	1.353
r(C═O)	1.219(12)	1.211(4)	1.218	1.232	1.224	1.215	1.229	1.221
r(C−H)	1.098(10)	1.12(fix)	1.100	1.098	1.098	1.099	1.097	1.097
r(N−H$_b$)	1.0016(30)	1.021(9)	1.001	1.025	1.006	1.004	1.023	1.004
r(N−H$_f$)	1.0015(30)	1.027(6)	1.003	1.004	1.003	1.002	1.002	1.002
θ(O═C−N)	124.7(3)	124.9(5)	124.7	125.1	123.9	124.6	125.1	123.9
θ(C−N−H$_b$)	118.5(5)		119.3	120.5	119.5	119.2	120.5	119.5
θ(C−N−H$_f$)	120.0(5)		121.1	119.3	120.8	121.1	119.4	120.8
θ(H−C−N)	113(2)		112.5	113.9	113.6	112.6	114.0	113.7
θ(N−H···O)				173.8			173.5	
θ(C−H···O)					141.5			140.9
R(N···O)				2.853	−		2.856	
R(C···O)					3.279			3.272
R(H···O)				1.833	2.348		1.837	2.346

[a] The values in columns 4−9 are calculated with the RIMP2 method and the aug-cc-pVTZ and aug-cc-pVQZ basis sets. Distances in Å, angles in deg. [b] Reference 35. [c] Reference 36.

The binding energies are extrapolated to the complete basis set (CBS) limit at the MP2 level for the two centrosymmetric dimers FA1 and FA5. These benchmark calculations allow the comparison of different schemes for obtaining CBS limit MP2 binding energies. We then evaluate the higher order contributions to the correlation energy at the CCSD(T) level using a series of basis sets up to aug-cc-pVTZ. These CCSD(T) binding energies allow us to assess whether the ΔCCSD(T) correction indeed approaches a limit value as a function of basis set size. The most efficient extrapolation scheme and ΔCCSD(T) correction method is then applied to the nonsymmetric formamide dimers[6] FA2, FA3, and FA4. Finally, we compare these benchmark energies to those calculated by density functional theory using the PW91, PBE, and B3LYP functionals.

Both π-stacked and hydrogen-bonded formamide tetramers have been optimized recently at the MP2/6-31+G(d,p) level.[20] The binding energies have been calculated on the optimized structures by using aVDZ and aVTZ basis sets and extrapolated to the CBS with use of a two-point extrapolation scheme.[21] The CBS limit energies have been compared to a series of newly designed density functionals and the PWB6K functional has been found to perform best;[20] however, the PW91 functional was not considered.

II. Methods

The formamide monomer and the FA1 to FA5 dimer structures were optimized by using the resolution of identity (RI) MP2 method and the augmented correlation-consistent polarized valence (aug-cc-pVXZ) basis sets,[10] using Turbomole 5.7.[22] The default auxiliary basis sets were used for density fitting; the core electrons were not correlated.[10−12] Geometries were converged to ≤0.00045 hartree/bohr. No symmetry constraints were applied; the FA1 and FA5 dimers converged to C_{2h} symmetry at all levels. Corrections for the basis set superposition error (BSSE) were calculated via the counterpoise (CP) correction scheme.[23]

To check the performance of the RIMP2 vs the MP2 method, the structures of formamide monomer and the FA1 and FA5 dimers were reoptimized by using the normal MP2 method and the same basis sets, using Gaussian 03.[24] The differences between the MP2 and RIMP2 methods for structures and binding energies are marginal, confirming earlier findings:[25] The bond lengths agree within ≤0.0001 Å and the binding energies within ≤0.003 kcal/mol. The computational time saving with the RIMP2 method is typically 1−2 orders of magnitude, relative to normal MP2 calculations.

The complete basis set (CBS) limit binding energies $D_{\mathrm{e,CBS}}$ were extrapolated by using both polynomial[26] (eq 1) and exponential[27] (eq 2) extrapolations for the CP-corrected and -uncorrected binding energies.

$$D_{\mathrm{e}}(X) = D_{\mathrm{e,CBS}} + \frac{A}{(X+1)^4} + \frac{B}{(X+1)^5} \qquad (1)$$

$$D_{\mathrm{e}}(X) = D_{\mathrm{e,CBS}} + A\cdot \mathrm{e}^{-(X-1)} + B\cdot \mathrm{e}^{-(X-1)^2} \qquad (2)$$

The MP2 limit binding energy $D^{\infty}_{\mathrm{e,MP2}}$ is taken to be the average of the extrapolated CP-corrected and -uncorrected binding energies, $D^{\mathrm{CPC}}_{\mathrm{e,CBS}}$ and $D_{\mathrm{e,CBS}}$, respectively. Half the difference between $D^{\mathrm{CPC}}_{\mathrm{e,CBS}}$ and $D_{\mathrm{e,CBS}}$ is taken as an estimate of the error of the extrapolation procedure. Tables with the MP2 optimized geometries and total energies are provided as Supporting Information.

Geometries and binding energies of the five formamide dimers were also calculated by using the three density functionals PW91,[28] PBE,[29] and B3LYP[30−32] with the 6-311++G-(d,p) basis set. Corrections for the basis set superposition error (BSSE) are calculated by applying the CP correction scheme.[23] Due to the accuracy aimed at in this work, maximal gradients of ≤0.000002 hartree/bohr were required for convergence.

III. Results and Discussion

A. Basis Set Extrapolation Tests. *Comparison of aVTZ to aVQZ optimized structures*: For formamide and the centrosymmetric dimers FA1 and FA5, optimized aVTZ as well as aVQZ structures were calculated. The important structure parameters are given in Table 1. Small but significant changes are noted for the aVQZ relative to the aVTZ structure: The intramolecular C−N and C═O bond lengths decrease by 0.002 to 0.003 Å and the C−H bond length by 0.001 Å. Use of the larger basis set increases the N···O hydrogen bond distance in FA1 by 0.003 Å, while the C···O distance of the FA5 dimer decreases by 0.007 Å.

CBS limit RIMP2 binding energies: For FA1 and FA5, the RIMP2 binding energies and counterpoise corrections were calculated with the aVXZ (X = D, T, Q, 5) basis sets at both the aVTZ and aVQZ minimum structures. We investigated the sensitivity of the CBS extrapolated binding energies to (i) the

TABLE 2: Summary of RIMP2 Complete Basis Set Extrapolated D_e^∞ (in kcal/mol) for the Formamide Dimers FA1 and FA5, Using Different Extrapolation Schemes

		exponential fit[a]		polynomial fit[b]	
structure	basis sets	FA1	FA5	FA1	FA5
aug-cc-pVTZ	D → Q	−14.80(9)	−5.19(4)	−14.79(7)	−5.17(3)
aug-cc-pVTZ	D → 5	−14.78(9)	−5.18(3)	−14.77(3)	−5.17(3)
aug-cc-pVQZ	D → Q	−14.80(9)	−5.19(4)	−14.79(7)	−5.17(3)
aug-cc-pVQZ	D → 5	−14.78(9)	−5.18(3)	−14.76(3)	−5.17(3)

[a] With eq 2.[27] [b] With eq 1.[26]

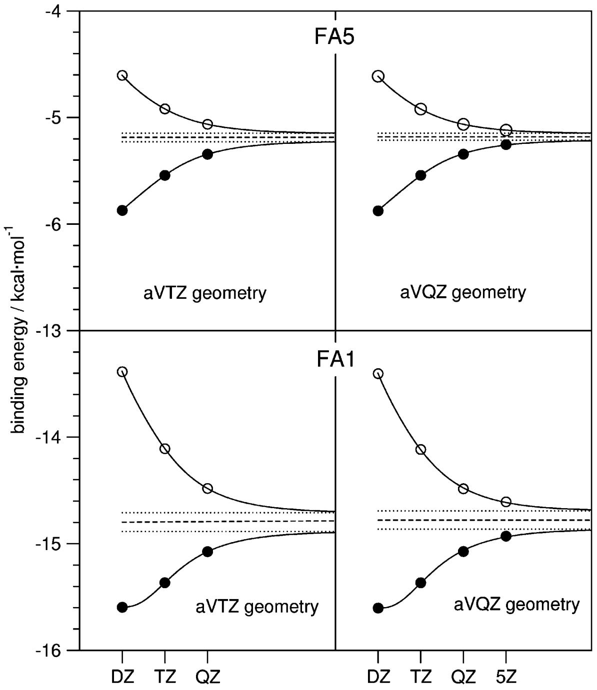

Figure 2. Complete basis set (CBS) binding energy extrapolations for the FA5 (top panel) and FA1 (lower panel) formamide dimers, see also Figure 1. Left panels: DZ → QZ extrapolations at the RIMP2/aug-cc-pVTZ optimized structures. Right panels: DZ → 5Z extrapolations at the RIMP2/aug-cc-pVQZ structures. The dotted lines mark the asymptotes of the counterpoise-corrected (○) and uncorrected (●) binding energies; the CBS limit $D_{e,MP2}^\infty$ is taken as the average and is indicated by a dashed line.

structure employed for the extrapolation, (ii) the length of the basis set extrapolations, and (iii) the extrapolation formula. The possible combinations give rise to eight different CBS extrapolation protocols. The resulting CBS limit binding energies $D_{e,MP2}^\infty$s are summarized in Table 2 and the exponential extrapolations are plotted in Figure 2. An additional table with all binding energies is given in the Supporting Information.

With respect to point (i) above, Table 2 shows that the CBS extrapolated binding energies of the FA1 and FA5 dimers depend only marginally (<0.005 kcal/mol) on whether the structure is optimized with the aVTZ or aVQZ basis set. From this we can also conclude that BSSE-corrected gradient optimizations with the larger basis set would have only a marginal influence on the binding energy.

With respect to point (ii), inclusion of the aV5Z binding energies into the extrapolation leads to slightly *smaller* CBS binding energies: these are 0.02 kcal/mol smaller for FA1, independent of the extrapolation scheme. For FA5 the decreases are 0.01 kcal/mol with exponential extrapolation and 0.004 kcal/mol with polynomial extrapolation.

With respect to point (iii), the choice of extrapolation function has a marginal influence on the binding energies of these dimers.

TABLE 3: ΔCCSD(T) Corrections[a] of the Binding Energies for the Hydrogen Bonded Formamide Dimers FA1 to FA5 (in kcal/mol, at the RIMP2/aug-cc-pVTZ minima)

basis set	FA1	FA2	FA3	FA4	FA5
6-31G(d,p)	0.15	−0.07	0.10	0.17	−0.24
6-31G*(0.25)	−0.37	−0.44	−0.17	−0.06	−0.49
6-31+G(d,p)	−0.09	−0.28	−0.14	−0.06	−0.42
aug-cc-pVDZ	−0.06	−0.26	−0.16	−0.06	−0.44
aug-cc-pVTZ	−0.10	−0.27	−0.13	−0.04	−0.42

[a] ΔCCSD(T) = $D_{e,CCSD(T)} - D_{e,MP2}$.

Polynomial extrapolation yields slightly (<0.14%) smaller binding energies and smaller uncertainties than the exponential extrapolation. However, since the polynomial extrapolation is often found to give counterpoise-corrected binding energies that are *larger* than the uncorrected ones, we have used the exponential extrapolation (eq 2) for the nonsymmetric isomers FA2, FA3, and FA4.

In summary, the eight different protocols lead to $D_{e,MP2}^\infty$ binding energies for FA1 that lie between 14.76 and 14.80 kcal/mol (0.3% relative difference). The $D_{e,MP2}^\infty$ of FA5 lie between 5.19 and 5.17 kcal/mol (0.4% relative difference). For all but one case, the protocol-dependent differences are smaller than the estimated uncertainties of the extrapolated binding energies. Specifically, the D → 5 extrapolation on aVQZ structures does not lead to significant changes in the extrapolated binding energies, nor to a significant decrease of the estimated error, when compared to the D → Q extrapolation on aVTZ structures—although the latter are computationally faster by at least 1 order of magnitude. This has important implications for the study of larger hydrogen bonded systems.[33,34]

B. The ΔCCSD(T) Correction Term. We have calculated the ΔCCSD(T) corrections of all five H-bonded formamide dimers, which span a large range of interaction strengths and H-bond geometries, using the 6-31G(d,p), 6-31G*(0.25), 6-31+G(d,p), aug-cc-pVDZ, and aug-cc-pVTZ basis sets. The CCSD(T) interaction energies were calculated at the MP2/aug-cc-pVTZ structures with Gaussian 03.[24] The correction terms ΔCCSD(T) = $D_{e,CCSD(T)} - D_{e,MP2}$ are given in Table 3.

Jurečka and Hobza have previously applied the basis sets 6-31G(d), 6-31G*(0.25), cc-pVDZ(0.25,0.15), aug-cc-pVDZ, and cc-pVTZ to the calculation of the ΔCCSD(T) correction term for the FA1 dimer.[7] Like them, we find the 6-31G(d,p) basis sets to give unreliable ΔCCSD(T) corrections. In contrast to their recommendation, we find that the 6-31G*(0.25) also gives unreliable results, ranging from 0.07 kcal/mol more stabilizing (for FA5) to 0.27 kcal/mol less stabilizing (for FA1), relative to the aug-cc-pVTZ result. By using the aug-cc-pVTZ result as a benchmark, the 6-31+G(d,p) and the aug-cc-pVDZ ΔCCSD(T) corrections are seen to behave in a satisfactory manner. All three basis sets give results that agree within ≤0.03 kcal/mol.

CCSD(T) intermolecular stretch potential: The intermolecular stretch potentials of FA1 and FA5 were calculated by displacing the rigid monomers along the H···O directions in steps of 0.1 Å over a range ±0.5 Å of the respective equilibrium distances. At each point, the SCF and RIMP2 interaction energy (both calculated with the aVTZ basis set) and the ΔCCSD(T)/6-31+G(d,p) correction term are displayed in Figure 3. As usual, the MP2 contribution becomes more strongly stabilizing toward smaller distances; hence the SCF+MP2 potential minimum is contracted relative to the SCF minimum.

For FA5, the ΔCCSD(T) correction leads to an additional stabilization of 0.42 kcal/mol, which is nearly distance independent over the range investigated. This leads to an additional

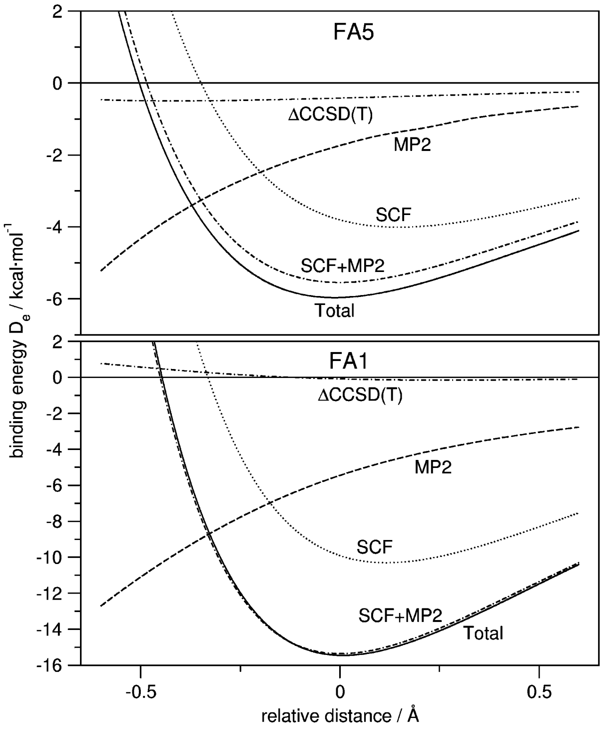

Figure 3. Hydrogen-bond stretching potentials for (a) FA5 and (b) FA1. The SCF, MP2, and ΔCCSD(T) contributions to the total binding energy at the aug-cc-pVTZ level are indicated. Note that the energy scale in the upper panel is magnified 2× relative to the lower panel. The ΔCCSD(T) correction is considerable for FA5 and shortens the H-bond distance for FA5.

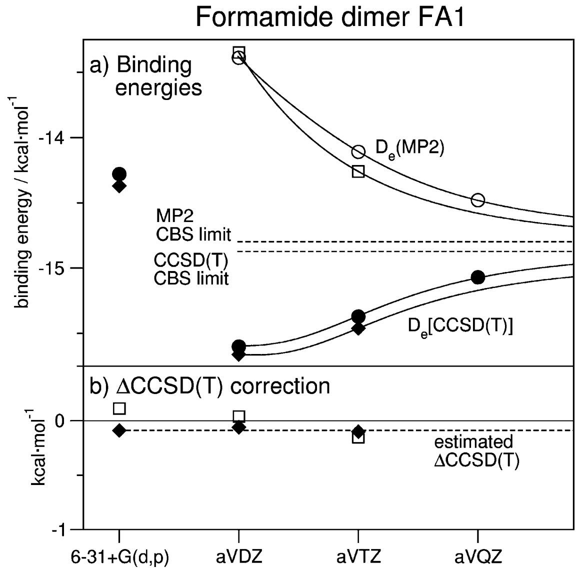

Figure 4. (a) Complete basis set binding energy extrapolations for the FA1 dimer at the RIMP2 and CCSD(T) levels. Open symbols denote CP-corrected and closed symbols CP-uncorrected values. (b) The CP-corrected and -uncorrected ΔCCSD(T) corrections as a function of basis set size.

slight contraction of the CCSD(T) H-bond length by ~0.01 Å relative to the SCF+MP2 minimum. For the more strongly bound FA1 dimer, the ΔCCSD(T) correction term is much smaller, in both absolute and relative terms.

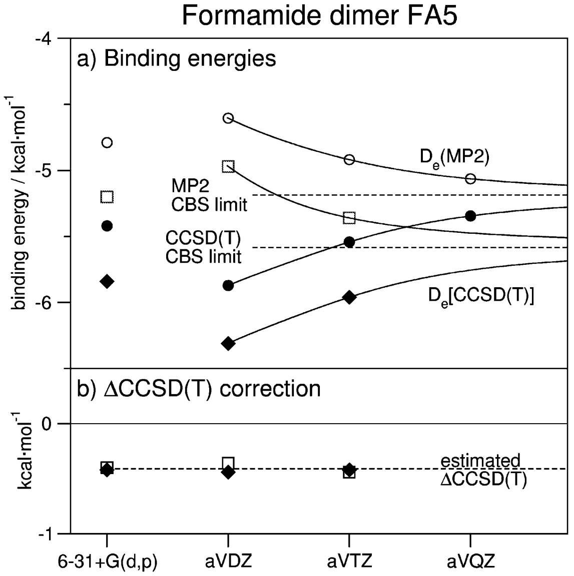

Figure 5. (a) Complete basis set binding energy extrapolations for the FA5 dimer at the RIMP2 and CCSD(T) levels. Open symbols denote CP-corrected, closed symbols CP-uncorrected values. (b) The CP-corrected and -uncorrected ΔCCSD(T) corrections as a function of basis set size.

TABLE 4: Calculated RIMP2, CCSD(T), and Density Functional H-Bond Binding Energies for the Five Formamide Dimers FA1−FA5

method	FA1	FA2	FA3	FA4	FA5
$D^{\infty}_{e,MP2}$ [a]	−14.80(9)	−10.01(6)	−7.68(5)	−6.97(5)	−5.19(4)
ΔCCSD(T)[b]	−0.10	−0.27	−0.13	−0.04	−0.42
$D^{\infty}_{e,MP2}$ + ΔCCSD(T)[a]	−14.90	−10.28	−7.81	−7.01	−5.61
B3LYP[c]	−13.21	−8.66	−6.21	−6.28	−4.04
PBE[c]	−14.41	−9.39	−7.04	−6.72	−4.32
PW91[c]	−15.00	−9.88	−7.44	−7.10	−4.72
BSSE					
B3LYP	0.42	0.39	0.25	0.27	0.15
PBE	0.47	0.32	0.31	0.26	0.15
PW91	0.53	0.38	0.26	0.30	0.18

[a] RIMP2/aug-cc-pVTZ. [b] At the RIMP2/aug-cc-pVTZ geometries. [c] 6-311++G(d,p) basis set.

C. RIMP2+ΔCCSD(T) Binding Energies of FA1−FA5. *Binding energies*: On the basis of the results presented in sections 3.A and 3.B, the binding energies of all five formamide dimers FA1−FA5 were calculated by using aVTZ structures and (D → Q) exponential extrapolations (eq 2). Figures 4 and 5 show the CBS extrapolations for the FA1 and FA5 dimers; the complete set of binding energies is collected in Table 4. The CBS extrapolated binding energies $D^{\infty}_{e,MP2}$ calculated in this work are about 3% larger than the core-correlated results of Vargas et al.[6]

Figures 4 and 5 also show the CCSD(T) binding energies and the ΔCCSD(T) corrections as a function of basis set size. Figure 5 emphasizes the large additional stabilization due to the CCSD(T) higher order correlation contributions for the FA5 dimer. Figures 4b and 5b show the near-constancy of the ΔCCSD(T) corrections as a function of the 6-31+G(d,p), aug-cc-pVDZ, and aug-cc-pVTZ basis sets, as discussed in the previous section.

For the centrosymmetric dimers FA1 and FA5, one can derive a single hydrogen bond energy of −7.40 kcal/mol for the N−H···O and −2.60 kcal/mol for the C−H···O hydrogen bond. FA2 exhibits both a N−H···O=C and a C−H···O=C hydrogen bond; its total binding energy of −10.28 kcal/mol is in good agreement with the sum of the FA1 and FA5 single H-bond energies (−10.0 kcal/mol). This suggests approximate transferability of these single hydrogen bond energies to larger formamide complexes.[5,9] As described above, the FA3 and FA4

TABLE 5: Calculated Intermolecular Structure Parameters and Intramolecular Bond Length Changes of upon Dimerization for the Formamide Dimers FA1−FA5, at the RIMP2/Aug-cc-pVTZ Level[a]

	formamide	FA1	FA2	FA3	FA4	FA5
			hydrogen bond parameters			
r(H···O)		1.834	1.869/2.270	1.968	1.933	2.346
r(N···O)		2.854	2.877	2.929	2.919	
r(C···O)			3.214			3.272
θ(N−H···O)		173.8	168.4	156.7	163.6	
θ(C−H···O)			142.9			140.940
			H-donor formamide			
r(N−H$_b$)	1.006	+0.018	+0.015	+0.010	+0.008	−0.000
r(C−N)	1.358	−0.018	−0.011	+0.004	−0.006	−0.003
r(C═O)	1.218	+0.014	+0.008	+0.001	+0.004	+0.006
τ(O−C−N−H$_b$)	0	0	0	−12.9	−3.8	0
τ(H−C−N−H$_b$)	0	0	0	14.9	3.3	0
			H-acceptor formamide			
r(N−H$_b$)′	1.006		0.000	+0.003	0.000	
r(C−N)′	1.358		−0.009	−0.008	−0.010	
r(C═O)′	1.218		+0.011	+0.008	+0.006	

[a] Distances in Å, angles in deg.

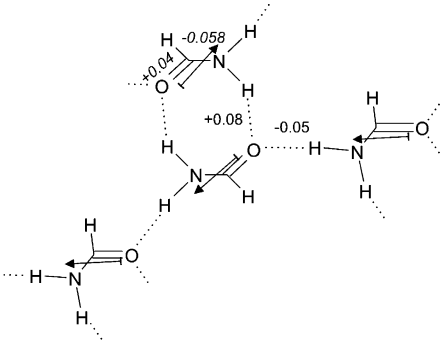

Figure 6. Hydrogen bond connectivity for the structure of crystalline formamide. Note the existence of H-bonded chains with FA4-type interactions, which are interconnected by FA1-type hydrogen bonds. The numbers indicate the change between the RIMP2 gas-phase and solid-state hydrogen bond lengths.

dimers exhibit a single strong N−H···O═C hydrogen bond, differing mainly with respect to the carbonyl oxygen lone pair to which the N−H group is H-bonded, see Figure 1. The $D^{\infty}_{\mathrm{e,MP2}}$ binding energy of FA3 is −7.68 kcal/mol, about 0.3 kcal larger than the N−H···O═C single hydrogen bond established above. This is due to an additional weak N−H···N hydrogen bond, see Figure 1. The single hydrogen bond of FA4 is 6.97 kcal/mol, which is about 95% of the single N−H···O hydrogen bond energy of FA1.

D. Structures and Structural Changes upon H-Bond Formation. Selected geometry parameters of the formamide monomer and the five dimers and their respective changes upon dimerization are given in Tables 1 and 5. The gas-phase bond lengths of the formamide monomer have been determined by microwave spectroscopy[35] and electron diffraction[36] and are given in Table 1. The interpretations of the experimental data[35,36] yield distances that differ by up to 0.015 Å and the in-plane bond angles up to 0.3°. We find that the calculated aug-cc-pVQZ equilibrium bond lengths are generally in better agreement with the microwave bond lengths, being 0.001 to 0.004 Å longer than the microwave values.

Intramolecular geometry changes: As expected, the N−H bond length increases upon hydrogen bonding: The largest increase is +0.018 Å for the FA1 dimer, diminishing to +0.007 Å for FA4. Note that the increases for FA1 to FA4 are *not* proportional to the hydrogen bond strengths, which have been shown to be very similar above. The increases are better correlated with the presence or absence of a second in-plane hydrogen bond, i.e., with synergic H-bond reinforcement.

In parallel, hydrogen bonding leads to longer C═O and shorter C−N intramolecular bonds. These changes reflect the incipient rearrangement of the intramolecular bonding from the keto toward the enol form induced by hydrogen bonding. Between the gas-phase formamide monomer and the formamide crystal, the C═O bond lengthens by +0.037 Å and the C−N bond shortens by 0.058 Å, see also the discussion below.

The suggested presence of a weak long-range N−H···N hydrogen bond in FA3 is supported by the increase of the N−H$_b$′ bond length by +0.004 Å. Also, the H-donor amine group in FA3 exhibits an out-of-plane or pyramidalization distortion of 18° that reflects the involvement of the amine lone-pair in this long-range N−H···N interaction. A slight pyramidalization of the amine group by 6° is also predicted for the FA4 dimer.

H-bond geometries: The N−H···O═C hydrogen bond distance in the most strongly H-bonded FA1 dimer is 1.834 Å. In the FA2 dimer, the N···O hydrogen bond distance increases by only 0.04 Å, relative to FA1. For the singly N−H···O═C bonded dimers FA3 and FA4 the N···O distance increase by 0.13 and 0.10 Å relative to the FA1 dimer. These increases are correlated with the absence of a second in-plane hydrogen bond, i.e., with the absence of synergic H-bond reinforcement. Note, however, the increases are *not* paralleled by decreases of the H-bond strengths (see previous section).

The C−H···O hydrogen bond distances in FA2 and FA5 are 2.270 and 2.346 Å, respectively; these are typical for C−H···O hydrogen bonds.[2] The fact that the C−H···O distance in FA2 is 0.076 Å shorter than that in FA5 is also traced to synergic reinforcement of the weak C−H···O by the strong N−H···O hydrogen bond.

E. Comparison to the Crystal Structure of Formamide. The formamide crystal structure[37] can be viewed as being built up from N−H···O hydrogen bonded zigzag chains.[37] The N−H···O bonds along the chains are analogous to that of the FA4 dimer; in the crystal, the N···O distances are 2.88 Å. The formamide chains are further interconnected sideways by double H-bonds between the formamide subunits, analogous to the H-bonds of the FA1 dimer. A schematic of the hydrogen bond topology and distances in the formamide crystal is shown in Figure 6. Interestingly, the FA1-type N···O distance in the crystal is 2.935 Å, which is 0.08 Å *longer* than that of FA1; in contrast the N···O distances along the chain are about 0.05 Å

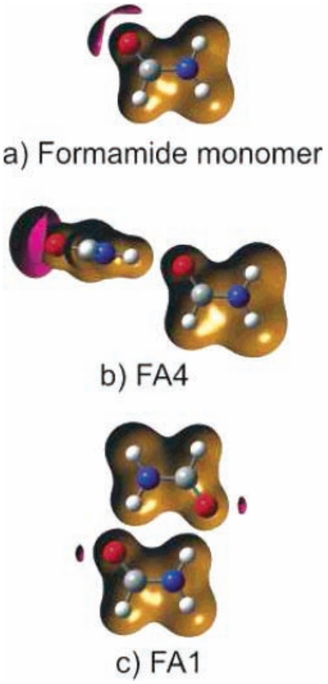

Figure 7. Isosurfaces of the molecular electrostatic potential at an energy of 50.2 kcal/mol of (a) formamide, (b) the linear formamide dimer FA4, and (c) the cyclic formamide dimer FA1. Purple surfaces show the negative regions while the brown surfaces show the positive regions.

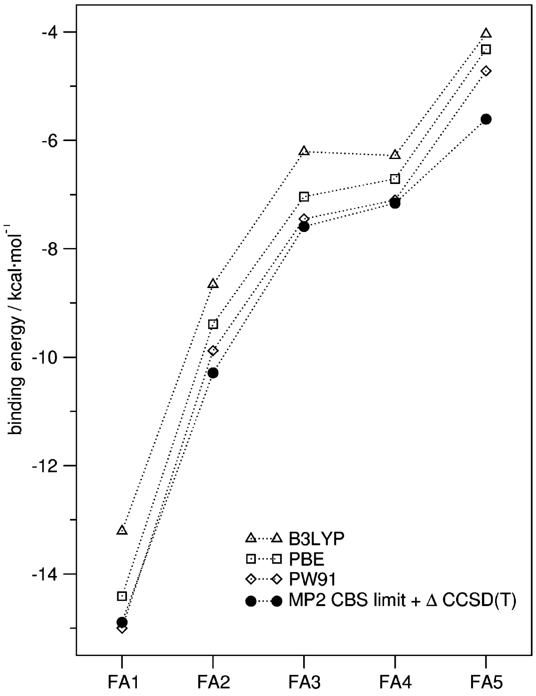

Figure 8. Comparison of the binding energies of the FA1-FA5 dimers calculated with the B3LYP, PBE, and PW91 density functional (CP-uncorrected). The RIMP2+ΔCCSD(T) benchmark values are also given (●).

shorter than the 2.93 H-bond length of FA4. This implies that the local H-bond strengths in the crystal are inverted relative to the gas phase (although the relationship between the H-bond distance and strength is not a simple one, as pointed out above).

The shortening of FA4-type hydrogen bonds along the chain is due to their cooperative reinforcement. This is predicted by a consideration of dipole−induced dipole interactions. These are expected to be large, since the local dipole moments of the formamide units are nearly collinear along the chain, as indicated in Figure 6. In contrast, the formamide monomer dipoles are in a repulsive orientation in the FA1-type dimer, implying that the inductive (dipole−induced dipole) interactions should also be weaker. Such effects have also been discussed in connection with early SCF/STO-3G calculations of chain-type and cyclic formamide clusters.[38]

A different rationalization of the inversion of the FA1 and FA4 hydrogen bond distances can be given in terms of the calculated molecular electrostatic potential (MEP) of formamide and the FA4 and FA1 dimers. Figure 7 shows the calculated MEP isosurfaces at a value of ±50.2 kcal/mol. These are the loci where the interaction energy of a proton is ±50.2 kcal/mol; the volumes enclosed in purple correspond to the negative MEP (stabilizing) regions. Figure 7a shows that there are two maximal interaction regions in the molecular plane of formamide. These correspond approximately to the two lone-pair regions of the carbonyl oxygen. Figure 7b shows that formation of the FA4 dimer strongly increases the spatial extent associated with the carbonyl oxygen lone pairs of the formamide molecule on the left. This will make the FA4-type dimer a stronger electron donor toward the next FA4-type N−H···O hydrogen bond.

As a comparison of parts a and b of Figure 7 shows, the FA4-type and FA1-type hydrogen bonds are mutually anticooperative: Formation of the FA4 dimer makes the oxygen lone pair binding site for a FA1-type “sideways” H-bond disappear, relative to the monomer. Parts a and b of Figure 7 show that formation of the FA1 dimer also makes it a weaker donor with respect to formation of a FA4 type H bond. However, the oxygen lone pair is small but still visible, showing that the anticooperativity is smaller when an FA1 dimer is formed, compared to the FA4 dimer.

F. Binding Energies by Density Functional Theory. Binding energies were also calculated by using the B3LYP, PBE, and PW91 functionals with the 6-311++G(d,p) basis set at the respective optimized geometries. Since the DFT binding energies are found to underestimate the CCSD(T) corrected RIMP2 binding energies, we give only the CP-uncorrected DFT binding energies in Table 4.

The DFT results are compared to the ab initio results in Figure 8. In every case, the best agreement is observed for the PW91 functional. The PW91/6-311++G(d,p) binding energies reproduce the RIMP2+ΔCCSD(T) values for FA1 and FA4 within 0.1 kcal/mol. For FA2 and FA3, the differences are still only 0.4 kcal/mol, while the C−H···O bonds of FA5 are underestimated by 0.89 kcal/mol or 16%. The second best agreement is observed for the PBE functional, with differences between 0.6 and 1.4 kcal/mol, followed by the B3LYP energies, which differ by 1.0−1.7 kcal/mol.

We have reported earlier that the PW91 functional can yield hydrogen bond binding energies close to those of high-level methods for the FA1 dimer.[8,39] It seems that this carries over to all the dimers that involve N−H···O hydrogen bonds. However, the strength of the C−H···O═C hydrogen bonds is underestimated even by the best functional.

IV. Conclusions

Benchmark hydrogen-bond binding energies for the five formamide dimers are calculated at the complete basis set MP2 level, including the ΔCCSD(T) correction. The best MP2+ΔCCSD(T) binding energies are −14.9 kcal/mol for FA1, −10.3 kcal/mol for FA2, −7.8 kcal/mol for FA3, −7.0 kcal/mol for FA4, and −5.6 kcal/mol for FA5. For both the

N—H···O and C—H···O bonded FA1 and FA5 dimers, the performance of DZ → QZ complete basis set extrapolations on optimized aVTZ structures is similar to those of the much more expensive DZ → 5Z extrapolations on aVQZ structures.

In agreement with previous work,[7,16−19] the ΔCCSD(T) correction term is found to converge quite rapidly with increasing basis set size. However, we find that at least a double-ζ plus diffuse function basis set [such as 6-31+G(d,p)] is necessary to obtain ΔCCSD(T) corrections that agree with the aug-cc-pVTZ corrections within 0.02 kcal/mol. Inclusion of the ΔCCSD(T) correction to the RIMP2 CBS limit increases the H-bond energies of the weakly C−H···O hydrogen bonded FA5 dimer by nearly 10%. This result is highly significant in view of the ubiquity of C—H···O═C hydrogen bonds in biological systems.[4,40] It suggests that long-range stabilization of C−H···O hydrogen bonds by higher order correlation contributions may be larger than expected.

For the N—H···O bound FA1, FA2, FA3, and FA4 complexes, the PW91 functional reproduces the binding energies to within 1% to 4%, i.e, with sufficient accuracy for applications to larger biological systems where CCSD(T) calculations would not be viable. However, the stabilizing contribution of the FA5-type C—H···O hydrogen bonds is underestimated by about 20%.

A comparison with the N···O hydrogen bond distances in crystalline formamide shows that a reversal of H-bond lengths occurs with respect to the gas-phase dimer: the slightly longer (and weaker) gas-phase FA4 hydrogen bond becomes shorter and the shorter gas-phase FA1 H-bond becomes longer in the crystal. This is attributed to mutually anticooperative effects of forming two simultaneous H-bonds at the two lone pairs of the carbonyl oxygen atom.

Acknowledgment. This work was supported by the Schweiz. Nationalfonds (project nos. 2000-68081.02 and 200020-105490).

Supporting Information Available: Tables of comparative MP2 and R1-MP2 calculations on formamide and formamide dimer FA1, selected structural parameters of the minimum energy structures of formamide and formamide dimer FA1, comparison of the binding energies at the complete basis set limit of the two isomers FA1 and FA3 of the formamide dimer, Cartesian coordinates of the R1-MP2 aug-cc-pVTZ structure of formamide and formamide dimers FA1, FA2, FA3, FA4, and FA5, and Cartesian corrdinates of the R1-MP2 aug-cc-pVQZ structures of formamide and formamide dimers FA1 and FA5. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer: Berlin, Germany, 1991.

(2) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond*; Oxford Science Publications: New York, 1999.

(3) Gu, Y.; Kar, T.; Scheiner, S. *J. Am. Chem. Soc.* **1999**, *121*, 9411.

(4) Scheiner, S. *J. Phys. Chem. B* **2005**, *109*, 16132.

(5) Torii, H.; Tasumi, M. *Int. J. Quantum Chem.* **1997**, *10*, 241.

(6) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 4963.

(7) Jurečka, P.; Hobza, P. *Chem. Phys. Lett.* **2002**, *365*, 89.

(8) Frey, J. A.; Leutwyler, S. *Chimia* **2005**, *59*, 511.

(9) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 6624.

(10) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(11) Frey, J. A.; Leutwyler, S. *J. Phys. Chem. A* **2005**, *109*, 6990.

(12) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2005**, *109*, 6991.

(13) Gauss, J.; Cremer, D.; Stanton, J. F. *J. Phys. Chem. A* **2000**, *104*, 1319.

(14) Bak, K. L.; Gauss, J.; Jorgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. *J. Chem. Phys.* **2001**, *114*, 6548.

(15) Pawlowski, F.; Jorgensen, P.; Olsen, J.; Hegelund, F.; Helgaker, T.; Gauss, J.; Bak, K. L.; Stanton, J. F. *J. Chem. Phys.* **2002**, *116*, 6482.

(16) Klopper, W.; Lüthi, H. P. *Mol. Phys.* **1999**, *96*, 559.

(17) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Phys. Chem. A* **1999**, *103*, 8265.

(18) Tsuzuki, S.; Lüthi, H. *J. Chem. Phys.* **2001**, *114*, 3949.

(19) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2002**, *124*, 104.

(20) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656.

(21) Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *294*, 45.

(22) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Klömel, C. *Chem. Phys. Lett.* **1989**, *162*, 165 (current version: see http://www.turbomole.de).

(23) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(24) Frisch, M. J.; Trucks, G. W.; Schlegel H. B.; et al. *GAUSSIAN 03*, Rev. A.1; Gaussian Inc., Pittsburgh, PA, 2003.

(25) Jurečka, P.; Nachtigall, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 4578.

(26) Klopper, W. *J. Chem. Phys.* **1995**, *102*, 6168.

(27) Peterson, K. A.; Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1994**, *100*, 7410.

(28) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.

(29) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(30) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(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) Frey, J. A.; Leist, R.; Leutwyler, S. *J. Phys. Chem. A* **2006**, *110*, 4188.

(34) Müller, A.; Frey, J. A.; Losada, M.; Leutwyler, S. *J. Phys. Chem. A* To be submitted.

(35) Hirota, E.; Sugisaki, R.; Nielsen, C.; Soerensen, G. *J. Mol. Struct.* **1974**, *49*, 251.

(36) Kitano, M.; Kuchitsu, K. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 67.

(37) Ladell, J.; Post, B. *Acta Crystallogr.* **1954**, *7*, 559.

(38) Hinton, J. F.; Harpool, R. D. *J. Am. Chem. Soc.* **1977**, *99*, 349.

(39) Müller, A.; Losada, M.; Leutwyler, S. *J. Phys. Chem. A* **2004**, *108*, 157.

(40) Scheiner, S.; Kar, T. *J. Phys. Chem. B* **2005**, *109*, 3681.