

## Energy analysis on small to medium sized H-bonded complexes

G. Dive, D. Dehareng, and J. M. Ghuysen

Centre d'Ingénierie des Protéines, Université de Liège, Institut de Chimie, B6, B-4000 Sart Tilman (Liège 1), Belgium

Received March 18, 1992/Accepted November 17, 1992

**Summary.** Dimers (water-methanol, guanine-cytosine) as well as trimers (methanol-water-imidazole, formamide-methylformate-formamide), are studied as H-bonded complexes of increasing complexity. All the investigated conformations have been fully optimized. In particular, it is the first time that all the intra- and intermolecular parameters of the guanine-cytosine complex are left variable. In minimal basis sets, the planar conformation has been found to be a first-order critical point. The minimal basis set MINI-1 has been adapted to provide nearly planar amides. The stability of the complexes is accounted for by four energy components of the same order: the first-order term (electrostatic + exchange), the polarization, the charge transfer and the correlation terms. In the case of the studied trimers, the energy components, apart from the electrostatic one, have been found to be nearly additive.

**Key words:** AM1-*ab initio* geometry optimization – H-bonded complexes – Basis set influence – MINI-1 new calibration – Guanine-cytosine – Energy decomposition

### 1. Introduction

Understanding the nature of weak, noncovalent interactions [1] has been of crucial interest for many years yet. These have been extensively studied for small systems and a very detailed review has been made by Chalasinski and Gutowski [2].

In particular, H-bonding plays a key role in the stability and folding of proteins [3–6] and has also been a matter of large discussions concerning DNA structure [6–8], as for instance Watson–Crick interactions [8–15] or stacking interactions [15–16]. A very well-documented historical overview of H-bonding can be found in the review by Reed et al. [17] as well as a comparison between their formalism, based on the natural bond orbitals, and that by Kitaura and Morokuma (KM) [18].

The aim of the present work is to analyse, in terms of optimized geometries, four weak complexes of increasing complexity: water-formamide, methanol-water-imidazole, methylformate-2-formamides, and guanine-cytosine. The dis-

cussion will also deal with the calculation level used, i.e., the basis set dependence at the ab initio SCF and post-SCF levels as well as the use of the AM1 semiempirical method.

## 2. Informatic environment

The geometry optimizations were performed using the GAUSSIAN86-88 [19] and MONSTERGAUSS [20] programs. The geometry scans were run with the GAUSSIAN link 108 package. Most of the calculations were made on a FPS264 attached processor (38 Mflops) linked to a VAX 11/780. At the MP2(FULL)/6-31G\*\* level, the geometry optimizations were done on an IBM 3090/180E with one VF running under VM/CMS. These calculations remain extremely heavy in computer time and scratch space disk.

## 3. Materials and methods

The decomposition of the interaction energy into its main terms was made according to the KM scheme [18] at the SCF level, for the optimized complex conformations. This method is based on the exchange perturbation theory and the first-order term  $E(1)$ , sum of the electrostatic and exchange repulsion terms, is explicitly included in the Tables hereafter presented.

For some cases, an evaluation of the BSSE (basis set superposition error) was performed by the counterpoise method by Boys and Bernardi [21], in which the energies of the monomers and the complex are calculated with the basis set of the complex.

The electronic correlation energy variation related with the interaction,  $E(\delta \text{ corr})$ , is evaluated as the difference in correlation energy for the complex and the monomers:

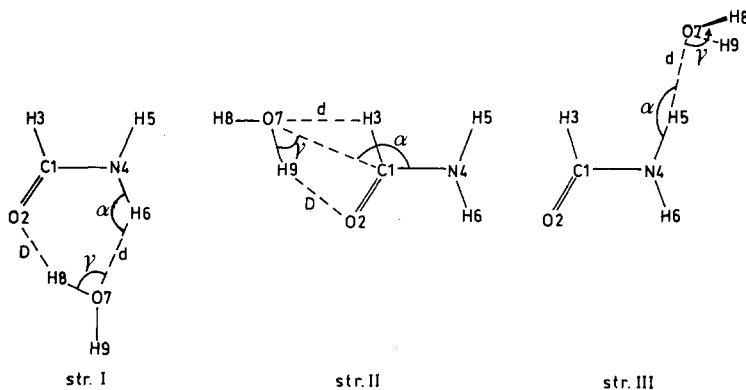
$$E(\delta \text{ corr}) = \Delta E_{\text{corr}}(\text{complex}) - \Delta E_{\text{corr}}(A) - \Delta E_{\text{corr}}(B).$$

The  $E(\delta \text{ corr})$  can be identified with the usual dispersion energy  $E(\text{disp})$  only for large inter-monomer distances. Otherwise it also includes the monomers electronic correlation variation due to the interaction with the other monomers.

### 3.1. Water-formamide complex

The water-formamide complex has been studied in three conformations (str. I, str. II, str. III; see Fig. 1).

They correspond to those obtained by Jasien and Stevens [22] and already served as a model for a detailed study on the BSSE correction [23]. The optimizations of the three conformations were carried out at the ab initio SCF level, within four minimal basis sets (STO-3G, STO-6G [24, 25]; MINI-1, MINI-4 [26]), three double  $\zeta$  basis sets (3-21G [27, 28], 4-31G [29], 6-31G [30]), one triple  $\zeta$  basis set (6-311G [31]), as well as at the ab initio correlated level (6-31G\*\* [32] MP2 [33]) or at the semiempirical AM1 level [34]. These optimizations were performed with the restraint that the complex remain planar. As a matter of fact, in opposition with the result of a full optimization in a double  $\zeta$  basis set, a full optimization within the minimal basis sets led to a nonplanar

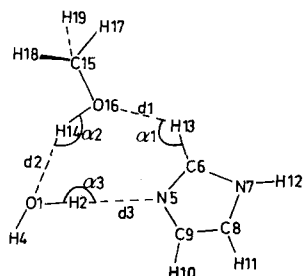


**Fig. 1.** The three conformational arrangements str. I, str. II, and str. III of the (water-formamide) complex

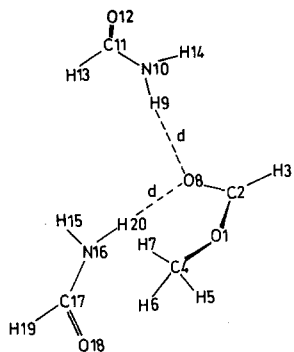
conformation. The complete analytic STO-3G and MINI-1 hessian matrices, at the  $C_s$  optimized conformation str. I, have two negative eigenvalues:  $-0.01487$  a.u.,  $-0.00675$  a.u. (STO-3G) and  $-0.00087$  a.u.,  $-0.00367$  a.u. (MINI-1). They correspond respectively to the  $NH_2$  bending and to the out-of-plane rotation of the water hydrogen H9 (Fig. 1). Since the formamide alone is already found nonplanar with STO-3G [35] as well as with MINI-1 [36], we tried to change the MINI-1 scaling factors of nitrogen (initially optimized for the  $N_2$  electronic ground state [26]) in order to obtain a planar formamide (see appendix). With this new basis set, noted MINI-1', the formamide in the complex is planar. However, the optimized minimum of the complex is still nonplanar. At the planar conformation, the only negative eigenvalue of the analytic hessian ( $-0.00453$  a.u.) corresponds to the out-of-plane rotation of H9.

### 3.2. Methanol-water-imidazole

The complex (methanol-water-imidazole) has been optimized in a cyclic arrangement (see Fig. 2), either at the semiempirical AM1 level, or at the ab initio SCF level within three minimal basis sets (STO-3G, STO-6G, MINI-1) and one double  $\zeta$  basis set (3-21G). Such a trimer was chosen as a minimal model of the nucleophilic agent of a serine peptidase: methanol would model the side chain of the serine, imidazole that of the nearby histidine.



**Fig. 2.** The cyclic conformation of the (methanol-water-imidazole) complex



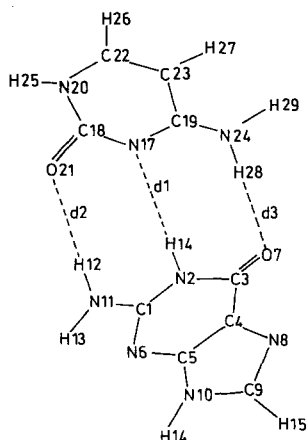
**Fig. 3.** The symmetrical  $C_s$  conformation of the (methylformate-2 formamides) complex. The symmetry plane contains the atoms O1, C2, H3, C4, O8, H5

### 3.3. Methylformate + 2 formamides

The complex (methylformate + 2 formamides) has been optimized in a symmetric  $C_s$  arrangement (see Fig. 3), again at both semiempirical AM1 and ab initio SCF levels. In the latter case, the basis sets used were STO-3G, MINI-1, MINI-1', MINI-4, 3-21G, and 6-31G. Such a geometrical arrangement was chosen for modelling a peptide oxyanion hole interacting with the oxygen of a carbonyl ligand group.

### 3.4. Guanine-cytosine

The complex (guanine-cytosine) has been optimized in its Watson-Crick arrangement (see Fig. 4), at the semiempirical AM1 as well as at the ab initio SCF level. In the former case, either full or planar-restrained optimizations were performed, leading respectively to a minimum or to a first-order critical point. In the latter case, a full optimization was performed within three minimal basis sets (STO-3G, MINI-1, and MINI-1') and a partial optimization of the 6 relative degrees of freedom was also performed within the STO-3G and 6-31G basis sets, the two monomers being frozen at their planar-restrained optimized geometry.



**Fig. 4.** The conformation of the (guanine-cytosine) complex

#### 4. Basis set dependence

SCF calculations with STO-3G, STO-6G, 3-21G, and, in a slighter degree MINI-1' and MINI-1, provide smaller H-bonds as compared with the 6-31G\*\* + MP2 ones. In contrast, MINI-4/SCF distances are too large, as well as the AM1 ones. The splitted basis sets 4-31G, 6-31G, and 6-311G distances very well agree with the 6-31G\*\* + MP2 ones.

##### 4.1. The water-formamide complexes

As far as the relative orientations are concerned, all the calculations agree very well for str. I. For the two other conformations, and especially for str. II, the relative orientations are much more dispersed versus the basis sets. This is due to the very flatness of the hypersurface. In the case of str. II, the minimal basis sets provide only one minimum of type "b" (the torsion  $\text{pl} = 0[\text{water}]-\text{C}-\text{N}-\text{O}[\text{formamide}] = 0$ ), as Jasien and Stevens [22], while splitted basis sets, except 3-21G, provide only minimum of type "a" ( $\text{pl} = 180.0$ ). With 3-21G, both minima of type "a" and "b" appear, whereas MINI-4 provides one minimum in-between the types "a" and "b". It seems that the description of a H-bonding of the type  $\text{O}-\text{C}-\text{H} \dots \text{O}-\text{H}$  is very sensitive to the basis set quality. It should be pointed out that the hydrogen of the formamide has a nonnegligible net charge of ca. 0.15 (from a Mulliken population analysis) compared with that of ca. 0.35 for each of the two hydrogens attached to the nitrogen. The AM1 optimization provides a nearly cyclic conformation for str. II, apparently making no distinction between a H-bond like  $\text{X}-\text{H} \dots \text{X}'$  ( $\text{X}, \text{X}' = \text{N}, \text{O}$ ) and like  $\text{C}-\text{H} \dots \text{X}$ .

##### 4.2. Methanol-water-imidazole

All the ab initio results are very similar (see Table 1). In contrast, the AM1 calculation is somewhat different. While, in ab initio, all the atoms H13-O16-H14-O1-H2-N5-C6 are nearly coplanar, in AM1 only H13-O16-O1-N5-C6 lie in the same plane; H14 is slightly above this plane, whereas H2 is beyond it. Moreover, the N5 lone pair does not point towards H2 any longer but in-between H2 and H4, the two distances H2-N5 and H4-N5 being nearly equal. However, if one excepts this discrepancy concerning the orientation of the water hydrogens, the three entities are well relatively positioned.

**Table 1.** Optimized geometrical parameters for the (methanol-water-imidazole) complex in a cyclic arrangement (Fig. 2), versus the level of calculation

Calculation level	$d_1$	$d_2$	$d_3$	$\alpha_1$	$\alpha_2$	$\alpha_3$
STO-3G	1.9202	1.5916	1.7999	150.02	167.84	158.33
STO-6G	1.8952	1.5577	1.7517	149.34	168.15	159.10
MINI-1	2.0866	1.7255	1.8468	150.19	167.20	160.60
3-21G	2.0266	1.6792	1.7856	150.99	166.60	159.19
AM1	2.1702	2.1132	2.6750	165.16	114.01	110.78

**Table 2.** Energies (kcal/mol) associated with the variation of four geometrical parameters defining the relative orientation of the water molecule in the complex (methanol-water-imidazole) (see text)

$\Delta E$	$r_1$	$r_2$	$t_1/\Delta d_2$	$\Delta d_3$
0.4	0	0	0/0	+0.1 Å
0.4	0	0	0/0	-0.1 Å
1.0	+25°	0	0/0	0
1.0	-35°	0	0/0	0
1.0	0	-10°	0/0	0
1.0	0	+10°	0/0	0
1.0	0	0	-1.5°/-0.1 Å	0
1.0	0	0	+4°/+0.15 Å	0
1.4	+30°	0	0/0	0
1.4	-60°	0	0/0	0
2.0	+35°	0	0/0	0
2.0	-140°	0	0/0	0
2.0	0	-15°	0/0	0
2.0	0	+15°	0/0	0
2.0	0	0	-3°/-0.15 Å	0
2.0	0	0	+6.5°/+0.32 Å	0

In order to calibrate the importance of well describing the relative orientation of the water molecule, four types of variables scans were performed, at the ab initio SCF level with the MINI-1 basis set:

(i)  $r_1$  = rotation of H4, from its equilibrium position, around the axis O1-H2,

(ii)  $r_2$  = rotation of H2, from its equilibrium position, around the axis O1-H4,

(iii)  $t_1$  = variation of the angle H2-N5-C6, directly related with  $\Delta d_2$ ,

(iv) the distance variation  $\Delta d_3$ ,

all the other internal coordinates being frozen at their optimized values. The results are reported in Table 2.

It is clear that small energy amounts (<2 kcal/mol) can make the water molecule move significantly by rotation and translation. Thus, it does not seem of crucial importance to describe very precisely its relative orientation.

#### 4.3. Methylformate-2 formamides

The two formamides were not found planar with the usual minimal basis sets (Table 3), whereas they are with double  $\zeta$  basis sets or at the AM1 level. Again, within the MINI-1' basis set, the two formamides are nearly planar.

The orientation of the formamides relative to the ester can substantially vary. This is due to the flatness of the energy hypersurface as confirmed by the very small hessian eigenvalues related to the torsions defining the orientation of the formamides. For instance, a scan of the two torsion angles  $p_1 = \text{N10-H9-O8-C2}$  and  $p_2 = \text{C11-N10-H9-O8}$ , at the SCF MINI-1' level, led to an energy variation less than 3 kcal for  $-35^\circ < p_1 < +35^\circ$  and less than 0.5 kcal for  $-55^\circ < p_2 < +55^\circ$ . Thus, large variations of the orientation of the two formamides result in small energetic differences.

**Table 3.** Optimized geometrical parameters for the (methylformate-2 formamides) complex in a symmetrical  $C_s$  arrangement (Fig. 3), versus the level of the calculation.Definition of the torsions:  $p1 = O1-C2-O8-H9$ ,  $p2 = H14-N10-H9-O8$ ,  $p3 = H14-N10-H9-C11$ ,  $p4 = C11-N10-H9-O8$ 

Calculation level	$d$	$p1$	$p2$	$p3$	$p4$
STO-3G	1.9463	-127.55	-140.53	125.75	93.72
MINI-1	2.0933	-126.89	-141.74	140.77	77.49
MINI-1'	2.0388	-128.87	-116.81	164.85	78.35
MINI-4	2.2760	-117.97	-135.34	164.58	60.08
3-21G	2.0524	-113.67	-135.76	178.24	46.00
6-31G	2.1365	-109.85	-138.47	179.14	42.39
AM1	2.1861	-121.10	-95.22	-178.42	83.21

#### 4.4. Cytosine, guanine and the complex (guanine-cytosine)

Both guanine and cytosine contain an amine group. A full optimization at the SCF level with the 6-31G basis set results in a minimum that is planar for the two molecules, though the starting geometries for the optimizations were chosen as nonplanar.

In contrast, the optimization with the minimal basis sets STO-3G and MINI-1 produce a nonplanar minimum for the two molecules (see Tables 4–5). The planar optimized geometries in these minimal basis sets revealed to be first-order transition states (TS), the hessian eigenvector associated with the negative eigenvalue being the out-of-plane deformation of the amine group.

The optimization was also performed within the MINI-1' basis set. Then, the minimum of the cytosine is found planar, but not yet that of the guanine that has become, however, much less bent.

The problem of nonplanarity of these two molecules, and certainly of the (guanine-cytosine) complex, was not frequently addressed as they were usually considered as planar [9–15, 37]. Nevertheless, Gould and Hillier [38] mentioned, at the SCF 6-31G\*\* level, that both bent and planar optimized geometries could be obtained for the cytosine that were very close in energy, the bent conformation being slightly more stable.

**Table 4.** Optimized torsions for the amine group of the cytosine minimum as a function of the calculation level.Definition of the torsions (see Fig. 4):  $p1 = H29-N24-C19-C23$ ,  $p2 = H29-N24-C19-H28$ ,  $p3 = H28-N24-C19-N17$ 

Calculation level	$p1$	$p2$	$p3$
STO-3G	-42.00	122.99	17.84
MINI-1	-30.67	135.54	16.16
MINI-1'	0.00	180.00	0.00
6-31G	0.00	180.00	0.00
AM1	-15.54	157.82	8.45

**Table 5.** Optimized torsions for the amine group of the guanine minimum as a function of the calculation level.

Definition of the torsions (see Fig. 4):  $p1 = \text{H13-N11-C1-N6}$ ,  
 $p2 = \text{H13-N11-C1-H12}$ ,  $p3 = \text{H12-N11-C1-N2}$

Calculation level	$p1$	$p2$	$p3$
STO-3G	-13.84	120.93	47.44
MINI-1	-14.15	130.55	37.76
MINI-1'	-5.32	160.75	14.81
6-31G	0.00	180.00	0.00
AM1	-9.65	136.15	36.79

**Table 6.** Deformation energies (kcal/mol) to make cytosine and guanine planar (first-order critical points, except in MINI-1' for cytosine)

Calculation level	Cytosine	Guanine	Total
STO-3G	2.477	2.987	5.464
MINI-1	0.641	1.076	1.717
MINI-1'	0.000	0.005	0.005
AM1	0.027	0.522	0.579

The deformation energies to make the monomers planar, from their nonplanar minimum, are shown in Table 6. The much greater values obtained in STO-3G can be related to the much more bent conformations of the monomers and the complex. Thus, the angular variations related to the torsional ones are greater and both then induce a greater deformation energy.

The fully optimized minimum for (guanine-cytosine) complex is nonplanar in STO-3G and MINI-1, while it is planar in MINI-1' (Table 7). Thus, the nonplanarity of the complex is directly related to the nonplanarity of the amine groups of the monomers. Moreover, the energy hypersurface is very flat, as confirmed by the small eigenvalues of the analytic hessian. For instance, in MINI-1', the two lowest eigenvalues, equal to 0.00871 and 0.00983 a.u., respectively, correspond to the variation of the relative orientation of the monomers and to the bending of the guanine amine hydrogens.

The flatness of the energy hypersurface and the coupling between the  $\text{NH}_2$  bending and the nonplanar relative orientation of the monomers have a greater incidence at the AM1 level than at the ab initio level. The arrangement of the two monomers is much more bent in the former case, as seen in Table 7 ( $p1$ ,  $p2$ ).

At the AM1 level, the TS obtained after a planar-restrained optimization is characterized by a small negative hessian eigenvalue ( $-0.00369$  a.u.), whose eigenvector main components are the out-of-plane motion of the guanine amine hydrogens and the out-of-plane motion of the cytosine relative to the guanine, thus clearly indicating the coupling between the planarity of the complex and the bending of the amine hydrogens.



**Table 7.** Optimized geometrical parameters corresponding to the minimum of the complex (guanine-cytosine), as a function of the calculation level. Definition of the torsions (see Fig. 4):  $p1 = N17-H14-N2-C3$ ,  $p2 = C18-N17-H14-C1$ ,  $p3 = H13-N11-C1-N6$ ,  $p4 = H13-N11-C1-H12$ ,  $p5 = H12-N11-C1-N2$ ,  $p6 = H29-N24-C19-C23$ ,  $p7 = H29-N24-C19-H28$ ,  $p8 = H28-N24-C19-N17$ . Definition of the distances:  $d1 = N17-H14$ ,  $d2 = O21-H12$ ,  $d3 = H28-O7$ .

Calculation level	$d1$	$d2$	$d3$	$p1$	$p2$	$p3$	$p4$	$p5$	$p6$	$p7$	$p8$
STO-3G	1.5466	1.6874	1.4423	48.38	-0.14	-18.02	139.44	25.26	-2.81	171.66	5.91
MINI-1	1.7460	1.8025	1.6645	37.22	-0.76	-9.57	158.17	13.75	-1.84	174.86	3.60
MINI-1'	1.7011	1.7632	1.6061	0.44	0.05	0.00	180.00	0.00	-0.02	-179.98	0.00
AM1	2.3474	2.1146	2.0872	98.65	45.26	-13.75	140.06	29.60	-8.14	166.85	6.31

## 5. Additivity of the energy components

### 5.1. Methanol-water-imidazole

It seems that the energy components could be considered as nearly additive properties (see Table 8). For instance, in the water + (methanol-imidazole) interaction, EX is approximately the sum of the EX for the interactions water-methanol and water-imidazole. This is true, within the variation of ca. 1 kcal/mol, for the other contributions. The nonadditive three-body term remains small for the other components though a little higher for E(ES) ( $\approx 1.5$  to 2 kcal/mol). Nevertheless, the observation of the *near additivity* of the energy terms within a range of 1 to 2 kcal/mol results in a *nonadditive* total interaction energy.

One point to be emphasized is that the dimeric methanol + imidazole interaction is nonnegligible at all though the H-bonding that occurs is not a "usual" one but a C6-H13---O16.

### 5.2. Methylformate + 2 formamides

A very interesting point that comes out from Table 9 concerns the nonadditivity character of the interaction energy components. On the basis of the comparison between the interaction E + (Fo-Fo) and twice (E-Fo) + Fo, three of the four energy terms, i.e. EX, PL, and CT, are additive within a threshold of 1 kcal/mol, while ES is quite nonadditive.

As for the other studied complexes, the total interaction between (E-Fo) and Fo comes from a sum of small similar contributions, E(1), E(PL) and E(CT). However, in the case of the interaction E + (Fo-Fo), those three contributions are no longer similar since E(1) becomes largely predominant due to the ES component. This difference is of course to be linked with the nonadditivity of the ES term.

## 6. Conclusions

The energy hypersurfaces generated for the studied complexes are very flat according to a few torsional angles thus leading to relative orientations of the monomers that can significantly differ as a function of the calculation level used. In such cases, a correlated level (MP2 or MP4) should be employed to determine the geometry for which the repulsive and attractive forces cancel. Nevertheless, the reliability and concomitantly the interest of the precision in the determination of those relative orientations seem to be of minor importance in account to the zero point energy that is available for the system to move. Thus, an SCF or even an AM1 calculation is found to give very satisfactory geometries for the stable conformations.

It is clear that a complex stability, as it is reflected by the associated interaction energy at its optimized conformation, is due to the four main components E(1), E(PL), E(CT), and E( $\delta$  corr), none of them being largely greater than the others. Nevertheless, it has to be noted that the absolute values of E(ES) and E(EX) are much greater than any of the other terms.

The two terms E(PL) and E( $\delta$  corr) were often found to be of the same order of magnitude. Nevertheless, they do not vary in the same way as different

**Table 8.** Energy components, at the SCF 3-21G and SCF 6-31G levels, at the optimized 3-21G geometry of the complex (methanol-water-imidazole) (see Fig. 2).

Energies are in kcal/mol.  $E(\delta \text{ corr})$  is calculated at the FC MP2 level.

Abbreviations: W = water; M = methanol; Im = imidazole.

For the CP correction for BSSE in the trimer, two possibilities were considered:  $\text{BSSE}_2(\text{A}-\text{B}-\text{C}) = \text{BSSE}(\text{A}-\text{B}) + \text{BSSE}(\text{C})$ ;  $\text{BSSE}_3(\text{A}-\text{B}-\text{C}) = \text{BSSE}(\text{A}) + \text{BSSE}(\text{B}) + \text{BSSE}(\text{C})$

	W - M	W - Im	M - Im	M + (W - Im)	Im + (M - W)	W + (Im - M)	M + W + Im
E(ES)	-15.179	-16.649	-6.527	-23.716	-25.330	-33.420	-
E(EX)	+14.741	+15.164	+5.305	+19.961	+19.433	+29.063	-
E(I)	-0.438	-1.485	-1.222	-3.755	-5.896	-4.357	-
E(PL)	-0.887	-1.790	-0.453	-1.982	-3.470	-3.538	-
E(CT)	-7.249	-5.720	-4.998	-13.609	-11.272	-13.964	-
E(res)	-2.195	-1.727	-0.915	-4.915	-3.576	-5.537	-
E(SCF)	-10.769	-10.721	-7.588	-24.261	-24.214	-27.395	-33.931
E(def)	+0.726	+0.710	+0.666	+1.051	+1.051	+1.051	+1.051
BSSE,2	-	-	-	+11.324	+9.142	+9.698	-
BSSE,3	+6.318	+6.074	+3.414	+13.308	+13.308	+13.308	+13.308
$E(\delta \text{ corr})$	-2.060	-1.354	-1.919	-4.510	-3.804	-3.945	-5.864

	W - M	W - Im	M - Im	M + (W - Im)	Im + (M - W)	W + (Im - M)
E(ES)	-17.495	-18.997	-7.938	-27.612	-28.730	-38.166
E(EX)	+17.140	+17.288	+6.793	+24.099	+23.321	+34.015
E(I)	-0.355	-1.709	-1.145	-3.513	-5.409	-4.151
E(PL)	-1.243	-2.828	-0.587	-2.577	-4.961	-5.280
E(CT)	-4.098	-4.890	-1.929	-7.056	-6.891	-9.722
E(res)	-0.396	+0.189	-0.214	-1.266	-0.296	-0.623
E(SCF)	-6.093	-9.238	-3.875	-14.412	-17.557	-19.775
BSSE,2	-	-	-	+2.720	+2.730	+2.970
BSSE,3	+3.370	+2.670	+2.280	+4.160	+4.160	+4.160

6-31G level

**Table 9.** Energy components (kcal/mol) for the (methylformate + 2 formamides) complex at its optimized geometry, as a function of the basis set.E( $\delta$  corr) is calculated at the FC MP2 level.

Abbreviations: E = methylformate ester; Fo = formamide

Decomposition scheme Basis set	E + (Fo - Fo)				(E - Fo) + Fo	
	STO-3G	MINI-1'	3-21G	6-31G	3-21G	6-31G
E(ES)	-7.132	-7.768	-11.353	-12.390	-3.980	-4.230
E(EX)	+10.937	+7.197	+6.012	+5.286	+2.918	+2.597
E(I)	+3.805	-0.571	-5.341	-7.104	-1.062	-1.633
E(PL)	-0.568	-0.762	-1.623	-1.892	-0.480	-0.555
E(CT)	-9.639	-5.413	-6.566	-2.189	-3.438	-1.165
E(res)	-0.039	-0.054	-1.176	-0.220	-0.299	-0.057
E(SCF)	-6.441	-6.800	-14.706	-11.404	-5.278	-3.409
E( $\delta$ corr)	-0.542	—	-1.099	-0.031	—	—

interactions are considered in the same complex (Table 8). Moreover, it seems that (Table 9) E( $\delta$  corr) more rapidly vanishes as the inter-fragment distance grows up than does E(PL). While E(PL) is a function of permanent and induced multipoles, E( $\delta$  corr) is a function of instantaneous and induced multipoles and certainly more sensitive to the relative distance of the interacting entities. Furthermore, the variation of E(PL) as a function of  $R$  may begin with  $R^{-4}$ , whereas that of E( $\delta$  corr) only begins with  $R^{-6}$ .

Short-distance interactions (EX and CT) as well as second-order long-distance ones (PL, disp) seem to be nearly additive within ca. 1 kcal/mol. In contrast, the first-order long-range interaction ES varies as powers of  $R^{-1}$ ,  $R^{-2}$ , ... and thus has a wide tail for three or four (or...) body interaction. Consequently, this term is by nature nonadditive.

Among all the minimal basis sets hereabove considered, it appears that MINI-1 provides the best results as well for the optimized conformations as for the energy decomposition. Furthermore, the "home-adapted" MINI-1' basis set seems even better to predict optimized conformations for amines and amides.

## Appendix

With reference to [35], STO-3G provides an optimized geometry of the formamide that is nonplanar while double  $\zeta$  basis sets agree with the experiments that find a planar formamide.

The complete optimization of the formamide in the MINI-1 basis set also results in a nonplanar conformation. Moreover, a planar constrained optimization in MINI-1 leads to a first-order critical point, associated with an eigenvector describing the bending of the amide hydrogens.

Starting from 1.017 and 1.061 respectively [26], a rough optimization of the scaling factors of the nitrogen  $2s$  and  $2p$  orbitals gives a planar formamide for values of 1.022 and 1.005, with optimized internuclear distances not very different from those obtained with the usual MINI-1. Moreover, this scaling factors choice leads to the lowest total energy. These results have been confirmed by calculations on  $\text{NH}_3$  and  $\text{N}_2$ .

*Note.* More detailed information about geometry parameters and decomposition energy tables is available on request.

*Acknowledgements.* This work was supported, in part, by the Belgian programme on Interuniversity Poles of attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming (PAI no. 19), an Action concertée with the Belgian Government (convention 86/91-90), the Fonds de la Recherche Scientifique Médicale (contract no. 3.4537.88), and a Convention tripartite between the Région wallonne, SmithKline Beecham, U.K., and the University of Liège. G. D. is Chercheur qualifié of the Fonds National pour la Recherche scientifique, Bruxelles.

## References

1. Hobza P, Zahradnik R (1980) *Weak intermolecular interactions in chemistry and biology.* Elsevier, Amsterdam
2. Chalasinski G, Gutowski M (1988) *Chem Rev* 88:943
3. Pauling L, Corey RB, Branson HR (1951) *Proc Natl Acad Sci US* 36:205
4. Pauling L, Corey RB (1951) *Proc Natl Acad Sci US* 37:251
5. Neurath H (1964) *The proteins.* vol 1, II, Academic Press, NY
6. Pullman B, Pullman A (1963) *Quantum biochemistry.* Wiley, NY
7. Marsch RE (1968) *Structural chemistry and molecular biology.* Freeman, San Fran
8. Watson JD, Crick FHC (1953) *Nature* 171:737; *ibid*, 171:964
9. De Voe H, Tinoco I (1962) *J Mol Biol* 4:500
10. Nash HA, Bradley DF (1965) *Biopolymers* 3:261
11. Nash HA, Bradley DF (1966) *J Chem Phys* 45:1380
12. Clementi E, Corongiu G, Detrich J, Chin S, Domingo L (1984) *Int J Quant Chem: Quant Chem Symp* 18:601
13. Aida M (1988) *J Comp Chem* 9:362
14. Hobza P, Sandorfy C (1987) 109:1302
15. Otto P (1986) *Int J Quant Chem* 30:275
16. Aida M, Nagata C (1986) *Int J Quant Chem* 29:1253
17. Reed AE, Curtiss LA, Weinhold F (1988) *Chem Rev* 88:899
18. Kitaura K, Morokuma K (1976) *Int J Quant Chem* 10:325
19. Programs GAUSSIAN86-88, Frisch MJ, Head-Gordon M, Schlegel HB, Raghavachari K, Binkley JS, Gonzalez C, Defrees DJ, Fox DJ, Whiteside RA, Seeger R, Melius CF, Baker J, Kahn LR, Stewart JJP, Fluder EM, Topiol S, Pople JA, Gaussian Inc, Pittsburgh PA
20. Program MONSTERGAUSS, Peterson MR, Poirier RA, Csizmadia IG (1980) Dept of Chemistry, Univ of Toronto, Toronto, Canada
21. Boys SF, Bernardi F (1970) *Mol Phys* 19:553
22. Jasien PG, Stevens WJ (1986) *J Chem Phys* 84:3271
23. Dehareng D, Dive G, Ghuysen JM (1992) *Theor Chim Acta* 81:281
24. Hehre WJ, Stewart RF, Pople JA (1969) *J Chem Phys* 51:2657
25. Hehre WJ, Ditchfield R, Stewart RF, Pople JA (1970) *J Chem Phys* 52:2769
26. Tatewaki H, Huzinaga S (1980) *J Comp Chem* 1:205
27. Binkley JS, Pople JA, Hehre WJ (1980) *J Amer Chem Soc* 102:939
28. Gordon MS, Binkley JS, Pople JA, Pietro WJ, Hehre WJ (1982) *J Amer Chem Soc* 104:2797
29. Ditchfield R, Hehre WJ, Pople JA (1971) *J Chem Phys* 54:724
30. Hehre WJ, Ditchfield R, Pople JA (1972) *J Chem Phys* 56:2257
31. Krishnan R, Binkley JS, Seeger R, Pople JA (1980) *J Chem Phys* 72:650
32. Frisch MJ, Pople JA, Binkley JS (1984) *J Chem Phys* 80:3265
33. Moller C, Plesset MS (1934) *Phys Rev* 46:618
34. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) *J Amer Chem Soc* 107:3902
35. Carlsen NR, Radom L, Riggs NV, Rodwell WR (1979) *J Amer Chem Soc* 101:2233
36. Unpublished results, available on request
37. Del Bene JE (1983) *J Phys Chem* 87:367
38. Gould IR, Hillier IH (1989) *Chem Phys Lett* 161:185