# **Molecular Mechanics (MM4) and ab Initio Study of Amide**-**Amide and Amide**-**Water Dimers**

## **Charles H. Langley\* and Norman L. Allinger\***

Department of Chemistry, Chemistry Annex, The University of Georgia, Athens, Georgia 30602-2526 *Recei*V*ed: July 31, 2002*

As an extension of earlier amide work, hydrogen bonding parameters in amides have been developed for the molecular mechanics force field MM4. These parameters were obtained from studies on several configurations for each of the following complexes: formamide dimer, formamide-water, *<sup>N</sup>*-methylacetamide dimer, and *<sup>N</sup>*-methylacetamide-water. MM4 hydrogen bonding distances and binding energies were compared and fit to ab initio calculations on these systems reported by previous authors, preferably those done at the MP2 level of theory using double- or triple-*ú* basis sets, in addition to calculations carried out in the present study at the MP2/6-311++ $G(2d,2p)$  level. The MM4 optimized structures for the lowest energy configurations of the formamide-water complex were also compared to the existing results from microwave studies on these systems.

#### **Introduction**

Hydrogen bonding is the general term used to describe the interaction between a weakly acidic hydrogen atom in a molecule to another atom or electron source, and the best known example occurs in liquid water, where multiple hydrogen bonds form between molecules and hold the entire assembly together. These special types of nonbonded interactions have been studied by experimental, ab initio, and molecular mechanics methods. A general broad-ranging study of the structures of organic and biological molecules requires that these interactions be taken into account. We have made, and continue to make, extensive efforts to better understand and describe them using our molecular mechanics force field.<sup>1-3</sup> We have prevously studied many examples $4-6$  of hydrogen bonding. One particular area in which hydrogen bonding is important, and which constitutes a basic field for the application of molecular mechanics, is in the field of proteins. The two major types of hydrogen bonding that occur in proteins are the intramolecular amide $\text{-}$ amide (N $\text{-}$  $H \cdot \cdot \cdot O=C'$  interactions and the intermolecular amide-water  $(O-H\cdots O=C'$  and  $N-H\cdots OH_2$ ) interactions. Both are essential to the stabilities of the secondary and tertiary structures of proteins in the aqueous phase. Therefore, it may be anticipated that in order to accurately model polypeptides and proteins by theoretical methods, especially by molecular mechanics methods, accurate hydrogen bonding potentials are crucial. This work is concerned with studying and modeling these two types of interactions using our latest molecular mechanics program  $MM4<sup>3</sup>$ 

*trans*-*N*-Methylacetamide is the smallest representative peptide unit in polypeptides and proteins. MM4 force-field parameters have already been developed for this and other amide compounds in a previous study.7 In the amide paper, only simple amides were studied, more specifically, their structures, conformational transition states, vibrational frequencies, heats of

formation, and heavy atom rotational barriers, including torsions about the peptide bond (known as the  $\Omega$  dihedral in protein papers). The MM4 modeling of  $\Phi$  and  $\Psi$  torsions from dipeptide analogues has not yet been carried out, because there exist in peptides internal  $N-H\cdots O=C'$  interactions which first need to be studied separately. Hydrogen bonding parameters for water, methanol, and ammonia were optimized in an earlier MM3 study<sup>5</sup> in which  $(H_2O)<sub>n</sub>$  ( $n = 2$ , 3, and 5), (MeOH)<sub>2</sub>,  $(NH_3)$ <sub>2</sub>, and  $NH_3\cdot H_2O$  were used as model systems. In a subsequent paper,<sup>6</sup> MM3 hydrogen bonding parameters for amides were optimized using formamide-formamide, formamide-water, and formamide-formaldehyde complexes as models. However, the parameters were derived from ab initio calculations on these systems at the MP2/6-31G\*\* level, which is now known to be insufficient for such systems, even when corrections for basis set superposition error are made.

In this study, more types of amide-amide and amide-water complexes and more binding configurations for each type were used in obtaining hydrogen bonding parameters for the MM4 force field. MM4 parameters for the  $N-H\cdots O=C'$  interactions between amides and the  $O-H \cdot \cdot \cdot O=C'$  and  $N-H \cdot \cdot \cdot \cdot OH_2$  interactions between amide and water molecules were obtained by studying different binding configurations for each of the following systems: formamide dimer, *N*-methylacetamide dimer, formamide-water, and *<sup>N</sup>*-methylacetamide-water. For water complexes with *N*-methylacetamide, both the trans and cis forms of the amide molecule were studied.

Numerous theoretical studies have been reported previously on nonbonded interactions in amide-amide and amide-water systems, more specifically, formamide dimer,<sup>8-22</sup> *N*-methylformamide dimer,<sup>10</sup> *N*-methylacetamide dimer,<sup>23-25</sup> formamidewater,<sup>26-28</sup> and *N*-methylacetamide-water.<sup>23,25,29-31</sup> Complexes involving two to three water molecules with formamide and *N*-methylacetamide have also been studied.<sup>29-31</sup> In obtaining molecular mechanics parameters, experimental data are preferred if available and reliable. However, there are only marginal experimental data $32-38$  for hydrogen bonding interactions involving amides. An X-ray crystal structure of the cyclic formamide dimer<sup>32</sup> and a microwave gas-phase structure of the

<sup>\*</sup> To whom correspondence should be addressed. Charles H. Langley: Phone: (706) 542-2044. Fax: (706) 542-2673. E-mail: clangley@decnla.chem. uga.edu. Norman L. Allinger: Phone: (706) 542-2043. Fax: (706) 542- 2673. E-mail: allinger@sunchem.chem.uga.edu.



**Figure 1.** Directional hydrogen bonding involving lone-pairs.

cyclic formamide-water complex<sup>33</sup> have been reported. Raman studies measuring shifts of the C'=O, C'-N, and N-H bands in the gaseous, neat liquid, solution, and solid phases of amides have suggested that there are significant differences in amideamide hydrogen bond strengths for different amide molecules.35,36 With 13C NMR spectroscopy, relative strengths of amide-amide and amide-water hydrogen bonds have been determined indirectly by measuring the kinetic barrier to cistrans isomerization of the prolyl peptide bond in various amide solvents and water.<sup>37</sup> However, in none of these studies were absolute values for hydrogen bond strengths determined.

At present, accurate molecular mechanics parameters for hydrogen bonds in amides (and in all other organic compounds for that matter) can best be obtained from ab initio or density functional calculations, so the question to be asked is, what level of theory provides adequate accuracy? MM4 parameters for the hydrogen bond lengths  $(l_0)$  and binding energies  $(\epsilon)$  were obtained from HF and MP2 calculations reported by previous authors12,17,18,23 using double- and triple-*ú* Dunning type basis sets, which minimize basis set superposition errors (BSSE). Unfortunately, calculations on formamide-water systems at these basis set sizes have not yet been reported. Gaussian type basis sets at the size of  $6-311++G(2d,2p)$  are nearly equivalent to triple-*ú* basis sets and computationally less intensive. Therefore, ab initio calculations at the MP2/6-311++ $G(2d,2p)$ level were carried out for the amide-water and formamide dimer systems looked at in the MM4 study.

#### **Hydrogen Bonding Potential**

In the early MM2 version of the molecular mechanics program, lone pairs were treated explicitly. This has its obvious advantages when directionality of hydrogen bonding is taken into account. Specific geometric arrangement of the lone pairs on the hydrogen bond acceptor atom can be set individually for each atom type. However, the lone pairs are counted along with the other atoms in the minimization process. Although this may be insignificant in simple amides, the calculation slows down significantly as one approaches the molecular sizes of proteins and even more so in simulations of proteins in the aqueous phase if lone pairs on the water molecules are also explicitly treated. Furthermore, the low mass of the lone pair presented additional problems in vibrational frequency calculations. Worse, the overall model becomes either unphysical, very complex, or both. In the MM3 program, the lone pairs were removed. Although the MM3 hydrogen bonding potential<sup>5,6</sup> had a cosine angular dependence term, which aligned the hydrogen between the two heavy atoms, the directionality due to the lone pairs was lost. In the MM4 program, this directionality was reintroduced, but the lone pairs in the hydrogen bonding arrangement shown in Figure 1 were treated implicitly rather than explicitly.

The energy  $E_{HB}$  (in kcal mol<sup>-1</sup>) for the hydrogen bonding interaction H····B in the MM4 force field is represented by the  $\leq 3.02$ 

 $> 3.02$ 

following equations:39

$$
E_{\rm HB} = \epsilon_{\rm HB} \bigg[ 1.84 \times 10^5 \exp\bigg( -12.0 \frac{R_{\rm HB}}{l_{0,\rm HB}} \bigg) - f
$$

$$
(\theta_{\rm H-A-B}, \theta_{\rm H-B-LP}, R_{\rm AH}) \frac{2.25 \left( l_{0,\rm HB} \right)^6}{D} \bigg]
$$
(1a)

 $l_{0,HB}$  $R_{\rm HB}$ 

when

$$
E_{\rm HB} = \epsilon_{\rm HB} \bigg[ 192.270 \bigg( \frac{l_{0,\rm HB}}{R_{\rm HB}} \bigg)^2 + 1706.96 - 1706.96 f(\theta_{\rm H-A-B}, \theta_{\rm H-B-LP}, R_{\rm AH}) \frac{1}{D} \bigg( \frac{l_{0,\rm HB}}{R_{\rm HB}} \bigg) \bigg] (1b)
$$

*l*0,HB  $R_{HB}$ 

when

where

$$
f(\theta_{H-A-B}, \theta_{H-B-LP}, R_{AH}) =
$$
  

$$
N[\cos(\theta_{H-A-B}) + 1]^n[\cos(\theta_{H-B-LP}) + 1]^n \left(\frac{R_{AH}}{l_{0,AH}}\right)
$$

and  $\epsilon$ <sub>HB</sub> and  $l_{0,HB}$  are the hydrogen bond energy and bond length parameters and *D* is the dielectric constant. Equation 1b is used to prevent atoms H and B from mathematically fusing in eq 1a when they get too close. In the angular dependence function *f*, *N* is a normalization factor. For the case of atom B being nitrogen,  $N = 1/4$  and  $n = 1$ . For the case of oxygen,  $N = 1/16$ and  $n = 2$ .

The hydrogen bond equations in the MM3 force field<sup>5</sup> have the same forms as eqs 1a and 1b except that the dielectric constant lies outside of the brackets in both equations and the angular dependence term is a function of only the hydrogen bond angle  $\theta_{\text{H-A-B}}$  and bond length  $R_{\text{HA}}$ :

$$
f(\theta_{\text{H-A-B}}, R_{\text{AH}}) = [\cos(\theta_{\text{H-A-B}})] \left(\frac{R_{\text{AH}}}{l_{0,\text{AH}}}\right)
$$

In the MM4 hydrogen bond equation, the directionality of the vector representing the lone pair in Figure 1 is determined by the angle  $\chi$ . If B is bound to only one atom, as in the case of  $B-X$  or  $B=X$ , the  $\chi$  angle is defined in reference to the bond. If B is connected to two or more atoms, as in the case of BXY or BXYZ,  $\gamma$  is defined in reference to the line connecting atom B to the midpoint of the line connecting atoms X and Y or to the centroid of the plane connecting the atoms X, Y, and Z. The value of  $\chi$  is set according to the the atom type. In the case of oxygen, *ø* is 65°, whether it is a carbonyl, alcohol, ether, or water oxygen. This value was derived from an average over representative samples of organic compounds containing oxygen. In the case of amines,  $\chi$  is 0°. The disadvantage of the directionality representation in the MM4 equation is that, in the oxygen case, the hydrogen bond interaction around the reference line axis defined by the cone in Figure 1 is the same. Any differences in the resulting hydrogen-bond geometry will be from steric effects from the attached atoms.

### **Method**

Having established the nature of the hydrogen-bond potential to be used, the next step was to determine the MM4 hydrogen

**TABLE 1: MM4 Hydrogen Bonding Parametersfor Amide**-**Amide and Amide**-**Water Systems**

hydrogen bond	atom types <sup><math>a</math></sup>	$\epsilon$ (kcal/mol)	$l_0(A)$
$N-H\cdots O=C'$	2879	8.370	1.800
$H_2O \cdots H-N$	628	5.700	1.865
$HO-H\cdots O=C'$	2179	7.000	1.905
$H-C' = 0 \cdots H-C' = 0$	79175	4.660	2.144
$H_2O \cdot H - C' = O$	6175	4.660	2.144

*<sup>a</sup>* Atom types: 28, amido H; 79, amide O; 9, amido N; 6, O water; 21, H water; 175, aldehydyl H in formamides; 147,  $sp<sup>2</sup>$  carbonyl C' in formamides; 152,  $sp^2$  carbonyl C' in acetamides.

bonding parameters, which are listed in Table 1. Most of the MM4 atom types listed are the same as those assigned in MM3. However, in MM4, there are separate atom types for the carbonyl carbons  $(C')$  in formamides (type 147) and acetamides (type 152) and a special atom type for the aldehydyl hydrogen in formamides (type 175). The justification for doing this is discussed in the amide paper.<sup>7</sup> The hydrogen bond energy  $(\epsilon)$ and bond length  $(l_0)$  parameters were initially fit to calculations reported by previous authors<sup>18,21,23</sup> at the MP2 level using augmented correllation-consisitent polarized Dunning-type double- or triple-*ú* valence basis sets (aug-cc-pVDZ and aug-ccpVTZ). The MM4 bond energies were fit to the BSSE corrected energies. For the formamide-water and *<sup>N</sup>*-methylacetamidewater complexes, there were no such data available. Instead,  $MP2/6-311++G(2d,2p)$  calculations on these complexes were carried out in this work, using the Gaussian 94 software program.40 BSSE corrections to the binding energies were made using the counterpoise method $41$  while taking into account the changes of the geometries of the monomers from their complexed to their totally isolated forms. The majority of the amide-amide and amide-water systems chosen for study had  $C_s$  symmetry. There were some complexes with  $C_1$  symmetry that were studied by ab initio methods by previous authors, but in most of these cases, the relative orientations between the two monomers were not reported, and a valid comparison between these and MM4 calculated structures would be difficult. Two exceptions to this were the cyclic formamide-water and cis-*<sup>N</sup>*-methylacetamide-water complexes, which both have *<sup>C</sup>*<sup>1</sup> symmetry. Both have the same  $H_{w2}$ -O- $H_{w1}$  $\cdots$ O=C' and  $H_2O \cdot$  $\cdot$ H-N double hydrogen bonding arrangement. The  $C_1$  symmetries are due to the remaining water hydrogen  $(H_{w2})$ projecting out the plane formed by this bonding arrangement.



**Figure 2.** MM4 optimized structures of formamide dimers.

A microwave study<sup>33</sup> was reported on the cyclic formamidewater complex and provided important structural information on the hydrogen bond lengths. In this case, the MM4 bond length parameters were weighted in favor of the microwave data. The  $H_{w2}$ -O- $H_{w1}$  $\cdots$ O=C' and  $H_2$ O $\cdots$ H-N (MM4 bond types  $21 \cdot \cdot \cdot 79$  and  $6 \cdot \cdot \cdot 28$ , respectively) hydrogen bond length parameters were obtained indirectly by fitting the MM4 calculated moments of inertia to the experimentally observed moments for this complex.

#### **Results and Discussion**

**Formamide Dimer.** The four different complexes of formamide dimer studied are shown in Figure 2. Their calculated equilibrium geometies (*r*e) are listed in Table 2. The MM4 structure  $(r_{\alpha})$  for the cyclic dimer (FAD1) is compared to the structure determined by X-ray diffraction<sup>32</sup> in Table 3, and their binding energies, calculated with and without BSSE correction, are listed in Table 4. The MM4 bond energy  $(\epsilon)$  parameters were weighted mostly to the estimated MP2 complete basis set (CBS) energies, extrapolated from MP2 calculations<sup>21</sup> at augcc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis set sizes. The MM4 binding energies are in very good agreement with the

**TABLE 2: Calculated Hydrogen Bonding Equilibrium (re) Geometries***<sup>a</sup>* **in Formamide Dimers**

					$N-H\cdots$ O=C' bond				$C'$ -H $\cdots$ O=C' bond	
dimer <sup>b</sup>	geometry	ref	$l$ H $\cdots$ O	$\theta$ N-H $\cdots$ O	$\theta$ C'=O…H	$l N \cdots Q$			$l$ H…O $\theta$ C'-H…O $\theta$ C'=O…H	$l \mathbb{C}' \cdots \mathbb{O}$
	$MP2/aug-cc-pVDZ$	21	1.836	174.2	120.2	2.863				
$FAD1(C_{2h})$	$MP2/aug-cc-pVTZ$	21	1.825	174.2	120.1	2.842				
	MP2/DZP	9	1.986	169.7		2.985				
	$MP2/ 6-311++G(2d,2p)$	this work	1.848	174.0	120.4	2.865				
	MM4		1.869	179.7	116.7	2.888				
	$MP2/aug-cc-pVDZ$	21	1.876	168.8	106.4	2.891	2.274	143.6	113.8	3.231
FAD2(C <sub>s</sub> )	MP2/aug-cc-pVTZ	21	1.857	168.9	105.7	2.863	2.234	144.9	113.1	3.190
	MP2/DZP	9	2.009	164.3		2.992	2.380	138.7		3.282
	$MP2/ 6 - 311 + + G(2d, 2p)$	this work	1.885	168.8	106.6	2.891	2.280	143.5	113.9	3.226
	MM4		1.886	169.1	108.4	2.893	2.280	140.8	116.7	3.218
FAD3(C <sub>s</sub> )	MP2/DZP	9	2.051	173.7	151.7	3.048				
	$MP2/ 6-311++G(2d,2p)$	this work	1.959	177.2	133.2	2.967				
	MM4		1.872	177.8	121.9	2.891				
	$MP2/aug-cc-pVDZ$	21					2.355	142.3	95.3	3.301
FAD4(C <sub>2h</sub> )	MP2/aug-cc-pVTZ	21					2.320	144.4	93.2	3.271
	MP2/DZP	9					2.462	134.0		3.317
	$MP2/ 6-311++G(2d,2p)$	this work					2.355	141.8	95.8	3.287
	MM4						2.324	137.3	100.7	3.228

*<sup>a</sup>* Bond lengths are in angstroms, and angles are in degrees. *<sup>b</sup>* Refer to Figure 2.



bond length $(\check{A})$	$X$ -ray <sup>a</sup> $(r_{\alpha})$	MM4 <sup>b</sup> $(r_{\alpha})$	bond angle (degrees)	$X$ -ray <sup><i>a</i></sup> $(r_{\alpha})$	MM4 <sup>b</sup> $(r_{\alpha})$
<b>CN</b>	1.326(4)	1.359	OCN	124.9(3)	125.1
CO	1.239(4)	1.207	CNH <sub>4</sub>	119(3)	119.2
NH <sub>4</sub>	1.01(5)	1.015	CNH <sub>5</sub>	118(3)	121.9
NH <sub>5</sub>	1.01(5)	1.002	$NCH_6$	116(3)	112.4
<b>CH</b>	1.09(5)	1.104	$CN \cdot \cdot \cdot$ O		118.6
$C_1 \cdots C_2$		4.091	$CO \cdot \cdot \cdot H$		116.6
$O \cdot \cdot \cdot H$	1.94(5)	1.889	NHO		179.1
		1.938c			
$0 \cdot \cdot \cdot N$	2.948(3)	2.904			
		2.942c			

*<sup>a</sup>* Reference 32. Measured at 90 K. *<sup>b</sup>* Simulated at 90 K and dielectric constant  $\epsilon = 1.5$ .  $c \epsilon = 4.0$ .

reported CBS values, which are to within  $0.4$  kcal mol<sup>-1</sup>. The MM4 equilibrium geometries for the formamide dimers FAD1, FAD2, and FAD4 are in reasonably good agreement with the  $MP2/aug-cc-pVTZ$  calculated geometries.<sup>21</sup> Most of the calculated hydrogen bond lengths differ by less than  $0.04 \text{ Å}$ , and the bond angles differ by no more than 5°. The MP2/DZP calculated bond lengths<sup>9</sup> are longer, by about  $0.10-0.15$  Å for most of the bonds listed. The MP2/6-311++ $G(2d,2p)$  bond lengths are somewhere in between.

Both theoretical<sup>8-22</sup> and experimental<sup>32,38</sup> studies show that the cyclic formamide dimer FAD1, having two N-H $\cdot\cdot\cdot$ O=C' bonds (MM4 bond type  $28...79$ ) arranged in a head-to-tail fashion and  $C_{2h}$  symmetry, is the most stable. FAD2, the next most stable, has *Cs* symmetry and a heterocyclic arrangement of one N-H $\cdot \cdot \cdot$ O=C' bond and one C'=O $\cdot \cdot \cdot$ H-C' bond (type 79 $\cdots$ 175). FAD3 ( $C_s$  symmetry) is the third most stable and has a single  $N-H \cdot \cdot \cdot O=C'$  hydrogen bond. For the FAD3 dimer,



**Figure 3.** MM4 optimized structures of *cis*- and *trans*-*N*-methylacetamide dimers.

for which there are no MP2 CBS values, the MM4 binding energy  $(-7.45 \text{ kcal mol}^{-1})$  is lower than the MP2/6-311++G-(2d,2p) BSSE corrected value  $(-5.97 \text{ kcal mol}^{-1})$ . However, by comparison of the MP2/CBS and MP2/6-311++G(2d,2p) energies for the other dimers listed, it is estimated that the MP2/ CBS binding energy for the FAD3 dimer should be about 0.7-  $1.0$  kcal mol<sup>-1</sup> lower, in better agreement with the MM4 value.

The least stable conformer studied, FAD4, has *C*<sup>2</sup>*<sup>h</sup>* symmetry with two  $C' = O \cdots H - C'$  bonds. However, the stability of this complex is greater than one might expect. Most of the binding energies range from  $-4.4$  to  $-5.0$  kcal mol<sup>-1</sup> (for both bonds) according to BSSE corrected MP2 calculations. A similar interaction involving an alkane hydrogen (type  $5 \cdots 79$ ) is much weaker and can be accounted for by the van der Waals terms alone in MM3 and MM4 force fields. However, according to preliminary density functional calculations done by us at the B3LYP/6-311++G(d,p) level, the C'-H $\cdots$ O interaction in formamide is 1.5 times stronger than the same type of interaction in the cyclic dimer of acetaldehyde (type  $5\cdots$ 7), with the BSSE corrected energies being  $-3.92$  kcal mol<sup>-1</sup> for the former and  $-2.37$  kcal mol<sup>-1</sup> for the latter. The MM4 binding energies for FAD4 and the acetaldehyde dimer are  $-4.92$  and  $-1.70$  kcal  $mol^{-1}$ , respectively, which are consistent with the DFT results. The partition of the MM4 binding energy in FAD4 is 72% van der Waals and 28% dipole-dipole interaction (induction effects





 $a$  The authors listed the deformation energy (1.73 kcal mol<sup>-1</sup>), the difference between the electronic energies of the monomers having the same geometries in the dimer and the energies of the fully relaxed monomers, separately from the interaction energies of the dimer. We have included the deformation energy in the values listed here. For the other reported BSSE corrected binding energies, the authors have already included the deformation energy contribution.

**TABLE 5: Calculated Hydrogen Bonding Equilibrium (***r***e) Geometries***<sup>a</sup>* **in** *N***-Methylacetamide Dimers**



*<sup>a</sup>* Bond lengths are in angstroms, and angles are in degrees. *<sup>b</sup>* Refer to Figure 3.

**TABLE 6: Equilibrium Binding Energies (∆***E***) for** *N***-Methylacetamide Dimers**

	optimized	<b>BSSE</b>			$\Delta E$ (kcal/mol)
level	geometry	corr.	ref	NMAD1	NMAD <sub>2</sub>
$MP2/aug-cc-pVDZ$	HF/DZP	no	23	$-16.4$	
MP2/aug-cc-pVDZ	HF/DZP	yes	23	$-14.0$	
$MP2/aug-cc-pVDZ$	$MP2/aug-cc-pVDZ$	no	21	$-17.9$	$-11.5$
$MP2/aug-cc-pVDZ$	MP2/aug-cc-pVDZ	yes	21	$-16.2$	$-10.2$
MP2/aug-cc-pVTZ	MP2/aug-cc-pVDZ	n <sub>0</sub>	21	$-19.1$	$-12.2$
MP2/aug-cc-pVTZ	MP2/aug-cc-pVDZ	yes	21	$-17.2$	$-10.8$
$MP2/6-311++G(2d,2p)$	$MP2/6-311++G(2d,2p)$	n <sub>0</sub>	this work	$-16.5$	
$MP2/6-311++G(2d,2p)$	$MP2/6-311++G(2d,2p)$	yes	this work	$-14.0$	
MM4	MM4			$-15.3$	$-7.9$





*<sup>a</sup>* Bond lengths are in angstroms, and angles are in degrees. *<sup>b</sup>* Refer to Figure 4.



Figure 4. MM4 optimized structures of formamide-water dimers.

included) between the two monomers, whereas in the acetaldehyde dimer, it is 37% van der Waals and 63% dipole-dipole interaction. However, the dipole-dipole interaction energies themselves for the formamide and acetaldehyde dimers are roughly the same, being  $-2.1$  and  $-1.5$  kcal mol<sup>-1</sup>, respectively. Therefore, the difference in binding energies lies mainly in the different vdW characteristics of the carbonyl oxygen and aldehydyl hydrogen atoms in the formamide and acetaldehyde molecules. Originally in MM3, the atom type for the aldehydyl hydrogen in formamides was the same as that for the corresponding hydrogen in aldehydes (atom type 5). Parameters for

aldehydes $42$  have already been optimized and fixed in MM4, including nonbonded parameters. Therefore, assigning a new the atom type to this hydrogen in formamides (type 175) was the only way of making this interaction strong enough in formamides without affecting the results from earlier MM4 work.

For the MM4 calculated  $r_\alpha$  structure for FAD1 in Table 3, agreement with the X-ray structure measured at 90 K is to within experimental error for most of the bond lengths and bond angles listed. The MM4 C'=O bond length is about 0.03 Å shorter than the experimentally measured bond length, and the MM4  $C'$  N bond length is about 0.03 Å longer. However, these discrepancies are in part due to errors in the experimental method itself. In X-ray diffraction experiments, the distances measured are between centers of electron density rather than between nuclei. For the carbonyl bond, the lone pairs on the oxygen shift the center of electron density away from the bonding region, and the measured bond length will be longer than the actual internuclear distance. For the  $C'$ -N bond, the center of electron density surrounding the nitrogen atom will be shifted toward the carbonyl carbon because delocalization of the amide bond, and the measured bond length will be shorter. In addition, there are significant but smaller errors of bond length measurement ranging from 0.003 to 0.015 Å introduced by thermal motions of the molecules in the crystal. Also, part of these discrepancies are due to limitations in the MM4 routine<sup>43</sup> for converting among the different types of bond lengths. The  $r_{z}$ - and  $r_{\alpha}$ -type values measured in microwave and X-ray diffraction experiments, respectively, are converted from the vibrationless  $r_{e}$ -type values calculated by ab initio and density functional methods. The routine cannot take into account the

∆*E* (kcal/mol)

**TABLE 8: Equilibrium Binding Energies (∆***E***) for Formamide**-**Water Dimers**

					$\Delta L$ ( <i>KCal/HIOT</i> )		
level	optimized geometry	<b>BSSE</b> corr.	ref	FAW1	FAW <sub>2</sub>	FAW3	
$MP2/6-311++G(2d,2p)$ $MP2/6-311++G(2d,2p)$ MM4	$MP2/6-311++G(2d,2p)$ $MP2/6-311++G(2d,2p)$ MM4	no yes	this work this work this work	$-9.94$ $-8.52$ $-8.87$	$-7.00$ $-5.99$ $-6.13$	$-5.57$ $-4.77$ $-5.55$	

**TABLE 9: Experimental and Calculated Moments of Inertia of Cyclic Formamide**-**Water Dimer (FAW1; units in** amu  $\AA^2$ <sup>*a*</sup>



*a* Conversion factors: 83920.99 (MHz)(10<sup>-39</sup>g cm<sup>2</sup>), 0.1660565  $(10^{-39} \text{g cm}^2) / (\text{amu Å}^2)$  2.7993028 (cm<sup>-1</sup>)(10<sup>-39</sup>g cm<sup>2</sup>).

electron density shifts described above in conversions to  $r_\alpha$ values. It also is expected to be less accurate for the hydrogen bond lengths which have a highly anharmonic vibration potential. For these reasons, the MM4 bond lengths were fit more toward the MP2 equilibrium bond lengths. Although these errors may also manifest themselves in the measured nonbonded distances, the discrepancies between the MM4 and X-ray values for the O...H and O...N bond lengths are mostly due to difficulty in simulating the dielectric environment of the crystal. The MM4 binding energy for the formamide dimer is partly composed of a dipole-dipole interaction, which diminishes when the dielectric constant is increased. When the dielectric constant in the MM4 calculation of the dimer is increased from 1.5 (the default for gas-phase) to 4.0 (an average value often used for crystals), the MM4 O…H and O…N bond lengths increase from 1.889 to 1.938 Å and from 2.904 to 2.942 Å, respectively. Increasing the dielectric constant to that for liquid formamide44 (109) increases the hydrogen bonding distance even further. The effective dielectric constant for the formamide crystal will be different and has not been measured, and the MM4 calculations cannot simulate the crystal environment with reliable accuracy, but they do demonstrate that the nonbonded interactions in the formamide dimer are quite sensitive to the dielectric constant.

*N***-Methylacetamide Dimer.** The MM4 calculated geometries and binding energies for two dimers of the cis forms of *N*-methylacetamide, shown in Figure 3, are listed and compared to MP2 results<sup>21,23</sup> in Tables 5 and 6, respectively. NMAD1 has the same *C*<sup>2</sup>*<sup>h</sup>* symmetry and cyclic hydrogen bonding geometry as the cyclic formamide dimer FAD1. NMAD2 (*Cs* symmetry) has a heterocyclic bonding geometry, having one N-H $\cdot \cdot \cdot$ O=C' (type 28 $\cdot \cdot \cdot$ 79) and one C-H $\cdot \cdot \cdot$ O=C' (type 5...79) hydrogen bond. MP2/aug-cc-pVDZ single-point energy calculations have been done for the HF/DZP optimized geometry of a *trans-N*-methylacetamide dimer.<sup>23</sup> The dimer had  $C_1$ symmetry. Although the hydrogen bond length was reported, the relative orientation between the two monomers was not, so this dimer was not used in the MM4 study. The MM4  $N-H$  $\cdot$ O=C' hydrogen bond lengths are in fair agreement with the MP2/aug-cc-pVTZ values, differing by about  $0.02-0.07$  Å. For the  $C-H\cdots O=C'$  bond length in NMAD2, the MM4 values are much longer, by more than 0.2 Å. This interaction is a very weak one and is mainly a van der Waals interaction. If the binding energies for the NMAD2 dimer are compared to onehalf the binding energies of the NMAD1 dimer, the energies are about the same. Most of the binding energy in the NMAD2 dimer comes from the  $N-H\cdots O=C'$  interaction, and the  $C-H\cdots$ 



**Figure 5.** MM4 optimized structure and moment of inertia axes for cyclic formamide-water complex.

 $\cdot$  O=C' only contributes about 0.3 kcal mol<sup>-1</sup> to the total binding energy. Although this discrepancy between the MM4 and MP2 bond lengths for the latter interaction is much larger than desired, it arises mainly from the MM4 vdW parameters for the alkane hydrogen and does not warrant a separate nonbonded parameters to correct it.

Of more concern are the MP2 binding energies for the NMAD1 dimer. The MP2/aug-cc-pVDZ BSSE corrected value  $(-16.2 \text{ kcal mol}^{-1})$  is 1.5 kcal mol<sup>-1</sup> lower than the corresponding BSSE corrected value for the formamide dimer FAD1  $(-14.7 \text{ kcal mol}^{-1})$  at the same level of theory. A similar difference of  $1.1$  kcal mol<sup>-1</sup> is seen in their BSSE corrected MP2/6-311++G(2d,2p) binding energies. In the *<sup>N</sup>*-methylacetamide dimer, induced dipole effects from the  $CH_3-C'$  and  $CH<sub>3</sub>-N$  bonds appear to play a stabilizing role. In MM4, the NMAD1 dimer  $(-15.3 \text{ kcal mol}^{-1})$  is more stable than FAD1 by 0.6 kcal mol<sup>-1</sup>, of which 0.4 kcal mol<sup>-1</sup> can be accounted for by inductive effects alone. However, they do not entirely account for the larger difference in the MP2 energies.

**Formamide**-**Water.** Three different configurations for the formamide-water dimer were studied and are shown in Figure

**TABLE 10: Microwave and MM4 (***rz***) Geometries for Cyclic Formamide**-**Water Dimer (FAW1)***<sup>a</sup>*



*<sup>a</sup>* Refer to Figure 5. Bond lengths are in angstroms, and bond angles are in degrees. *<sup>b</sup>* Reference 33. *<sup>c</sup>* Fixed value by authors. *<sup>d</sup>* Angle of O-Hw2 bond with respect to the  $O-H_{w1} \cdots O'$  plane.





*<sup>a</sup>* Bond lengths are in angstroms, and angles are in degrees. *<sup>b</sup>* Refer to Figure 6.

**TABLE 12: Equilibrium Binding Energies (∆***E***) for** *<sup>N</sup>***-Methylacetamide**-**Water Dimers**

				$\Delta E$ (kcal/mol)				
level	optimized geometry	BSSE corr.	ret	NMAW1	NMAW <sub>2</sub>	NMAW3	NMAW4	NMAW <sub>5</sub>
$MP2/aug-cc-pVDZ$ $MP2/aug-cc-pVDZ$ $MP2/6-311++G(2d,2p)$ $MP2/6-311++G(2d,2p)$ MM4	HF/DZP HF/DZP $MP2/6-311++G(2d,2p)$ $MP2/6-311++G(2d,2p)$ MM4	no yes no yes	23 23 this work this work this work	$-5.20$ $-4.09$ $-6.28$	$-8.27$ $-7.01$ $-6.84$	$-8.24$ $-6.98$ $-6.89$	$-10.2$ $-8.9$ $-10.62$ $-8.99$ $-9.25$	$-8.0$ $-7.0$ $-8.15$ $-6.99$ $-6.22$



**Figure 6.** MM4 optimized structures of *<sup>N</sup>*-methylacetamide-water dimers.

4. The calculated equilibrium (*r*e) geometries and hydrogen bond energies are listed in Tables 7 and 8, respectively. Not surprisingly, the most stable dimer is the cyclic doubly bonded form, FAW1, having one  $C' = O \cdots H-O-H$  and one N-H $\cdots$ 

OH2 bond. This dimer has been studied experimentally in the gas phase by microwave (MW) spectroscopy.33 The MW determined moments of inertia are listed and compared to the MM4 calculated moments in Table 9. The MW determined zeropoint geometries  $(r<sub>z</sub>)$  for the complex are listed and compared to the MM4 geometries in Table 10. The MM4 structure for the cyclic formamide-water dimer, superimposed with the moment of inertia axes, is shown in Figure 5.

The MM4 calculated binding energies for the dimers are in good agreement with the BSSE corrected MP2/6-311++G(2d,- 2p) energies. The MM4 energies are somewhat lower, by 0.1- 0.8 kcal mol<sup>-1</sup>. The MM4 equilibrium  $(r_e)$  values for the H-O- $H \cdot \cdot \cdot O=C'$  bond (type 21 $\cdot \cdot \cdot$ 79) are longer than the MP2 values, by 0.06–0.12 Å, and the MM4 values for the  $H_2O^{\bullet} \rightarrow H$  (type  $6\cdots 28$ ) bonds are shorter, by 0.05–0.06 Å. The types  $21\cdots 79$ and 6…28 hydrogen bond length parameters were weighted far more heavily on the microwave data. The MM4 moments of inertia listed in Table 9 are in very good agreement with the experimental values, with errors being within  $\pm 0.6$ %. Consequently, the MM4  $r<sub>z</sub>$  geometries are in close agreement with the MW determined geometries for this complex, for which the hydrogen bond lengths differ by  $0.01-0.02$  Å. The negative value for the MM4 calculated inertial defect  $\Delta = I_c - I_a - I_b$ 

 $=$  -0.89 amu  $\AA^2$  is indicative of a nonplanar complex. It is seven times larger than the microwave value  $(-0.12 \text{ amu A}^2)$ but still relatively small. If the MM4 complex were totally planar, the defect would be an order of magnitude lower. (For a planar vibrationless system,  $\Delta \equiv 0$ .) The larger MM4 defect is due to the steeper tilt angle ( $\phi = 56^{\circ}$ ) at which the nonbonded water hydrogen  $(H_{w2})$  projects out of the O- $H_{w1}$   $\cdots$  O' plane in Figure 5. The MM4 angle is a mainly a result of the implicit  $sp<sup>3</sup>$  arrangement of the lone pairs of the water molecule in MM4. The tilt angle is much smaller in the MW geometry ( $\phi = 15^{\circ}$ ) but not nearly as much so in the MP2/6-311++ $G(2d,2p)$ optimized geometry ( $\phi = 35^{\circ}$ ).

*<sup>N</sup>***-Methylacetamide**-**Water Dimer.** The MM4 hydrogen bonding geometries and energies for five complexes of *N*methylacetamide and water, shown in Figure 6, are listed in Tables 11 and 12, respectively. Dimers NMAW1, NMAW2, and NMAW3 are water complexes with the trans form of *N*-methylacetamide, and NMAW4 and NMAW5 are complexes with the cis form. NMAW4 is the double-bonded cyclic form analogous to the formamide-water complex FAW1 described previously and is similarly the most stable. The MM4 binding energies are in agreement with the BSSE corrected MP2/6-  $311++G(2d,2p)$  energies to within 0.8 kcal mol<sup>-1</sup>, except for NMAW1, for which the MM4 energy is lower by 2.2 kcal  $mol^{-1}$ . The discrepancies in the MM4 hydrogen bond lengths relative to the MP2/6-311++ $(2d,2p)$  values show a pattern similar to that found in the formamide-water geometries in Table 7. The MM4 bond lengths show a closer agreement (by  $0.001 - 0.02$  Å) with the HF/DZP values<sup>23</sup> listed, but this is fortuitous.

It deserves mentioning that the presence of methyl groups *N*-methylacetamide does not appear to play as much of a stabilizing role in the complexation of the molecule with water, as was suspected in the case of *N*-methylacetamide dimer discussed earlier. NMAW4 and FAW1, which have the similar hydrogen bond configurations, also have similar MP2/6-  $311++G(2d,2p)$  (BSSE corrected) binding energies of  $-8.99$ and  $-8.52$  kcal mol<sup>-1</sup>, respectively. In the three analogous dimers, NMAW2, NMAW5 and FAW2, the stabilization effect of the methyl groups appears to be more pronounced, the binding energies of the first two being  $1.0$  kcal mol<sup>-1</sup> lower than that of the third. However, reverse trend is seen for the NMAW1 and FAW3 analogue pair.

#### **Conclusion**

The MM4 amide-amide and amide-water hydrogen bonding parameters were fit, for the most part, to within good agreement to ab initio calculations carried at the MP2 level using double and triple-*ú* basis sets, the highest level of theory available for these systems. Where such data were lacking, the MM4 binding energies were in satisfactory agreement with the MP2/6-  $311++G(2d,2p)$  energies. Better agreement was achieved with the microwave data for the cyclic-formamide dimer, which gives the most accurate structural information, albeit indirectly. Overall, the MM4 hydrogen bonding parameters are regarded as sufficiently optimized that they can be utilized in the ongoing MM4 study of peptides, which involves these interactions, both internally and externally.

**Acknowledgment.** The authors thank Tripos Inc. in St. Louis, MO for their financial support of this work and the University Computing and Networking Resources at the University of Georgia for the computing facilities used in this research.

#### **References and Notes**

(1) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982.

(2) The MM3 program is available to all users from Tripos Inc., 1699 South Hanley Road, St. Louis, MO 63144. Academic users, and others from nonprofit organizations, may alternatively obtain the program from the Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN 47405. The program is described in the following references: (a) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551. (b) Lii, J.-H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566, 8576.

(3) MM4 will be available from the same source in due course and is described in the following references: (a) Allinger, N. L.; Chen, K.; Lii, J.-H. *J. Comput. Chem.* **1996**, *17*, 642. (b) Nevins, N.; Chen, K.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 669. (c) Nevins, N.; Lii, J.-H.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 695. (d) Nevins, N.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 730. (e) Allinger, N. L.; Chen, K.; Katzenellenbogen, J. A.; Wilson, S. R.; Anstead, G. M. *J. Comput. Chem.* **1996**, *17*, 747.

(4) Lii, J.-H.; Allinger, N. L.; Kok, R. A.; Imam, M. R. *J. Comput. Chem.* **1988**, *9*, 591.

(5) Lii, J.-H.; Allinger, N. L. *J. Phys. Org. Chem.* **1994**, *7*, 591.

(6) Lii, J.-H.; Allinger, N. L. *J. Comput. Chem.* **1998**, *19*, 1001.

(7) Langley, C. H.; Allinger, N. L. *J. Phys. Chem. A* **2002**, *106*, 5638.

(8) Novoa, J. J.; Whangbo, M. H. *J. Am. Chem. Soc.* **1991**, *113*, 9017.

(9) Neuheuser, T.; Hess, B. A.; Reutel, C.; Weber, R. *J. Phys. Chem.* **1994**, *98*, 6459.

(10) McGrady, J. E.; Mingos, D. M. P. *J. Chem. Soc.*, *Perkins Trans. 2* **1995**, *12*, 2287.

(11) Floria´n, J.; Leszczynski, J.; Johnson, B. *J. Mol. Struct.* **1995**, *349*, 421.

(12) Floria´n, J.; Johnson, B. G. *J. Phys. Chem.* **1995**, *99*, 5899.

(13) Adalsteinsson, H.; Maulitz, A. H.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 7689.

(14) Suhai, S. *J. Phys. Chem.* **1996**, *100*, 3950.

(15) Hobza, P.; Sponer, J. *J. Mol. Struct.* (*THEOCHEM*) **1996**, *338*, 115.

(16) Hobza, P.; Havlas, Z. *Theor. Chem. Acc.* **1998**, *99*, 372.

(17) Colominas, C.; Luque, F. J.; Orozco, M. *J. Phys. Chem. A* **1999**, *103*, 6200.

(18) Sponer, J.; Hobza, P. *J. Phys. Chem. A* **2000**, *104*, 4592.

(19) Brdarski, S.; Astrand, P.-O.; Karlstrom, G. *Theor. Chem. Acc.* **2000**, *105*, 7.

(20) Kang, Y. K. *J. Phys. Chem.* **2000**, *104*, 8321.

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

(22) Tsuzuki, S.; Lu¨thi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949.

(23) Dixon, D. A.; Dobbs, K. D.; Valentini, J. J. *J. Phys. Chem.* **1994**,

*98*, 13435. (24) Qian, W.; Mirkin, N. G.; Krimm, S. *Chem. Phys. Lett.* **1999**, *315*, 125.

(25) Kang, Y. K. *J. Phys. Chem. B* **2000**, *104*, 8321.

(26) Jasien, P. G.; Stevens, W. J. *J. Chem. Phys.* **1986**, *84*, 3271.

(27) Sim, F.; Stamant, A.; Papai, I.; Salahub, D. R. *J. Am. Chem. Soc.* **1992**, *114*, 4391.

(28) Sathyan, N.; Santhanam, V.; Sobhanadri J. *J. Mol. Struct.* (*THEOCHEM*) **1995**, *333*, 179.

(29) Guo, H.; Karplus, M. *J. Phys. Chem.* **1992**, *96*, 7273.

- (30) Guo, H.; Karplus, M. *J. Phys. Chem.* **1994**, *98*, 7104.
- (31) Han, W.-G.; Suhai, S. *J. Phys. Chem.* **1996**, *100*, 3942.
- (32) Stevens, E. D. *Acta Crystallogr.* **1978**, *B34*, 544.

(33) Lovas, F. J.; Suenram, R. D.; Fraser, G. T.; Gillies, C. W.; Zozom, J. *J. Chem. Phys.* **1988**, *88*, 722.

(34) Engdahl A.; Nelander, B.; Astrand, P.-O. *J. Chem. Phys.* 1993, *99*, 4894.

(35) Triggs, N. E.; Valentini, J. J. *J. Phys. Chem.* **1992**, *96*, 6922.

(36) Triggs, N. E.; Bonn, R. T.; Valentini, J. J. *J. Phys. Chem.* **1993**, *97*, 5535.

(37) Eberhardt, E. S.; Raines, R. T. *J. Am. Chem. Soc.* **1994**, *116*, 2149.

(38) Desfrancois, C.; Periquet, V.; Carles, S.; Schermann, J. P.; Andamowicz, L. *Chem. Phys.* **1998**, *239*, 475.

(39) Lii, J.-H. Hydrogen Bonding: 2. In *The Encyclopedia of Computational Chemistry*, Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III., Schreiner, P. R., Eds.; John Wiley

& Sons: Chichester, 1998; p 1271. (40) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.;

Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski,

V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.;

Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, revision C.3; Gaussian,

Inc.: Pittsburgh, PA, 1995. (41) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(42) Langley, C. H.; Lii, J.-H.; Allinger, N. L. *J. Comput. Chem.* **2001**, *22*, 1396.

(43) Bartelli, L. S.; Kuchitsu, K.; Seip, H. M. *Acta Crystallogr.* **1976**, *32*, 1013.

(44) *CRC Handbook of Chemistry and Physics*, 74th ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, FL, 1993.