

# Supplementary material available: A semiempirical quantum model for hydrogen bonded nucleic acid base pairs

Timothy J. Giese, Edward C. Sherer, Christopher J. Cramer  
and Darrin M. York\*,

Department of Chemistry, University of Minnesota, 207 Pleasant St. SE, Minneapolis,  
MN 55455-0431, USA.

## Contents

<b>1</b>	<b>Nucleic acid base monomers</b>	<b>3</b>
<b>2</b>	<b>Nucleic acid base dimers</b>	<b>4</b>
2.1	Hydrogen bonds . . . . .	4
2.2	Binding enthalpies . . . . .	6
<b>3</b>	<b>Transferability to molecules not in the parameterization set</b>	<b>7</b>

## List of Figures

A-1 Giese <i>et al.</i> , J. Chem. Theory Comput. . . . .	11
A-2 Giese <i>et al.</i> , J. Chem. Theory Comput. . . . .	12

## List of Tables

A-I Comparison of the error in the semiempirical bond lengths relative to corresponding DFT values for nucleic acid bases <sup>a</sup> . . . . .	13
A-II Comparison of the error in the semiempirical bond angles relative to corresponding DFT values for nucleic acid bases <sup>a</sup> . . . . .	14
A-III Comparison of the error in the semiempirical torsion angles relative to corresponding DFT values for nucleic acid bases <sup>a</sup> . . . . .	15
A-IV Comparison of semiempirical and DFT dipole moments for nucleic acid bases <sup>a</sup> . . . . .	16
A-V Comparison of semiempirical and DFT hydrogen bond lengths for nucleic acid base pairs <sup>a</sup> . . . . .	17
A-VI Comparison of semiempirical and DFT hydrogen bond angles for nucleic acid bases <sup>a</sup> . . . . .	19
A-VII Comparison of semiempirical and DFT binding enthalpies for nucleic acid base dimers <sup>a</sup> . . . . .	21
A-VIII Comparison of semiempirical and DFT dimerization energies of hydrogen bonded complexes <sup>a</sup> . . . . .	22
A-IX Comparison of semiempirical and DFT dipole moments of neutral hydrogen bonded complexes and their monomers <sup>a</sup> . . . . .	24
A-X Comparison of semiempirical and DFT hydrogen-heavy atom hydrogen bond distances <sup>a</sup> . . . . .	26

# 1 Nucleic acid base monomers

The fundamental unit of a nucleic acid base pair is, of course, the nucleic acid base. Consequently, results for the internal geometry and dipole moments for cytosine, guanine, adenine, thymine and uracil nucleotide bases are briefly summarized. The geometries of the semiempirical nucleic acid bases are compared to the DFT results of Sherer *et al.*<sup>1</sup> All of the semiempirical methods perform reasonably well for the internal geometries of the base monomers.

The root-mean-square errors (RMSEs) of the semiempirical bond lengths (Table A-I) from the DFT values range from 0.015 Å (the PM3<sub>BP</sub>, MNDO and MNDO/H methods) to 0.022 and 0.021 Å for AM1 and PM3, respectively. The RMSEs for bond angles (Table A-II) are all less than 3 degrees, the smallest values (2.0 degrees) for MNDO and MNDO/H, and the largest (2.9 degrees) for PM3. The RMSE for the heavy-atom torsion angles (Table A-III) are also quite reasonable, the largest value occurring for PM3 (2.7 degrees) and the smallest values for AM1 and PM3<sub>BP</sub> (0.7 degrees). The origin of the particularly large RMSE (and MAXE magnitude) of PM3 derives mainly from the uracil structure that is predicted by PM3 to have considerable deviation from planarity for the heavy-atom torsion angles involving nitrogen.

The semiempirical dipole moments for the nucleic acid bases are compared with the density-functional calculations of Sherer *et al.*<sup>1</sup> and the high basis-set level (B3LYP/cc-pVTZ) calculations of Li *et al.*<sup>2</sup> in Table A-IV and are illustrated in Figure A-1. The

PM3<sub>BP</sub> method performs the best with respect to the high basis DFT results, with a mean signed error of -0.002 D and an RMSE of 0.101 D. The MNDO/H method has the largest RMSE (0.717 D) of the semiempirical methods. It is of interest to note that the DFT dipole moments calculated with the smaller basis set (*m*PWPW) have an RMSE (0.761 D) with respect to the higher basis-set (B3LYP/cc-pVTZ) values that is larger than any of the semiempirical RMSE values.

## 2 Nucleic acid base dimers

### 2.1 Hydrogen bonds

Nucleic acid base pairs can form dimers in a variety of ways.<sup>3–6</sup> A detailed discussion of nucleic acid base pairing derived from density-functional calculations has been presented elsewhere.<sup>1</sup> The fundamental intermolecular interactions that stabilize nucleic acid base pairs are hydrogen bonding and base stacking.<sup>7</sup> The present work focuses on the problem of intermolecular hydrogen bonding between nucleic acid bases for which reasonably reliable *ab initio* data are available. Future work will address the problem of base stacking, which will require the introduction of a dispersion energy correction.<sup>8–10</sup> Tables A-V and A-VI compare the density-functional and semiempirical hydrogen bond lengths and angles, respectively, for 31 types of nucleic acid base pairs.

The standard MNDO method does not predict stable hydrogen bonding. This is a well-known problem, and there has been considerable effort to develop new semiempirical forms that correct this deficiency.<sup>11,12</sup> It is not the purpose here to provide a com-

prehensive comparison of all the semiempirical methods. The methods compared here all share a common form.<sup>13</sup> The AM1 method predicts hydrogen bond lengths systematically too long compared with the density-functional values with a mean signed error (MSE) of 0.571 Å and RMSE value of 0.665 Å. The PM3 method performs considerably better for hydrogen bond lengths with MSE and RMSE values of 0.084 and 0.165 Å, respectively. The MNDO method predicts unreasonably long hydrogen bond lengths, whereas the MNDO/H method corrects this so that the MSE and RMSE values are -0.250 and 0.310 Å, respectively. The PM3<sub>BP</sub> method agrees most closely with the density-functional results with MSE and RMSE values of 0.006 and 0.151 Å, respectively.

Results for the hydrogen bond angles (Table A-VI) are qualitatively similar in that errors are most significant for the MNDO and AM1 methods. The AM1 method predicts hydrogen bonds to be too bent, having large MSE and RMSE values of -18.5 and 23.3 degrees, respectively. The hydrogen bond angles are considerably improved with the PM3, MNDO/H and PM3<sub>BP</sub> methods. The MSE and RMSE values for hydrogen bond angles are -3.8 and 5.5 degrees, respectively, for PM3, -1.9 and 6.0 degrees, respectively, for MNDO/H and -3.4 and 5.5 degrees, respectively, for the PM3<sub>BP</sub> method. Overall, for intermolecular hydrogen bond geometries, the PM3, MNDO/H and PM3<sub>BP</sub> method perform quite well, with the PM3<sub>BP</sub> method slightly better overall. As is demonstrated in the next section, however, that the AM1, PM3 and MNDO methods all severely underestimate the binding enthalpies of the base pairs, and the MNDO/H method overestimates the binding enthalpies.

## 2.2 Binding enthalpies

Table A-VII compares the binding enthalpies of 31 nucleic acid base pairs calculated with density-functional theory<sup>1</sup> and semiempirical models. The binding enthalpies reflect not only the number and stability of local hydrogen bonds, but also intermolecular dipole-dipole interactions, and to a lesser extent dispersion effects. Dispersion effects are more important for base stacking interactions and for proper folding of DNA and RNA in polar environments.<sup>14</sup> The most strongly bound dimer is the CG Watson-Crick base pair, which is the only base pair with 3 strong hydrogen bonds and has a DFT (*mPWPW*) binding enthalpy of -22.4 kcal/mol. All other base pairs have only 2 hydrogen bonds (with the exception of GG<sub>1</sub> that has 2 strong and 2 very weak hydrogen bonds), and the DFT binding enthalpies range from -7.5 to -17.6 kcal/mol. The MNDO method performs very poorly as discussed earlier. In general, the AM1 and PM3 methods underestimate the DFT binding enthalpy by over 5 kcal/mol (MSE values of 5.1 and 5.5 kcal/mol, respectively) leading to RMSE values of 5.7 and 5.8 kcal/mol, respectively. Relative errors for these methods are typically in the range of 30-60% corresponding to a factor of roughly 1.5-3. The MNDO/H method, on the other hand, over-predicts the binding enthalpy (MSE value of -6.2 kcal/mol) and leads to an RMSE value of 6.5 kcal/mol. The PM3<sub>BP</sub> method performs significantly better (Figure A-2). The MSE is -0.7 kcal/mol, and the RMSE (1.3 kcal/mol) is reduced by over a factor of 4 with respect to AM1 and PM3 models. For Watson-Crick base pairing, the PM3<sub>BP</sub> method predicts -21.4, -12.4 and -12.6 kcal/mol

for CG, AT and AU<sub>WC</sub> base pairs, respectively; these values compare well (and retain the same trend in order of stability) with the DFT values of -22.4, -11.3 and -11.4 kcal/mol, respectively. These results offer promise that semiempirical methods might be developed that have predictive value for biological hydrogen bonded systems.

### **3 Transferability to molecules not in the parameterization set**

Tables A-VIII-A-X provide the complete set of dimerization enthalpies, dipole moments, and hydrogen bond lengths for molecules not considered in the parameterization procedure to examine the transferability of the PM3<sub>BP</sub> method.

## References

- [1] Sherer, E. C.; York, D. M.; Cramer, C. J. *J. Comput. Chem.* **2003**, *24*, 57–67.
- [2] Li, J.; Xing, J.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **1999**, *111*, 885–892.
- [3] Saenger, W. *Principles of nucleic acid structure*. Springer-Verlag, New York, **1984**.
- [4] Donohue, J.; Trueblood, K. N. *J. Mol. Biol.* **1960**, *2*, 363–371.
- [5] Poltev, V. I.; Shulyupina, N. V. *J. Biomol. Struct. Dyn.* **1986**, *3*, 739–765.
- [6] Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, *99*, 3247–3276.
- [7] Maki, A.; Brownell, F. E.; Liu, D.; Kool, E. T. *Nucleic Acids Res.* **2003**, *31*, 1059–1066.
- [8] Šponer, J.; Leszczyński, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590–5596.
- [9] Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149–5155.
- [10] Giese, T. J.; York, D. M. *in preparation*.
- [11] Thiel, W. In *Adv. Chem. Phys.*, Prigogine, I.; Rice, S. A., *edts*, volume 93, 703–757.  
John Wiley and Sons, New York **1996**.
- [12] Thiel, W. In *Handbook of Molecular Physics and Quantum Chemistry*, Wilson, S., *edt*, volume 2 (John Wiley and Sons, Chichester, 2003).

[13] The MNDO/d, AM1, PM3 and MNDO/H methods all share a common general form with the only difference involving the specific treatment of core-core terms. In this way, we have restricted the scope of our comparison to methods of this form.

[14] Šponer, J. E.; Špačková, N.; Kulhánek, P.; Leszczynski, J.; Šponer, J. *J. Phys. Chem. A* **2005**, *109*, 2292–2301.

- Figure A-2: Regression of semiempirical binding enthalpies for nucleic acid base dimers with DFT *mPWPW*<sup>1</sup> (x-axis reference) values. A linear fit for each method produces intercept (b), slope (m), and correlation coefficient (c) values of ( $\text{PM3}_{BP}$ :  
 $b=-0.521 \text{ kcal/mol}$ ,  $m=0.983$ ,  $c=0.963$ ), ( $\text{AM1}$ :  $b=-0.970 \text{ kcal/mol}$ ,  $m=0.971$ ,  $c=0.934$ ),  
( $\text{PM3}$ :  $b=0.810 \text{ kcal/mol}$ ,  $m=0.615$ ,  $c=0.670$ ), ( $\text{MNDO}$ :  $b=1.013 \text{ kcal/mol}$ ,  $m=0.596$ ,  
 $c=0.892$ ), ( $\text{MNDO/H}$ :  $b=1.833$ ,  $m=0.273$ ,  $c=0.734$ ).
- Figure A-1: Regression of semiempirical and DFT *mPWPW*<sup>1</sup> dipole moments for nucleic acid bases with B3LYP/cc-pVTZ<sup>2</sup> (x-axis reference) values. A linear fit for each method produces intercept (b), slope (m), and correlation coefficient (c) values of (DFT:  $b=0.253 \text{ D}$ ,  $m=0.809$ ,  $c=0.982$ ), ( $\text{PM3}_{BP}$ :  $b=0.081 \text{ D}$ ,  $m=0.981$ ,  
 $c=1.000$ ), ( $\text{AM1}$ :  $b=0.116 \text{ D}$ ,  $m=0.918$ ,  $c=0.996$ ), ( $\text{PM3}$ :  $b=0.289 \text{ D}$ ,  $m=0.818$ ,  
 $c=0.988$ ), ( $\text{MNDO}$ :  $b=0.349 \text{ D}$ ,  $m=0.807$ ,  $c=0.985$ ), ( $\text{MNDO/H}$ :  $b=0.367$ ,  $m=0.805$ ,  
 $c=0.985$ ).

Figure A-1: Giese *et al.*, J. Chem. Theory Comput.

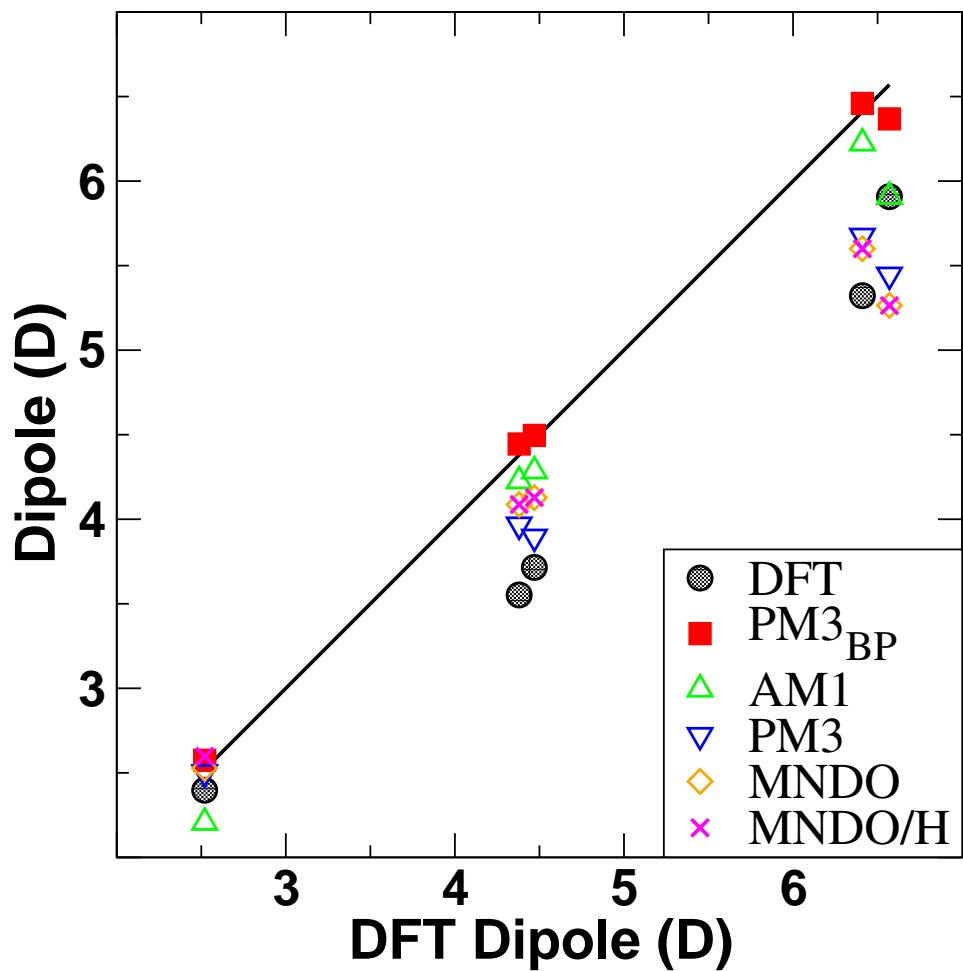


Figure A-2: Giese *et al.*, J. Chem. Theory Comput.

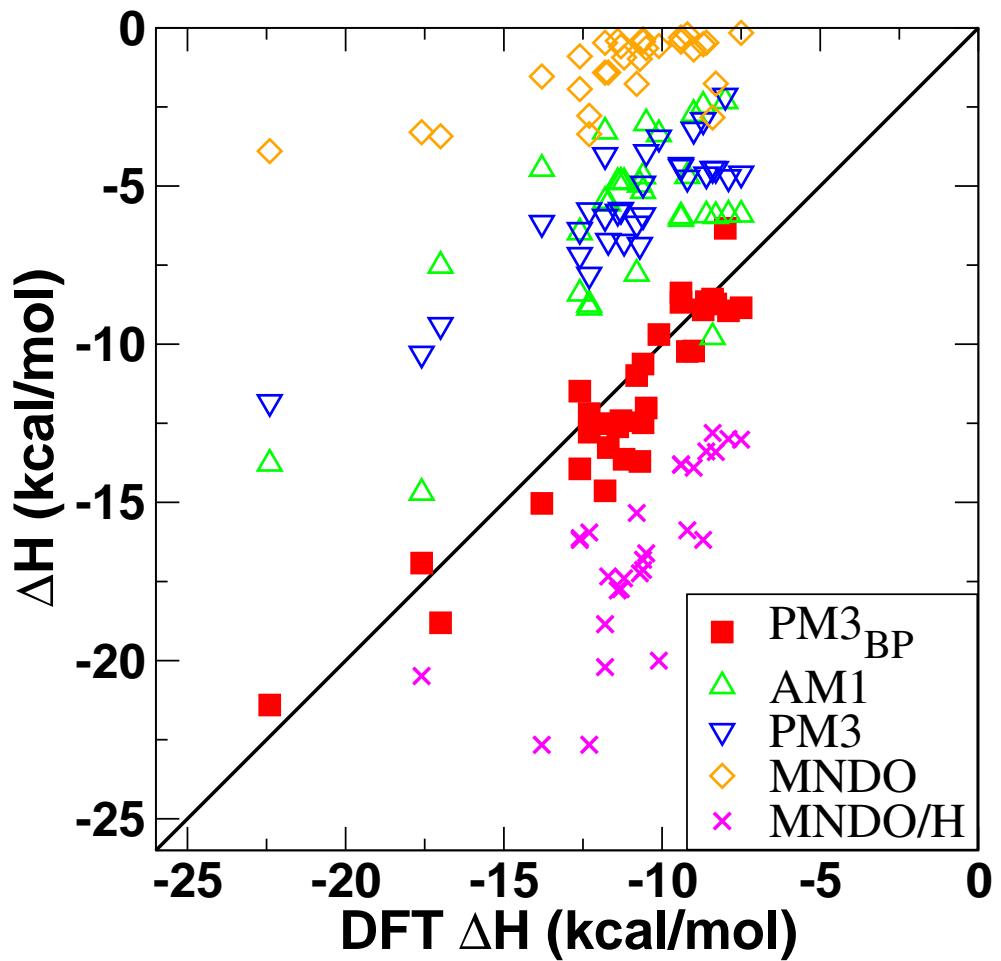


Table A-I: Comparison of the error in the semiempirical bond lengths relative to corresponding DFT values for nucleic acid bases<sup>a</sup>

Molecule	PM3 <sub>BP</sub>	AM1	PM3	MNDO	MNDO/H
<i>A</i>	0.016	0.024	0.020	0.016	0.016
<i>T</i>	0.012	0.017	0.017	0.012	0.012
<i>C</i>	0.014	0.021	0.022	0.017	0.018
<i>G</i>	0.017	0.026	0.024	0.017	0.017
<i>U</i>	0.013	0.018	0.018	0.013	0.013
<b>MSE</b>	<b>0.000</b>	<b>-0.003</b>	<b>-0.004</b>	<b>-0.003</b>	<b>-0.003</b>
<b>MUE</b>	<b>0.012</b>	<b>0.017</b>	<b>0.016</b>	<b>0.011</b>	<b>0.011</b>
<b>RMSE</b>	<b>0.015</b>	<b>0.022</b>	<b>0.021</b>	<b>0.015</b>	<b>0.015</b>
<b>MAXE</b>	<b>-0.029</b>	<b>0.046</b>	<b>-0.043</b>	<b>-0.035</b>	<b>-0.035</b>

<sup>a</sup> Comparison of bond lengths ( $\text{\AA}$ ) for nucleic acid base (monomers) from semiempirical (PM3<sub>BP</sub>, AM1, PM3, MNDO and MNDO/H) and DFT (*mPWPW*)<sup>1</sup> calculations. Shown are the average errors for each individual base, and summarized at the bottom are the error metrics (bold) for the semiempirical results with respect to the corresponding DFT *mPWPW* results of Sherer *et al.*<sup>1</sup>

Table A-II: Comparison of the error in the semiempirical bond angles relative to corresponding DFT values for nucleic acid bases<sup>a</sup>

Molecule	PM3 <sub>BP</sub>	AM1	PM3	MNDO	MNDO/H
<i>A</i>	2.6	2.1	2.5	2.0	1.9
<i>T</i>	2.5	2.5	3.0	2.0	2.0
<i>C</i>	2.1	2.9	2.2	2.2	2.2
<i>G</i>	2.8	2.5	2.8	1.9	1.9
<i>U</i>	2.9	3.0	3.5	2.1	2.1
<b>MSE</b>	<b>0.3</b>	<b>0.4</b>	<b>0.1</b>	<b>0.0</b>	<b>0.0</b>
<b>MUE</b>	<b>2.2</b>	<b>1.9</b>	<b>2.3</b>	<b>1.6</b>	<b>1.6</b>
<b>RMSE</b>	<b>2.6</b>	<b>2.6</b>	<b>2.9</b>	<b>2.0</b>	<b>2.0</b>
<b>MAXE</b>	<b>-6.4</b>	<b>7.6</b>	<b>-7.9</b>	<b>-6.0</b>	<b>-5.7</b>

<sup>a</sup> Comparison of bond angles (degrees) for nucleic acid base (monomers) from semiempirical (PM3<sub>BP</sub>, AM1, PM3, MNDO and MNDO/H) and DFT *mPWPW*<sup>1</sup> calculations. Shown are the average errors for each individual base, and summarized at the bottom are the error metrics (bold) for the semiempirical results with respect to the corresponding DFT *mPWPW* results of Sherer *et al.*<sup>1</sup>

Table A-III: Comparison of the error in the semiempirical torsion angles relative to corresponding DFT values for nucleic acid bases<sup>a</sup>

Molecule	PM3 <sub>BP</sub>	AM1	PM3	MNDO	MNDO/H
<i>A</i>	0.7	0.9	1.0	1.3	1.2
<i>T</i>	0.1	0.3	0.5	0.6	0.6
<i>C</i>	1.1	0.5	1.1	1.5	1.5
<i>G</i>	0.9	0.9	1.1	1.1	1.1
<i>U</i>	0.2	0.1	6.8	0.1	0.1
<b>MSE</b>	<b>0.0</b>	<b>-0.0</b>	<b>-0.1</b>	<b>0.0</b>	<b>0.0</b>
<b>MUE</b>	<b>0.5</b>	<b>0.4</b>	<b>1.3</b>	<b>0.7</b>	<b>0.7</b>
<b>RMSE</b>	<b>0.7</b>	<b>0.7</b>	<b>2.7</b>	<b>1.1</b>	<b>1.1</b>
<b>MAXE</b>	<b>2.3</b>	<b>-2.7</b>	<b>-12.0</b>	<b>3.4</b>	<b>3.2</b>

<sup>a</sup> Comparison of heavy atom (non-hydrogen) torsion angles (degrees) for nucleic acid base (monomers) from semiempirical (PM3<sub>BP</sub>, AM1, PM3, MNDO and MNDO/H) and DFT *mPWPW*<sup>1</sup> calculations. Shown are the average errors for each individual base, and summarized at the bottom are the error metrics (bold) for the semiempirical results with respect to the corresponding DFT *mPWPW* results of Sherer *et al.*<sup>1</sup>

Table A-IV: Comparison of semiempirical and DFT dipole moments for nucleic acid bases<sup>a</sup>

Molecule	B3LYP	<i>m</i> PWPW	PM3 <sub>BP</sub>	AM1	PM3	MNDO	MNDO/H
<i>A</i>	2.520	2.397	2.574	2.206	2.502	2.528	2.597
<i>T</i>	4.380	3.552	4.445	4.226	3.967	4.087	4.087
<i>C</i>	6.410	5.322	6.459	6.224	5.675	5.600	5.600
<i>G</i>	6.570	5.908	6.368	5.908	5.445	5.264	5.264
<i>U</i>	4.470	3.715	4.496	4.285	3.895	4.128	4.128
<b>MSE</b>		<b>-0.691</b>	<b>-0.002</b>	<b>-0.300</b>	<b>-0.573</b>	<b>-0.548</b>	<b>-0.566</b>
<b>MUE</b>		<b>0.691</b>	<b>0.079</b>	<b>0.300</b>	<b>0.573</b>	<b>0.552</b>	<b>0.535</b>
<b>RMSE</b>		<b>0.761</b>	<b>0.101</b>	<b>0.355</b>	<b>0.680</b>	<b>0.716</b>	<b>0.717</b>
<b>MAXE</b>		<b>-1.088</b>	<b>-0.202</b>	<b>-0.662</b>	<b>-1.125</b>	<b>-1.306</b>	<b>-1.306</b>

<sup>a</sup> Comparison of dipole moments (Debye) for nucleic acid base (monomers) from from semiempirical (PM3<sub>BP</sub>, AM1, PM3, MNDO and MNDO/H) and DFT *m*PWPW<sup>1</sup> (*m*PWPW) and B3LYP/cc-pVTZ<sup>2</sup> (B3LYP) calculations. Shown are the magnitude of the dipole moments for each individual base, and summarized at the bottom are the error metrics (bold) for the semiempirical and DFT (*m*PWPW) results of Sherer *et al.*<sup>1</sup> with respect to the corresponding high basis-set level DFT (B3LYP) calculations of Li *et al.*<sup>2</sup>

Table A-V: Comparison of semiempirical and DFT hydrogen bond lengths for nucleic acid base pairs<sup>a</sup>

Molecule	<i>m</i> PWPW	PM3 <sub>BP</sub>	AM1	PM3	MNDO	MNDO/H
<i>CG</i> <sub>WC</sub>	1.624	1.723	2.060	1.805	3.209	1.446
<i>CG</i> <sub>WC</sub>	1.821	1.761	2.105	1.837	3.088	1.447
<i>CG</i> <sub>WC</sub>	1.761	1.719	2.073	1.778	3.284	1.565
<i>CG</i> <sub>1</sub>	1.734	1.720	2.799	1.796	3.642	1.403
<i>CG</i> <sub>1</sub>	1.829	1.726	3.421	1.815	4.019	1.432
<i>CG</i> <sub>2</sub>	2.002	1.792	2.760	1.879	3.675	1.560
<i>CG</i> <sub>2</sub>	1.864	1.766	2.219	1.827	3.825	1.597
<i>AT</i> <sub>WC</sub>	1.631	1.695	2.167	1.781	3.287	1.414
<i>AT</i> <sub>WC</sub>	1.784	1.742	2.101	1.823	3.719	1.493
<i>AT</i> <sub>RWC</sub>	1.630	1.696	2.084	1.783	3.317	1.421
<i>AT</i> <sub>RWC</sub>	1.827	1.741	2.122	1.817	3.855	1.498
<i>AT</i> <sub>H</sub>	1.634	1.677	2.056	1.766	3.190	1.443
<i>AT</i> <sub>H</sub>	1.852	1.748	2.145	1.828	3.897	1.527
<i>AT</i> <sub>RH</sub>	1.631	1.681	2.475	1.767	3.159	1.449
<i>AT</i> <sub>RH</sub>	1.894	1.749	2.134	1.826	3.868	1.537
<i>CC</i>	1.783	1.711	2.353	1.797	3.287	1.399
<i>CC</i>	1.782	1.711	2.330	1.794	3.130	1.399
<i>TT</i> <sub>1</sub>	1.718	1.729	2.106	1.800	3.283	1.512
<i>TT</i> <sub>1</sub>	1.709	1.738	2.107	1.805	3.361	1.519
<i>TT</i> <sub>2</sub>	1.692	1.735	2.092	1.805	3.251	1.501
<i>TT</i> <sub>2</sub>	1.689	1.735	2.116	1.804	3.364	1.501
<i>TT</i> <sub>3</sub>	1.733	1.734	2.111	1.800	7.778	1.522
<i>TT</i> <sub>3</sub>	1.733	1.734	2.111	1.800	7.683	1.523
<i>UU</i> <sub>1</sub>	1.712	1.732	2.099	1.807	6.977	1.521
<i>UU</i> <sub>1</sub>	1.722	1.736	2.118	1.799	7.038	1.512
<i>UU</i> <sub>2</sub>	1.689	1.733	2.104	1.801	4.208	1.509
<i>UU</i> <sub>2</sub>	1.690	1.733	2.101	1.805	3.724	1.509
<i>UU</i> <sub>3</sub>	1.743	1.734	2.111	1.801	7.425	1.518
<i>UU</i> <sub>3</sub>	1.742	1.735	2.117	1.801	7.173	1.519
<i>GU</i> <sub>1</sub>	1.652	1.726	2.038	1.794	3.658	1.500
<i>GU</i> <sub>1</sub>	1.656	1.730	2.122	1.801	5.495	1.500
<i>GU</i> <sub>2</sub>	1.694	1.727	2.048	1.790	3.740	1.507
<i>GU</i> <sub>2</sub>	1.683	1.732	2.125	1.806	5.630	1.508
<i>GG</i> <sub>1</sub>	2.246	2.977	2.411	3.066	4.605	2.903
<i>GG</i> <sub>1</sub>	1.620	1.721	2.109	1.802	3.287	1.486
<i>GG</i> <sub>1</sub>	1.621	1.724	2.098	1.790	3.323	1.482
<i>GG</i> <sub>1</sub>	2.245	2.985	2.534	2.975	3.496	2.894
<i>GG</i> <sub>3</sub>	1.714	1.738	2.480	1.816	3.330	1.519
<i>GG</i> <sub>3</sub>	2.156	2.448	2.138	2.481	3.129	1.631
<i>GG</i> <sub>4</sub>	1.829	1.739	2.583	1.818	3.692	1.419
<i>GG</i> <sub>4</sub>	1.830	1.743	2.603	1.821	3.612	1.415

*Table A-V (continued)*

Molecule	<i>mPWPW</i>	$\text{PM3}_{BP}$	AM1	PM3	MNDO	MNDO/H
$AG_1$	1.687	1.737	2.106	1.819	3.534	1.491
$AG_1$	1.793	1.710	2.385	1.774	3.519	1.446
$AG_2$	1.818	1.719	2.598	1.811	3.462	1.453
$AG_2$	1.838	1.743	2.608	1.846	3.764	1.474
$AG_3$	1.778	1.742	2.136	1.821	3.524	1.537
$AG_3$	1.710	1.705	2.335	1.789	3.155	1.458
$AG_4$	1.785	1.728	2.552	1.812	3.801	1.412
$AG_4$	1.827	1.728	2.623	1.836	3.401	1.425
$CU_1$	1.707	1.732	2.091	1.810	3.986	1.483
$CU_1$	1.738	1.721	2.588	1.806	3.338	1.419
$CU_2$	1.744	1.736	2.102	1.811	3.788	1.495
$CU_2$	1.759	1.723	2.589	1.808	3.371	1.429
$AU_{WC}$	1.630	1.695	2.051	1.780	3.245	1.416
$AU_{WC}$	1.777	1.743	2.121	1.818	4.119	1.490
$AC_1$	1.747	1.717	2.428	1.803	3.254	1.419
$AC_1$	1.768	1.712	2.585	1.805	3.518	1.408
$AC_2$	1.779	1.742	2.591	1.829	3.364	1.466
$AC_2$	1.806	1.704	4.237	1.802	3.471	1.447
$AA_1$	1.783	1.722	2.608	1.817	3.458	1.425
$AA_1$	1.783	1.722	2.544	1.813	3.575	1.425
$AA_2$	1.819	1.737	2.553	1.828	3.691	1.471
$AA_2$	1.807	1.713	2.626	1.807	3.308	1.460
$AA_3$	1.849	1.736	2.574	1.835	3.799	1.506
$AA_3$	1.849	1.737	2.660	1.837	3.328	1.507
<b>MSE</b>	<b>0.006</b>	<b>0.571</b>	<b>0.084</b>	<b>2.174</b>	<b>-0.250</b>	
<b>MUE</b>	<b>0.084</b>	<b>0.571</b>	<b>0.095</b>	<b>2.174</b>	<b>0.290</b>	
<b>RMSE</b>	<b>0.151</b>	<b>0.665</b>	<b>0.165</b>	<b>2.479</b>	<b>0.310</b>	
<b>MAXE</b>	<b>0.738</b>	<b>2.431</b>	<b>0.820</b>	<b>6.044</b>	<b>0.657</b>	

<sup>a</sup> Comparison of hydrogen bond lengths (Å) for nucleic acid base pairs from semiempirical ( $\text{PM3}_{BP}$ , AM1, PM3, MNDO and MNDO/H) and DFT *mPWPW*<sup>1</sup> (*mPWPW*) calculations. Shown are the hydrogen bond lengths for each base pair, and summarized at the bottom are the error metrics (bold) for the semiempirical values with respect to the corresponding DFT *mPWPW* results of Sherer *et al.*<sup>1</sup>

Table A-VI: Comparison of semiempirical and DFT hydrogen bond angles for nucleic acid bases<sup>a</sup>

Molecule	<i>m</i> PWPW	PM3 <sub>BP</sub>	AM1	PM3	MNDO	MNDO/H
<i>CG<sub>WC</sub></i>	176.3	174.0	163.7	173.3	141.5	176.9
<i>CG<sub>WC</sub></i>	175.4	176.1	175.5	176.6	170.4	175.0
<i>CG<sub>WC</sub></i>	176.0	177.8	174.7	176.2	173.3	173.2
<i>CG<sub>1</sub></i>	177.8	174.3	93.1	172.1	102.7	178.0
<i>CG<sub>1</sub></i>	177.0	175.8	146.8	177.8	167.1	176.8
<i>CG<sub>2</sub></i>	176.9	177.7	136.0	177.1	126.6	177.1
<i>CG<sub>2</sub></i>	174.0	168.2	138.5	168.0	112.8	167.7
<i>AT<sub>WC</sub></i>	178.8	176.2	177.7	175.7	172.5	179.0
<i>AT<sub>WC</sub></i>	178.4	175.0	176.9	177.1	163.7	172.9
<i>AT<sub>RWC</sub></i>	179.0	176.0	173.7	174.7	170.9	177.7
<i>AT<sub>RWC</sub></i>	177.2	175.6	170.0	177.2	159.0	173.2
<i>AT<sub>H</sub></i>	173.6	175.8	171.4	175.6	169.0	175.7
<i>AT<sub>H</sub></i>	168.8	165.4	153.1	164.8	150.8	173.3
<i>AT<sub>RH</sub></i>	174.2	176.2	176.5	176.7	165.7	177.8
<i>AT<sub>RH</sub></i>	169.5	164.7	147.3	163.9	157.0	173.3
<i>CC</i>	174.8	175.8	157.7	175.0	148.5	177.8
<i>CC</i>	174.8	175.8	160.5	175.3	146.1	177.8
<i>TT<sub>1</sub></i>	178.7	171.7	160.4	170.9	155.9	171.3
<i>TT<sub>1</sub></i>	178.0	170.6	161.0	169.0	151.9	169.1
<i>TT<sub>2</sub></i>	179.4	170.8	162.2	170.8	155.4	172.6
<i>TT<sub>2</sub></i>	180.0	170.8	161.4	170.8	153.0	172.6
<i>TT<sub>3</sub></i>	177.3	171.0	160.3	169.4	151.1	168.8
<i>TT<sub>3</sub></i>	177.3	171.0	160.3	169.4	153.3	168.8
<i>UU<sub>1</sub></i>	178.9	170.6	159.8	171.2	57.5	170.4
<i>UU<sub>1</sub></i>	178.0	170.6	161.4	169.3	42.1	169.3
<i>UU<sub>2</sub></i>	179.9	170.3	160.9	170.7	113.7	170.7
<i>UU<sub>2</sub></i>	179.9	170.3	160.9	170.5	129.0	170.7
<i>UU<sub>3</sub></i>	177.0	170.6	159.9	169.7	90.8	169.6
<i>UU<sub>3</sub></i>	177.0	170.6	159.9	169.8	81.1	169.6
<i>GU<sub>1</sub></i>	171.6	171.6	148.8	172.4	101.7	173.3
<i>GU<sub>1</sub></i>	177.9	171.7	172.6	171.8	104.9	171.3
<i>GU<sub>2</sub></i>	172.6	172.3	151.0	171.7	101.5	172.6
<i>GU<sub>2</sub></i>	178.8	171.8	174.5	170.8	106.7	170.0
<i>GG<sub>1</sub></i>	137.9	125.4	140.3	126.4	122.9	120.8
<i>GG<sub>1</sub></i>	162.4	172.1	136.8	165.8	158.4	172.3
<i>GG<sub>1</sub></i>	162.4	171.5	143.4	163.4	95.8	172.3
<i>GG<sub>1</sub></i>	137.9	125.4	140.2	128.4	125.5	120.6
<i>GG<sub>3</sub></i>	169.7	171.8	124.3	171.1	113.7	173.9
<i>GG<sub>3</sub></i>	164.7	169.0	153.7	166.8	157.8	174.0
<i>GG<sub>4</sub></i>	175.9	173.7	145.0	173.9	135.6	178.3
<i>GG<sub>4</sub></i>	176.2	173.0	160.4	172.4	161.5	178.6

Table A-VI (continued)

Molecule	<i>m</i> PWPW	PM3 <sub>BP</sub>	AM1	PM3	MNDO	MNDO/H
<i>AG</i> <sub>1</sub>	178.8	174.5	171.0	175.2	166.7	170.3
<i>AG</i> <sub>1</sub>	177.6	174.2	151.0	173.3	147.5	174.9
<i>AG</i> <sub>2</sub>	174.1	169.4	148.1	167.0	128.6	173.4
<i>AG</i> <sub>2</sub>	170.4	164.5	143.1	164.1	171.2	171.5
<i>AG</i> <sub>3</sub>	168.5	166.4	146.6	166.6	171.5	177.6
<i>AG</i> <sub>3</sub>	172.2	175.1	150.5	175.5	176.4	176.6
<i>AG</i> <sub>4</sub>	177.9	175.4	143.9	175.0	139.2	179.4
<i>AG</i> <sub>4</sub>	176.7	175.1	159.2	178.7	174.0	176.9
<i>CU</i> <sub>1</sub>	178.3	170.7	157.3	170.6	151.8	172.9
<i>CU</i> <sub>1</sub>	174.3	169.7	152.6	169.3	165.0	174.4
<i>CU</i> <sub>2</sub>	179.6	171.9	160.4	172.0	145.5	172.6
<i>CU</i> <sub>2</sub>	172.8	169.9	153.4	168.6	163.0	173.7
<i>AU</i> <sub>WC</sub>	178.8	175.7	174.6	175.4	162.6	178.3
<i>AU</i> <sub>WC</sub>	178.4	174.9	170.7	176.5	165.0	173.4
<i>AC</i> <sub>1</sub>	175.7	175.5	169.9	177.2	154.0	177.3
<i>AC</i> <sub>1</sub>	179.9	178.0	167.7	179.2	140.4	179.5
<i>AC</i> <sub>2</sub>	169.4	166.0	115.0	165.3	145.1	169.3
<i>AC</i> <sub>2</sub>	175.4	172.4	154.1	169.5	174.1	172.9
<i>AA</i> <sub>1</sub>	178.8	177.6	166.2	178.0	171.1	178.7
<i>AA</i> <sub>1</sub>	178.8	177.9	169.5	178.0	163.4	178.6
<i>AA</i> <sub>2</sub>	169.5	163.4	136.6	164.0	154.3	168.7
<i>AA</i> <sub>2</sub>	175.9	171.8	156.1	170.8	173.0	173.6
<i>AA</i> <sub>3</sub>	168.3	159.9	136.2	158.8	157.7	168.1
<i>AA</i> <sub>3</sub>	168.3	160.3	140.5	159.7	159.0	168.2
<b>MSE</b>	<b>-3.4</b>	<b>-18.5</b>	<b>-3.8</b>	<b>-30.3</b>	<b>-1.9</b>	
<b>MUE</b>	<b>4.5</b>	<b>18.7</b>	<b>4.5</b>	<b>30.6</b>	<b>4.4</b>	
<b>RMSE</b>	<b>5.5</b>	<b>23.3</b>	<b>5.5</b>	<b>42.5</b>	<b>6.0</b>	
<b>MAXE</b>	<b>-12.4</b>	<b>-84.7</b>	<b>-11.4</b>	<b>-136.0</b>	<b>-17.3</b>	

<sup>a</sup> Comparison of hydrogen bond angles (degrees) for nucleic acid base pairs from semiempirical (PM3<sub>BP</sub>, AM1, PM3, MNDO and MNDO/H) and DFT *m*PWPW<sup>1</sup> (*m*PWPW) calculations. Shown are the hydrogen bond angles for each base pair, and summarized at the bottom are the error metrics (bold) for the semiempirical values with respect to the corresponding DFT *m*PWPW results of Sherer *et al.*<sup>1</sup>

Table A-VII: Comparison of semiempirical and DFT binding enthalpies for nucleic acid base dimers<sup>a</sup>

Molecule	<i>m</i> PWPW	PM3 <sub>BP</sub>	AM1	PM3	MNDO	MNDO/H
<i>CG</i> <sub>WC</sub>	-22.4	-21.4	-13.8	-11.8	-3.9	-29.2
<i>CG</i> <sub>1</sub>	-12.3	-12.2	-8.8	-5.8	-2.8	-22.7
<i>CG</i> <sub>2</sub>	-8.4	-8.6	-9.8	-4.5	-2.8	-12.8
<i>AT</i> <sub>WC</sub>	-11.3	-12.4	-4.9	-5.8	-0.6	-17.7
<i>AT</i> <sub>RWC</sub>	-10.6	-12.5	-4.7	-5.9	-0.4	-17.1
<i>AT</i> <sub>H</sub>	-11.2	-13.6	-4.9	-6.8	-0.9	-17.4
<i>AT</i> <sub>RH</sub>	-10.7	-13.7	-5.0	-6.9	-1.0	-17.2
<i>CC</i>	-17.0	-18.8	-7.5	-9.4	-3.4	-26.6
<i>TT</i> <sub>1</sub>	-8.6	-8.7	-5.9	-4.6	-0.5	-13.4
<i>TT</i> <sub>2</sub>	-9.4	-8.4	-6.0	-4.4	-0.4	-13.8
<i>TT</i> <sub>3</sub>	-7.9	-8.9	-5.9	-4.7	0.2	-13.0
<i>UU</i> <sub>1</sub>	-8.3	-8.7	-6.0	-4.5	-1.8	-13.4
<i>UU</i> <sub>2</sub>	-9.4	-8.7	-6.0	-4.3	-0.3	-13.8
<i>UU</i> <sub>3</sub>	-7.5	-8.8	-5.9	-4.6	-0.2	-13.0
<i>GU</i> <sub>1</sub>	-12.6	-11.5	-8.4	-6.4	-1.9	-16.1
<i>GU</i> <sub>2</sub>	-10.8	-11.0	-7.8	-6.2	-1.8	-15.3
<i>GG</i> <sub>1</sub>	-17.6	-16.9	-14.7	-10.3	-3.3	-20.5
<i>GG</i> <sub>3</sub>	-12.3	-12.8	-8.8	-7.8	-3.4	-16.0
<i>GG</i> <sub>4</sub>	-8.0	-6.3	-2.3	-2.1	0.4	-18.8
<i>AG</i> <sub>1</sub>	-12.6	-13.9	-6.5	-7.2	-0.9	-16.2
<i>AG</i> <sub>2</sub>	-8.7	-8.9	-2.5	-2.9	-0.5	-16.2
<i>AG</i> <sub>3</sub>	-11.7	-13.3	-5.6	-6.7	-1.4	-17.3
<i>AG</i> <sub>4</sub>	-10.1	-9.7	-3.4	-3.5	-0.6	-20.0
<i>CU</i> <sub>1</sub>	-10.6	-10.6	-5.2	-4.9	-0.4	-16.8
<i>CU</i> <sub>2</sub>	-9.2	-10.2	-4.7	-4.7	-0.2	-15.9
<i>AU</i> <sub>WC</sub>	-11.4	-12.6	-4.9	-5.8	-0.4	-17.8
<i>AC</i> <sub>1</sub>	-13.8	-15.0	-4.5	-6.2	-1.5	-22.7
<i>AC</i> <sub>2</sub>	-11.8	-14.6	-5.5	-6.0	-1.4	-18.8
<i>AA</i> <sub>1</sub>	-11.8	-12.5	-3.3	-4.0	-0.5	-20.2
<i>AA</i> <sub>2</sub>	-10.5	-12.0	-3.0	-3.9	-0.6	-16.6
<i>AA</i> <sub>3</sub>	-9.0	-10.2	-2.7	-3.2	-0.7	-13.9
<b>MSE</b>	<b>-0.7</b>	<b>5.1</b>	<b>5.5</b>	<b>10.0</b>	<b>-6.2</b>	
<b>MUE</b>	<b>1.1</b>	<b>5.2</b>	<b>5.5</b>	<b>10.0</b>	<b>6.2</b>	
<b>RMSE</b>	<b>1.3</b>	<b>5.7</b>	<b>5.8</b>	<b>10.3</b>	<b>6.5</b>	
<b>MAXE</b>	<b>-3.0</b>	<b>9.5</b>	<b>10.6</b>	<b>18.5</b>	<b>-10.8</b>	

<sup>a</sup> Comparison of binding enthalpies (kcal/mol) for nucleic acid base dimers from semiempirical (PM3<sub>BP</sub>, AM1, PM3, MNDO and MNDO/H) and DFT *m*PWPW<sup>1</sup> (*m*PWPW) calculations. Shown are the binding enthalpy values for each base pair, and summarized at the bottom are the error metrics (bold) for the semiempirical values with respect to the corresponding DFT *m*PWPW results of Sherer *et al.*<sup>1</sup>

Table A-VIII: Comparison of semiempirical and DFT dimerization energies of hydrogen bonded complexes<sup>a</sup>

Dimer <sup>b</sup>	B3LYP <sup>c</sup>	PM3 <sub>BP</sub>	AM1	PM3	MNDO/H
H <sub>2</sub> O/H <sub>2</sub> O	-3.01	-4.11	-5.47	-3.50	-4.59
H <sub>2</sub> CO/H <sub>2</sub> O	-2.86	-3.38	-3.41	-2.74	-4.60
CH <sub>3</sub> OH/H <sub>2</sub> O	-2.94	-3.75	-5.10	-3.21	-4.91
CH <sub>3</sub> CHO/H <sub>2</sub> O	-4.05	-4.41	-4.73	-3.65	-4.90
CH <sub>3</sub> COCH <sub>3</sub> /H <sub>2</sub> O	-4.16	-4.83	-5.27	-3.97	-4.85
CH <sub>3</sub> CONHCH <sub>3</sub> /H <sub>2</sub> O	-7.37	-5.96	-7.73	-4.01	-4.91
CH <sub>3</sub> CH <sub>2</sub> OH/H <sub>2</sub> O	-2.61	-5.54	-6.65	-4.90	-4.26
CH <sub>3</sub> CONH <sub>2</sub> /H <sub>2</sub> O	-2.74	-3.14	-3.85	-2.05	-4.25
(CH <sub>3</sub> ) <sub>2</sub> CHOH/H <sub>2</sub> O	-2.23	-5.37	-6.58	-4.70	-3.89
NH <sub>3</sub> /NH <sub>3</sub>	-1.59	-1.34	-2.38	-0.94	-3.78
CH <sub>3</sub> NHCH <sub>3</sub> /NH <sub>3</sub>	-1.12	-0.88	-2.07	-0.66	-4.32
NH <sub>3</sub> /H <sub>2</sub> O	-4.33	-4.83	-2.91	-3.04	-5.92
NH <sub>3</sub> /CH <sub>3</sub> OH	-4.40	-4.65	-1.71	-2.88	-6.84
H <sub>2</sub> CO/NH <sub>3</sub>	-1.16	-1.22	-2.69	-1.03	-2.82
CH <sub>3</sub> CHO/NH <sub>3</sub>	-1.94	-2.31	-3.55	-3.02	-3.04
CH <sub>3</sub> COCH <sub>3</sub> /NH <sub>3</sub>	-1.48	-2.18	-3.46	-2.86	-3.02
H <sub>2</sub> CO/CH <sub>3</sub> OH	-2.79	-2.94	-3.17	-2.39	-4.84
CH <sub>3</sub> OH/CH <sub>3</sub> OH	-3.30	-3.02	-3.71	-2.60	-3.83
CH <sub>3</sub> CHO/CH <sub>3</sub> OH	-3.97	-3.96	-3.70	-3.30	-5.16
CH <sub>3</sub> COCH <sub>3</sub> /CH <sub>3</sub> OH	-4.13	-4.31	-3.80	-3.54	-5.10
CH <sub>3</sub> CONH <sub>2</sub> /CH <sub>3</sub> OH	-3.03	-2.31	-3.56	-2.82	-3.32
CH <sub>3</sub> CONHCH <sub>3</sub> /CH <sub>3</sub> OH	-5.98	-5.13	-4.20	-3.84	-5.84
CHOOH/CHOOH	-14.00	-11.26	-6.43	-8.64	-15.64
NH <sub>2</sub> CHO/NH <sub>2</sub> CHO	-11.22	-10.72	-8.14	-4.75	-12.67
CH <sub>3</sub> COOH/CH <sub>3</sub> NH <sub>2</sub>	-9.06	-7.40	-3.78	-4.61	-8.51
CH <sub>3</sub> COOH/CH <sub>3</sub> NC(NH <sub>2</sub> ) <sub>2</sub>	-15.52	-11.42	-5.07	-6.86	-15.20
CH <sub>3</sub> CONH <sub>2</sub> /CH <sub>3</sub> CONH <sub>2</sub>	-4.52	-4.75	-4.72	-2.61	-5.30
CH <sub>3</sub> CONHCH <sub>3</sub> /CH <sub>3</sub> CONHCH <sub>3</sub>	-7.01	-4.92	-4.48	-3.56	-5.81

*Table A-VIII (continued)*

Dimer <sup>b</sup>	B3LYP <sup>c</sup>	PM3 <sub>BP</sub>	AM1	PM3	MNDO/H
CH <sub>3</sub> NH <sub>3</sub> <sup>+</sup> /H <sub>2</sub> O	-16.64	-14.07	-14.40	-12.80	-21.90
CH <sub>3</sub> NHC(NH <sub>2</sub> ) <sub>2</sub> <sup>+</sup> /CH <sub>3</sub> OH	-14.66	-10.18	-12.55	-8.93	-14.30
CH <sub>3</sub> NHC(NH <sub>2</sub> ) <sub>2</sub> <sup>+</sup> /H <sub>2</sub> O	-13.33	-11.54	-13.03	-10.20	-15.87
CH <sub>3</sub> O <sup>-</sup> /H <sub>2</sub> O	-22.92	-21.26	-16.45	-19.13	-19.00
CH <sub>3</sub> COO <sup>-</sup> /H <sub>2</sub> O	-14.13	-14.86	-15.91	-13.62	-16.28
CH <sub>3</sub> COO <sup>-</sup> /CH <sub>3</sub> OH	-15.87	-15.04	-10.20	-13.77	-17.14
(CH <sub>3</sub> ) <sub>2</sub> CHO <sup>-</sup> /H <sub>2</sub> O	-20.50	-20.72	-16.39	-18.45	-18.59
(CH <sub>3</sub> ) <sub>3</sub> CCOO <sup>-</sup> /H <sub>2</sub> O	-13.38	-14.30	-15.33	-13.07	-15.32
C <sub>6</sub> H <sub>5</sub> O <sup>-</sup> /H <sub>2</sub> O	-14.58	-13.83	-12.14	-12.19	-12.58
<b>MSE</b>	<b>0.34</b>	<b>0.81</b>	<b>1.61</b>	<b>-0.93</b>	
<b>MUE</b>	<b>1.11</b>	<b>2.37</b>	<b>2.04</b>	<b>1.63</b>	
<b>RMSE</b>	<b>1.58</b>	<b>3.27</b>	<b>2.87</b>	<b>1.91</b>	
<b>MAXE</b>	<b>4.47</b>	<b>10.45</b>	<b>8.65</b>	<b>-5.26</b>	

<sup>a</sup> Comparison of semiempirical (PM3<sub>BP</sub>, PM3, AM1, and MNDO/H) and DFT (B3LYP) dimerization enthalpies (kcal/mol).

<sup>b</sup> The notation X/Y represents the dimer formed from monomers X and Y.

<sup>c</sup> The DFT enthalpies were obtained from B3LYP/6-311++G(3df,2p)//B3LYP/6-31++G(d,p) energies corrected for zero-point vibrational energy and thermal corrections to the enthalpy from a frequency analysis at the B3LYP/6-31++G(d,p) level of theory. The dimerization enthalpy is defined as the difference in enthalpy between the dimer and the sum of the isolated monomers.

Table A-IX: Comparison of semiempirical and DFT dipole moments of neutral hydrogen bonded complexes and their monomers<sup>a</sup>

Molecule <sup>b</sup>	B3LYP <sup>c</sup>	PM3 <sub>BP</sub>	AM1	PM3	MNDO/H
H <sub>2</sub> O/H <sub>2</sub> O	2.89	2.83	2.26	2.47	3.19
H <sub>2</sub> CO/H <sub>2</sub> O	1.46	2.84	0.68	2.63	3.39
CH <sub>3</sub> OH/H <sub>2</sub> O	2.87	2.65	2.19	2.23	3.04
CH <sub>3</sub> CHO/H <sub>2</sub> O	2.69	3.37	2.73	3.07	3.73
CH <sub>3</sub> COCH <sub>3</sub> /H <sub>2</sub> O	3.11	3.67	3.67	3.39	3.88
CH <sub>3</sub> CONHCH <sub>3</sub> /H <sub>2</sub> O	3.47	3.26	3.80	3.51	2.48
CH <sub>3</sub> CH <sub>2</sub> OH/H <sub>2</sub> O	2.78	2.54	2.37	2.15	2.96
CH <sub>3</sub> CONH <sub>2</sub> /H <sub>2</sub> O	6.72	6.39	5.98	5.36	5.59
(CH <sub>3</sub> ) <sub>2</sub> CHOH/H <sub>2</sub> O	2.56	2.42	2.48	2.07	2.80
NH <sub>3</sub> /NH <sub>3</sub>	2.51	3.80	3.78	2.87	3.41
CH <sub>3</sub> NHCH <sub>3</sub> /NH <sub>3</sub>	2.68	2.67	3.16	1.63	3.22
NH <sub>3</sub> /H <sub>2</sub> O	3.59	4.05	3.80	3.48	3.68
NH <sub>3</sub> /CH <sub>3</sub> OH	3.54	3.87	3.08	3.28	3.61
H <sub>2</sub> CO/NH <sub>3</sub>	1.16	0.63	0.63	2.02	2.97
CH <sub>3</sub> CHO/NH <sub>3</sub>	1.60	3.50	2.69	2.76	3.19
CH <sub>3</sub> COCH <sub>3</sub> /NH <sub>3</sub>	2.13	4.04	3.21	3.16	3.49
H <sub>2</sub> CO/CH <sub>3</sub> OH	1.61	2.82	0.93	2.54	3.27
CH <sub>3</sub> OH/CH <sub>3</sub> OH	2.85	2.52	1.05	2.21	2.88
CH <sub>3</sub> CHO/CH <sub>3</sub> OH	2.93	3.37	1.23	3.00	3.61
CH <sub>3</sub> COCH <sub>3</sub> /CH <sub>3</sub> OH	3.27	3.70	2.03	3.32	3.78
CH <sub>3</sub> CONH <sub>2</sub> /CH <sub>3</sub> OH	6.68	6.14	5.57	5.24	5.30
CH <sub>3</sub> CONHCH <sub>3</sub> /CH <sub>3</sub> OH	5.05	5.14	3.13	4.10	4.78
CHOOH/CHOOH	0.00	0.00	0.00	0.00	0.00
NH <sub>2</sub> CHO/NH <sub>2</sub> CHO	0.01	0.73	0.02	1.26	0.90
CH <sub>3</sub> COOH/CH <sub>3</sub> NH <sub>2</sub>	2.18	2.43	1.54	1.96	2.60
CH <sub>3</sub> COOH/CH <sub>3</sub> NC(NH <sub>2</sub> ) <sub>2</sub>	3.55	2.22	0.77	1.69	1.70
CH <sub>3</sub> CONH <sub>2</sub> /CH <sub>3</sub> CONH <sub>2</sub>	9.13	8.98	7.91	7.24	7.33
CH <sub>3</sub> CONHCH <sub>3</sub> /CH <sub>3</sub> CONHCH <sub>3</sub>	9.48	9.31	7.17	7.05	7.81

*Table A-IX (continued)*

Molecule <sup>b</sup>	B3LYP <sup>c</sup>	PM3 <sub>BP</sub>	AM1	PM3	MNDO/H
H <sub>2</sub> O	1.91	1.90	1.86	1.74	1.78
H <sub>2</sub> CO	2.44	2.38	2.32	2.16	2.16
CH <sub>3</sub> OH	1.69	1.63	1.62	1.49	1.48
CH <sub>3</sub> CHO	2.83	2.75	2.60	2.46	2.34
CH <sub>3</sub> COCH <sub>3</sub>	3.07	3.07	2.85	2.73	2.48
CH <sub>3</sub> CONHCH <sub>3</sub>	4.25	4.07	3.81	3.39	3.20
CH <sub>3</sub> CH <sub>2</sub> OH	1.61	1.61	1.55	1.45	1.40
CH <sub>3</sub> CONH <sub>2</sub>	3.95	3.90	3.70	3.27	3.11
(CH <sub>3</sub> ) <sub>2</sub> CHOH	1.65	1.70	1.62	1.53	1.43
NH <sub>3</sub>	1.49	1.89	1.85	1.55	1.75
CH <sub>3</sub> NHCH <sub>3</sub>	0.97	1.40	1.23	1.27	1.17
H <sub>2</sub> CO	2.44	2.38	2.32	2.16	2.16
HCOOH	1.58	1.59	1.48	1.51	1.49
NH <sub>2</sub> CHO	4.02	3.80	3.70	3.11	3.10
CH <sub>3</sub> COOH	1.82	1.90	1.89	1.84	1.68
CH <sub>3</sub> NH <sub>2</sub>	1.28	1.62	1.49	1.40	1.48
CH <sub>3</sub> NC(NH <sub>2</sub> ) <sub>2</sub>	2.39	2.28	1.92	1.96	1.56
<b>MSE</b>	<b>0.17</b>	<b>-0.36</b>	<b>-0.25</b>	<b>0.01</b>	
<b>MUE</b>	<b>0.40</b>	<b>0.63</b>	<b>0.61</b>	<b>0.69</b>	
<b>RMSE</b>	<b>0.63</b>	<b>0.90</b>	<b>0.83</b>	<b>0.91</b>	
<b>MAXE</b>	<b>1.91</b>	<b>-2.78</b>	<b>-2.43</b>	<b>1.93</b>	

<sup>a</sup> Comparison of semiempirical (PM3<sub>BP</sub>, PM3, AM1, and MNDO/H) and DFT (B3LYP) dipole moments (Debye).

<sup>b</sup> The notation *X/Y* represents the dimer formed from monomers *X* and *Y*.

<sup>c</sup> The DFT dipole moments were obtained from B3LYP/6-311++G(3df,2p)//B3LYP/6-31++G(d,p) level of theory.

Table A-X: Comparison of semiempirical and DFT hydrogen-heavy atom hydrogen bond distances<sup>a</sup>

Molecule <sup>b</sup>	Type <sup>c</sup>	B3LYP <sup>d</sup>	PM3 <sub>BP</sub>	AM1	PM3	MNDO/H
H <sub>2</sub> O/H <sub>2</sub> O	O-H···O	1.92	1.76	2.23	1.81	1.65
H <sub>2</sub> CO/H <sub>2</sub> O	O-H···O	1.98	1.77	2.20	1.82	1.67
CH <sub>3</sub> OH/H <sub>2</sub> O	O-H···O	1.93	1.77	2.30	1.81	1.63
CH <sub>3</sub> CHO/H <sub>2</sub> O	O···H-O	1.93	1.76	2.27	1.81	1.66
CH <sub>3</sub> COCH <sub>3</sub> /H <sub>2</sub> O	O···H-O	1.90	1.76	2.24	1.81	1.67
CH <sub>3</sub> CONHCH <sub>3</sub> /H <sub>2</sub> O	O···H-O	1.84	1.76	2.22	1.81	1.64
CH <sub>3</sub> CONHCH <sub>3</sub> /H <sub>2</sub> O	N-H···O	2.06	1.77	2.16	2.78	1.90
CH <sub>3</sub> CH <sub>2</sub> OH/H <sub>2</sub> O	O-H···O	1.94	1.77	2.32	1.82	1.64
CH <sub>3</sub> CONH <sub>2</sub> /H <sub>2</sub> O	N-H···O	2.03	1.78	2.16	1.84	1.65
(CH <sub>3</sub> ) <sub>2</sub> CHOH/H <sub>2</sub> O	O-H···O	1.96	1.77	2.34	1.82	1.64
NH <sub>3</sub> /NH <sub>3</sub>	N-H···N	2.23	2.99	2.83	2.84	1.63
CH <sub>3</sub> NHCH <sub>3</sub> /NH <sub>3</sub>	N-H···N	2.23	2.67	3.50	2.86	1.59
NH <sub>3</sub> /H <sub>2</sub> O	N···H-O	1.93	1.76	2.64	1.83	1.61
NH <sub>3</sub> /CH <sub>3</sub> OH	N···H-O	1.93	1.77	2.58	1.83	1.57
H <sub>2</sub> CO/NH <sub>3</sub>	O···H-N	2.34	2.67	2.22	2.70	1.72
CH <sub>3</sub> CHO/NH <sub>3</sub>	O···H-N	2.26	2.66	2.44	2.91	1.72
CH <sub>3</sub> COCH <sub>3</sub> /NH <sub>3</sub>	O···H-N	2.21	2.66	2.44	2.87	1.72
H <sub>2</sub> CO/CH <sub>3</sub> OH	O···H-O	2.00	1.77	2.20	1.82	1.65
CH <sub>3</sub> OH/CH <sub>3</sub> OH	O-H···O	1.90	1.77	2.24	1.82	1.66
CH <sub>3</sub> CHO/CH <sub>3</sub> OH	O···H-O	1.94	1.77	2.19	1.82	1.64
CH <sub>3</sub> COCH <sub>3</sub> /CH <sub>3</sub> OH	O···H-O	1.91	1.77	2.15	1.82	1.65
CH <sub>3</sub> CONH <sub>2</sub> /CH <sub>3</sub> OH	O···H-O	2.00	1.79	2.18	1.86	1.68
CH <sub>3</sub> CONHCH <sub>3</sub> /CH <sub>3</sub> OH	O···H-O	1.85	1.76	2.12	1.81	1.63
CHOOH/CHOOH	O-H···O	1.66	1.73	2.10	1.78	1.54
CHOOH/CHOOH	O···H-O	1.66	1.73	2.10	1.78	1.54
NH <sub>2</sub> CHO/NH <sub>2</sub> CHO	O-H···O	1.87	1.74	2.07	1.82	1.54
NH <sub>2</sub> CHO/NH <sub>2</sub> CHO	O···H-O	1.87	1.74	2.07	1.82	1.54
CH <sub>3</sub> COOH/CH <sub>3</sub> NH <sub>2</sub>	O-H···N	1.69	1.73	2.58	1.79	1.53
CH <sub>3</sub> COOH/CH <sub>3</sub> NC(NH <sub>2</sub> ) <sub>2</sub>	O···H-N	1.88	1.75	2.19	1.82	1.54
CH <sub>3</sub> COOH/CH <sub>3</sub> NC(NH <sub>2</sub> ) <sub>2</sub>	O-H···N	1.56	1.71	2.51	1.78	1.48
CH <sub>3</sub> CONH <sub>2</sub> /CH <sub>3</sub> CONH <sub>2</sub>	N-H···O	1.99	1.77	2.19	1.84	1.65
CH <sub>3</sub> CONHCH <sub>3</sub> /CH <sub>3</sub> CONHCH <sub>3</sub>	N-H···O	1.98	1.77	2.15	1.84	1.65

*Table A-X (continued)*

Molecule <sup>b</sup>	Type <sup>c</sup>	B3LYP <sup>d</sup>	PM3 <sub>BP</sub>	AM1	PM3	MNDO/H
CH <sub>3</sub> NH <sub>3</sub> <sup>+</sup> /H <sub>2</sub> O	N-H···O	1.70	1.74	2.26	1.78	1.45
CH <sub>3</sub> NHC(NH <sub>2</sub> ) <sub>2</sub> <sup>+</sup> /CH <sub>3</sub> OH	N-H···O	1.99	1.81	2.09	1.86	1.51
CH <sub>3</sub> NHC(NH <sub>2</sub> ) <sub>2</sub> <sup>+</sup> /CH <sub>3</sub> OH	N-H···O	2.00	1.80	2.09	1.85	2.90
CH <sub>3</sub> NHC(NH <sub>2</sub> ) <sub>2</sub> <sup>+</sup> /H <sub>2</sub> O	N-H···O	2.03	1.80	2.09	1.85	1.50
CH <sub>3</sub> NHC(NH <sub>2</sub> ) <sub>2</sub> <sup>+</sup> /H <sub>2</sub> O	N-H···O	2.04	1.79	2.09	1.84	2.86
CH <sub>3</sub> O <sup>-</sup> /H <sub>2</sub> O	O···H-O	1.43	1.66	2.04	1.69	1.48
CH <sub>3</sub> COO <sup>-</sup> /H <sub>2</sub> O	O···H-O	1.72	1.72	2.74	1.75	1.53
CH <sub>3</sub> COO <sup>-</sup> /CH <sub>3</sub> OH	O···H-O	1.62	1.69	1.99	1.72	1.47
(CH <sub>3</sub> ) <sub>2</sub> CHO <sup>-</sup> /H <sub>2</sub> O	O···H-O	1.48	1.67	2.06	1.70	1.48
(CH <sub>3</sub> ) <sub>3</sub> CCOO <sup>-</sup> /H <sub>2</sub> O	O···H-O	1.73	1.72	2.74	1.76	1.54
C <sub>6</sub> H <sub>5</sub> O <sup>-</sup> /H <sub>2</sub> O	O···H-O	1.64	1.71	2.12	1.75	1.56
<b>MSE</b>		<b>-0.04</b>	<b>0.39</b>	<b>0.05</b>	<b>-0.24</b>	
<b>MUE</b>		<b>0.19</b>	<b>0.39</b>	<b>0.19</b>	<b>0.32</b>	
<b>RMSE</b>		<b>0.23</b>	<b>0.49</b>	<b>0.26</b>	<b>0.37</b>	
<b>MAXE</b>		<b>0.76</b>	<b>1.27</b>	<b>0.72</b>	<b>0.90</b>	

<sup>a</sup> Comparison of semiempirical (PM3<sub>BP</sub>, PM3, AM1, and MNDO/H) and DFT (B3LYP) hydrogen bond distances (Å).

<sup>b</sup> The notation *X/Y* represents the dimer formed from monomers *X* and *Y*. The hydrogen bond distance is measured from the hydrogen to the heavy atom to which it hydrogen bonds.

<sup>c</sup> The hydrogen bond type (Type) indicates the monomeric hydrogen bond donor and acceptor and the heavy atoms involved, e.g., D-H···A indicates that monomer *X* donates a hydrogen bond through a hydrogen bound to *D* and monomer *Y* accepts the hydrogen bond to an atom of type *A*.

<sup>d</sup> The DFT hydrogen bond lengths correspond to optimized geometries at the B3LYP/6-31++G(d,p) level of theory.