Distance Dependent Hydrogen Bond Potentials for Nucleic Acid Base Pairs from *ab Initio* Quantum Mechanical Calculations (LMP2/cc-pVTZ)

Ken Brameld, Siddharth Dasgupta,* and William A. Goddard III*

Materials and Process Simulation Center, Beckman Institute (139-74), Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received: January 13, 1997; In Final Form: March 17, 1997[⊗]

Hydrogen bonding between base pairs in nucleic acids is a key determinant of their structures. We have examined the distance dependence of the hydrogen bonding of AT-WC (Watson-Crick), GC-WC, and AT-H (Hoogsteen) base pairs using *ab initio* quantum mechanics, LMP2/cc-pVTZ(-f) energies at HF/cc-pVTZ(-f) optimized geometries. From these curves, we have extracted Morse potentials between the H atoms and the acceptor atoms that accurately reproduce the quantum mechanical energies for a range of geometries. Using these parameters, we have calculated the complexation energies of the remaining 26 possible pairwise combinations, and the agreement with previously reported *ab initio* calculations is excellent. We have also extracted off-diagonal Lennard-Jones 12-6 parameters to be used with the popular AMBER95 and CHARMM95 force fields that significantly improve their descriptions of the base-pairing energy and optimum geometry.

1. Introduction

Hydrogen bonds play a key role in maintaining structure and specificity of biological systems. In particular, the base pairing of nucleic acids (stemming from the specific formation of hydrogen bonds between Watson−Crick base pairs) is essential for the transfer of genetic information. We examine here the three of greatest biological relevance; the Watson−Crick base pairs, in which adenine hydrogen bonds to thymine (AT-WC) or guanine hydrogen bonds to cytosine (GC-WC), and the Hoogsteen adenine−thymine pair (AT-H).

Given the importance of hydrogen bonding in biological systems, considerable theoretical attention has been focused on exploring the nature and strength of these interactions. Extensive calculations have been carried out on the nucleic acid base pairs using semiempirical or *ab initio* quantum mechanical (QM) methods. For a recent review covering the current state of the field and discussing the extensive progress from semiempirical to *ab initio* QM methods, see ref 1. It is now established that the determination of accurate hydrogen bond energies and geometries requires a large diffuse basis set and the inclusion of electron correlation.[2,3] Many of the earlier studies of the nucleic acid base pairs did not include electron correlation effects, or estimated correlation energies with empirical methods. Recently, a series of studies have been reported in which dispersion energies have been evaluated using second-order Møller−Plesset perturbation theory (MP2)[4−8] or density functional theory (DFT).[9]

Experimentally, nucleic acid base pairing has been difficult to study, and there is little data with which theoretical results may be compared. Gas phase association energies have been reported for some systems,[10] and solution studies have been undertaken in nonpolar solvents.[11] The limited experimental data leads to uncertainty in the accuracy of empirical potentials.

We report *ab initio* QM calculations (LMP2/cc-pVTZ(-f)) on the complexation energies of the AT-WC, GC-WC, and AT-H nucleic acid base pairs. In addition to determining the minimum energy and geometry of these complexes, we used QM to calculate the strength of these interactions as a function of intermolecular distance. These potential energy curves were used to determine distance dependent hydrogen bond functions (described as Morse potentials). We find that force fields (FFs) incorporating these hydrogen bond potentials in conjunction with standard van der Waals and Coulombic terms accurately reproduce the full QM potential curve. This differs from the approach of current standard FFs (AMBER95,[12] CHARMM,[13] and OPLS[14]), which have been parametrized to reproduce only the energy and geometry at the bottom of the potential curve.

2. Methods

Calculations involving the nucleic acid base pairs were carried out with the PS-GVB software package.[15] Optimum geometries (HF/6-31G**) of the free bases 1-(hydroxymethyl)cytosine, 9-(hydroxymethyl)guanine, 1-(hydroxymethyl)thymine, 9-(hydroxymethyl)adenine, 1-methylthymine, and 9-methyladenine were determined while maintaining planar C_s symmetry. This constraint was also imposed for the optimization of the AT-WC, GC-WC, and AT-H base pairs. For the unpaired bases containing an exocyclic amino group, the optimum geometry has been reported to be nonplanar.[16] However, it has been shown that for the three base-paired structures we examine, the optimum structure is planar,[5b] and hence the errors due to the symmetry constraint are negligible.

Initially we obtained the optimum base pair geometries by gradient optimization (HF/6-31G**) while keeping internal base coordinates fixed in the optimum monomer conformation. These HF/6-31G** results proved to be inadequate and yielded geometries that did not match experimental structures. Full optimization with a more diffuse cc-pVTZ(-f) basis set greatly improved the results.[17] For these calculations no symmetry constraints were used.

Energetics were obtained using local second-order Møller−Plesset perturbation theory (LMP2)[18] with frozen core orbitals for geometries determined at the HF/cc-pVTZ(-f) level. The LMP2 method differs from canonical MP2 by limiting each occupied orbital's space of correlating virtual orbitals to be localized on the atoms of the occupied HF orbital. A Boys localization[19] of the HF reference wave function is used to determine the local MP2 reference wave function. This local correlation method has been reported to be essentially free from

*To whom correspondence should be addressed. E-mail: sdg@wag.caltech.edu and wag@wag.caltech.edu.

[⊗] Abstract published in *Advance ACS Abstracts,* May 15, 1997.

S1089-5647(97)00199-5 CCC: $14.00

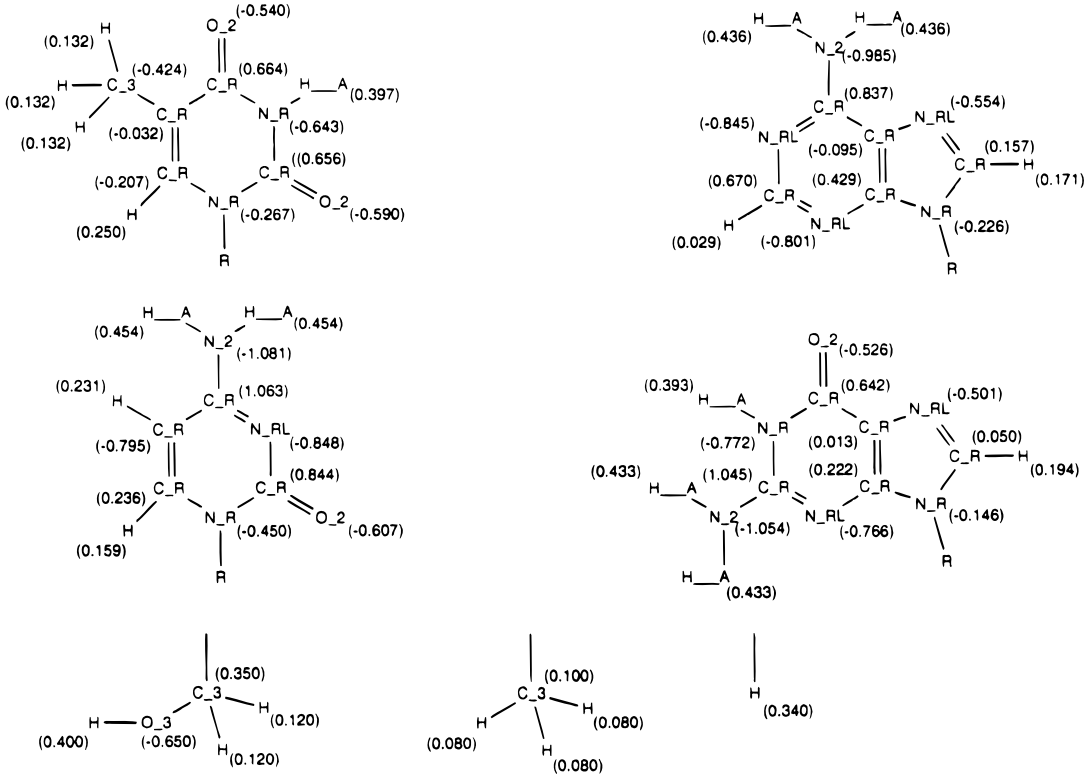

Figure 1. Diagram of the nucleic acid bases; thymine, adenine, cytosine, and guanine including force field atom types and partial charges. Charges were determined from LMP2/cc-pVTZ(-f) quantum mechanical calculations. The methyl and methoxy groups in the last row are common to all bases and are abbreviated as R in the base structures.

correlation basis set superposition error (BSSE) and offers significant computational savings.[18c]

Starting from the optimum geometry, energies were calculated at the LMP2/cc-pVTZ(-f) level as the central N−H···N hydrogen bond length was systematically increased. This yielded a potential energy curve as a function of distance for each base pair.[19] The HF BSSE correction was calculated for each geometry, using the counterpoise method of Boys and Bernardi.[20]

To simplify the calculations, we want to replace the deoxyribose sugar with either a methyl or hydroxymethyl substitution on the base. To establish which is best, we carried out a calculation on 2′-deoxyadenosine, including the sugar explicitly. Comparing the potential-derived charges for the 2′-deoxyadenosine case with those of 9-methyladenine and 9-(hydroxymethyl)adenine, we found that the hydroxymethyl substitution gave results closer to that with the complete sugar (see Figure 1). Since these hydrogen bond potentials are intended for parametrization of a force field (FF) that will include the sugar, the hydroxymethyl-substituted bases were used for all subsequent calculations. This hydroxymethyl substitution is expected to lead to a complexation energy similar to that of the full sugar because the charges for all atoms in close contact are quite similar.

Experimental hydrogen bond energies have been measured for the methyl-substituted base pairs. In order to compare our calculations directly with these experiments, we also examined the methyl substituted AT Hoogsteen base pair. The effects of this substitution are discussed in greater detail in sections 3.1.2 and 3.1.3.

The complexation enthalpies for base pair dimerization have been measured experimentally using mass spectrometry.[10] To compare between these experimental values and theoretical calculations, several corrections must be made. The free bases undergo a small conformational change upon forming base-paired complexes, resulting in a positive deformation energy associated with base pairing. Thus, the QM energy may be defined as

$$\Delta E_{QM} = \Delta E_{HF} - BSSE_{HF} + E_{cor} + E_{def} \qquad (1a)$$

where ΔE_{HF} is the HF energy, $BSSE_{HF}$ is the basis set superposition error, E_{cor} is the LMP2 correlation energy, and E_{def} the deformation energy. Finally, ΔH_{300} may be determined from

$$\Delta H_{300} = \Delta E_{QM} + \Delta ZPE + \Delta H_{0\rightarrow 300} \qquad (1b)$$

in which ΔZPE is the change in zero-point energy between the free bases and the dimer, while $\Delta H_{0\rightarrow 300}$ is the temperature dependence of ΔH from 0 to 300 K. ΔZPE and $\Delta H_{0\rightarrow 300}$ were calculated from the vibrational frequencies obtained from the FF. The two terms have opposite sign, with ΔZPE decreasing the value of ΔE while $\Delta H_{0\rightarrow 300}$ increases ΔE.

Atomic point charges were determined from the electrostatic potential (ESP) derived from the electron density distribution (constrained to reproduce the molecular monopole and dipole moments) calculated from the converged LMP2/cc-pVTZ(-f) wave functions (Figure 1).[21] MM calculations were performed with the POLYGRAF[22] software package using the Dreiding FF[23] with an exponential-6 van der Waals potential (2).

$$E_{vdw} = Ae^{-CR} - BR^{-6} \qquad (2)$$

The procedure for determining the *ab initio* derived hydrogen bond potentials was as follows:

1. The total nonbonded potential energy for the complex was determined for each geometry *excluding* the specific van der Waals contribution between hydrogen-bonding atoms. This steeply attractive potential matches the hydrogen bond energy at large distances.

2. The difference between the "electrostatic" potential and the hydrogen bond curve determined from QM is fit with a simple Morse function (3).

$$E_{\text{morse}} = D_0[\chi^2 - 2\chi] \tag{3}$$

$$\chi = \exp\left[-\frac{\zeta}{2}\left(1 - \frac{R}{R_0}\right)\right]$$

This off-diagonal Morse function is specific for each pair of hydrogen bond donor and acceptor atoms and enables the MM calculations to reproduce energies and geometries obtained from QM over the full range of hydrogen bond distances.

Thus, the nonbonded interactions involve superposing two-body van der Waals (E_{vdw}), hydrogen bond (E_{morse}) and electrostatic terms:

$$E_{\text{nonbond}} = E_{\text{vdw}} + E_{\text{morse}} + (332.0637)\frac{Q_iQ_j}{R_{ij}} \tag{4}$$

Hydrogen bond curves for the FF were obtained analogous to the QM methods. The hydrogen-bonded complexes were optimized while keeping the internal degrees of freedom fixed at the HF/6-31G** optimum geometry. The intermolecular distance was then increased and single-point energies calculated. We chose to keep the internal base geometry fixed so as to compare to the nonbonded potentials of the AMBER95,[12] CHARMM,[13] and OPLS[14] FFs without correcting for valence distortions.[24]

3. Results and Discussion

3.1. Quantum Mechanical Calculations for Nucleic Acid Base Pairs. *3.1.1. Level of Theory.* QM calculations on systems as large as nucleic acid base pairs are still computationally intensive. Our initial approach was to carry out all calculations at the HF level with a 6-31G** basis set. Earlier optimizations reported for the 6-31G* basis set yielded hydrogen bond lengths that were an average of 0.1 Å too long.[8] We obtained similar results with the 6-31G** basis set. Using the cc-pVTZ(-f) basis set gave much better geometries when compared to experiment. HF/6-31G** energies were determined for all points and corrected for BSSE. These energies were found to underestimate the ΔH of dimerization. To improve upon both the optimum geometries and energies obtained from QM, we used a larger basis set and included electron correlation for the single-point energy calculations (LMP2/cc-pVTZ(-f)). The results of these HF/cc-pVTZ(-F)// LMP2/cc-pVTZ(-f) calculations with BSSE corrections are discussed in detail in sections 3.1.3 and 3.1.4.

3.1.2. Geometry Optimization. The optimum hydrogen bond lengths for the AT-WC, GC-WC, and AT-H base pairs are reported in Table 1. The HF/6-31G** geometries lead to hydrogen bond lengths that are consistently too long by 0.15 Å. Increasing the basis set size [cc-pVTZ(-f)] gave much better results with an average difference of only 0.05 Å from the experimental geometry (some differences may be expected between the gas phase geometries of isolated bases and the crystal structure of a two-base-pair dimer).

As discussed in section 2, hydroxymethyl was used to replace the full ribose or deoxyribose sugar. For the AT-H base pair, both the methyl and hydroxymethyl substitutions were examined. The effect this substitution has on the HF/6-31G** geometry is worth noting. The optimum distances for N7···(H)N3 and N6(H)···O4 hydrogen bonds in the hydroxymethyl case are 3.09 and 2.98 Å, respectively, whereas in the methyl case they are 2.96 and 3.12 Å, respectively. This marks a reversal in the the relative lengths for both hydrogen bonds, yet this effect is not observed for the HF/cc-pVTZ optimized structures. These observations further support the need for a large basis set to correctly describe nonbonded interactions.

TABLE 1: Hydrogen Bond Lengths of the DNA Base Pairs (Å)

atom pair	HF/3-21G[a]	HF/6-31G**[b]	HF/cc-pVTZ(-f)[c]	X-ray[d]
		GC Watson−Crick		
O6−N4	2.77	2.91	2.83	2.91
N1−N3	2.91	3.08	2.95	2.95
N2−O2	2.86	3.11	2.92	2.86
		AT Watson−Crick		
N6−O4	2.96	3.17	3.06	2.95
N1−N3	2.78	3.02	2.92	2.82
		AT Hoogsteen		
N6−O4	2.99	2.98	2.91	2.86
N7−N3	2.75	3.09	2.95	2.93

[a] As reported by Gould et al.[8] [b] Internal geometry held fixed. [c] Full geometry optimization with no constraints. [d] From experimental X-ray crystallography data.[37]

TABLE 2: Root Mean Square Comparsion between *ab Initio* [HF/cc-pVTZ(-f)] and Experimental Nucleic Acid Base Structures

base	coordinate (Å)	bond (Å)	angle (deg)
1-methylcytosine	0.022	0.024	1.448
1-methylthymine	0.020	0.020	1.433
9-methyladenine	0.049	0.040	4.121
9-methylguanine	0.049	0.032	3.573

TABLE 3: DNA Base Pair Energies for HF Optimized Geometries (kcal/mol)

	GC Watson−Crick	AT Watson−Crick	AT Hoogsteen
HF/6-31G**	−23.8	−12.0	−11.5
HF/6-31G**[BSSE]	−21.4	−9.8	−9.4
HF/cc-pVTZ(-f)	−24.2	−10.3	−10.5
HF/cc-pVTZ(-f)[BSSE]	−22.8	−9.2	−9.4
LMP2/cc-pVTZ(-f)	−24.9	−12.0	−12.6
LMP2/cc-pVTZ(-f)-[BSSE]	−23.8	−11.1	−11.7
ΔE_{QM}[a]	−22.4	−10.8	−11.0
ΔZPE[b]	1.42	0.84	0.77
$\Delta H_{0\rightarrow 300}$[c]	−0.17	−0.23	−0.25
ΔH_{300}[d]	−21.2	−10.2	−10.5
ΔH_{exp}[e]	−21.0		−12.1[f]

[a] ΔE_{QM}, the LMP2/cc-pVTZ(-f) interaction energy at the HF/cc-pVTZ(-f) minimum geometry corrected for the deformation energy, Δ_{def}. [b] ΔZPE, the zero-point energy correction determined from vibrational frequencies calculated with molecular mechanics. [c] $\Delta H_{0\rightarrow 300}$, the difference in temperature dependence of ΔH. [d] $\Delta H_{300} = \Delta E_{\text{QM}} + \Delta\text{ZPE} + \Delta H_{0\rightarrow 300}$. [e] ΔH_{exp}, experimental ΔH from mass spectrometry data.[10] [f] This experimental ΔH has been corrected to reflect the effects of the multiple conformations possible for the AT base pair.[25]

The individual bases were also optimized (HF/cc-pVTZ(-f)) so as to determine internal strain energies associated with base pairing. These geometries were found to compare favorably with crystal structures of the free bases. Table 2 lists the rms difference in coordinates, bond lengths, and angles for each base. The average rms difference in coordinates for the pyrimidine bases was 0.02 Å and for the purine bases 0.05 Å.

3.1.3. Base Pair Complexation Energies. HF/6-31G** complexation energies were calculated for the optimized base pairs and corrected for BSSE. The base pair energies are listed in Table 3 and found to be substantially weaker than the reported experimental ΔH, even before taking ΔZPE corrections into account. The BSSE was found to be 2.1 and 2.4 kcal/mol for

the AT-WC and GC-WC base pairs, respectively. This compares favorably with the reported results of Gould *et al.*[8] in which the effects of BSSE corrections are discussed in detail for various methods and basis sets.

Using a cc-pVTZ(-f) basis set improves upon these results. However the inclusion of correlation energies is necessary to adequately describe the system. The HF/cc-pVTZ(-f) BSSE is reduced to 1.1 and 1.3 kcal/mol for the AT-WC and GC-WC base pairs respectively. With the LMP2 method, there is no additional correction to the dispersion energy due to BSSE.[18c] This offers a distinct advantage over standard MP2 methods, for which BSSE corrections are difficult to estimate and computationally intensive.

As discussed earlier, LMP2/cc-pVTZ(-f) energies can be compared to experimental complexation enthalpies only after a number of factors have been taken into account. Table 3 lists ΔE_{QM}, ΔZPE, and $\Delta H_{1\rightarrow 300}$, the sum of which ($\Delta H_{300}$) may be compared directly to experimentally determined enthalpies (ΔH_{exp}). For the GC-WC base pair, the calculated value of -21.2 kcal/mol is in excellent agreement with the experimental value of -21.0 kcal/mol. The calculated hydrogen bond energy of the AT-H base pair is considerably lower than the experimental results, -10.5 versus -13.0 kcal/mol. These energies differ from those of Gould *et al.*,[8] who found good agreement with the AT-H base pair but reported a GC-WC base pair energy 4.4 kcal/mol too negative when compared with experiment. The ΔH energies reported by Sponer *et al.*[5c] are within 1 kcal/mol of the energies we report.

We examined the AT-H complexation enthalpies in greater detail to understand the discrepancy between the theoretical and experimental energies. A methyl versus hydroxymethyl substitution for the full sugar makes no difference in the complexation energy. The HF/cc-pVTZ(-f) energy for methyl-AT-H is -9.45 kcal/mol compared to -9.42 kcal/mol for hydroxymethyl substitution. This is also in agreement with the experimental report that ΔH is independent of methylation.[10]

The AT base pair may assume one of four conformations; Watson−Crick, reverse Watson−Crick, Hoogsteen, or reverse Hoogsteen, all of which have been reported to have energies within 1 kcal/mol of each other.[5c] Experimental gas phase complexation enthalpies were not corrected for contributions from all four possible orientations.[25] At 60 °C, this leads to an overestimation of ΔH by 0.9 kcal/mol. Thus the experimental AT complexation enthalpy should be -12.1 kcal/mol. While this correction reduces the difference between our theoretical energies and experiments, the discrepancy of 1.6 kcal/mol is still significant.

Although stabilization between base pairs is dominated by electrostatics, dispersion energies contribute significantly more to the AT-H and AT-WC base pairs versus GC-WC. For the GC-WC base pair dispersion energies account for only 3% of the total complexation energy, whereas for the AT-WC and AT-H pairs this contribution is increased to 15% and 18%, respectively. It may be the case that an LMP2 geometry optimization of the AT and GC base pairs would be found to preferentially stabilize the AT pair.

3.1.4. Distance Dependent Potential Energy Surfaces. Given an optimum geometry for each of the Watson−Crick and Hoogsteen base pairs, it is possible to determine the complexation energy as a function of distance. For each base pair, the distance between atoms of the central (N···H−N) hydrogen bond was increased and single-point energy calculations were performed followed by BSSE corrections. Complete HF/6-31G** curves were obtained for all the base pairs. The AT-WC and GC-WC curves are shown in Figure 2. As discussed above, these hydrogen bond lengths are slightly too long and the energies too weak to be used to fit a FF. For completeness, the energies of the HF/6-31G** calculations may be found in the Supporting Information Table S1.

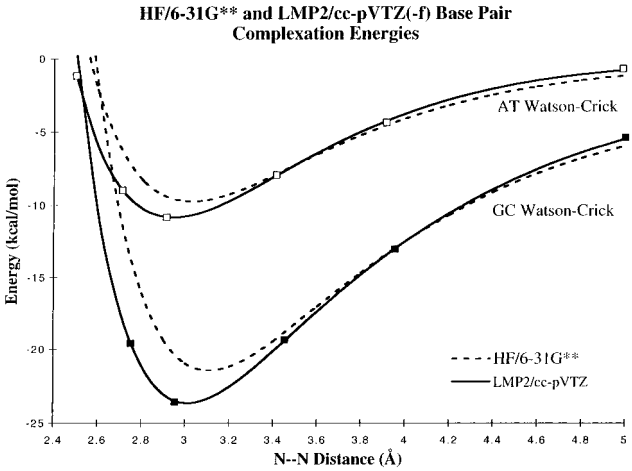

Figure 2. Potential energy curves for the AT-WC and GC-WC base pairs. The base pair separation is defined as the distance between heteroatoms of the central hydrogen bond, atoms N1 of the pyrimidines, and N3 of the purines. Note the overestimation of the complexation by the Hartree−Fock calculation at distances greater than 4 Å.

TABLE 4: LMP2/cc-pVTZ(-f) Total and Component Energies (kcal/mol) for the GC-WC Base Pair at Variable N1(H)···N3 Distances (Å)

N1(H)···N3	E_{HF}	$BSSE_{HF}$	E_{cor}	E_{LMP2}
2.50	+5.053	1.660	−4.884	+1.829
2.75	−18.473	1.226	−2.328	−19.575
2.95	−24.172	1.347	−0.737	−23.562
3.45	−20.611	0.718	+0.530	−19.363
3.95	−14.516	0.493	+0.957	−13.066
5.00	−6.840	0.227	+1.176	−5.437

TABLE 5: LMP2/cc-pVTZ(-f) Total and Component Energies (kcal/mol) for the AT-WC Base Pair at Variable N1···(H)N3 Distances (Å)

N1···(H)N3	E_{HF}	$BSSE_{HF}$	E_{cor}	E_{LMP2}
2.50	+1.862	1.539	−4.577	−1.176
2.72	−7.420	1.316	−2.876	−8.980
2.92	−10.347	1.157	−1.630	−10.820
3.42	−8.054	0.947	−0.863	−7.970
3.92	−4.872	0.705	−0.212	−4.379
5.00	−1.360	0.304	+0.339	−0.717

TABLE 6: LMP2/cc-pVTZ(-f) Total and Component Energies (kcal/mol) for the AT-H Base Pair at Variable N7···(H)N3 Distances (Å)

N7···(H)N3	E_{HF}	$BSSE_{HF}$	E_{cor}	E_{LMP2}
2.50	+3.119	1.469	−6.789	−2.201
2.75	−6.975	1.287	−5.111	−10.799
2.95	−10.750	1.327	−1.884	−11.307
3.45	−8.741	0.958	−0.775	−8.558
3.95	−5.345	0.675	−0.059	−4.729
5.00	−1.561	0.332	+0.469	−0.760

For each base pair, six single-point LMP2/cc-pVTZ energies were calculated: one at the HF/cc-pVTZ optimum geometry and five additional points up to a hydrogen bond length of 5 Å. A final distance of 5 Å between heteroatoms was chosen, as the dispersion energies are nearly zero and the total potential energy is largely determined by electrostatics. For each geometry, the component and total energies are listed in Tables 4−6 for the GC-WC, AT-WC, and AT-H base pairs, respectively. To compare with the HF/6-31G** results, the points for the AT-WC and GC-WC pairs are plotted in Figure 2.

TABLE 7: Dependence of Correlation Energies (kcal/mol) for the H_2O Dimer on O···O Distances (Å) and Level of Theory, Using the cc-pVTZ Basis Set[a]

O···O	MP2	MP3	MP4	LMP2[b]
2.50	−2.018	−1.834	−1.794	−1.917
2.66	−1.481	−1.363	−1.337	−1.659
2.97	−0.765	−0.712	−0.695	−1.055
4.00	−0.036	−0.021	+0.001	−0.119
5.00	+0.036	+0.037	+0.056	+0.005
6.00	+0.031	+0.030	+0.042	+0.030
7.00	+0.021	+0.020	+0.028	+0.023

[a] All calculations use a frozen core approximation. The optimum geometry has an O···O distance of 2.97 Å. [b] LMP2 calculations were carried out with the cc-pVTZ(-f) basis set, as discussed in the Methods section.

Examination of the correlation energies reveals a *positive* contribution to the total energy at intermediate distances. Calculations on the CC and GG dimers, the only other DNA systems for which distance dependent hydrogen bond potentials have been reported, show the same tendencies.[5c] The LMP2 electron correlation energies may be decomposed as in (6).[26,27]

$$E_{\mathrm{cor}}^{\mathrm{LMP2}} = E_{\mathrm{disp}} + \Delta E_{\mathrm{rep}} \tag{6}$$

where E_{disp} is the dispersion energy and is always negative, while ΔE_{rep} represents the remaining repulsive terms, which include the intrasystem correlation correction to the electrostatic effect, the induction correlation, and exchange terms. MP2 calculations have been reported for simple hydride dimers such as $(HF)_2$.[26] These calculations indicate that ΔE_{rep} is slightly smaller in magnitude than E_{disp}, with the two terms largely canceling each other, leading to a small total correlation correction. A fine balance is maintained between these components, which individually are sensitive to the intermolecular separation, yet vary slowly once summed. At intermediate to long hydrogen bond distances, where E_{disp} becomes negligible, ΔE_{rep} still contributes slightly to the total energy, resulting in a positive E_{cor}. Thus the HF calculation overestimates the electrostatic energies and has a lower relative energy for these points. Similarly, a FF which incorporates charges determined from isolated monomers will overestimate the electrostatic energy at points where there is a positive change in electrostatic energy, with decreasing distance, due to electron correlation.

To examine whether such positive correlation energies observed using MP2 or LMP2 methods are general, we carried out extensive calculations on a water dimer system. The optimum dimer geometry used in these calculations was the same as that determined by Saebo *et al.*[18] The intermolecular distance was then varied, and for each geometry the LMP2/cc-pVTZ(-f) energies were compared to MP2, MP3, and (SDTQ)MP4/cc-pVTZ energies after full BSSE corrections. The correlation energies as a function of distance are listed in Table 7, and the total energies plotted in Figure 3. All levels of theory show the same distance dependent behavior with a positive correlation energy at intermediate intermolecular distances. The complexation energy determined with LMP2 (4.67 kcal/mol) is within the range of experiment,[28] 5.4 ± 0.7 kcal/mol (MP2 with augmented cc-pVTZ basis set leads to 5.58 and 4.65 kcal/mol when corrected for BSSE and monomer relaxation).[29]

3.2. Force Field Parametrization. Using the curves calculated from QM, we sought a simple FF description for use in molecular mechanics (MM) calculations that could correctly describe both the inner wall of the minimum energy well and the curvature as the hydrogen bond length is increased. Examination of the hydrogen bond potentials determined from QM reveals that at large distances the energies are dominated by electrostatics. We find that using charges derived from QM for all atoms, it is only necessary to define a repulsive Morse potential between the atom pairs directly involved in hydrogen bonding to reproduce the QM curves. This method has been used successfully in the parametrization of a specialized FF for nylon.[30]

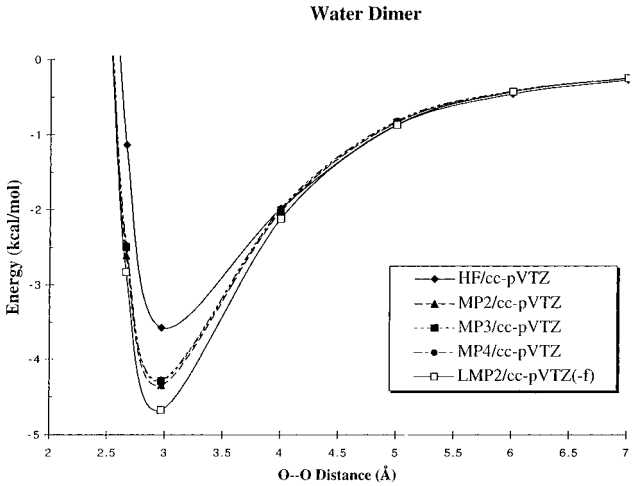

Figure 3. Potential energy curves for the H_2O dimer at various levels of theory. The LMP2/cc-pVTZ(-f) minimum energy is closer to the experimental association energy of 5.4 ± 0.7 kcal/mol. In addition, all calculations that include electron correlation effects show the same behavior at intermediate distances, with association energies that are weaker than the Hartree−Fock calculation predicts.

Parametrization of the Morse potentials is closely related to the van der Waals description of the FF used. For these calculations we have used the Dreiding FF with the exponential-6 (Exp-6) van der Waals (vdW) option. All vdW interactions are included explicitly except the specific vdW interaction between two atoms that form a hydrogen bond. This energy is set to zero. For each geometry, a MM energy is calculated with nonbonded energies comprised of an Exp-6 vdW term and a standard Coulombic function. The difference between the QM energy and the FF Exp-6 + Coulomb energy may be fit with a Morse potential. The Morse potential is specific for the selected pair of atoms forming the hydrogen bond and replaces the normal Exp-6 vdW energy calculated from combination rules.

We report two sets of Morse parameters for the MSC FF, which differ in accuracy and generality. The best parameters (MSC1 FF) require different potentials between the AT (-WC and -H) base pairs and the GC-WC base pair. While this gives an excellent fit to the QM data, it is not as useful for studies of uncommon base pair geometries or non-natural bases. A satisfactory fit using a global set of parameters (MSC2 FF) for all the base pairs was also obtained. The MSC2 FF leads to systematic differences from the QM data, probably because of cooperative effects in hydrogen bonding. The AT-WC and AT-H base pair energies are too strong by 0.6 and 0.4 kcal/mol, respectively, and the GC-WC energy is underestimated by 1.1 kcal/mol. The parameters for both MSC1 FF and MSC2 FF are listed in Table 8. The optimum energies and geometries obtained using these new FF are listed with the QM results in Tables 9 and 10.

3.3. Molecular Mechanics Force Field Evaluations. *3.3.1. Distance Dependent Potential Energy Surfaces.* Hydrogen bonding has been addressed in a number of ways in various FFs. Earlier versions of AMBER,[31] CHARMM, and Dreiding used a Lennard-Jones (LJ) 12-10 potential for hydrogen-bonding atom pairs. The current releases of AMBER (AMBER95),[12] CHARMM,[13] and OPLS[14] all use a LJ 12-6 potential, but differ in the derivation of atomic charges and the specific LJ

TABLE 8: Morse Potential Parameters for the DNA Base Pairs

atom pair	atom types	R_0	D_0	ζ
MSC1 FF GC				
O6−N4/N2−O2	O_2−H___A	2.48	0.987	7.55
N1−N3	N_RL−H___A	2.32	0.711	8.83
MSC1 FF AT				
N6−O4	O_2−H___A	2.58	0.241	9.92
N1−N3	N_RL−H___A	2.62	0.192	10.46
MSC2 FF (all pairs)				
	O_2−H___A	2.55	0.200	9.00
	N_RL−H___A	2.70	0.200	9.40

TABLE 9: Optimum Hydrogen Bond Lengths of the DNA Base Pairs (Å) As Determined by Molecular Mechanics with the AMBER95.1*, CHARMM*, and MSC Force Fields; Internal Base Geometries Were Defined by the Optimum HF/6-31G Structures and Held Fixed during Optimization by Each Force Field**

atom pair	AMBER95.1*[a]	CHARMM*[b]	MSC1	MSC2	X-ray[c]
		GC Watson−Crick			
O6−N4	2.91	2.88	2.94	2.84	2.91
N1−N3	2.97	2.95	3.02	2.91	2.95
N2−O2	2.89	2.88	2.95	2.84	2.86
		AT Watson−Crick			
N6−O4	2.99	2.97	3.02	2.87	2.95
N1−N3	2.89	3.01	2.93	2.88	2.82
		AT Hoogsteen			
N6−O4	3.08	3.03	3.07	2.91	2.86
N7−N3	2.92	2.93	2.91	2.86	2.93

[a] AMBER95* force field parameters and charges from ref 12 with specific off-diagonal hydrogen bond van der Waals terms added. [b] CHARMM*[13] with specific off-diagonal hydrogen bond van der Waal terms added. [c] From experimental X-ray crystallography data.[37]

TABLE 10: Optimum Complexation Energies for the DNA Base Pairs (Å) As Determined by Molecular Mechanics with the AMBER95.1*, CHARMM*, and MSC Force Fields; Internal Base Geometries Were Defined by the Optimum HF/6-31G Structures and Held Fixed during Optimization by Each Force Field**

	GC Watson−Crick	AT Watson−Crick	AT Hoogsteen
AMBER95.1*[a]	−23.5	−11.0	−11.4
CHARMM*[b]	−23.1	−11.3	−11.4
MSC1	−23.7	−10.9	−12.0
MSC2	−22.7	−11.7	−12.1
QM[c]	−23.8	−11.1	−11.7

[a] AMBER95* force field parameters and charges from ref 12 with specific off-diagonal hydrogen bond van der Waals terms added. [b] CHARMM*[13] with specific off-diagonal hydrogen bond van der Waal terms added. [c] LMP2/cc-pVTZ(-f)[BSSE] quantum mechanical energies for the HF/cc-pVTZ(-f) optimized structure.

parameters. AMBER95 and CHARMM both use QM-derived charges. AMBER95 uses the restrained electrostatic potential (RESP)[32] method and alters the LJ parameters of the hydrogen-bonding atoms to fit experimental energies and geometries. The OPLS FF uses general LJ parameters based on element type only and uses a charge-fitting scheme which reproduces the energy and geometries of water-based complexes as determined from *ab initio* quantum mechanics. We find that these standard MM FFs lead to an inaccurate description of the curvature of the hydrogen bond potential and usually do not predict correct complexation energies and geometries.

For each FF (including MSC1 FF) we have evaluated the base pair complexation energies as a function of distance. The internal geometry of each base was held fixed in the HF/6-31G** optimum conformation. Intermolecular geometries were determined by gradient optimization of the nonbonded energies,

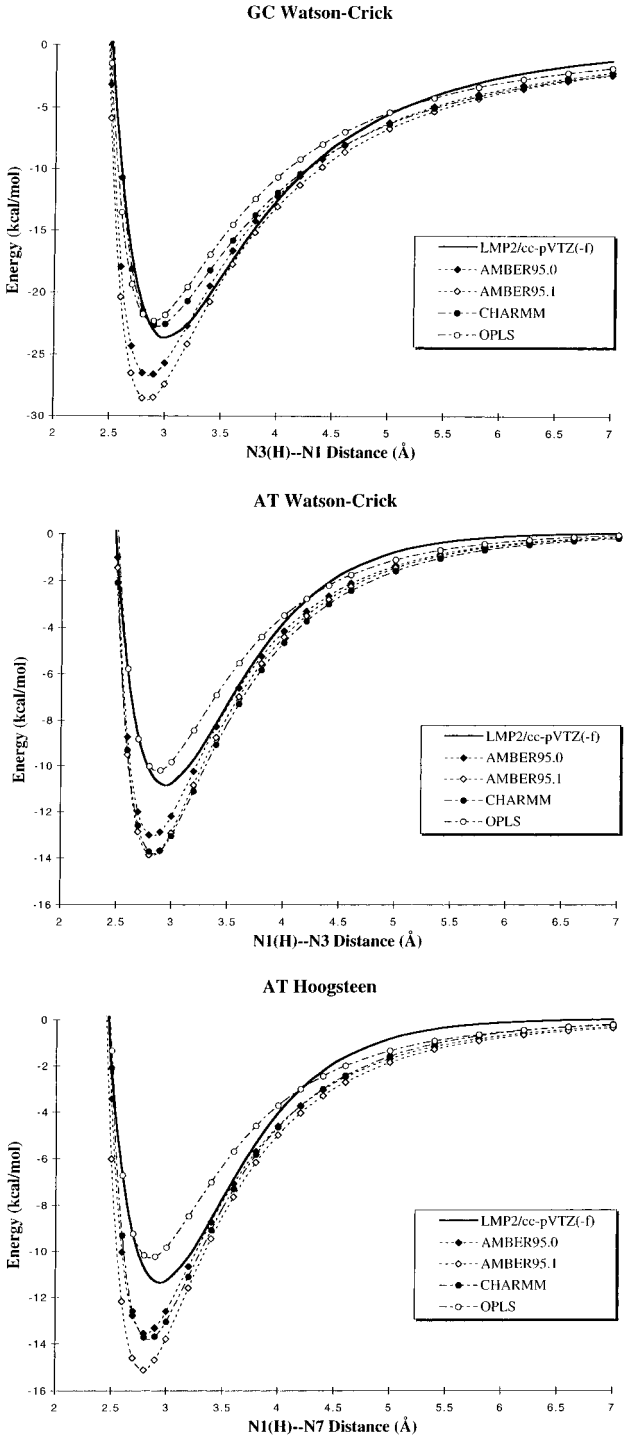

Figure 4. Comparison between potential energy curves for the DNA base pairs (a) GC-WC, (b) AT-WC, and (c) AT-H as determined with different molecular mechanics force fields and quantum mechanical results. The AMBER95.0 FF uses charges from ref 33, while AMBER95.1 uses charges from ref 12.

followed by single-point energy calculations at fixed intermolecular distances. Two sets of charges are available for the AMBER95 FF: (i) AMBER95.0 denotes calculations using the original RESP charge scheme (listed in the supplementary materials of ref 33), for which the complexation energies are reported in the FF paper, ref 12, and (ii) AMBER95.1[34] denotes the charge scheme released with the full FF,[12] which does not match the energies in ref 33.

The results of this analysis, plotted in Figure 4a−c and listed in Tables S2−S4, reveal that FFs that employ *ab initio* QM-derived charges are more successful at reproducing the QM

energies. For all three cases examined, AMBER95.0 and AMBER95.1 overestimate the hydrogen bond energies by 3−5 kcal/mol. CHARMM describes GC-WC to within 1 kcal/mol, but overestimates the AT-WC and AT-H energies by 3 and 2.5 kcal/mol, respectively. The OPLS FF underestimates the complexation energies by 1−2 kcal/mol and is considerably worse at reproducing the QM energies away from the optimum geometry. All FF incorporating a LJ6-12 description suffer from the same shortcoming intrinsic to the form of the potential. Specifically, the attractive portion of the potential is too steep, creating a minimum energy well that is too narrow. In comparison, the Morse potential used with the MSC FF leads to a softer inner wall and a broader minimum energy well that better describes the true curve.

In spite of this improvement, the MSC1 Morse potential does not reproduce the exact QM energies for the AT-H and AT-WC pairs at longer hydrogen bond distances (see Figure 5b−c). As discussed earlier, inclusion of electron correlation decreases the electrostatic energy, leading to a weaker bonding at the longer distances. An electrostatic model based on independent monomers correctly describes the system around the minimum, but overestimates the stabilization energy at longer distances. For the GC-WC pair, where dispersion contributes less to the total bonding, this effect is negligible, but for AT complexes, which depend more strongly upon dispersion, it is significant.

3.3.2. Additional Parameters for the CHARMM and AMBER95.1 Force Fields. In addition to determining optimum off-diagonal nonbonded parameters for the MSC FF, we carried out similar optimizations for CHARMM and AMBER95.1. For both FFs we report the optimum LJ12-6 parameters needed to define the vdW potentials between the heteroatoms involved in hydrogen bonding. These off-diagonal van der Waals terms replace the standard R_e and D_e determined using combination rules. No charges or valence terms have been altered. The parameters are listed in Table 11. We denote calculations using these modified terms as CHARMM* and AMBER95*. The new optimum geometries and energies are compared in Tables 9 and 10, respectively. Full potential curves are plotted in Figure 5a−c.

Incorporating these special off-diagonal terms greatly improves both FFs. For AMBER95.1* (which uses the charge scheme distributed with AMBER95), the minimum geometry and energy for all three base pairs could be adjusted to reproduce the QM results with quite satisfactory distance dependence. Due to the additional number of atom types used by the CHARMM FF, all three base pair complexes could be fit independently; however, this was not sufficient to attain a satisfactory fit for AT-WC. Here the final optimum hydrogen bond lengths are too long by 0.1 Å and the energies too strong by 0.4 kcal/mol. As with AMBER95.1*, both the GC-WC and AT-H complexes could be fit and yielded similar distance dependent curves shown in Figure 5a−c. These off-diagonal van der Waals parameters are simple to include during MM and dynamics simulations and greatly improve the accuracy of these two popular FFs.

3.3.3. Base Pair Complexation Energies of 26 Possible Geometries. The stabilization energies of 26 possible geometries for the DNA base pairs, in which a minimum of two hydrogen bonds are formed, have been evaluated using *ab initio* QM (MP2/6-31G**).[5c] These energies have been compared to those obtained from empirical potentials such as AMBER95, CHARMM, and OPLS FFs.[35] A linear regression analysis of each of these potentials compared to the MP2 results simplifies the task of assessing which FF best reproduce the QM energies. Such an analysis has been previously reported for the potentials

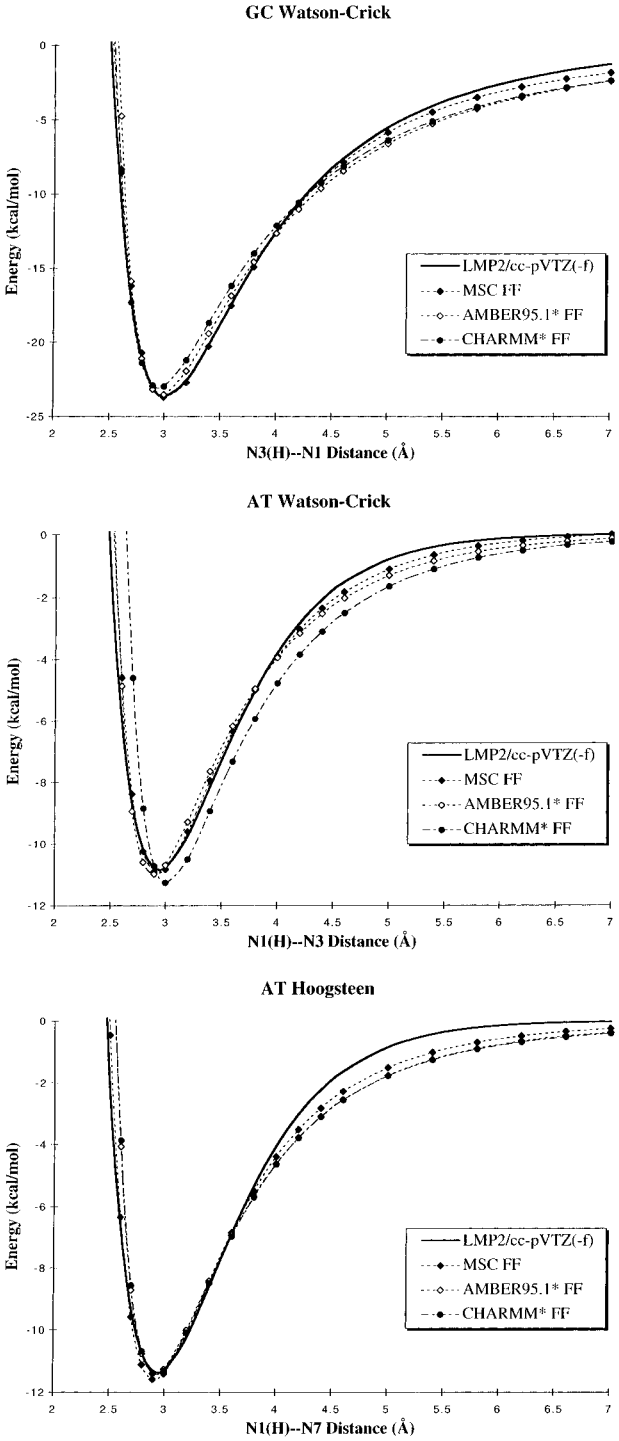

Figure 5. Comparison between potential energy curves for the DNA base pairs (a) GC-WC, (b) AT-WC and (c) AT-H as determined by *ab initio* quantum mechanics, the MSC FF, and modified versions of AMBER95 and CHARMM. The AMBER95* force field parameters and charges are from ref 12 with the specific off-diagonal hydrogen bond van der Waals terms listed in Table 11 included. The CHARMM* FF is standard CHARMM[13] with the specific off-diagonal hydrogen bond van der Waal terms listed in Table 11 included.

listed above,[35] and we have now extended this work to include the general MSC2 FF.

The MP2 and MSC2 FF energies of each complex are plotted in Figure 6a, and the linear regression plot is shown in Figure 6b. The MSC2 FF energies correlate very well with the QM results. The specific linear regression parameters for each of the FFs are listed in Table 12. The correlation coefficient is an overall measure of how well the MM potentials and QM methods are correlated, with a value closest to 1.0 representing

TABLE 11: LJ12-6 Parameters for the DNA Base Pairs Used To Correct the AMBER95.1 and CHARMM Force Fields

atom types	R_0	D_0
	AMBER95.1	
NA−NC	4.395	0.020
N2−O	4.410	0.017
NA−NB	4.500	0.020
	CHARMM	
NN3−NN2G	3.38	0.550
NN3A−NN2U	4.01	0.130
NN4−NN2U	4.15	0.060
NN1−ON1	3.59	0.180

TABLE 12: Linear Regression Analysis of Complexation Energies for 26 Hydrogen-Bonded DNA Base Pair Geometries Using the AMBER95.1, CHARMM, and MSC Force Fields Compared to *ab Initio* Energies at the MP2/6-31G level:[a] The Equation Is $Y = A + B\,X$ where $X \equiv \Delta E^{MP2}$**

	MSC2 FF	AMBER95[b]	CHARMM[b]	OPLS[b]
R[c]	0.98	0.96	0.93	0.96
SD[d]	0.83	1.06	1.60	1.11
A	0.60	0.35	0.66	−1.34
B	0.98	0.98	0.96	0.92
AAE[e]	0.7	1.0	1.0	2.4

[a] MP2/6-31G** energies reported by Hobza *et al.*[5c] [b] Linear regression analysis reported by Hobza *et al.*[35] [c] Correlation coefficient. [d] Standard deviation (kcal/mol). [e] Average absolute error (kcal/mol). AAE $= (1/26)\sum_i^{26}|\Delta E^{MP2} - \Delta E^{FF}|$.

the highest possible correlation. In addition, the standard deviation reflects the difference in energies between the two methods for each geometry. The MSC2 FF has the highest correlation coefficient (0.98) of the FFs examined and the smallest standard deviation (0.83 kcal/mol). Other parameters may also be examined such as the intercept (A) or slope (B) of the linear fit, which should be close to 0.0 and 1.0, respectively. Finally, the MCS2 FF leads to the smallest average absolute error.

Since none of the MP2 energies were used to parametrize the MSC2 FF, this represents an independent test of both the method of parametrization and the quality of LMP2 complexation energy calculations compared to the MP2 method. It is evident that a generalized hydrogen bond potential derived from the full LMP2 potential energy surfaces of three base pair geometries is sufficient to accurately describe all possible base pair interactions.

4. Conclusion

We examined the distance dependent hydrogen bond energies of the AT-WC, GC-WC, and AT-H base pairs using *ab initio* QM (LMP2/cc-pVTZ(-f)//HF/cc-pVTZ(-f)). Using these energies as a standard, an alternative hydrogen bond potential for use in MM and dynamics simulations is presented. This new FF incorporates a Morse potential to describe the nonbonded energies between hydrogen bonding atoms and reproduces the QM energies for a range of geometries. In addition to parametrizing the new FF, we examined the accuracy of two commonly used FFs, AMBER95 and CHARMM95. Neither FF was found to correctly describe the optimum geometry or energy of the base-paired complexes. The accuracy of these FFs was greatly improved by replacing the LJ12-6 potential between the heteroatoms directly involved in hydrogen bonding with an off-diagonal LJ12-6 potential tuned for the base pairs. The new parameters and the resulting complex energies and geometries are reported.

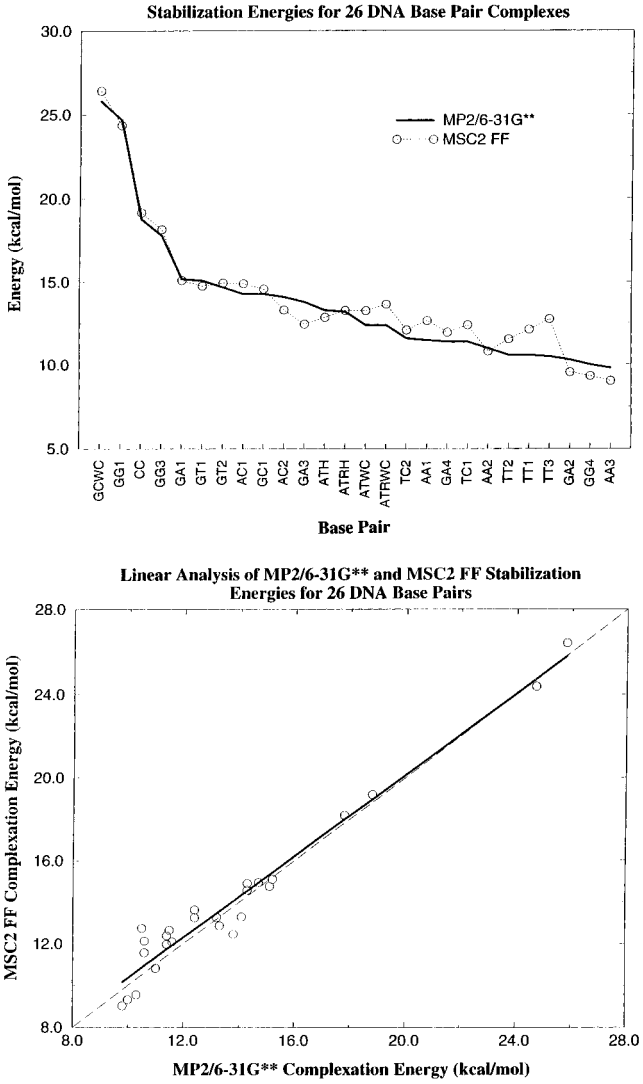

Figure 6. (a) Complexation energies for 26 possible base pair conformations determined with *ab initio* quantum mechanics (MP2/6-31G**) and the general MSC2 FF. See ref 5c for specific geometries. (b) For each complex, the MSC2 FF energies are plotted against the MP2 energies. The dashed line represents the ideal case in which the two energies are equal. The solid line is the least squares best fit through all points.

These *ab initio* calculations and analytical hydrogen bond potential energy functions are part of an ongoing effort to develop general first principles FF for biological systems. The goal of this work is to provide a relatively simple set of functions which have been fit to reproduce *ab initio* calculations on small model compounds. Parameters thus defined are general and may be easily extended to new systems including non-natural amino acids and nucleic acid bases. Such a new generation FF may be extended using QM theory and is thus independent of experimental parameters.

Acknowledgment. This research was funded by DOE-BCTR (DE-FG36-93CH105 81). The facilities of the MSC are also supported by grants from NSF (CHE 95-22179 and ASC 92-100368), Chevron Petroleum Technology Co., Aramco, Asahi Chemical, Owens-Corning, Chevron Chemical Company, Asahi Glass, Chevron Research and Technology Co., Hercules, BP Chemical, and Beckman Institute. Some calculations were carried out at the NSF San Diego Supercomputer Center (SDSC).

Supporting Information Available: Tables of HF/6-31G** complexation energies and MM evaluation of the complexation energies (4 pages). Ordering information is given on any current masthead page.

References and Notes

(1) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117−135.

(2) (a) Delbene, J. E.; Shavitt, I. *J. Mol. Struct. (THEOCHEM)* **1994**, *307*, 27−34. (b) Delbene, J. E. *J. Phys. Chem.* **1993**, *97*, 107−110. (c) Delbene, J. E.; Shavitt, I. *Int. J. Quantum Chem.* **1989**, 445−452. (d) Delbene, J. E. *Int. J. Quantum Chem.* **1987**, 27−35. (e) Delbene, J. E. *J. Comput. Chem.* **1987**, *8*, 810−815. (f) Delbene, J. E. *J. Chem. Phys.* **1987**, *86*, 2110−2113.

(3) (a) Szczesniak, M. M.; Latajka, Z.; Scheiner, S. *J. Mol. Struct. (THEOCHEM)* **1986**, *28*, 179−188. (b) Latajka, Z.; Scheiner, S. *J. Chem. Phys.* **1984**, *81*, 2713−2716. (c) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. *J. Comput. Chem.* **1983**, *4*, 294−301. (d) Spitznagel, G. W.; Clark, T.; Chandrasekhar, J.; Schleyer, P. V. R. *J. Comput. Chem.* **1982**, *3*, 363−371.

(4) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Comput. Chem.* **1996**, *17*, 841−850.

(5) (a) Sponer, J.; Hobza, P. *Chem. Phys.* **1996**, *204*, 365−372. (b) Sponer, J.; Florian, J.; Hobza, P.; Leszczynski, J. *J. Biomol. Struct. Dyn.* **1996**, *13*, 827−833. (c) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965−1974.

(6) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590−5596.

(7) Sponer, J.; Leszczynski, J.; Vetterl, V.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *13*, 695−706.

(8) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493−2499.

(9) Hutter, M.; Clark, T. *J. Am. Chem. Soc.* **1996**, *118*, 7574−7577.

(10) Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. *Biopolymers* **1979**, *18*, 1149−1170.

(11) Mayevsky, A. A.; Sukhorukov, B. I. *Nucleic. Acids Res.* **1980**, *8*, 3029.

(12) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 2309−2309.

(13) Mackerell, A. D.; Wiorkiewiczkuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.

(14) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810−2819.

(15) (a) Ringnalda, M. N.; Langlois, J-M.; Greeley, B. H.; Murphy, R. B.; Russo, T. V.; Cortis, C.; Muller, R. P.; Marten, B.; Donnelly, R. E.; Mainz, D. T.; Wright, J. R.; Pollard, W. T.; Cao, Y.; Won, Y.; Miller, G. H.; Goddard, W. A., III; Friesner, R. A. PS-GVB 2.34 from Schrödinger Inc., 1996. (b) Muller, R. P.; Langlois, J-M.; Ringnalda, M. N.; Friesner, R. A.; Goddard, W. A., III. *J. Chem. Phys.* **1994**, *100*, 1226. (c) Murphy, R. B.; Friesner, R. A.; Ringnalda, M. N.; Goddard, W. A., III. *J. Chem. Phys.* **1994**, *101*, 2986. (d) Langlois, J.-M.; Yamasaki, T.; Muller, R. P.; Goddard, W. A., III. *J. Phys. Chem.* **1994**, *98*, 13498. (e) Greeley, B. H.; Russo, T. V.; Mainz, D. T.; Friesner, R. A.; Langlois, J.-M.; Goddard, W. A., III; Donnelly, R. E.; Ringnalda, M. N. *J. Chem. Phys.* **1994**, *101*, 4028.

(16) Sponer, J.; Burcl, R.; Hobza, P. *J. Biomol. Struct. Dyn.* **1994**, *11*, 1357−1376.

(17) The PS-GVB implementation of this triple-ζ correlation-consistent Dunning[36] basis set does not include diffuse f functions on second-row elements [thus the notation cc-pVTZ9-(f)] and d functions on hydrogen; however this has been shown to have little effect on hydrogen bond energies.[2a]

(18) (a) Saebo, S.; Pulay, P. *Theor. Chim. Acta* **1986**, *69*, 357. (b) Saebo, S.; Pulay, P. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213. (c) Saebo, S.; Tong, W.; Pulay, P. *J. Chem. Phys.* **1993**, *98*, 2170−2175.

(19) Foster, J. M.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300.

(20) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *25*, 553.

(21) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M. N.; Goddard, W. A., III; Honig, B. *J. Am. Chem. Soc.* **1994**, *116*, 11875.

(22) Distributed by Molecular Simulations Inc., San Diego, CA.

(23) Mayo, S. L.; Olafson, B. D.; Goddard, W. A., III. *J. Phys. Chem.* **1990**, *94*, 8897−8909.

(24) Quantum mechanical calculations were performed on either HP 9000/735 CRX workstations, Cray C90 (San Diego Super Computer Center), or IBM SP2 (Cornell Theory Center). Molecular mechanics calculations were performed on either Silicon Graphics 4D/480 SMP server or Silicon Graphics Indigo2 Extreme workstations.

(25) If four conformations of the AT base pair contribute equally, the experimentally measured association constant (K_{obs}) will be overestimated by a factor of 4. That is $K_{obs} = 4K_{true}$. A substitution for K where $K = Ae^{-(\Delta H/KT)}$ leads to $Ae^{-(\Delta H_{obs}/KT)} = 4(Ae^{-(\Delta H_{true})})$. If $T = 333$, then $KT = 0.662$ kcal/mol leading to $\Delta H_{obs} + KT \ln(4) = \Delta H_{true}$. If $\Delta H_{obs} = -13.0$ kcal/mol, this leads to $\Delta H_{true} = -13.0 + 0.92 = -12.08$ kcal/mol.

(26) Latajka, Z. *J. Mol. Struct. (THEOCHEM)* **1991**, *251*, 245−260.

(27) Latajka, Z. *J. Mol. Struct. (THEOCHEM)* **1992**, *253*, 225−241.

(28) Curtiss, L. A.; Frurip, D. J.; Blander, M. *J. Chem. Phys.* **1979**, *71*, 2703.

(29) Xantheas, S. S.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *99*, 8774.

(30) Dasgupta, S.; Hammond, W. B.; Goddard, W. A., III. *J. Am. Chem. Soc.* **1996**, *118*, 12291−12301.

(31) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230−252.

(32) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1357−1377.

(33) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.

(34) The charges for HC1′ were set to 0.0800, and the charge at C1′ was adjusted so as to have a net neutral charge for each base.

(35) Hobza, P.; Hubalek, F.; Kabelac, M.; Mejzlik, P.; Sponer, J.; Vondrasek, J. *Chem. Phys. Lett.* **1996**, *257*, 31−35.

(36) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(37) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, Berlin, Heidleberg, Tokyo, 1984; pp 123−124, and references therein.