Accurate Structures and Binding Energies for Stacked Uracil Dimers

Matthew L. Leininger,[†] Ida M. B. Nielsen,*[,†] Michael E. Colvin,[‡] and Curtis L. Janssen[†]

Sandia National Laboratories, P.O. Box 969, Livermore, California 94551, and Computational Biology Group, Lawrence Livermore National Laboratory L-452, Livermore, California 94550

Received: October 17, 2001; In Final Form: January 29, 2002

The face-to-face and face-to-back stacked uracil dimers have been investigated by second-order Møller−Plesset (MP2) perturbation theory and by the coupled-cluster singles and doubles method augmented with a perturbative contribution from connected triple substitutions [CCSD(T)]. Full MP2 geometry optimizations were performed with a TZ2P(f,d)++ basis and with the 6-31G* basis for which harmonic vibrational frequencies were computed as well. Complete basis set MP2 binding energies were obtained from basis set extrapolations using the correlation-consistent basis sets cc-pVXZ (X = D−5) and aug-cc-pVXZ (X = D−Q). Higher-order correlation effects were gauged by computing the MP2 → CCSD(T) shift in the counterpoise-corrected binding energy using a modified 6-31G* basis set. By adding this correction to the infinite basis set limit MP2 binding energies, final estimates of 9.7 and 8.8 kcal mol^{-1} are obtained for the binding energies of the face-to-face and face-to-back structures, respectively.

1. Introduction

The structure of the DNA double helix has long been known to result from a balance of many intra- and intermolecular energy factors;[1] however, the relative strength of these different factors is still not quantitatively known. Because of the scientific and practical value of precise stacking energies for DNA and RNA bases (for example, in the design of synthetic DNA analogues[2]), this remains a topic of considerable interest. In particular, the energy factors contributing to nucleic acid base stacking are being evaluated by both experimental[3,4] and theoretical methods.[5,6] The application of ab initio quantum chemical methods to study base stacking is complicated by the need to include electron correlation to describe the dispersion interaction and the need for large basis sets to avoid basis set superposition error.[7] Note that dispersion is not properly included in any common density functionals, although functionals empirically including long-range dispersion interactions are under development.[8]

In this work, we compute accurate structures and binding energies of two stacked uracil dimers by means of second-order Møller−Plesset (MP2) perturbation theory and the coupled-cluster singles and doubles method augmented with a perturbative contribution from connected triple substitutions [CCSD(T)]. The two structures considered here are the face-to-face and face-to-back stacked uracil dimers, depicted in Figure 1. The face-to-face and face-to-back uracil dimers were previously investigated by Hobza and Šponer[9] in a study of stacked nucleic acid base pairs. The structures were optimized at the 6-31G* MP2 level of theory, and counterpoise-corrected binding energies were computed at the MP2 level using a 6-31G*(0.25) basis constructed by replacing the *d*-type polarization functions in the 6-31G* basis by more diffuse *d*-functions with exponents of 0.25. The nature of the stationary points was not ascertained by ab initio theory, but stacked uracil dimers corresponding to the face-to-face and face-to-back structures were found to be minima in an investigation of the uracil dimer potential energy surface by Kratochvíl et al.[10] using the AMBER 4.1 force field.

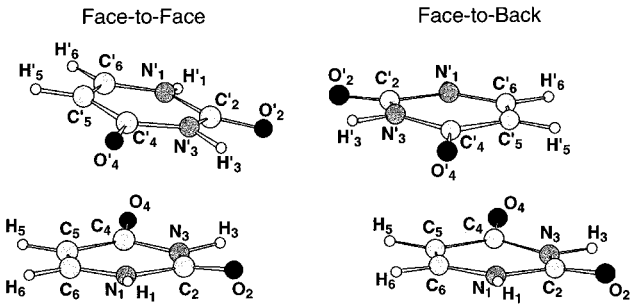

Figure 1. Face-to-face and face-to-back uracil dimers.

Using the MP2 method, we here compute geometries and harmonic vibrational frequencies with a 6-31G* basis and perform full geometry optimizations with an extended triple-ζ basis set with multiple diffuse and polarization functions. Moreover, using two series of correlation-consistent basis sets, ranging in size from 132 to 2020 basis functions, we study the basis set convergence of the MP2 binding energy and obtain complete basis set binding energies from basis set extrapolations. Finally, we investigate the effect of correlation beyond MP2 by computing counterpoise-corrected binding energies at the CCSD(T) level using the 6-31G*(0.25) basis set.

2. Computational Details

Stationary points were located at the MP2[11] level using a 6-31G* basis[12,13] and a triple-ζ quality basis augmented with polarization and diffuse functions, denoted TZ2P(f,d)++. The TZ2P(f,d)++ basis set was constructed by augmenting the Huzinaga−Dunning set of triple-ζ Gaussian functions[14] with two sets of *p*-type and one set of *d*-type functions on all hydrogen atoms and two sets of *d*-type and one set of *f*-type polarization functions on each first-row atom. The exponents for the polarization functions are $\alpha_p(H) = 1.50, 0.375;$

* E-mail: ibniels@ca.sandia.gov.
† Sandia National Laboratories.
‡ Lawrence Livermore National Laboratory.

10.1021/jp013866f CCC: $22.00
Published on Web 03/19/2002

$\alpha_d(H) = 1.00$; $\alpha_d(C) = 1.50, 0.375$; $\alpha_f(C) = 0.80$; $\alpha_d(N) = 1.60, 0.40$; $\alpha_f(N) = 1.00$; $\alpha_d(O) = 1.70, 0.425$; and $\alpha_f(O) = 1.40$. A diffuse *s*-type function (H,C,N,O) and a *p*-type function (C,N,O) were added according to the prescription of Lee and Schaefer[15]

$$\alpha_{\text{diffuse}} = \frac{1}{2}\left(\frac{\alpha_1}{\alpha_2} + \frac{\alpha_2}{\alpha_3}\right)\alpha_1 \quad (1)$$

where α_1, α_2, and α_3 are the three smallest Gaussian orbital exponents of the *s*- or *p*-type primitive functions for a given atom ($\alpha_1 < \alpha_2 < \alpha_3$). The final TZ2P(f,d)++ basis set contains 15 functions per H atom and 35 functions per C, N, or O atom for a total of 680 basis functions. For all located stationary points, the maximum component of the gradient was less than 10^{-4} hartree/bohr. The nature of the stationary points was ascertained by computation of MP2 6-31G* harmonic vibrational frequencies by finite differences of analytic gradients.

The MP2 binding energies for the face-to-face and face-to-back uracil dimers were computed at the optimum TZ2P(f,d)++ MP2 geometries using the correlation-consistent basis sets cc-pVXZ[16] (X = D, T, Q, 5) and the augmented sets aug-cc-pVXZ[16] (X = D, T, Q). For the monomers, MP2 computations were also carried out with the cc-pV6Z, aug-cc-pV5Z, and aug-cc-pV6Z sets to gauge the convergence toward the infinite basis set limit. Binding energies were also computed at the CCSD(T)[17] level using the 6-31G*(0.25) basis set. All binding energies were corrected for basis set superposition error by means of the counterpoise correction of Boys and Bernardi.[18] The uncorrected binding energy (D_e), the counterpoise correction (CPC), and the counterpoise-corrected binding energy (D_e^{CPC}) can be expressed as

$$D_e = 2E_m - E_d \quad (2)$$

$$\text{CPC} = \sum_{i=1}^{2} [E(i, d) - E(i, m)] \quad (3)$$

$$D_e^{\text{CPC}} = D_e + \text{CPC} \quad (4)$$

where $E(i, m)$ and $E(i, d)$ denote the energy of monomer i at the optimum geometry in the dimer using the monomer (m) and dimer (d) basis sets, respectively, E_m is the energy of the monomer at the optimum monomer geometry using the monomer basis, and E_d is the energy of the dimer.

The MP2 binding energies at the infinite basis set limit were estimated by extrapolating both the Hartree−Fock (HF) and MP2 correlation energies. The HF energies were extrapolated using the exponential form[19,20]

$$E_X^{\text{HF}} = E_\infty^{\text{HF}} + A\exp(-BX) \quad (5)$$

where X is the highest angular momentum represented in the basis set. MP2 correlation energies were extrapolated using the two-point formula[21]

$$E_\infty^{(2)} = \frac{X^3 E_X^{(2)} - (X-1)^3 E_{X-1}^{(2)}}{X^3 - (X-1)^3} \quad (6)$$

employing two basis sets with cardinal numbers $X - 1$ and X, respectively.

The coupled-cluster computations were performed with the PSI3 program,[22] and all other computations were carried out with the massively parallel quantum chemistry (MPQC) package.[23−26] The frozen core approximation was employed throughout.

TABLE 1: Selected TZ2P(f,d)++ MP2 Optimum Geometrical Parameters (Å) for the Uracil Monomer and the Face-to-face and Face-to-back Dimers[a]

monomer		face-to-face		face-to-back	
$r(N_1H_1)$	1.0065	$r(N_1H_1)$	1.0079	$r(N_1H_1)$	1.0084
$r(C_2O_2)$	1.2151	$r(C_2O_2)$	1.2166	$r(C_2O_2)$	1.2154
$r(N_3H_3)$	1.0109	$r(N_3H_3)$	1.0137	$r(N_3H_3)$	1.0113
$r(C_4O_4)$	1.2188	$r(C_4O_4)$	1.2232	$r(C_4O_4)$	1.2233
$r(C_5H_5)$	1.0755	$r(C_5H_5)$	1.0758	$r(C_5H_5)$	1.0767
$r(C_6H_6)$	1.0791	$r(C_6H_6)$	1.0790	$r(C_6H_6)$	1.0793
$r(N_1C_2)$	1.3847	$r(N_1C_2)$	1.3810	$r(N_1C_2)$	1.3817
$r(C_2N_3)$	1.3794	$r(C_2N_3)$	1.3820	$r(C_2N_3)$	1.3827
$r(N_3C_4)$	1.4033	$r(N_3C_4)$	1.3999	$r(N_3C_4)$	1.3993
$r(C_4C_5)$	1.4525	$r(C_4C_5)$	1.4500	$r(C_4C_5)$	1.4490
$r(C_5C_6)$	1.3475	$r(C_5C_6)$	1.3491	$r(C_5C_6)$	1.3502
$r(C_6N_1)$	1.3707	$r(C_6N_1)$	1.3679	$r(C_6N_1)$	1.3682
		$r(N_1C'_4)$	3.2632	$r(N_1C'_4)$	3.2355
		$r(C_2N'_3)$	2.9872	$r(C_2C'_5)$	3.2695
		$r(N_3C'_2)$	2.9874	$r(N_3C'_6)$	3.3261
		$r(C_4N'_1)$	3.2638	$r(C_4N'_1)$	3.2355
		$r(C_5C'_6)$	3.6449	$r(C_5C'_2)$	3.2695
		$r(C_6C'_5)$	3.6446	$r(C_6N'_3)$	3.3261
		$r(H_1O'_4)$	3.0188	$r(H_1O'_4)$	3.0091
		$r(O_2H'_3)$	2.6358	$r(O_2H'_5)$	3.2292
		$r(H_3O'_2)$	2.6365	$r(H_3H'_6)$	3.2925
		$r(O_4H'_1)$	3.0199	$r(O_4H'_1)$	3.0091
		$r(H_5H'_6)$	3.9511	$r(H_5O'_2)$	3.2292
		$r(H_6H'_5)$	3.9505	$r(H_6H'_3)$	3.2925

[a] See Figure 1 for numbering of atoms.

3. Results and Discussion

Selected optimum TZ2P(f,d)++ MP2 geometrical parameters for the monomer and the two uracil dimers are given in Table 1. Cartesian geometries are available upon request from the authors. The monomer geometries in the face-to-face structure are nearly the same as in the face-to-back structure, and the two dimers are distinguished mainly by the relative orientation of the monomers, either face-to-face or face-to-back (cf. Figure 1). The face (back) side of a monomer ring is defined as the side from which the direction of the atoms N_1–C_2–N_3 is counterclockwise (clockwise).

The face-to-face structure was optimized in C_1 symmetry, producing a structure very close to C_2 symmetry. All symmetry-related bond distances differ by less than 10^{-4} Å; for intermonomer O−H, H−H, C−C, and C−N distances, symmetry related distances differ by up to about 0.001 Å. A tighter optimization in C_1 with the smaller 6-31G* basis set yielded a near-C_2 structure, which, when symmetrized by averaging symmetry-related coordinates, was found to be a minimum (vide infra) and which had a maximum gradient component of $2.3 \cdot 10^{-5}$ hartree/bohr. The two rings in the face-to-face dimer form an open V structure, thereby reducing the intermonomer $O_2-H'_3$ and $H_3-O'_2$ distances, which take on a value of 2.64 Å. The $H_1-O'_4$ and $O_4-H'_1$ distances are 3.02 Å, and the closest distances between heavy atoms in opposite rings range from 2.99 to 3.64 Å. The face-to-back dimer has C_i symmetry, and the two rings are parallel. The intermonomer O−H distances are longer than in the face-to-face structure, namely 3.01, and 3.23 Å, and all the intermonomer C−C and C−N distances are similar, ranging from 3.24 to 3.37 Å. While the uracil monomer is planar, the monomers in the dimers are distorted somewhat from planarity. Thus, in both the face-to-face and face-to-back dimers, the rings are slightly puckered in a way that increases

TABLE 2: Harmonic Vibrational Frequencies (cm^{-1}) for the Uracil Monomer and Dimers Computed at the 6-31G* MP2 Level of Theory[a]

monomer (C_s)				face-to-back (C_i)				face-to-face (C_2)			
A'		A''		A_g		A_u		A		B	
ω_1	3657			ω_1	3630	ω_{34}	3630	ω_1	3637	ω_{35}	3637
ω_2	3613			ω_2	3605	ω_{35}	3605	ω_2	3567	ω_{36}	3568
ω_3	3306			ω_3	3297	ω_{36}	3297	ω_3	3303	ω_{37}	3303
ω_4	3266			ω_4	3266	ω_{37}	3266	ω_4	3267	ω_{38}	3267
ω_5	1864			ω_5	1858	ω_{38}	1861	ω_5	1851	ω_{39}	1849
ω_6	1821			ω_6	1805	ω_{39}	1814	ω_6	1805	ω_{40}	1816
ω_7	1710			ω_7	1704	ω_{40}	1705	ω_7	1709	ω_{41}	1707
ω_8	1533			ω_8	1533	ω_{41}	1538	ω_8	1538	ω_{42}	1534
ω_9	1447			ω_9	1450	ω_{42}	1452	ω_9	1456	ω_{43}	1447
ω_{10}	1436			ω_{10}	1434	ω_{43}	1435	ω_{10}	1440	ω_{44}	1437
ω_{11}	1411			ω_{11}	1406	ω_{44}	1406	ω_{11}	1420	ω_{45}	1418
ω_{12}	1270			ω_{12}	1273	ω_{45}	1273	ω_{12}	1273	ω_{46}	1271
ω_{13}	1236			ω_{13}	1234	ω_{46}	1241	ω_{13}	1237	ω_{47}	1242
ω_{14}	1112			ω_{14}	1115	ω_{47}	1116	ω_{14}	1115	ω_{48}	1115
ω_{15}	999			ω_{15}	1004	ω_{48}	1002	ω_{15}	1006	ω_{49}	1006
ω_{16}	989			ω_{16}	991	ω_{49}	991	ω_{16}	996	ω_{50}	994
		ω_{22}	920	ω_{17}	928	ω_{50}	926	ω_{17}	923	ω_{51}	924
		ω_{23}	787	ω_{18}	791	ω_{51}	792	ω_{18}	792	ω_{52}	789
ω_{17}	782			ω_{19}	782	ω_{52}	781	ω_{19}	783	ω_{53}	781
		ω_{24}	725	ω_{20}	730	ω_{53}	730	ω_{20}	772	ω_{54}	742
		ω_{25}	709	ω_{21}	718	ω_{54}	717	ω_{21}	715	ω_{55}	715
		ω_{26}	686	ω_{22}	691	ω_{55}	692	ω_{22}	692	ω_{56}	690
		ω_{27}	559	ω_{23}	593	ω_{56}	587	ω_{23}	586	ω_{57}	579
ω_{18}	561			ω_{24}	561	ω_{57}	561	ω_{24}	560	ω_{58}	561
ω_{19}	541			ω_{25}	544	ω_{58}	543	ω_{25}	541	ω_{59}	540
ω_{20}	518			ω_{26}	509	ω_{59}	510	ω_{26}	512	ω_{60}	511
ω_{21}	383			ω_{27}	398	ω_{60}	396	ω_{27}	395	ω_{61}	398
		ω_{28}	371	ω_{28}	377	ω_{61}	377	ω_{28}	373	ω_{62}	375
		ω_{29}	159	ω_{29}	178	ω_{62}	190	ω_{29}	197	ω_{63}	182
		ω_{30}	134	ω_{30}	156	ω_{63}	163	ω_{30}	177	ω_{64}	158
				ω_{31}	101	ω_{64}	78	ω_{31}	103	ω_{65}	65
				ω_{32}	62	ω_{65}	37	ω_{32}	70	ω_{66}	17
				ω_{33}	31	ω_{66}	31	ω_{33}	43		
								ω_{34}	23		

[a] The monomer frequencies are lined up with the corresponding dimer frequencies.

the intermonomer C−C and C−N distances. Also, the exocyclic N−H bonds across from an oxygen atom in the opposite monomer point toward the opposite monomer, thus reducing the intermonomer O−H distances. All N−H, C−O, and C−N bonds in the dimers are longer than their monomer counterparts, except for the C_2–N_3 bond, and the dimer C−C bonds are shorter than the corresponding C−C bonds in the monomer.

The previously reported ab initio geometries for the face-to-face and face-to-back dimers were computed at the 6-31G* MP2 level.[9] We note that, in the face-to-face dimer, the intermonomer distances O_2−H'_3 and H_3−O'_2 computed with the 6-31G* basis set are significantly shorter than those obtained with the much larger TZ2P(f,d)++ set, viz., 2.48 vs 2.64 Å. A similar trend, though less pronounced, is observed for the face-to-back structure, where the H_1−O'_4 and O_4−H'_1 distances obtained with the 6-31G* and TZ2P(f,d)++ sets are 2.90 and 3.01 Å, respectively.

Harmonic vibrational frequencies were computed for the monomer and both dimers at the MP2 6-31G* level. The frequencies and their assignments are listed in Tables 2 and 3. In the dimers, each monomer mode is split into two modes, one symmetric and one antisymmetric, and six new modes appear; the six new modes, corresponding to intermonomer vibrations, have very low frequencies, around 100 cm^{-1} or less. For both dimers, most frequency pairs are shifted by less than 10 cm^{-1} relative to the corresponding monomer frequency. The frequencies for which larger shifts are encountered include the N−H stretches (ω_1, ω_2) and C−O stretches (ω_5, ω_6) as well as

TABLE 3: Assignments of the Vibrational Modes for the Uracil Monomer and Dimers[a]

assignment	monomer	face-to-back	face-to-face
N_1H_1 stretch	ω_1	ω_1, ω_{34}	ω_1, ω_{35}
N_3H_3 stretch	ω_2	ω_2, ω_{35}	ω_2, ω_{36}
C_5H_5 stretch	ω_3	ω_3, ω_{36}	ω_3, ω_{37}
C_6H_6 stretch	ω_4	ω_4, ω_{37}	ω_4, ω_{38}
C_2O_2 stretch	ω_5	ω_5, ω_{38}	ω_5, ω_{39}
C_4O_4 stretch	ω_6	ω_6, ω_{39}	ω_6, ω_{40}
C_5C_6 stretch	ω_7	ω_7, ω_{40}	ω_7, ω_{41}
in-plane deformation	ω_8−ω_{16}	ω_8−ω_{16}, ω_{41}−ω_{49}	ω_8−ω_{16}, ω_{42}−ω_{50}
	ω_{17}	ω_{19}, ω_{52}	ω_{19}, ω_{53}
	ω_{18}−ω_{21}	ω_{24}−ω_{27}, ω_{57}−ω_{60}	ω_{24}−ω_{27}, ω_{58}−ω_{61}
out-of-plane deformation	ω_{22}, ω_{23}	ω_{17}, ω_{18}, ω_{50}, ω_{51}	ω_{17}, ω_{18}, ω_{51}, ω_{52}
	ω_{24}−ω_{27}	ω_{20}−ω_{23}, ω_{53}−ω_{56}	ω_{20}−ω_{23}, ω_{54}−ω_{57}
	ω_{28}−ω_{30}	ω_{28}−ω_{30}, ω_{61}−ω_{63}	ω_{28}−ω_{30}, ω_{62}−ω_{64}
intermonomer vibrations		ω_{31}−ω_{33}	ω_{31}−ω_{34}
		ω_{64}−ω_{66}	ω_{65}−ω_{66}

[a] For previous assignments of the monomer vibrational modes, see, elsewhere.[28−31]

several lower-frequency modes representing out-of-plane ring deformations coupled to "wagging" of N−H and C−O bonds (ω_{24}, ω_{27}, ω_{29}, ω_{30}), and an in-plane vibration (ω_{21}) involving wagging of the C−O bonds and bending of the C_2–N_3–C_4 angle; the numbers here refer to the monomer normal modes.

The binding energies of the face-to-face and face-to-back structures were computed at the MP2 and CCSD(T) levels at the optimum TZ2P(f,d)++ MP2 geometries. MP2 binding energies were computed using the cc-pVXZ, X = D−5, and aug-cc-pVXZ, X = D−Q, basis sets, and the 6-31G*(0.25) basis set was employed at the CCSD(T) level. The MP2 binding energies in the infinite basis set limit were computed for both series of correlation-consistent basis sets using eqs 5 and 6 for the HF and MP2 correlation energies, respectively. The computed binding energies and counterpoise corrections are listed in Tables 4 and 5.

Considering first the face-to-face binding energies computed with the aug-cc-pVXZ basis sets, we note that the uncorrected MP2 binding energy is rather sensitive to basis set improvement, assuming values of 15.57, 12.83, 11.53 kcal mol^{-1} for X = D, T, and Q, respectively. The counterpoise-corrected MP2 energy converges more rapidly, taking on values of 9.14, 10.08, and 10.29 kcal mol^{-1} for X = D, T, and Q, respectively. The MP2 counterpoise correction is sizable, −6.43 kcal mol^{-1}, at the aug-cc-pVDZ level, but it decreases by more than a factor of 2 with each basis set improvement and assumes a value of −1.24 kcal mol^{-1} with the aug-cc-pVQZ basis. The HF binding energy is close to the infinite basis set limit at the aug-cc-pVQZ level, as indicated by the aug-cc-pVQZ counterpoise correction of only −0.13 kcal mol^{-1} as well as the apparent convergence of the counterpoise-corrected HF binding energy which equals −3.00 and −3.02 kcal mol^{-1} with the aug-cc-pVTZ and aug-cc-pVQZ basis sets, respectively. These values agree well with the infinite basis set limit of −3.04 kcal mol^{-1} obtained by basis set extrapolation. Using the extrapolated HF binding energy of −3.04 kcal mol^{-1} in conjunction with an extrapolated MP2 correlation energy of 13.76 kcal mol^{-1}, we arrive at a final estimate of 10.7 kcal mol^{-1} for the complete basis set MP2 binding energy of the face-to-face structure.

For the nonaugmented basis sets, the uncorrected MP2 binding energies appear to converge faster than for the augmented series. This is due to a fortuitous cancellation of errors: without diffuse functions, the correlation contribution to the binding energy is far too small, but the basis set superposition error inherent in the uncorrected results raises the

TABLE 4: MP2 Binding Energies (kcal mol^{-1}) for the Face-to-Face and Face-to-Back Structures of the Uracil Dimer[a]

	face-to-face			face-to-back		
basis set[b]	D_e	CPC	D_e^{CPC}	D_e	CPC	D_e^{CPC}
cc-pVDZ (264)	11.41 (0.06)	−6.72 (−3.36)	4.69 (−3.30)	10.32 (−0.82)	−5.69 (−2.80)	4.62 (−3.61)
cc-pVTZ (592)	11.69 (−1.82)	−3.44 (−1.24)	8.25 (−3.06)	11.05 (−2.38)	−3.09 (−1.13)	7.96 (−3.51)
cc-pVQZ (1120)	11.18 (−2.58)	−1.52 (−0.45)	9.66 (−3.03)	10.65 (−3.03)	−1.37 (−0.41)	9.28 (−3.45)
cc-pV5Z (1896)	10.83 (−2.93)			10.35 (−3.35)		
∞[c]	10.66 (−3.11)			10.21 (−3.52)		
aug-cc-pVDZ (440)	15.57 (−1.14)	−6.43 (−1.72)	9.14 (−2.86)	15.21 (−1.48)	−6.36 (−1.77)	8.84 (−3.25)
aug-cc-pVTZ (920)	12.83 (−2.50)	−2.75 (−0.50)	10.08 (−3.00)	12.30 (−2.94)	−2.63 (−0.47)	9.68 (−3.41)
aug-cc-pVQZ (1648)	11.53 (−2.89)	−1.24 (−0.13)	10.29 (−3.02)	11.04 (−3.31)	−1.12 (−0.12)	9.92 (−3.43)
∞[d]	10.72 (−3.04)			10.26 (−3.44)		

[a] Hartree−Fock results are given in parentheses. [b] The number of basis functions is given in parentheses. [c] Extrapolation to the infinite basis set limit (see text) using X = T, Q, 5 for the HF energies and X = Q, 5 for the MP2 correlation energies. [d] Extrapolation to the infinite basis set limit (see text) using X = D, T, Q for the HF energies and X = T, Q for the MP2 correlation energies.

TABLE 5: Binding Energies (kcal mol^{-1}) for the Face-to-Face and Face-to-Back Structures of the Uracil Dimer Computed with the 6.31G*(0.25) Basis Set

	face-to-face			face-to-back		
method	D_e	CPC	D_e^{CPC}	D_e	CPC	D_e^{CPC}
SCF	2.79	−5.12	−2.32	1.12	−4.10	−2.99
MP2	19.42	−11.39	8.03	17.57	−9.98	7.58
CCSD	16.27	−10.98	5.29	14.01	−9.51	4.50
CCSD(T)	18.65	−11.60	7.05	16.21	−10.08	6.14

binding energy. When the counterpoise correction is applied to the nonaugmented results, the convergence, as expected, is slower than that observed for the augmented series. Extrapolation of the HF and MP2 correlation energies yield an MP2 binding energy for the face-to-face structure of 10.66 kcal mol^{-1} in excellent agreement with that obtained for the augmented series.

The trends observed for the face-to-back structure closely parallel those for the face-to-face structure discussed above. Again, the HF binding energy appears to be almost converged at the aug-cc-pVQZ level, assuming a value of −3.43 kcal mol^{-1}, which is close to the extrapolated value of −3.44 kcal mol^{-1}. The aug-cc-pVQZ MP2 binding energy is 11.04 kcal mol^{-1}, and after extrapolation, a final estimate for the infinite basis set MP2 binding energy of 10.26 kcal mol^{-1} is obtained for the face-to-back uracil dimer. Extrapolation to the infinite basis set limit using the nonaugmented basis sets yields an MP2 binding energy of 10.21 kcal mol^{-1}, in close agreement with the result obtained for the augmented basis sets.

In addition to the basis set extrapolations employed in Table 4, several other extrapolations were performed, again using eqs 5 and 6, including both smaller basis sets and the cc-pV6Z, aug-cc-pV5Z, and aug-cc-pV6Z sets that are applicable for the monomer (ranging in size from 1336 to 2020 basis functions). The results are listed in Table 6. For the binding energy, the employed extrapolation schemes agree to within about 0.2 kcal mol^{-1}, except for schemes using a DZ basis set for the MP2 correlation energy. The employed extrapolation schemes that produce similar binding energies nonetheless differ in their ability to reproduce total energies. Considering the HF energies for the monomer, the nonaugmented (TZ,QZ,5Z) sets produce monomer energies within ca. 0.7 kcal mol^{-1} of the best scheme (QZ,5Z,6Z), whereas the augmented (DZ,TZ,QZ) scheme produces a monomer energy about 1.9 kcal mol^{-1} below the best energy. Likewise, for the monomer MP2 correlation energy, the nonaugmented (QZ,5Z) scheme produces an energy within 0.6 kcal mol^{-1} of the (5Z,6Z) result, whereas the augmented (TZ,QZ) extrapolation differs from the augmented (5Z,6Z) energy by about 5.1 kcal mol^{-1}.

Comparing our best estimates for the complete basis set MP2 binding energies for the face-to-face and face-to-back uracil dimers, we note that the correlation parts of the binding energy are almost equal for the two structures, viz., 13.77 and 13.73 kcal mol^{-1}. Thus, the difference in binding energies is mainly due to differences in the HF binding energies, assuming values of −3.04 and −3.44 kcal mol^{-1}, respectively, for the face-to-face and face-to-back structures. Our MP2 binding energies of 10.7 and 10.3 kcal mol^{-1} for the face-to-face and face-to-back

TABLE 6: HF and MP2 Correlation Energies[a] Extrapolated to the Infinite Basis Set Limit using Eqs 5 and 6

	total energy			binding energy	
basis sets	monomer	face-to-face	face-to-back	face-to-face	face-to-back
		HF			
cc-pVXZ; X = D, T, Q	−412.658847	−825.312958	−825.312324	−2.972	−3.370
cc-pVXZ; X = T, Q, 5	−412.657539	−825.310127	−825.309475	−3.107	−3.516
cc-pVXZ; X = Q, 5, 6	−412.656484				
aug-cc-pVXZ; X = D, T, Q	−412.659559	−825.314271	−825.313641	−3.042	−3.437
aug-cc-pVXZ; X = T, Q, 5	−412.657232				
aug-cc-pVXZ; X = Q, 5, 6	−412.656498				
		MP2			
cc-pVXZ; X = D, T	−1.593472	−3.209928	−3.209891	14.423	14.399
cc-pVXZ; X = T, Q	−1.653445	−3.329110	−3.328987	13.944	13.867
cc-pVXZ; X = Q, 5	−1.662644	−3.347235	−3.347159	13.772	13.724
cc-pVXZ; X = 5, 6	−1.663592				
aug-cc-pVXZ; X = D, T	−1.607383	−3.238266	−3.238093	14.746	14.638
aug-cc-pVXZ; X = T, Q	−1.656063	−3.334051	−3.333960	13.758	13.701
aug-cc-pVXZ; X = Q, 5	−1.662927				
aug-cc-pVXZ; X = 5, 6	−1.664190				

[a] Total energies in hartrees, binding energies in kcal mol^{-1}.

structures are significantly higher than the 6-31G*(0.25) MP2 values of 7.7 and 7.4 kcal mol^{-1}, respectively, obtained by Hobza and co-workers.[9,10] Our binding energies were computed at optimum TZ2P(f,d)++ MP2 geometries, whereas those of Hobza and co-workers were computed at 6-31G* MP2 geometries. To investigate the effect of the improved geometry, we therefore computed counterpoise-corrected TZ2P(f,d)++ MP2 binding energies at both the 6-31G* and TZ2P(f,d)++ MP2 optimum geometries. The counterpoise-corrected MP2 binding energies thus obtained for the face-to-face and face-to-back structures were, respectively, 8.86 and 8.48 kcal mol^{-1} for the 6-31G* geometries and 9.02 and 8.68 kcal mol^{-1} for the TZ2P-(f,d)++ geometries, showing a small effect of about 0.2 kcal mol^{-1} of improving the geometry beyond the 6-31G* MP2 level.

To estimate higher-order correlation effects, counterpoise-corrected binding energies were computed at the 6-31G*(0.25) CCSD(T) level using the TZ2P(f,d)++ MP2 optimum geometries (Table 5). The 6-31G*(0.25) basis set was chosen because it is of a relatively modest size (slightly smaller than the cc-pVDZ basis) but provides counterpoise-corrected MP2 binding energies roughly comparable with those obtained with the larger aug-cc-pVDZ or cc-pVTZ basis sets. Higher-order correlation effects are found to destabilize the dimers: The counterpoise-corrected binding energies decrease by 0.98 and 1.45 kcal mol^{-1} for the face-to-face and face-to-back structures, respectively, when improving the level of correlation from MP2 to CCSD-(T). We note that a previous study of several stacked aromatic complexes smaller than the uracil dimer also found the complexes to be destabilized upon improving the correlation level from MP2 to CCSD(T).[27] Adding the 6-31G*(0.25) MP2 → CCSD(T) shift in the counterpoise-corrected binding energies to the infinite basis set MP2 limits computed above, we obtain final estimates of 9.7 and 8.8 kcal mol^{-1} for the binding energy of the face-to-face and face-to-back stacked uracil dimers, respectively.

Concluding Remarks

Accurate structures and binding energies have been computed for the face-to-face and face-to-back stacked uracil dimers using the MP2 and CCSD(T) methods. MP2 optimum geometries were obtained with a triple-ζ basis set with multiple diffuse and polarization functions [TZ2P(f,d)++] and with a 6-31G* basis set for which harmonic vibrational frequencies were computed as well, and both structures were found to be minima. MP2 binding energies, including a counterpoise correction, were computed using the correlation-consistent basis sets cc-pVXZ (X = D, T, Q, 5) and aug-cc-pVXZ (X = D, T, Q), and basis set extrapolation yielded complete basis set MP2 binding energies of 10.7 and 10.3 kcal mol^{-1} for the face-to-face and face-to-back structures, respectively. A basis set of at least aug-cc-pVTZ quality was found to be required to obtain a counterpoise-corrected binding energy within 1 kcal mol^{-1} of the infinite basis set limit. Binding energies, computed at the TZ2P-(f,d)++ MP2 level, were found to differ by only ca. 0.2 kcal mol^{-1} when computed at optimum 6-31G* and optimum TZ2P-(f,d)++ geometries, demonstrating that the effect on the binding energies of using extended basis sets for geometry optimizations is small. The effect of correlation beyond MP2 was investigated by means of the CCSD(T) method. By use of a 6-31G*(0.25) basis set, the MP2 → CCSD(T) shift in the counterpoise-corrected binding energies was found to be −0.98 and −1.45 kcal mol^{-1} for the face-to-face and face-to-back structures, respectively. Adding this correction to the estimated infinite basis set MP2 binding energies, our final estimate for the binding energies of the face-to-face and face-to-back stacked uracil dimers is 9.7 and 8.8 kcal mol^{-1} respectively.

Acknowledgment. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000. The work at Lawrence Livermore National Laboratory was performed under the auspices of the U.S. Department of Energy, Contract Number W-7405-ENG-48. The authors acknowledge a generous allotment of computing time on the IBM SP ASCI White computer from Livermore Computing at Lawrence Livermore National Laboratory.

References and Notes

(1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(2) Kool, E. T. *Chem. Rev.* **1997**, *97*, 1473.
(3) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X.-F.; Sheils, C. J.; Tahmassebi, D. C.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 2213.
(4) McKay, S. L.; Haptonstall, B.; Gellman, S. H. *J. Am. Chem. Soc.* **2001**, *123*, 1244.
(5) Luo, R.; Gilson, H. S. R.; Potter, M. J.; Gilson, M. K. *Biophys. J.* **2001**, *80*, 140.
(6) Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, *99*, 3247.
(7) Rappé, A. K.; Bernstein, E. R. *J. Phys. Chem. A* **2000**, *104*, 6117.
(8) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149.
(9) Hobza, P.; Šponer, J. *Chem. Phys. Lett.* **1998**, *288*, 7.
(10) Kratochvíl, M.; Engkvist, O.; Šponer, J.; Jungwirth, P.; Hobza, P. *J. Phys. Chem. A* **1998**, *102*, 6921.
(11) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
(12) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
(13) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
(14) Dunning, T. H., Jr. *J. Chem. Phys.* **1971**, *55*, 716.
(15) Lee, T. J.; Schaefer, H. F. *J. Chem. Phys.* **1985**, *83*, 1784.
(16) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
(17) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.
(18) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(19) Jensen, F. *J. Chem. Phys.* **1999**, *110*, 6601.
(20) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437.
(21) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.
(22) Crawford, T. D.; Sherrill, C. D.; Valeev, E. F.; Fermann, J. T.; Leininger, M. L.; King, R. A.; Brown, S. T.; Janssen, C. L.; Seidl, E. T.; Yamaguchi, Y.; Allen, W. D.; Xie, Y.; Vacek, G.; Hamilton, T. P.; Kellogg, C. B.; Remington, R. B.; Schaefer, H. F., III. *PSI 3.0* PSITECH Inc.: Watkinsville, GA, 1999.
(23) Colvin, M. E.; Janssen, C. L.; Whiteside, R. A.; Tong, C. H. *Theor. Chim. Acta* **1993**, *84*, 301.
(24) Janssen, C. L.; Seidl, E. T.; Colvin, M. E. In *Parallel Computing in Computational Chemistry*; Mattson, T. G., Ed.; ACS Symposium Series, Vol. 592; American Chemical Society: Washington, DC, 1995; p 47.
(25) Nielsen, I. M. B.; Seidl, E. T. *J. Comput. Chem.* **1995**, *16*, 1301.
(26) Nielsen, I. M. B. *Chem. Phys. Lett.* **1996**, *255*, 210.
(27) Šponer, J.; Hobza, P. *Chem. Phys. Lett.* **1997**, *267*, 263.
(28) Bencivenni, L.; Ramondo, F.; Pieretti, A.; Sanna, N. *J. Chem. Soc., Perkin Trans. 2* **2000**, 1685.
(29) Colarusso, P.; Zhang, K.; Guo, B.; Bernath, P. F. *Chem. Phys. Lett.* **1997**, *269*, 39.
(30) Estrin, D. A.; Paglieri, L.; Corongiu, G. *J. Phys. Chem.* **1994**, *98*, 5653.
(31) Leś, A.; Adamowicz, L.; Nowak, M. J.; Lapinski, L. *Spectrochim. Acta, Part A* **1992**, *48*, 1385.