Ab Initio Study of Stacking Interactions in A- and B-DNA

Cristóbal Alhambra,[†] Francisco J. Luque,*[,‡] Federico Gago,[§] and Modesto Orozco*[,†]

Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain, Departament de Farmàcia, Unitat Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avgda Diagonal s/n, Barcelona 08028, Spain, Departamento de Fisiología y Farmacología, Universidad de Alcalá de Henares, 28871 Madrid, Spain

Received: August 23, 1996; In Final Form: December 30, 1996[⊗]

High level *ab initio* methods have been used to study stacking interactions in standard A- and B-DNA, as well as in selected B-type crystal structures. A combination of quantum mechanical and classical mechanical methods has been used to examine cooperative contributions to stacking in DNA polymers. The results provide valuable information on intrastrand, interstrand, and cooperative contributions to stacking in DNA, and are useful for understanding the role of stacking interactions in the stabilization of DNA structures.

Introduction

Stacking of aromatic systems is an essential interaction in a large number of chemical and biochemical processes. These interactions contribute significantly to the structure and stability of biomacromolecules. The classic example is the stacking between nucleic acid bases, which determines the structure and flexibility of DNA and RNA.[1] In this context, detailed knowledge of stacking interactions may help in the design of drugs intercalating within DNA.[2] Stacking interactions are also important for understanding of DNA sequence-dependent polymorphism, which is mainly determined by the properties of individual base pairs.[3] Stacking is also known to influence the physicochemical properties of DNA fibers, such as bending propensity, resistance to opening, and the ability to form intercalation sites. All of these considerations justify the great research effort devoted to understanding the nature of stacking interactions. A detailed knowledge of the physical factors governing this interaction will give deeper insight into the structural and functional implications of aromatic interactions in biomolecules, particularly the nucleic acids, and to the design of new intercalating drugs with potential therapeutic value.

Theoretical techniques, in particular high-level quantum mechanical (QM) *ab initio* methods, are valuable for the analysis of the basic energy contributions of intermolecular interactions, such as hydrogen-bond pairing or stacking interactions between nucleic acid bases. However, a QM approach is severely limited by (i) the large size of the system (up to 24 second-row atoms for the stacking between base-pair dimers), (ii) the need to use basis sets sufficiently extended and flexible enough to describe π-electron distributions, (iii) the importance of correlation effects, and (iv) the correction of the basis set superposition error (BSSE). An additional difficulty is that the study of stacking complexes for nucleic acid bases taken directly from B- or A-DNA crystal structures are not real minima on the gas phase potential energy surface. As a result, most of the theoretical studies to date have been performed using classical force fields,[4] and few QM studies have been reported.[5] These studies have yielded valuable information on the nature of base stacking, but a complete description of this interaction has not yet been obtained.

In this paper, a systematic QM study of the stacking of nucleic acid bases in standard B- and A-DNA conformations, as well as in a few B-type crystal structures, is presented. *Ab initio* methods have been used to determine the interaction energies for the stacking between the bases and the hydrogen-bonded base-pair dimers in standard B- and A-DNA structures. Results using different classical force fields have been compared with the *ab initio* data and were used to explore the contribution of 1−3 stacking interactions (intrastrand interactions between two bases separated by another base) to the DNA stability. Finally, calculations on crystal structures were used to explore the importance of local geometrical changes in the stability of stacked complexes. The results provide a valuable picture of stacking interactions in two of the canonical conformations of DNA.

Methods

The geometries of methylated pyrimidine (N1) and purine (N9) bases were fully optimized in the gas phase at the HF/6-31G(d) level. The optimized geometries were not further reoptimized in B- and A-DNA-type-stacked complexes to avoid artefactual changes in geometry due to the neglect of the DNA environment, as well as to effects of BSSE error.[5e] For the same reasons, the inter-base geometry was fixed to the values of standard B- and A-DNA fibers,[6] as implemented in the biopolymer module of the Insight-II computer program.[7] The methyl group was oriented to mimic the carbon C1′ of the 2′-deoxyribose of the DNA fiber. In stacked complexes containing thymine, the orientation of the methyl group at C5 was optimized by systematic rotation around the C4−CH_3 bond at the classical level.

The intrastrand stacking energy between bases was determined as the difference in the energy of the complex and the energy of the isolated bases (Figure 1 and ref 5f). Single-point calculations were performed at the MP2/6-31G(d) level[8] and the BSSE error was corrected using the standard counterpoise method.[9] MP2 calculations were necessary due to the importance of dispersion interactions in these systems. However, these were extremely expensive from a computational point of view since they involved nearly 400 basis functions, which implies several CPU days in a SGI-R8000 workstation for each individual computation. No further enlargements of the basis set or increase in the level of theory to capture correlation effects is possible with current computational resources. However, test

[†] Departament de Bioquímica i Biologia Molecular.
[‡] Departament de Farmàcia.
[§] Departamento de Fisiología y Farmacología.
* Author to whom correspondence should be addressed.
[⊗] Abstract published in *Advance ACS Abstracts,* March 15, 1997.

S1089-5647(96)02626-0 CCC: $14.00

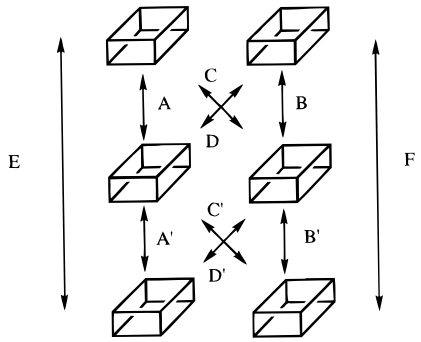

Figure 1. Schematic representation of the different energy contributions that modulate stacking between base-pair dimers. See text for more details.

calculations at the self-consistent field (SCF) level were performed for several dimers to verify the suitability of the 6-31G(d) basis by comparison with the results derived from the 6-311+G(d,p) basis.[10] After the counterpoise correction, the small difference (a few tenths of a kcal/mol) between results estimated with the two basis sets gives relative confidence in the preciseness of the Hartree−Fock (HF) results reported here (see below). Unfortunately, there is more concern on the suitability of the 6-31G(d) basis set to account for dispersion interactions, since it has been discussed[5e−h,11] that the standard 6-31G(d) is not flexible enough to represent properly dispersion interactions. In order to estimate the expected range of error in our results due to the low flexibility of the 6-31G(d) basis set, we performed additional test calculations, where the results derived at the MP2/6-31G(d) level were compared with the values determined from MP2/6-311+G(d) calculations, as well as from MP2/diffuse 6-31G(d)[12] data (see below).

The stacking energy of hydrogen-bonded base-pair dimers was computed by assuming additivity[5d] of the corresponding intrastrand and interstrand contributions (Figure 1), this latter term being the cross-interaction between the bases in different chains of the DNA. The interstrand contribution was determined for the different base combinations in B- and A-DNA conformations at the MP2/6-31G(d) level. The *ab initio* results were compared with the estimates obtained using two popular force fields for the study of nucleic acids: the optimized potential for liquid simulation (OPLS) force field reported by Jorgensen and co-workers,[13] and the AMBER-95 force field developed by Kollman's group.[14] The force field which best reproduced the QM values was subsequently used to study cooperative effects by examination of the magnitude of 1−3 stacking interactions (Figure 1).

Ab initio calculations were performed with the Gaussian 92-DFT computer program.[15] Modeling and building of structures was carried out with the Insight-II computer program.[7] Calculations were performed on Cray YMP and IBM-SP2 computers of the Centre de Supercomputació de Catalunya (CESCA), as well as with HP and SGI workstations in our laboratory.

Results and Discussion

As noted in Methods, we were concerned on the preciseness of the MP2/6-31G(d) calculations, since several authors [5e−h,11] have suggested that the 6-31G(d) basis set is too small to represent properly dispersion effects. These authors have suggested the use of a diffuse 6-31G(d) basis set, where the exponents of the d-type orbitals are reduced from their standard optimum values.[5e−h,11−12] In this work we decided to use the standard 6-31G(d) basis set, because it is more popular, leads to smaller BSSE errors, and is more extensively tested. It is difficult to determine the ability of the 6-31G(d) and the diffuse 6-31G(d) basis sets to reproduce dispersion effect, since the reference calculation should be performed with a very large basis set, which is not affordable for the present systems.[11] However, a rough estimate of the errors in our calculations can be obtained by comparison of 6-31G(d) data with diffuse 6-31G(d) results, as well as with data derived from a much larger basis set. For this purpose, we computed the HF and MP2 estimates of intrastrand stacking energies (C−C and T−C dimers) (C = cytosine and T = thymine) using the 6-31G(d), the diffuse 6-31G(d), and the large 6-311+G(d) basis sets (see Table 1).

TABLE 1: SCF and MP2 Interaction Energies (kcal/mol) Using the 6-31G(d), the Diffuse 6-31G(d), and the 6-311+G(d) Basis Set for C−C and T−C dimers in B- and A-Type Stackings

	B-DNA		A-DNA	
basis	SCF	MP2	SCF	MP2
C−C				
6-31G(d)	8.83	1.46	7.30	0.79
d-6-31G(d)	8.51	−1.90	7.40	−1.70
6-311+G(d)	8.46	−0.83	7.21	−0.88
T−C				
6-31G(d)	4.41	−1.78	5.40	−0.99
d-6-31G(d)	4.30	−4.50	5.28	−3.41
6-311+G(d)[a]	4.24	−3.20	5.37	−2.08

[a] Due to computational limitations, the 6-31G basis was used for hydrogen atoms in T−C complexes.

Table 1 shows that the increase in the size of the basis set, or the modification of d-exponent (from 0.8 to 0.25), does not lead to an significant change in the HF values, as previously found by comparison of HF calculations with the basis 6-31G-(d) and 6-311+G(d,p) (data not shown). On the contrary, the modification of the basis set has a notable impact in the determination of dispersion effects, as noted in the values of the MP2 interaction energies. The results demonstrate the importance of the diffuse polarization functions to cover the intermolecular correlation interaction energy. The larger 6-311+G(d) basis is not diffuse enough and still underestimates the stabilization with respect to the diffuse 6-31G(d) basis by about 1−1.5 kcal/mol. Let us note that even the diffuse 6-31G-(d) basis set does not fully cover the intermolecular electron correlation contribution.[5e,f] Fortunately, recent studies[16] indicate that the contribution to the dispersion energy from higher order correlation contributions is repulsive. This would make present estimates of dispersion energy closer to reality. Therefore, even caution is needed from a quantitative point of view, present results are valuable for qualitative discussion of sequence-stacking preferences.

Intrastrand Stacking

The SCF and MP2/6-31G(d) interaction energies for B- and A-DNA-type base-stacked complexes are reported in Table 2. With the exception of the stacking between two cytosines (C−C), the MP2/6-31G(d) interaction energy is negative for all B-DNA complexes, indicating that these interactions are energetically favored in an anhydrous environment. Comparison of SCF and MP2 results reveals the crucial role of correlation effects in the stability of base stacking. In fact, SCF results indicate that formation of a stacked complex is unfavorable, mainly resulting from electrostatic interactions, which are the leading components in SCF energies.[5c] Inclusion of correlation effects stabilizes base stacking by an average of 7.6 kcal/mol, thus inverting the sign of SCF stacking energies, which agree

TABLE 2: SCF and MP2 Interaction Energies (kcal/mol) Using the 6-31G(d) Basis Set for All 16 Possible Pairs of Stacked Bases in Standard B- and A-DNA Conformations

sequence (5′→3′)	B-DNA		A-DNA	
	SCF	MP2	SCF	MP2
A−A	4.74	−3.34	8.07	−1.39
A−G	3.52	−4.10	6.26	−2.34
G−A	2.98	−5.67	5.82	−3.72
G−G	8.19	−0.67	12.00	1.82
T−T	7.38	−0.90	18.28	7.05
C−T	6.55	−1.26	15.91	6.13
T−C	4.41	−1.78	5.40	−0.99
C−C	8.83	1.46	7.30	0.79
A−T	7.58	−2.31	18.92	4.98
G−T	9.80	−1.19	20.50	5.51
A−C	7.09	−1.68	13.21	2.69
G−C	2.99	−5.86	6.03	−4.37
T−A	2.41	−3.37	1.23	−2.89
T−G	2.41	−2.86	3.08	−1.36
C−A	4.35	−1.48	3.97	−1.04
C−G	0.25	−3.39	1.90	−2.01

with previous high-level calculations.[5f] It is worth noting that the SCF term, despite its repulsive character, is important for determination of the relative stability of stacked complexes, as noted by inspection of Table 2. All these findings agree well with most previous theoretical results obtained from low-, and high-level *ab initio* and empirical calculations.[4,5] Our *ab initio* data supports the importance of dispersion forces in aromatic stacking in opposition with recent results by Newcomb and Gellman.[17]

Purine−purine interactions are in general most favored (around −3.4 kcal/mol, on average), followed by purine−pyrimidine and pyrimidine−purine (−2.8 kcal/mol), and finally by pyrimidine−pyrimidine (−0.6 kcal/mol). This ordering agrees with experimental information for stacking in small oligonucleotides.[18] In contrast with recent theoretical values determined from classical simulations,[4] present results do not show any systematic difference between purine−pyrimidine and pyrimidine−purine stackings. This is in agreement with previous *ab initio* studies,[5] and with UV- and circular dichroism (CD)-based experimental results,[18] which show no clear energy difference between these inverted base stackings. The lower stability of the stacking between pyrimidines arises mainly from unfavorable electrostatics, which yield large positive SCF interaction energies. The stability of purine−purine stackings stems from a moderate electrostatic destabilization and a favorable dispersion interactions.[19] Dispersion also determines the general stability of purine−purine complexes over pyrimidine−purine ones, these latter being on average less stable by around 1 kcal/mol.

The average stacking energies for complexes with either adenine or guanine are −3.1 and −3.3 kcal/mol, respectively and are −2.0 and −1.8 kcal/mol for complexes containing cytosine and thymine. This shows the greater stacking preference of purines over pyrimidines and the small difference in stacking preferences between different purines or between different pyrimidines. The larger size of purines, which increases the magnitude of dispersion interactions, is likely responsible for this difference. Figure 2A illustrates the qualitative relationship between dispersion contribution and "overlap" area, which was estimated from the difference between the solvent-accessible surface of the stacked complex and that of the isolated bases. However, it is also clear that there exist quantitative discrepancies for some cases. No correlation is observed between the stacking interaction and the "overlap" area (see Figure 2B). This reveals the complex nature

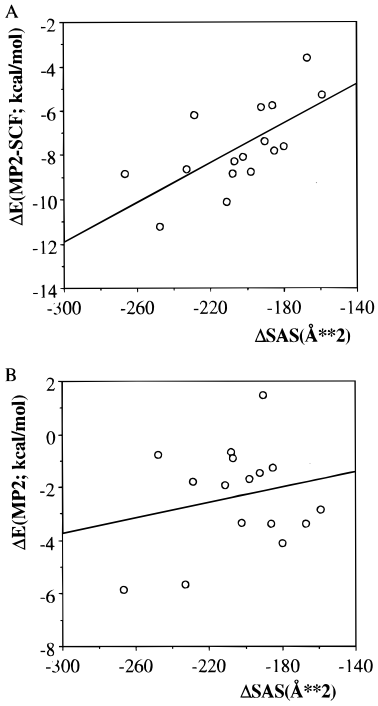

Figure 2. Representation of the dispersion contribution (estimated from the difference between MP2/6-31G(d) and HF/6-31G(d) energies) and the total MP2/6-31G(d) stacking energy *versus* the overlap surface between the bases in the stacking complexes.

of stacking interactions, which cannot be accurately described in terms of simple empirical parameters related to "overlap" areas.

The most stable complexes, G−A and G−C (G = guanine and A = adenine), are moderately destabilized at the SCF level and have a large, favorable dispersion interaction (1 kcal/mol larger than the average dispersion term). This latter contribution reflects the large overlap (37 and 72 Å^2 for G−A and G−C) between the bases, which is greater than the average value for all other purine−purine and purine−pyrimidine stackings. These findings agree with previous *ab initio* results by Aida and co-workers,[5c] and with thermal studies of dinucleotide stacking.[18p] There is also qualitative agreement with recent theoretical results obtained from classical simulations,[4k] although for G−C the interaction energy (nearly 20 kcal/mol) seems too negative according to present *ab initio* results or the different experimental estimates.[18]

In addition to destabilizing interaction for C−C stacking, complexes G−G, T−T, and G−T exhibit poor stability, mainly because of unfavorable electrostatic interactions. Our results agree qualitatively with the values reported by Aida and Nagata[5a−d] from lower level calculations, who also reported a positive interaction energy for C−C stacking in B-DNA. Similar qualitative trends were found in recent classical calculations[4k] of base−base stacking. Experimental data is unclear, stemming in uncertainties in the real conformation of the complexes, and because of the large range of energy values derived from different techniques.[18] Furthermore, most experimental data were obtained in aqueous solution, which might significantly influence the strength of stacking interactions.[4h,l−n] However, despite the difficulties in making a rigorous comparison between QM data on standard DNA base stacking and experimental data for the base stacking of small nucleotides in solution, a general qualitative agreement is found. Thus, the experimental data[18] reveal that both C−C and G−G are among the least stable base-stacked complexes, which agrees with present results. The theoretical finding of the low stability for

T−T and G−T complexes is more difficult to verify because of the lack of experimental information on thymine stacking.

An interesting finding is the great sensitivity of stacking interactions not only to base composition, but also to base sequence. This is clearly illustrated by comparison of the interaction energies for the complexes 5′-X-Y-3′ and 5′-Y-X-3′. The stackings 5′-Y-T-3′ (T: thymine) are less stable than 5′-T-Y-3′, which probably stems from unfavorable steric contacts between the methyl group of thymine in 3′ and the aromatic ring of the nucleobase at 5′. An interesting finding is the great stability (around 2.5 kcal/mol) of 5′-G-C-3′ relative to 5′-C-G-3′. This mainly results from dispersion, in agreement with the greater surface overlap (more than 100 Å^2) for G-C relative to C-G.

The SCF results indicate that A-DNA-type stackings are clearly destabilized relative to B-DNA structures (around 4 kcal/mol less favored than the corresponding B-DNA complexes, in average). Again, dispersion has a stabilizing effect (around 1 kcal/mol greater than for B-DNA stackings), although in seven cases the complex is energetically disfavored. Consequently, near all A-DNA-type complexes are less stable than B-DNA ones. This is because of unfavorable steric contacts owing to the reduced helical rise in A-DNA conformations derived from fiber diffraction. The nature of these steric contacts is mainly electrostatic (nuclei-nuclei repulsion), as revealed by the SCF values.

The stacking interactions in A-DNA are, on average, −1.4, −1.8, 2.2, and 3.2 kcal/mol for purine−purine, pyrimidine−purine, purine−pyrimidine and pyrimidine−pyrimidine complexes. Therefore, in contrast to the results found for B-DNA-type complexes, the most stable interaction in A-DNA is for pyrimidine−purine stacking, whose stability is similar to that of the corresponding B-DNA complex, and the worst is for purine−pyrimidine stacking, which is greatly destabilized at the SCF level in A-DNA (nearly 8 kcal/mol less stable than in B-DNA).

Comparison of the results for B- and A-DNA complexes suggests classification into four categories: (i) C−C, which prefers A-DNA conformation; (ii) complexes T−C, T−A, and C−A, with slight preference (less than 1 kcal/mol) for B-DNA conformation; (iii) complexes A−A, A−G, G−A, G−G, G−C, T−G, and C−G, with clear preference (near 2 kcal/mol) for B-DNA conformation; and (iv) complexes T−T, C−T, A−T, G−T, and A−C, extremely unstable in A-DNA. No definitive rules emerge to explain these preferences, although it is clear that a thymine in 3′ destabilizes A-DNA stacking, while the stabilities between B- and A-DNA stackings become more comparable for 5′ thymine or when there is a cytosine in either 5′ or 3′.

Classical simulations are frequently used to study nucleic acids. The last generation of force fields successfully reproduces both conformational flexibility of nucleosides and interactions between hydrogen-bonded base pairs. However, less is known about the ability of these force fields to represent stacking interactions. Figure 3 shows the comparison of OPLS[13] and AMBER[14] interaction energies relative to the MP2/6-31G(d) results. Both force fields reasonably reproduce the sequence dependence of stacking interactions. AMBER correlates better with our *ab initio* data. However, they fail in quantitative description of the interaction energies, which are largely overestimated (especially in AMBER calculations) with respect to both the *ab initio* values and the available experimental information.[18] A similar trend is noted by inspection of recent results derived from the CHARMM force field.[4k,20] Inspection of the components of the classical interaction energy (data not shown) reveals that the van der Waals contribution reproduces quite well the dispersion term determined at the 6-31G(d) level. In contrast, the electrostatic term is positive, but generally too small.[21]

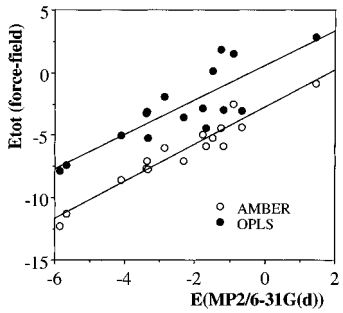

Figure 3. Comparison of AMBER-95 and OPLS interaction energies *versus* the MP2/6-31G(d) results for the base-stacked complexes.

TABLE 3: MP2/6-31G(d) (Plain) and OPLS (Italics) Interaction Energies (kcal/mol) for All 10 Possible Cross-Interactions in Stacked Dimers in Standard B- and A-DNA Conformations

sequence (5′−X_1−Y_1−3′) (3′−Y_2−X_2−5′)	B-DNA			A-DNA		
	5′−5′	3′−3′	total	5′−5′	3′−3′	total
A−A	−0.42	−1.19	−1.61	−0.46	−2.10	−2.56
T−T	*−0.91*	*−2.05*	*−2.96*	*−0.48*	*−3.43*	*−3.91*
A−G	0.27	0.11	0.38	0.31	1.24	1.55
T−C	*−0.36*	*−1.88*	*−2.24*	*0.01*	*−1.39*	*−1.38*
G−A	0.09	0.68	0.77	0.53	0.32	0.85
C−T	*0.91*	*−0.03*	*0.88*	*0.41*	*−0.94*	*−0.53*
G−G	−3.16	−2.21	−5.37	−2.45	−2.41	−4.86
C−C	*−3.63*	*−3.55*	*−7.18*	*−2.63*	*−4.32*	*−6.95*
A−T	−0.09	0.42	0.33	−0.39	0.39	0.00
T−A	*−1.39*	*0.04*	*−1.35*	*−0.92*	*−0.59*	*−1.48*
G−T	−2.23	−0.30	−2.53	−2.17	−1.78	−3.95
C−A	*−2.54*	*−0.16*	*−2.70*	*−1.90*	*−1.76*	*−3.66*
G−C	1.60	2.56	4.16	2.70	1.36	4.06
C−G	*0.73*	*3.44*	*4.17*	*2.51*	*2.22*	*4.73*
T−A	0.29	−0.79	−0.50	0.35	2.20	2.55
A−T	*0.21*	*−2.87*	*−2.66*	*0.36*	*−1.33*	*−0.97*
T−G	−0.34	−2.82	−3.16	−0.34	−1.64	−1.98
A−C	*−0.50*	*−5.97*	*−6.47*	*−0.29*	*−3.75*	*−4.04*
C−G	0.99	−2.35	−1.36	−0.84	0.29	−0.55
G−C	*1.17*	*−4.97*	*−3.80*	*1.10*	*−0.48*	*0.62*

Interstrand Stacking

Inspection of standard nucleic acid models shows that significant overlap between bases in complementary strands may occur depending on the sequence. Therefore, interstrand 5′↔5′ and 3′↔3′ interactions (Figure 1) may be important [5d,e] for stabilization of stacking between base-pair dimers. Table 3 reports *ab initio* MP2/6-31G(d) and OPLS estimates of interstrand interactions in both B- and A-DNA conformations. Results show that these interactions are significant for the stability of stacked base pairs in B-DNA conformations, even before consideration of possible distortion in the planarity of the bases.[22] This finding suggests that this term cannot be neglected, as was done in several previous works. In contrast with the intrastrand interactions, the dispersion component is less important (data not shown), so that interstrand interactions are largely modulated by electrostatic contributions.

The interstrand stabilization is especially relevant for G:C−G:C, G:C−T:A, and T:A−G:C, which have interaction energies larger than 2 kcal/mol. In particular, G:C−G:C is stabilized by −5.4 kcal/mol, while the corresponding intrastrand stacking energy derived from addition of the MP2/6-31G(d) values for G−G and C−C reported in Table 2 is 0.9 kcal/mol. In most

TABLE 4: OPLS Interaction Energies (kcal/mol) Between Bases Separated by One Step in the Polynucleotide Chain in Standard A- and B-DNA Conformations

sequence (5′→3′)	B-DNA	A-DNA
A−X−A	−0.03	−0.03
A−X−G	0.05	0.36
G−X−A	−0.23	−0.21
G−X−G	0.22	1.11
T−X−T	−0.10	−0.24
C−X−T	−0.18	−0.39
T−X−C	0.13	−0.70
C−X−C	0.20	−0.41
A−X−T	−0.09	−0.45
G−X−T	0.00	0.05
A−X−C	−0.02	−0.62
G−X−C	−0.56	−1.10
T−X−A	−0.04	−0.11
T−X−G	−0.40	−0.01
C−X−A	0.20	−0.01
C−X−G	0.14	−0.19

cases electrostatic effects (mainly dipole-dipole interactions) are responsible for the stabilization energy. Interstrand interactions are destabilizing for A:T−G:C, G:C−A:T, A:T−T:A, and especially for G:C−C:G (around 4.2 kcal/mol), resulting from unfavorable electrostatics.

Table 3 also reports the interstrand interactions for A-DNA conformations. The results have similar trends as for those reported for B-DNA complexes. Again, the interaction energy is mainly determined by the electrostatic component. Furthermore, the best interaction is for G:C−G:C, and the worst for G:C−C:G. Comparison of the results reveals that, on average, interstrand interactions are around 0.4 kcal/mol more favorable for B-DNA type complexes. Only A:T-A:T, and G:C−T:A have better interstrand interactions in A-DNA conformations. The interaction energy is similar for G:C−A:T, A:T−T:A, and G:C−C:G, while for the remaining dimers interstrand interactions favor the B-DNA conformation.

The agreement between *ab initio* and force field derived results for the interstrand interactions is better that that found for the intrastrand interactions. Only the OPLS values are given in Table 3, but both OPLS and AMBER results agree reasonably well, although large errors occur in some cases. Accordingly, it seems prudent that the next generation of force fields be developed to maintain a proper balance between inter- and intrastrand stacking interactions.

1−3 Stacking Interactions

Stacking in DNA fibers are also influenced by cooperative effects, which may be entropic or enthalpic in nature. Entropic effects arise from the reduced configurational space accessible to the bases upon stacking of neighboring bases, while enthalpic contributions are expected to arise from long-range 1−3 interactions, *i.e.*, intrastrand stacking interactions between bases separated by one step (Figure 1). The importance of this interaction, which should be not *a priori* neglected in theoretical or experimental studies of base stacking (as it is usually done), should be assessed. Analysis of standard B- and A-DNA structures reveals that the distance between base pairs separated by one step ranges from 5 to 7 Å. Therefore, dispersion forces are expected to be small, if not negligible, but long-range electrostatic effects may still be important at these distances.

Table 4 contains all the 1−3 base-base energetic results for standard A- and B-DNA models as determined from OPLS calculations, which like most last generation force fields is expected to accurately estimate long-range electrostatic interactions. For B-DNA stackings, 1−3 interactions are small, the largest value being for G−X−C (around −0.5 kcal/mol). These interactions have a small effect, on average, on stability of the B-DNA fiber (the average energy amounts to −0.04 kcal/mol). As expected, 1−3 interactions are larger in A-DNA, reaching values around ±1 kcal/mol. This results from the smaller rise in fiber structure of A-DNA, as reflected in a reduction of around 2 Å in the distance between 1−3 bases with respect to the B-DNA conformation. On average, the 1−3 interactions slightly stabilize (−0.2 kcal/mol) the A-DNA structure. Overall, the energy contribution of 1−3 stacking interactions seems to have a very slight influence on the stability of DNA. This suggests that cooperative effects of DNA folding should arise mainly from entropic contributions.

TABLE 5: Total Interaction Energies (kcal/mol) for All 10 Possible Stacked Dimers in Standard A- and B-DNA Conformation. Values are Derived from Addition of MP2/6-31G(d) Intrastrand Base Stacking and Cross-Interaction Energies Reported in Tables 2 and 3. OPLS Values in Parentheses

sequence (5′-X_1-Y_1-3′) (3′-Y_2-X_2-5′)	B-DNA	A-DNA
A−A T−T	−5.85 (−4.24)	3.10 (5.66)
A−G T−C	−4.98 (−5.36)	5.34 (3.79)
G−A C−T	−6.68 (−7.45)	−3.86 (4.71)
G−G C−C	−4.58 (0.79)	−2.25 (2.61)
A−T T−A	−4.29 (−4.62)	9.96 (9.96)
G−T C−A	−5.40 (−2.87)	4.25 (8.20)
G−C C−G	−7.56 (−11.72)	−4.68 (−8.74)
T−A A−T	−7.24 (−6.74)	−3.23 (−5.78)
T−G A−C	−7.50 (−4.34)	−4.38 (−2.40)
C−G G−C	−8.14 (−6.78)	−4.57 (−4.02)

Dimer Stacking Interactions

Because of the small magnitude of 1−3 interactions, total stacking energy between two base-pair dimers can be estimated from addition of intrastrand (5′↔3′) and interstrand (5′↔5′, 3′↔3′) interactions (Figure 1). The *ab initio* MP2/6-31G(d) energies for B- and A-DNA conformations, which are shown in Table 5, strongly suggest that B-type stacking is more favorable than A-type, despite the lower rise of this latter conformation. Owing to the large energy differences (more than 6 kcal/mol, on average) this finding is not expected to be an artifact of these simulations. In contrast with previous QM studies that reported A-type stacking to be preferred for some base-pair dimers,[5a−d] the preference for B-type stacking extends to all possible base-pair dimers. The discrepancy with previous theoretical results[5d] might arise from the use of different rise and inclination values in models used previously to study A-DNA stacking. Present results which have been obtained for the standard fiber model of A-DNA[1,6,7] suggest that the existence of the A-DNA conformation for some base sequences under partial anhydrous conditions is not due to an intrinsic stabilizing effect related to base stacking, when the A-DNA geometry is that reported in fiber. Actually, our *ab initio* data

TABLE 6: B- and A-DNA Crystal Structures Extracted from the Brookhaven Data Bank. For Sequences Solved by More than One Author, Only One Reference is Included. No Drug DNA or Protein DNA Structures were Considered in the Search

sequence	type	ref
d(GGGATCCC)	A	23c
d(GGTATACC)	A	23a
d(CTCTAGAG)	A	23d
d(GTACGTAC)	A	23e
d(GGGTACCC)	A	23f
d(GTGTACAC)	A	23g
d(GTCTAGAC)	A	23h
d(GGATGGAG)	A	23b
d(CCGTACGTACGG)	A	23i
d(GCGTACGTACGC)	A	23j
d(CGATCGATCG)	B	24d
d(CGATTAATCG)	B	24e
d(CGATATATCG)	B	24f
d(CATGGCCATG)	B	24h
d(CCATTAATGG)	B	24j
d(CGCGAATTCGCG)	B	24a
d(CGCATATATGCG)	B	24b
d(CGTGAATTCACG)	B	24c
d(CGCAAATTTGCG)	B	24g
d(CGTAGATCTACG)	B	24i

suggest that the A-DNA conformation as determined from fiber analysis is in general unstable. It can be suggested based on our results that A-DNA-like stacking stability might be gained by increasing the rise value of A-DNA to values similar to those of B-DNA stacking.

The largest differences between A- and B-type conformations occur for sequences having thymine in 3′, as expected from the results of intrastrand stacking. This difference is especially large for A:T−T:A (stabilized by more than 14 kcal/mol), probably resulting from the bulky methyl group. These results agree with experimental evidence obtained from analysis of crystallographic data, which reveals a strong propensity for sequences with 3′ thymine to adopt B-DNA conformation, particularly for the 5′−A−T−3′ sequence (see Table 6). This effect is not expected to occur in RNA, where the A-type conformation is very important,[1] since thymine is substituted by uracil. The smallest difference (in absolute terms) in stability between A- and B-DNA stackings is found for G:C−G:C. Again, the agreement with X-ray data is good (Table 6; ref 22 and 23), since polyG:polyC sequences are the most commonly found in A-DNA crystal structures. Results in Table 5 reveal that this is not due to a large stability of the A-type stacking, but rather to a low stability of B-DNA stacking, reflecting the poor base−base stacking interaction in both G−G and C−C structures (see above).

Inspection of the tables also shows the nonnegligible influence of interstrand stacking interactions, which in some cases introduce notable differences between the intrastrand contribution and the total stacking energy. This is the case for G:C−T:A, T:A−G:C, G:C−C:G, and G:C−G:C. The two latter cases are especially relevant, since for G:C−G:C interstrand interactions stabilize the stacking by around 5 kcal/mol while they play a destabilizing effect for G:C−C:G.

The best B-DNA stackings are for pyrimidine:purine-purine:pyrimidine (C:G−G:C, T:A−G:C, T:A−A:T) dimers (−7.6 kcal/mol, on average), while both purine:pyrimidine-purine:pyrimidine (A:T−A:T, A:T−G:C, G:C−A:T, G:C−G:C) or purine:pyrimidine-pyrimidine-purine (A:T−T:A, G:C−T:A, G:C−C:G) dimers have almost the same average stabilization energy (around −5.6 kcal/mol). It is worth noting that this ordering is quite different from that inferred only from intrastrand base stacking.

TABLE 7: Intrastrand, Interstrand, and Total Interaction Energies (kcal/mol) Determined at the MP2/6-31G(d) Level for Stacked Dimers in B-DNA Crystal Structures

sequence ($5'$-X_1-Y_1-$3'$) ($3'$-Y_2-X_2-$5'$)	ref	intra	inter	total
G−C C−G	22a	−10.84	4.23	−6.21
C−G G−C	22a	−8.42	0.27	−8.15
A−T T−A	22a	−8.16	1.76	−6.40
T−A A−T	22b	−6.71	0.53	−6.18
G−T C−A	22c	−4.28	−1.02	−5.30
T−G A−C	22g	−5.21	−2.98	−8.19

G:C rich sequences (GC, GG, and CG) have on average better stacking energies (1 kcal/mol for each stacked dimer) than A:T rich sequences (AA, AT, and TA), which agrees with the greater thermal stability of DNA regions with high G:C content. The most stable dimer stacking is C:G−G:C, followed by G:C−C:G. In contrast, A:T−T:A is the least stable followed by G:C−G:C. Comparison with experimental data is difficult, since most experimental studies performed on single-strand oligonucleotides, and thus provide information only on interstrand base stacking. An analogous situation occurs for previous theoretical studies, which in most cases consider only intrastrand stacking interactions.

Dependence of Stacking on Local Helix Conformation

Previous results are based on standard B- and A-DNA conformations as determined from fiber diffraction analysis.[6] These average helical parameters give only a rough representation of DNA structure, since sequence-specific changes are neglected. It is beyond the scope of this study to analyze how stacking interactions may be influenced by changes of local helicity and interbase parameters. Nevertheless, it is interesting to estimate the relevance of changes in local conformation on the magnitude of stacking interactions in the physiologically-relevant conformation, *i.e.*, the B-DNA. For this reason, the stacking energy for selected base-pair dimers was computed at the MP2/6-31G(d) level using high-resolution crystal geometries.[24] The test dimers are G:C−C:G,[24a] C:G−G:C,[24a] A:T−T:A,[24a] T:A−A:T,[24b] G:C−T:A,[24c] and T:A−G:C.[24g]

Results in Table 7 reveals a general similarity with simulation results obtained from fiber B-DNA conformation. The RMS deviation between the total energy for crystal stacking (Table 7), and the total energy for fiber stacking (Table 5) is 1.1 kcal/mol, indicating that the changes between fiber and crystal geometries do not induce remarkable alterations in the strength of stacking interactions. Furthermore, the average absolute difference between results obtained for crystal and fiber conformations is only 0.1 kcal/mol, which demonstrates that sequence-dependent structural changes in DNA do not necessarily increase the strength of dimer stacking. The only apparent exception to this rule is the A:T−T:A sequence, whose stacking stabilization increases more than 2 kcal/mol when changing from fiber to crystal geometry, due probably to the strong perturbation in the DNA sequence induced by the A:T−T:A sequences. It is suggested that large variations might be found when comparing fiber results for A-DNA with crystal data of DNA sequences which resembles the canonical A-DNA conformation, since in this case rise difference between fiber and crystal is bigger.

Conclusions

Knowledge of the factors that modulate stacking of nucleic acid bases is essential to understand DNA structure and function. Unfortunately, these factors are very difficult to determine experimentally and theoretically. Thus, experimental techniques have difficulties in distinguishing between the pure stacking contribution, and other interactions arising from hydration, hydrogen bonds, local conformational changes, and other effects. Furthermore, small dinucleotide models which are often used for experimental study of stackings might have stacking patterns very different than those of polymeric DNA. On the other hand, a rigorous theoretical description of stacking is difficult because of the need to perform high-level QM calculations for very large molecular systems. Furthermore, the non-additivity of interactions made in the present study can be nonnegligible for quantitative discussion of the different energy contributions. Finally, theoretical calculations cannot properly represent the very complex molecular environment of DNA, which makes it difficult to extrapolate results derived from gas phase calculations to physiological conditions.

In spite of the caution needed to interpret simulation results, a few general trends are clear from these *ab initio* calculations, which may be valuable in understanding the DNA structure and dynamics, and for the design of DNA-intercalating compounds with potential therapeutic applications. Regarding the energy contributions that influence stacking between bases, the results show that, in addition to the role of electrostatics in determining stacking sequence specificity, the stacking is modulated by the relevant stabilizing contribution of dispersion. Indeed, the results reveal that stacking stabilization between base-pair dimers is determined by two main contributions: the intrastrand base stacking and the interstrand cross-interactions, the contribution of this latter term often being nonnegligible. The energy contributions arising from one-step separated bases (1−3 interactions) are, in general, very small, suggesting an entropic nature for DNA folding cooperativity. It is also found that base-pair dimer stacking in the B-DNA conformation is much more stable than that for A-DNA-type complexes for fiber geometries. This result holds true for all possible dimer stacking sequences. The largest differences between B- and A-DNA-stacking complexes is observed for sequences containing adenine and thymine, whereas the smallest differences are found for sequences having guanine and cytosine. This agrees with the experimentally observed crystal structural preference of B-DNA for regions with high A−T content and of A-DNA for regions with high G−C content. Comparison of stacking energies in fiber and monocrystal structures suggests that small changes in stability of B-DNA sequences are expected to arise from local sequence-dependent structural changes.

Acknowledgment. We are indebted to Prof. J. Sánchez and Prof. G. C. Shields for valuable comments. We also thank the reviewer for very helpful comments. This work has been supported by the Centre de Supercomputació de Catalunya (CESCA, Molecular Recognition Project), and by the Dirección Gerneral de Investigación Científica y Técnica (Grants PB93-0779 and PB94-0940).

References and Notes

(1) (a) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1988; Chapter 6. (b) Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chemistry and Biology*; IRL Press: Oxford, 1990; Chapter 1.

(2) (a) Neidle, S., Waring, M. J., Ed. *Molecular Aspects of Anti-Cancer Drug Action*; Macmillan Press: London, 1983; Vol. 3. (b) Wang A. H. In *Nucleic Acids and Molecular Biology*; Eckstein, F., Lilley, D. M. J., Eds.; Springer-Verlag: Berlin, 1987; Vol. 1, pp 53−69. (c) Neidle, S.; Pearl, L. H.; Skelly, J. V. *Biochem. J.* **1987**, *243*, 1. (d) Cheng, C. C. In *Progress in Medicinal Chemistry*; Ellis, G. P., West, G. B., Eds.; Elsevier: Amsterdam, 1988; Vol. 25, Chapter 2. (e) Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chemistry and Biology*; IRL Press: Oxford, 1990; Chapter 7.

(3) (a) Calladine, C. R. *J. Mol. Biol.* **1982**, *161*, 343. (b) Calladine, C. R.; Drew, H. R. *J. Mol. Biol.* **1984**, *178*, 773. (c) Calladine, C. R.; Drew, H. R. *Understanding DNA*; Academic Press, London, 1992.

(4) (a) Ornstein R. L.; Rein, R.; Breen, D. L.; Macelroy, R. D. *Biopolymers* **1978**, *17*, 2341 and references therein. (b) Rein, R. In *Intermolecular Intersactions: From Diatomics to Biopolymers;* Pullman, B., Ed. John Willey & Sons: Chichester, 1978; and references therein. (c) Langlet, J.; Claverie, P.; Caron, F.; Boeuve, J. C. *Int. J. Quant. Chem.* **1981**, *19*, 299. (d) Haran, T. E.; Berkowich-Yellin, Z.; Shakked, Z. *J. Biomol. Struct. Dyn.* **1984**, *2*, 397. (e) Srinivasan, A. R.; Torres, R.; Clark, W.; Olson, W. K. *J. Biomol. Struct. Dyn.* **1987**, *5*, 459. (f) Sarai, A.; Mazur, J.; Nussinov, R.; Jernigan, R. L. *Biochemistry* **1988**, *27*, 8498. (g) Sarai, A.; Mazur, J.; Nussinov, R.; Jernigan, R. L. *Biochemistry* **1989**, *28*, 7842. (h) Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 503. (i) Friedman, R. A.; Honig, B. *Biopolymers* **1992**, *32*, 145. (j) Hunter, C. A. *J. Mol. Biol.* **1993**, *230*, 1025. (j) Hunter, C. A. *Chem. Soc. Rev.* **1994**, 101. (k) Norberg, J.; Nilsson L. *J. Am. Chem. Soc.* **1995**, *117*, 10832. (l) Norberg, J.; Nilsson L. *J. Phys. Chem.* **1995**, *99*, 13056. (m) Friedman, R. A.; Honig, B. *Biophys. J.* **1995**, *69*, 1528. (n) Norberg, J.; Nilsson, L. *Biophys. J.* **1995**, *69*, 2277.

(5) (a) Aida, M.; Nagata, C. *Chem. Phys. Lett.* **1982**, *86*, 44, and references therein. (b) Aida, M.; Nagata, C. *Int. J. Quant. Chem.* **1986**, *29*, 1253. (c) Nagata, C.; Aida, M. *J. Mol. Struct. (THEOCHEM)* **1988**, *179*, 451. (d) Aida, M. *J. Theor. Biol.* **1988**, *130*, 327. (e) Hobza, P.; Sponer, J.; Polásek, M. *J. Am. Chem. Soc.* **1995**, *117*, 792. (f) Sponer, J.; Leszcyznski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590. (g) Sponer, J.; Leszcyznski, J.; Vetter, V.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *13*, 695. (h) Sponer, J.; Leszcyznski, J.; Hobza, P. *J. Comput. Chem.* **1996**, *17*, 841.

(6) Arnott, S.; Hukins, D. W. L. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1504.

(7) Insight II Computer Program, Biosym Technology, San Diego CA. 1991. The preciseness of the fiber models obtained by Insight II was verified by comparison with those generated by NUCGEN-95 program developed by Kollman's research group at UCSF.

(8) (a) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1978**, *28*, 213. (b) Möller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(9) Boys, S. F., Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(10) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

(11) (a) Hobza, P.; Zahradnik, R. *Chem. Rev.* **1988**, *88*, 871. (b) Tsuzuki,S.; Uchimaru, T.; Tanabe, K. *J. Mol. Struct. (THEOCHEM)* **1994**, *307*, 107.

(12) The diffuse 6-31G(d) basis set is defined by using the standard basis set substituting the standard exponent of the d functions by 0.25.

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

(14) 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.* **1995**, *117*, 5179.

(15) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schelgel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkey, J. S.; Gonzalen, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P., Pople, J. A. *Gaussian 92* Gaussian Inc. Pittsburgh, PA, 1992.

(16) (a) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790. (b) Hobza, P.; Selzle, H. L.; Schlag, E. W. *Chem. Rev.* **1994**, *94*, 11767.

(17) Newcomb, L. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4993.

(18) (a) Ts'o, P. O. P.; Melvin, I. S.; Olson, A. C. *J. Am. Chem. Soc.* **1963**, *85*, 1289. (b) Warshaw, M. M.; Tinoco, I. *J. Mol. Biol.* **1966**, *20*, 29. (c) Leng, M.; Felsenfeld, G. *J. Mol. Biol.* **1966**, *15*, 455. (d) Applequist, J.; Damle, V. *J. Am. Chem. Soc.* **1966**, *88*, 3895. (e) Broom, A. D.; Schweizer, M. P.; Ts'o, P. O. P.; *J. Am. Chem. Soc.* **1967**, *89*, 3612. (f) Nakano, N. I.; Igarashi, S. J. *Biochemistry* **1970**, *9*, 577. (g) Powell, J. T.; Richards, E. G.; Gratzer, W. B., *Biopolymers* **1972**, *11*, 235. (h) Guschlbauer, W.; Frick, I.; Holy, A. *Eur. J. Biochem.* **1972**, *31*, 1. (i) Mitchell, P. R.; Sigel, H. *Eur. J. Biochem.* **1978**, *88*, 149. (j) Ogasawara, N.; Inoue, Y., *J. Am. Chem. Soc.* **1976**, *98*, 7054. (k) Watts, M. T.; Tinoco, I. *Biochemistry* **1976**, *17*, 2455. (l) Lee, C.-H.; Ezra, F. S.; Kondo, N. S.; Sarma, R. H.; Danyluk, S. S. *Biochemistry* **1977**, *16*, 3627. (m) Ezra, F. S.; Lee, C.-H.; Kondo, N. S.; Danyluk, S. S.; Sarma, R. H. *Biochemistry* **1977**, *16*, 1977. (n) Lee, C.-H.; Tinoco, I. *Biochemistry* **1977**, *16*, 5403. (o) Suurkuusk, J.; Alvarez, J.; Freire, E.; Biltonen, R. *Biopolymers* **1977**, *16*, 2641. (p) Frechet, D.; Ehrlich, R.; Remy, P. *Nucl. Acid Research* **1979**, *7*, 1981. (q) Sowers, L. C.; Shaw, B. R.; Sedwick, W. D. *Biochem. Biophys. Rese. Commun.* **1987**, *148*, 790. (r) Morcillo, J.; Gallego, E.; Peral, F. *J. Mol. Struct.* (THEOCHEM) **1987**, *157*, 353. (s) Iza, N.; Montero, J. L.; Morcillo, J. *J. Mol. Struct.* (THEOCHEM) **1988**, *175*, 19.

(19) We have assumed in this paper that the MP2−HF difference is due only to dispersion interactions. This is an obvious simplification, but it is useful for discussion.

(20) (a) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S. Karplus, M. *J. Comput. Chem.*, **1983**, *4*, 187. (b) MacKerell, A. D.; Wiórkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem.Soc*. **1995**, *117*, 11946.

(21) Overestimation of stacking stability in classical calculations might be smaller than that deduced from inspection of Figure 3. This is due to the tendency of MP2/6-31G(d) calculations to underestimate dispersion interactions.

(22) Coll, M.; Frederick, C. A.; Wang, A. H. J.; Rich, A. J. *Proc. Natl. Acad. Sci. U.S.A.*, **1987**, *84*, 8385.

(23) (a) Shakked, Z.; Rabinovich D.; Kennard, O.; Cruse, W. B.; Salisbury, S. A.; Viswamitra, M. A. *J. Mol. Biol.* **1983**, *166*, 183. (b) McCall, M.; Brown, T.; Hunter, W. N.; Kennard, O. *Nature* **1986**, *322*, 661. (c) Lauble, H.; Frank, R.; Blocker, H.; Heinemann, U. *Nucleic Acid Research* **1988**, *16*, 7799. (d) Hunter, W. N.; D'Estaintot, B. L.; Kennard, O. *Biochemistry* **1989**, *28*, 2444. (e) Takusagawa, F. *J. Biomol. Struct. Dynam.* **1990**, *7*, 795. (f) Eisenstein, M.; Frolow, F.; Shakked, Z.; Rabinovich, D. *Nucleic Acids Research* **1990**, *18*, 3185. (g) Jain, S.; Zon, G.; Sundaralingam, M. *Biochemistry* **1991**, *30*, 3576. (h) Cervi, A.; Lagnlois D'Estaintot, B.; Hunter, W. N. *Acta Crystallogr., Sect. B.: Struct. Sci.* **1992**, *48*, 714. (i) Bingman, C. A., Zon, G., Sundaralingam, M. *J. Mol. Biol.* **1992**, *227*, 738. (j) Bingman, C.; Jain, S.; Zong, G.; Sundaralingam, M. *Nucleic Acids Research* **1992**, *20*, 6637.

(24) (a) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179. (b) Yoon, C.; Prive G. G.; Goodsell, D. S.; Dickerson R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6332. (c) Narayama, N.; Genel, S. L.; Russu, I. M.; Berman, H. M. *Biochemistry* **1991**, *30*, 4449. (d) Grzeskowiak, K.; Yanagi, K.; Prive, G. G.; Dickerson R. E.; *J. Biol. Chem.* **1991**, *266*, 8861. (e) Quintana, J. R.; Grzeskowiak, K.; Yanagi, K.; Dickerson, R. E. *J. Mol. Biol.* **1992**, *225*, 379. (f) Yuan, H.; Quintana, J.; Dickerson, R. E. *Biochemistry* **1992**, *31*, 8009. (g) Edwards, K. J.; Brown, D. G.; Spink, N.; Skelly, J. V.; Neidle, S. *J. Mol. Biol.* **1992**, *226*, 1161. (h) Goodsell, D. S.; Kopka, M. L.; Cascio, D.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 2930. (i) Leonard, G. A.; Hunter, W. N. *J. Mol. Biol.* **1993**, *234*, 198. (j) Goodsell, D. S.; Kaczor-Grzeskowiak, M.; Dickerson, R. E. *J. Mol. Biol.* **1994**, *239*, 79.