Cyclic Adenine-, Cytosine-, Thymine-, and Mixed Guanine−Cytosine-Base Tetrads in Nucleic Acids Viewed from a Quantum-Chemical and Force Field Perspective

Michael Meyer,*,† Christoph Schneider,‡ Maria Brandl,§ and Jürgen Sühnel*,§

Revotar Biopharmaceuticals AG, Neuendorfstrasse 24b, D-16761 Hennigsdorf, Germany, Accelrys Incorporated, Inselkammerstrasse 1, D-82008 Unterhaching, Germany, and Institut für Molekulare Biotechnologie, Beutenbergstrasse 11, D-07745 Jena, Germany

Received: July 5, 2001; In Final Form: October 2, 2001

We have carried out a systematic study of hydrogen-bonded cyclic A, C, T, and mixed GCGC tetrads resembling conformations occurring in experimental tetraplex structures, using the B3LYP hybrid density functional (DFT) method and the MMFF and AMBER force fields to determine tetrad structures and interaction energies. The results are compared to G and U tetrads analyzed previously, thereby presenting a comprehensive overview of all cyclic tetrads formed from one base type only, in addition to the GCGC data. The DFT calculations indicate that the C tetrad is planar and the GCGC tetrad is nearly planar and correspond to local minima at C_{4h} and C_i symmetry, respectively. For A tetrads with N6−H6···N7 and with N6−H6···N1 H-bonds and for T tetrads, nonplanar structures are more stable than the planar ones, and among the nonplanar structures, S_4-symmetric conformations are more stable than C_4 structures. Minima confirmed by frequency calculations are found for the planar C tetrad, the C_i-symmetric GCGC tetrad, the C_4-symmetric structure of the A tetrad with N6−H6···N7 H-bonds, and the S_4-symmetric T tetrad, in addition to the already known local minima of the S_4-symmetric structure of the G tetrad and the planar U tetrad with C−H···O hydrogen bonds. The interaction energies (ΔE^T) corrected for deformation of the bases in the tetrads range from −69.47 kcal/mol for the GCGC tetrad with three pairs of strong Watson−Crick type hydrogen bonds to −11.09 kcal/mol in the planar A tetrad having only a single N6−H62···N1 interbase hydrogen bond. With more than 18%, the largest cooperativity contribution to ΔE is found for the C and G tetrads. The interaction energies derived from the MMFF and AMBER force field are similar to the DFT data for those tetrads having high interaction energies, whereas the relative deviations are much larger for weakly H-bonded tetrads.

Introduction

It has been known for a long time that DNA and RNA can form helices consisting of more than two strands.[1] Currently, however, the interest in multistranded nucleic acid structures is growing.[2−4] For several reasons, tetraplex structures are of particular importance. They occur, for example, in the telomeres at the ends of linear chromosomes. Telomeres consist of tandem repeats of guanine (G)-rich sequences associated with various proteins including telomerase, an RNA-dependent DNA polymerase. Proposed telomere functions are maintenance of the structural integrity of the genome and ensurance of complete replication at the chromosome termini. It has been shown that the G-rich sequences form G tetrads and that G tetraplex structures inhibit telomerase activity.[5] Therefore, G tetraplex-based telomerase inhibitors may be important for cancer therapy and aging research. Similar sequence motifs do also occur in regulatory regions of oncogenes. Recently, it has been proposed that targeted control elements of the *c-myc* oncogene adopt an intrastrand foldback G tetraplex.[6] Finally, a tetraplex structure has also been found for a centromere sequence.[7]

Starting with the first G tetraplex structures of telomeric DNA[8,9] published in 1992, our knowledge on base tetrads has been greatly improved by a variety of further three-dimensional (3D) structures reported in recent years.[3] The strands can be oriented in a parallel or antiparallel manner, and in many of these structures, NH_4^+ and mono- or divalent metal ions located either in the central cavity of the base tetrads or between the base planes play an important role for tetraplex structure and function.

In addition to these experimental investigations, theoretical studies of G tetraplexes and G tetrads have been carried out by means of molecular dynamics (MD) simulations[10−12] and quantum-chemical calculations,[13−17] respectively.

Much less is known about other bases than G in tetraplexes. Even though the telomeric sequences are G-rich, they also contain other bases and therefore tetrads with non-G bases are principally possible. Indeed, very recently in 3D DNA structures containing various telomere repeat motifs, non-G base tetrads have been observed (Figure 1). An A tetrad has been found in the truncated human telomeric DNA sequence d-AGGGT,[18] and a C tetrad has been observed in d-TGGGCGGT containing a repeat sequence from the SV40 viral genome.[19] A T tetrad has been found in a parallel-stranded tetraplex formed by the DNA sequence d-TGGTGGC containing two repeats of *Saccharomyces cerevisiae* telomere DNA.[20] GCGC tetrads have been found in DNA tetraplexes with the repeat sequence d(GGGC-

* To whom correspondence should be addressed. E-mail: m.meyer@revotar-ag.de. Fax: +49 33022025030. E-mail: jsuehnel@imb-jena.de. Fax: +49 3641656210.
† Revotar Biopharmaceuticals AG.
‡ Accelrys Incorporated.
§ Institut für Molekulare Biotechnologie.

10.1021/jp012546t CCC: $20.00
Published on Web 11/27/2001

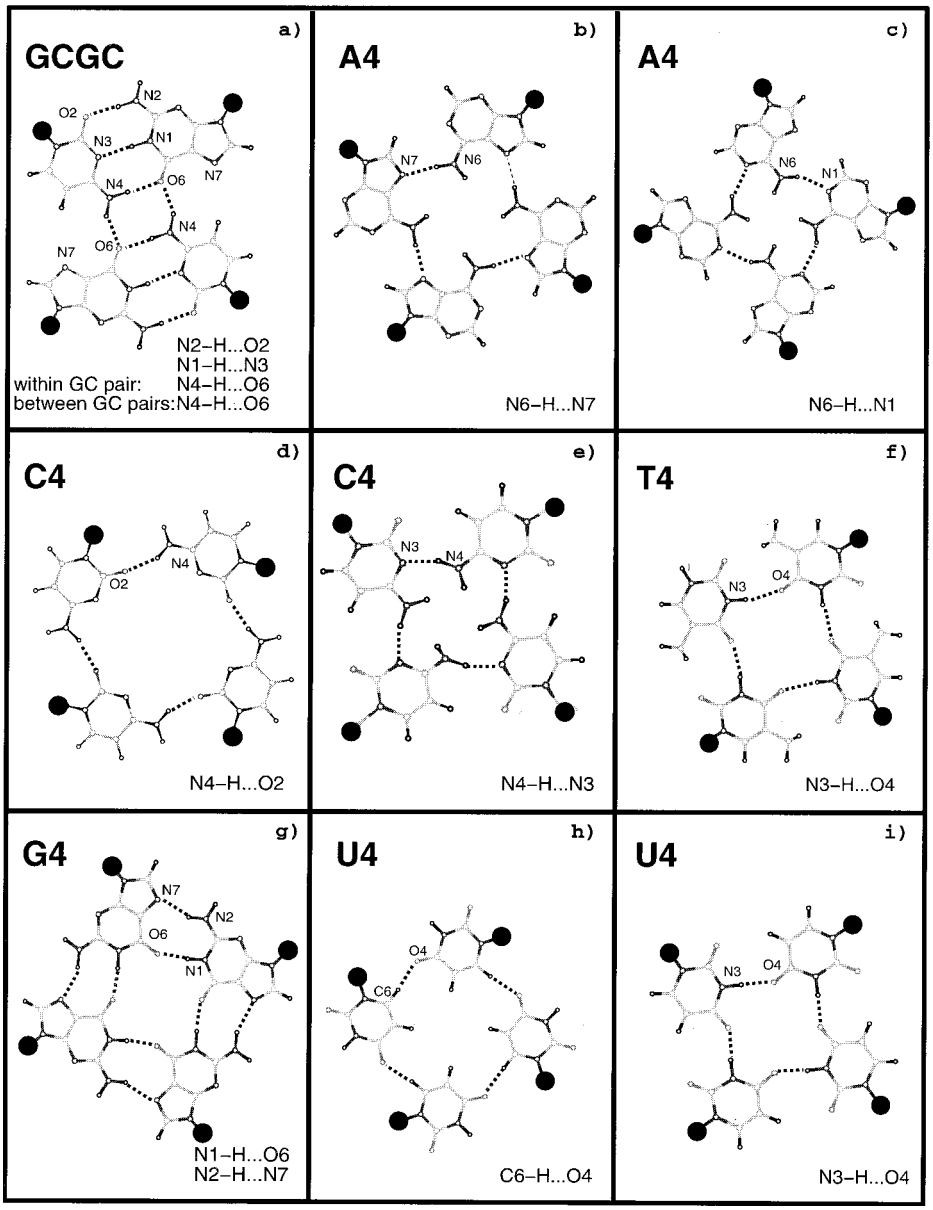

Figure 1. H-bond patterns of cyclic base tetrads found in experimental tetraplex structures. (a) GCGC (1A8N, 1A8W), (b) A4: N6−H···N7 H-bonds (1EVN), (c) A4: N6−H···N1 H-bonds (1EVM), (d) C4: N4−H···O2 H-bonds (1EVO), (e) C4: N4−H···N3 H-bonds (1EMQ), (f) T4 (1EMQ), (g) G4 (1D59), (h) U4: C6−H···O4 H-bonds (1RAU), (i) U4: N3−H···O4 H-bonds (1RAU). The full circles indicate the links to the sugar−phosphate backbone in nucleic acid tetraplexes. The PDB codes are given in parentheses.

T_4-GGGC) that is observed in adeno-associated viral DNA and in human chromosome 19 and with the sequence d(GGGC-T_3-GGGC) found in the fragile X syndrome triplet repeat.[21] A further GCGC tetrad occurs in a cyclic octamer.[22] The formation of GCGC tetrads by dimerization of Watson−Crick GC may also be relevant to the pairing of helical stems occurring in genetic recombination.[23] A TGGT tetrad has been found in a structure containing a DNA box in human centromeric α satellite,[7] and an ATAT tetrad occurs as a structural element in a synthetic DNA octamer.[22]

From a topological point of view, base tetrads may exhibit a cyclic, linear, or branched base−base interaction pattern. Almost all base tetrads known thus far adopt a cyclic structure (Figure 1). The only exception we are aware of is the TGGT tetrad structure mentioned above. In this case, each thymine is linked to one of the guanines of the central GG pair. This corresponds to a linear topology. To our knowledge, branched tetrads have not been found up until now.

Like for nucleic acid base pairs[24−26] and for the G tetrads,[13−16] quantum-chemical studies on these new base tetrads can be expected to yield valuable information in addition to the experimental data. Thus far, quantum-chemical studies on non-G tetrads include the ATAT,[27] AGAG,[27] GCGC,[28] U,[15] and very recently the A[17] tetrad. MD studies have been carried out for tetraplex structures containing GCGC tetrads.[29]

Here, we report on a quantum-chemical analysis of the properties of A, T, C, and GCGC tetrads with a cyclic topology resembling the tetrad H-bond patterns occurring in experimentally known 3D DNA structures. Furthermore, we include in the discussion results on U and G tetrads previously reported.[15] The work is focused on geometries with H-bond patterns similar to the ones occurring in nucleic acid structures. It is not aimed at a comprehensive exploration of the tetrad energy hypersurface.

The quantum-chemical analysis presented in this paper provides new data on geometrical, energetic, and charge

distribution features of nucleic acid base tetrads and includes, for example, information on cooperativity as a possible source for the enhancement of base pairing interaction energies and on the intrinsic planarity or nonplanarity of the base complexes. It should be noted that incorporation of nucleic acid base complexes into a nucleic acid environment may lead either to a deviation form planarity for intrinsically planar structures or to a (partial) planarization of intrinsically nonplanar structures. With the results reported here, it should be possible to take a step toward a comprehensive classification of nucleic acid base tetrads.

Even though all calculations described in this paper do not take into account metal ions, the discussion of the results is performed in the light of potential interactions with a metal ion, since this is an important aspect for many tetrad structures.

Base tetrads are relatively large systems for quantum-chemical calculations with the inclusion of electron correlation. It would thus be advantageous if the less time-consuming density functional theory (DFT) could be applied. We and others have recently shown that this approach is indeed suitable for the treatment of H-bonded base complexes,[15,27−32] unusual nucleic acid analogues,[33] interactions between metal ions and nucleic acids,[14,15,34] and other aromatic bioorganic compounds.[35,36] Therefore, we adopt the DFT approach for the calculations performed in this study.

As complete medium-sized biopolymers for which 3D structures are available still cannot be studied by nonempirical quantum-chemical methods, the results for the tetrad geometries and interaction energies are compared to force field calculations. The assessment of the empirical force field parameters by means of results obtained from quantum-chemical methods is important for the reliability of all approaches using these force fields. Therefore, we have selected both the AMBER force field[37] widely used for biopolymer structures and the MMFF force field designed for organic molecules in general.[38]

Methods

Initial structures of the tetrads have been generated from the tetraplex coordinates with A (1EVM, 1EVN),[18] C (1EVO,[19] 1EMQ[20]), T (1EMQ),[20] and GCGC tetrads (1A8N,[21] 1A8W[23]) deposited at the Protein Databank (PDB). The PDB codes are given in parentheses. The backbone has been deleted, and the bases have been capped with hydrogen atoms. It has been shown previously that capping of bases has no significant influence on hydrogen-bonding (H-bonding) between the bases in GC pairs.[30] We have started the calculations with geometries closely related to the experimental geometries (Figure 1). For the A tetrad, two different conformations termed N7 and N1 structures with either N6−H61···N7 or N6−H62···N1 H-bonds have been taken into account, since both conformations have been found in the human telomere DNA nuclear magnetic resonance (NMR) structure (1EVM, 1EVN).[18] Similarly, two C tetrad structures with H-bonds from the hydrogen atoms at N4 to O2 (1EVO) and N3 (1EMQ) have been studied. The tetrads located in helix centers are often approximately coplanar whereas tetrads at the helix termini may also adopt nonplanar geometries. Incorporation of nonplanar base tetrads into a helix environment leads very likely to a partial planarization. To get information on the energy required for this process for intrinsically nonplanar base tetrads, we have therefore also studied the planar structures. To save computation time, the geometry optimizations have been performed imposing symmetry constraints taking into account planar (C_{4h}: A_4, T_4, C_4; C_{2h}: CGCG) and nonplanar (S_4: A_4, T_4, C_4; C_4: A_4, T_4, C_4; C_i: GCGC) structures. Minor deviations of the tetrad geometries from the selected symmetries have been removed prior to optimizations with UNICHEM.[39]

Monomers and tetrads have been studied using the B3LYP hybrid DFT method[40,41] and the DFT-optimized DZVP and TZVP basis sets with the contraction schemes {H (5s) → [2s], B−Ne (9s, 1p, 1d) → [3s, 2p, 1s]} and {H (5s, 1p) → [3s, 1p], C−F (10s, 6p, 1d) → [4s, 3p, 1d]}, respectively.[42,43] Unless otherwise stated, all quantum-chemical data presented in this work refer to TZVP calculations. The DZVP results are listed in the Supporting Information for comparison. The optimized structures have been checked with subsequent frequency calculations, and interaction energies have been corrected for the basis set superposition error with the standard counterpoise method.[44] The DFT and supplementary semiempirical calculations were carried out with GAUSSIAN94 and 98.[45] For comparison, we have also optimized the structures both with the MMFF94 force field[36] implemented in Sybyl 6.6[46] and with the Cornell et al. force field[37] of the AMBER package.[47]

In both cases, a dielectric constant of 1.0 was used and the convergence criteria for the forces were assumed to be 0.0001 in Sybyl and 0.000 02 kcal/mol Å in AMBER, respectively. For the AMBER calculations, partial charges of the capping hydrogen atoms have been obtained by neutralizing the total charge; for the MMFF94 calculations, the default charges of this force field have been used.

The interaction energies are calculated according to a previously described scheme.[15] The most important points are briefly summarized below. The total interaction energy of the tetrads was calculated according to eq 1, where E(B4) denotes the energy of the tetrad consisting of four bases and E(B) is the energy of a single base in the full tetrad-centered basis.

$$\Delta E = E(\mathrm{B4}) - 4E(\mathrm{B}) \tag{1}$$

For the tetrads with four symmetry equivalent bases, the total interaction energy can also be expressed according to eq 2

$$\Delta E = 4\Delta E^{\mathrm{n}} + 2\Delta E^{\mathrm{d}} + \Delta E^{\mathrm{c}} \tag{2}$$

ΔE^{n} is the interaction energy between neighbor base pairs, the interaction energy between diagonal opposite base pairs is given by ΔE^{d}, and ΔE^{c} is the cooperative contribution, which is adjusted so that the sum equals ΔE. For the GCGC tetrad, two different neighbor interaction energies exist for the WC and the non-WC base pair interactions, as can be seen from the structure shown in Figure 4. Similarly, there are no longer two equivalent interaction energies between the diagonally opposite bases. Instead, two terms have to be introduced for GG and CC

$$\Delta E = 2\Delta E^{\mathrm{nGC,WC}} + 2\Delta E^{\mathrm{nGC,nWC}} + \Delta E^{\mathrm{d,GG}} + \Delta E^{\mathrm{d,CC}} + \Delta E^{\mathrm{c}} \tag{3}$$

Each base can be deformed from its ideal monomer geometry when complexes are formed, and the corresponding deformation energy (ΔE^{def}) is the energy difference between the structure adopted by a single base in the complex and the optimized structure of this base.

$$\Delta E^{\mathrm{T}} = \Delta E + 4\Delta E^{\mathrm{def}} \tag{4}$$

Furthermore, the zero point vibration energy difference (ΔZPE) between the tetrad and the four individual bases contributes to the ΔE_0 defined as

$$\Delta E_0 = \Delta E^{\mathrm{T}} + \Delta\mathrm{ZPE} \tag{5}$$

TABLE 1: *E* (H) and ZPE (H) of the Monomers Determined with the B3LYP/TZVP Method

	A	C	G	T
E	−467.461 35	−395.061 77	−542.723 79	−454.284 58
ZPE	0.111 21	0.097 81	0.116 14	0.114 25

For the GCGC tetrad, different deformation and zero point energies of G and C were considered. We discuss the interaction between all bases of structures corresponding to local minima at the level of ΔE_0; interaction energies of other structures are discussed in terms of ΔE^T.

Electrostatic potentials have been calculated according to the standard approach implemented in the GAUSSIAN package. Graphical representations of the molecular structures have been generated with InsightII;[48] for the electrostatic potential, UNICHEM[39] has been used.

Results

DFT Calculations. The calculated energies for the monomers are listed in Table 1, structural parameters for the base−base H-bonds within the tetrad structures are given in Table 2, and the tetrad energies are listed in Table 3. The tetrad structural formula is shown in Figure 1, and structures with optimized geometries are displayed in Figure 2.

Monomers. We restrict the monomer analysis to those aspects that are important for tetrad structures and energies. The total and zero point energies of the monomers are summarized in Table 1. In the most favorable thymine methyl group conformation, the in-plane hydrogen is pointing away from O4. The energy of this conformation is 1.22 (DZVP) and 1.27 kcal/mol (TZVP) lower than the one with the in-plane hydrogen rotated toward O4.

The amino groups of A, C, and G may adopt pyramidal or planar structures. The energy differences between the planar and the nonplanar structures corresponding to the inversion barriers are 0.12, 0.17, and 0.16 kcal/mol at DZVP level for A, C, and G, respectively. With the TZVP basis, energy differences of −0.01, 0.00, and 0.56 kcal/mol are predicted with B3LYP. This means that at the DZVP level the optimized structures of A, C, and G have nonplanar amino groups. However, in passing to the TZVP basis set for A and C, there is almost no energy difference between the planar and the nonplanar structures and only G remains nonplanar with an even larger energy difference between the preferred nonplanar and the planar structures. The angles between the amino group plane and the least squares plane of all base atoms except the amino hydrogen atoms are 25, 15, and 41° for DZVP and 22, 15, and 36° for TZVP calculations of the pyramidal structures of A, C, and G.

A Tetrad. The A tetrad has been analyzed in detail for different relative base−base orientations adopting a planar C_{4h} symmetry or nonplanar C_4 and S_4 symmetries. The bases may interact for all symmetries either via N6−H62···N1 or via N6−H61···N7 H-bonds to form cyclic H-bonded tetrads. It turned out, however, that neither the planar structures nor the S_4-symmetric ones correspond to local energy minima. Frequency calculations indicate that only the C_4-symmetric structure with N6−H61···N7 H-bonds corresponds to a local minimum.

For the structures with N6−H61···N7 H-bonds, the planar C_{4h}-symmetric and the nonplanar C_4-symmetric conformations are 0.33 and 0.15 kcal/mol less stable than the S_4-symmetric conformation. For the structures with N6−H62···N1 H-bonds, these energy differences are 6.17 and 3.07 kcal/mol and thus much larger. This shows that for both H-bond patterns S_4-symmetric structures are more stable than bowl type structures with C_4 symmetry. The energy difference between the N7 and the N1 conformations is 8.95 kcal/mol for the planar and 6.02 and 3.10 kcal/mol for the nonplanar structure C_4 and S_4 symmetries. Hence, the S_4-symmetric N7 conformation of the A tetrad is the most stable one among the investigated structures, but it does not correspond to a local energy minimum in contrast to the somewhat less stable C_4-symmetric structure. Similar to the total energy, the interaction energies for this conformation are almost independent of the symmetry (Table 3). The neighbor base interaction energy for the nonplanar S_4-symmetric structures is −5.33 kcal/mol, and the total interaction energy (ΔE^T) is −20.01 kcal/mol. The nonadditive contribution for the four A tetrad conformations is below 14% of ΔE. For the H62···N1 H-bond pattern, the magnitude of the neighbor base interaction energy (ΔE^n) and the total interaction energy (ΔE^T) increase from C_{4h} over C_4 to S_4 symmetry. At the latter symmetry, ΔE^n was calculated to be −4.63 kcal/mol and ΔE^T is −16.71 kcal/mol.

The geometrical analysis shows that the less stable N1 conformation is characterized by an unusual small N6−

TABLE 2: Selected H-Bond Distances (*r*, Å) and Angles (∠, deg) as Well as Plane Angles between Diagonal Bases (*ω*, deg) in Tetrads Derived from DFT and Force Field Calculations

		B3LYP/TZVP	MMFF	AMBER
		A Tetrad (H61···N7)		
r (H61···N7)	C_{4h}	1.994	1.885	1.868
	S_4	1.990	1.978	*a*
	C_4	2.001	*a*	*a*
∠ (N6−H61···N7)	C_{4h}	175.9	170.4	169.1
	S_4	175.8	154.0	*a*
	C_4	172.4	*a*	*a*
ω	S_4	153	126	*a*
	C_4	131	*a*	*a*
		A Tetrad (H62···N1)		
r (H62···N1)	C_{4h}	2.168	2.066	1.910
	S_4	2.109	1.933	*a*
	C_4	2.094	2.015	*a*
∠ (N6−H62···N1)	C_{4h}	152.0	152.0	158.9
	S_4	153.0	153.0	*a*
	C_4	161.3	164.8	168.3
ω	S_4	108	157	*a*
	C_4	97	99	112
		C Tetrad		
r (H41···O2)	C_{4h}	1.844	1.873	1.746
r (H42···N3)	C_{4h}	2.691	2.751	2.630
r (H41···N3)	S_4	1.963	1.905	*a*
∠ (N4−H41···O2)	C_{4h}	163.2	160.5	164.4
∠ (N4−H42···N3)	C_{4h}	101.7	103.4	95.63
∠ (N4−H41···N3)	S_4	159.3	153.3	*a*
ω	S_4	134	140	*a*
		T Tetrad		
r (H3···O4)	C_{4h}	2.133	1.953	1.917
	S_4	1.847	1.686	1.748
	C_4	1.962	1.802	1.808
∠ (N3−H3···O4)	C_{4h}	171.0	167.7	166.9
	S_4	176.5	174.4	175.5
	C_4	165.7	160.8	164.6
ω	S_4	124	118	126
	C_4	139	97	94
		GCGC Tetrad		
r (H21···O2)	C_i	1.842	1.843	1.722
r (H1···N3)		1.905	1.828	1.837
r (O6···H41)		1.874	1.872	1.777
r (N7′···H42)		3.583	3.755	3.869
r (O6′···H42)		1.899	1.778	1.757
∠ (N4−H2···O6′)		148.7	162.6	161.8
∠ (N4−H2···N7′)		151.3	144.0	145.1

[a] Force field structure does not correspond to DFT structure.

TABLE 3: Comparison of the Calculated Tetrad Interaction and Deformation Energies (kcal/mol) Determined with B3LYP/TZVP[a]

tetrad/symmetry	ΔE^{n}	ΔE^{d}	ΔE^{c}	ΔE	ΔE^{def}	ΔE^{T}	ΔE^{ZPE}	ΔE_0	H-bonds
A (N6−H62···N1)/C_{4h}	−2.94	0.24	−1.77	−13.05	0.65	−11.09	*b*	*b*	1
A (N6−H62···N1)/C_4	−3.77	0.01	−1.49	−16.55	0.72	−13.67	*b*	*b*	1
T/C_{4h}	−4.10	0.00	−0.54	−16.94	0.37	−15.46	*b*	*b*	1
A (N6−H62···N1)/S_4	−4.63	0.51	−0.97	−18.47	0.44	−16.71	*b*	*b*	1
T/C_4	−5.23	−0.09	−0.70	−21.80	0.66	−19.16	*b*	*b*	1
A (N6−H61···N7)/C_{4h}	−5.19	0.01	−0.73	−21.47	0.49	−19.51	*b*	*b*	1
A (N6−H61···N7)/C_4	−5.33	−0.02	−0.56	−21.92	0.56	−19.92	3.46	−16.46	1
A (N6−H61···N7)/S_4	−5.33	0.00	−0.61	−21.93	0.48	−20.01	*b*	*b*	1
U[c] (C6−H6···O4)/C_{4h}	−4.39	−0.48	−3.39	−21.91	0.24	−20.95	2.08	−18.87	0 (C−H···O)
T/S_4	−5.50	−0.06	−1.50	−23.62	0.37	−22.14	2.16	−19.08	1
U[c] (N3−H3···O4)/C_{4h}	−6.06	−0.02	−2.79	−27.07	0.53	−24.95	*b*	*b*	1
C/S_4	−8.11	−0.91	−7.58	−41.84	1.07	−37.56	*b*	*b*	1
C/C_{4h}	−10.49	−1.86	−10.08	−55.76	1.65	−49.16	4.76	−44.40	1
G[d]/C_{4h}	−13.79	−2.25	−15.30	−74.96	2.67	−64.28	*b*	*b*	2
G[d]/S_4	−13.42	−2.06	−15.37	−73.17	2.15	−64.57	3.29	−61.28	2
GCGC/C_i	−29.14[e]	1.85[g]	−5.31	−79.05	2.88[i]	−69.47	4.89	−64.58	3[e]
	−9.32[f]	1.68[i]			1.91[j]				1[f]

[a] The B3LYP/TZVP total tetrad energies in Hartrees are as follows: A (N1): −1869.865 32 (C_{4h}), −1869.875 16 (S_4), −1869.870 22 (C_4). A (N7): −1869.879 58 (C_{4h}), −1869.880 10 (S_4), −1869.879 82 (C_4). C: −1580.327 94 (C_{4h}), −1580.310 59 (S_4). T: −1817.173 35 (C_{4h}), −1817.184 64 (S_4), −1817.179 70 (C_4). GCGC: −1875.685 69 (C_i). [b] No local minimum. [c] Ref 15. [d] B3LYP/DZVP, ref 15. [e] Watson−Crick type GC interaction (see eq 3). [f] Non-Watson−Crick GC. [g] GG. [h] CC. [i] G. [j] C.

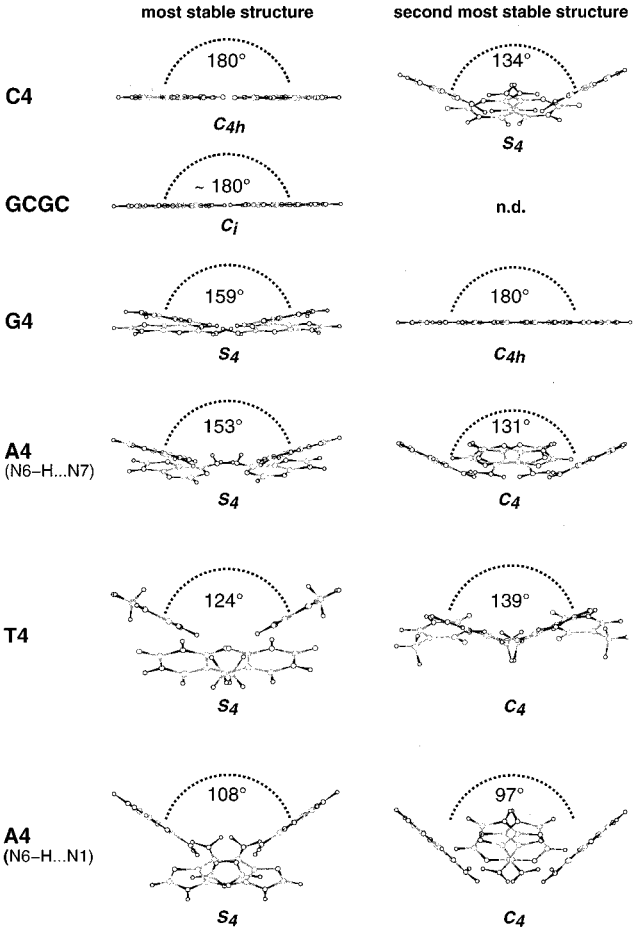

Figure 2. Geometries of the most stable and the second most stable tetrad structures calculated with the B3LYP method. The structures are characterized by the angle between the planes of opposite bases and by symmetry. The following structures correspond to local energy minima. C_4: A tetrad (N6−H61···N7); S_4: G and T tetrad; C_i: GCGC tetrad; C_{4h}: C tetrad.

H62···N1 angle of 152 or 153° in the C_{4h} and S_4 structures. This value is smaller than the almost linear geometry found frequently for H-bonds in base pairs. On the other hand, the corresponding angle for the N6−H61···N7 H-bond is only slightly smaller than 180° for both the planar and nonplanar structures.

A further basic difference between the two conformations refers to the distance between neighbor amino groups pointing to the tetrad center. The distance between these amino group nitrogen atoms is extended from 3.31 Å in the planar H62···N1 structure to 3.71 Å at S_4 symmetry and 3.86 Å at C_4 symmetry. A similar but weaker effect is observed in the N6−H61···N7 structure (4.20, 4.32, and 4.24 Å). The deviation from planarity is much more pronounced for the N1 conformation than for the N7 structure. The root mean square (rms) deviation of all atoms from a plane is only 0.52 and 0.63 Å for the H61···N7 structure at S_4 and C_4 symmetry, respectively, whereas rms deviations of 1.82 and 1.22 Å occur in the N6−H62···N1 structures. The angle between the planes of diagonal opposite bases is 150 and 110° in the N7 and N1 conformations at S_4 symmetry and 131 and 97° at C_4 symmetry, respectively.

Even though the interaction and deformation energies of the previously investigated G tetrad are much larger than the ones of the A tetrad N7, the small energy difference of 0.33 kcal/mol between the C_{4h}- and the S_4-symmetric conformations of this A tetrad is comparable to the G tetrad energy difference of 0.39 kcal/mol.[15] The A tetrad H62···N7 energy difference of 0.18 kcal/mol between the S_4 and the bowl type C_4-symmetric structure indicates a high degree of flexibility, since several conformations correspond to very similar energies. Our calculations show that the C_4-symmetric energy minimum is not the global minimum, since the S_4-symmetric structure with a small imaginary vibrational frequency of 14 cm^{-1} is more stable. The absence of an energy minimum at S_4 symmetry is also a fundamental difference relative to the Hoogsteen type G tetrads.

During finalization of this work, we became aware of HF and B3LYP/6-311G(d,p) calculations on the N7 and N1 conformations of the A tetrad.[17] The results obtained from these calculations are similar to the data we report here, but they do not include pair interaction energies, frequencies, and all analyses of S_4-symmetric structures.

C Tetrad. The C tetrad has been optimized at C_{4h} symmetry starting from two different structures with H-bonds from the N4 amino group to O2 and N3, respectively (Figures 1d,e and 3a,b). The structure with the N4−H···N3 H-bonds (Figure 3a)

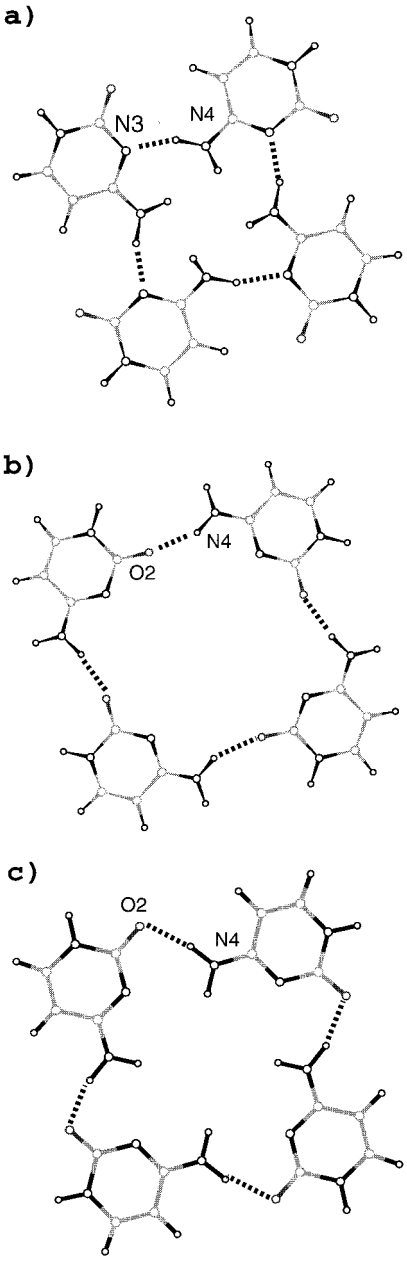

Figure 3. Different H-bond patterns in C tetrads. (a) C tetrad located in the center of a quadruplex structure with the SV40 repeat sequence GGGCGG (PDB code: 1EMQ). (b) C tetrad located at the end of a quadruplex with the *S. cerevisiae* sequence (PDB code: 1EVO). (c) C tetrad structure calculated with the DFT (B3LYP) method.

was not stable but converged to a structure with N4−H···O2 H-bonds (Figure 3c). Interestingly, in this structure, a different hydrogen of the amino group is involved in the H-bond as compared to the experimental structure shown in Figure 3b. This leads to a relatively short distance between the second amino group hydrogen atom and the N3. Frequency calculations indicate that the C tetrad has a local energy minimum for this planar complex geometry. In contrast, S_4-symmetric structures show H-bonds between H41 and N3. This tetrad structure is 10.89 kcal/mol less stable than the C_{4h}-symmetric one, and it does not correspond to a local energy minimum. Initial tetrad geometries with C_4 symmetry converged almost to the planar C_{4h} symmetry. We have excluded this structure from the discussion since it is only 0.13 kcal/mol less stable than the C_{4h}-symmetric conformation, and the rms deviation of these structures is only 0.2 Å.

At C_{4h} symmetry, the monomers are linked by H-bonds between N4−H41 and O2 with a length of 1.844 Å (Table 2). The distance between the other amino group hydrogen atom H42 and the acceptor atom N3 is 2.768 Å, and the angle formed by the donor, hydrogen, and acceptor atoms has a value of 101.7°. Both parameters are outside the limits usually accepted or adopted for H-bond interactions. The distance between the amino group nitrogen atoms is 4.75 Å. This is larger than the corresponding distances of 3.31 and 4.20 or 3.71 and 4.32 found for the nonplanar and planar N1 and N7 conformations of the A tetrad. In addition, the distance between H-atoms pointing to the tetrad center is with 3.35 Å much larger than the sum of the corresponding van der Waals radii. At S_4 symmetry, there is only a single H-bond with a distance of 1.963 Å between H41 and N3 of the neighbor bases; a short distance between the other amino group hydrogen atom and O2 is missing. This leads to weaker interaction between neighbor bases.

The interaction energy (ΔE) of the most stable C_{4h}-symmetric structure amounts to −55.76 kcal/mol with the TZVP basis; the one calculated with DZVP is not much different. As expected, the main contribution to ΔE arises from the four interactions between H-bonded neighbor bases, the interaction energy between the diagonally opposite bases being much smaller but nevertheless slightly stabilizing. At S_4 symmetry, ΔE is significantly smaller (−41.84 kcal/mol) since the pair interaction energy between the neighbor bases is much weaker than for C_{4h} symmetry (Table 3). The cooperative energy (ΔE^c) contributes −10.08 kcal/mol. This corresponds to 18% of ΔE. The total interaction (ΔE^T) energy that takes into account the deformation energies of the four monomers is 6.60 kcal/mol less negative than ΔE. Finally, the zero point energy change of 4.76 kcal/mol leads to a value of −44.40 kcal/mol for ΔE_0.

T Tetrad. Like for the A tetrads and contrary to the C tetrad, the planar structure does not correspond to a local energy minimum as indicated by imaginary frequencies up to 29 cm^{-1}. In the preferred methyl group orientation, the in-plane C−H group is pointing away from O4. The alternative planar conformation with all methyl groups rotated by 180° is 5.42 kcal/mol less stable. This is close to four times the energy difference of 1.27 kcal/mol for a single T. The interaction energy ΔE^T of −15.46 kcal/mol is clearly dominated by interactions between the H-bonded neighbors. The lack of a second H-bond connecting the monomers is the reason for the small pair contribution of −4.10 kcal/mol.

At S_4 symmetry, the T tetrad structure corresponds to a local energy minimum and is 6.96 kcal/mol more stable than the planar and 3.10 kcal/mol more stable than the C_4-symmetric tetrad. In addition, the H-bonds are shorter at S_4 (1.847 Å) than at C_{4h} symmetry (2.133 Å), whereas the distance between the O2 and the methyl carbon increases from 3.089 to 3.525 Å. At C_4 symmetry, these distances adopt the intermediate values of 1.962 and 3.361 Å, but this structure does not correspond to an energy minimum, even though the imaginary frequencies are smaller than the ones of the coplanar structure.

GC Pair and GCGC Tetrad. The optimized GC pair parameters are listed in Table 4. The TZVP interaction energy (ΔE_0) between both bases connected by three H-bonds is −23.59 kcal/mol. In contrast to LMP2 calculations,[49] the amino groups do not remain considerably pyramidal in the complex. Therefore, the GC pair is almost planar. The rms deviation of all atoms from a plane is 0.01 Å.

Similarly, the GCGC tetrad is approximately planar. The C_i-symmetric structure is somewhat more stable than the completely planar structure optimized at C_{2h} symmetry. The WC

TABLE 4: Summary of the Calculated Data for the Watson−Crick GC pair: E (H), ZPE (H), ΔE^{def} (kcal/mol), ΔE^{T} (kcal/mol), and ΔE_0 (kcal/mol) for the GC Pair and H-Bond Distances (Å) at the B3LYP/TZVP Level

E	−937.827 24	ΔZPE	1.64
ZPE	0.216 56	ΔE_0	−23.59
ΔE	−28.69	r (H21···O2)	1.912
ΔE^{def}	1.91[a], 1.55[b]	r (H1···N3)	1.913
ΔE^{T}	−25.32	r (O6···H41)	1.757

[a] G. [b] C.

type H-bonds between G and C in the tetrad differ from those in the GC pair (Tables 2 and 4). The distances H21···O2 and H1···N3 are shorter in the tetrad, whereas O6···H41 is longer. The latter H-bond is the shortest one in both systems.

GCGC may be considered as a set of two strongly bonded WC GC pairs ($\Delta E^{nGC,WC} = -29.14$ kcal/mol) interacting by two H-bonds between both H42 and O6 atoms ($\Delta E^{nGC,nWC} = -9.32$ kcal/mol). The GG and CC interactions are even slightly repulsive. The interaction energies (ΔE_0) of the GCGC tetrad is by 17.40 kcal/mol more negative than for the sum of two isolated GC pairs.

A comparison of the interaction energy between the single GC pair ($\Delta E = -28.69$ kcal/mol) with the Watson−Crick GC pairwise interaction energy in the tetrad ($\Delta E^{n,GC,WC} = -29.14$ kcal/mol) indicates that the presence of a second GC pair leads to a slight strengthening of the base−base interaction energy. The cooperative interaction energy contributes further to the tetrad stability, whereas the deformation energies are increased as compared to the GC pair.

Finally, it should be noted that an additional energy minimum exists for the GCGC tetrad when both WC GC pairs are translated relative to each other and form H-bonds between N4−H42 and N7 instead of O6. This minimum is, however, by 2.27 kcal/mol less stable.

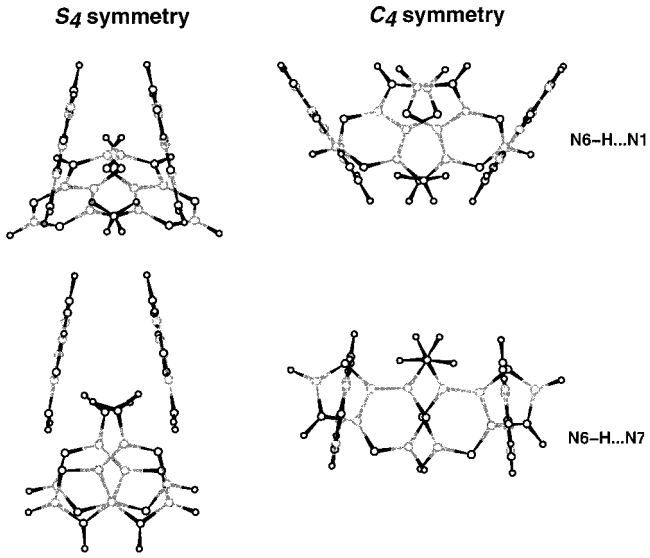

Figure 4. Structures of A tetrads optimized with the AMBER force field.

Force Field Calculations. Geometrical data of the molecular mechanical calculations with the MMFF and AMBER force fields on the tetrads are listed in Table 2, and the interaction energies determined according to eqs 1 and 4 are reported in Table 5. For A tetrads, the AMBER geometries differ substantially from the DFT results. These structures are shown in Figure 4. Both force fields yield for T the same preferred orientation of the methyl groups as the DFT approach. For example, a 180° rotation of the methyl group is 1.81 kcal/mol less favorable according to a MMFF calculation. The MMFF force field takes into account the pyramidal amino groups whereas the AMBER force field does not. Therefore, the AMBER calculations yield

TABLE 5: Force Field Interaction Energies ΔE and ΔE^{T} and ΔE^{def} (kcal/mol)[a]

tetrad			ΔE	ΔE^{n}	ΔE^{d}	ΔE^{def}	ΔE^{T}
A	(N6−H62···N1)	C_{4h}	−14.2	−3.6	0.1	0.9	−10.6
			−24.0	−6.1	0.2	0.3	−22.8
U	(C6−H6···O4)	C_{4h}	−14.0	−3.2	−0.6	0.0	−14.0
			−16.2	−3.9	−0.3	0.0	−16.2
A	(N6−H61···N7)	C_{4h}	−18.6	−4.7	0.1	0.8	−15.4
			−29.2	−7.3	0.0	0.2	−28.4
A	(N6−H62···N1)	C_{4h}	−21.6	−5.4	0.0	1.1	−17.2
A	(N6−H62···N1)	S_{4h}	−23.2	−6.2	0.8	0.4	−21.6
A	(N6−H61···N7)	S_{4h}	−27.6	−7.4	1.2	0.3	−26.4
T		C_{4h}	−27.6	−6.8	−0.2	0.4	−26.4
			−29.6	−7.3	−0.2	0.6	−27.2
T		S_4	−32.2	−7.7	−0.7	0.2	−26.4
			−36.0	−8.8	−0.4	0.1	−35.6
T		C_4	−30.6	−7.4	−0.5	0.8	−27.2
			−36.6	−8.9	−0.5	0.6	−34.2
C		C_{4h}	−39.0	−9.0	1.5	1.5	−36.0
			−50.4	−12.3	−0.6	0.7	−47.6
		S_{4h}	−44.8	−10.7	−1.0	0.8	−41.6
U	(N3−H3···O4)	C_{4h}	−34.4	−8.4	−0.4	0.1	−34.0
			−34.6	−8.5	−0.3	0.1	−34.2
GCGC		C_i	−64.0	−22.6[b], −10.1[c]	−0.2[d], 1.2[e]	0.6[f], 0.6[f]	−61.7
			−78.0	−29.5[b], −11.1[c]	0.7[d], 2.5[e]	0.6[f], 0.6[g]	−75.7
G		C_{4h}	−77.0	−17.7	−3.1	0.3	−75.8
			−70.2	−17.0	−1.1	0.5	−68.2
		S_{4h}	−75.8	−17.5	−2.9	1.1	−71.4

[a] MMFF first line, AMBER second line for each molecule. [b] $\Delta E^{nGC,WC}$. [c] $\Delta E^{nGC,nWC}$. [d] $\Delta E^{d,GG}$. [e] $\Delta E^{d,CC}$. [f] $E^{def,C}$. [g] $E^{def,G}$. [h] Force field structure does not correspond to DFT structure.

planar structures for A, C, and G. The inversion barriers are 0.7, 1.0, and 1.1 kcal/mol for A, C, and G according to the MMFF calculations. The angles formed between the plane of the amino group atoms and the other base atoms including the amino nitrogen atoms are 41, 43, and 44°.

As the energies are additive, no cooperative term like in eq 2 is taken into account. A comparison between the B3LYP, MMFF, and AMBER data is presented below.

Discussion

Monomer Geometries and Energies. Previous ab initio studies have shown that the amino groups of nucleic acids and other organic molecules have a nonplanar pyramidal structure.[25,49] In contrast to the geometry, the inversion barriers are sensitive to the basis. MP2/6-311G(2df,p) calculations provided an estimate of 0.13, 0.15, and 1.12 kcal/mol for A, C, and G, whereas MP2/6-31G(d) barriers are somewhat larger.[25] B3LYP/TZVP agrees with these data insofar, as the barrier for G is the highest one (0.56 kcal/mol), and for A and C, almost no barrier exists. The observation that the G inversion barrier from B3LYP calculations seems somewhat too small corresponds to the underestimation of the coplanar torsional barrier of aromatic compounds.[36] It has been proposed that DFT calculations tend to overestimate the stability of the coplanar conformations of aromatic compounds.[50] A detailed comparison with extended basis sets and correlation techniques would, however, be necessary to analyze the inversion barriers in detail. This is outside of the scope of our work.

As expected, the MMFF94 force field leads to a pyramidal structure of amino nitrogen atoms in the monomers. The inversion barriers, however, are somewhat larger than the DFT results. On the other hand, the AMBER minimizations lead to planar amino groups. This different behavior of force field and DFT calculations has to be taken into account when analyzing the tetrad geometries. In summary, the amino group of guanine is the only case for which a significant deviation from a planarity is well established.

Base Tetrad Geometries. Ideal DNA conformations are characterized by planar and parallel base pair planes. In real nucleic acid structures, however, base pairs deviate more or less from this ideal conformation. These deviations may be either an intrinsic property of base pairs or enforced by interactions with the nucleic acid backbone, neighboring base pairs/bases, metal ions, or the solvent. It may happen that both intrinsically planar base pairs become nonplanar in nucleic acid structures and nonplanar base complexes are subject to a (partial) planarization upon incorporation in a nucleic acid. The quantum-chemical studies reported in this paper provide information on the intrinsic properties of base tetrads and therefore represent the starting point for a full understanding of base tetrad geometries in nucleic acid structures.

The optimized base tetrad geometries exhibit marked differences. The planar C tetrad corresponds to a local energy minimum, and also, the GCGC tetrad is practically planar and represents a local minimum. In the U tetrad, the bases interact either via N3−H···O4 (U4−NHO) or via C6−H···O4 (U4−CHO) H-bonds. In contrast to the U4−CHO structure, the more stable planar U4−NHO structure does not correspond to a local minimum. The energy difference between the planar C_{4h}-symmetric and the slightly nonplanar S_4-symmetric G tetrads with Hoogsteen type H-bonds is very small and is difficult to estimate with high accuracy.[16,17] On the contrary, for the T tetrad, the nonplanar S_4 structure is significantly more stable than the C_{4h} and C_4 structures, the energy differences being 7.08 and 3.10 kcal/mol. In the S_4 structure, the H-bonds between neighbor bases are shorter and the distance between O2 and the methyl carbon of neighbor bases is longer than in the other conformations. For both A tetrads with N6−H···N7 and N6−H···N1 H-bonds, again the nonplanar S_4 structures are more stable than the C_{4h}- and C_4-symmetric ones. However, in this case, neither the planar C_{4h} nor the nonplanar S_4 structures correspond to local minima.

How can the different structural features of the tetrads be understood? The formation of cyclic tetrads requires the interaction of each single base with the two neighbor bases via two different base edges. If the angle between the base edges adopts values around 90° and if the H-bond donor/acceptor sites have an appropriate orientation, an effective base−base H-bond interaction within a plane is possible. This requirement is best met by the G tetrad. Here, the Watson−Crick and the Hoogsteen edges allow for an interaction between neighboring bases via two H-bonds. The slight nonplanarity is very likely due to the nonplanar amino group.

In the A Hoogsteen edge, the O6 atom found in G is replaced by an amino group and there is no amino group in 2-position. Therefore, for A, an interaction pattern analogous to G is not possible. If one does not take into account the C2−H group as a possible donor, then only the amino group in 6-position remains as a donor. It can interact with either N7 or N1. This leads to a completely different interaction pattern as compared to the G tetrad. Especially in the N1 conformation, the distance between the amino groups is relatively small and this probably induces a deviation from planarity. We have also searched for alternative conformations, but at C_4 symmetry, we have found no structure that is more stable by means of the DFT approach. However, the AMBER force field optimization yields two extremely nonplanar structures with relatively normal H-bond geometries (H···N distance $\sim$ 1.9 Å, N−H···N angle $\sim$ 155°) but with much different relative orientations of the bases as compared to the DFT structure (Figure 4). These structures do not show the H-bond patterns found in the experimental structures of the corresponding nucleic acid. MMFF predicts also that this structure is more stable than the DFT-like ones, but this minimum was not reached from the DFT structure as a starting point for energy minimization. The N7 conformation is characterized by the fact that each amino group forms two H-bonds to the N1 atom of one neighbor base and to the N7 atom of the other neighbor. This is the most stable structure we have found by means of force field calculations. There is an additional and less stable structure with the N6−H··N1 pattern but again with the unusual orientation of neighbor base planes. The results on the A tetrads bear resemblance to the AGAG tetrad obtained by DFT calculations where also an extremely distorted V-shaped conformation has been found.[27]

Cytosine is the pyrimidine base in which the donor group amino group in 4-position and the acceptor group O2 is relatively far away. Therefore, the tetrad structure with N4−H···O H-bonds is planar. Contrary to G, the amino group is planar in the monomer and in the C tetrad.

In thymine, the alternating pattern of neighboring N3−H and O4 groups involved in the base−base H-bond interaction is similar to the N1−H and O6 sites in the G tetrad and to the N3−H and O4 sites in the U tetrad. However, contrary to the G tetrad, there is no additional H-bond interaction between the bases. The bulky methyl group of T leads to the rather long H-bond distance of 2.133 Å in the planar structure and is the main reason for the greater stability of the nonplanar conforma-

tion. In passing from the planar to the most stable nonplanar S_4 structure, the H-bond length is reduced to 1.847 Å and the distance between O2 and the methyl group carbon is increased to 3.525 Å.

The lack of the bulky methyl groups leads to shorter N3−H3···O4 H-bonds of 1.780 Å in the coplanar U tetrad. However, this conformation does not correspond to a local energy minimum.[15]

In summary, we find that the cyclic tetrad geometries with H-bonding patterns occurring in experimental nucleic acid structures are planar for the GCGC and C tetrads, slightly nonplanar for the G tetrad (S_4, $\omega = 159°$), and deviate significantly from planarity for the T tetrad (S_4, $\omega = 124°$). These structures correspond to local energy minima. The nonplanarity of the A tetrads is more pronounced in the structures with H62···N1 H-bonds (S_4, $\omega = 108°$; C_4, $\omega = 97°$) than in the structures with H61···N7 H-bonds (S_4, $\omega = 153°$; C_4, $\omega = 131°$). From these structures, only the latter corresponds to a local energy minimum, even though it is not the most stable one. For the nonplanar structures of the T tetrad and the A tetrad with H62···N1 H-bonds, a comparison of the pair interaction energies between neighbor bases (Table 3) shows that unfavorable repulsions preventing these structures from being planar can be better avoided in the S_4-symmetric than in the C_4-symmetric conformation. For the less nonplanar structures of the A tetrad with N6−H···N1 H-bonds and for the T tetrad, the S_4-symmetric structures are also more stable than the C_4 structures. However, in these cases, the cooperativity contribution is the relevant quantity.

Base Tetrad Stabilities, Interaction Energies, and Cooperativity. In nucleic acids, stacking interactions with other bases, the backbone, and the presence of metal ions and solvent effects do affect nucleic acid structure and stability in addition to interactions within the base tetrads studied in this paper. Therefore, the energies obtained for the isolated base tetrads cannot be directly compared to thermodynamic data for nucleic acids. Nevertheless, the interaction energies of the base tetrads constitute one factor of the overall stability and are thus useful for a detailed understanding of this property.

When discussing the interaction energies of tetrad structures, one should realize that many of these structures obtained under symmetry restraints do not correspond to local minima (Table 3). In addition, as recently shown for the Watson−Crick AT pair,[51] it may happen that the global minimum structure cannot occur in nucleic acids because the nitrogen linked to the sugar−phosphate backbone is involved in an H-bond. This explains also the fact that in some cases the most stable conformation of the conformational subset studied here does not correspond to a local minimum. In the following discussion, we take into account all structures and indicate, however, if the structure discussed represents a local minimum or not.

Nonplanar structures may be (partly) planarized when incorporated in a nucleic acid environment. Therefore, the energy difference between the optimal nonplanar structures and the planar ones provides some information whether this planarization can easily occur. For the G tetrad[15] and for the N7 conformation of the A tetrad, the energy differences between the planar and the nonplanar conformations are smaller than 0.5 kcal/mol (Table 3), whereas the other A conformations and the T tetrad show significant energy differences: T tetrad, 3.86 kcal/mol (C_4) and 6.96 kcal/mol (S_4); A tetrad (N1 conformation), 3.07 kcal/mol (C_4) and 6.17 kcal/mol (S_4). In general, the S_4-symmetric structures are more stable than the C_4-symmetric ones.

In addition to the geometries, the interaction energies enable a comparison of different tetrads found in tetraplexes. In the case of local energy minima, ΔE_0 can be used for a discussion; in other cases, ΔE^T is appropriate. Local energy minima exist for the practically planar C and GCGC tetrads and, in addition, for the C_4-symmetric A tetrad with N6−H61···N7 H-bonds and for the S_4-symmetric T tetrad. Previously, local energy minima have also been found for the Hoogsteen type G tetrad with two H-bonds between H22 and N7 and H1 and O6, respectively, and for the U tetrad with C6−H6···O4 interactions. The absolute values of interaction energies (ΔE_0) for the most stable planar or nonplanar structures decrease in the order GCGC > G ≫ C ≫ T > U (C6−H6···O4) > A (N6−H61···N7) (Table 3). These interaction energies cover a range between −64.58 and −16.46 kcal/mol. The corresponding sequence for ΔE^T of the planar structures is GCGC > G ≫ C ≫ U (N3−H3···O4) > U (C−H6···O4) > A (N6−H61···N7) > T > A (N6−H62···N1), and the same ordering is also obtained when for all tetrads the most stable symmetry is taken into account. Thus, as expected, tetrads having only one H-bond or CH···O interactions between the bases have the lowest interaction energy, whereas the almost planar GCGC with three H-bonds in the Watson−Crick GC pairs has the highest interaction energy. Finally, if all planar and nonplanar conformations are taken into account, then ΔE^T covers a range between 11.09 kcal/mol for the planar A tetrad in the N1 conformation and −69.47 kcal/mol for the GCGC tetrad. The total energies indicate that the GCGC tetrad is more stable than two isolated WC GC pairs. The interaction energy (ΔE^T) of the U tetrad with N3−H3···O4 H-bonds exceeds the one of the corresponding T tetrad by about 10 kcal/mol. This is in agreement with shorter H-bonds in U tetrads.

The interaction energy (ΔE) has been determined recently for the AGAG and TATA tetrads.[27] These energies of −29.41 and −32.46 kcal/mol are smaller than the ones of the GCGC, G, and C tetrads but comparable to our results for the A, T, and U tetrads.

The order of tetrad interaction energies parallels the deformation energies. This is easy to understand because a strong interaction between neighboring bases should have a large geometrical effect. The main contribution to the interaction energy (ΔE) arises from the interactions between neighbor bases. The pair interaction between diagonally opposite bases is much smaller though still attractive in most cases. For the T tetrad, this term is almost zero, and in the GCGC tetrad, there is even a repulsive interaction between diagonal bases.

Relative and absolute contributions of the cooperative effect (ΔE^c) to the total interaction energy (ΔE) exhibit marked differences (Table 3). The GCGC tetrad shows the largest interaction energy, but ΔE^c contributes only 7% to ΔE. The largest absolute values of ΔE^c are found for the G and C tetrads with −15.30 and −10.08 kcal/mol, respectively. This corresponds to 21 and 18% of ΔE. In these two cases, the nonadditive contribution reaches the magnitude of the neighbor base−base interaction energies (ΔE^n). Among the most stable tetrad structures corresponding to local minima, the C_4-symmetric A tetrad (N6−H61···O4) and the S_4-symmetric T tetrad have the smallest relative cooperativity contribution with 3 and 6%, respectively.

Comparing Calculated Tetrad Geometries with Tetrad Structures Occurring in Nucleic Acids. The comparison of optimized base tetrad geometries obtained from quantum-chemical studies with the geometry of tetrads occurring in experimental nucleic acid structures yields information whether the tetrad geometries are governed by intrinsic tetrad properties

alone or are affected by other factors such as solvent effects, the presence of metal ions, or additional interactions within the nucleic acid structure.

For Watson−Crick base pairs and some other noncanonical base pairs, it has been found that the calculated and experimental geometries are similar, but an influence of the environment is noticeable.[30] However, for base tetrads, we come across a different situation. In this case, the presence of metal ions is a sine qua non for the formation of tetraplex structures. Therefore, a detailed comparison of the calculated geometrical parameters of isolated tetrads without metal ions with experimental structural data that do only form in the presence of metal ions is not appropriate. For example, it is well-known that the O6···O6 distance in G tetrads and even their deviation from planarity is heavily affected by the type of metal ion on.[15] Ions may also change the tetrad H-bond pattern.

The only two geometrical aspects we want to discuss here in relation to experimental structures are planarity vs nonplanarity and the qualitative H-bond pattern. Stacking interactions with the two neighbor base pairs or tetrads are especially effective if all base complexes adopt a planar geometry. Therefore, it is not unlikely that intrinsically nonplanar base tetrads become more planar upon incorporation in a helix center. On the other hand, at helix ends, this force can only act from side; therefore, this planarization should not occur to the same extent. One should be aware of the fact, however, that both in the center of relatively normal helix structures base pairs or tetrads and at helix ends intrinsically planar base complexes may show significant deviations from planarity.

An example for the different geometry of tetrads in the helix center and end is found in a parallel-stranded DNA tetraplex formed by *S. cerevisiae* telomere sequence d-TGGTGGC (PDB code, 1EMQ).[20] The central T tetrad atoms (except for methyl hydrogen atoms) show an rms deviation from the least squares plane of only 0.12 Å in the NMR average structure, whereas the corresponding rms value for the terminal T tetrad is 0.35 Å. A similar pattern is found for the angles between the planes of diagonal bases ω. This angle is 172° in the central tetrad and 157−159° at the helix end. Note, however, that for the experimental structures the angle ω may have a different meaning if these structures do not exhibit an at least approximate S_4 symmetry. Therefore, as a more general measure of deviation from planarity, the rms deviation from the tetrad least squares plane is used for the experimental structures.

As compared to the rms value of the experimental structure for the T tetrad located at the helix end of 0.35 Å, the calculated results obtained with B3LYP, MMFF, and AMBER, respectively, are 1.43, 1.59, and 1.20 Å for S_4 symmetry and 0.46, 0.64, and 0.68 Å for C_4 symmetry. This may indicate that even for base tetrads located at a helix end a partial planarization does occur.

The C tetrad adopts a nearly coplanar conformation in the central part of the tetraplex structure with the SV40 repeat sequence d-GGGCGG (PDB code, 1EVO).[19] However, a more detailed comparison of the NMR and the planar DFT structure reveals a striking difference even though the calculations have been started with the experimental H-bond pattern (Figure 3). In the optimized DFT model, the bases are rotated relative to each other so that N4−H41···O2 H-bonds are formed, which enables also a relatively short distance between N4−H42 and N3 (Figure 3c). Instead, in the NMR tetraplex structure, N4−H42···O2 H-bonds are formed. Whether this structural difference is due to the effects of metal ions remains to be clarified. This is also true for the fact that the C tetrad conformation with N4−H···N3 H-bonds found in the structure with the PDB code 1EMQ is not stable in the DFT calculations but converges to the conformation characterized above.

TABLE 6: Mulliken Charges of Selected Monomer and Tetrad Atoms

	C	C Tetrad[a]
C2	0.43	0.53
O2	−0.48	−0.58
H41	0.23	0.30
H42	0.25	0.24
	T	**T Tetrad[a]**
N3	−0.39	−0.39
H3	0.27	0.36
O4	−0.46	0.52
C4	0.47	0.49

	A	A Tetrad[a] H61···N7	A Tetrad[a] H62···N1
N1	−0.36	−0.38	−0.46
N7	−0.34	−0.44	−0.34
N6	−0.47	−0.49	−0.45
H61	0.24	0.21	0.31
H62	0.25	0.32	0.19

	G or C	GC Pair	GCGC Tetrad[b]
O6	−0.46	−0.56	−0.66
H1	0.25	0.34	0.34
H21	0.22	0.32	0.33
H22	0.25	0.23	0.23
O2	−0.48	−0.57	−0.59
N3	−0.39	−0.54	−0.56
H41	0.25	0.33	0.29

[a] C_{4h} symmetry. [b] C_i symmetry.

It would be interesting to know if nonplanar base tetrads in experimental nucleic acid structures prefer conformations resembling either C_4 or S_4 symmetries. Unfortunately, this analysis is hampered by the fact that the asymmetric environment prevents tetrad structures to adopt exactly conformations with these symmetries. Nevertheless, we found that the experimental structure of the A tetrad with H61···N7 H-bonds (PDB code, 1EVN) converges to a C_4-symmetric structure when optimized at the semiempirical PM3 level.

Partial Charges. Mulliken charges can be used to provide a crude estimate of the change of the partial charges when complex formation happens. Previously, it has been shown that there is a considerable charge transfer in G upon complex formation with metal ions.[15,52] As the monomers are symmetry-related in the neutral A, C, G, and T tetrads discussed here, all monomers remain neutral as well. A change of partial atomic charges within G can be observed when comparing G monomers with G tetrads.[15] Similarly, in the AT pair, Mulliken charges of the atoms involved in H-bonding change.[33]

This can also be observed in the C tetrad showing a considerable cooperativity. For example, the negative charge of O2 increases from −0.48 in the monomer to −0.58 in the tetrad, whereas the positive charge of C2 increases from 0.43 to 0.53 (Table 6). In other words, the carbonyl group is more polar in the tetrad. The charge of the hydrogen atom H41 involved in the H-bond increases from 0.23 to 0.30, but the charge of H42 remains almost constant.

In the T tetrad, the partial charges of the carbonyl group and of H3 change in the same way but to a smaller extent. A comparison of both A tetrads with the monomer indicates that only for the nitrogen atom acting as H-bond acceptor the negative charge increases, whereas the other ones are less

affected. Like in the C tetrad, the positive charge increases only for the one amino group hydrogen atom that is involved in H-bonding.

In the GC pair and in the GCGC tetrad, the negative charges of the acceptor atoms and the positive charges of the corresponding hydrogen atoms increase upon complex formation. In the GCGC tetrad, the negative charge of O6 exceeds the one of the GC pair since this atom is involved in an additional H-bond. Similarly, the positive charge of H42 increases when the tetrad is formed from two GC pairs. In the GC pair, there is a charge transfer of 0.01e from C to G; in the GCGC tetrad, it is 0.02e. This charge transfer is much smaller than the one between G tetrads and metal ions.[15]

It should be noted, however, that Mulliken charges are basis set-dependent. Therefore, we have calculated electrostatic potential-derived charges for comparison. According to this method, the charge transfer within the GC pair and the GCGC tetrad is also 0.02e and the negative charge at O6 in the base pair exceeds the one of the monomer. For the tetrad, an O6 charge with a magnitude between the monomer and the base pair was obtained.

Electrostatic Potentials. It is well-known that metal ions or NH_4^+ are essential for the formation of G tetrad structures.[2−4,15] Also, the tetrads studied in this work do occur in tetraplexes formed in Na^+ or K^+ solutions. However, a detailed understanding of the interaction between metal ions, bases, and solvent is still missing. Binding sites for metal ions characterized by an unusually negative electrostatic potential have been predicted in RNA structures by Poisson−Boltzmann calculations.[53] In small model systems, quantum-chemical calculations may be used to derive qualitative conclusions. Such calculations reveal indeed a negative electrostatic potential at the central cavity of G tetrads.

We have calculated the molecular electrostatic potential for the C tetrads retaining the experimental geometry with N4−H42···O2 H-bonds found in the tetraplex structure with the PDB code 1EVO. It turns out that the C tetrad indeed possesses a large negative potential (Figure 5), but in contrast to G tetrads, it is not focused on a small region in the central cavity. We have recently shown that in the G tetrad complex with K^+ a structure with the ion located above the tetrad plane is more stable than the structure with the metal ion in the plane.[15] Because the C tetrad cavity size is larger than the one of the G tetrads, it is likely that in C tetrad/metal ion complexes larger ions than sodium may be located in the tetrad plane. Even though we have started geometry optimization with the experimental geometry, the final optimized structures have a different H-bond pattern (N4−H41···O2). This structure has no strongly electronegative potential in the tetrad center. If the interaction energy between the bases and the metal ion exceeds the one between the bases such as in G tetrads,[15] one might expect that metal ions induce a change of the H-bond pattern.

Planar U tetrads with N3−H3···O4 H-bonds have a negative electrostatic potential located at the central cavity,[15] while the electrostatic potential formed by the T tetrad is only moderately negative and spread over a large region between the bases (Figure 6). Therefore, the interaction of both tetrads with metal ions might be quite different. In addition to the weak interaction between the bases, the weak interaction between T and metal ions might also contribute to a preference of T loop formation relative to formation of four tetrads in d(GGGC-T4-GGGC) (1A8N).

The A tetrad with H62···N1 H-bonds (Figure 1) does not possess a central cavity of sufficient size for alkaline ions such as G tetrads. In the alternative structure with H61···N7 H-bonds, a cavity exists, but the regions with the most negative electrostatic potential are located outside close to the nitrogen atoms. Therefore, A tetrads with this type of structure cannot interact favorably with metal ions in the central cavity.

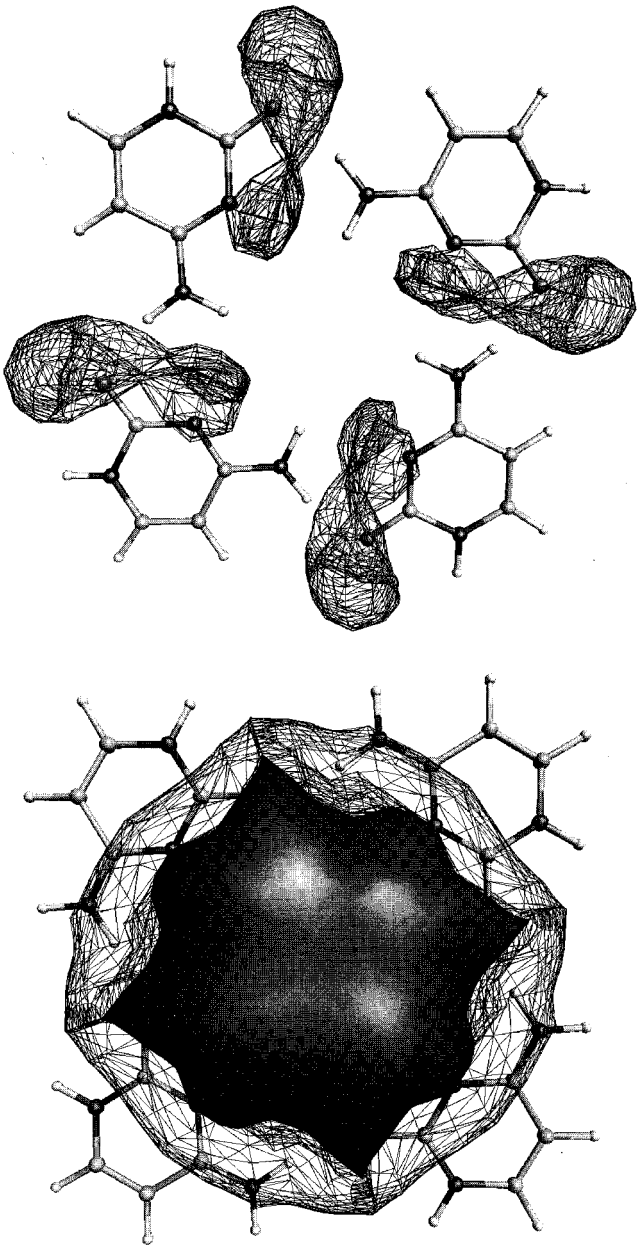

Figure 5. Electrostatic potential of the optimized C tetrad structure (a) and the biopolymer type structure occurring in the tetraplex structure 1EVO (b) at −50 (opaque surface) and −25 au (wireframe).

It should be noted, however, that in the presence of metal ions the tetrad conformation may be changed. In addition, different metal ions may induce different conformations. For example, in a tetraplex structure with GCGC tetrads, a conformational switch has been observed in passing from NaCl to KCl solution. In KCl solution, the non-Watson−Crick H-bonds between both GC pairs are removed and this leads to a much less planar tetrad (PDB code, 1A8W) as compared to the structure in NaCl (PDB code, 1A8N)[23] Interestingly, this conformational switch is also seen in an MD simulation on this structure.[29]

In addition to metal ion binding, calculations of the electrostatic potential are useful for the analysis of the hydrogen

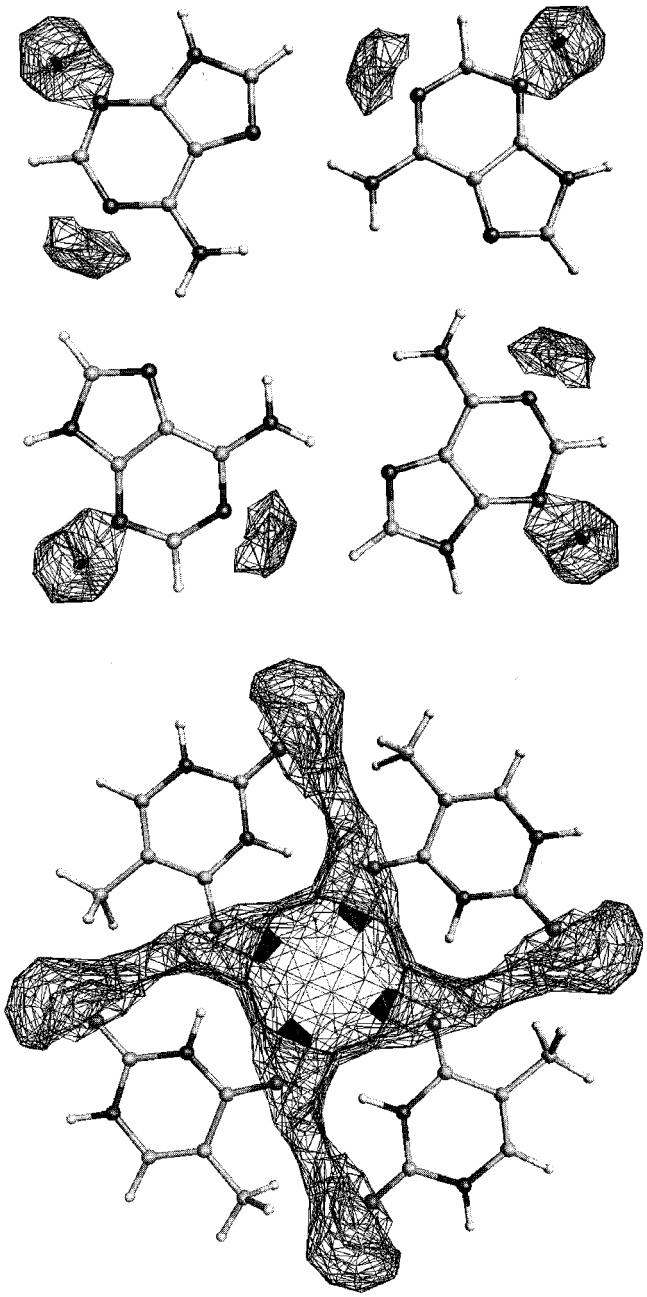

Figure 6. Electrostatic potential of the optimized C_{4h}-symmetric A tetrad (a) and T tetrad (b).

bonding pattern in tetrads. The C and T tetrad show a negative potential close to the acceptor atoms, whereas the weakly interacting A tetrad does not have this property (Figures 5 and 6). Similar to the G tetrad, there is a negative electrostatic potential outside of the tetrad close to N1 that might provide additional weak interaction sites for cations.

Performance of DFT and Force Field Calculations. DFT calculations have provided recently much information about the structure and interaction energies in nucleic acid bases.[13,25−28,30−32] The advantage of DFT calculations is primarily due to their efficiency as compared to the Møller−Plesset perturbation theory of second order (MP2), which has been used often only as single point calculation subsequently to Hartree−Fock geometry optimizations. Comparative calculations have recently shown a closer agreement of B3LYP with MP2 calculations than with Hartree−Fock calculations.[31] For the GC base pair, the DZVP and TZVP interaction energies of −23.99 and −25.32 kcal/mol (Table 4) are close to −23.8 kcal/mol from MP2/6-31G(0.25)//HF/6-31G(d,p) single point calculations.[24] A similar agreement has been found for the AF pair with the T analogue difluorotoluene,[33] water-mediated base pairs,[31] and complexes with metal ions.[34] BLYP calculations lead to somewhat smaller interaction energies than B3LYP calculations.[15]

For the basis sets, we note that there is an excellent agreement for the calculated interaction energies between TZVP (Table 3) and the much smaller DZVP basis (Supporting Information). For the H-bond geometries, TZVP provides somewhat shorter distances between the hydrogen and the acceptor atom for the C tetrad, whereas they are longer for the T tetrad. In general, DZVP appears to be an efficient alternative for the larger TZVP and other established basis sets. This makes the investigation of larger nucleic acid complexes feasible.

As the size of biopolymers is too large for DFT methods, force field methods that are able to handle these systems are required. As expected, MMFF94 can reproduce the pyramidal structure of amino nitrogen atoms in the monomers. The inversion barriers, however, are somewhat larger than the DFT results. Therefore, MMFF94 calculations lead to larger deviations from planarity as compared to DFT, for example, in the G tetrad.[15]

For the GCGC, C, and T tetrads, the force field calculations yield in a qualitative sense similar structures as the DFT calculations (Table 1). The differences in H-bond lengths do not usually exceed 0.15 Å, and the corresponding differences in H-bond angles are smaller than 15°. A substantial difference between force field and DFT calculations among these tetrads is found for the T tetrad. Here, the force field structures are much more nonplanar (ω, 97 and 94° for MMFF and AMBER, respectively) than the DFT tetrad geometry (ω, 139°). For the A tetrads, similar differences are found. However, AMBER force field calculations, and for C_4 symmetry also MMFF calculations, result in A tetrad structures with an extreme deviation from planarity no longer resembling the DFT geometries (Figure 4). The only A tetrad for which AMBER and DFT calculations result in a similar structure is the C_4-symmetric conformation with N6−H···N1 H-bonds (Figures 2 and 4). Contrary to the AMBER structures, the MMFF geometries remain close to the ones obtained from the DFT calculations for S_4 symmetry. However, MMFF predicts also that the structures obtained with AMBER are more stable, when they are used as input.

In agreement with the DFT calculations, MMFF94 predicts that the A tetrad N6−H62···N1 is less stable than the N6−H61···N7 structure, but the energy difference of 4.5 kcal/mol at C_{4h} symmetry is smaller than 8.95 kcal/mol from B3LYP/TZVP. The energy difference of 5.4 kcal/mol between the C_{4h}- and the S_4-symmetric T tetrad structures is in better agreement with the B3LYP result of 7.08 kcal/mol. Contradicting relative stabilities have been obtained for the C tetrad. B3LYP/TZVP predicts that the planar structure is 10.89 kcal/mol more stable than the S_4-symmetric conformation, whereas MMFF94 predicts the latter to be 8.3 kcal/mol more stable. AMBER energies have not been taken into account since the nonplanar structures are distorted as compared to the DFT structures. A comparison between the interaction energies from DFT calculations listed in Table 3 with the corresponding force field data given in Table 5 indicates that MMFF ranks them in the same order as B3LYP. There is a close correspondence of ΔE^T for the A tetrad H62···N1, the G tetrad with C_{4h} symmetry, and the GCGC tetrad. In contrast, ΔE^T is about 10 kcal/mol too high for the T and U tetrad with N3−H3···O4 H-bonds and too low for the C

tetrad, a complex having a high cooperative energy contribution. Good estimates of the deformation energies are provided for the C, G (C_{4h}), T, and A tetrad H61···N7; for the GCGC tetrad, ΔE^{def} is clearly underestimated for both components. The differences in the deformation energies are in the range from 0.0 to 2.4 kcal/mol, and the absolute total interaction energies differ between 1 and 12 kcal/mol. For the AMBER force field, the deformation energies deviate up to 2.3 kcal/mol from the B3LYP data. The total interaction energies differ between 3 and 15 kcal/mol. Similar to MMFF, the interaction energies are strongly overestimated for the T tetrad. For both force fields, relative deviations from the B3LYP interaction energies are large for tetrads with weak interaction energies and small for tetrads with strong interaction energies such as the G and GCGC tetrads. Previously, a deviation of 0.5−3.4 kcal/mol has been reported for the deformation energy and 2−8 kcal/mol for the total interaction energy of neutral trimers for the AMBER force field.[54] This force field also performed best in a systematic study of a series of hydrogen-bonded and stacked base pairs.[55]

Conclusions

Cyclic tetrad structures resembling conformations occurring in experimental tetraplex structures show marked differences in their geometries and interaction energies. Whereas the C, U, and GCGC tetrads are planar, the G tetrad and the A tetrad with N6−H···N7 H-bonds are slightly nonplanar, and the T tetrad and especially the A tetrad N1 conformation show marked deviations from planarity. Most of the optimized tetrads have an H-bond pattern that is similar to the experimental structures. An exception is the C tetrad for which the optimized tetrad geometry does not agree with the experimental structure. The nonplanar tetrad structures may be partially planarized in helix centers.

The interaction energies (ΔE^{T}) also vary significantly between the various tetrads. With −11.09 and −13.67 kcal/mol, the planar and C_4-symmetric A tetrad conformation with N6−H62···N1 H-bonds has the smallest values and the most negative values below −60 kcal/mol are found for the GCGC and G tetrads.

A comparison of partial atomic charges determined by a Mulliken population analysis indicates that there is a substantial change of these charges for those atoms involved in H-bonding. In the planar C tetrad, a negative potential exists in the central cavity suggesting that this might be a binding site for cations such as in G tetrads and, perhaps, in U tetrads.

We found that from the MMFF force field derived Hartree−Fock and Møller−Plesset perturbation theory provides in most cases structures close to the DFT calculations. The same is true for the AMBER tetrad structures, except for the weakly bound A tetrads. In the latter case, the AMBER structures bear no resemblance to the DFT results. The interaction energies derived from the MMFF and AMBER force field agree reasonably for those tetrads having high interaction energies, whereas the relative deviations are much larger for weakly H-bonded tetrads. For a quantitative understanding of these large differences for weakly bounded complexes, high level quantum-chemical calculations are desirable, which is beyond the scope of our study. It should be noted that some tetrads have quite large nonadditive contributions to the interaction energies, which is not covered in standard force fields.

Supporting Information Available: Tetrad structures and interaction energies at B3LYP/DZVP level. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Saenger. W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(2) Williamson, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 703−730.
(3) Gilbert, D. E.; Feigon, J. *Curr. Opin. Struct. Biol.* **1999**, *9*, 305−314.
(4) Han, H.; Hurley, L. H. *Trends Pharmacol. Sci.* **2000**, *21*, 136−142.
(5) Zahler, A. M.; Wiliamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991**, *350*, 718−720.
(6) Simonsson, T.; Pecinka, P.; Kubista, M. *Nucleic Acids Res*. **1998**, *26*, 1167−1172.
(7) Gallego, J.; Chou, S.-H.; Reid, B. R. *J. Mol. Biol.* **1997**, *273*, 840−856.
(8) Kang, Z.; Zhang, X.; Ratcliff, R.; Moyzis, R.; Rich, A. *Nature* **1992**, *356*, 126−131.
(9) Smith, F. W.; Feigon, J. *Nature* **1992**, *356*, 164−168.
(10) Špačkovà, N.; Berger, I.; Šponer, J. *J. Am. Chem. Soc.* **1999**, *121*, 5519−5534.
(11) Chowdhury, S.; Bansal, M. *J. Biomol. Struct. Dyn*. **2000**, *18*, 11−28.
(12) Štefl, R.; Špačkovà, N.; Berger, I.; Šponer, J. *Biophys. J.* **2001**, *80*, 455−468.
(13) Gu, J.; Leszczynski, J.; Bansal, M. *Chem. Phys. Lett*. **1999**, *311*, 209−214.
(14) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 6308−6313.
(15) Meyer, M.; Steinke, T.; Brandl, M.; Sühnel, J. *J. Comput. Chem.* **2001**, *22*, 109−124.
(16) Meyer, M.; Brandl, J.; Sühnel, J. *J. Phys. Chem. A* **2001**, *105*, 8223−8225.
(17) Gu, J.; Leszczynski, J.; Bansal, M. *Chem. Phys. Lett*. **2001**, *335*, 465−474.
(18) Patel, P. K.; Koti, A. S. R.; Hosur, R. V. *Nucleic Acids Res*. **1999**, *27*, 3836−3843.
(19) Patel, P. K.; Bhavesh, N. S.; Hosur, R. V. *Biochem. Biophys. Res. Commun.* **2000**, *270*, 967−971.
(20) Patel, P. K.; Hosur, R. V. *Nucleic Acids Res.* **1999**, *27*, 2457−2464.
(21) Kettani, A.; Bouaziz, S.; Gorin, A.; Zhao, H.; Jones, R. A.; Patel, D. J. *J. Mol. Biol*. **1998**, *282*, 619−636.
(22) Escaja, J.; Pedroso, E.; Rico, M.; González, C. *J. Am. Chem. Soc.* **2000**, *122*, 12723−12742.
(23) Bouaziz, S.; Kettani, A.; Patel, D. J. *J. Mol. Biol*. **1998**, *282*, 637−652.
(24) Šponer, J.; Leszcynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117−135.
(25) Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, *99*, 3247−3276.
(26) Schuster, P.; Wolschann, P. *Chem. Mon.* **1999**, 947−960.
(27) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 1898−1904.
(28) Gu, J.; Lesczynski, J. *J. Phys. Chem. A* **2000**, *104*, 7353−7358.
(29) Špackova, N.; Berger, I.; Šponer, J. *J. Am. Chem. Soc.* **2001**, *123*, 3295−3307.
(30) Guerra, C. F.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *J. Am. Chem. Soc*. **2000**, *122*, 4117−4128.
(31) Brandl, M.; Meyer, M.; Sühnel, J. *J. Phys. Chem. A* **2000**, *104*, 11177−11187.
(32) Brandl, M.; Meyer, M.; Sühnel, J. *J. Am. Chem. Soc.* **1999**, *121*, 2605−2606.
(33) Meyer, M.; Sühnel, J. *J. Biomol. Struct. Dyn.* **1997**, *15*, 619−624.
(34) Šponer, J.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1999**, *17*, 61−77.
(35) Meyer, M. *THEOCHEM* **1997**, *417*, 163−168.
(36) Meyer, M. *Int. J. Quantum Chem*. **2000**, *76*, 724−732.
(37) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(38) Halgren, T. A. *J. Comput. Chem.* **1995**, *17*, 490−519.
(39) UNICHEM, Oxford Molecular Group, 1999.
(40) Becke, A. D. *J. Chem. Phys*. **1993**, *98*, 5648−5652.
(41) Lee, C.; Yang, G.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(42) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560−571.
(43) EMSL Basis set library.
(44) Boys, S. F.; Bernardi, F. *Mol. Phys*. **1970**, *19*, 553−577.
(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.;

Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A5; Gaussian, Inc.: Pittsburgh, PA, 1998.

(46) Sybyl 6.6, Tripos Associates, St. Louis.

(47) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER 5. University of California: San Francisco, 1997.

(48) InsightII, Accelrys Inc., San Diego.

(49) Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 730−748.

(50) Karpfen, A.; Choi, C. H.; Kertesz, M. *J. Phys. Chem. A* **1997**, *101*, 7426−7433.

(51) Kratochvil, M.; Šponer, J.; Hobza, P. *J. Am. Chem. Soc.* **2000**, *122*, 3495−3499.

(52) Burda, J. V.; Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* **1997**, *101*, 9670−9677.

(53) Chin, K.; Sharp, K. A.; Honig, B. *Nat. Struct. Biol.* **1999**, *6*, 1055−1061.

(54) Šponer, J.; Burda, J. V.; Meijzlík, P.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1997**, *14*, 613−628.

(55) Hobza, P.; Kabeláč, M.; Šponer, J.; Meijzlík, P.; Vondrásek, J. *J. Comput. Chem.* **1197**, *18*, 1136−1150.