Crystal Structure of d(GGCCAATTGG) Complexed with DAPI Reveals Novel Binding Mode[†,‡]

Dominique Vlieghe,[§] Jiri Sponer,[||] and Luc Van Meervelt*[,§]

Department of Chemistry, Katholieke Universiteit Leuven, Celestijnenlaan 200F, B-3001, Heverlee, Belgium, and J. Heyrovsky Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejskova 3, 182 23 Prague, Czech Republic

Received April 5, 1999; Revised Manuscript Received July 12, 1999

ABSTRACT: The single-crystal X-ray structure of the complex between the minor groove binder 4′,6-diamidino-2-phenylindole (DAPI) and d(GGCCAATTGG) reveals a novel way of off-centered binding, with an unique hydrogen bond between the minor groove binder and a CG base pair. Application of crystal engineering and cryocooling techniques helped to extend the resolution to 1.9 Å, resulting in an unambiguous determination of drug conformation and orientation. The structure was refined to completion using SHELXL-93, resulting in a residual factor R of 18.0% for 3562 reflections with $F_o > 4\sigma(F_o)$ including 81 water molecules. As the bulky NH_2-group on guanine is believed to prevent drug binding in the minor groove, the nature and stability of the CG−DAPI contact was further addressed in full detail using ab initio quantum chemical methods. The amino groups involved in the guanine−drug interaction are substantially nonplanar, resulting in an energy gain of about 5 kcal/mol. The combined structural and theoretical data suggest that the guanine NH_2-group does not destabilize the drug binding to an extent that it prevents complexation.

Since it first was used for its antitrypanosomal activity (*1*), the synthetic antibiotic 4′,6-diamidino-2-phenylindole (DAPI)[1] has become one of the important DNA agents because of its numerous applications as antibiotic, antiviral, and anticancer drug. Binding of DAPI is known to alter various biomolecular processes such as replication, repair, transcription, and topology, because it inhibits binding of the necessary enzymes to the DNA double helix (*2*−*5*). It is also used as a fluorescent dye in biochemical and cytochemical studies (*6*, *7*).

The structure of DAPI (see Figure 1) resembles the classical minor groove binders such as netropsin and distamycin because of its crescent shape and two positively charged end groups, but the DAPI molecule is shorter. This leads to the general assumption that DAPI also binds in the minor groove of AT-rich DNA segments. Footprinting studies showed that DAPI appears to bind to AT-rich regions

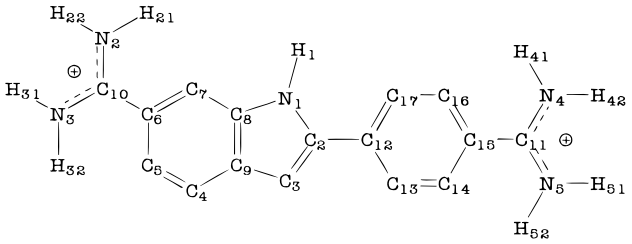

FIGURE 1: Structure and numbering scheme of 4′,6-diamidino-2-phenylindole (DAPI). Only hydrogens attached to nitrogens are shown.

containing at least four such base pairs (*8*). The crystal structure determination of the d(CGCGAATTCGCG)−DAPI complex (12-DAPI) (*9*) and the solution structure of the d(GCGATCGC)−DAPI complex (*10*) confirmed this mode of binding. In both structure determinations, the indole N1 nitrogen is believed to hydrogen bond to the two O2 atoms of the thymines of the central AT base pairs. Modeling studies also established that 2:1 drug−DNA complexes are possible for the DAPI molecule (*11*).

Minor groove sequence selectivity for AT-rich sequences can have several reasons. First, it was thought to be a direct consequence of base pair sequence: CG base pairs have an additional bulky NH_2-group at the floor of the minor groove which can prevent binding and sequence selective hydrogen bonding. In recent years, however, indirect sequence alterations are held responsible for AT base pair binding. The minor groove width, for example, influences the extent of van der Waals' interactions between the drug and the floor and walls of the minor groove (*12*). Electrostatic interactions between the negatively charged minor grooves and the minor groove binders having positively charged end groups also

[†] This work was partly supported by the Human Capital and Mobility Program of the European Community, by the Vlaams Instituut voor de Bevordering van het Wetenschappelijk-Technologisch Onderzoek in de Industrie (IWT), by the Fund for Scientific Research (Flanders) and by Grant 203/97/0029 GA CR.

[‡] Atomic coordinates and structure factors have been deposited in the NDB (entry code DD0002) and PDB (entry code 432D).

* To whom correspondence should be addressed. Phone: (32)16 327609. Fax: (32)16 327990. E-mail: Luc.VanMeervelt@chem.kuleuven.ac.be.

[§] Katholieke Universiteit Leuven.

[||] Academy of Sciences of the Czech Republic.

[1] Abbreviations: 10-DAPI, current crystal structure of the d(GGCCAATTGG)−DAPI complex; 12-DAPI, crystal structure of the d(CGCGAATTCGCG)−DAPI complex (9); CSD, Cambridge Structural Database; DAPI, 4′,6-diamidino-2-phenylindole; HF, Hartree−Fock; MP2, second-order Møller−Plesset perturbational theory; NMR, nuclear magnetic resonance.

10.1021/bi9907882 CCC: $18.00
Published on Web 11/18/1999

play a key factor in complex formation: AT sequences have a more negative minor groove which can explain the sequence selectivity (*13*).

Depending on the sequence, other methods of interaction were also observed. Using footprinting results, it was established that complexes of DAPI with regions containing GC or mixed GC and AT base pairs display a classical intercalator behavior (*14*). DAPI intercalates at AU sites of RNA (*15*). Depending on the position of CG base pairs at the center or the extremities of a DNA sequence, two different intercalation modes were observed in an NMR structure (*16*). Recently, simultaneous intercalation and minor groove binding was assessed in a solution structure of DAPI with d(CGATCG) (*17*). It is clear that, to date, no conclusion can be drawn concerning the nature of the DNA−DAPI interaction.

Crystal engineering techniques can be used to mimic triple helical fragments in the crystal lattice of d(GGCCAATTGG) and at the same time to improve the resolution of the obtained diffraction data (*18*, *19*). The resolution of most DNA-minor groove binder complexes is limited to 2.2−2.5 Å, with one important exception of a DNA-Hoechst 33342 complex diffracting to 1.4 Å resolution (*20*). We now have determined the crystal structure of the d(GGCCAATTGG)−DAPI complex and found that DAPI binds in the minor groove. Compared to the expected central position in the AATT segment, DAPI is shifted one base pair away from the center, forming a hydrogen bond to the neighboring GC base pair. This novel interaction is also characterized by ab initio methods.

EXPERIMENTAL PROCEDURES

Crystallization and Data Collection. The DNA decamer d(GGCCAATTGG) was purchased from Oswel DNA service (University of Southampton, U.K.), DAPI from Sigma-Aldrich (Bornem, Belgium). Crystals were grown using the hanging drop method from a typical solution of 43 mM $MgCl_2$, 3% MPD, 2 mM spermine, 0.4 mM ssDNA, and 0.2 mM DAPI (pH 6.9) against a 50% MPD stock solution. Crystals of dimensions 0.5 × 0.1 × 0.02 mm grow in 1 day.

Intensity data were collected at 120 K on a 90 mm MAR imaging plate detector at beamline 5.2 of the synchrotron Elettra at Trieste (λ = 1.000 Å) over a 105° φ range and increment of 3° using cryocooling techniques with a crystal-to-detector distance of 120 mm. The crystal did not show any sign of decay during the data collection. Data were processed using the DENZO/SCALEPACK (*21*) suite of programs. Data collection statistics are given in Table 1.

Structure Solution and Refinement. Unit cell parameters and space group indicated isomorphism with the native decamer structure, which was used as a starting model (NDB entry BD0006, except solvent molecules) for further refinement on F^2 using SHELXL-93 (*22*). The nucleotides of strand 1 are labeled G1−G10 in the 5′ to 3′ direction and G11−G20 on strand 2. After determination of the weighting factor and positional adjustment of parts of the DNA structure, the *R*-factor dropped to 27.9% [$F_o > 4\sigma(F_o)$]. Prior to DAPI addition, 20 water molecules were added, but not in the minor groove region. At this stage of the refinement, the DAPI molecule was added as indicated by the ($F_o - F_c$) Fourier difference map (see Figure 2). At first, the molecule was kept planar in accordance with the conformation observed in the 12-DAPI crystal structure (*9*). Inspection of the ($2F_o - F_c$) electron density map clearly showed that this was not the case. From this point on, the DAPI molecule was considered to consist of four planar groups which are allowed to rotate relative to each other: two positively charged amidinium end groups, a phenyl group and an indole group. A search in the Cambridge Structural Database (version 5.15, April 1998) (*23*) for the DAPI substructures in Figure 3 clearly revealed that the groups could indeed be twisted relative to each other. Fragment 1 was found 15 times in the CSD, but only 12 structures were retained because the others showed conformational changes related to inter- and intramolecular interactions. The absolute value of the N−C−C−C torsion angle ranges 12−44°. Fragment 2 appeared 28 times in the CSD, again three structures were not included in the parameter calculations (e.g., because the fragment is part of a full ring system). The absolute value of the N−C−C−C torsion angle ranges again from 24 to 89°.

Table 1: Data Collection Statistics of the d(GGCCAATTGG)−DAPI Complex

space group	$P2_12_12_1$
unit cell (Å)	a = 25.616, b = 36.565, c = 52.961
total number of measured reflections	24 368
no. of independent reflections	4115
resolution (Å)	1.9
multiplicity to 1.9 Å	5.9
χ^2 to 1.9 Å	1.137
χ^2 in 1.93−1.9 Å shell	1.145
R_{symm} to 1.9 Å (%)	7.3
R_{symm} in 1.93−1.90 Å shell (%)	29.1
completeness to 1.9 Å (%)	97.1
completeness in 1.93−1.90 Å shell (%)	95.1
mean $I/\sigma(I)$	13.8
reflections with $I > 3\sigma(I)$ (%)	78.6
reflections with $I > 3\sigma(I)$ in 1.93−1.90 Å shell (%)	54.2

Analysis of these fragments showed the need for a new dictionary of bond distances and angles (or 1,3-distances) for DAPI. On the basis of the CSD fragments, a new set of parameters was derived (Table 2) and used for the subsequent refinement. The main differences between these parameters and previously used ones are that (a) the new parameters reflect more the resonance in the aromatic rings and (b) bond lengths C6−C10 and C11−C15 exhibit more single bond character, and hence both amidinium groups should also be allowed to rotate.

After DAPI addition, water molecules were gradually added during several further refinement cycles. During the conjugate-gradient refinement 1,2- and 1,3-distance, planarity, antibumping, and chiral volume restraints were applied, but no torsion angle or hydrogen bond restraints. At the end, 81 water molecules were found, with *B*-factors ranging 17−59 Å^2, with an average of 39 Å^2. Water molecules show good hydrogen-bonding geometry to neighboring molecules. Final refinement parameters

$$R = \Sigma\, ||F_o| - |F_c||/\Sigma|F_o|$$

$$wR_2 = \{\Sigma[w(F_o^2 - F_c^2)^2]/\Sigma[w(F_o^2)^2]\}^{1/2}$$

for this structure are R = 18.0% [$F_o > 4\sigma(F_o)$, 3562

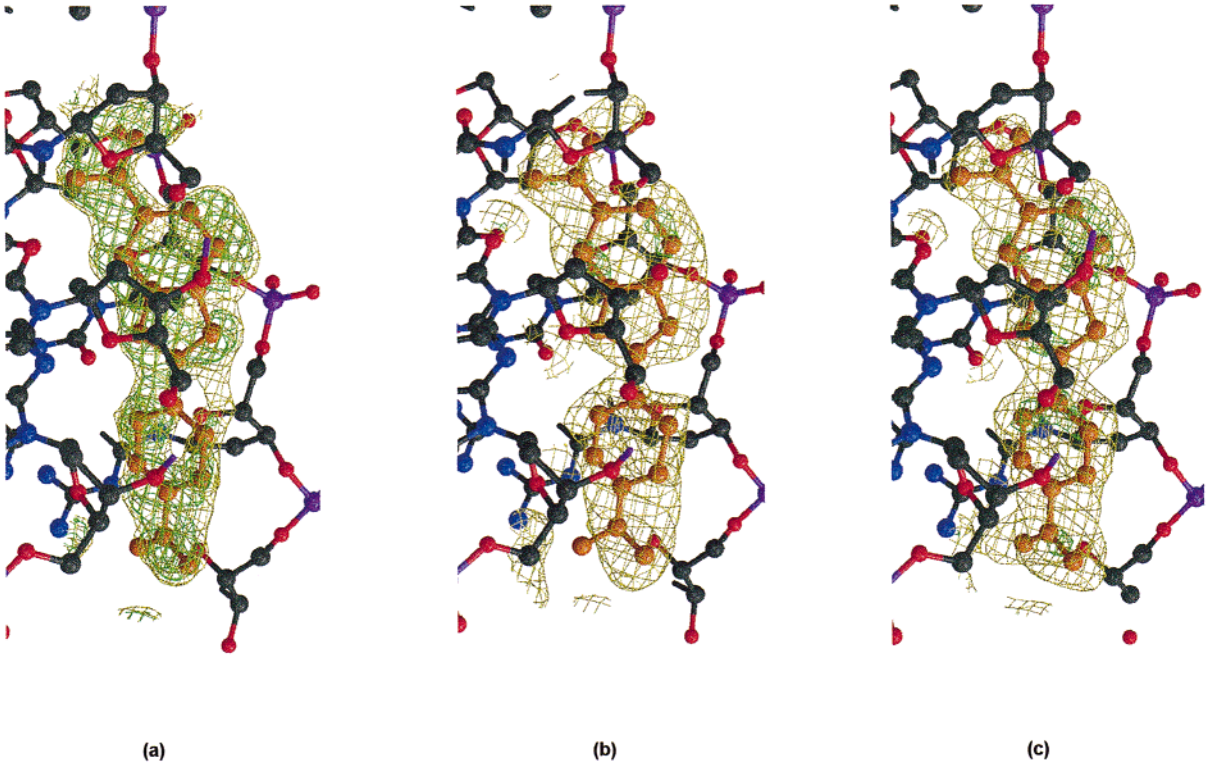

FIGURE 2: Electron density map contoured in the minor groove of the crystal structure of the d(GGCCAATTGG)−DAPI complex: (a) ($F_o - F_c$) difference map prior to DAPI addition contoured at 1.5 and 2.0σ, (b) ($2F_o - F_c$) density map after a first cycle of refinement with added drug (contoured at 1.5 and 3.0σ), (c) final ($2F_o - F_c$) electron density map. The final DAPI structure is shown in all three maps. The figure was prepared with Bobscript (*42*) and Raster3D (*43*).

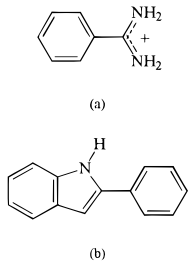

FIGURE 3: Model fragments for the DAPI molecule used in a CSD search. (a) Fragment 1, (b) fragment 2.

reflections], $R = 19.2\%$ (all 4106 reflections), $wR_2 = 47.1\%$, resolution limits 10−1.9 Å. The final rms deviation on bond lengths and angles is 0.013 and 0.029, respectively.

Quantum Chemical Calculations. Quantum chemical calculations have been done within the Hartree−Fock (HF) approximation or with inclusion of electron correlation effects using the second-order Møller−Plesset perturbational theory (MP2). Geometries of amino groups were investigated using the 6-31G basis set augmented by a standard d-shell on the amino group nitrogen atoms only [designated as 6-31G(NH_2*)] to make the amino group flexibility closer to the actual values (*24*, *25*). Interaction energies have been evaluated using the 6-31G basis set augmented by diffuse d-polarization functions to all second-row elements [exponents of 0.25, designated as 6-31G*(0.25)] to properly account for the dispersion attraction (*26*−*28*). We have also utilized standard 6-31G(d,p) and 6-311G(2d,p) basis sets. All calculational methods used in the present study have been extensively used in recent studies of base pairing (*26*, *29*), base stacking (*26*−*28*), close amino group contacts (*24*, *30*), bifurcated hydrogen bonds (*24*), and other interactions of bases. Reliability of predictions has been verified, for reviews see refs *31* and *32*. All quantum chemical calculations have been carried out using the Gaussian 94 program suite (*33*).

The structure of an isolated phenyl amidinium fragment of DAPI was optimized using several ab initio methods. Table 3 summarizes the optimal values of the dihedral angles between the phenyl and amidinium groups and the energy difference between the planar molecule and its global nonplanar minimum. The results are well converged, and the best MP2/6-311G(2d,p) estimate should be close to absolute predictions for this type of structure analysis. The calculations show convincingly that this segment of DAPI is intrinsically nonplanar and quite flexible. The nonplanarity in the crystal structures is reduced compared to the gas phase, as expected. The amino groups of the amidinium fragment show no pyramidalization, since the positive charge prevents a partial sp^3 pyramidalization of the nitrogen atoms.

RESULTS AND DISCUSSION

General Structural Features. In this complex (Figure 4), the decamer, which is designed to form a Watson−Crick base-paired octamer duplex structure with two overhanging bases at the 5′-end of both strands, has a conformation which closely resembles the native structure (*18*, *19*). The sugar−phosphate backbone of the G11−G12 overhang is a continuous extension of the octamer duplex backbone and forms a parallel triplex fragment with a neighboring duplex within the same column, while the G1−G2 overhang swings out to form an antiparallel triplex with a neighboring duplex of another column.

Table 2: Comparison between New DAPI Dictionary Values for 1,2- and 1,3-Distances Based on the Cambridge Structural Database and Those Used for the Refinement of the d(CGCGAATTCGCG)−DAPI Structure (12-DAPI) (*9*)[a]

1,2 distances (Å)				1,3 distances (Å)			
distance	12-DAPI	mean CSD distance	σ^b	distance	12-DAPI	mean CSD distance	σ^b
N1−C2	1.41	1.385	0.014	N2−N3	2.26	2.275	0.014
C2−C3	1.30	1.381	0.013	N2−C6	2.33	2.407	0.011
C3−C9	1.38	1.438	0.013	N3−C6	2.33	2.411	0.011
C4−C5	1.40	1.383	0.018	C10−C7	2.36	2.485	0.010
C5−C6	1.39	1.396	0.014	C10−C5	2.39	2.481	0.016
C6−C7	1.39	1.382	0.017	C6−C4	2.41	2.419	0.024
C7−C8	1.42	1.400	0.013	C5−C9	2.44	2.404	0.018
C8−C9	1.36	1.405	0.011	C4−C8	2.37	2.416	0.016
C8−N1	1.36	1.376	0.009	C9−C7	2.42	2.465	0.020
C4−C9	1.39	1.407	0.013	C8−C6	2.43	2.366	0.017
C6−C10	1.35	1.472	0.010	C7−C5	2.41	2.428	0.027
C10−N2	1.32	1.310	0.011	C7−N1	2.52	2.503	0.012
C10−N3	1.33	1.312	0.010	C4−C3	2.55	2.627	0.018
C2−C12	1.48	1.478	0.009	C8−C2	2.21	2.249	0.009
C12−C13	1.37	1.392	0.010	C8−C3	2.20	2.281	0.020
C13−C14	1.40	1.389	0.015	C9−C2	2.19	2.269	0.014
C14−C15	1.43	1.379	0.022	C9−N1	2.21	2.253	0.011
C15−C16	1.36	1.372	0.014	C3−N1	2.21	2.248	0.017
C16−C17	1.38	1.392	0.018	C3−C12	2.55	2.605	0.024
C17−C12	1.42	1.392	0.014	N1−C12	2.49	2.477	0.016
C15−C11	1.34	1.472	0.010	C2−C13	2.52	2.493	0.018
C11−N4	1.32	1.310	0.011	C2−C17	2.45	2.494	0.023
C11−N5	1.35	1.312	0.010	C12−C14	2.38	2.417	0.020
				C13−C15	2.46	2.396	0.022
				C14−C16	2.39	2.387	0.027
				C15−C17	2.39	2.397	0.024
				C16−C12	2.42	2.414	0.021
				C17−C13	2.41	2.396	0.025
				C14−C11	2.39	2.485	0.010
				C16−C11	2.37	2.481	0.016
				C15−N4	2.40	2.407	0.011
				C15−N5	2.32	2.411	0.011
				N4−N5	2.20	2.275	0.014

[a] In the calculation of σ, n is the number of distance observations, d_i the mean distance, and d the current distance.

[b]
$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(d_i - d)^2}{(n-1)}}$$

Table 3: Energy Difference between Fully Optimized and Planar Phenyl Amidinium Groups of DAPI and N−C−C−C Dihedral Angle

method/basis set	ΔE (kcal/mol)	N−C−C−C (deg)
HF/6-31G**	−3.39	139.8
MP2/6-31G**	−4.30	139.0
MP2/6-311G(2d,p)	−3.18	141.1

The global helical parameters of the octamer duplex closely resemble those obtained for the native structure: the helical twist is 34.5° (compared to 35.2°) and a rise for both of 3.4 Å, values in the range for B-DNA structures. A least-squares fit between the native 1.15 Å resolution structure and this structure shows an overall rms deviation of 0.65 Å. A conformational change is observed for the phosphate group of C13 and is correlated with the difference in resolution and not with the application of the flash cooling technique. Indeed, in the native structure at 2.0 Å resolution and room temperature this group is oriented as in the 10-DAPI structure. The differences in backbone torsion angles illustrate this clearly: in both the lower resolution structures (native and 10-DAPI) ϵ12, α13, and γ13 are in the normal trans, −sc, +sc range, while for the high-resolution structure they are in the −ac, +sc, trans. The relation between the obtained resolution and the conformation of phosphate group C13 is not clear. The change in conformation has no influence on the position of the G11−G12 overhanging residues involved in triplet formation. In all three structure determinations, there

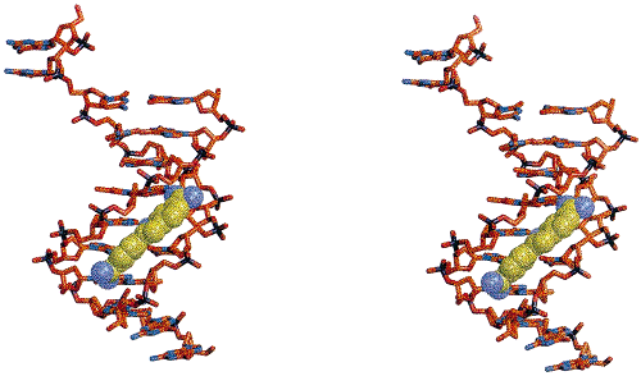

FIGURE 4: Stereoview of the d(GGCCAATTGG)−DAPI structure, with DAPI shown as CPK (prepared with Bobscript (*42*) and Raster3D (*43*)].

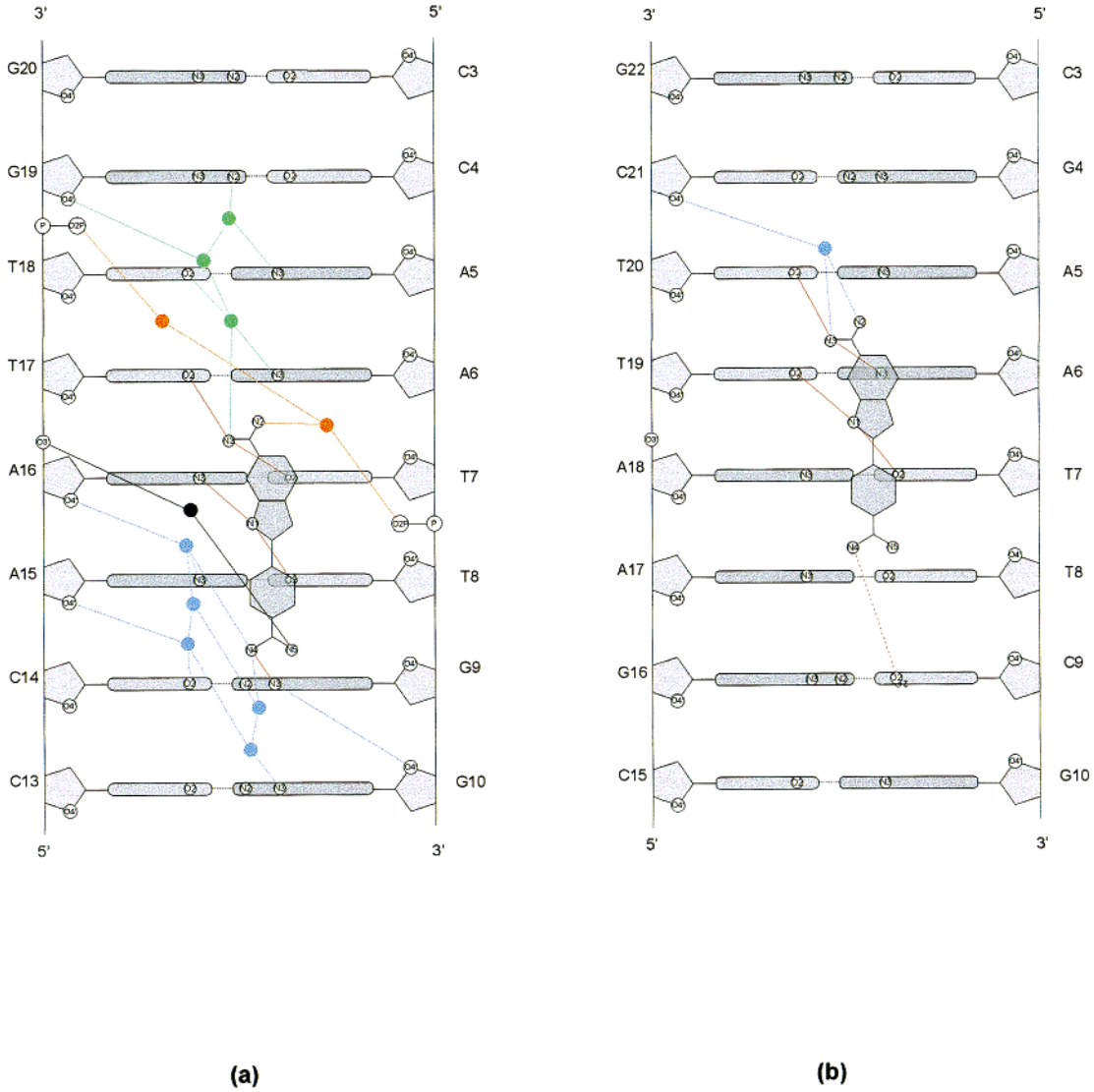

FIGURE 5: Schematic view of the DAPI position and hydration in the minor groove of (a) the d(GGCCAATTGG)−DAPI and (b) the d(CGCGAATTCGCG)−DAPI structures (*9*). For the 10-DAPI structure the following color scheme is applied: direct drug-base pair hydrogen bonds are shown in red, the short hydration spine is shown in green, the extended spine in orange, the double spine in blue and the DAPI-solvent-backbone interaction in black.

is a strict conservation of triplet conformations, resulting in similar helical parameters, hydrogen-bonding distances, and angles between base planes.

In the DAPI molecule, the final dihedral angle around C6−C10 is 21°, around C2−C12 it is 24°, and around C15−C11 it is 15°, in contrast with the dihedral angles in the 12-DAPI structure, which were constrained to be 0° (*9*). In the gas phase, the optimal dihedral angle around C2−C12 at the HF/6-31G** level is 36° (whole DAPI molecule optimized). The structure with this dihedral angle constrained to be 0° is 1.7 kcal/mol less favorable, suggesting that DAPI is nonplanar and flexible.

Position of DAPI Molecule in the Crystal Structure. The DAPI molecule is positioned in the minor groove of the central d(AATT) sequence, as previously observed in the 12-DAPI structure (*9*). There is, however, a difference in position of the DAPI molecule relative to the base pairs (Figure 5). The drug is shifted one base pair toward the 3′-end of the first strand. The central N1 atom of the drug forms bifurcated hydrogen bonds with O2(T8) and N3(A16), which catalogs the drug to belong to class I (*34*), while in 12-DAPI, this atom is hydrogen bonded to the two O2 atoms of the thymine bases of the central AT steps. In the 10-DAPI structure, these atoms form a hydrogen bond with N3 of the amidinium group positioned on the indole ring. The N4 of the terminal group on the phenyl ring is hydrogen bonded to N3(G9). In total, there are five direct drug−DNA hydrogen bonds in this structure. Compared to 12-DAPI, four base hydrogen bond acceptors are conserved (some belong to different residues) and one [O2(C9)] is replaced by N3(G9). Apart from these hydrogen bonds, there are several stabilizing van der Waals' contacts between the O4′ atoms of the DNA sugars and DAPI atoms (Table 4).

In the literature, there is a tendency to exclude GC base pairs for DAPI-binding sites, because the bulky NH_2-group on guanine is believed to prevent binding. Due to the partial double bond character of C15−C11, the terminal amidinium group is not completely free to rotate, so the N2(G9)⋯N4(DAPI) distance is rather small (3.14 Å) and can cause a destabilization of the binding of DAPI to DNA. In the 12-DAPI crystal structure, 27 contacts with a distance less than 3.8 Å are observed between the minor groove binder and

Table 4: Comparison of Close Contacts of Atoms in the Minor Groove between d(GGCCAATTGG) or d(CGCGAATTCGCG) and DAPI Atoms (*9*)[a]

	d(GGCCAATTGG) + DAPI (10-DAPI)			
C4		G19		0
A5		T18	O2, O4′, C4′, C5′	4
A6	C2	T17	O2, O4′, C4′, C5′	5
T7	O2, C1′	A16	C2, N3, O4′	5
T8	O2, O4′, C4′, C5′	A15		4
G9	N2, C2, N3, C1′, O4′, C4′, C5′	C14		7
G10	O4′	C13		1
	d(CGCGAATTCGCG) + DAPI (12-DAPI)			
G4		C21	O4′,C4′, C5′	3
A5		T20	O2, O4′, C4′, C5′	4
A6	C2, N3	T19	O2, C1′	4
T7	O2, C1′, O4′, C4′	A18	C2, N3, O4′	7
T8	O2, O4′, C4′	A17	C2	4
C9	O2, C1′, O4′, C4′, C5′	G16		5
G10		C15		0

[a] Distances smaller than 3.8 Å are considered to be close contacts.

the DNA duplex. In the current structure, 26 contacts were identified, indicating that the DAPI molecule is not pushed away from the floor of the minor groove (Table 4). However, the distribution of the close contacts is different: for 10-DAPI, more close contacts are situated at the 3′-end of the first strand (G9−C14), while for the 12-DAPI structure they are situated at the central base pairs. This means that the phenyl amidinium part is pushed somewhat away, which is compensated by a closer fitting of the indole amidinium part of the molecule. This is in agreement with the NMR structure of DAPI with d(GCGATCGC), which contains only a two base pair AT-track making close G-NH_2···DAPI contacts inevitable (*10*).

Quantum Chemical Analysis of the CG−DAPI Contact. The crystal structure shows a close contact between the guanine amino group and one of the amino groups of a DAPI amidinium group. Such interactions are usually considered as very repulsive steric clashes. However, B-DNA crystal structures contain a large amount of close amino group contacts (*24*, *30*, *35*) contradicting this view. It has been convincingly demonstrated that amino groups of nucleic acid bases are intrinsically nonplanar and very flexible (*25*, *31*, *32*, *36*). They can be involved in stabilizing interactions with out-of-plane orientation of the amino group hydrogen atoms and even in weak amino-acceptor interactions including amino-amino contacts (*24*, *30*, *31*).

The geometry and energetics of the guanine and DAPI amino groups in the contact area of our structure were investigated. The calculations have been done at the HF/6-31G(NH_2*) approximation in a similar way as used in previous papers for related interactions (*24*, *30*). The fragment studied involved the G9C14 base pair and the phenyl amidinium fragment of the drug (Figure 6 a) to which hydrogen atoms were added at calculated places. The intermolecular geometry of this complex was constrained according to the crystal structure using six parameters fixing position of the N−C−C(ring) segment of DAPI with respect to the C2−N3−C4 fragment of guanine. The N−C−C−C torsional angles of the drug were constrained as in the crystal, while the phenyl ring was kept planar. Other parameters (intramolecular geometries of monomers and the base pair) were relaxed in the course of the ab initio calculations, except the constraints imposed on selected amino groups during some calculations. First, we have optimized the structure while keeping the amino groups of DAPI and guanine planar.

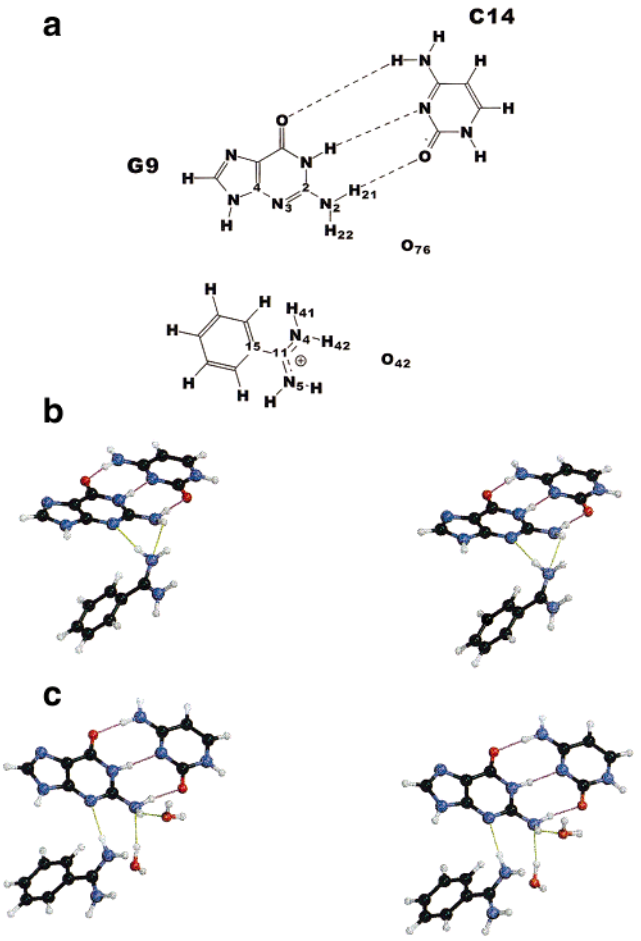

FIGURE 6: (a) Fragment used for quantum chemical analysis of the CG-DAPI contact showing atom labels used during discussion. (b) Optimized geometry of the GC···phenyl amidinium interaction by HF/6-31G(NH_2*) calculations without and (c) with two water molecules. The stereo drawings were prepared with Bobscript (*42*) and Raster3D (*43*).

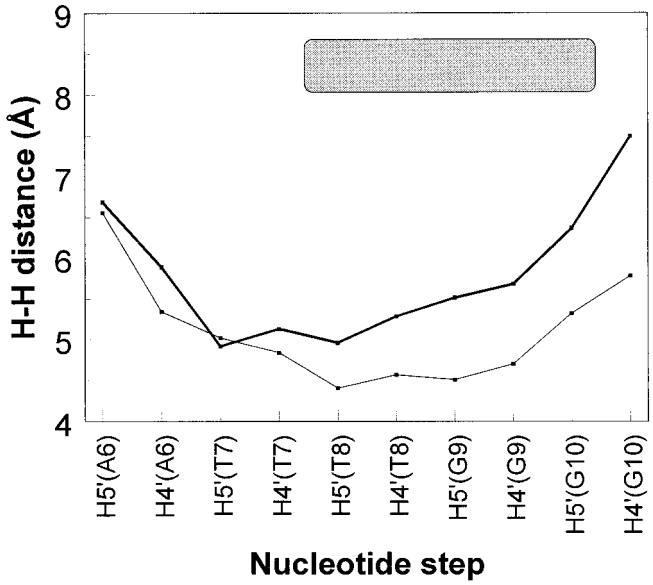

FIGURE 7: Minor groove width (Å) based on the H4′−H5′ distances (see ref *12*) for the d(GGCCAATTGG)−DAPI complex (bold) and the native decamer (light) (*18*). The position of DAPI is indicated by the gray rectangle.

Second, we have completely relaxed both amino groups of DAPI. The most striking feature of the latter structure is a marked pyramidalization of the amino group interacting with guanine. The H41−N4−C11−C15 torsion angle reaches a value of 25°, and hydrogen H41 is directed toward the N3 nitrogen atom of guanine. This is accompanied by a very significant energy improvement of 4.9 kcal/mol. Finally, we have relaxed the guanine amino group (Figure 6b). The DAPI amino group nitrogen atom remains oriented toward N3 (H41−N4−C11−C15 torsion angle of 22°) while the proximal guanine amino group hydrogen atom H22 is bent away from the contact area having a torsion angle with respect to the ring (H22−N2−C2−N3) of −27°. The other guanine amino hydrogen atom (H21) does not move much due to its involvement in the base pairing. Surprisingly, the substantial geometrical relaxation of the guanine amino group improves the energy only by 0.1 kcal/mol, the structure being now by 5.0 kcal/mol more stable than the structure with all amino groups planar. To get further insight into the role of the H22 hydrogen atom of guanine, we have reoptimized the structure once again, keeping the H22−N2−C2−N3 dihedral angle at a large value of −50°. This constraint deteriorates the energy of the system by 0.3 kcal/mol only. Therefore, the results can be summarized in the following way. The DAPI−guanine contact is characterized by a strong amidinium···N3 hydrogen bond, and the nonplanarity of the DAPI amino group is essential. Despite the short distance between the amidinium and guanine amino groups, there is no clear interaction comparable to previously observed close amino group contacts (*24*, *30*, *31*), since this interaction is not competitive with the amidinium···N3 hydrogen bond. Nevertheless, the guanine amino group is greatly influenced by the contact: it can reach extreme degrees of nonplanarity without any energy penalty and its outer (unpaired) hydrogen is predicted to be exceptionally flexible. In the course of previous ab initio studies, large flexibilities of amino groups of bases have been observed in some interactions (*24*, *31*, *32*). However, we have never seen an amino group hydrogen atom with such a degree of directional flexibility as predicted here for guanine. The actual geometry of the amino group will be very sensitive to the distribution of other chemical groups around the contact area.

There are two water molecules close to the guanine amino group in the crystal. Therefore, we have repeated the calculations by including these two water molecules; the positions of their oxygen atoms were constrained according to the crystal data, and the hydrogen atoms were fully relaxed. The final structure (Figure 6c) additionally shows two clear hydrogen-bonding contributions. The H22 atom is not coplanar with the attached base (torsion angle H22−N2−C2−N3−46°) and serves as a hydrogen bond donor to water molecule O76 located above the guanine plane. This allows the development of an accessible negative lone electron pair region below the amino group nitrogen atom which acts as hydrogen bond acceptor for the second water molecule O42 at the other side of the guanine plane. Imposing planarity on the guanine amino group results in a significant destabilization of 3.1 kcal/mol with respect to the relaxed amino group. Thus, we suggest that the two water molecules observed in the crystal near the contact area are actively involved in the stabilization of the local structure and are oriented in a concerted manner together with the guanine amino group.

The above calculations analyze the amino group geometries in the DAPI−guanine contact area. Besides that, we have also calculated [MP2/6-31G*(0.25) method] absolute values of the interaction energies between guanine and the phenyl amidinium fragment in the crystal geometry. There is a significant attraction of −8.7 kcal/mol between DAPI and guanine. This is comparable to base stacking and hydrogen-bonding energies (*27*−*29*). To estimate the net effect of the guanine amino group, we have replaced guanine by 6-oxopurine (applying the same optimization procedure as above), improving the interaction energy moderately to −10.3 kcal/mol. This is in agreement with the crystal data and indicates that the amino group is indeed destabilizing for the drug binding, but does not represent any ultimate obstacle for the binding. Interestingly, adenine has a better binding energy of −17.7 kcal/mol in this position. The difference between adenine and guanine is due to a more favorable electrostatic interaction (i.e., magnitude and orientation of molecular dipoles, see, e.g., ref *29*, Figure 1) of adenine with DAPI, compared to guanine and 6-oxopurine [guanine and 6-oxopurine have very similar electrostatic potentials (*29*)]. The N2 amino group is not so important.

Comparison with Other Minor Groove Binding Drugs. In this sequence, DAPI has the freedom to target a 6 base pair sequence with a central d(AATT) tract; however, it shifts its position by one base pair, changing the base specificity from AAT in the dodecamer to ATTG in the decamer. This is in contrast with all DAPI−DNA complexes so far studied, whether they were examined using X-ray, NMR, or modeling techniques. Footprinting results indicate that G(C) nucleotides are almost never protected by binding of a single DAPI molecule (*8*). It is also in conflict with the general assumption that minor groove binders specifically bind to AT-tracts. Netropsin, which is also known to target for the minor groove, shows interaction with CG base pairs, because it is too long to fit in the AATT sequence and needs at least five base pairs (*37*). Other shorter minor groove binders with the same end groups as DAPI such as pentamidine (*38*), γ-oxapentamidine (*39*), and propamidine (*40*), also bind in the central AATT tract. The question arises why DAPI is not centrally positioned. There are two main differences between the 10-DAPI and the 12-DAPI structure: crystal packing and DNA sequence. Packing differences bring the backbone of two symmetry-related molecules in 10-DAPI in relatively close contact with the minor groove of the asymmetric unit. The more exposed phenyl amidinium end group of DAPI is in relatively close electrostatic contact (<10 Å) with six phosphate oxygens of symmetry-related molecules. In the 12-DAPI structure, there is only one such contact. At this point, one should remember that the packing in 10-DAPI is controlled by the triple helix formation, and definitely not by the presence of the drug molecule. A second difference is the sequence. In the 10-DAPI structure, the GC base pair in 12-DAPI is replaced by a C·G pair at both sides of the d(AATT) tract, breaking down the 5′-$(pu)_3(py)_3$-3′ sequence. This brings, together with the intramolecular DAPI twisting, the phenyl amidinium group in hydrogen-bonding position with N3 of G9, a contact, which is not a priori destabilizing as shown by our theoretical calculations. In the crystal structure of the d(CGCAATTGCG)−netropsin com-

plex, a similar hydrogen bond is observed (*41*). It is thus tentative to hold the sequence responsible for the DAPI shift.

Influence of the DAPI on DNA Conformation. Addition of a minor groove-binding molecule is believed to open the minor groove, which is also confirmed by this structure. Compared to the native structure, a widening of about 1.5 Å is observed, as shown in Figure 7. A somewhat different trend can be seen when defining the minor groove width by different atoms. However, the minor groove width based on C4′ atoms reflects best the groove width at the position of the DAPI molecule. The groove width is symmetric for the native structure. On addition of the drug, a widening is observed for the whole groove, but the part at the 3′-end of the first strand is opened more as a consequence of drug addition. Helical parameters and backbone torsion angles are comparable with the values observed for the native decamer, and no large systematic changes are observed to explain the minor groove widening.

Hydration. In contrast to the 12-DAPI crystal structure, where at least parts of the minor groove are occupied by symmetry-related molecules, an extensive minor groove hydration at both sides of the drug is observed in 10-DAPI (Figure 5). Binding of DAPI in the minor groove displaces the typical spine of hydration of the d(AATT) sequence and typical O2(T)/N3(A)···O(water) hydrogen bonds are replaced by O2(T)/N3(A)···N1/3/4(DAPI) contacts, mimicking thus the spine of hydration. Because of the relative shortness of the groove binder and the shift in the minor groove toward the 3′-end of the first strand, a small part of the spine consisting of three water molecules remains as shown in green in Figure 5. This short spine is prolonged by DAPI via a hydrogen bond with N3. The second amino group (N2) of this end group is hydrogen bonded to a water molecule, which is further connected through several solvent molecules with various hydrogen bond acceptors and donors of the DNA molecule and a symmetry equivalent one. It is also part of a direct two-water interstrand connection of phosphate groups T8 and G19, forming one branch of the so-called extended spine of hydration (shown in orange in Figure 5) before only observed in the high-resolution structure of the native decamer (*19*). The N4/N5 end group at the phenyl group forms three direct hydrogen bonds with solvent molecules, of which one is further connected with the phosphate backbone (shown in black in Figure 5). Two water molecules are part of a short double spine of hydration observed at the C·G base pairs at the 3′-end of the first strand of the molecule (shown in blue in Figure 5), which makes various direct contacts with DNA bases and sugars. A double hydration spine can thus stabilize binding of a drug in the wider minor groove of a GC base pair.

In contrast to the native decamer, no systematic pattern of hydration for the major groove can be observed. This is due to the relative low resolution of this structure determination. Only a small number of first hydration shell waters is present. In addition to the water-mediated interstrand contact analogous to the extended hydration spine, some intrastrand [e.g., O2P(T18)···W85···W61···O1P(G19)] and numerous interstrand contacts between symmetry equivalent duplexes are observed. There are also two water-mediated contacts between backbone oxygens and the DAPI molecule [O3′(T17)···W56···N5(DAPI) and O2P(C14)···W72···N4/N5(DAPI of a symmetry equivalent molecule)]. One direct duplex−duplex contact is also present: between O1P(C14) and O3′(G20) of a symmetry equivalent duplex. Most of the interduplex contacts are a consequence of triplex formation, which brings not only the bases but also the sugar−phosphate backbones in close proximity.

Conclusions. The interaction of DAPI with DNA covers a broad spectrum of complex formation. The current crystal structure adds a novel minor groove binding mode for this drug, with mixed AT/GC sequence selectivity. The driving force for this is unclear: differences in sequence, packing, hydration, and conformation can all be responsible. Our results emphasize the need for more structural details of minor groove binders that form complexes with different DNA sequences. To date, about 80% of the DNA sequences in minor groove complexes is directly related to the Dickerson−Drew dodecamer d(CGCpuXXXXpyGCG) (with XXXX = AATT, ATAT or TTAA). Changing the central hexamer sequence from GAATTC to CAATTG already has a major influence on the position of the drug in the minor groove, albeit direct or indirect, and makes a direct hydrogen bond between a GC base pair and one of the rigid end groups of DAPI possible, despite the bulky guanine NH_2-group. This is probably a general feature and was also established in the netropsin-d(CGCAATTGCG) structure, the only other structure so far with a central CAATTG sequence (*41*).

A strong interaction between N3 of G and one of the amidinium end groups of the drug is observed. This interaction is characterized by the nonplanar behavior of the amino groups involved, especially the amino group of the amidinium part directly involved in hydrogen bonding. Further stabilization of the local structure is achieved by the addition of two solvent molecules. Analysis of absolute interaction energies, obtained from quantum chemical calculations of DAPI−GC interactions clearly shows that the presence of the G-NH_2 does not destabilize the drug binding to an extent that it prevents interaction. Although a small destabilization is observed compared to the DAPI−AT interaction, the overall interaction energy still favors complexation.

The 1.9 Å resolution structure determination of the complex of DAPI with d(GGCCAATTGG) was a first possible application of the use of triplex formation as an additional crystal stabilization and hence resolution enhancing tool. Together with the application of cryocooling techniques and synchrotron radiation, the resolution could be improved by 0.5 Å [compared to 12-DAPI (*9*)]. It is hoped that, in the future, atomic resolution can be achieved for complexes of DNA and minor groove binders, which generally do not diffract further than about 2.2 Å. This could help in elucidating the correct or simultaneous existence of both orientations of the drug molecule in the groove, a problem that still exists for some low resolution structure determinations.

ACKNOWLEDGMENT

We thank the staff of the Elettra X-ray diffraction beamline at Trieste for assistance.

REFERENCES

1. Dann, O., Bergen, G., Demant, E. and Volz, G. (1971) *Justus Liebigs Ann. Chem. 749*, 68−89.

2. Straney, D. C., and Crothers, D. M. (1987) *Biochemistry 26*, 1987−1995.
3. Störl, K., Störl, J., Zimmer, Ch., and Lown, J. W. (1993) *FEBS Lett. 317*, 157−162.
4. Chiang, S.-Y., Welch, J., Rausher, F. J., III, and Beerman, T. A. (1994) *Biochemistry 33*, 7033−7040.
5. Welch, J. J., Rausher, F. J., III, and Beerman, T. A. (1994) *J. Biol. Chem. 269*, 31051−31058.
6. Kapuscinski, J., and Szer, W. (1979) *Nucleic Acids Res. 6*, 3519−3534.
7. Cavatorta, P., Masotti, L., and Szabo, A. G. (1985) *Biophys. Chem. 22*, 11−16.
8. Portugal, J., and Waring, M. J. (1988) *Biochim. Biophys. Acta 949*, 158−168.
9. Larsen, T. A., Goodsell, D. S., Cascio, D., Grzeskowiak, K., and Dickerson, R. E. (1989) *J. Biomol. Struct. Dyn. 7*, 477−491.
10. Trotta, E., D′Ambrosio, E., Del Grosso, N., Ravagnan, G., Cirilli, M., and Paci, M. (1993) *J. Biol. Chem. 268*, 3944−3951.
11. Mohan, S., and Yathindra, N. (1991) *J. Biomol. Struct. Dyn. 9*, 695−704.
12. Neidle, S. (1992) *FEBS Lett. 298*, 97−99.
13. Pullman, B. (1983) *J. Biomol. Struct. Dyn. 1*, 773−794.
14. Wilson, W. D., Tanious, F. A., Barton, H. J., Jones, R. L., Fox, K., Wydra, R. L., and Strekowski, L. (1990) *Biochemistry 29*, 8452−8461.
15. Tanious, F. A., Veal, J. M., Buczak, H., Ratmeyer, L. S., and Wilson, W. D. (1992) *Biochemistry 31*, 3103−3112.
16. Trotta, E., D'Ambrosio, E., Ravagnan, G., and Paci, M. (1995) *Nucleic Acids Res. 23*, 1333−1340.
17. Trotta, E., D'Ambrosio, E., Ravagnan, G., and Paci, M. (1996) *J. Biol. Chem. 271*, 27608−27614.
18. Vlieghe, D., Van Meervelt, L., Dautant, A., Gallois, B., Précigoux, G., and Kennard, O. (1996) *Science 273*, 1702−1705.
19. Vlieghe, D., Turkenburg, J., and Van Meervelt, L. (1999) *Acta Crystallogr., Sect. D 55*, 1495−1502.
20. Robinson, H., Gao, Y.-G., Bauer, C., Roberts, C., Switzer, C., and Wang, A. H.-J. (1998) *Biochemistry 37*, 10897−10905.
21. Otwinowski, Z. (1993) in *Proceedings of the CCP4 study weekend: data collection and processing, 29−30 January* (Sawyer, L., Isaacs, N., and Bailey, S., Eds.) pp 56−62, SERC Daresbury Laboratory, England.
22. Sheldrick, G. M. (1993) *SHELXL-93. Program for Crystal Structure Refinement*, Göttingen, Germany.
23. Allen, F. H., and Kennard, O. (1993) *Chem. Des. Aut. News 8*, 1 and 31−37.
24. Sponer, J., and Hobza, P. (1994) *J. Am. Chem. Soc. 116*, 709−714.
25. Bludsky, O., Sponer, J., Leszczynski, J., Spirko, V., and Hobza, P. (1996) *J. Chem. Phys. 105*. 11042−11050.
26. Hobza, P., Sponer, J., and Polasek, M. (1995) *J. Am. Chem. Soc. 117*, 792−798.
27. Sponer, J., Leszczynski, J., and Hobza, P. (1996) *J. Phys. Chem. 100*, 5590−5596.
28. Sponer, J., Gabb, H. A., Leszczynski, J., and Hobza, P. (1997) *Biophys. J*. 73, 76−87.
29. Sponer, J., Hobza, P., and Leszczynski, J. (1996) *J. Phys. Chem. 100*, 1965−1974.
30. Luisi, B., Orozco, M., Sponer, J., Lugue, F. J., and Shakked, Z. (1998) *J. Mol. Biol. 279*, 1123−1136.
31. Sponer, J., Leszczynski, J., and Hobza, P. (1996) *J. Biomol. Struct. Dyn. 14*, 117−136.
32. Sponer, J., and Hobza, P. (1998) in *Encyclopedia of Computational Chemistry* (Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaeger, H. F., III, and Schreiner, P. R., Eds.) pp 777−789, John Wiley & Sons, Chichester, U.K.
33. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. G., Robb, M. A., Cheeseman, J. R., Keith, T., Petersson, G. A., Montgomery, J. A., Raghavachari, K., Al-Laham, M. A., Zakrzewski, W. G., Ortiz, J. V., Foresman, J. B., Peng, C. Y., Ayala, P. Y., Chen, W., Wong, M. W., Andres, L. J., Replogle, E. S., Gomperts, R., Martin, R. L, Fox, D. J., Binkley, J. S., Defrees, D. J., Baker, J., Stewart, J. J. P., Head-Gordon, M., Gonzalez, C., and Pople, J. A. (1995) Gaussian, Inc., Pittsburgh, PA.
34. Goodsell, D. S., Kopka, M. L., and Dickerson, R. E. (1995) *Biochemistry 34*, 4983−4993.
35. Sponer, J., and Kypr, J. (1994) *Int. J. Biol. Macromol. 16*, 3−6.
36. Sponer, J., and Hobza, P. (1994) *J. Phys. Chem. 98*, 3161−3164.
37. Sriram, M., van der Marel, G. A., Roelen, H. L. P. F., van Boom, J. H., and Wang, H.-J. (1992) *Biochemistry 31*, 11823−11834.
38. Edwards, K. J., Jenkins, T. C., and Neidle, S. (1992) *Biochemistry 31*, 7104−7109.
39. Nunn, C. M., Jenkins, T. C., and Neidle, S. (1994) *Eur. J. Biochem. 226*, 953−961.
40. Nunn, C. M., and Neidle, S. (1995) *J. Med. Chem. 38*, 2317−2325.
41. Nunn, C. M., Garman, E., and Neidle, S. (1997) *Biochemistry 36*, 4792−4799.
42. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 133−138.
43. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr., Sect. D 50*, 869−873.

BI9907882