Intercalation of an Acridine−Peptide Drug in an AA/TT Base Step in the Crystal Structure of $[d(CGCGAATTCGCG)]_2$ with Six Duplexes and Seven Mg^{2+} Ions in the Asymmetric Unit†

Lucy Malinina,‡,§,|| Montserrat Soler-López,‡,|| Joan Aymamí,‡ and Juan A. Subirana*,‡

Departament d'Enginyeria Química, ETSEIB, Universitat Politècnica de Catalunya, Diagonal 647, E-08028 Barcelona, Spain, and Laboratory of X-Ray Crystallography, Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov Street 32, 117984 Moscow, Russia

Received February 15, 2002

ABSTRACT: We present the crystal structure of an acridine drug derivatized at carbon 9, [N^{α}-(9-acridinoyl)-tetraarginine], intercalated within the dodecamer [d(CGCGAATTCGCG)]$_2$. The presence of a lateral chain at the central carbon 9 atom differentiates this compound from most acridine drugs hitherto studied, which are usually derivatized at carbon 4. The DNA:drug interaction we observe differs from that observed in previous studies, which primarily involves shorter, mainly hexameric sequences, in two important regards: the acridine intercalates within an AA/TT base step, rather than within a CG/CG base step; and the binding site is located at the center of the sequence, rather than at one end of the duplex. In addition, we observe a novel crystal packing arrangement, with six dodecamer duplexes and seven hydrated magnesium ions in the asymmetric unit of a large ($66.5 \times 68.4 \times 77.4$ Å^3) unit cell in space group $P2_12_12_1$. The duplexes are organized in layers parallel to the *ab* plane, with consecutive layers crossing each other at right angles.

Derivatives of the planar nitrogen heterocyclic compound acridine are used (*1*) as antineoplastic agents (e.g., amsacrine) and are being developed as HIV antivirals (*2*), following the report that interactions with nucleic acids can be mimicked and improved when the ring system of the acridine derivative replaces the aliphatic side chains of the HIV regulatory proteins. However, the acridine alone is not sufficient for a strong and specific DNA binding. Therefore, peptides and other compounds have been added at the central carbon 9, usually through an amino group. Biochemical and NMR analysis of the adducts (*3*−*5*) have shown that they do interact with DNA, although no data are available on the crystal structure of these compounds when intercalated in DNA. Recently, the structure of one of these compounds was reported (*6*), but it lies in an extrahelical position in a complex crystal structure. In fact, very little data are available on acridines intercalated in DNA. The intercalation of the related compound proflavine was studied in structural detail in dinucleotides (*7*−*9*). Recently, the structure of an intercalated antitumor acridine derivatized at carbon 4, DACA, was reported (*10*). However, as with most previous structures of DNA-intercalated drug complexes, the structure was of a hexameric DNA with intercalation occurring at the terminal base steps. Related compounds have been found to intercalate in extrahelical positions (*11*, *12*). Here we present the structure of N^{α}-(9-acridinoyl)-tetraarginine intercalated in the dodecamer duplex d(CGCGAATTCGCG). The structure of the drug as it appears in the crystal structure is shown in Figure 1. It is the first reported structure of an intercalated acridine derivatized at the central carbon 9 and the first report of a drug intercalated in an AA/TT base step in the central region of a comparatively long DNA fragment. Intercalation in AA/TT was unknown up to now. It should be noted that almost 100 crystal structures have been determined for drugs intercalated in DNA (*13*). Practically all of them prefer alternating base steps (mainly CG/CG). Only recently (*14*) was another planar drug found to intercalate in a CC/GG base step.

An additional unexpected finding was the complex packing arrangement observed in the crystal structure. The dodecamer d(CGCGAATTCGCG) used in this study was the first B-DNA fragment to be analyzed (*15*) by single crystal X-ray diffraction methods. That work directly confirmed the double helical structure of B-DNA and allowed the study of the details (*16*, *17*) of B-DNA structure, including bending, the influence of base sequence (*18*), and hydration (*19*). High-resolution studies (*20*−*22*) have provided greater detail on the structure of this dodecamer and its ionic and hydration atmosphere. Many other dodecanucleotides with a related sequence and their complexes with minor groove binding drugs crystallize isomorphously in the same space group ($P2_12_12_1$), as can be ascertained by inspecting the Nucleic Acid Database (*13*). Some exceptions exist in which the dodecamers crystallize in related space groups (*23*−*26*), but

† This work was supported in part by the Dirección General de Enseñanza Superior e Investigación Científica (Grant PM98-0135), the Generalitat de Catalunya (Grant 1999SGR-00144), and the Training and Mobility of Researchers (Large Scientific Facilities) Program at the European Molecular Biology Laboratory, Hamburg Outstation (ERBFMGECT-980134).

* To whom correspondence should be addressed. FAX: +34 934 017 150. E-mail: juan.a.subirana@upc.es.

‡ Universitat Politècnica de Catalunya.

§ Russian Academy of Sciences.

|| Both have contributed equally to this work.

10.1021/bi020135c CCC: $22.00
Published on Web 06/27/2002

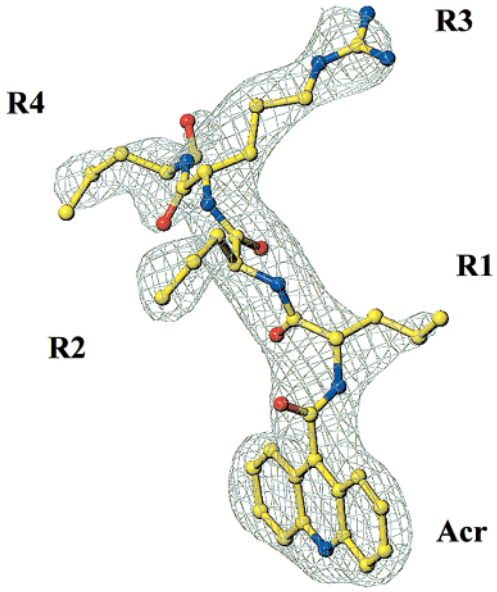

FIGURE 1: Electron density map of the whole drug (1σ level). The peptide main chain and the entire side chain of arginine 3 are clearly visible. The guanidinium groups of the other three arginine residues are disordered.

the packing interactions are only slightly modified. A different packing arrangement for the Drew−Dickerson dodecamer has been reported in two space groups of the hexagonal family: *R*3 in the Ca^{2+} salt (*27*) and $P32_12$ in the Ba^{2+} salt (*28*). In the latter case, the chemical symmetry of the dodecamer is maintained in the crystal: the two halves of the duplex have an identical conformation, whereas in all other cases, the whole dodecamer duplex or even two duplexes are found in the asymmetric unit in the crystal.

In this paper, we describe a new packing arrangement of d(CGCGAATTCGCG), with six duplexes in the asymmetric unit. Surprisingly only one of them contains an intercalated drug. The molecules are organized in columns, interacting among themselves in a similar way to that observed in the original $P2_12_12_1$ structure (*29*, *30*), except that in our case, the columns in consecutive layers cross each other at 90°. Such a perpendicular organization of neighboring duplexes is seldom found in B-form DNA. Only two related decamers which have been crystallized in a tetragonal space group (*31*, *32*) are also organized as layers of perpendicular columns of duplexes.

The new structure we describe here was crystallized in the absence of spermine. It is stabilized by Mg^{2+} ions. Despite the comparatively low resolution of our data (2.7 Å), we were able to unambiguously localize seven Mg^{2+} ions. It is likely that additional ions contribute to stabilize the crystal matrix. The analysis and comparison of the six different duplexes in the crystal with the standard Drew−Dickerson dodecamer (*15*−*22*) allows us to draw some conclusions on ionic interactions and their influence on DNA structure.

EXPERIMENTAL PROCEDURES

Synthesis. The drug−peptide conjugate was synthesized by solid-phase methods. After the Boc-tetraarginine sequence was assembled on a *p*-methylbenzhydrylamine resin, the N-terminus was deprotected, and 9-acridinecarboxylic acid was coupled by means of PyBOP (*33*) and diisopropylethylamine

Table 1: Crystal Data and Refinement Statistics

space group	$P2_12_12_1$
unit cell dimensions	$a = 66.47$ Å
	$b = 68.36$ Å
	$c = 77.36$ Å
	$\alpha = \beta = \gamma = 90°$
asymmetric unit contents	6 DNA duplexes
	1 N^{α}-(9-acrydinoyl)-tetraarginine
	4 MgO6
	3 MgO5
	61 H_2O
total	3052 non-hydrogen atoms
no. of unique reflections	8086
redundancy factor	3.5
$\langle I\rangle/\langle\sigma I\rangle$ (last shell)	13.2 (1.8)
data completeness (last shell)	89.4% (88.6%)
R_{merge} (last shell)	0.070 (0.483)
refinement resolution range	20−2.7 Å
test dataset size, %	10.2
work/test reflections	7165/921
R_{work}/R_{free}, %	22.00/26.61
RMS (bonds, Å; angles, deg)	0.013, 2.409

Table 2: Characteristics of Molecular Replacement Search and Model Splitting

no. of rigid bodies in asymmetric unit[a]		1	2	3	4	5	6	8	73
model 1[b]	CC, %[c]	46.7	50.1	54.9	57.7	61.6	63.8	66.2	74.5
	R-factor, %	59.0	57.9	55.7	54.6	53.4	52.4	51.3	46.4
model 2[b]	CC, %[c]	55.6	58.8	64.0	66.2	70.5	72.4	75.4	81.1
	R-factor, %	55.1	53.7	50.4	49.2	47.2	46.3	44.4	40.0

[a] Six duplexes were added one by one. Then, one of them was appropriately splitted, and the intercalated molecule was inserted between two parts. At the last step, all duplexes were splitted into base-pairs. See text for details. [b] The resolution range 12−2.7 Å was used for a molecular replacement search. Models 1 and 2 were duplexes of dodecamer CGCGAATTCGCG available from NDB (*13*) as entries BDL084 and BDL001. [c] Correlation coefficient values.

(5, 5 and 10 equivalents, respectively) in 1:1 DMSO-DMF for 1 h. Acidolysis with HF/anisole (9:1, 0 °C, 1 h) provided the target compound in highly homogeneous form (>95% by HPLC) and with a correct MALDI-TOF mass spectrum.

Crystallization. The dodecamer was crystallized at 20 °C, using a batch method, in sitting drops containing 0.5 mM duplex, 1mM drug-peptide adduct, 20 mM sodium cacodylate buffer, pH 7, 100 mM $MgCl_2$, and 45% MPD. Crystals appear in about 1 week. Crystals of approximately 0.2 × 0.3 × 0.3 mm^3 were used for data collection.

Data Collection and Analysis. Single crystals were flash-frozen in liquid nitrogen at 120 K. An initial data set containing 29 388 reflections was collected at EMBL beamline BW7A (DESY, Hamburg) to 2.54 Å resolution using a Mar Research image-plate detector and a wavelength of 1.1 Å. A final data set with 8086 reflections (35.0−2.7 Å) was used for structure refinement. Determination of unit-cell parameters, space group, and integration of reflection intensities was performed using DENZO (*34*), and the data were scaled using SCALEPACK. The results obtained are summarized in Table 1.

Structure Determination. The structure was determined by molecular replacement with the AMoRe program (*35*). The progress of the molecular replacement search is listed in Table 2. Since the dodecamer CGCGAATTCGCG is known to form B-DNA, we assumed that it adopted the B conformation in our crystal as well. NDB (*13*) entries BDL084

and BDL001 served as two trial models of the dodecamer duplex. Although structure BDL084 had been studied at higher resolution, we succeeded in getting a solution by using the BDL001 structure (*16*) and then reproduced appropriate orientations and positions for model BDL084. Table 2 demonstrates the difference between correlation coefficients and *R*-factor values for both models at various steps of the search. The main line of strategy was to search for the next molecule while previously found duplexes were kept fixed in the unit cell. The actual search, however, was more complicated: the solution for the first DNA molecule was obtained from a model constructed by two consequent B-DNA dodecamers forming a column through N2−N3 interactions between terminal guanine residues, as found in the usual crystal packing arrangement for this dodecamer (*29*). For the other DNA molecules, a dodecamer with one or two terminal base pairs removed was used as a model. However, the principles of an appropriate packing arrangement were also used during the search (*30*, *36*): (a) B-DNA dodecamers should form continuous columns through N2−N3 interactions between terminal guanine residues; (b) in our case, such columns should be oriented along the *a* and *b* directions of unit cell; (c) the general arrangement of duplexes along the *a* and *b* axes should be similar to each other; (d) there could be 4−6 duplexes in the asymmetric unit. Since we found the first duplex to be located on a 2_1 screw axis, we concluded that the asymmetric unit should contain six duplexes. The preliminary (2Fo-Fc) map calculated after getting a solution for all six duplexes showed that the last molecule had an intercalated drug at an AA/TT base step. Splitting of this molecule and an appropriate insertion of the acridine drug improved the values of the correlation coefficient and *R*-factor (Table 2). The last step with the AMoRe program was to split all duplexes into base pairs and fit them into the structure, with a total number of 73 rigid bodies.

Refinement. The model with six dodecamer duplexes and one acridine molecule was first refined in XPLOR 3.851 (*37*) using a data range from 8 to 3 Å. 2Fo-Fc and Fo-Fc maps calculated after some cycles of rigid-body refinement (the number of rigid-groups was increased step by step from 73 to 421) allowed adding to the model the oligopeptide $(Arg)_4$ covalently linked to acridine. Although most of the guanidinium groups of $(Arg)_4$ could not be located, the main chain was clearly traced in both maps, as well as parts of the side chains. Seven hydrated magnesium ions were also apparent in the map. The location of these ions in logical positions indicates that the refinement has been carried out adequately. Then the resolution range was extended, and the model was refined using the "TLS & Restrained refinement" option in REFMAC (*38*, *39*). The final refinement converged at R_{work} 0.22 and R_{free} 0.266 for all data in the range 20−2.7 Å. The asymmetric unit contains six double-stranded DNA molecules, one drug-peptide molecule, seven hydrated magnesium ions, and 61 additional water molecules. Examples of the electron density map (2Fo-Fc) are shown in Figures 1 and 2. Final coordinates have been deposited in the Nucleic Acid Database (NDB code DD0037).

RESULTS

General Packing Features. The structure of this dodecamer crystal has two main unique features: the complex molecular packing and the pattern of drug intercalation. Both are related, but for the sake of clarity, we will first analyze the packing features of the crystal.

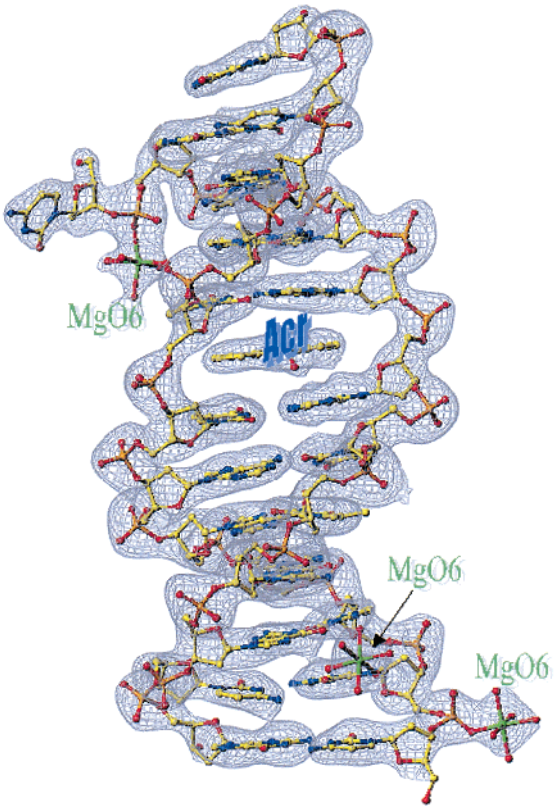

FIGURE 2: Electron density map ($2F_0-F_c$) of duplex **6** (1σ level). The intercalated drug is visible at the center of the duplex, as well as the terminal cytosine 24 which occupies an extrahelical position. The peptide which is attached to acridine has been removed for clarity. Three hydrated magnesium ions are also shown.

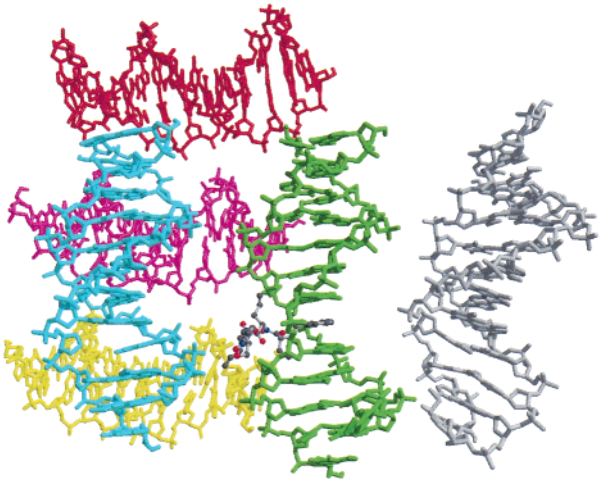

FIGURE 3: View of the six duplexes in the asymmetric unit. They are color-coded: **1**, red; **2**, yellow; **3**, gray; **4**, cyan; **5** pink; **6**, green. The intercalated drug in molecule **6** is represented as spheres with color-coded atoms. Note that the duplexes which are on the same plane are not exactly parallel.

The six dodecamers in the asymmetric unit (*40*) are shown in Figure 3. Only one of them has an intercalated drug. A simplified view of their packing arrangement is given in Figure 4. The dodecamers are organized in continuous columns similar to those observed in the standard packing arrangement with a single duplex in the asymmetric unit in

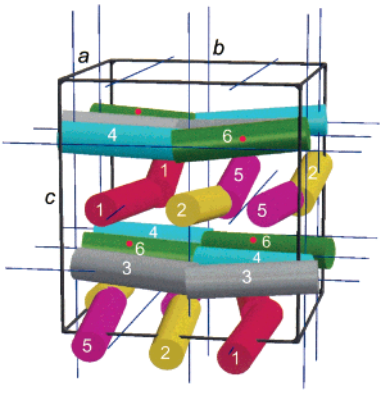

FIGURE 4: Schematic view of the organization of the unit cell. Each duplex is color coded as in Figure 3. The molecules form infinite columns of different types in the crystal which are horizontal in this drawing. The red dots indicate the position of the intercalated drug.

space group $P2_12_12_1$ (*16*, *28*, *29*). In all cases, each column is stabilized by guanine−guanine interactions between the two terminal base pairs of the duplexes. One set of oligonucleotide columns has the molecular axes oriented along the *y* axis, whereas another set has the molecules oriented along *x*. Both sets are organized as pleated sheets on planes perpendicular to the *z* axis. In each plane the columns are approximately parallel, the average distance among neighbor DNA columns is 22.5 Å. This organization is similar to the classical $P2_12_12_1$ structure of the dodecamer in planes perpendicular to the *z* axis. However, in the latter case, all oligonucleotides have their axes oriented in the same direction, whereas in our case the molecules have their axes alternatively oriented in the *x* and *y* directions, which results in a different overall organization of the crystal. It should also be noted that neighbor molecules in each plane are not related by an exact translation, as it can be easily appreciated in Figure 3.

In the crystal, there are two different types of columns. Duplexes **1** (shown in red) and **3** (shown in gray) are perpendicular to each other and form continuous columns of identical duplexes, related by one of the screw axes of the crystal. Another type of column is the one formed by duplexes **2** (in yellow) and **5** (in purple) where the two types of duplexes present in each column occupy alternate positions. A similar stack is formed by duplexes **4** (in cyan) and **6** (in green).

In other reported packing arrangements, the molecules are also parallel but displaced about 24.6 Å in one case (*17*), whereas in a different setting (*27*), neighbor molecules are rotated 120° around their axis and mutually displaced 24.2 Å. Thus, the average distance between neighbor columns is smaller in our case (22.5 Å).

In summary, all packing arrangements of d(CGCGAATTCGCG) crystals described up to now have all molecules approximately parallel, organized on planes. In our case, the molecules are also organized in planes perpendicular to the *z* axis, but alternate planes have the molecular axes in approximately perpendicular orientations.

Intercalated Drug. The intercalation we have found shows several novel features. Practically all known crystallographic structures of intercalated drugs have their substituents at the extremities of the aromatic system of rings, whereas the acridine derivative we have studied has the arginines attached

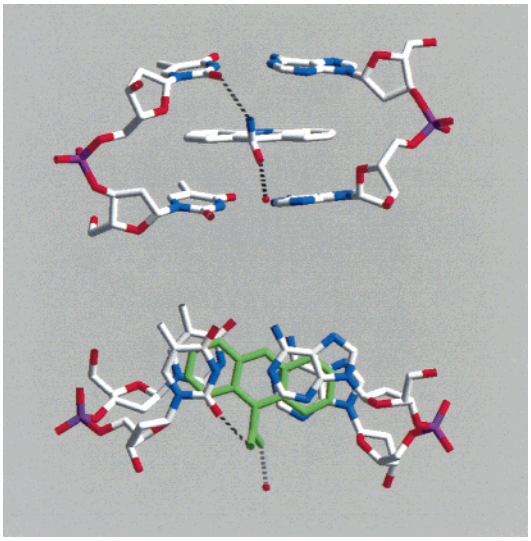

FIGURE 5: Interactions of the intercalated acridine drug. Only the first amide group of the lateral peptide chain is shown. It forms a hydrogen bond (dashed lines) with the O2 atom of thymine. Another hydrogen bond is formed with a water molecule, shown as a small red sphere. In the lower frame, the drug is shown in green.

at the central carbon 9. As a result of this difference in chemical structure, the acridine ring is in another position: it is intercalated between residues A5 and A6 as shown in Figure 2, with its peptide moiety lying in the minor groove. The AA/TT base step has a smaller twist (18.9°), as it is usually found in intercalated drugs.

This is the first aromatic drug found to intercalate into an AA/TT step. Previous analysis of several intercalating drugs with DNA had only shown selective interactions with alternating sequences (*41*), mainly poly(dG−dC). Details of the interaction of the drug with the AA/TT base step are shown in Figure 5. The drug intercalates with a standard stacking distance of about 3.4 Å to both A•T base pairs. Its major axis is approximately parallel to the C1′−C1′ vector of one of the two A•T base pairs, A6T19. This is the usual orientation for acridines. The peptide side chain attached to acridine lies in the minor groove. The positively charged arginines are probably responsible for such behavior. The presence of a hydrogen bond between the exocyclic amide of the drug and the O2 atom of thymine contributes to stabilizing the interaction. Such stabilization is possible since the acridine drug described here has the peptide attached through a carbonyl group, in contrast with other drugs which are amino conjugated. Both facts determine the local interactions at the intercalation site, as it is apparent in Figure 5.

The electron density map (Figure 1) of the peptide−drug adduct clearly shows the main chain of the tetra-arginine moiety. However, the peptide side chains are ill defined and could not be unambiguously located. Only the complete side chain of Arg 3 was visible. It appears that the other arginine residues may occupy multiple positions since they have several charged phosphate groups in their neighborhood. Despite this local disorder, it is likely that all arginine side chains contribute to stabilize the crystal lattice. This is particularly evident in the arginine side chain 3, as shown in Figure 6. Its guanidinum group is hydrogen bonded to residues from three different duplexes in the crystal, including a 5′ terminal OH group and two phosphates.

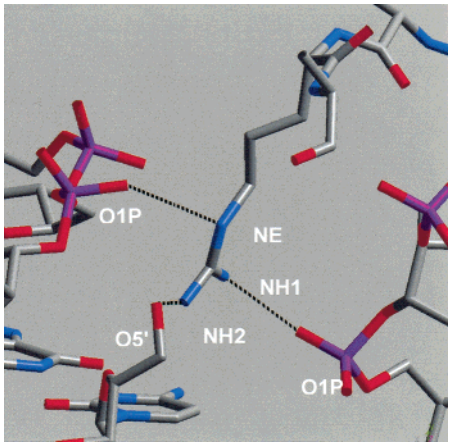

FIGURE 6: Interactions of the guanidinium group of arginine 3. Hydrogen bonds are formed with phosphates which belong to molecules **4** and **6**. An additional hydrogen bond is formed with the 5′ terminal OH of a symmetry-related molecule **6**.

Central Role of the Drug-Binding Molecule. The drug-binding molecule **6** (shown in green in Figures 3 and 4) plays a central role in the packing arrangement of the crystal. It appears to form a seeding point around which a complex ensamble of DNA duplexes is organized. It is very clearly defined in the electron density map and shows strong interactions with molecule **1**, which is approximately perpendicular to molecule **6** (Figure 7). It is surrounded by six hydrated Mg^{2+} ions, whose location will be analyzed in detail below. Interaction with molecule **1** involves three Mg^{2+} ions. Furthermore, a terminal cytosine base (shown in black in Figure 7) swings out from the duplex and penetrates into the minor groove of molecule **1**. The cytosine conformation is stabilized by van der Waals interactions, in particular by stacking over a sugar ring of the phosphodiester backbone of molecule **1**. It also forms one hydrogen bond with a thymine of molecule **1**. It should be noted that if the presence of drug imposes a perpendicular orientation of neighbor oligonucleotides, then it follows that the asymmetric unit must contain several duplexes, as we have found, to fill space efficiently.

Mg^{2+} Ions Are Major Groove Binders. Usually dodecanucleotides are crystallized at low Mg^{2+} concentrations (4−20 mM) in the presence of spermine. In our case, crystals were obtained at a much higher Mg^{2+} concentration (100 mM). Spermine was not used; instead, the acridine drug was present. Probably all these features are instrumental in order to create the crystal structure we observe. We located seven hydrated Mg^{2+} ions in the asymmetric unit. One of them is only associated with phosphate groups, which belong to molecules **1** and **6**, but the other six ions are major groove binders. Five of them are associated with a GC base step, as it is usually found in the conventional $P2_12_12_1$ unit cell of this dodecamer (*21*,*42*). Another one appears in a GA/TC base step, a position also present in a nonamer with a related sequence (*43*). Other Mg^{2+} major groove binding sites have also been recently reported (*22*), always involving a C•G base pair. All these ions contribute to stabilize the crystal lattice, since they also interact with phosphate groups of neighbor molecules. Nevertheless, it appears that the interaction sites represented in Figure 8 are favored locations for DNA−Mg interactions, since they are present in several cases. They are not always visible in DNA crystals, probably due to a high mobility. Only when there is a phosphate residue in a near position do they appear to be stabilized in a well-defined location.

The contribution of the Mg^{2+} ions to stabilize the crystal lattice is clearly apparent from Figure 7. The interaction between molecules **1** and **6** is stabilized by three Mg^{2+} ions, two of which create base−phosphate bridges. The other three ions shown in Figure 7, at the right and lower part of the figure, also generate base−phosphate bridges between molecule **6** and molecules **2** and **5**. Thus, all six ions shown in Figure 7 create connections between molecule **6** and molecules found in the neighbor planes in the crystal. On the other hand, molecules **3** and **4** do not show any Mg^{2+}-mediated interaction with molecules in neighboring planes, although the seventh Mg^{2+} ions creates a base−phosphate bridge among the latter two molecules. In fact, the latter is the only located Mg^{2+} ion which defines an interaction among parallel molecules within one of the molecular planes of the structure.

Many more Mg^{2+} ions should be present in the structure in order to neutralize all the phosphate charges, but they have not been localized. In the crystal, there are many phosphate−phosphate distances between duplexes which are shorter than 7 Å, and counterions should be present in their neighborhoods. In particular, it is striking that we have only located one ion within the planes formed by parallel molecules perpendicular to the *z* axis (Figure 4), despite the comparatively short distance (22.5 Å) between parallel molecules. In any case, it is clear that Mg^{2+} ions should cement the helices together (*22*) and stabilize the crystal lattice.

Oligonucleotide Conformation. The resolution obtained in our crystal structure is not sufficient to warrant a detailed comparison and analysis of the conformation of the six dodecamers present in the asymmetric unit. The average twists are given in Table 3. From them and from an inspection of the electron density map, it is clear that dodecamer **5** has a less well defined conformation with greater local disorder. Nevertheless, it is clear from Figure 3 and from the average twist of the molecules (given in Table 3) that the overall conformation of each of the six dodecamers in the asymmetric unit is similar to the conformation of the same dodecamer when crystallized in other space groups.

Inspection of the conformation angles shows that the standard BI form predominates in all bases in each duplex, with the exception of the G10 and G22 nucleotides, which have the BII form in most cases. A few other bases are found to lie outside the conformational BI region, but given the low resolution of our structure, a detailed discussion is not warranted. Thus, the overall conformation of all six dodecamers is similar to that described (*17*) in the standard $P2_12_12_1$ space group, which also has G10 and G22 in the BII conformation.

Comparison of the twist values of the base steps T7 to G12 (half of the molecule) shows a clear alternation of twist in all six dodecamers (data not shown), as it is found in the other crystal structures obtained from the same molecule. A comparison was recently published showing this trend (*28*). It was also reported that the first half of the molecule (C1 to A6) was much more variable, a feature also found in our case. The asymmetry of the molecules appears to be related with their end-to-end interactions (*29*, *30*), which stabilize the columns of duplexes in the crystal. Only in one crystal

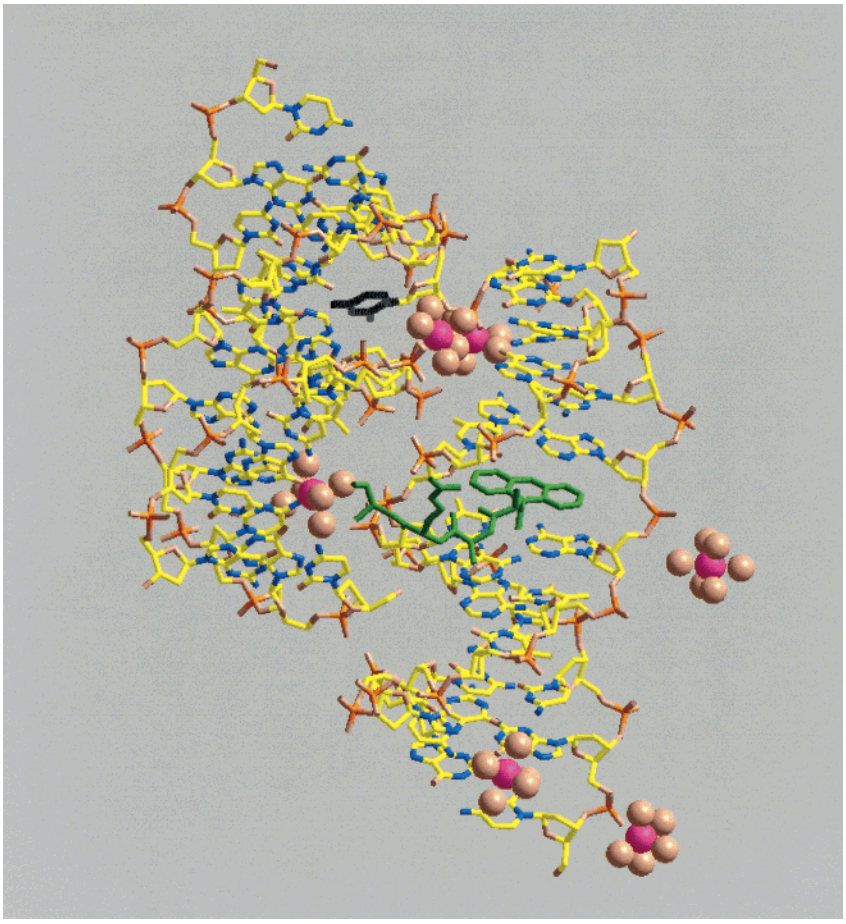

FIGURE 7: View of the interaction of molecule **6** (at right) with molecule **1**. The intercalated drug is shown in green and the extrahelical cytosine in black. Six magnesium ions are shown as pink spheres with their coordinate waters in brown. All of them interact with molecule **1** as described in the text. Only the complete side chain of arginine 3 is shown. For the other arginine residues, only the β carbon is included. For better visibility, the duplexes are shown inclined; they are approximately perpendicular among themselves.

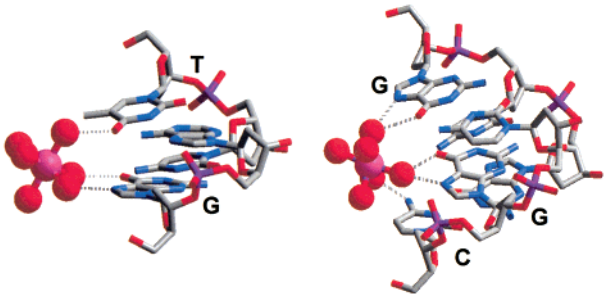

FIGURE 8: Two types of interaction of hydrated Mg^{2+} ions with the major groove of duplexes. At left is shown the interaction with a GA/TC step, at right with a CGC sequence. The latter is the most common. Mg^{2+} ions interact mainly with both guanines in the GC step, and the hydrogen bond with cytosine is not always found. The bases which form hydrogen bonds with the hydrated Mg^{2+} ions are labeled.

structure in a trigonal space group (*28*), where the end-to-end interactions are different, do the molecules show a symmetric conformation.

The ends of each dodecamer interact with a neighbor dodecamer through hydrogen bonds between the N2 and N3 atoms of guanines in the two terminal base steps, as occurs in the classical $P2_12_12_1$ structure (*29*, *30*). However, cytosine 1 is disordered in molecule **4** and does not appear in the electron density map, a feature already observed in other cases (*26*, *27*, *31*, *32*). Furthermore, cytosine 1 in molecule **6** swings out from the duplex as shown in Figure 7.

Table 3: Average Twist of the Six Duplexes in the Asymmetric Unit[a]

molecule	twist (deg)
1	36.1 (3.4)
2	35.7 (6.0)
3	35.9 (5.8)
4	37.5 (4.7)
5	35.4 (6.8)
6	37.3 (4.3)
standard (*14*)	35.8 (4.1)

[a] Standard deviations are given in parentheses. For molecule **6**, the base step AA/TT, which contains the intercalated drug, has been excluded from the calculation, and it has a twist of 18.9°.

Nevertheless, in both cases, the guanine−guanine interactions between neighbor duplexes are maintained.

A surprising feature of our structure is the absence of the hydration spine (*19*) in all of the six duplexes in the asymmetric unit, although this is a very stable feature of the dodecamer structure (*40*). This peculiarity may be due to the different conditions of crystallization, in particular the high Mg^{2+} concentration.

DISCUSSION

The oligodeoxynucleotide used in this study provided the first detailed image of the DNA double helix in its B-form (*15*). It was later studied in great detail (*17*−*21*, *27*−*30*), as well as other duplexes with related sequences (*13*, *22*−*26*,

43). It is the reference structure for many theoretical studies and it has been extensively used for the analysis of groove-binding drugs. Here we have reported a new feature of this classic molecule: the first case in which a drug has been intercalated into the Dickerson−Drew dodecamer. Thus, the use of this compound as a reference structure is further extended. New avenues for the discovery of drugs with different types of interaction with DNA are thus available.

The dodecamer d(CGCGAATTCGCG) had been previously crystallized in the $P2_12_12_1$ space group (*16*) in the presence of spermine and a low Mg^{2+} concentration. It has also been crystallized in the R3 space group (*27*) in the presence of Ca^{2+}. In both cases, the asymmetric unit contains one duplex molecule. Recently another structure was reported (*28*) in the presence of Ba^{2+} ions which has a higher symmetry, with a single strand in the asymmetric unit. In our case, we have observed a much more complex structure. Nevertheless, the conformation of the individual duplexes does not change significantly. This more complicated structure probably appears as a result of the higher concentration of Mg^{2+} in the crystallization medium and the presence of drug. The strong interactions between molecule **1** and **6**, shown in Figure 7, might form a nucleation center around which the crystal is built. Alternatively, the high Mg^{2+} concentration might favor a crystal structure with duplex planes crossing at 90° angles into which the drug penetrates. In this sense, it should be noted that some of the GC steps in which we find Mg^{2+} ions are occupied by a spermine molecule in the classical $P2_12_12_1$ structure (*19*, *21*).

A puzzling feature of our results is the presence of a single drug molecule in the asymmetric unit, since it contains 12 potential AA/TT binding sites. It could be expected that there were at least six drug molecules bound, one in each duplex. However, it appears that the affinity of the drug for the duplex is not sufficient to achieve complete intercalation. Probably a single site is stabilized through the complex interaction that involves both duplexes **1** and **6**, as shown in Figure 7. Additional disordered drug molecules may be present in interstitial spaces among the duplexes in the crystal. Attempts to cocrystallize this drug with an oligonucleotide with the related sequence GCGAATTCG had been carried out before (*43*) under practically identical ionic conditions, but no drug was found in the crystal. That result confirms that this drug does not easily form cocrystals containing an intercalated drug.

The structure of another complex containing DACA, also an acridine drug, differs significantly from ours (*10*). As with most previous structures of DNA-intercalated drug complexes (reviewed by Williams et al., *45*), the structure is that of a hexameric DNA with intercalation occurring at the terminal base steps. It should be noted that the terminal base steps have a much greater conformational freedom than internal base steps, where base intercalation may occur more easily. In contrast, because the acridine-peptide drug is inserted in the center of the dodecanucleotide, we believe that our structure is more related to intercalation in full length DNA. Aside from the similar orientation of the acridine ring, the DACA amino derivative shows a different pattern of interactions. The lateral chain, now attached to carbon 4 of acridine, is found in the major groove. The interaction is stabilized through several hydrogen bonding interactions.

The acridine drug used in this study is intercalated in an AA/TT base step. Recently another drug was reported to intercalate in a GG/CC step (*14*). Most drugs previously studied (*10*, *13*) were observed to intercalate in alternating purine-pyrimidine base steps. From all these findings, the requirements for sequence specific intercalation start to emerge. Thus, the way is open to develop sequence specific drugs by linking acridines (or other aromatic drugs) with chains of an appropriate length. Thus, polyfunctional drugs may be developed which can bind to different base steps placed at a designed distance.

ACKNOWLEDGMENT

We thank Dr. D. Andreu and co-workers for the synthesis of the acridine conjugate.

REFERENCES

1. Budavari, S. (1969) The Merck Index, 12th ed., p 639, Merck & Co., Whitehouse Station, NJ.
2. Hamy, F., Brondani, V., Flörsheimer, A., Stark, W., Blommers, M. J. J., and Klimkait, T. (1988) *Biochemistry 37*, 5086−5095.
3. Bailly, F., Bailly, C., Helbecque, N., Pommery, N., Colson, P., Houssier, C., and Hénichart, J. P. (1992) *Anti-Cancer Drug Des. 7*, 83−100.
4. Takenaka, S., Iwamasa, K., Takagi, M., Nishino, N., Mihara, H., and Fujimoto, T. J. (1996) *Heterocyclic Chem. 33*, 2043−2047.
5. Bourdouxhe-Housiaux, C., Colson, P., Houssier, C., Waring, M. J., and Bailly, C. (1996) *Biochemistry 35*, 4251−4264.
6. Yang, X.-L., Robinson, H., Gao, Y.-G., and Wang, A. H.-J. (2000) *Biochemistry 39*, 10950−10957.
7. Berman, H. M., Stallings, W., Carrell, H. L., Glusker, J. P., Neidle, S., Taylor, G., and Achari, A. (1979) *Biopolymers 18*, 2405−2429.
8. Shieh, H.-S., Berman, H. M., Dabrow, M., and Neidle, S. (1980) *Nucleic Acids Res. 8*, 85−97.
9. Sakore, T. D., Bhandary, K. K., and Sobell, H. M. (1984) *J. Biomol. Struct. Dyn. 1*, 1219−1227.
10. Adams, A., Guss, J. M., Collyer, C. A., Denny, W. A., and Wakelin, L. P. G. (1999) *Biochemistry 38*, 9221−9233.
11. Thorpe, J. H., Hobbs, J. R., Todd, A. K., Denny, W. A., Charlton, P., and Cardin, C. J. (2000) *Biochemistry 39*, 15055−15061.
12. Adams, A., Guss, J. M., Collyer, C. A., Denny, W. A., and Wakelin, L. P. G. (2000) *Nucleic Acids Res. 28*, 4244−4253.
13. Berman, H. M., Olson, W. K., Beveridge, D. T., Westbrook, J., Gelvin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R., and Sneider B. (1992) *Biophys. J. 63*, 751−759.
14. Lisgarten, J. N., Coll, M., Portugal, J., Wright, C. W., and Aymamí, J. (2002) *Nat. Struct. Biol. 9*, 57−60.
15. Wing, R., Drew, H., Takano, T., Broka, C., Takano, S., Itakura, K., and Dickerson, R. E. (1980) *Nature 287*, 755−758.
16. Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S., Itakura, K., and Dickerson, R. E. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2179−2183.
17. Fratini, A. V., Kopka, M. L., Drew, H. R., and Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686−14707.
18. Dickerson, R. E., and Drew, H. R. (1981) *J. Mol. Biol. 149*, 761−786.
19. Drew, H. R., and Dickerson, R. E. (1981) *J. Mol. Biol. 151*, 535−556.
20. Egli, M., Tereshko, V., Teplova, M., Minasov, G., Joachimiak, A., Sanishvili, R., Weeks, C. M., Miller, R., Maier, M. A., An, H., Cook, P. D., and Manoharan, M. (2000) *Biopolymers 48*, 234−252.
21. Shui, X., Sines, C. C., McFail-Isom, L., VanDerveer, D., and Williams, L. D. (1998) *Biochemistry 37*, 16877−16887.
22. Chiu, T. K., and Dickerson, R. E. (2000) *J. Mol. Biol. 301*, 915−945.
23. DiGabriele, A. D., and Steitz, T. A. (1993) *J. Mol. Biol. 231*, 1024−1039.
24. Leonard, G. A., and Hunter, W. N. (1993) *J. Mol. Biol. 234*, 198−208.
25. Hunter, W. N., Langlois d'Estaintot, B., and Kennard, O. (1988) *J. Mol. Biol. 202*, 921−922.

26. Tereshko, V., Urpí, L., Malinina, L., Huynh-Dinh, T., and Subirana, J. A. (1996) *Biochemistry 35*, 11589−11595.
27. Liu, J., and Subirana, J. A. (1999) *J. Biol. Chem. 274*, 24749−24752.
28. Johansson, E., Parkinson, G., and Neidle, S. (2000) *J. Mol. Biol. 300*, 551−561.
29. Dickerson, R. E., Goodsell, D. S., Kopka, M. L., and Pjura, P. E. (1987) *J. Biomol. Struct. Dyn. 5*, 557−579.
30. Tereshko, V., and Subirana, J. A. (1999) *Acta Crystallogr., Sect. D 55*, 810−819.
31. Abrescia, N. G. A., Malinina, L., Fernandez, L. G., Huynh-Dinh, T., Neidle, S., and Subirana, J. A. (1999) *Nucleic Acids Res. 27*, 1593−1599.
32. Abrescia, N. G. A., Huynh-Dinh, T., and Subirana, J. A. (2002) *J. Biol. Inorg. Chem. 7*, 195−199.
33. PyBOP: benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate.
34. Otwinowsky, Z, and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
35. Navaza, J. (1994) *Acta Crystallogr., Sect. D 50*, 157−163.
36. Malinina, L. V., and Ivaninskii, S. V. (1999) *Mol. Biol. 33*, 453−459.
37. Brunger, A. T. (1992) *XPLOR Manual*, Yale University Press, New Haven, CT.
38. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr., Sect. D 53*, 240−255.
39. Winn M. D., Isupov, M. N., and Murshudov, G. N. (2001) *Acta Crystallogr., Sect. D 57*, 122−133.
40. The duplexes in the asymmetric unit are numbered **1**−**6** throughout this paper.
41. Waring, M. J., and Bailly, C. J. (1994) *Mol. Recognit. 7*, 109−122.
42. Tereshko, V., Minasov, G., and Egli, M. (1999) *J. Am. Chem. Soc. 121*, 470−471.
43. Soler-López, M., Malinina, L., and Subirana, J. A. (2000) *J. Biol. Chem. 275*, 23034−23044.
44. Soler-López, M., Malinina, L., Liu, J., Huynh-Dinh, T., and Subirana, J. A. (1999) *J. Biol. Chem. 274*, 23683−23686.
45. Williams, L. D., Egli, M., Gao, Q., and Rich, A. (1992) in *Structure & Function, Vol. I: Nucleic Acids* (Sarma, R. H., and Sarma, M. H., Eds.) pp 107−125, Adenine Press, Schenectady, NY.

BI020135C