Targeting the Minor Groove of DNA: Crystal Structures of Two Complexes between Furan Derivatives of Berenil and the DNA Dodecamer d(CGCGAATTCGCG)$_2$

John O. Trent,[†] George R. Clark,[‡] Arvind Kumar,[§] W. David Wilson,[§] David W. Boykin,[§] James Edwin Hall,[∥] Richard R. Tidwell,[∥] Byron L. Blagburn,[▽] and Stephen Neidle*,[†]

The CRC Biomolecular Structure Unit, The Institute of Cancer Research, Sutton, Surrey SM2 5NG, U.K., Chemistry Department, University of Auckland, Auckland, New Zealand, Department of Chemistry, Georgia State University, Atlanta, Georgia 30303, Department of Pathology, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, and Department of Pathobiology, College of Veterinary Medicine, Auburn University, Auburn, Alabama 36849

Received June 21, 1996[⊗]

Crystal structures are reported of complexes of two novel furan derivatives of berenil with alkyl benzamidine groups bound to the DNA sequence d(CGCGAATTCGCG)$_2$. They have both been determined to 2.2 Å resolution and refined to *R* factors of 16.9% and 18.6%. In both structures the alkyl substituents, cyclopropyl and isopropyl, are found to be orientated away from the floor of the minor groove. The drugs are located in the minor groove by two strong amidinium hydrogen bonds, to the O2 of the thymines situated at the 5′ and 3′ ends of the AT-rich region. The isopropyl-substituted derivative has a tight hydrogen-bonded water network in the minor groove at one amidine site, which alters the orientation of the isopropyl substituent. This compound has superior DNA-binding properties and activity against *Pneumocystis carinii* and *Cryptosporidium parvum* infections *in vivo* compared to the cyclopropyl derivative, which in turn is superior to the parent furan compound. We suggest that the nature and extent of the interactions of these compounds in the DNA minor groove play an important role in these activities, possibly in conjunction with a DNA-binding protein. The overall effect of these alkyl benzamidine substitutions is to increase the binding of the drugs to the minor groove.

Introduction

There is now considerable evidence that the superfamily of drugs exemplified by netropsin, distamycin, berenil, pentamidine, and Hoechst 33258 exerts their biological effects by means of noncovalent binding to the AT-rich regions in the minor groove of DNA.[1−3] Interference with transcription control has been demonstrated in several instances,[4] especially with the TATA-box binding proteins at the initiation of transcription.[5] A number of these drugs have activity against parasitic or fungal diseases; in these cases, inhibition of the functioning of parasitic or fungal DNA topoisomerases (especially topoisomerase I) may be the basis for their selectivity.[6] Antitumor effects have been found for Hoechst 33258 itself and a number of alkylating analogues,[7] as well as for the alkylating distamycin analogue tallimustine,[8] where inhibition of basal transcription has been proposed as the basis for its action. Other distamycin alkylating analogues have been shown to display increased cytotoxicity with enhanced sequence selectivity.[9]

The bis(phenylamidinium) compound pentamidine (**1**) has established activity against the *Pneumocystis carinii* pathogen (PCP) which affects a high proportion of AIDS-infected patients and is a major cause of mortality in these individuals. This drug, which is one of the two currently used in the clinic, has toxic side effects, which has prompted a search for related compounds with less toxicity and greater efficacy. A correlation between anti-PCP activity and DNA binding has been established in several studies,[10] which may be due to selective interference with the normal processing of drug−DNA complexes by the pathogen's DNA topoisomerases.[11] X-ray crystallographic studies on oligonucleotides complexed to pentamidine and several analogues[12] have established that the drugs are minor groove binders, confirming DNA footprinting and other biophysical studies.[13]

We have recently reported that several dicationic diarylfuran compounds, which have a strong resemblance to pentamidine and berenil (**2**), have significant activity against PCP in animal models.[14] The parent compound in this series, furamidine (**3**), which is significantly more active and less toxic than pentamidine in the rat PCP model, also binds more strongly to duplex DNA. We have shown[15] by a combined X-ray crystallographic, molecular modeling, and biophysical study that this enhanced DNA affinity is in large part due to additional nonbonded interactions between the furan group and the walls of the narrow AT-rich minor groove.

This study reports structural and molecular modeling studies on two members of a further series of *N*-alkylamidine furans which have the alkyl substituents cyclopropyl (compound **4**) and isopropyl (compound **5**) directly attached to the amidinium groups (Figure 1). We examine here the effects of these substituents on DNA binding and biological activity.

* Corresponding author. Tel: (44) 181 643 8901, ext. 4251. Fax: (44) 181 643 1675. E-mail: steve@iris5.icr.ac.uk.
[†] The Institute of Cancer Research.
[‡] University of Auckland.
[§] Georgia State University.
[∥] The University of North Carolina at Chapel Hill.
[▽] Auburn University.
[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1996.

S0022-2623(96)00448-7 CCC: $12.00

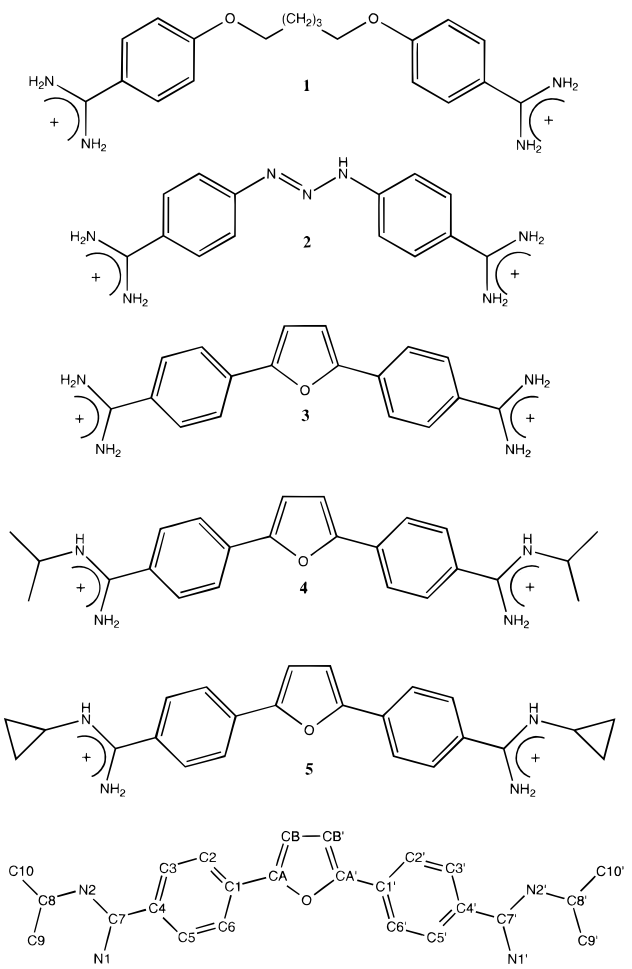

Figure 1. Structures of **1**−**5** with the numbering scheme of **4**.

Results and Discussion

Biological Results. Table 1 contains the results from a study of the interaction of **3**−**5** with DNA and from evaluation of these compounds in animal models for *P. carinii* and *Cryptosporidium parvum* infections. The binding to poly(dA·dT) was studied, and their DNA complexes melt at >90 °C. In order to rank the relative binding affinities for interactions with DNA, the duplex oligomer d(GCGCAATTGCGC)$_2$ (A2T2) was employed. The melting temperature range for this oligomer allows measurement of ΔT_m values for these compounds, even though they exhibit strong binding affinities. The increases in melting temperature on complex formation with the dicationic molecules are related to their binding affinities with nucleic acids.[16] It is notable that introduction of an alkyl group on the amidine nitrogen results in an increase in the ΔT_m values; compare the results for **3** with that for **4** and **5**. We have previously suggested that increases in ΔT_m values are due to involvement of increased van der Waals interactions.[17] The *in vivo* activity for the furan dications against *P. carinii* roughly parallels the DNA affinity of these compounds. The furans **3**−**5** are significantly more active and less toxic than pentamidine in the rat model. These results suggest that other alkylfurans should be investigated; such studies are underway. Blagburn and co-workers[18] noted that analogues of pentamidine showed activity in a neonatal mouse model for *C. parvum* which employs an oral dosage regimen. In view of the fact that no effective drug is available for treatment of *C. parvum* infections, we evaluated **3**−**5** in the neonatal mouse model. The data in Table 1 demonstrate that the three furans significantly reduce the oocyst counts in this screen. Compound **5** is of comparable activity to pentamidine, and **3** and **4** are more active in this model for *C. parvum* infections. Additional studies in other *C. parvum* models are in progress. Collectively these results, and earlier ones,[14,19] suggest that a strategy based upon design of molecules which selectively recognize the DNA minor groove is a promising approach for development of effective anti-OI agents.

Crystallographic Analysis. DNA and Drug Structures. The crystal structures of the complexes of compounds **4** and **5** with the d(CGCGAATTCGCG)$_2$ duplex (Figure 2) are both isomorphous with other drug−dodecanucleotide crystal structures, including those with compounds **1**−**3**. The DNA is in the right-handed B form. Both **4** and **5** fit well into the minor grooves at the center of the AT-rich regions in each duplex. In each structure, in principle, there are two possible orientations for the alkyl-substituted amidine groups: one with the alkyl group located at the floor of the groove and the other with it directed to the mouth of the groove. It is notable that despite the alkyl substitutions, both molecules fit deeply into the minor groove, with the amidine alkyl groups oriented away from the floor of the minor groove, as shown by their unambiguous fit to the electron density in both cases (Figure 3). The fit of the molecules to the density shows that the diphenylfuran moieties are highly coplanar, appreciably more so than they are in the dodecanucleotide complex of molecule **3**. This is shown by the average furan−phenyl dihedral angles (O−CA−C1−C6 and O−CA′−C1′−C6′), being 1°, 1°, and 10°, for **4**, **5**, and **3**, respectively. As has been observed previously with compounds containing the same core aryl rings, for example, the d(CGCGAATTCGCG)$_2$:**3** complex, both **4** and **5** form strong hydrogen bonds from N1 to O2 of T20 and from N1′ to O2 of T8. We would expect the overall effect of **4** and **5**, being more planar than **3**, to be a narrowing of the minor groove in the center of the AT region of their complexes, compared with the corresponding complex with **3**. The alkyl groups should not widen the groove significantly since they could form favorable hydrophobic interactions with the walls of the groove. There are two possible orientations that the alkyl groups can possess with respect to the minor groove axis direction, one parallel and one perpendicular. The minor groove can accommodate alkyl groups of this size (and larger ones) if they are orientated parallel to the groove, having minimal unfavorable interactions with the wall of the groove or with the alkyl groups pointing out perpendicular to the groove. In the case of the complex with **5**, only the perpendicular orientation is observed for the cyclopropyl group since it is more rigid and spherical in cross section and cannot have as narrow a profile as the isopropyl group of **4**. However, both orientations are observed for the isopropyl group in the complex with **4**, as seen in Figure 1. It is apparent that a water hydration network is responsible for the perpendicular orientation, in this case due to hydrogen bonding between the water molecules and amidinium groups (see below).

Table 1. *In Vivo* Activity of Dicationic Diarylfurans against *P. carinii* and *P. parvum*

		P. carinii			*C. parvum*	
compd	ΔT_m[a] *d*(GCGCAATTGCGC)$_2$	dosage[b] (μmol/kg/day)	toxicity[c]	cyst/g of lung[d] (% of control)	dosage[e] (μmol/kg/day)	oocysts[f] (% of control)
saline			0	100.0		
pentamidine	4.8	22.1	2	1.47	62	20.8
3	11.7	13.3	0	0.71	37.9	15
4	14.4	10.8	0	0.24	36.9	3
5	12.4	10.9	1	0.02	72.2	17

[a] Increase in thermal melting of the oligomer d(GCGCAATTGCGC)$_2$, see ref 4. [b] Intravenous dosage. [c] See ref 5 for detailed explanation of toxicity scale; generally, the larger the value the more severe the toxicity, and values greater than 2 indicate death of some animals. [d] Counted cysts in lung tissue; using a blinded protocol reported as percentage of saline-treated controls, see ref 5. [e] Oral dosage; no adverse reactions resulting from treatment with the test compounds were observed. [f] Counted oocysts obtained from alimentary tract tissue; using a blinded protocol reported as percentage of controls, see ref 3.

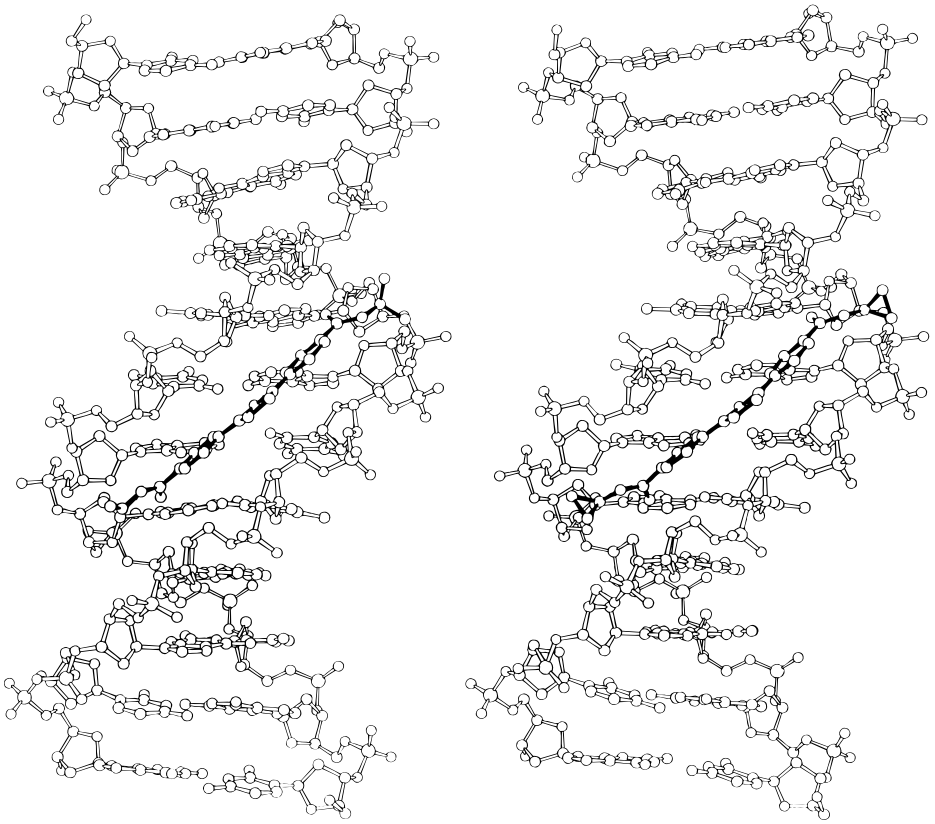

Figure 2. Structures of the complexes of (left) **4** and (right) **5** with d(CGCGAATTCGCG)$_2$. Water molecules have been removed for clarity.

Although the three core aryl rings in both complexes are nearly planar, the amidinium groups themselves are twisted significantly out from the aryl ring planarity. This is to be expected and is consistent with molecular mechanics calculations showing that the optimal *in vacuo* deviation between the phenyl and alkyl-substituted amidine groups is ca. 42° but close to 0° for unsubstituted amidines (a search of the small-molecule Cambridge Crystallographic Database for compounds with similar functionality confirms these orientations). Since the crystal structures of these complexes are not symmetric (the duplex plus drug molecule being the asymmetric unit), the amidinium groups tend to have varying conformations as shown by the N1−C7−C4−C3 and N1′−C7′−C4′−C3′ dihedral angles: A2T2:**4**, 18°, 40°; A2T2:**5**, 26°, 30°; and A2T2:**3**, 20°, 24°. These values are altered from those expected for the drug alone due to the strong thymine−drug hydrogen bonding overcoming low-energy barriers for dihedral angle (O−CA−C1−C6 and O−CA′−C1′−C6′) rotations. In each case the three aryl-linked rings are forced near to coplanarity in order to maintain the optimum orientation and position of the amidine ends. Despite these differing dihedral angles, the amidine nitrogen atoms are in nearly the same position in all three complexes (Figure 4), with about the same overall twist between the two amidine ends of each molecule in the complexes. The N1−C7−C7′−N1′ torsion angles are for **4**, 51°; **5**, 57°; and **3**, 60°; the intramolecular drug N1−N1′ distances are for **4**,12.5 Å; **5**,12.5 Å; and **3**,12.8 Å.

The locations of the drugs in the three furan complexes are almost identical with respect to the A2T2 region of the minor groove. This is not surprising since the two N1 and N1′ hydrogen bonds (to O2 of T8 and T20) fix the location of the amidine groups. It is notable that the minor groove can tolerate the binding of ligands with extended regions of planar structure; the three aryl rings of **4** and **5** have little specific interactions with the groove other than van der Waals contacts.

DNA Structure. The shape of the groove in the structures of the A2T2 complexes of **4** and **5** appears to have changed subtly from that in the complex with **3**. The duplexes in the first two complexes have a narrower groove floor, while the mouth of the groove stays

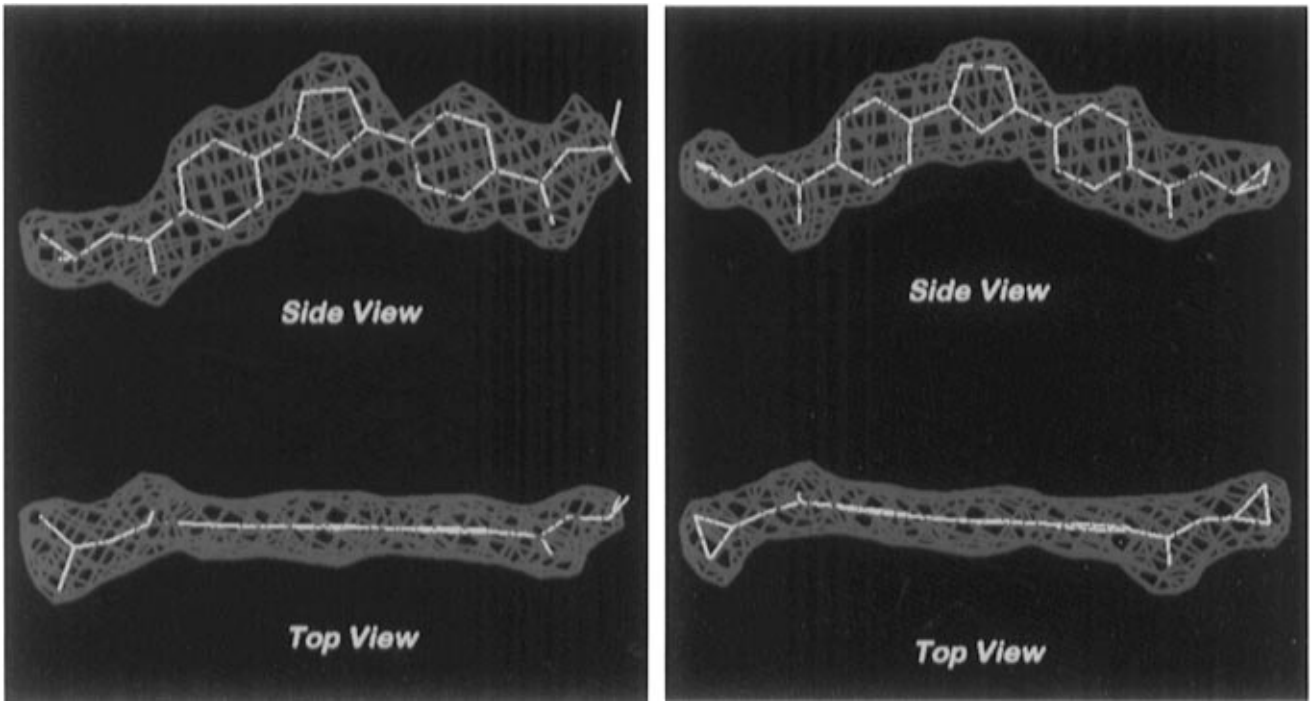

Figure 3. $F_o - F_c$ electron density maps (drug atoms have been omitted from the calculation) contoured at the 1.0 s level showing the "fit" of **4** and **5** to the density: (top left) side view of **4**, (bottom left) top view of **4**, (top right) side view of **5**, and (bottom right) top view of **5**.

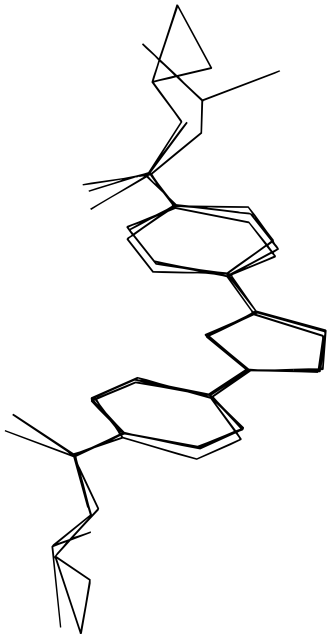

Figure 4. Superposition of bound drugs **3**−**5** showing phenyl ring rotation and amidine rotation and position. Atoms used in the superposition were C7, C4, C1, CA, CA′, C1′, C4′, and C7′.

approximately at the same width. The effect of this can be seen in the close contacts of the drugs and DNA (Table 2). The close contacts (<3.5 Å) for **4** and **5** are with the walls of the groove but with significantly fewer interactions than in the A2T2:**3** complex. The three aryl rings of **3** have a noncoplanar arrangement which can more effectively contact the groove walls.

The N1 and N1′ atoms of the drugs are involved in several nonbonded interactions in addition to the hydrogen bonds to O2 of T8 and T20; these are each with three other charge-complementary atoms: A6 O4′, A6 N3, C21 O4′; and A18 O4′, A18 N3, C9 O4′, respectively (Table 2). The flexibility of the base pair movement and the rotation of the amidine group mean that these nonbonding arrangements are variable, but there is a preference for a close contact to the sugar ring oxygen atoms C21 O4′ or C9 O4′.

The isopropyl and cyclopropyl groups have few close contacts (Table 2) with the minor groove when in the perpendicular orientation. The perpendicular oriented isopropyl carbon atoms C9 and C10 of the complex A2T2:**4** generally weakly interact with the phosphate backbone. The parallel oriented isopropyl group has completely different interactions: the C9′ atom situated in the minor groove interacts with the wall of the groove, while the C10′ atom, situated in the mouth of the minor groove, has no contacts under 4 Å. The cyclopropyl C9′ and C10′ atoms of the complex A2T2:**5** have weak interactions with the phosphate backbone that are similar to the perpendicular oriented isopropyl group in the A2T2:**4** complex; however, the cyclopropyl C9 and C10 atoms have only one interaction with the phosphate backbone, despite being in the same orientation.

The two alkyl-substituted amidine structures are very similar to each other, with a rms difference on superposition of all the DNA atoms of just 0.39 Å (Figure 5). The least variable region is in the AATT sequence where the drugs **4** and **5** effectively hold it relatively rigid. A comparison of helical parameters (calculated by the program CURVES[20]) for the A2T2 complexes of **3**−**5** with the native dodecamer[21] (Nucleic Acid Data Base[22] code bdl001) shows that in terms of propeller twist the two new structures with **4** and **5** are closer to the native than to the A2T2:**3** complex (Table 3). This is somewhat surprising due to the structural similarity of **4** and **5** to **3**, although it appears that the large twists between the phenyl ring and the alkyl-substituted amidine groups in **4** and **5** follow the isohelicity of the minor groove to a greater extent and with less unfavorable interactions than **3**, thus not causing a significant deformation to

Table 2. Drug−DNA Close Contacts (<3.5 Å) of **3**−**5** in Complexes with d(CGCGAATTCGCG)$_2$[a]

drug atom	contact base	contact atom	distance (Å)
	A2T2:**4**		
N1	A6	O4′	3.9
N1	A6	N3	3.8
N1	T20	O2	2.9[b]
N1	C21	O4′	3.3
C5	T20	O2	2.9
C5	A5	C2	3.3
C6′	T19	O2	3.3
CA′	T7	C4′	3.4
C5′	T19	O4′	2.9
C5′	A18	N3	3.2
C5′	A18	C2	3.4
C5′	T8	O2	3.1
N1′	A18	O4′	3.8
N1′	A18	N3	3.4
N1′	T8	O2	3.2[b]
N1′	C9	O4′	3.7
C9	A6	O3′	3.3
C9	T7	O2P	3.9
C9	T7	C5′	4.0
C10	C21	O3′	3.8
C10	G22	C5′	4.0
C9′	A17	C4′	3.6
C9′	A17	O3′	3.6
C9′	A17	O4′	3.8
C9′	A18	C5′	3.9
	A2T2:**5**		
N1	A6	O4′	3.8
N1	A6	N3	3.3
N1	T20	O2	3.0[b]
N1	C21	O4′	3.8
C3	C21	C5′	3.3
C5	T20	O2	3.0
CA	T20	C5′	3.4
CA′	T8	C5′	3.4
C6′	T19	O2	3.4
C5′	T19	O4′	2.9
C5′	A18	N3	2.8
C5′	A18	C2	3.1
C5′	T8	O2	2.9
N1′	A18	O4′	4.0
N1′	A18	N3	3.5
N1′	T8	O2	2.9[b]
N1′	C9	O4′	3.0
C9	A6	O3′	3.6
C9′	C9	O3′	4.0
C9′	G10	O5′	3.7
C10′	A18	O3′	3.6
C10′	A18	C4′	3.7
	A2T2:**3**		
N1	A6	O4′	3.5
N1	A6	N3	3.4
N1	T20	O2	3.2[b]
N1	C21	O4′	3.5
C6	T20	O2	3.2
C6	T7	O4′	3.3
C6	A5	N3	3.3
C5	T7	O4′	3.4
C5	T20	O4′	3.3
CA′	T8	C5′	3.4
C1′	T8	O4′	3.3
C6′	T8	O4′	3.0
C6′	T19	O4′	3.3
C5′	T19	O4′	3.0
C5′	T8	O2	3.0
C4′	T19	O4′	3.2
C7′	C9	O4′	3.2
N1′	A18	O4′	3.9
N1′	A18	N3	3.8
N1′	T8	O2	3.1[b]
N1′	C9	O4′	3.0

[a] The longer contacts of interest (≤4.0 Å) have been included for the internal amidine nitrogen atoms N1 and N1′ and the alkyl substituent carbon atoms C9, C10, C9′, and C10′. [b] Indicates the thymine hydrogen bonds at the 5′ and 3′ ends.

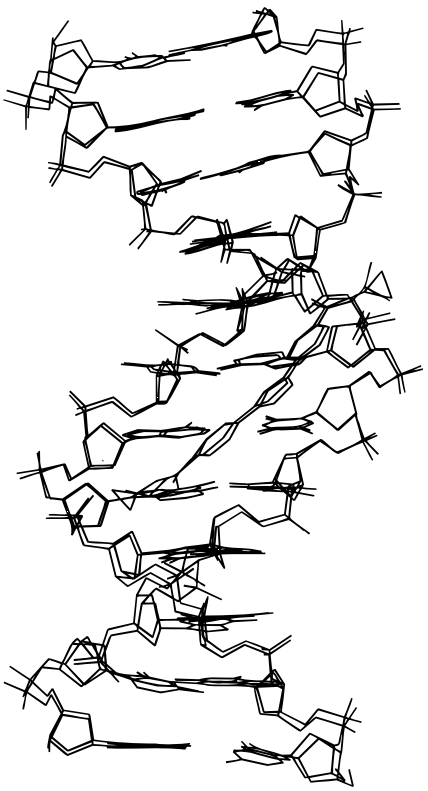

Figure 5. Superposition of all DNA atoms of the complexes A2T2:**4** and A2T2:**5.** The total rms difference is 0.39 Å.

Table 3. Propeller Twists (Calculated using CURVES) for Complexes of **3**−**5** with d(CGCGAATTCGCG)$_2$ and for the Native Dodecamer[21]

base pair	A2T2:**4**	A2T2:**5**	A2T2:**3**	native
C1-G24	−8	−13	3	−17
G2-C23	−8	−12	−8	−13
C3-G22	−12	−12	0	−4
G4-C21	−15	−16	−10	−17
A5-T20	−17	−21	−10	−27
A6-T19	−28	−24	−14	−27
T7-A18	−18	−23	−11	−24
T8-A17	−18	−14	−16	−28
C9-G16	−15	−11	−4	−25
G10-C15	−8	−8	5	−9
C11-G14	−16	−25	−18	−27
G12-C13	2	8	23	−5
average	−13	−14	−5	−18

the DNA. It is difficult to comparatively analyze groove widths for all three as width is dependent on the defining atom chosen. In spite of this complication, there are several clear trends: (i) all three structures have a wider groove width in the AT-rich region than the native dodecamer (Figure 6), (ii) the isopropyl complex produces the greatest amount of groove widening, and (iii) the cyclopropyl complex produces the least. These effects are most apparant towards the 3′ end of the AT region. It is possible to ascribe the differences produced by the drugs to their dihedral twists and the size of the substituent. The differences in groove widening produced by the isopropyl and cyclopropyl groups parallel their differences in steric volume, of 303 and 282 Å^3, respectively.

The average thermal parameters for the complexes of **3**−**5** with A2T2 (Table 4) show that the DNA follows the expected trend of phosphates having more motion than sugars, which in turn are more mobile than bases. The thermal motions of the **4** and **5** molecules in their

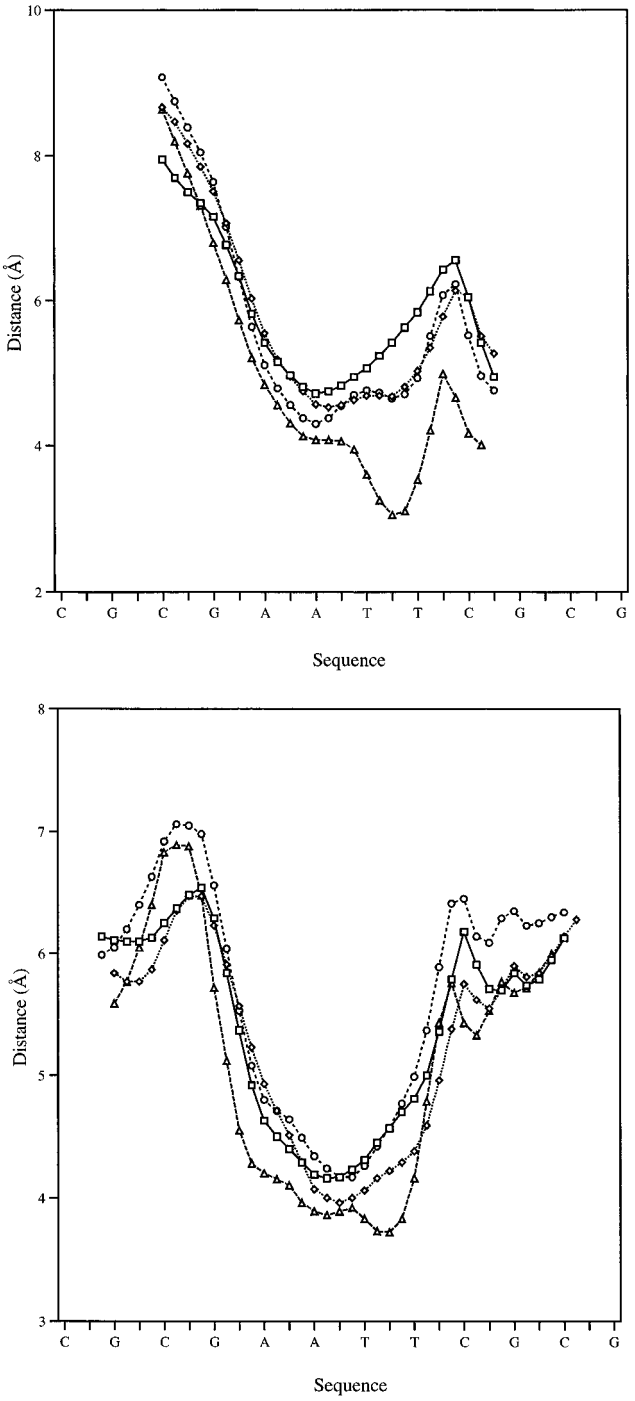

Figure 6. Minor groove widths (calculated using CURVES) for the **3**−**5** complexes with $(CGCGAATTCGCG)_2$ and the native dodecamer: (top) using P atoms to measure width and (bottom) using C1′ atoms to measure width; (□) A2T2:**4**, (◇) A2T2:**5**, (○) A2T2:**3**, and (△) native A2T2.

Table 4. Average Temperature Factors for A2T2:**4**, A2T2:**5**, and A2T2:**3** (in Å^2)

group	A2T2:**4**	A2T2:**5**	A2T2:**3**
phosphates	37.3	49.2	33.0
sugars	30.1	39.5	27.6
bases	21.6	29.7	21.9
drug	28.7	43.8	37.6
water	41.5	55.9	39.9

complexes are comparable to the average motion of the DNA (calculated by averaging the phosphate, sugar, and base temperature factors), while **3** has a higher motion, indicating a less complementary fit to the DNA and hence reduced overall binding. Ratios of the average temperature factors for the drug to DNA for **3**−**5** complexes are 1.4, 1.0, and 1.1, respectively.

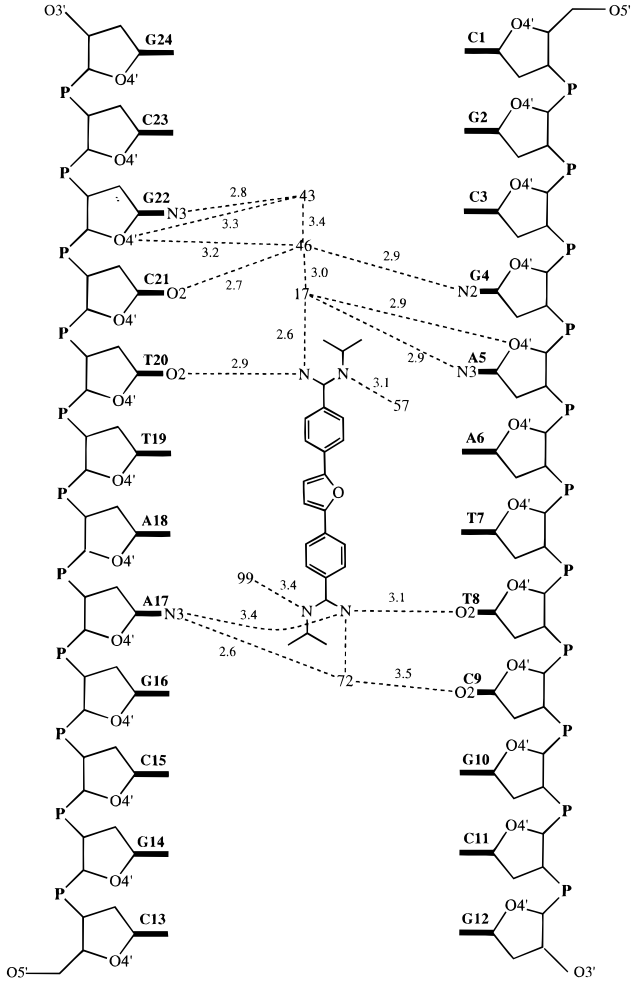

Figure 7. Schematic representation of the A2T2:**4** complex showing the drug hydrogen bonds and the associated minor groove water structure. The numbers represent the oxygen atoms of the water molecules. Distances are in angstroms.

Hydration of A2T2:4 and A2T2:5. There is extensive hydration of the DNA backbones and drug amidine groups in both structures. The amidine nitrogen atoms N2 and N2′ along the outer spine of the drug are hydrogen-bonded to water molecules as are the amidine N1 and N1′ atoms in the minor groove in the A2T2 complex with **4**. There is a tight hydrogen-bonded three-water network at the G4 end (N1 and N2 of **4**) in the minor groove that prevents the isopropyl group from being parallel to the groove. These water molecules interact with the A5 and G22 sugar O4′ and the G4, A5, C21, and G22 bases (Figure 7). At the C9 end (N1′ and N2′ of **4**) one water molecule interacts with the C9 and A17 bases with no further water network being observed in the electron density maps.

The same water network is not observed in the A2T2:**5** complex, with only the amidine nitrogen atoms N1, N2, and N1′ being hydrated by single waters, which are not involved in further hydrogen-bonded networks.

Comparison with Structures from Molecular Modeling. A molecular modeling study was performed prior to the crystallographic analysis to investigate the effect of alkyl substitution on the binding and to identify possible orientations for the drug in the minor groove.

Low-energy conformations of the separate molecules, obtained by a torsion angle (C5−C4−C7−N2, C4−C7−N2−C8, C7−N2−C8−C9, C5′−C4′−C7′−N2′, C4′−C7′−N2′−C8′, C7′−N2′−C8′−C9′) systematic search procedure, were placed in an ideal B-DNA molecule (generated by Macromodel 5.0) in a similar position to **3** in its A2T2 complex. A full continuum-solvated torsion search of the same angles, with minimization of the duplex−drug complex at each point, gave the global minima. This was checked by taking the two global minima for complexes of **4** and **5** and subjecting them to a standard mechanics/dynamics/mechanics protocol which yielded the drugs in the same orientation as **3**. The two global minima have the alkyl groups in the parallel position, which correctly predicts the A2T2:**5** complex structure but only part of the A2T2:**4** structure. The GB/SA continuum model[23] uses implicit waters, so it is not surprising that it did not correctly identify the unusual perpendicular orientation of the isopropyl group, which is tightly held in position by the network of waters in the minor groove.

Implications for Ligand Design. The three diarylfuran compounds **3**−**5** show a graduation in DNA binding abilities that parallel their *in vivo* activities against *both P. carinii* and *C. parvum* (Table 1). The isopropyl analogue **4** is consistently the superior compound in the series providing credence to the hypothesis that the minor groove drugs act by directly interfering with DNA function and processing.[4] The observation from the three X-ray crystallographic structures is that the principal structural differences between the DNA complexes of **3**−**5** are the extent of minor groove deformation and nonbonded drug−DNA contacts, with compound **4** producing the largest structural changes compared to the native DNA dodecanucleotide structure. This in turn suggests that it may primarily be the involvement of the widened DNA minor groove in ternary complexes with a regulatory or processing protein and drug that ultimately results in the therapeutic responses detailed here. The superior DNA binding of **4** is due to the larger size of the isopropyl group and consequently its greater surface area and larger number of nonbonded contacts with the minor groove surface. The involvement of a network of tightly bound minor groove water molecules in the binding of one end of the drug may also play a role in enhancing the binding.

The two structures presented in this paper reveal three features significant for the future design of ligands: (i) the minor groove can tolerate a planar orientation for the linked aromatic moieties of a drug if they are energetically favored, i.e., the drug molecule itself may not be in a global minimum but the increase in drug−DNA binding energy can compensate for this; (ii) the orientation of the amidine alkyl substituents places them in a specific location which may be used as the starting point for the design of other substituents that increase binding and extend base pair selectivity; and (iii) the water network in the A2T2:**4** complex may be exploited to devise new ligands that employ similar interactions to increase binding and base pair selectivity.

Experimental Section

Crystallography. The DNA dodecamer d(CGCGAATTCGCG)$_2$ (A2T2) was purchased from Oswell DNA Service (University of Edinburgh) and annealed before use. The drugs **4** and **5** were prepared as their hydrochloric salts.

The complexes were grown from hanging drops at 286 K as colorless prismatic blocks. The crystal of the A2T2:**4** complex used for data collection was grown from a drop containing 1 μL of 50% (w/v) 2-methylpentane-2,4-diol, 2 μL of 5 mM **4**, 3 μL of 15 mM $MgCl_2$, 1 μL of 2.5 mM spermine, and 3 μL of 2 mM A2T2 equilibrated against a reservoir containing 1 mL of 55% (w/v) 2-methylpentane-2,4-diol. The crystal of the A2T2:**5** complex used for data collection was grown from a drop containing 3 μL of 30% (w/v) 2-methylpentane-2,4-diol, 3 μL of 5 mM **5**, 3 μL of 50 mM $MgCl_2$, 3 μL of 2.5 mM spermine, and 3 μL of 2 mM A2T2 equilibrated against a reservoir containing 1 mL of 50% 2-methylpentane-2,4-diol. The A2T2 solution was prepared using 30 mM sodium cacodylate buffer at pH 7.0. The crystals employed for the X-ray studies were obtained after about 3 months of growth.

Crystallographic Data Collection. The colorless crystals of A2T2 complexes of **4** and **5** used for data collection were of approximate dimensions 0.18 × 0.18 × 0.12 and 0.60 × 0.28 × 0.26 mm, respectively, and were mounted in 0.5 mm Lindemann glass capillaries with small amounts of mother liquor. Intensity data were collected at 287 K using a Siemens-Xentronics multiwire area detector with a rotating anode X-ray generator and a graphite monochromator. A crystal-to-detector distance of 10 cm and a swing angle of 15° were used to collect data to a maximum possible resolution of 2.2 Å. Data were collected with χ set at 45°, while the crystal was rotated through 100° in ω at ϕ values of 0° and 60°; 180 s frames were recorded every 0.20° step. The crystals did not suffer any observable decay during the data collection. Data processing was carried out using the program package XENGEN, version 1.3. After merging, the data for the complex with **4** comprised 3146 of a possible 3592 unique reflections to 2.2 Å (87.6%) with a merging R value of 7.3%, and the data for **5** comprised 3427 of a possible 3748 unique reflections to 2.2 Å (91.4%) with a merging R value of 8.8%.

Structure Refinement. The unit cell dimensions of the A2T2 complexes of **4** and **5** were a = 24.60 Å, b = 40.07 Å, c = 65.45 Å, and a = 25.43 Å, b = 40.66 Å, c = 66.13 Å, respectively, in the space group $P2_12_12_1$. These cells are close to that reported for the native d(CGCGAATTCGCG)$_2$ dodecamer (Dickerson and Drew, 1981; a = 24.87 Å, b = 40.39 Å, c = 66.2 Å) and other groove-bound dodecamer−drug complexes, suggesting that the present crystals are isomorphous with them. The dodecamer coordinates used as starting models for the structure refinements were those of the native d(CGCGAATTCGCG)$_2$ structure.[21]

The crystallographic refinement was carried out using the program X-PLOR,[24] version 3.1. Rigid body refinement of the DNA molecule as one constrained group was performed with the resolution range of the data increased from 8.0−4.0 Å (522 and 566 reflections) to 8.0−3.5 Å (804 and 874 reflections) for data for A2T2:**4** and A2T2:**5**, respectively. The resolution range was gradually increased from 8.0−3.5 to 8.0−2.2 Å (2591 and 2817 reflections), the R factors being 26.6% and 25.1% for A2T2:**4** and A2T2:**5**, respectively. Electron density maps were calculated and displayed using the graphics package TOM/FRODO,[25] version 3.2.

The DNA molecules fitted the density well, and a lobe of density was clearly visible in the minor groove in each structure. The lobes of density were in the positions predicted for both **4** and **5** from molecular modeling using Macromodel,[26] version 4.5. In the case of the structure of A2T2:**4**, there was a second small lobe of density in the minor groove close to the large lobe, and all possible orientations of the alkyl amidine group were fitted to it. Electrostatic charges for **4** and **5** were calculated by MOPAC,[27] version 6.0, using the electrostatic potential option. Force field parameters were interpolated from previous molecular modeling studies in this laboratory.

The drug molecules were included, and the refinement was repeated stepwise from low- to high-resolution data. The R factor at this point for the structure A2T2:**5** was 22.7%. It was possible to unambiguously determine the orientation of the alkyl amidine group for structure A2T2:**4** by the difference

in *R* factors, 24.2% for the alkyl group oriented deep in the minor groove and 23.8% for the alkyl group in an opposing orientation, at the mouth of the groove. The later was used for further refinement.

Solvent positions were included toward the end of the refinement and assigned as water molecules. The criteria for acceptance of solvent were proximity to the DNA−drug complex (within 4 Å), peak height > 3σ in difference maps, potential hydrogen-bonded neighbors (2.2−3.4 Å), and sensible thermal parameters.

At the end of the refinements a total of 81 and 51 water molecules had been included and gave final *R* factors of 16.9% and 18.6% for structures A2T2:**2** and A2T2:**3**, respectively, for all data in the range 8.0−2.2 Å. Coordinates and structure factors have been submitted to the Nucleic Acid Data Base (identification numbers GDL044 and GDL045).

Molecular Modeling. The conformation spaces of **4** and **5** were searched extensively by rotating torsion angles OA−CA−C1−C6, C5−C4−C7−N1, and C7−N1−C8−C9 in 5° increments through 360°. Macromodel, v4.5, running on an SGI Challenge Workstation, using the modified AMBER* force field with a distance-dependent dielectric constant of 4, gave very similar global minima structures for the two drugs in the minor groove. The AMBER* structures of **4** and **5** were placed in a generated ideal B-form d(CGCGAATTCGCG)$_2$ duplex using the position suggested by the crystal structure of A2T2:**3**. A torsion angle search and minimization using AMBER*, similar to the previous study on the drugs alone but eliminating the bad-contact structures, was used to obtain the starting structure for molecular dynamics calculations. A 40 ps molecular dynamics run at 300 K with a 1.5 fs time step was sampled 100 times, and the average structure was further minimized using a standard protocol.

Synthesis. 2,5-Bis{[4-(*N*-isopropyl)amindino]phenyl}furan Dihydrochloride (4). Dry isopropylamine (0.47 g, 0.008 mol) was added to a suspension of imidate ester (1.3 g, 0.003 mol), prepared as previously described[20] from 2,5-bis(4-cyanophenyl)furan,[28] in 45 mL of absolute ethanol. Within 0.5 h the imidate ester dissolved. After ca. 3 h a white solid precipitated; the slurry was stirred overnight at room temperature. The solvent was removed under reduced pressure, diluted with water, filtered, and washed with water. After the solid was dried, it was recrystallized from an ethanol/ether mixture to yield a white solid (0.9 g, 78%): mp 233−4 °C; IR (KBr) 3249, 3069, 2997, 1600, 1383, 1210 cm^{-1}; ^{1}H NMR (DMSO-d_6/60 °C) δ 7.79 (brs, 8H), 7.11 (s, 2H), 6.25 (br, 4H), 3.81 (br, 2H), 1.14 (d, 6H, J = 5.9 Hz); ^{13}C NMR (DMSO-d_6/60 °C) δ 159.6, 152.4, 136.6, 130.4, 126.8, 122.8, 108.7, 43.5, 22.8; MS *m/e* 388(M$^+$).

The free base (0.78 g, 0.002 mol) was dissolved in 10 mL of absolute ethanol, treated with 10 mL of ethanol saturated with hydrogen chloride, and warmed for 2 h. The mixture was reduced in volume to 5 mL. Addition of 20 mL of dry ether produced a bright yellow precipitate which was filtered, washed with 3 × 5 mL of dry ether, and dried *in vacuo* at 65 °C for 2 h to yield 0.8 g (87%): mp 276−7 °C dec; IR (KBr) 3407, 3132, 3065, 1669, 1611, 1388, 1128 cm^{-1}; ^{1}H NMR (DMSO-d_6) δ 9.72 (s, lH), 9.69 (s, lH), 9.57 (s, 2H), 9.24 (s, 2H), 8.06 (d, 4H, J = 8.1 Hz), 7.86 (d, 4H, J = 8.1 Hz), 7.47 (s, 2H), 4.14 (sep, 2H, J = 6.6 Hz), 1.29 (d, 12H, J = 6.6 Hz); ^{13}C NMR (DMSO-d_6) δ 161.1, 152.3, 133.6, 129.2, 127.7, 123.5, 111.3, 45.1, 2l.l. Anal. Calcd ($C_{24}H_{28}N_4O$·2HCl·1.25H_2O): C, 59.57; H, 6.76; N, 11.57. Found: C, 60.00; H, 6.80; N, 11.52.

2,5-Bis{[4-(*N*-cyclopropylamidino)phenyl]}furan Dihydrochloride (5). The free base was prepared from cyclopropylamine and the imidate ester as described for **4** in a yield of 79%: mp 185−186 °C dec; IR (KBr) 3464, 3320, 3080, 1610, 1510, 1364, 1022, 848, 791 cm^{-1}; ^{1}H NMR ($CDCl_3$) δ 7.71 (brs, 8H), 6.78 (s, 2H), 5.3 (v br, 4H), 2.6 (brm, 2H), 0.87−0.81 (m, 4H), 0.67−0.62 (m, 4H); ^{13}C NMR ($CDCl_3$ + DMSO-d_6) δ 159.6, 152.2, 134.8, 130.7, 126.4, 122.6, 107.7, 25.7, 6.04; MS *m/e* 384 (M$^+$).

The free base was converted to the salt in a yield of 80% (yellow solid): mp >310 °C dec; IR (KBr) 3369, 3181, 3037, 1665, 1607, 1502, 1032, 782, 674 cm^{-1}; ^{1}H NMR (DMSO-d_6) δ 10.24 (s, 2H), 9.86 (s, 2H), 9.27 (s, 2H), 8.06 (d, 4H, J = 7.94 Hz), 7.95 (d, 4H, J = 8.54 Hz), 7.42 (s, 2H), 2.87 (brm, 2H), 1.09−0.85 (m, 8H); ^{13}C NMR (DMSO-d_6) δ 163.9, 152.3, 133.7, 129.1, 126.6, 123.5, 111.3, 24.7, 6.5. Anal. Calcd ($C_{24}H_{24}N_4O$·2HCl): C,63.02; H, 5.73; N, 12.25. Found: C, 62.89; H, 5.95; N, 12.00.

Acknowledgment. This work was supported by the Cancer Research Campaign Programme Grant SP1384 (to S.N.) and by NIH Grant AI-33363 (to D.W.B.).

References

(1) Brown, D. G.; Sanderson, M. R.; Garman, E.; Neidle, S. Crystal structure of a berenil-d(CGCGAAATTTCGCG)$_2$ complex: An example of drug-DNA recognition based on sequence-dependent structural features. *J. Mol. Biol.* **1992**, *226*, 481−490.

(2) Nunn, C. M.; Jenkins, T. C.; Neidle, S. Crystal structure of d(CGCGAATTCGCG)$_2$ complexed with propamidine, a short-chain homologue of the drug pentamidine. *Biochemistry* **1993**, *32*, 13838−13843.

(3) Edwards, K. J.; Jenkins, T. C.; Neidle, S. Crystal structure of a pentamidine-oligonucleotide complex: Implications for DNA-binding properties. *Biochemistry* **1992**, *31*, 7104−7109.

(4) (a) Welch, J. J.; Rauscher, F. J., III; Beerman, T. A. Targeting DNA-binding drugs to sequence-specific transciption factor DNA complexes. *J. Biol. Chem.* **1994**, *269*, 31051−31058. (b) Wong, S. S. C.; Sturm, R. A.; Michel, J.; Zhang, X.-M.; Danoy, P. A. C.; Mcgregor, K.; Jacobs, J. J.; Kaushal, A.; Dong, Y.; Dunn, I. S.; Parsons, P. G. Transcriptional regulation of differentiation, selective toxicity and ATGCAAAT binding of bisbenzimidazole derivatives in human melanoma cells. *Biochem. Pharmacol.* **1994**, *47* (5), 827−837. (c) Henderson, D.; Hurley, L. H. Molecular struggle for transcription control. *Nature Med.* **1995**, *1*, 525−527. (d) Mote, J., Jr.; Ghanouni, P.; Reines, D. A minor groove-binding ligand both potentiates and arrests transcription by RNA polymerase II. *J. Mol. Biol.* **1994**, *236*, 725−737.

(5) Chiang, S.-Y.; Welch, J.; Rauscher, F. J., III; Beerman, T. A. Effects of minor groove binding drugs on the interaction of TATA box binding protein and TFIIA with DNA. *Biochemisty* **1994**, *33*, 7033−7040.

(6) (a) Chen, A. Y.; Yu, C.; Gatto, B.; Liu, L. F. DNA minor groove-binding ligands: A different class of mammalian DNA topoisomerase I inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8131−8135. (b) Beerman, T. A.; McHugh, M. M.; Sigmund, R.; Lown, J. W.; Rao, K. E.; Bathini, Y. Effects of analogues of the DNA minor groove binder Hoechst 33258 on topoisomerase II and I mediated activities. *Biochim. Biophys. Acta* **1992**, *1131*, 53−61.

(7) (a) Sun, Q.; Gatto, B.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. Structure activity of topoisomerase I poisons related to Hoechst 33342. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2871−2876. (b) Atwell, G. J.; Yaghi, B. M.; Turner, P. R.; Boyd, M.; O'Conner, C. J.; Ferguson, L. R.; Baguley, B. C.; Denny, W. A. Synthesis, DNA interactions and biological activity of DNA minor groove targeted polybenzamide-linked nitrogen mustards. *Bioorg. Med. Chem.* **1995**, *3*, 679−691. (c) Gravatt, G. L.; Baguley, B. C.; Wilson, W. R.; Denny, W. A. DNA-directed alkylating agents. 6. Synthesis and antitumor activity of DNA minor groove-targeted aniline mustard analogues of Pibenzimol (Hoechst 33258). *J. Med. Chem.* **1994**, *37*, 4338−4345.

(8) (a) D'Incalci, M. DNA-minor-groove alkylators, a new class of anticancer agents. *Ann. Oncol.* **1994**, *5*, 877−878. (b) Bellorini, M.; Moncollin, V.; D'Incalci, M.; Mongelli, N.; Mantovani, R. Distamycin A and tallimustine inhibit TBP binding and basal *in vitro* transcription. *Nucleic Acids Res.* **1995**, *23*, 1657−1663.

(9) Wyatt, M. D.; Lee, M.; Garbiras, B. J.; Souhami, R. L.; Hartley, J. A. Sequence specificity of alkylation for a series of nitrogen mustard-containing analogues of distamycin of increasing binding site size: evidence for increased cytotoxicity with enhanced sequence specificity. *Biochemistry* **1995**, *34*, 13034−13041.

(10) (a) Tidwell, R. R.; Jones, S. K.; Geratz, J. D.; Ohemeng, K. A.; Cory, M.; Hall, J. E. Analogues of 1,5-bis(4-amidinophenoxy)-pentane (pentamidine) in the treatment of experimental *Pneumocystis carinii* pneumonia. *J. Med. Chem.* **1990**, *33*, 1252−1257. (b) Cory, M.; Tidwell, R. R.; Fairley, T. A. Structure and DNA binding activity of analogues of 1,5-bis(4-amidinophenoxy)-pentane (pentamidine). *J. Med. Chem.* **1992**, *35*, 431−438. (c) Donker, I. O.; Tidwell, R. R.; Jones, S. K. Pentamidine congeners. 2. 2-Butene-bridged aromatic diamidines and diimidazolines as potential anti- *Pneumocystis carinii* pneumonia agents. *J. Med. Chem.* **1994**, *37*, 4554−4557.

(11) Tidwell, R. R.; Jones, S. K.; Naiman, N. A.; Berger, L. C.; Brake, W. B.; Dykstra, C. C.; Hall, J. E. Activity of cationically substituted bis-benzimidazoles against experimental *Pneumocystis carinii* pneumonia. *Antimicrob. Agents Chemother.* **1993**, *37*, 1713−1716.

(12) (a) Nunn, C. M.; Jenkins, T. C.; Neidle, S. Crystal structure of γ-oxapentamidine complexed with d(CGCGAATTCGCG)$_2$. The effects of drug structural change on DNA minor-groove recognition. *Eur. J. Biochem.* **1994**, *226*, 953−961. (b) Nunn, C. M.; Neidle, S. Sequence-dependent drug binding to the minor groove of DNA: crystal structure of the DNA dodecamer d(CGCAAATTTGCG)$_2$ complexed with propamidine. *J. Med. Chem.* **1995**, *38*, 2317−2325.

(13) (a) Fox, K. R.; Sansom, C. E.; Stevens, M. F. G. Footprinting studies on the sequence-selective binding of pentamidine to DNA. *FEBS Lett.* **1990**, *266*, 150−154. (b) Greenidge, P. A.; Jenkins, T. C.; Neidle, S. DNA minor groove recognition properties of pentamidine and its analogues: A molecular modeling study. *Mol. Pharmacol.* **1993**, *43*, 982−988. (c) Jenkins, T. C.; Lane, A. N.; Neidle, S.; Brown, D. G. NMR and molecular modeling studies for the interaction of berenil and pentamidine with d(CGCAAATTTGCG)$_2$. *Eur. J. Biochem.* **1993**, *213*, 1175−1184.

(14) Boykin, D. W.; Kumar, A.; Spychala, J.; Zhou, M.; Lombardy, R. J.; Wilson, W. D.; Dykstra, C. C.; Jones, S. K.; Hall, J. E.; Tidwell, R. R.; Laughton, C.; Nunn, C. M.; Neidle, S. Dicationic diarylfurans as Anti-*Pneumocystis carinii* Agents. *J. Med. Chem.* **1995**, *36*, 912−916.

(15) Laughton, C.; Tanious, F.; Nunn, C. M.; Boykin, D. W.; Wilson, W. D.; Neidle, S. A crystallographic and spectroscopic study of the complex between d(CGCGAATTCGCG)$_2$ and 2,5-bis(4-guanylphenyl)furan and an analogue of berenil. Structural origins of enhanced DNA-binding affinity. *Biochemistry* **1996**, *35*, 5655−5661.

(16) Crothers, D. M. Statistical thermodynamics of nucleic acid melting transtions with coupled binding equilibria. *Biopolymers* **1971**, *10*, 2147−2160.

(17) Czarny, A.; Boykin, D. W.; Wood, A. A.; Nunn, C. M.; Neidle, S.; Zhao, M.; Wilson, W. D. Analysis of van der Waals and electrostactic contributions in the interactions of minor groove binding benzimidazoles with DNA. *J. Am. Chem. Soc.* **1995**, *117*, 4716−4718.

(18) Blagburn, B. L.; Sundermann, C. A.; Lindsay, D. S.; Hall, J. E.; Tidwell, R. R. Inhibition of *Cryptosporidium parvum* in neonatal Hsd:(lCR)BR swiss mice by polyether ionophores and aromatic amidines. *Antimicrob. Agents Chemother.* **1991**, *35*, 1520−1523.

(19) (a) Lombardy, R. L.; Tanious, F.; Ramachandran, K.; Tidwell, R. R.; Wilson, W. D. Synthesis and DNA interactions of benzimidazole dications which have activity against opportunistic infections. *J. Med. Chem.* **1996**, *39*, 1452−1462. (b) Tidwell, R. R.; Jones, S. K.; Naiman, N. A.; Berger, L. C.; Brake, W. B.; Dykstra, C. C.; Hall, J. E. Activity of cationically substituted bis-benzimidazoles against experimental *Pneumocystis carinii* pneumonia. *Antimicrob. Agents Chemother.* **1993**, *37*, 1713−1716.

(20) Lavery, R.; Sklenar, H. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91.

(21) Drew, H. R.; Samson, S.; Dickerson, R. E. Structure of a B-DNA dodecamer at 16-K. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4040−4044.

(22) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.* **1992**, *63*, 751−759.

(23) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.

(24) Brunger, A. T.; Kuriyan, J.; Karplus, M. Crystallographic *R*-factor refinement by molecular dynamics. *Science* **1987**, *235*, 458−460.

(25) Cambillau, C.; Horjales, E. TOM: a FRODO subpackage for protein-ligand fitting with interactive energy minimization. *J. Mol. Graph.* **1987**, *5*, 174−177.

(26) MacroModel V4.5: Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. Macromodel - an integrated software system for modeling organic bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **1990**, *11*, 440−467.

(27) Stewart, J. P. P. Mopac 6.0 (QCPE), available from the Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN 47405.

(28) Bajic, M.; Kumar, A.; Boykin, D. W. Synthesis of 2,5-Bis(4-cyanophenyl)furan. *Heterocycl. Commun.*, in press.

JM9604484