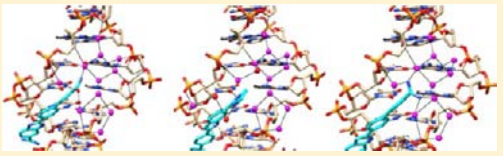
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Small-molecule Binding to the DNA Minor Groove Is Mediated by a Conserved Water Cluster

DengGuo Wei,[†] W David Wilson,[§] and Stephen Neidle[†,*]

[†]UCL School of Pharmacy, 29-39 Brunswick Square, London WC1N 1AX, United Kingdom
[§]Department of Chemistry, Georgia State University, Atlanta, Georgia 30303, United States

ABSTRACT: High-resolution crystal structures of the DNA duplex sequence d(CGCGAATTCGCG)$_2$ complexed with three minor-groove ligands are reported. A highly conserved cluster of 11 linked water molecules has been found in the native and all 3 ligand-bound structures, positioned at the boundary of the A/T and G/C regions where the minor groove widens. This cluster appears to play a key structural role in stabilizing noncovalently binding small molecules in the AT region of the B-DNA minor groove. The cluster extends from the backbone phosphate groups along the mouth of the groove and links to DNA and ligands by a network of hydrogen bonds that help to maintain the ligands in position. This arrangement of water molecules is distinct from, but linked by, hydrogen bonding to the well-established spine of hydration, which is displaced by bound ligands. Features of the water cluster and observed differences in binding modes help to explain the measured binding affinities and thermodynamic characteristics of these ligands on binding to AT sites in DNA.



INTRODUCTION

The hydration of nucleic acids is central to their structural stability.[1] In the case of B-form DNA, the existence of a well-ordered array of water molecules in AT-rich regions (the "spine of hydration") has been demonstrated initially by crystallographic means,[2,3] and subsequently by a range of biophysical[4] and computational methods.[5−7] The spine is disrupted and effectively displaced when small molecules of appropriate size and shape bind in the minor groove of B-DNA.[8,9] We report here that in addition to the spine there exists in and adjacent to the minor groove a well-structured cluster of eleven water molecules, which are preserved on the noncovalent binding of linked heterocyclic small molecules. This cluster plays a key role in mediating between ligand and DNA by means of an array of hydrogen bonds.

The biological properties of many of these DNA-binding small molecules, especially the clinically important antiparasitic activity shown by a number of compounds in this general class, are generally related to their affinities for duplex DNA.[10−14] A number of studies have examined the molecular basis of their interactions with DNA.[9,15−20] A structural requirement for these ligands to possess curvature complementing the minor groove concavity, "isohelicity", was initially postulated as an absolute requirement,[8] especially for strong-binding ligands such as the oligoamide compounds netropsin and distamycin. Compounds with either too great or insufficient curvature in general show a dramatic reduction in DNA binding affinity. The isohelical concept has been supported by a large number of cocrystal structures involving ligand and dodecanucleotide duplex sequences especially d(CGCGAATTCGCG)$_2$ (A2T2); see for example refs 15−18, 20, 21. The isohelicity requirement for strong binding was subsequently challenged by the findings from biophysical studies,[19,22] which demonstrated that the nonisohelical linear heterocyclic molecule DB921 (Figure 1, Table 1) in a complex with A2T2 shows surprisingly strong binding affinity to the DNA minor groove. This was found in a parallel crystallographic analysis (PDB ID 2B0K) to be due to the existence of a single bound water molecule that mediates between the amidinium group of the DB921 molecule and the N3 atom of A5 in the A2T2 sequence.[22] This bridging water molecule in effect completes the curvature of the bound ligand, resulting in a stable complex.

The crystal structure at 2.05 Å of the A2T2 complex with the similar but isohelical compound DB911 (PDB ID: 2NLM) has shown that this binds in an inverted orientation compared to DB921.[23] Closer complementation with the DNA minor groove is achieved by DB921 with the inner-facing nitrogen atom of the benzimidazole ring anchoring the ligand by a pair of bifurcated hydrogen bonds to the O2 atoms of T7 and T19. The terminal amidinium group attached to the benzimidazole further stabilizes the binding by hydrogen bonding to the O2 atom of T8. DB911 is not able to participate in any of these interactions and hence its DNA binding is much weaker.

We have determined the crystal structures of several new complexes of A2T2 with three ligands related to DB921 (Figure 1: DB1804, DB1883 and DB1963). These ligands incorporate systematic modifications to the terminal recognition groups. Quantitative binding data for these (Table 1) has been obtained by surface plasmon resonance techniques.[22] All three compounds appear to have closely similar linearity compared to DB921 and would be expected to have closely similar affinities for A2T2, yet the experimental data indicates substantial differences. Compound DB1963, with an amide replacing the amidinium group, has duplex DNA binding affinity less than one-third of that of DB921. This is a surprising

Received: September 12, 2012
Published: December 31, 2012

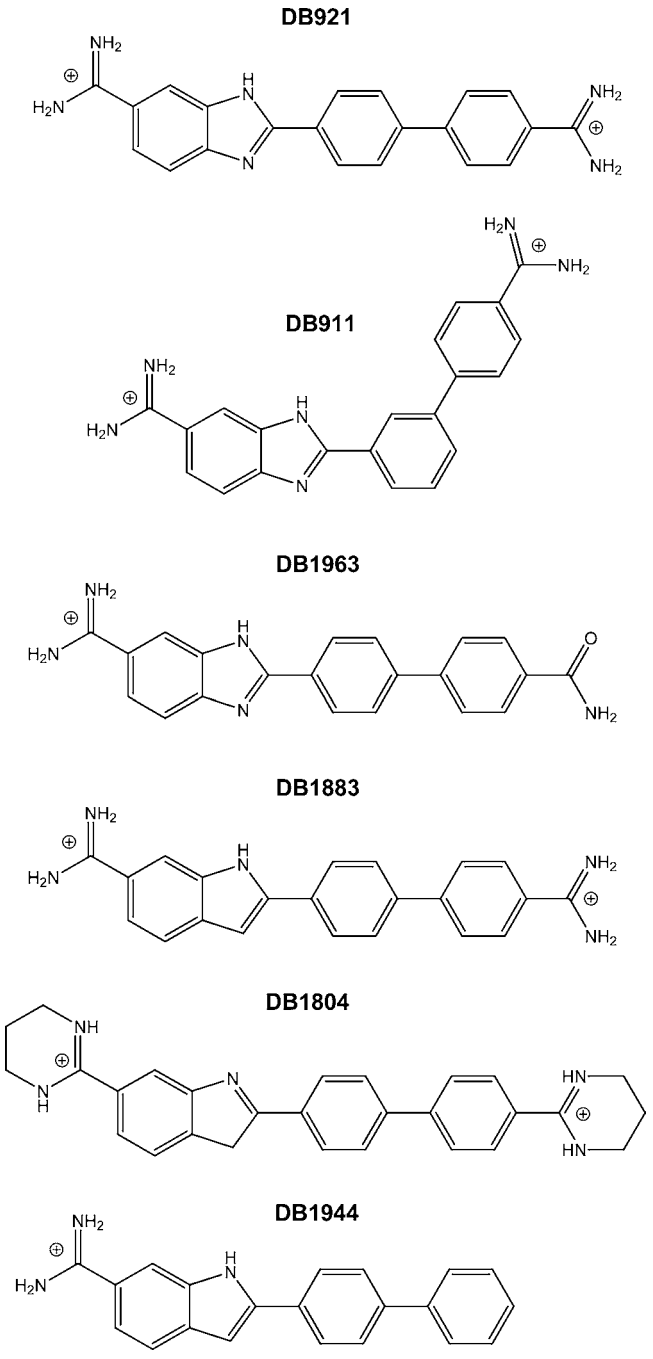


Figure 1. Structures of the minor-groove ligands analyzed in this study.

difference, because there is no major structural difference between the capability of amidinium and amide groups to hydrogen-bond with a single bridging water molecule, although one factor could be the higher basicity of the amidinium group and consequent enhanced electrostatics contributions to binding affinity. However the 3-fold higher affinity of DB1804 compared to DB921 suggests that electrostatics considerations alone do not explain these differences since the terminal modifications, to a carbocyclic structure, still maintain the high p*K* values of the amidines. Indeed, on the basis of the earlier crystal structure of the A2T2:DB921 complex, the cyclic amidine terminal rings of DB1804 were assumed to interfere with water-mediated binding on steric grounds; however, this compound has the highest affinity in the series. Removal of one nitrogen atom in the benzimidazole ring results in the indole-containing compound DB1883, with almost twice the A2T2 affinity of DB921. Removal of one terminal group (DB1944) results in a complete loss of A2T2 affinity. An initial goal of this study was to address these questions in a fundamental manner with high-resolution structural data. Subsequent analysis has revealed an extensive structured water arrangement as a more general feature of this DNA and its minor-groove ligand complexes, which is the major focus of this manuscript.

MATERIALS AND METHODS

General. Syntheses of compounds DB1883 and DB1055 (Figure 1) have been published, and the syntheses of compounds DB1963 and DB1804 will be published elsewhere. NMR and elemental analysis were used to verify compound purity. DB1883, DB1963 and DB1804 were used as the hydrochloride salts and DB1055 as the acetate salt.

X-ray Crystallography. The HPLC-purified oligonucleotide d(CGCGAATTCGCG) was purchased from Eurofins MWG Operon (Germany) and used without further purification. A 6 mM single-stranded DNA solution in 20 mM sodium cacodylate buffer at pH 7.0 was heated to 358 K before annealing by slow cooling to room temperature. Crystals were grown by the hanging drop vapor diffusion method. One μL of a premixed drop solution containing 6% MPD, 60 mM magnesium chloride, and 40 mM sodium cacodylate at pH 6.5, was added to a 1 μL solution of 0.75 mM A2T2 DNA and 1.0 mM of a ligand solution in water. The drop solution was equilibrated against a 50% MPD well solution at 283 K. Suitable complex crystals took up to 20 days to appear. The DB1883 and DB1804:A2T2 complexes form rod-like colorless crystals, whereas those of DB1963:A2T2 form yellow rods. Both rod-like and boulder-like colorless crystals were obtained for the DB1055:A2T2 complex, albeit in different drops. All crystals were flash-frozen in liquid nitrogen, and diffraction data sets were collected at 105 K at the Diamond synchrotron facility (UK). Unit cell parameters and data collection information are given in Table 2. Indexing and data processing was carried out automatically using the XIA2 package.[24] For the DB1963, DB1883 and DB1804 complexes, the resolution is ~1.20 Å. The crystals of the DB1055:A2T2 complex diffracted to 1.4 and 1.47 Å, respectively, compared to 1.65 Å in an earlier study (PDB ID: 2I5A). However the ligand in these more recent data sets was found to be disordered, occupying two indistinct orientations.

Structure Solution and Refinement. Structures were solved by molecular replacement using the PHASER program[25] from the CCP4 package,[24] and the DNA structure extracted from the complex structure (PDB ID: 2B3E) was used as the starting point. Electron density and difference maps were used to locate the position of the ligands, magnesium ions and water molecules. REFMAC[26] from the CCP4 package was used in the refinement, and the crystallographic data is detailed in Table 2. Electron density maps were visualized with the COOT program[27] and structures analyzed and drawn with the UCSF CHIMERA program[28] and the PYMOL programme (www.pymol.org).

dx.doi.org/10.1021/ja308952y | J. Am. Chem. Soc. 2013, 135, 1369−1377

Table 1. Thermodynamic Characteristics of the Minor Groove Ligands Reported Here[a]

ligand	K_a ($\times 10^7$)	ΔG (kcal/mol)	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	PBD id	resolution (Å)
DB921	14.2	−11.1	−4.5	−6.6	2B0K	1.64
DB911	1.2	−9.6	−2.6	−7.0	2NLM	2.05
DB1963	3.9	−10.3	−1.7	−8.6	***3U08***	***1.25***
DB1883	24.6				***3U0U***	***1.24***
DB1804	48.2				***3U05***	***1.27***
DB1944	<0.01					
A2T2	n/a				***3U2N***	***1.25***

[a] K_a and ΔG values were determined by surface plasmon resonance;[21] ΔH values were determined by ITC.[21] $-T\Delta S$ and ΔG values were derived from these ΔH values. PDB numbers in bold-italics refer to crystal structures reported here, together with the maximum resolution of the refined structures. Structures with PDB numbers in normal typeface have been previously determined; that of DB911 has been deposited but not reported in detail.[23] No structural data is currently available for a DB1944:A2T2 complex. There is no ITC data for compounds DB1883, 1804, and 1944 due to aggregation problems during ITC determinations.

Table 2. Crystallographic Data Collection and Refinement Statistics

	DB1883:A2T2	DB1963:A2T2	DB1804:A2T2	Native A2T2
		Data Collection		
Space group		$P2_12_12_1$		
Cell; dimensions a, b, c (Å)	23.94, 39.69, 65.64	24.07, 39.76, 65.70	24.10, 39.88, 65.79	23.92, 39.64, 65.39
Wavelength (Å)	0.9796	0.9796	0.9796	0.9795
Resolution (Å)	1.24−39.69	1.25−24.07	1.27−39.9	1.25−39.64
R_{merge}	0.034	0.033	0.037	0.044
I/σ	19.2	19.9	21	15.1
Completeness (%)	92.9	90.3	95.2	97.1
Redundancy	4.1	4.2	6.1	4.4
Total no. of reflections	68335	68294	100697	77025
No. of unique reflections	16826	16455	16503	17420
		Refinement		
Resolution (Å)	1.24−15.16	1.25−15.33	1.27−34.1	1.25−15.26
No. of reflections	15901	15560	15596	16498
R_{work}/R_{free}	0.21/0.242	0.203/0.243	0.200/0.236	0.193/0.239
no. of Mg^{2+} ions	3	2	1	9
no. of water molecules	142	149	154	215
overall B factor (Å^2)	18.21	18.46	19.24	18.92
<rmsd> in bond length (Å)	0.012	0.009	0.009	0.007
<rmsd> in bond angle (°)	1.236	1.347	1.10	1.214
PDB ID	3U0U	3U08	3U05	3U2N

RESULTS

The availability of near-atomic resolution high-quality synchrotron X-ray data (collected on Diamond beamline I24) of the A2T2-ligand complexes has enabled these structures to be examined in detail. In order to ensure meaningful comparisons with the native A2T2 structure (which has been extensively studied previously[1−3]), the water arrangement in the native structure has been redetermined with high-resolution synchrotron data collected and refined in the same manner as for the complexes themselves. All structures crystallize in the space group $P2_12_12_1$ and are isomorphous with a large number of existing A2T2 native and ligand complex structures[10] (Table 2). The structures were solved by molecular replacement.

The initial ($F_o - F_c$) difference electron-density maps calculated immediately following the refinement of A2T2 alone in each structure, are shown in Figures 2a−d. The native structure displays the characteristic spine of hydration,[2] with the first and second shell of waters being observed, as in earlier high-resolution crystal structures.[3] Comparison of all three ligand complexes and the native structure shows that they closely overlay each other (Figure 3). These are heavily hydrated structures and a pronounced cluster of water molecules is apparent, which is also readily visible when comparing the difference electron density maps for each structure (Figures 4a−d). Comparative analysis of these maps has revealed a large group of water molecules hydrating the DNA and the ligand-DNA interface in the minor groove, which form a conserved water network around one of the termini of these ligands. This involves 11 water molecules in each of the native A2T2, DB1963:A2T2, DB1883:A2T2 and DB1804:A2T2 structures. All of these waters are common to, and therefore conserved in the structures (Tables 3, 4). In each complex the interactions between ligand terminus and the hydrogen bond acceptors/donors on the DNA are mediated by the solvent water cluster rather than by a single water molecule, as was suggested by the earlier lower-resolution crystallographic data.

The water arrangement is linked into, but spatially distinct from, the spine of hydration and its second coordination shell of waters. Figures 4a−d and 5a−d show that the cluster of eleven conserved waters comprises two adjacent groups of water molecules. One group, comprising W1−W4, effectively links the end of the spine of hydration at the 5′ end of the duplex, to the phosphate group of T7, forming a pentagonal arrangement of hydrogen bonds at the mouth of the groove.

dx.doi.org/10.1021/ja308952y | J. Am. Chem. Soc. 2013, 135, 1369−1377

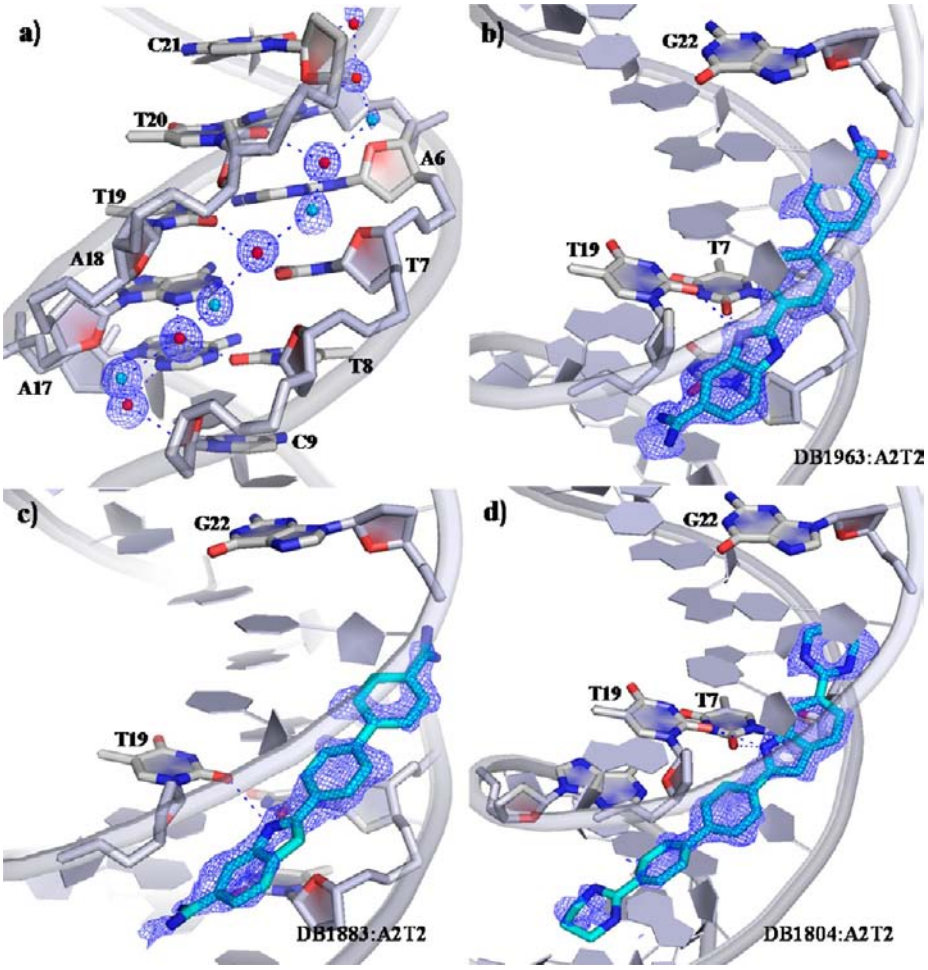


Figure 2. Initial, unbiased difference electron density calculated with only DNA in the structures: (a) Native A2T2 DNA; (b) DB1963:A2T2; (c) DB1883:A2T2; (d) DB1804:A2T2. The density in Figure 1a is contoured at the 1.5 σ level, and that in the other three figures are contoured at the 1.0 σ level. The view in each figure is looking into the minor groove. In (a), the water molecules hydrogen-bonding to DNA in the minor groove are shown as red spheres, and the bridging water molecules are colored cyan. DNA bases and sugars are drawn in stick format and the backbones are in ribbon representation.

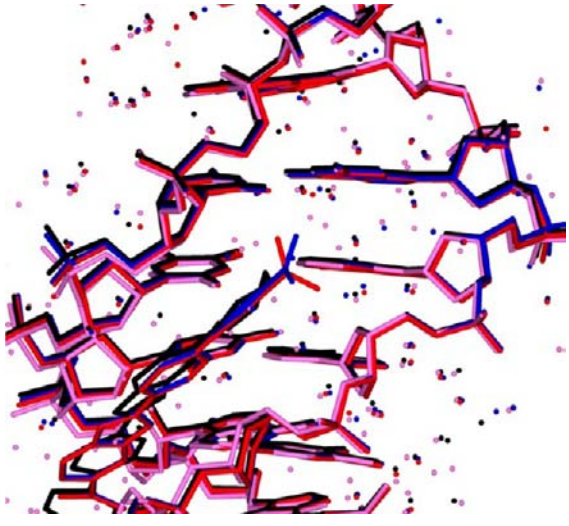


Figure 3. Overlay of the three ligand complexes and the native A2T2 structure. The native A2T2 structure is colored pink, DB1804:A2T2 is black, DB1883:A2T2 is red and DB1963:A2T2 is blue.

Water W5 is hydrogen-bonded to the phosphate group of G22 on the complementary DNA strand at the point where the minor groove starts to widen at the start of the G/C sequence. In the native structure W5 is also positioned at the groove mouth so that it is linked to the spine of hydration (Figures 4a, 5a). In the DB1883:A2T2 and DB1963:A2T2 structures (Figure 5b, c), water W5 plays a key role in bridging between this phosphate group and the amide/amidinium group of the ligand. It is balanced by the W1−W4-phosphate pentagon, which also hydrogen bonds to the amide/amidinium group, via W4. The second group of conserved waters (W6−11) form a network extending in the 5′ direction, hydrogen bonding to bases, O4′ sugar ring atoms and to phosphate groups. Water molecules W6−9 form a distorted square arrangement in the complexes, although a subtly distinct water structure is formed in the native structure where waters W8−9 are also part of the hydration sphere of a Mg^{2+} ion, and water W10 coordinates to G4 in the native structure − this ion is not present at the same position in the ligand complexes. Table 3 shows that the pattern of mobility for an individual water molecule is consistent across all four structures, with W1, W2, W6, W7 and W11 being the least mobile water molecules. Neither the amide group in the DB1963 complex nor the amidinium group in the DB1883 complex form direct DNA contacts. Instead W4, W5 and W6 play important roles in bridging to the DNA base edges and phosphate groups.

dx.doi.org/10.1021/ja308952y | J. Am. Chem. Soc. 2013, 135, 1369−1377

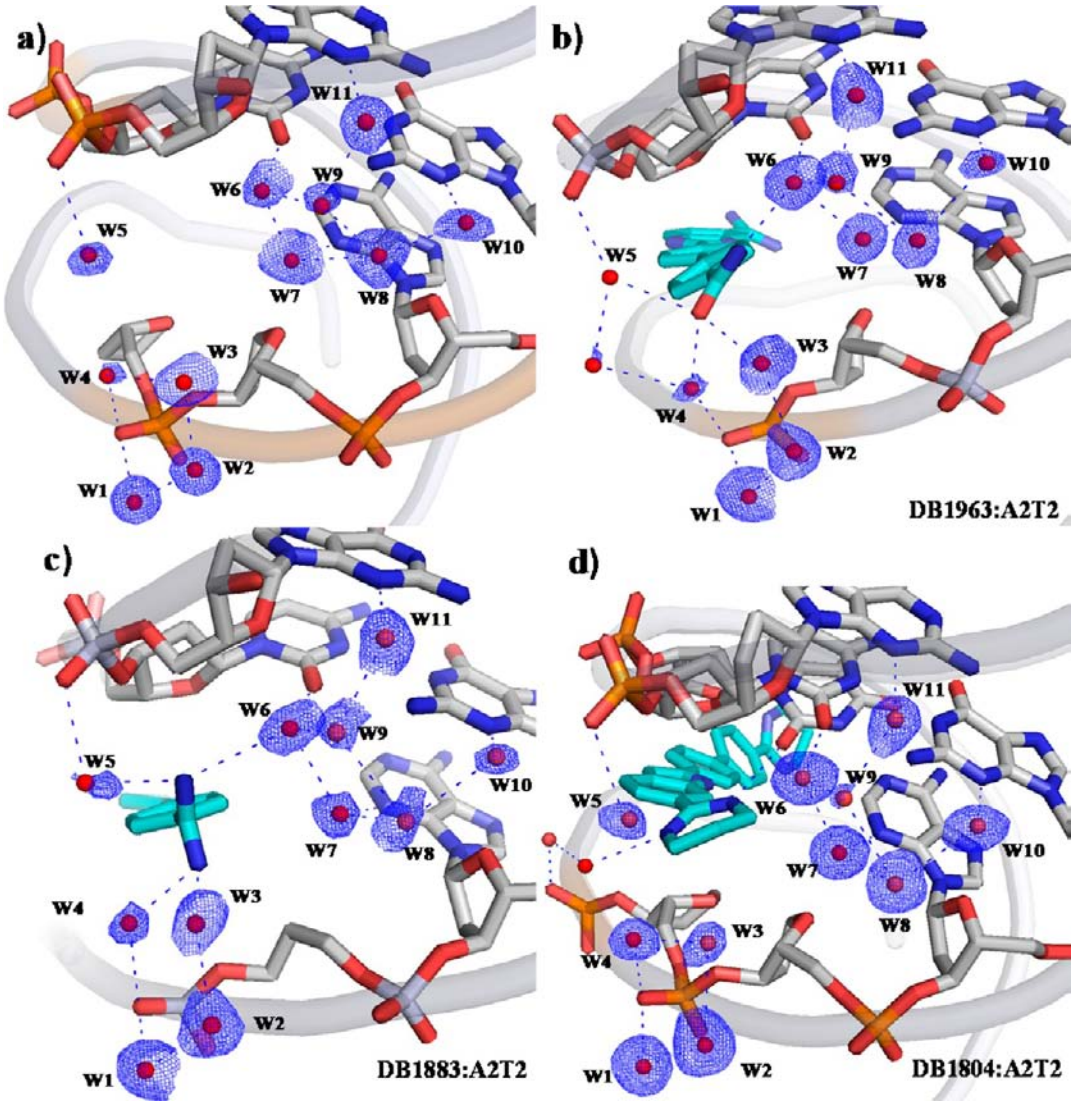


Figure 4. Difference electron density maps for the four structures, viewed in an identical orientation looking down the minor groove, showing the closely similar positions of water molecules in the cluster. The difference electron density is contoured at the 1.0 σ level.

Table 3. Individual Temperature Factors (B Factors in Å^2), for the Eleven Water Molecules Forming the Conserved Water Cluster in Each Crystal Structure.[a]

	DB1883:A2T2	DB1963:A2T2	DB1804:A2T2	native A2T2
B–W1	12.5	12.8	13.6	14.0
B–W2	17.0	17.9	15.8	8.5
B–W3	34.9	32.3	16.9	28.5
B–W4	25.7	31.4	15.2	25.4
B–W5	33.0	46.9	18.1	29.3
B–W6	22.5	16.0	18.9	22.3
B–W7	21.0	17.7	17.0	20.7
B–W8	34.4	26.3	18.7	44.1
B–W9	28.5	27.6	20.8	26.8
B–W10	23.7	21.6	19.4	21.8
B–W11	18.8	19.3	22.2	15.2
<B>	28.5	30.2	31.2	30.1
W_{tot}	142	149	154	215

[a] <B> is the mean water B factor, in Å^2. W_{tot} is the total number of water molecules observed in each structure.

Table 4. RMSD Values (Å) for the Conserved Cluster of Eleven Water Molecules Comparing Their Positions in the A2T2 and Three A2T2–Ligand Complex Crystal Structures

	DB1963:A2T2	DB1883:A2T2	DB1804:A2T2
Native A2T2	0.53	0.46	0.50
DB1963:A2T2		0.20	0.21
DB1883:A2T2			0.34

The structures of the three ligand complexes show that all have the characteristic bifurcated interstrand hydrogen-bonding involving the inner-facing nitrogen atom of the benzimidazole ring and the O2 atoms of T7 and T19 (Figures 1, 2, 6). For the crystal structures with compounds DB1963 and DB1883, the biphenyl and benzimidazole rings can be clearly observed in the electron density maps, and the amidinium group attached to each benzimidazole ring was able to be unambiguously identified (Figures 2b,c). These two compounds adopt the same orientation as previously found in the crystal structure of the A2T2 complex with DB921.[22] For both ligands the absence in initial electron density maps of complete electron density for the nitrogen atoms of the amidinium and amide groups linking

dx.doi.org/10.1021/ja308952y | J. Am. Chem. Soc. 2013, 135, 1369–1377

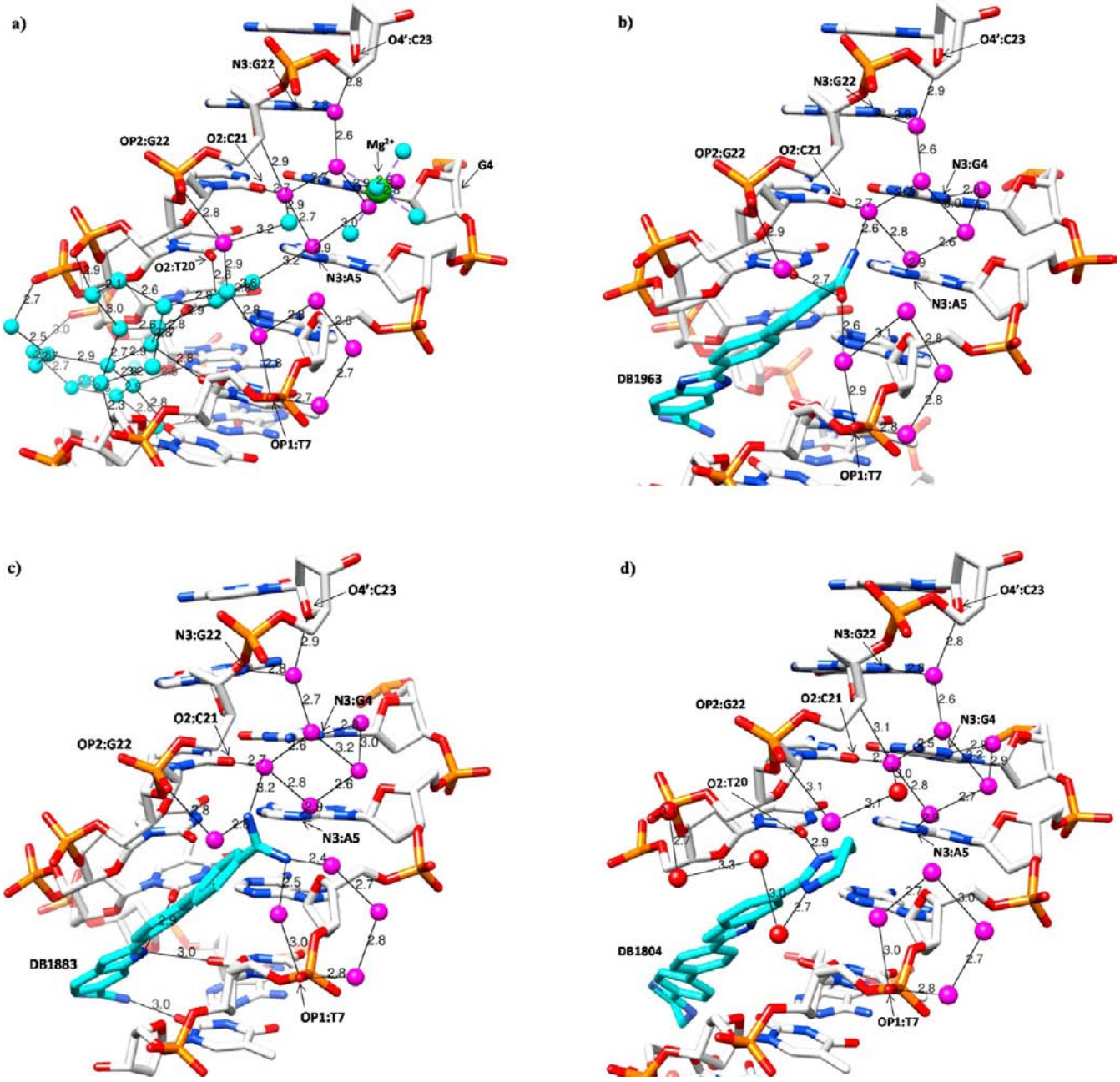


Figure 5. (a) Water clusters in the native A2T2 structure. Water molecules in the cluster are colored mauve and those in the spine of hydration are cyan. Hydrogen bonds are indicated by thin lines. (b) Water clusters in the DB1963:A2T2 complex. Carbon atoms in the ligand are colored cyan. (c) Water clusters in the DB1883:A2T2 complex. (d) Water clusters in the DB1804:A2T2 complex. Noncluster water molecules are colored red.

to the phenyl ring may be due to the greater flexibility of these end groups and the absence of direct nonbonding interactions with the DNA.

For the DB1804:A2T2 complex crystal, the initial ($F_o - F_c$) difference electron density map unambiguously showed the position of the phenyl and benzimidazole rings and the cyclic amidinium group attached to the benzimidazole (Figure 2d). This indicates that DB1804 adopts an inverted orientation in the groove compared to the DB1883 and DB1963 complexes, and thus the cyclic amidinium group occupies the position of the phenyl ring in these two complexes (Figure 6). However for the A2T2 complex with DB1055, the electron density could not be unambiguously interpreted even at higher resolution, to show the exact position for the benzimidazole and biphenyl rings. In particular, there was no clear density apparent for the terminal phenyl ring and its attached amidinium group at the corresponding position in the earlier lower-resolution DB1055-DNA crystal structure. This suggests that DB1055 may be present in two orientations as found in solution for several other minor-groove ligands.[29]

The localization of an extensive structured water environment surrounding the complexes provides additional understanding for the observed differences in affinities between the ligands. The hydrogen bonding of the amidinium group in DB1883 is identical to that of the amide group in DB1963 so the greater basicity of the former group is likely a significant factor in its greater DNA duplex affinity. In addition, the change from a benzimidazole (in DB1963) to an indole group has

dx.doi.org/10.1021/ja308952y | J. Am. Chem. Soc. 2013, 135, 1369−1377

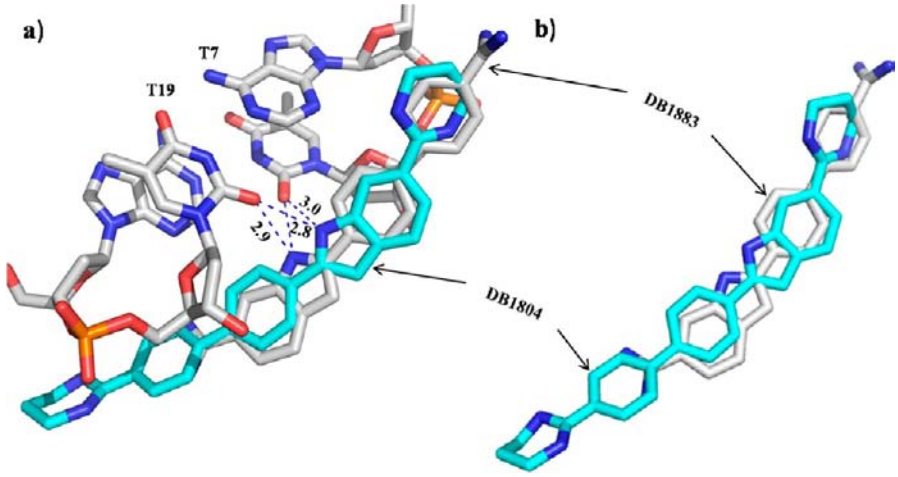


Figure 6. Overlaid structures of the DB1804 and DB1883:A2T2 complexes highlighting the bifurcated hydrogen bonding to O2 atoms of T7 and T19 and the distinct orientations of the two ligands in the minor groove. DB1804 is colored cyan and DB1883 gray.

resulted in a loss of the hydrogen-bonding to a water molecule by the outer-facing nitrogen atom in the benzimidazole ring of DB1883. Figure 7 shows the arrangement in the DB1963

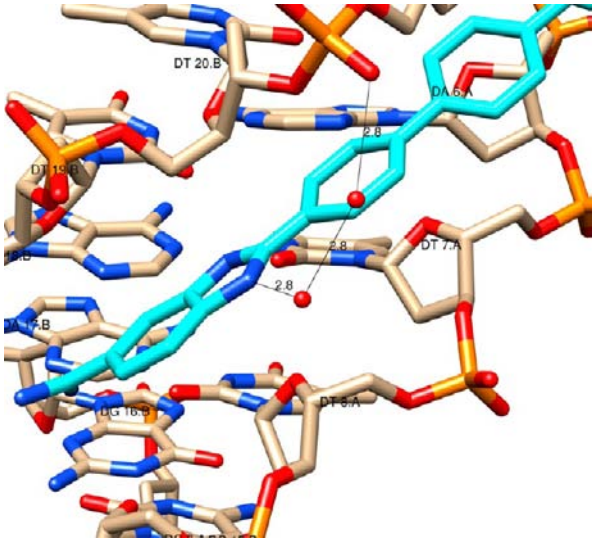


Figure 7. (a) Hydrogen bonding of the outer edge of the benzimidazole moiety of the ligand (colored cyan) in the DB1963:A2T2 complex, to a water molecule. This links to a second water molecule, bridging between the outer-facing nitrogen atom of the benzimidazole ring and a phosphate oxygen atom of C21. All other waters have been removed for clarity.

complex: in DB1883 the equivalent carbon atom in the indole has a nearest-neighbor water molecule 4.7 Å distant. This suggests that the entropic loss of the local water chain is likely to be greater than the gain in enthalpy from hydration, which results in the higher binding affinity of DB1883.

The cyclic amidinium group in DB1804 does not produce any significant change in the water cluster, even though there are no longer any direct hydrogen bonds between the cluster and ligand. Instead the inner-facing amidinium nitrogen atom is able to directly contact a base edge atom, O2 of T20 (Figure 5d), possibly contributing to the enhanced affinity of DB1804 (see Discussion). Interestingly the water cluster remains intact, with the W1−W4 pentagon close to, but not quite contacting structure-specific waters (shown as red spheres) that start at hydrated phosphate oxygen atoms, one of which ends at the outer-facing amidinium nitrogen atom.

DISCUSSION

The binding affinity between small molecules and DNA depends on changes in enthalpy and entropy during the binding process. By contrast with the ΔH values, the consistently greater $T\Delta S$ values (Table 1), suggests that the binding is driven in large part by changes in entropy. This is consistent with the removal of large amount of water molecules in the minor groove upon the binding of the small-molecule ligands discussed here. However, the relative binding affinity of two ligands depends on the difference of the sum of ΔH and $T\Delta S$ changes. The ΔH value results from the competition between ligands and the water molecules in the DNA minor groove. The interactions between the nonisohelical linear heterocyclic molecule and A2T2 DNA mainly comprises van der Waals interactions between the small molecules with four pairs of sugar rings (C21-A6, T20-T7, T19-T8, A18-C9), together with hydrogen bonding to four pairs of bases (17A-9C, 18A- 8T, 19T-7T, 20T-6A). The origins of the differences in

Table 5. Solvent-accessible Surface Areas (SA) of Ligands in Various A2T2 Minor Groove Complexes, in Å^2

ligand	DB921	DB911	DB1963	DB1883	DB1804	netropsin
PDB ID	2B0K	2NLM	3U08	3U0U	3U05	101D
SA ligand	3192.2	3192.2	***3190.0***	***3203.6***	***3928.2***	3588.8
SA ligand complex	869.8	1090.7	***879.6***	***849.7***	***1184.1***	1171.1
SA net buried surface	2322.4	2101.5	***2310.4***	***2353.9***	***2744.0***	2417.6
SA net buried surface/SA ligand	0.73	0.66	***0.72***	***0.73***	***0.70***	0.67

Numbers in bold-italics refer to the crystal structures reported here. The netropsin:A2T2 structure was reported in ref 17.

dx.doi.org/10.1021/ja308952y | J. Am. Chem. Soc. 2013, 135, 1369−1377

binding affinity are not immediately apparent for compounds with equivalent van der Waals contacts and hydrogen bonding modes.

For compounds DB921 and DB1963, the hydrogen bonding and van der Waals contacts in the minor groove are equivalent. The difference in binding free energy is 3.2 kcal/mol (Table 1), and the difference in entropy change is only 0.4 kcal/mol, which is consistent with the observed structured water molecules and the water cluster in the two complex crystal structures. It may be concluded that the difference in binding affinities is mainly mediated by ΔH. The difference in ΔH is 2.8 kcal/mol, which dominates 87.5% of the difference in ΔG (3.2 kcal/mol). This difference in ΔH may come from the greater electrostatic component to the total binding energy from the higher basicity of the amidinium group in DB921 interacting with the water cluster compared to the neutral terminal amide group in DB1963.

The binding affinity and the binding mode of DB1804 is consistent with the water cluster. Table 5 shows that DB1804 has ~20% greater solvent-accessible surface area than the other ligands and a correspondingly greater net solvent-accessible buried surface area in the minor groove binding site. This strongly suggests that the higher binding affinity of DB1804 arises from enhanced van der Waals interactions between its phenyl ring and sugar rings of A18 and C9, together with the effect of one hydrogen bond between a cyclic amidinium group nitrogen atom and O2 of T20. If DB1804 were to adopt the same orientation in the A2T2 minor groove as DB1883, the cyclic amidinium group connecting to the benzimidazole ring would be in close van der Waals contact with the sugar rings of A18 and C9, and the amidinium nitrogen atom still could form a hydrogen bond, with O2 of T8. However, the cyclic amidinium group at the other end of the ligand would then be in close proximity to and thus likely to destabilize the water cluster rather than being oriented away from the minor groove. Any van der Waals interactions between DNA atoms and this amidinium group would not compensate for the loss of hydration energy of the water cluster. The suggestion from the structures reported here, is that the cluster is a stable arrangement and ligands adopt an appropriate orientation around it.

As a result of the high-resolution crystal structures reported here, a new paradigm for recognition of A-tract minor groove binding sites, involving three major components, is apparent, that extends beyond the classical concept of isohelicity. Compounds such as the isohelical DB911 bind about 10-fold weaker than the distinctly nonisohelical DB921, which uses an interfacial water molecule to complete interaction with DNA bases at the floor of the groove. The compound substituents that interact with minor groove base edges form a second major component. DB921 and DB1963, for example, interact with the groove in a very similar pattern but have a 3−4 fold difference in binding affinity. DB1804 with a cyclic amidine has a K_a of 2 nM and is one of the strongest AT-specific minor groove binders yet discovered. A critical third component, a highly conserved cluster of additional water molecules, which is reported here, plays a distinctive role in attaching compounds to the minor groove and provides the finishing stabilization motif for the ligand:DNA-water network. A number of previous studies have reported assemblies of structured water molecules (see for example refs 30−33), with cyclic water arrangements being observed in the major groove of B-DNA[30] and around intercalating drug molecules.[33] The existence of analogous water pentagons, hexagons and other motifs in bulk water as well as ice is now well-established.[34] The minor-groove cluster reported here indicates ways in which such individual water motifs can be clustered together. It is also hoped that this finding will lead to new insights into designing DNA-binding molecules, as well as leading to further experimental and computational insights into how DNA structure is maintained in a stable arrangement.

AUTHOR INFORMATION

Corresponding Author

s.neidle@ucl.ac.uk

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are grateful to Cancer Research UK for support (Programme grant No. C129/A4489 to S.N. and a CRUK China Fellowship to D.W.), and to the Diamond Light Source for access to Synchrotron facilities. Research at Georgia State University is supported by NIH grant AI 064200 (to W.D.W. and D.W.B.). Drs. Nancy Campbell, Min Yang, and Stephan Ohnmacht and Mr. Emmanuel Samuels are thanked for useful discussions. We thank Professor David Boykin for the compounds used in this study.

REFERENCES

(1) Franklin, R. E.; Gosling, R. G. *Acta Crystallogr.* **1953**, *6*, 673−677. Berman, H. M. *Curr. Opin. Struct. Biol.* **1991**, *1*, 423−427. Fuller, W.; Forsyth, T.; Mahendrasingam, A. *Phil. Trans. R. Soc. Lond.* **2004**, *B359*, 1237−1247. Neidle, S. *Principles of Nucleic Acid Structure*; Academic Press: London, 2008.

(2) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535−556. Kopka, M. L.; Fratini, A. V.; Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1983**, *163*, 129−146. 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.; Manoharan, M. *Biopolymers* **1998**, *48*, 234−252. Vlieghe, D.; Turkenburg, J. P.; Van Meervelt, L. *Acta Crystallogr.* **1999**, *D55*, 1495−1502.

(3) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341−8355. Woods, K. K.; Maehigashi, T.; Howerton, S. B.; Sines, C. C.; Tannenbaum, S.; Williams, L. D. *J. Am. Chem. Soc.* **2004**, *126*, 15330−15331. Arai, S.; Chatake, T.; Ohhara, T.; Kurihara, K.; Tanaka, I.; Suzuki, N.; Fujimoto, Z.; Mizuno, H.; Niimura, N. *Nucleic Acids Res.* **2005**, *33*, 3017−3024.

(4) Tereshko, V.; Minasov, G.; Egli, M. *J. Am. Chem. Soc.* **1999**, *121*, 470−471.

(5) Wilton, D. J.; Ghosh, M.; Chary, K. V.; Akasaka, K.; Williamson, M. P. *Nucleic Acids Res.* **2008**, *36*, 4032−4037.

(6) Yonetani, Y.; Kono, H. *Biophys. J.* **2009**, *97*, 1138−1147. Fenn, T. D.; Schnieders, M. J.; Brunger, A. T.; Pande, V. S. *Biophys. J.* **2010**, *98*, 2984−2992.

(7) Madhumalar, A.; Bansal, M. *Biophys. J.* **2003**, *85*, 1805−1816. Jana, B.; Pal, S.; Bagchi, B. *J. Phys. Chem. B* **2010**, *114*, 3633−3638.

(8) Dixit, S. B.; Mezei, M.; Beveridge, D. L. *J. Biosci.* **2012**, *37*, 399−421. Xhu, X.; Schatz, G. C. *J. Phys. Chem. B* **2012**, *116*, 13672−13681.

(9) Goodsell, D.; Dickerson, R. E. *J. Med. Chem.* **1986**, *29*, 727−733.

(10) Nguyen, B.; Neidle, S.; Wilson, W. D. *Acc. Chem. Res.* **2009**, *42*, 11−21.

(11) Das, B. P.; Boykin, D. W. *J. Med. Chem.* **1977**, *20*, 531−536.

(12) Hall, J. E.; Kerrigan, J. E.; Ramachandran, K.; Bender, B. C.; Stanko, J. P.; Jones, S. K.; Patrick, D. A.; Tidwell, R. R. *Antimicrob. Agents Chemother.* **1998**, *42*, 666−674.

(13) Ismail, M. A.; Batista-Parra, A.; Miao, Y.; Wilson, W. D.; Wenzler, T.; Brun, R.; Boykin, D. W. *Bioorg. Med. Chem.* **2005**, *13*, 6718−6726.

(14) Wilson, W. D.; Tanious, F. A.; Mathis, A.; Tevis, D.; Hall, J. E.; Boykin, D. W. *Biochimie* **2008**, *90*, 999−1014.
(15) Paine, M. F.; Wang, M. Z.; Generaux, C. N.; Boykin, D. W.; Wilson, W. D.; De Koning, H. P.; Olson, C. A.; Pohlig, G.; Burri, C.; Brun, R.; Murilla, G. A.; Thuita, J. K.; Barrett, M. P.; Tidwell, R. R. *Curr. Opin. Invest. Drugs* **2010**, *11*, 876−883.
(16) Coll, M.; Aymami, J.; Van der Marel, G. A.; Van Boom, J. H.; Rich, A.; Wang, A. H. J. *Biochemistry* **1989**, *28*, 310−320.
(17) Goodsell, D. S.; Kopka, M. L.; Dickerson, R. E. *Biochemistry* **1995**, *34*, 4983−4993.
(18) Nguyen, B.; Lee, M. P.; Hamelberg, D.; Joubert, A.; Bailly, C.; Brun, R.; Neidle, S.; Wilson, W. D. *J. Am. Chem. Soc.* **2002**, *124*, 13680−13681.
(19) Mallena, S.; Lee, M. P.; Bailly, C.; Neidle, S.; Kumar, A.; Boykin, D. W.; Wilson, W. D. *J. Am. Chem. Soc.* **2004**, *126*, 13659−13669.
(20) Miao, Y.; Lee, M. P.; Parkinson, G. N.; Batista-Parra, A.; Ismail, M. A.; Neidle, S.; Boykin, D. W.; Wilson, W. D. *Biochemistry* **2005**, *44*, 14701−14708.
(21) Goodwin, K. D.; Lewis, M. A.; Tanious, F. A.; Tidwell, R. R.; Wilson, W. D.; Georgiadis, M. M.; Long, E. C. *J. Am. Chem. Soc.* **2006**, *128*, 7846−7854.
(22) Liu, Y.; Kumar, A.; Depauw, S.; Nhili, R.; David-Cordonnier, M. H.; Lee, M. P.; Ismail, M. A.; Farahat, A. A.; Say, M.; Chackal-Catoen, S.; Batista-Parra, A.; Neidle, S.; Boykin, D. W.; Wilson, W. D. *J. Am. Chem. Soc.* **2011**, *133*, 10171−10183.
(23) Lee, M. P.; Wilson, W. D.; Neidle, S. to be published.
(24) Winn, M. D.; et al. *Acta Crystallogr.* **2011**, *D67*, 235−242.
(25) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658−674.
(26) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *Acta Crystallogr.* **2011**, *D67*, 355−367.
(27) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr.* **2010**, *D66*, 486−501.
(28) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(29) Rettig, M.; Germann, M. W.; Ismail, M. A.; Batista-Parra, A.; Munde, M.; Boykin, D. W.; Wilson, W. D. *J. Phys. Chem. B* **2012**, *116*, 5620−5627.
(30) Minasov, G.; Tereschko, V.; Egli, M. *J. Mol. Biol.* **1999**, *291*, 83−99.
(31) Soler-López, M.; Malinina, L.; Subirana, J. A. *J. Biol. Chem.* **2000**, *275*, 23034−23044.
(32) Kennard, O.; Cruse, W. B.; Nachman, J.; Prange, T.; Shakked, Z.; Rabinovich, D. *J. Biomol. Struct. Dynam.* **1986**, *3*, 623−647.
(33) Neidle, S.; Berman, H. M.; Shieh, H. S. *Nature* **1980**, *288*, 129−133.
(34) Pérez, C.; Muckle, M. T.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Science* **2012**, *336*, 897−901.