Recognition of ATGA Sequences by the Unfused Aromatic Dication DB293 Forming Stacked Dimers in the DNA Minor Groove†

Christian Bailly,*,‡ Christelle Tardy,‡ Lei Wang,§ Bruce Armitage,||,⊥ Katherine Hopkins,§ Arvind Kumar,§ Gary B. Schuster,|| David W. Boykin,*,§ and W. David Wilson*,‡,§

INSERM U-524 et Laboratoire de Pharmacologie Antitumorale du Centre Oscar Lambret, IRCL, Place de Verdun, 59045 Lille, France, Department of Chemistry and Laboratory for Chemical and Biological Sciences, Georgia State University, Atlanta, Georgia 30303, and School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400

Received April 25, 2001; Revised Manuscript Received June 19, 2001

ABSTRACT: Furamidine and related diamidines represent a promising series of drugs active against widespread parasites, in particular the *Pneumocystic carinii* pathogen. In this series, the phenylfuranbenzimidazole diamidine derivative DB293 was recently identified as the first unfused aromatic dication capable of forming stacked dimers in the DNA minor groove of GC-containing sequences. Here we present a detailed biochemical and biophysical characterization of the DNA sequence recognition properties of DB293. Three complementary footprinting techniques using DNase I, Fe^{II}-EDTA, and an anthraquinone photonuclease were employed to locate binding sites for DB293 in different DNA restriction fragments. Two categories of sites were identified by DNase I footprinting: (i) 4/5 bp sequences containing contiguous A·T pairs, such as 5′-AAAA and 5′-ATTA; and (ii) sequences including the motif 5′-ATGA·5′-TCAT. In particular, a 13-bp sequence including two contiguous ATGA motifs provided a highly preferential recognition site for DB293. Quantitative footprinting analysis revealed better occupancy of the 5′-ATGA site compared to the AT-rich sites. Preferential binding of DB293 to ATGA sites was also observed with other DNA fragments and was confirmed independently by means of hydroxyl radical footprinting generated by the Fe^{II}-EDTA system, as well as by a photofootprinting approach using the probe anthraquinone-2-sulfonate (AQS). In addition, this photosensitive reagent revealed the presence of sites of enhanced cutting specific to DB293. This molecule, but not other minor groove binders such as netropsin, induces specific local structural changes in DNA near certain binding sites, as independently shown by DNase I and the AQS probe. Recognition of the ATGA sequence by DB293 was investigated further using melting temperature experiments and surface plasmon resonance (SPR). The use of different hairpin oligonucleotides showed that DB293 can interact with AT sites via the formation of 1:1 drug−DNA complexes but binds much more strongly, and cooperatively, to ATGA-containing sequences to form 2:1 drug−DNA complexes. DB293 binds strongly to ATGA sequences with no significant context dependence but is highly sensitive to the orientation of the target sequence. The formation of 2:1 DB293/DNA complexes is abolished by reversing the sequence 5′-ATGA→3′-ATGA, indicating that directionality plays an important role in the drug−DNA recognition process. Similarly, a single mutation in the A[T→G]GA sequence is very detrimental to the dimer interactions of DB293. From the complementary footprinting and SPR data, the 5′-ATGA sequence is identified as being a highly favored dimer binding site for DB293. The data provide clues for delineating a recognition code for diamidine-type minor groove binding agents, and ultimately to guide the rational design of gene regulatory molecules targeted to specific sites of the genetic material.

Furamidine (DB75) is a synthetic diphenylfuran diamidine showing promising activity against a variety of microorganisms, including *Pneumocystic carinii* and other infectious diseases that can cause severe health problems to immune-compromised populations (*1*). A bis-amidoxime prodrug of furamidine has shown excellent activity in animals (*2*) and is currently undergoing clinical trials for the treatment *Pneumocystic carinii* pneumonia. The biological activity of this category of molecules is thought to result from direct interaction with DNA and subsequent inhibition of DNA-dependent enzymes (topoisomerases, polymerases, nucleases) or possibly by direct inhibition of transcription (*3*−*8*).

Furamidine and related unfused aromatic dications strongly bind to the minor groove of AT-rich DNA sequences (*9*−*11*). The two cationic terminal groups of the molecule play a major role in the complex formation with DNA (*12*).

† This work was supported by Research Grant GM 61587 (to W.D.W. and D.W.B.) from the National Institutes of Health and by a grant (to C.B.) from the Ligue Nationale Française Contre le Cancer (Comité du Nord). W.D.W. is currently the recipient of an INSERM "Poste Orange" fellowship. The BIAcore 2000 instrumentation was purchased through funds from the Georgia Research Alliance.

* Address correspondence to either of these authors. E-mail: chewdw@panther.gsu.edu, chedwb@panther.gsu.edu (for W.D.W. and D.W.B., respectively); bailly@lille.inserm.fr (for C.B.).

‡ INSERM U-524.

§ Georgia State University.

|| Georgia Institute of Technology.

⊥ Present address: Department of Chemistry, Carnegie Mellon University, 4400 Fifth Ave., Pittsburgh, PA 15213-3890 (army@cyrus.andrew.cmu.edu).

10.1021/bi0108453 CCC: $20.00
Published on Web 07/25/2001

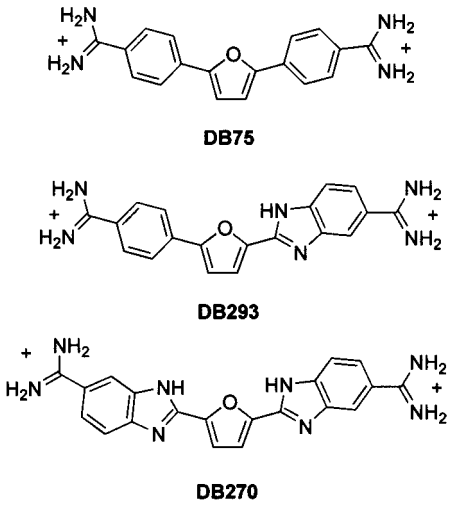

FIGURE 1: Structure of the drugs used in this study: DB75, DB270, and DB293.

Replacement of the amidine ends with imidazolines shifts the mode of binding to DNA and confers intercalative properties in GC base pairs containing sequences (*13*). Alkylation of the amidine substituents also affects considerably and diversely the strength of the interaction with DNA and the sequence preference (*7*). For example, introduction of a cyclopentyl group on the amidine reinforces DNA interaction whereas a 3-pentyl group reduces binding (*12*). Similarly, substitution of guanidino groups for the amidino groups of furamidine alters the DNA recognition properties and modifies the anti-parasitic activities of the molecules (*14*). Recent studies have revealed that hydrophobic interactions between the drug and the DNA receptor are essential to the drug−DNA complex formation and its stability (*15*).

The central aromatic core of furamidine also contributes significantly to the DNA binding process. Methylation of the central furan ring and incorporation of a benzoxazole ring in place of one of the phenyl nuclei both modify significantly the DNA interaction. However, these structural variations do not profoundly change the sequence selectivity of the compound. Among the hundreds of furamidine analogues synthesized thus far, almost all synthetic derivatives maintain a significant preference for AT-rich DNA sequences (*9, 11, 12, 14*). The presence of a G·C base pair in a sequence of contiguous A·T pairs is generally sufficient to exclude drug binding to the minor groove of this sequence. The lack of G·C recognition was a major obstacle to the development of furamidine-type molecules as sequence recognition elements. Fortunately, this hurdle has been removed with the recent discovery of the phenylfuranbenzimidazole analogue DB293 (Figure 1).

In the course of a program aimed at rationally designing small molecules capable of reading the genetic information of DNA, we identified DB293 as a nonpolyamide, aromatic dication capable of forming strong and stable complexes with DNA sequences containing both A·T and G·C base pairs (*16*). Unlike other diamidines related to furamidine, DB293 was found to bind GC-containing sites in DNA more strongly than to pure AT sequences. This is unprecedented in the furan-based set of derivatives and is quite rare for minor groove binding agents in general. NMR and SPR[1] studies unambiguously revealed that the interaction of this unique compound with oligonucleotides is highly cooperative and, most importantly, it involves the formation of drug dimers stacked in the minor groove of DNA (*16*). The formation of 2:1 drug−DNA minor groove complexes has been previously described with monocationic drugs such as distamycin and related polyamides (*17*−*20*) but not with dications or nonpolyamides. DB293 is the first unfused aromatic dication capable of forming stacked dimers in the DNA minor groove of GC-containing sequences (*16*). As such, it provides a paradigm for design of new categories of gene-regulatory small molecules.

It is clearly essential to determine the details of DNA recognition by this new dimer complex, but, to date, our understanding of the molecular basis for the exquisite recognition of DNA sequences by DB293 remains relatively sparse. However, a few important aspects of the DB293−DNA recognition process have been elucidated. Recently, structure binding relationship studies have revealed that the formation of the minor groove stacked dimer is very sensitive to compound structure (*21*). For example, removal of one amidine or addition of an alkyl group strongly inhibits dimer formation. Similarly, changing the benzimidazole ring of DB293 to a phenyl or a benzofuran ring also abolished dimer formation of the derivatives studied up to this time. Only the conversion of the amidines to guanidinium groups preserves the cooperative dimer (*21*). Interestingly, we discovered that DB340, which is a tetracationic derivative of DB293 incorporating a dimethylaminopropyl side chain on the amidine, binds to a specific loop-containing RNA from HIV-1 as a dimer (*22*), suggesting that this class of phenylfuranbenzimidazole cationic molecules may be useful for the targeting of both DNA and RNA sequences (*23*).

The use of various synthetic analogues of DB293 has greatly helped to identify the drug substituents involved in the DNA recognition process. In contrast, we know relatively little concerning the types of DNA sequences preferentially recognized by DB293. Our initial study combining NMR, SPR, and DNase I footprinting showed that DB293 forms an antiparallel, stacked dimer in complex with some DNA sites that contain G·C base pairs (*21*), but no specific binding site was identified. Here we report studies aimed at delineating more precisely the nature of the DNA sequences recognized preferentially by DB293 and the relationships between the sequence of the binding site and the stoichiometry of the drug−DNA interaction. The results reported here, obtained by complementary chemical and enzymatic footprinting as well as by SPR methodologies, demonstrate that the sequence ATGA provides a highly preferential binding site for the DB293 dimer.

MATERIALS AND METHODS

Compounds, Biochemicals, Buffers, and DNA Oligomers. The benzimidazole derivatives DB270 and DB293 were prepared as reported previously (*24*). Netropsin was purchased from Serva. The purity of all compounds was verified by NMR and elemental analysis. Nucleoside triphosphates labeled with ^{32}P (α-dATP and γ-ATP) were obtained from Amersham (3000 Ci/mmol). Unlabeled dATP and the restriction endonucleases *Eco*RI, *Ava*I, and *Pvu*II were from

[1] Abbreviations: SPR, surface plasmon resonance; RU, response unit; T_m, thermal melting temperature; TBE, Tris−borate−EDTA.

Boehringer Mannheim (Germany). Alkaline phosphatase, T4 polynucleotide kinase, and avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim) were used according to the supplier's recommended protocol in the activity buffer provided. Bovine pancreatic deoxyribonuclease I (DNase I, Sigma Chemical Co.) was stored as a 7200 units/mL solution in 20 mM NaCl, 2 mM $MgCl_2$, 2 mM $MnCl_2$, pH 8.0. The stock solution of DNase I was kept at −20 °C and freshly diluted to the desired concentration immediately prior to use. All other chemicals were analytical grade reagents, and all solutions were prepared using doubly deionized, Millipore filtered water. MES buffers contained 0.01 M MES and 10^{-3} M EDTA with the pH adjusted to 6.2 with NaOH and sodium chloride added to adjust the ionic strength to the desired value. The different 5′-biotin-labeled hairpin oligonucleotides used for the SPR experiments were obtained HPLC-purified and desalted from Midland Certified Reagent Co. Oligomer concentrations were determined optically using extinction coefficients per mole of strand at 260 nm determined by the nearest-neighbor procedure (*25*).

Purification and Radiolabeling of the DNA Substrates for Footprinting. Plasmids pBS (Stratagene, La Jolla, CA) and pUC19 (Boehringer Mannheim, Germany) and pKMp27 (*26*) were prepared from *E. coli* according to standard procedures employing sodium dodecyl sulfate−sodium hydroxide lysis followed by purification using Qiagen columns. The 232- and 265-bp DNA fragments were prepared by 3′-^{32}P-end-labeling of the *Eco*RI−*Pvu*II double digest of plasmids pUC19 and pBS, respectively, using [α-^{32}P]dATP and AMV reverse transcriptase. The 265-mer fragment was also prepared by 5′-^{32}P-end-labeling of the *Eco*RI/alkaline phosphatase-treated plasmid using [γ-^{32}P]ATP (6000 Ci/mmol) and T4 polynucleotide kinase followed by treatment with *Pvu*II. The *tyr* T DNA was obtained by digestion of plasmid pKMp27 with *EcoR*I and *Ava*I and was labeled specifically at the 3′-end of the *Eco*RI site. In all cases, the singly end-labeled DNA fragments were then purified by preparative nondenaturing polyacrylamide gel electrophoresis (6.5% acrylamide, 1.5 mm thick, 200 V, 2 h, in TBE buffer: 89 mM Tris base, 89 mM boric acid, 2.5 mM Na_2EDTA, pH 8.3). After autoradiography, the requisite band of DNA was excised, crushed, and soaked in water overnight at 37 °C. This suspension was filtered, and the DNA was precipitated with ethanol. Following washing with 70% ethanol and vacuum-drying of the precipitate, the labeled DNA was resuspended in 10 mM Tris adjusted to pH 7.0 containing 10 mM NaCl.

DNase I Footprinting. DNase I footprinting experiments were performed essentially as previously described (*27*). Briefly, samples (3 μL) of the labeled DNA fragments were incubated with 5 μL of the buffered solution containing the ligand at the appropriate concentration. After 30 min incubation at 37 °C to ensure equilibration of the binding reaction, the digestion was initiated by the addition of 2 μL of a DNase I solution whose concentration was adjusted to yield a final enzyme concentration of about 0.01 unit/mL in the reaction mixture. After 3 min, the reaction was stopped by freeze-drying. Samples were lyophilized and resuspended in 5 μL of an 80% formamide solution containing tracking dyes. The DNA samples were then heated at 90 °C for 4 min and chilled in ice for 4 min prior to electrophoresis on a 8% polyacrylamide gel under denaturating conditions (8 M urea). A Molecular Dynamics 425E PhosphorImager was used to collect data from the storage screens exposed to dried gels overnight at room temperature. Baseline-corrected scans were analyzed by integrating all the densities between two selected boundaries using ImageQuant version 3.3 software. Each resolved band was assigned to a particular bond within the DNA fragments by comparison of its position relative to sequencing standards generated by treatment of the DNA with dimethyl sulfate followed by piperidine-induced cleavage at the modified guanine bases in DNA (G-track).

DNA Thermal Melting. Thermal melting experiments were conducted with Cary 3 or Cary 4 spectrophotometers interfaced to microcomputers as previously described (*28*). A thermistor fixed into a reference cuvette was used to monitor the temperature. The DNA or oligomer was added to 1 mL of buffer (MES with 0.1 M NaCl added) in 1 cm path length reduced volume quartz cells, and the concentration was determined by measuring the absorbance at 260 nm. Experiments were generally conducted at a concentration of 3×10^{-6} M in duplex. For the complex T_m experiments, a ratio of 1 compound per oligomer duplex was used.

Determination of Binding Constants by Surface Plasmon Resonance. Surface plasmon resonance (SPR) measurements were performed with a four-channel BIAcore 2000 system and streptavidin-coated sensor chips (SA), as previously described (*16*). To prepare the sensor chips for use, they were conditioned with three consecutive 1-min injections of 1 M NaCl in 50 mM NaOH followed by extensive washing with buffer. 5′-Biotinylated DNA (25 nM) in MES buffer with 0.1 M NaCl was immobilized on the surface by noncovalent capture. Three flow cells were used to immobilize DNA samples, and the fourth flow cell was left blank as a control. Manual injection was used with a flow rate of 2 μL/min to achieve long contact times with the surface and to control the amount of DNA bound to the surface. Samples of the drug were prepared in filtered and degassed buffer by serial dilutions from stock solutions. All procedures for binding studies were automated as methods using repetitive cycles of sample injection and regeneration. All the drug samples were injected from 7 mm plastic vials with pierceable plastic crimp caps at flow rate of 20 μL/min using the KINJECT command. For the DB293−DNA complexes, buffer flow alone is sufficient to dissociate the drug from DNA for surface regeneration. An array of different drug concentrations was used in each experiment.

RESULTS AND DISCUSSION

Footprinting

Footprinting experiments were carried out using three DNA fragments having different arrangements of base pairs and three footprinting probes with varied mechanisms of DNA binding and cleavage. In each case, the products of digestion by the probe in the absence and presence of DB293 were resolved by polyacrylamide gel electrophoresis. For the sake of convenience, we will consider the results obtained with each fragment in turn.

265-mer from pBS. DNase I footprinting experiments were initially performed using an *Eco*RI−*Pvu*II restriction fragment cut out from plasmid pBS. This 265-bp DNA fragment,

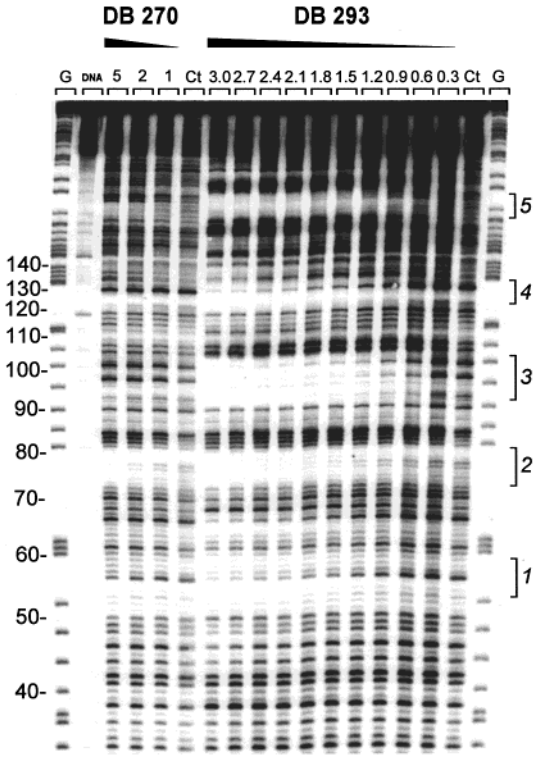

FIGURE 2: DNase I footprinting of DB270 and DB293 on the *Eco*RI−*Pvu*II 265-bp fragment from pBS. The DNA was 3′-end-labeled at the *Eco*RI site with [α-^{32}P]dATP in the presence of AMV reverse transcriptase. The products of nuclease digestion were resolved on an 8% polyacrylamide gel containing 8 M urea. The concentration (μM) of the drug is shown at the top of the appropriate gel lanes. Control tracks (Ct) contained no drug. The tracks labeled "G" represent dimethyl sulfate−piperidine markers specific for guanines. Numbers on the left side of the gel refer to the standard numbering scheme for the nucleotide sequence of the DNA fragment.

3′-labeled at the *Eco*RI site, was incubated with increasing concentrations of DB293 for 30 min at room temperature to establish equilibrium, and the samples were exposed to DNase I. DB293 concentrations varying from 0.3 to 3 μM were chosen so as to generate complexes in which the fractional occupancy of DNA binding sites varied from a few percent to near 100% under the conditions of the experiments (Figure 2). Parallel experiments were performed with the symmetric compound DB270 containing a benzimidazole ring in place of the phenyl ring of DB293 (Figure 1). The results support our earlier conclusion that, unlike DB293, the dibenzimidazolefuran compound does not form minor groove dimers and its binding is strictly restricted to AT tracts (*21*). A similar AT binding preference has been observed with the diphenylfuran derivative DB75.

The autoradiogram in Figure 2 reveals the presence of five clear binding sites for DB293 (numbered 1−5) whereas fewer and weaker sites are detected with DB270. Sites 4 and 5 at the top of the gel were not accessible to densitometric analysis (because the bands are not well separated), and therefore the binding sequences could not be precisely identified. In contrast, the densitometric analysis of the phosphorimage allowed us to localize precisely the position of sites 1−3 and also to estimate the relative strength of DB293 binding to these sites under the gel conditions. Both DB293 and DB270 were found to interact with the sequences 5′-AAAA (site 1) and 5′-ATTA (site *2*). But only DB293

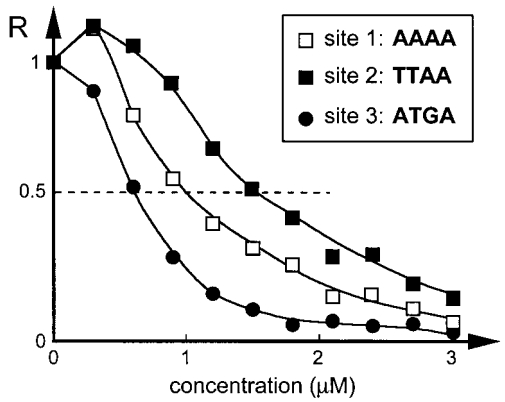

FIGURE 3: Footprinting plots for the binding of DB293 to sites 1−3. The relative band intensity R corresponds to the ratio I_c/I_o, where I_c is the intensity of the band at the ligand concentration c and I_o is the intensity of the same band in the absence of DB293.

interacted with the sequence 5′- TATGACCATGATTA (site 3 around nucleotide position 100). To estimate the relative binding affinity of DB293 to sites 1−3, we considered the variation of band intensities as a function of the drug concentration. The footprinting plots in Figure 3 show that for these sites the cleavage intensity decreases rapidly with increasing DB293 concentration. The decrease reflects the direct effect of the DNA-bound drug molecules which impede access of DNase I to specific sites. In each case, we determined the drug concentration required for half-maximal footprinting. The C_{50} values are 1.08, 1.57, and 0.62 μM for sites 1, 2, and 3, respectively. Under the conditions of the footprinting experiments, a large fraction of the ligand must be free, so that these C_{50} values may approximate dissociation constants for binding to individual sites. There is no doubt that DB293 binds more strongly to site 3 composed of two ATGA motifs than to sites 1 and 2 containing only A·T bp. Quantitative BIAcore binding experiments with this and an AT reference sequence are described below.

Next we repeated the footprinting analysis with the same DNA fragment labeled on the complementary strand, at the 5′-end of the *Eco*RI site. In this case, the typical minor groove binder netropsin was used as a reference. Here again, it can be seen in the gel shown in Figure 4A that the rate of cleavage by DNase I is subject to considerable modulation through local binding of DB293. Visual inspection of the gel clearly indicates that DB293, but not netropsin, binds specifically to the ATGA-containing sequence around nucleotide positions 90−100. The differential cleavage plots presented in Figure 4B reveal unambiguously that the footprinting profiles are different for the two drugs. For netropsin, the binding sites are confined to AT-rich sequences such as 5′-TTT, 5′-TTTA, 5′-AATTT, and 5′-TAAT near positions 54, 65, 79, and 93. DB293 can also bind to these AT-rich sequences, but, in addition, it recognizes the sequence 5′-AATCATGGTCATA characterized by the two adjacent 5′-ATGA·5′-TCAT motifs (underlined). Detailed analysis of this site is difficult because of its large size. The bidentate aspect of the differential cleavage plot suggests that this impressive site, which is by far the preferred binding site of DB293 in the 265-mer fragment, may correspond to an overlap of two binding subsites, but this is difficult to ascertain by means of DNase I footprinting. For this reason, complementary experiments were performed with the metal

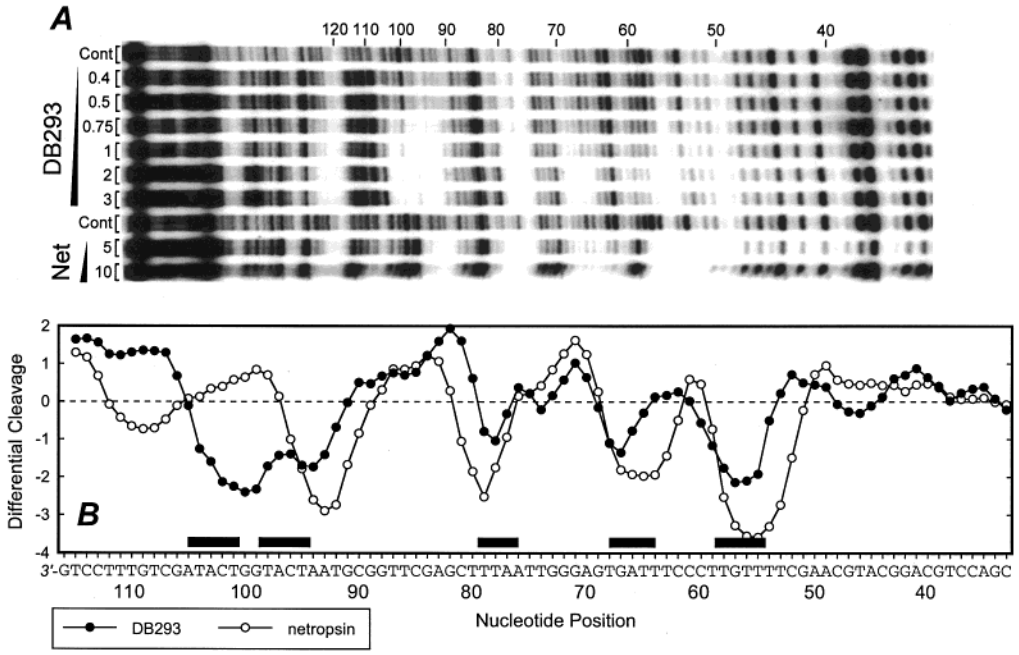

FIGURE 4: (A) DNase I footprinting of netropsin and DB293 on the 265-bp fragment from pBS. The DNA was 5′-end-labeled at the *Eco*RI site with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase. Other details as for Figure 2. Panel B shows differential cleavage plots comparing the susceptibility of the fragment to DNase I cutting in the presence of 10 μM netropsin and 1.5 μM DB293. Negative values correspond to a ligand-protected site, and positive values represent enhanced cleavage. Vertical scales are in units of $\ln(f_a) - \ln(f_c)$, where f_a is the fractional cleavage at any bond in the presence of the drug and f_c is the fractional cleavage of the same bond in the control, given closely similar extents of overall digestion. Only the region of the restriction fragment analyzed by densitometry is shown. Black boxes indicate the positions of inhibition of DNase I cutting in the presence of the drugs.

complex EDTA·Fe^{II} in order to visualize at nucleotide resolution the position of individual DB293 binding sites in the ATGA-containing sequence to aid in elucidation of the determinants of sequence recognition.

Hydroxyl radicals generated by EDTA·Fe^{II} were used to cleave the 265-bp *Eco*RI−*Pvu*II restriction fragment (3′-labeled) in the absence and presence of 10 μM DB293. In this case, the regions of attenuated cleavage by EDTA·Fe^{II} are more difficult to identify visually, and a densitometric analysis of the autoradiogram is required to permit any accurate estimation of the location and approximate size of DB293 binding sites. The densitometric traces presented in Figure S1 (see Supporting Information) show that the cleavage rate by EDTA·Fe^{II} is subject to modulation by local binding of DB293 to the two adjacent ATGA sites. Thus, a good correlation is found between DNase I and OH$^{\bullet}$ footprinting, which tends to validate both sets of observations. These observations are strongly supported by SPR results (below) that show two molecules of DB293 bound to ATGA sequences.

232-mer from pUC19. To investigate further the sequence selectivity of DB293, additional footprinting experiments were performed using the photosensitive probe AQS (anthraquinone-2-sulfonate) (*29*−*31*). Irradiation of AQS in a chloride-containing solution gives efficient and nonselective cleavage of DNA presumably through generation of chlorine atoms due to photooxidation of chloride by the excited AQS, producing a relatively uniform cleavage ladder (*32*).

The DNA used for these experiments was a 232-bp *Eco*RI−*Pvu*II restriction fragment from plasmid pUC19. Like the 265-bp fragment from pBS, this fragment also contains a sequence composed of two adjacent ATGA sites separated by two C residues. An example autoradiogram

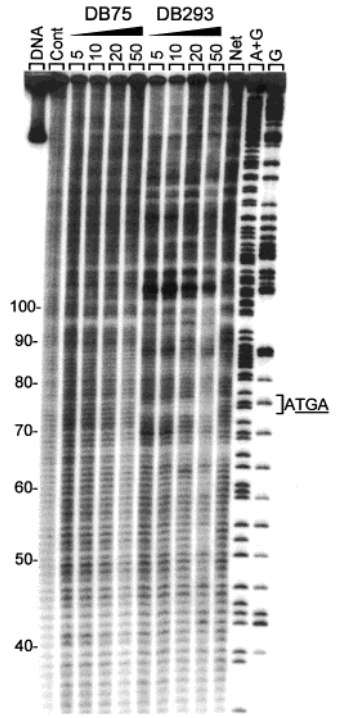

FIGURE 5: Anthraquinone photofootprinting of DB75 and DB293 bound to the 232-bp *Eco*RI−*Pvu*II restriction fragment from plasmid pUC19. Netropsin (Net) was used at 50 μM. Other details as for Figure 2.

resulting from this photofootprinting reaction is presented in Figure 5. Here again, it can be seen clearly that the cleavage of the DNA is significantly affected by binding of DB293, whereas under the same conditions the cleavage is only weakly modified by furamidine (DB75). Although this method does not produce footprints that are as strong as with

DNase I, areas of decreased band intensity in the presence of DB293, compared to the control lane, can readily be seen at several positions along the DNA sequence. For example, a close inspection of the autoradiograph reveals the important result that the drug inhibits the anthraquinone from cleaving DNA at the sequence ATGA at positions 73−76. The underlined triplet in the ATGACCATGA is with no doubt protected by DB293 from cleavage by the probe. This is in agreement with the DNase I footprinting experiments performed with the same DNA fragment. With the enzyme, the two adjacent ATGA motifs were strongly protected by DB293 but not by netropsin which was found to bind a vicinal ATTA tetrad (Figure S2).

The main reason for use of the photofootprinting probe was to analyze the perturbation of the DNA helix induced by the ligand. The striking result, seen in Figure 5, is that several sequences that are close to DB293 binding sites are cut with higher efficiency (cleavage enhancement) in the presence of DB293 but such enhancements are not seen at any sites with DB75 (Figure 5). Interestingly, the sites of enhanced cutting by the photoreagent are quite specific to DB293 bound to ATGA and are not detected at all with netropsin or DB270 bound at any AT sites (Figure S3). For example, upon binding to DNA, DB293 strongly potentiates the reaction of the anthraquinone probe at sites G68, G86, GG104−105, and GG112−113. The accumulation of oxidative DNA damage is particularly strong at the AGGC site (nucleotide positions 103−106), even with a low concentration of DB293 (5 μM). This suggests that the distortion of the helix, which is sensed by the AQS probe, varies locally according to the sequence to which the drug is bound. The results suggest that DB293 induces specific local structural changes in DNA near certain binding sites.

160-mer from pKMp27. The *tyr*T fragment has been extensively used in the past to footprint a vast number of DNA binding small molecules including both intercalators and groove binders (*33*). In particular, it has been used to identify the binding sites for netropsin and distamycin (*34*−*37*) as well as various diphenylfuran derivatives of DB293 (*7*). Figure 6 shows a gel obtained using increasing concentrations of DB293 with the *tyr*T fragment. Four major sites of drug protection can be discerned at base positions 22, 46, 60, and 83 (respectively sites 1, 2, 3, and 4). Two additional sites can be seen at the top of the gel around positions 110 and 130, but they lie beyond the region accessible to densitometry. The protected zones extend for 6−10 bases. Parallel footprinting experiments performed with DB270 showed only two binding sites centered around positions 30 and 48, both corresponding to pure AT sites as revealed by the differential cleavage plots presented in Figure 7. Site 2 is common to DB270, DB293, and netropsin: it corresponds to the sequence 5′-AAAAAT. In sharp contrast, sites 1, 3, and 4 are specific to DB293, and it is remarkable to observe that two of them (1 and 4) coincide with the position of 5′-ATGA·TCAT motifs. In particular, the prominent binding site between nucleotide positions 80−92 corresponds to the sequence 5′-TCATATCAAATGA. This finding corroborates the previous set of data obtained with the 265- and 232-mer fragments suggesting that the sequence ATGA provides a strongly favored binding site for DB293. This binding motif is with no doubt a prime target for DB293. Nevertheless, this drug can also interact with other types of non-AT

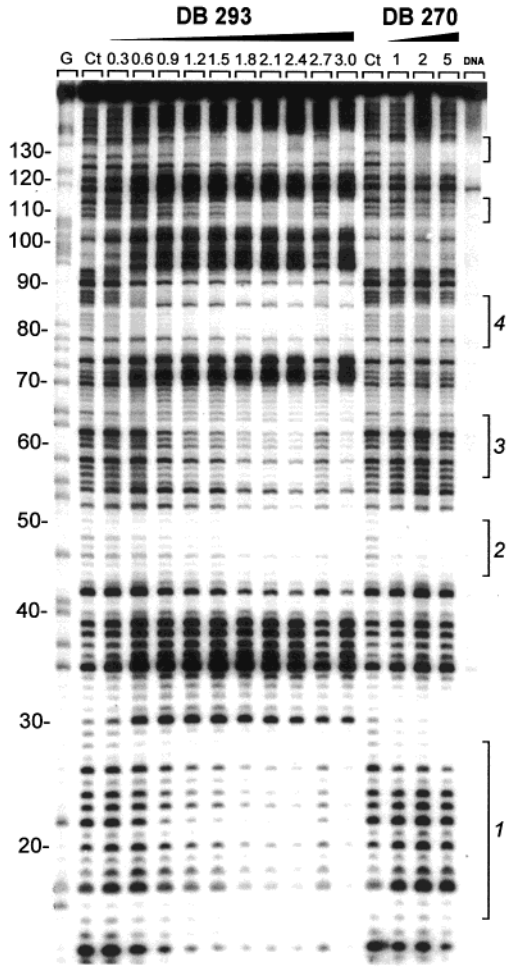

FIGURE 6: DNase I footprinting of DB270 and DB293 on the 160-bp *tyr*T fragment. Numbers on the left side of the gel refer to the standard numbering scheme for the nucleotide sequence of the DNA fragment, as indicated in Figure 7. Other details as for Figure 2.

sequences such as that corresponding to site 3 which encompasses the sequence TGTTACGTTG. DB293 can interact with this type of sequence but with a lower affinity compared to an ATGA site, as indicated by the footprinting plots shown in Figure 8.

As with the other DNA fragments, there are for the *tyr*T DNA a number of regions adjoining the drug-protected sites where the DNase I cutting rate has been substantially enhanced relative to the control. This is particularly evident for the T residue at position 30 flanking the 5′-ATGA·TCAT motif in site 1, suggesting again that DB293 induces specific structural changes of the DNA double helix. Altogether, the footprinting data are mutually consistent and reveal that DB293 interacts strongly with the mixed base pair sequence 5′-ATGA·TCAT and that upon binding, the drug introduces unusual specific structural changes in adjacent regions of the double helix.

SPR and T_m Analysis

Oligomers Based on Footprinting of the 265-mer pBS DNA Fragment. A pivotal finding in our investigations of the unusual interactions of DB293 with DNA is the footprint in the 90−100-bp region of the pBS 265-mer fragment (site 3 in Figure 2). This mixed base-pair sequence does not show a similar footprint with other closely related furan derivatives

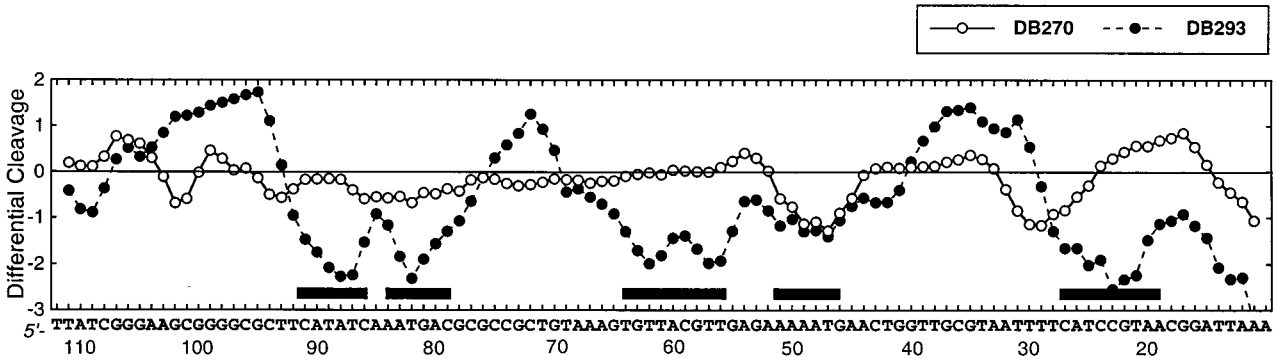

FIGURE 7: Differential cleavage plots comparing the susceptibility of the 160-mer *tyr*T fragment to DNase I cutting in the presence of 5 μM DB270 and 1.5 μM DB293. Other details as for Figure 4B.

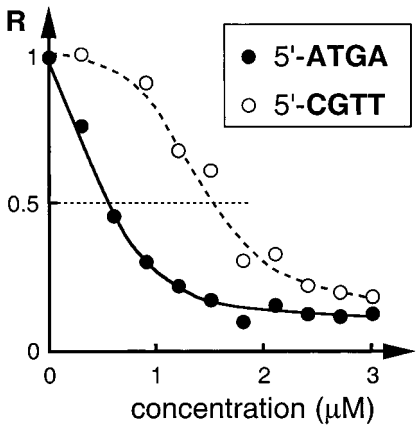

FIGURE 8: Footprinting plots for the binding of DB293 to sites ATGA and CGTT. Other details as for Figure 3.

or with classical minor groove binding agents such as netropsin. To obtain more molecular detail on this unusual interaction, we investigated the binding of DB293 and several related analogues to oligo2 (Figure 9), whose sequence is based on site 3 in the pBS fragment (Figure 2). T_m experiments demonstrated strong and highly cooperative binding of several DB293 molecules to the oligomer. The T_m curves are biphasic with a derivative peak near the free DNA T_m and a peak for complex melting at high temperature (not shown). The free DNA peak did not disappear, and the upper peak increased in amplitude until a ratio of approximately four molecules of DB293 were added to oligo2. This is a relatively short DNA sequence, and the binding of so many molecules of DB293 is surprising and suggests some type of compound stacking in the complex.

Because of the complicated binding observed with oligo2, we sought to simplify the system for quantitative analysis. To accomplish this, oligo2 was divided into two similar smaller oligomers, oligo2-1 and oligo2-2 from the 5′- and 3′-ends of oligo2, respectively (Figure 9). Both of these oligomers show similar, strong binding of DB293, and a sensorgram of the results with oligo2-2 is shown in Figure 10 (a sensorgram of the results with oligo2-1 has been shown in Supporting Information to reference *16*). For reference, a sensorgram of DB293 binding to the same amount of oligo1 in an SPR experiment is also shown in Figure 10. The binding of DB293 to both oligo2-1 and 2-2 saturates at an RU value that indicates two molecules of DB293 are bound per oligomer hairpin and that the binding is highly cooperative. Only one molecule of DB293 binds to oligo1 over the same concentration range. As previously described (*15*, *16*, *21*, *38*), fitting of the SPR results yields equilibrium constants for the interaction of DB293 with the oligomers, and *K* values for binding of DB293 to oligo1 and to oligo2-1 and oligo2-2 are collected in Table 1. Binding of DB293 to both oligo2-1 and oligo2-2 is very similar and exhibits pronounced positive cooperativity with $K2$ significantly greater than $K1$ in both cases. The $K2$ values obtained by fitting the oligo1 SPR results are much less, and saturation of any secondary binding sites is not reached in the accessible concentration range. Any binding of DB293 after the primary AATT site is saturated probably represents weak electrostatic interactions of the dication with the oligomer phosphate groups. The difference in $K2$ values then serves as the key feature to distinguish strong cooperative dimer binding from other types of weaker and less specific and noncooperative interactions.

As described above, footprinting of DB293 is also observed in more classical DNA AT sequences, and to compare this type of interaction with that observed in the oligo2 sequences, an oligomer with a central ATAT sequence was also investigated. As with the AATT sequence (Figure 10), the results are much simpler than with oligo2-1 and oligo2-2 and indicate that one molecule of DB293 binds strongly to the ATAT sequence (Table 1). Only weak secondary binding is observed after the strong site is saturated. The highly cooperative binding observed with the ATGA sequences is not present with the AT oligomer sequences. Similar results with these sequences are obtained with classical minor groove binding agents as expected for monomer minor groove interactions. The diphenylfuran derivative, DB75, and the dibenzimidazolefuran, DB270, that are closely related to DB293 also show strong binding to the AT sequence DNA oligomers, but they show only weak binding to the ATGA sequences, as previously illustrated with DB270 (*16*).

Context Dependence of the Binding of DB293 to ATGA Sequences. The similar feature of oligomers 2-1 and 2-2, both of which show strong 2:1 binding of DB293, is a central ATGA sequence. Both oligomers have the same length duplex stem with a connecting hairpin loop the same distance from the ATGA sequence. To determine the effect of the position of the hairpin loop on binding of DB293 to ATGA, 2 bp were added to oligo2-1 between the ATGA sequence and the loop (oligo21-9bp in Figure 9). There was little change in binding as a result of these changes, suggesting that the length of the stem and the proximity of the ATGA binding site to the hairpin loop are not major features of the cooperative 2:1 binding of DB293 to oligomers 2-1 and 2-2.

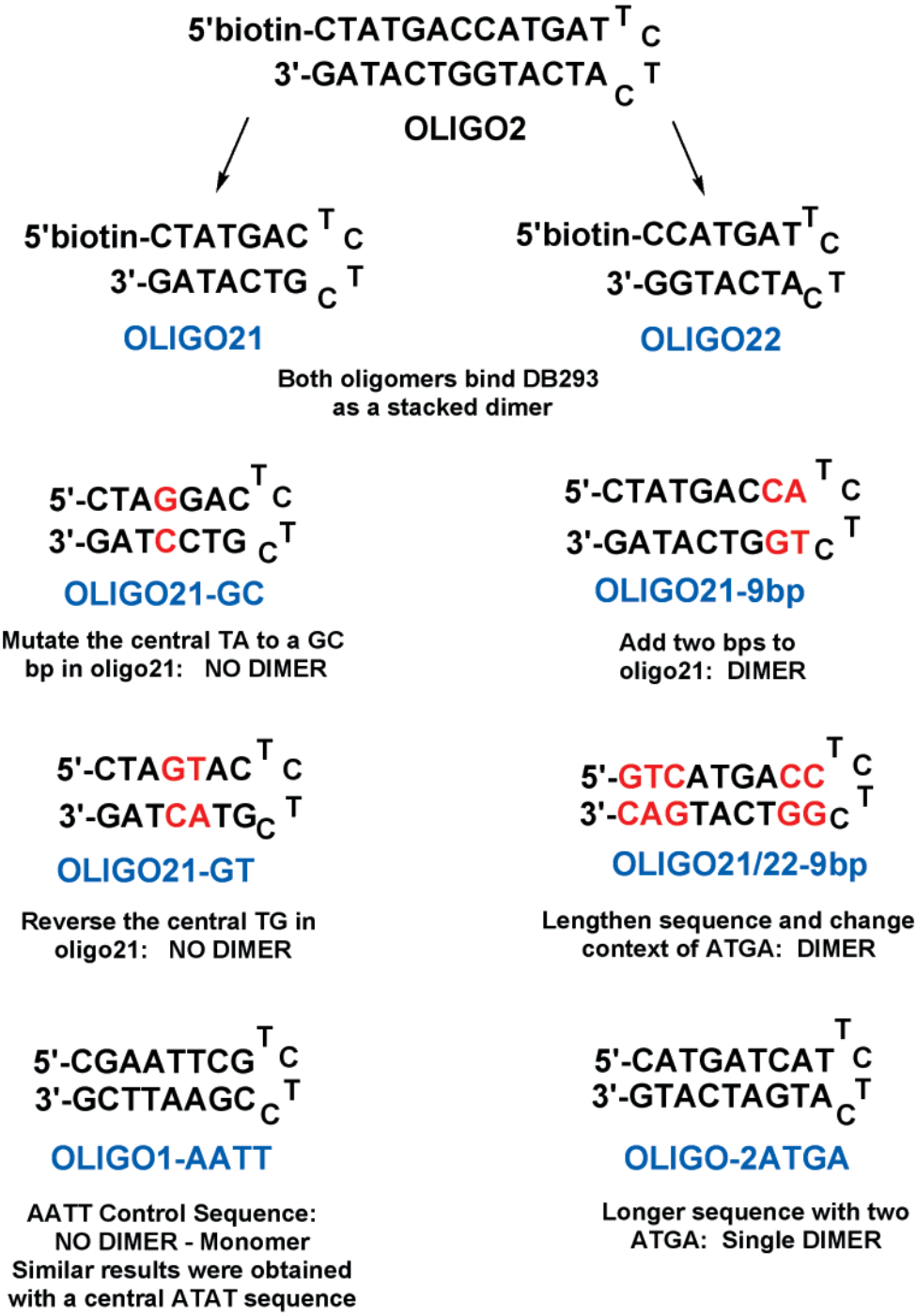

FIGURE 9: Oligonucleotide sequences for the DNAs used in BIAcore SPR experiment are shown.

The next context question is based on the fact that the ATGA sequence is flanked by T...C in oligo2-1 while the flanking bases in oligo2-2 are C...T. This observation raises the question of whether the dimer binding at ATGA in the minor groove can occur in the context of two flanking CG base pairs. To answer this question, the sequence CATGAC was incorporated into oligo21/22-9bp (Figure 9). Again, cooperative dimer binding with $K2$ significantly greater than $K1$, as with oligomers 2-1 and 2-2, was observed with this modified sequence (Table 1). These SPR binding results provide clear support for the footprinting observations that DB293 binds strongly to ATGA sequences with no significant context dependence for sequences investigated up to this time.

Effects of Reversing the ATGA Sequence to AGTA on the 2:1 Binding Mode of DB293. The results presented above show that the strong and unusual footprints observed with DB293 as well as the cooperative binding in SPR experiments are correlated with the presence of an ATGA sequence. An important question that is not answered by these results is whether DB293 can also read the same sequence but in the reverse direction (5′-AGTA-3′). Is there a directionality to the binding of DB293 to ATGA? To begin to answer this question, we first looked at the effect of mutating the T base of ATGA to G to convert a TA to a GC base pair (oligo21-GC in Figure 9). This sequence shows a much lower $K2$ value with DB293 than oligo2-1 and oligo2-2. The strong positive cooperativity is not present, and it is clear that this

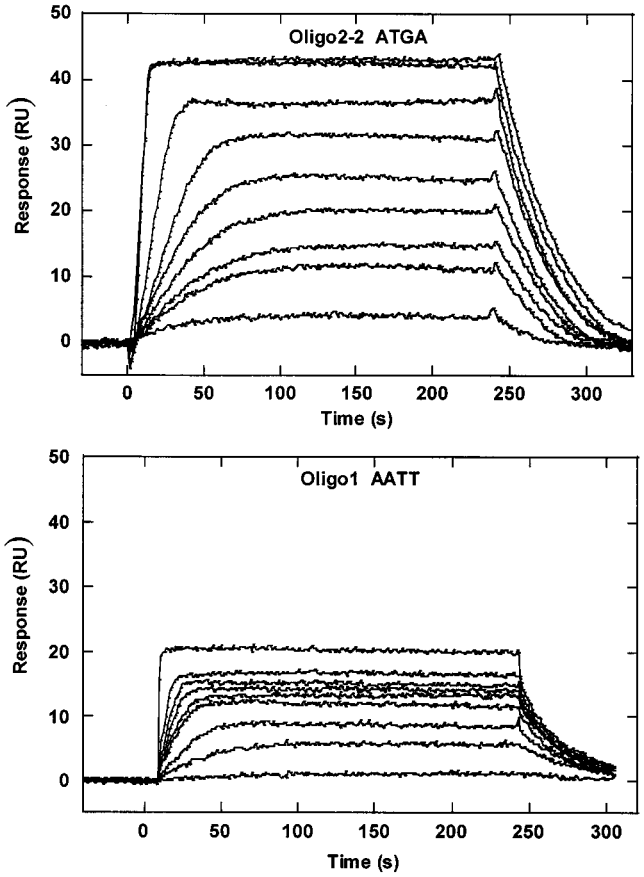

FIGURE 10: SPR sensorgrams for binding of DB293 to the oligo2-2 and oligo1 sequences (Figure 9) at 25 °C. The unbound DB293 concentrations in the flow solutions range from 1 nM in the lowest curve to 1 μM in the top curve. MES buffer with 0.1 M NaCl was used in the experiment.

Table 1: SPR Binding Constants of DB293 with Different Oligomers[a]

sequence	$K1$ (M^{-1})	$K2$ (M^{-1})
CTATGAC	1.8e+6	6.2e+7
CCATGAT	1.3e+6	7.0e+7
CATGATCAT	0.4e+6	4.1e+7
GTCATGACC	0.3e+6	9.5e+7
CTAGGAC	1.4e+6	2.6e+6
CTAGTAC	1.0e+6	1.6e+5
CGATATCG	7e+6	1e+5
CGAATTCG	2.3e+7	0.7e+5

[a] The DNA sequences show one strand of the duplex stem of the hairpin used in the BIAcore SPR experiments. All of the hairpin DNA samples have a TCTC loop sequence. The SPR experiments were conducted in MES buffer with 0.1 M NaCl at 25 °C.

single mutation in the ATGA sequence is very detrimental to the dimer interactions of DB293.

The next question addressed is whether converting the second G in the mutated AGGA sequence to a T to create the reverse of oligo2-1 (oligo21-GT in Figure 9) would restore cooperative binding of DB293. In fact, the DB293 $K2$ value for interaction with the reverse sequence is significantly weaker than for its binding to the single mutation sequence AGGA, and it is clear that no significant dimer interactions occur with the reverse sequence. DB293 thus reads ATGA but only from the 5′ to 3′ direction. NMR studies clearly show that the phenyl group of DB293 is at the 5′-end of ATGA while the benzimidazole is directed to the 3′-GA end of the sequence in oligo2-1, and this is the only binding orientation of DB293. The dimer binding mode is thus highly specific for the DNA ATGA sequence and the compound orientation.

CONCLUSION

This study constitutes the first detailed biochemical and biophysical characterization of the DNA sequence recognition properties of the phenylfuranbenzimidazole diamidine derivative DB293 which we recently identified as the first unfused aromatic dication capable of forming stacked dimers in the DNA minor groove of GC-containing sequences (*16, 21*). The perfect complementarity between the footprinting and SPR results demonstrates that DB293 binds preferentially as a dimer to the minor groove of sequences containing the ATGA motif. Although the drug can also interact with conventional AT-rich sequences in a 1:1 complex, binding to ATGA-containing sequences is highly preferred via the cooperative formation of 2:1 drug−DNA complexes. Substitution of AGGA for ATGA disrupts the formation of 2:1 drug−DNA complexes, leading thus to considerably weaker DNA interactions. The drug is also very sensitive to the orientation of the target sequence: 5′-ATGA but not 3′-ATGA accommodates two DB293 molecules in the minor groove. This observation suggests that directionality plays an important role in the drug−DNA recognition process.

The various chemical and enzymic footprinting experiments also reveal that the drug dimer introduces specific local perturbation of the DNA helix in the vicinity of the ATGA recognition sequence. The exact nature of these specific DB293-induced structural changes is not yet known. They may correspond to an opening of the minor groove required to accommodate two stacked drug molecules. Experiments using different chemical probes (such as $KMnO_4$ and diethyl pyrocarbonate) are in progress to investigate further the effect of drug binding on local DNA structure (*39, 40*).

The very specific stoichiometry and strong binding of DB293 to oligo2-1 obtained in the experiments described here suggested that the complex would be excellent for structure determination. Initial NMR structural results (unpublished data) show that the two molecules of DB293 bind in a stacked, antiparallel manner to the minor groove of the ATGA sequence. The stacking of two molecules of DB293 for binding to DNA is extremely important for DNA recognition. Since DNA is a double helix, each of the DB293 molecules in the stacked complex can recognize bases on one of the DNA strands of the double helix. The G base on the ATGA strand of the duplex is recognized by the benzimidazole isomer of one molecule of the dimer with the NH pointed out of the groove while the A base on the 3′-TACT-5′ complementary strand is recognized by the benzimidazole isomer of the other molecule of the dimer with the NH pointed into the groove. The two AT base pairs on the ends of the ATGA sequence form specific hydrogen bonds with the terminal amidine groups on the benzimidazole rings of the two DB293 molecules in the stacked complex. Each base pair of the recognition sequence thus has specific hydrogen bond interactions with at least one of the molecules of the dimer. Our ongoing NMR, SPR, and footprinting experiments will hopefully permit identification of the molecular rules for 1:1 and 2:1 DB293−DNA recognition,

as it exists with the pyrrole/imidazole-containing polyamides (*41*, *42*). Our rational drug design strategy may provide a second class of small molecules capable of direct readout of predetermined sequences of double-stranded DNA. Drugs such as furamidine, DB293, and similar unfused aromatic dications, which appear to enter the cell easily to accumulate in the nucleus (unpublished results), represent attractive candidates for the control of gene expression in human cells. Beyond its promising therapeutic potential for the treatment of infectious diseases such as *Pneumocystis carinii* pneumonia, DB293 can serve as a template for future generations of dicationic sequence-specific binding dimers.

ACKNOWLEDGMENT

We thank William Laine and Brigitte Baldeyrou for the plasmid preparations and radioactive labeling of the DNA fragments.

SUPPORTING INFORMATION AVAILABLE

Three figures showing EDTA·Fe^{II} footprinting of DB293 on the 265-bp *Eco*RI−*Pvu*II restriction fragment, DNase footprinting of DB293 and netropsin bound to the 232-bp *Eco*RI−*Pvu*II restriction fragment from pUC19, and anthraquinone footprinting of netropsin and DB270 bound to the 232-bp *Eco*RI−*Pvu*II restriction fragment from plasmid pUC19 (4 pages). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Blagburn, B. L., Drain, K. L., Land, T. M., Moore, P. H., Kinard, R. G., Lindsay, D. S., Kumar, A., Shi, J., Boykin, D. W., and Tidwell, R. R. (1998) *J. Parasitol. 84*, 851−856.
2. Rahmathullah, S. M., Hall, J. E., Bender, B. C., McCurdy, D. R., Tidwell, R. R., and Boykin, D. W. (1999) *J. Med. Chem. 42*, 3994−4000.
3. Bell, C. A., Dykstra, C. C., Aiman, N. A. I., Cory, M., Fairley, T. A., and Tidwell, R. R. (1993) *Antimicrob. Agents Chemother. 37*, 2668−2673.
4. Dykstra, C. C., McClernon, D. R., Elwell, L. P., and Tidwell, R. R. (1994) *Antimicrob. Agents Chemother. 38*, 1890−1898.
5. Neidle, S., Kelland, L. R., Trent, J. O., Simpson, I. J., Boykin, D. W., Kumar, A., and Wilson, W. D. (1997) *Bioorg. Med. Chem. 7*, 1403−1408.
6. Hildebrandt, E., Boykin, D. W., Kumar, A., Tidwell, R. R., and Dystra, C. C. (1998) *J. Eukaryotic Microbiol. 45*, 112−121.
7. Bailly, C., Dassonneville, L., Carrasco, C., Lucas, D., Kumar, A., Boykin, D. W., and Wilson, W. D. (1999) *Anti-Cancer Drug Des. 14*, 47−60.
8. Fitzgerald, D. J., and Anderson, J. N. (1999) *J. Biol. Chem. 274*, 27128−27138.
9. Boykin, D. W., Kumar, A., Spychala, J., Zhou, M., Lombardi, R. L., Wilson, W. D., Dykstra, C. C., Jones, S. K., Hall, J. E., Tidwell, R. R., Laughton, C., Nunn, C. M., and Neidle, S. (1995) *J. Med. Chem. 38*, 912−916.
10. Trent, J. O., Clark, G. R., Kumar, A., Wilson, W. D., Boykin, D. W., Hall, J. E., Tidwell, R. R., Blagburn, B. L., and Neidle, S. (1996) *J. Med. Chem. 39*, 4554−4562.
11. Francesconi, I., Wilson, W. D., Tanious, F. A., Hall, J. E., Bender, B. C., Tidwell, R. R., McCurdy, D., and Boykin, D. W. (1999) *J. Med. Chem. 42*, 2260−2265.
12. Boykin, D. W., Kumar, A., Xiao, G., Wilson, W. D., Bender, B. C., McCurdy, D. R., Hall, J. E., and Tidwell, R. R. (1998) *J. Med. Chem. 41*, 124−129.
13. Wilson, W. D., Tanious, F. A., Ding, D., Kumar, A., Boykin, D. W., Colson, P., Houssier, C., and Bailly, C. (1998) *J. Am. Chem. Soc. 120*, 10310−10321.
14. Stephens, C. E., Tanious, F., Kim, S., Wilson, W. D., Schell, W. A., Perfect, J. R., Franzblau, S. G., and Boykin, D. W. (2001) *J. Med. Chem. 44*, 1741−1748.
15. Mazur, S., Tanious, F. A., Ding, D., Kumar, A., Boykin, D. W., Simpson, I. J., Neidle, S., and Wilson, W. D. (2000) *J. Mol. Biol. 300*, 321−337.
16. Wang, L., Bailly, C., Kumar, A., Ding, D., Bajic, M., Boykin, D. W., and Wilson, W. D. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 12−16.
17. Pelton, J. G., and Wemmer, D. E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5723−5727.
18. Geierstanger, B. H., Mrksich, M., Dervan, P. B., and Wemmer, D. E. (1994) *Science 266*, 646−650.
19. Geierstanger, B. H., and Wemmer, D. E. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 463−493.
20. Wemmer, D. E., and Dervan, P. B. (1997) *Curr. Opin. Struct. Biol. 7*, 355−361.
21. Wang, L., Carrasco, C., Kumar, A., Stephens, C. E., Bailly, C., Boykin, D. W., and Wilson, W. D. (2001) *Biochemistry 40*, 2511−2521.
22. Li, K., Davis, T. M., Bailly, C., Kumar, A., Boykin, D. W., and Wilson, W. D. (2001) *Biochemistry 40*, 1150−1158.
23. Ratmeyer, L., Zapp, M. L., Green, M. R., Vinayak, R., Kumar, A., Boykin, D. W., and Wilson, W. D. (1996) *Biochemistry 35*, 13689−13696.
24. Hopkins, K. T., Wilson, W. D., Bender, B. C., McCurdy, D. R., Hall, J. E., Tidwell, R. R., Kumar, A., Bajic, M., and Boykin, D. W. (1998) *J. Med. Chem. 41*, 3872−3878.
25. Fasman, G. D. (1995) Nucleic Acids. in *Handbook of Biochemistry and Molecular Biology* (Fasman, G. D., Ed.) Vol. I, CRC Press, Cleveland, OH.
26. Drew, H. R., Weeks, J. R., and Travers, A. A. (1985) *EMBO J. 4*, 1025−1032.
27. Bailly, C., and Waring, M. J. (1995) *J. Biomol. Struct. Dyn. 12*, 869−898.
28. Wilson, W. D., Tanious, F. A., Fernandez-Saiz, M., and Rigl, C. T. (1997) *Methods Mol. Biol. 90*, 219−240.
29. Armitage, B., Yu, C., Devadoss, C., and Schuster, G. B. (1994) *J. Am. Chem. Soc. 116*, 9847−9859.
30. Breslin, D. T., and Schuster, G. B. (1996) *J. Am. Chem. Soc. 118*, 2311−2319.
31. Henderson, P. T., Armitage, B., and Schuster, G. B. (1998) *Biochemistry 37*, 2991−3000.
32. Armitage, B., and Schuster, G. B. (1997) *Photochem. Photobiol. 66*, 164−170.
33. Waring, M. J., and Bailly, C. (1994) *J. Mol. Recognit. 7*, 109−122.
34. Fox, K. R., and Waring, M. J. (1984) *Nucleic Acids Res. 12*, 9271−9285.
35. Portugal, J., and Waring, M. J. (1987) *Eur. J. Biochem. 167*, 281−289.
36. Portugal, J., and Waring, M. J. (1987) *FEBS Lett. 225*, 195−200
37. Fox, K. R., Yan, Y., and Gong, B. (1999) *Anti-Cancer Drug Des. 14*, 219−230.
38. Davis, T. A., and Wilson, W. D. (2000) *Anal. Biochem. 284*, 348−353.
39. Fox, K. R. (1992) in *Advances in DNA Sequence Specific Agents*, Vol. 1, pp 167−214, JAI Press Inc., Stamford, CT.
40. Bailly, C., and Waring, M. J. (1997) in Drug−DNA Interactions: Methods, Case Studies, and Protocols. *Methods Mol. Biol. 90*, 51−79.
41. Dervan, P. B., and Bürli, R. W. (1999) *Curr. Opin. Chem. Biol. 3*, 688−693.
42. Urbach, A. R., and Dervan, P. B. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 4343−4348.

BI0108453