J|A|C|S
ARTICLES
Published on Web 09/29/2004

Thiophene-Based Diamidine Forms a "Super" AT Binding Minor Groove Agent

Sirish Mallena,[†] Michael P. H. Lee,[§] Christian Bailly,*[,‡] Stephen Neidle,*[,§] Arvind Kumar,[†] David W. Boykin,*[,†] and W. David Wilson*[,†]

Contribution from the Department of Chemistry, Georgia State University, P.O. Box 4098, Atlanta, Georgia 30302-4098, INSERM U-524 et Laboratoire de Pharmacologie Antitumorale du Centre Oscar Lambret, IRCL, Place de Verdun, 59045 Lille, France, and Cancer Research UK Biomolecular Structure Group, The School of Pharmacy, University of London, London WC1N 1AX, United Kingdom

Received March 30, 2004; E-mail: wdw@gsu.edu

Abstract: The DNA minor groove is the interaction site for many enzymes and transcription control proteins and as a result, development of compounds that target the minor groove is an active research area. In an effort to develop biologically active minor groove agents, we are preparing and exploring the DNA interactions of a systematic set of diamidine derivatives with a powerful array of methods including DNase I footprinting, biosensor-SPR methods, and X-ray crystallography. Surprisingly, conversion of the parent phenyl-furan-phenyl diamidine to a phenyl-thiophene-benzimidazole derivative yields a compound with over 10-fold-increased affinity for the minor groove at AT sequences. Single conversion of the furan to a thiophene or a phenyl to benzimidazole does not cause a similar increase in affinity. X-ray results indicate a small bond angle difference between the C−S−C angle of thiophene and the C−O−C angle of furan that, when amplified out to the terminal amidines of the benzimidazole compounds, yields a very significant difference in the positions of the amidines and their DNA interaction strength.

Introduction

The DNA minor groove is an important target site for enzymes and transcription control proteins, and it has been a particularly attractive target site for development of synthetic agents for therapeutic purposes and for sequence selective recognition of DNA.[1−7] The minor groove binding diamidine, pentamidine, for example, has been successfully used in the clinic against a range of human parasitic diseases for many years.[1,8−10] Furamidine (DB75 in Figure 1) is a synthetic diamidine that has significantly better activity against a number of microbial parasites with lower toxicity than pentamidine.[1,10,11]

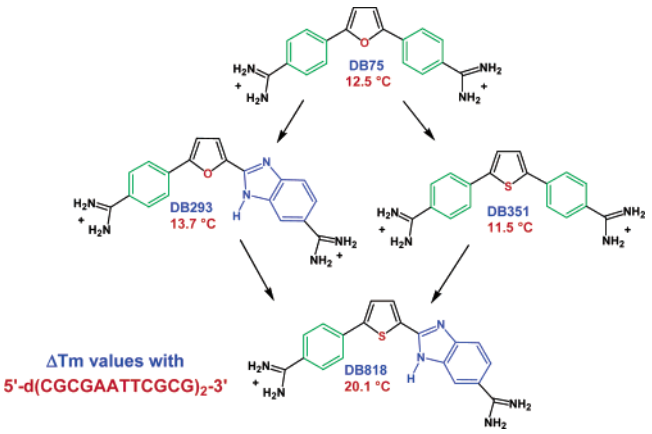

Figure 1. Structures of the compounds and ΔTm values with a DNA oligomer are shown.

A limitation on the use of all diamidines is their lack of oral bioavailability but in a very exciting development an orally active prodrug of DB75 has been developed and is currently in phase II clinical trials as an antitrypanosomal agent.[1,10] The prodrug is readily converted into DB75 after passing through the intestinal membrane barrier.

In addition to the therapeutic applications described above, the minor groove has been the recent target of choice for development of synthetic agents for sequence selective binding of DNA. Very successful development of sequence-selective polyamide analogues of distamycin[4] has been described by the Dervan,[2] Lown,[7] and Lee[12] groups. Bruice and co-workers have

[†] Department of Chemistry, Georgia State University.
[‡] IRCL, Place de Verdun, 59045 Lille, France.
[§] Biomolecular Structure Unit, University of London.

(1) Tidwell, R. R.; Boykin, D. W. In *DNA and RNA Binders: from Small Molecules to Drugs*; Demeunynck, M., Bailly, C., Wilson, W. D., Eds.; Wiley-VCH: 2003; Vol. 2, Chapter 16, pp 414−460.
(2) (a) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Bio.* **2003**, *13*, 284−299. (b) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215−2235.
(3) Neidle, S. *Nat. Prod. Rep.* **2001**, *18*, 291−309.
(4) Wemmer, D. E. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 439−461.
(5) Bailly, C.; Chaires, J. B. *Bioconjugate Chem.* **1998**, *9*, 513−538.
(6) Chaires, J. B. *Curr. Opin. Struct. Biol.* **1998**, *8*, 314−320.
(7) Sondhi, S. M.; Reddy, B.; Lown, J. W. *Curr. Med. Chem.* **1997**, *4*, 313−358.
(8) Cory, M.; Tidwell, R. R.; Fairley, T. A. *J. Med. Chem.* **1992**, *35*, 431−438.
(9) Fairley, T. A.; Tidwell, R. R.; Donkor, I.; Naiman, N. A.; Ohemeng, K. A.; Lombardy, R. J.; Bentley, J. A.; Cory, M. *J. Med. Chem.* **1993**, *36*, 1746−1753.
(10) (a) Bouteille, B.; Oukem, O.; Bisser, S.; Dumas, M. *Fund. Clin. Pharmacol.* **2003**, *17*, 171−181. (b) Fairlamb, A. H. *Trends Parasitol.* **2003**, *19*, 488−494. (c) Ismail, M. A.; Brun, R.; Easterbrook, J. D.; Tanious, F. A.; Wilson, W. D.; Boykin, W. D. J. *J. Med. Chem.* **2003**, *46*, 4761.
(11) Das, B. P.; Boykin, D. W. *J. Med. Chem.* **1977**, *20*, 531−536.

shown that some polyamide-benzimidazole hybrid compounds can recognize long base pair sequences of DNA in a 1:1 complex.[13] Armitage and co-workers have made the exciting discovery that selectively designed cyanine dyes can stack as dimers in the minor groove in AT sequences to form a tandem helical array that has the potential to generate new nanomaterials.[14] For the heterocyclic diamidines described above, the DNA complexes of DB75 and the related furan-benzimidazole derivative DB293 (Figure 1) have been thoroughly studied and they display many interesting, useful and quite different properties.[1,3,15,16] Both compounds bind strongly as monomer complexes in the DNA minor groove in AT sequences of at least four base pairs in length. Both have a significant hydrophobic energetic component in formation of their AT complexes but also depend on other energetic terms such as van der Waals, electrostatic and H-bonding to drive complex formation.[17,18] DB75 binds much more strongly to AT rich sequences than pentamidine and, because of the fact that many parasitic microorganisms have AT rich DNA, the AT binding appears to be correlated with their biological activity.[1] DB75 is an important compound because of its therapeutic properties but it is essential to develop additional compounds in case of drug resistance as well as to develop compounds with activities against other organisms. Compounds that exploit the nature of the AT rich DNAs in both the nucleus and particularly in the kinetoplast of organisms such as trypanosomes are of particular interest.

DB293 is also one of the most exciting new minor groove binding compounds discovered recently since it can form a unique dimer structure that stacks in an antiparallel arrangement in the minor groove to recognize sites that contain GC as well as AT base pairs.[1,15,16,18] In addition to its classical binding monomer binding mode in AT sequences, this compound offers an entirely new dimer interaction mode for sequence specific recognition of mixed DNA sequences. Since it is readily taken up into cells,[19] it obviously offers unique opportunities to develop new types of sequence-selective DNA targeted therapeutic agents. For these reasons, we are engaged in a series of synthetic, biological, and biophysical studies to understand the DNA complexes of these compounds in detail and to use this information to design new compounds with additional and improved therapeutic and DNA recognition abilities. As part of this effort, thiophene derivatives with amidine cationic groups (Figure 1), which are directly related to DB75 and DB293, have been prepared and their DNA complexes characterized with a broad array of different methods including X-ray determination of a complex structure with DB818 bound at an AATT minor groove site. We show in this paper that although the thiophene derivatives differ by only a single atom from the parent furans, they have significant shape changes that can make a large difference in DNA complex formation depending on the other groups in the molecular system and their interactions with DNA.

Materials and Methods

Buffers and Oligonucleotides. The double stranded polymers poly(dAT)$_2$ and poly(dGC)$_2$ were purchased from Pharmacia (U. S.), and were used in UV−Vis titration experiments. In Biacore and CD experiments three 5′-biotin labeled hairpin duplexes namely d(Biotin-CGAATTCGTCTCCGAATTCG) (AATT hairpin), d(Biotin-CATATATA CCCCTATATATG) (AT hairpin), d(Biotin-CGCGCGC TTTTGCGCGCG) (Figure 1) were purchased with HPLC purification (Midland Certified Reagent Co., Midland, TX). The MES buffer used in these experiments contained 0.010 M [2-(N-morpholino) ethanesulfonic acid] (MES), 0.001 M EDTA, 0.1 M NaCl, pH 6.25.

Compound Synthesis: 2-{2-[5(6)-Amidino]benzimidazolyl}-5-[(4-amidino)phenyl] thiophene trihydrochloride (DB818, Figure 1). A solution of potassium carbonate (4.4 g, 0.32 mole) in 10 mL water was added to a stirred solution of 5-bromo-2-(dimethoxymethyl)thiophene19 (3.3 g, 0.014 mole), 4-cyanophenylboronic acid (2.35 g, 0.016 mole) in 35 mL dioxane, under nitrogen, and stirring was continued for 15 min. After addition of Pd (PPh_3)$_4$ (0.32 g, 2mole%) the mixture was heated at reflux for 8 h (TLC monitored), cooled, diluted with water and extracted with 2 × 35 mL chloroform. The chloroform layer was washed with water and stirred with 20% HCl (40 mL) for 35 min. The organic layer was washed with water, 5% sodium bicarbonate, water, brine, and dried over sodium sulfate, passed through a bed of silica, dried and the solvent reduced to yield 5-(4-cyanophenyl)-2-thiophene carboxaldehyde as a white solid 2.35 g (77%), mp 208−210 °C dec, ^{1}H NMR (DMSO-d_6): δ 9.95 (s, 1H), 8.06 (d, 1H, J = 4 Hz), 7.98 (d, 2H, J = 8.8 Hz), 7.91 (d, 1H, J = 8.8 Hz), 7.88 (d, 1H, J = 4 Hz); ^{13}C NMR (DMSO-d_6): δ 184.0, 149.6, 143.5, 138.4, 136.5, 133.0, 127.2, 126.7, 118.0, 111.4; MS: *m*/*e* 213 (M+). Anal. Calcd. for $C_{12}H_7NOS$: C, 67.58; H, 3.30; N, 6.56. Found: C, 67.31; H, 3.41; N, 6.44.

A mixture of the thiophene aldehyde (2.13 g, 0.01 mol), 3,4-diaminobenzonitrile (1.33 g, 0.01 mol) and 1,4-benzoquinone (1.08 g, 0.01 mol) in 50 mL dry ethanol was heated under reflux (N_2 atmos.) for 8 h, cooled, diluted with ether, and filtered. The solid was collected and stirred with 1:3 mixture of EtOH and ether for 20 min., filtered, washed with ether, and vacuum-dried at 70 °C for 12 h to yield 2.10 g (60%) of the yellow

(12) Lacy, E. R.; Madsen, E. M.; Lee, M.; Wilson, W. D. In *Small Molecule DNA and RNA Binders*; Demeunynck, M., Bailly, C., Wilson, W. D., Eds.; Wiley-VCH: 2003; Vol. 2, pp 384−413.
(13) (a) Reddy, P. M.; Jindra, P. T.; Satz, A. L.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 7843−8. (b) Satz, A. L.; Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 86.
(14) (a) Garoff, R. A.; Litzinger, E. A.; Connor, R. E.; Irene, F.; Armitage, B. A. *Langmuir* **2002**, *18*, 6330−6337. (b) Wang, M.; Silva, G. L.; Armitage, B. A. *J. Am. Chem. Soc.* **2000**, *122*, 9977−9986.
(15) (a) Tanious, F. A.; Hemmelberg, D.; Bailly, C.; Czarny, A.; Boykin, D. W.; Wilson, W. D. *J. Am. Chem. Soc.* **2004**, *126*, 143−153. (b) Nguyen, B.; Hammelberg, D.; Bailly, C.; Colson, P.; Stanek, J.; Brun, R.; Neidle, S.; Wilson, W. D. *Biophys. J.* **2004**, *86*, 1028−1041. (c) Trent, J. O.; Clark, G. R.; Kumar, A.; Wilson, W. D.; Boykin, D. W.; Hall, J. E.; Tidwell, R. R.; Blagburn, B. L.; Neidle, S. *J. Med. Chem.* **1996**, *39*, 4554−4562.
(16) (a) Wang, L.; Bailly, C.; Kumar, A.; Ding, D.; Bajic, M.; Boykin, D. W.; Wilson, W. D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12−16. (b) Wang, L.; Carrasco, C.; Kumar, A.; Stephens, C. E.; Bailly, C.; Boykin, D. W.; Wilson, W. D. *Biochemistry* **2001**, *40*, 2511−2521. (c) Nguyen, B.; Tardy, C.; Bailly, C.; Colson, P.; Houssier, C.; Kumar, A.; Boykin, D. W.; Wilson, W. D. *Biopolymers* **2002**, *63*, 281−297. (d) Tanious, F. A.; Wilson, W. D.; Wang, Lei.; Kumar, A.; Boykin, D. W.; Marty, C.; Baldeyrou, B.; Bailly, C. *Biochemistry* **2003**, *46*, 13576−13586. (e) Bailly, C.; Tardy, C.; Wang, L.; Armitage, B.; Hopkins, K.; Kumar, A.; Schuster, G. B.; Boykin, D. W.; Wilson, W. D. *Biochemistry* **2001**, *40*, 9770−9779. (f) Bailly, C.; Dassonneville, L.; Carrosco, C.; Lucas, D.; Kumar, A.; Boykin, D. W.; Wilson, W. D. *Anti-Cancer Drug Des.* **1999**, *14*, 47−60.
(17) Mazur, S.; Tanious, F. A.; Ding, D.; Kumar, A.; Boykin, D. W.; Simpson, J. J.; Neidle, S.; Wilson, W. D. *J. Mol. Biol.* **2000**, *300*, 321−337.
(18) (a) Wang, L.; Kumar, A.; Boykin, D. W.; Bailly, C.; Wilson, W. D. *J. Mol. Biol.* **2002**, *317*, 361−374. (b) Lansiaux, A. L.; Tanious, F.; Mishal, Z.; Dassonneville, L.; Kumar, A.; Stephens, C. E.; Hu, Q.; Wilson, W. D.; Boykin, D. W.; Bailly, C. *Cancer Res.* **2002**, *62*, 7219−7229.
(19) (a) Lansiaux, A. L.; Dassonneville, L.; Faxompre, M.; Kumar, A.; Stephans, C. E.; Bajic, M.; Tanious, F. A.; Wilson, W. D.; Boykin, D. W.; Bailly, C. *J. Med. Chem.* **2002**, *45*, 1994−2002. (b) Lansiaux, A.; Tanious, F. A.; Mishal, Z.; Dassonneville, L.; Kumar, A.; Stephens, C. E.; Hu, O.; Wilson, W. D.; Boykin, D. W.; Bailly, C. *Cancer Res.* **2000**, *62*, 7219−7229.

brown 2-{2-[5(6)-cyano(benzimidazolyl)]}-5-(4-cyanophenyl) thiophene, mp 315−17 °C dec, ^{1}H NMR (DMSO-d_6/D_2O): δ 8.03 (d, 1H, J = 1.5 Hz), 7.91 (d, 2H, J = 8.4 Hz), 7.90 (d, 1H, J = 3.9 Hz), 7.85 (d, 2H, J = 8.4 Hz), 7.77 (d, 1H, J = 3.9 Hz), 7.70 (d, 1H, J = 8.4 Hz), 7.55 (d, 1H, J = 1.5 Hz, J = 8.4 Hz); ^{13}C NMR (DMSO-d_6): δ 149.1, 144.1, 137.0, 133.4, 132.8, 129.1, 126.9, 125.9, 125.6, 119.5, 118.2, 110.3, 104.1; MS: *m/e* 326 (M+). Anal. Calcd. for $C_{19}H_{10}N_4S\cdot1.25H_2O$: C, 65.40; H, 3.61; N, 16.06. Found: C, 65.32; H, 3.67; N, 16.47.

The bis-nitrile (1.8 g, 0.0055 mole) in 75 mL ethanol was saturated with HCl gas at 0 °C and stirred at room temperature for 8 days (TLC monitored). Ether was added and the resulting yellow imidate ester was filtered, washed with ether, and vacuum-dried at 30 °C for 5 h. The imidate ester hydrochloride [2.43 g (84%). 0.65 g (0.0012 mole)] was suspended in 65 mL ethanol, saturated with ammonia at 0 °C, and stirred at room temperature for 24 h. The solvent was removed, ether: ethanol (6:1) added and the solid filtered. The yellow amidine was suspended in 30 mL ethanol and saturated with HCl gas. The yellow crystalline amidine hydrochloride salt was filtered, washed with ether and dried in a vacuum at 70 °C for 12 h, to yield 0.39 g (68.5%), mp >350 °C dec, ^{1}H NMR (DMSO-d_6/D_2O): δ 8.05 (d, 1H, J = 1.5 Hz), 7.93 (d, 1H, J = 3.6 Hz), 7.92 (d, 2H, J = 8.4 Hz), 7.84 (d, 2H, J = 8.4 Hz), 7.45 (d, 1H, J = 3.6 Hz), 7.73 (d, 1H, J = 8.4 Hz, 7.63 (dd, 1H, J = 1.5 Hz, J = 8.4 Hz); ^{13}C NMR (DMSO-d_6): δ 166.6, 165.8, 149.3, 146.7, 141.2, 138.4, 138.3, 131.9, 131.3, 129.7, 127.9, 127.7, 126.9, 123.9, 123.0, 116.3, 115.8. FABMS: *m/e* 361-(M++1). Anal. Calcd. for $C_{19}H_{16}N_6S\cdot3HCl\cdot0.25H_2O$: C, 48.11; H, 4.14; N, 17.71. Found: C, 48.18; H, 4.56; N, 17.77.

Purification Of DNA Restriction Fragments And Radiolabeling for Footprinting. Plasmids were isolated from *E. coli* by a standard sodium dodecyl sulfate-sodium hydroxide lysis procedure and purified by banding in CsCl−ethidium bromide gradients. The 117-bp and 265-bp DNA fragments were prepared by 3′-[^{32}P]-end labeling of the *Eco*RI−*Pvu*II double digest of the pBS plasmid (Stratagene) using α-[^{32}P]-dATP (Amersham, 3000 Ci/mmol) and AMV reverse transcriptase (Roche). The 160-bp *tyrT* fragment was obtained from plasmid pKMΔ98 after digestion with the restriction enzymes *Eco*RI and *Ava*I. In each case, the labeled digestion products were separated on a 6% polyacrylamide gel under nondenaturing conditions in TBE buffer (89 mM Tris-borate pH 8.3, 1 mM EDTA). After autoradiography, the requisite band of DNA was excised, crushed, and soaked in water overnight at 37 °C. This suspension was filtered through a Millipore 0.22 μm filter 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, Electrophoresis, and Quantitation. 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. Stock solutions of DNase I were kept at −20 °C and freshly diluted to the desired concentration immediately prior to use. Footprinting experiments were performed essentially as recently described.[15,16] Briefly, reactions were conducted in a total volume of 10 μL. Samples (3 μL) of the labeled DNA fragments were incubated with 5 μL of the buffered solution containing the ligand at appropriate concentration. After 30 min incubation at 37 °C, the digestion was initiated by the addition of 2 μL of a DNase I solution 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 samples were then heated at 90 °C for 4 min and chilled in ice for 4 min prior to electrophoresis. DNA cleavage products were resolved by polyacrylamide electrophoresis under denaturing conditions (0.3 mm thick, 8% acrylamide containing 8 M urea). After electrophoresis (about 2.5 h at 60 W, 1600 V in Tris-Borate-EDTA buffered solution), gels were soaked in 10% acetic acid for 10 min, transferred to Whatman 3 MM paper, and dried under vacuum at 80 °C. 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 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 dimethylsulfate followed by piperidine-induced cleavage at the modified guanine bases.

Thermal Melting. Thermal melting experiments were conducted with Cary 3 or Cary 4 spectrophotometers interfaced to microcomputers as previously described.[15,16] A thermistor fixed into a reference cuvette was used to monitor the temperature. The DNA is added to 1 mL of buffer (MES with 0.1 M NaCl added) in 1 cm path length quartz cells and the concentration determined by measuring the absorbance at 260 nm. Experiments were generally conducted at a concentration of 5×10^{-5} M base pairs for poly(dAT), poly(dAT)$_2$ and 5×10^{-5} M duplex for d(CGCGAATTCGCG)$_2$. For experiments with complexes a ratio of 0.3 compound per base pair for poly(dAT) and poly-(dAT)$_2$ and a ratio of one compound per oligomer duplex for d(CGCGAATTCGCG)$_2$ was used.

Spectrophotometric Titrations. CD spectra were obtained with a computer controlled Jasco J-710 Spectrophotometer in a 1 cm quartz cell in MES buffer. The desired ratios of compound to DNA were obtained by adding compound to the cell containing a constant amount of DNA. Absorption spectra were collected on a Cary 4 spectrophotometer under the same conditions. During the spectrophotometric titrations, the spectra of the complex at different DNA (bp) to compound ratios were collected in the region above 300 nm to avoid DNA absorption interference. The DNA was added to a cell containing a constant amount of compound until all compound was bound.

Biosensor-Surface Plasmon Resonance. Experiments were performed on a BIAcore 3000 instrument and BIAcore sensorchips with streptavidin linked to dextran were used for DNA immobilization by using established methods.[17,18,20] DNA immobilization was performed at a flow rate of 5 μL/min with DNA concentrations of 500−1000 pM. The amounts of DNA hairpin immobilized on the flow cells were in the range of 400 RUs and one flow cell was always left as a blank for reference. The sensorgrams (SPR angle in resonance units, RU, vs time) were collected at flow rate of 10, 20, or 50 μL/min with the

(20) (a) Davis, T. M.; Wilson, W. D. *Anal. Biochem.* **2000**, *284*, 348−353. (b) Davis, T. M.; Wilson, W. D. *Methods Enzymol.* **2001**, *340*, 22−51.

samples dissolved in degassed MES10 buffer with 5×10^{-3} % v/v Surfactant P20. 10 mM glycine, pH 2.0 solution was also used in surface regeneration. The RU values were averaged over a one-minute time span in the steady state response region and were converted to r (moles of compound bound per mole of DNA duplex) for binding analysis as previously described.[15−18] The binding constants of the compounds were obtained from fitting a plot of r vs free compound concentration (from the flow solution with one- and two-site interaction models (for the one-site model, $K_2 = 0$)

$$r = (K_1 \cdot C_{free} + 2 \cdot K_1 \cdot K_2 \cdot C^2_{free})/(1 + K_1 \cdot C_{free} + K_1 \cdot K_2 \cdot C^2_{free}) \quad (1)$$

where K_1 and K_2 are the equilibrium binding constants; C_{free} is the concentration of the compound in the flow solution. This concentration is fixed by the flow solution concentration because the solution is continuously refreshed by flow over the chip surface; $r = RU_{eq}/RU_{max}$ and represents moles of compound bound per mole of DNA hairpin, where RU_{eq} is the averaged response at the steady-state level, and RU_{max} is the maximum response for binding one molecule of compound and DNA per binding site and is predicted with the following equation

$$RU_{max} = (RU_{DNA} \cdot MW_{compound} \cdot RII_r)/MW_{DNA} \quad (2)$$

where RU_{DNA} is the amount in response units of DNA immobilized, MW is the molecular weight of compound and DNA, respectively, and RII_r is the refractive index increment ratio of compound to refractive index increment of DNA.[20] The RII_r values for these compounds are between 1.25 and 1.40.

Isothermal Titration Calorimetry (ITC). The ITC experiments were carried out with a MicroCal VP−ITC (MicroCal Inc., USA) isothermal titration calorimeter interfaced to a PC for instrument control and data collection with Origin 5.0 software. In a typical titration 10 μL of a 0.15 mM compound solution was added every 300 s to a total of 29 injections to DNA in the sample cell at 10 μM duplex. The heat associated with each injection is observed as a peak that corresponds to the power required to keep the sample and reference cells at identical temperature.[21] The areas of the peaks produced over the course of a titration can be converted to heat output per injection by integration and correction for cell volume, sample concentration, and subtraction of the heats of dilution. Titrations of compound into buffer were carried out to determine the contribution to the binding enthalpy from the heat of compound dilution.[17,18,21] Each titration was repeated 3 to 5 times. Finally, binding enthalpies were obtained by fitting the corrected data to an appropriate binding model.[17,18,21]

Crystallization and Data Collection. The oligonucleotide d(CGCGAATTCGCG) (HPLC-purified, Eurogentec Ltd.) was annealed at 85 °C for 6 min in 30 mM sodium cacodylate buffer at pH 7.0 before use. Crystals were grown by vapor diffusion at 285 K from hanging drops. The crystal used for data collection was obtained from a 6 μL drop containing 10 mM $MgCl_2$, 0.5 mM double-stranded DNA, 0.75 mM DB818 compound, 5% v/v (±)-2-methyl-2,4-pentanediol (MPD), 1.0 mM spermine tetrahydrochloride, in 30 mM sodium cacodylate buffer at pH 7.0, equilibrated against a reservoir of 700 μL of 50% v/v MPD solution. Crystals grew as large rods within three weeks. Typical crystals were approximately 1.0 × 0.3 × 0.3 mm in dimension.

Table 1. Crystallographic Data

	DB818−DNA complex
DNA sequence	$d(CGCGAATTCGCG)_2$
space group	$P2_12_12_1$
unit cell dimensions (Å)	a=25.677, b=39.921, c=65.185
resolution (Å)	17−1.77
wavelength (Å)	1.54178
V_m (Å^3/Da)	2.80
measured reflections > 1σ	43881
unique reflections	6667
completeness (%)	95.2
in the outer shell (%)	92.5 (1.83−1.77 Å)
I/σ (I) for the data set	70.18
I/σ (I) for the outer shell	13.16
data redundancy	6.58
R_{merge} (%)	4.3
R_{merge} for the outer shell (%)	13.5
refinement	
resolution range	8.0−1.77
reflections	5894
reflections with $F_o > 4\sigma(F_o)$	5495
parameters refined	2307
restraints	2538
R-factor {$F_o > 4\sigma(F_o)$} (%)	21.50
R_{free} {$F_o > 4\sigma(F_o)$} (%)	29.06
rms deviations	
bond lengths (Å)	0.008
bond angle distances (Å)	0.028
no. of atoms	
DNA	486
DB818 ligand	26
magnesium−water complex	7
full occupancy waters	54
mean temperature factor of atoms (Å^2)	
DNA	24.8
DB818 ligand	29.6
magnesium−water complex	21.5
waters	33.0

A crystal was mounted in a loop and flash-frozen in a nitrogen gas stream at 103 K and maintained at that temperature using an Oxford Cryostream (Oxford Cryosystems). X-ray diffraction data were collected on a Rigaku RAXIS IV image plate detector with CuK_α radiation from a Rigaku RU200 rotating anode generator (100 mA, 50 kV) and an Osmic focusing mirror system. The crystal-to-detector distance was set at 110 mm. Each frame was exposed for 35 min at 2.5° oscillation. Seventy-three frames were collected for the structure, from 0° to 180°, which resulted in a maximum resolution of 1.77 Å. Indexing and data processing were carried out using the program package DENZO and SCALEPACK.[22]

Structure Solution and Refinement. The structure was solved by molecular replacement. A suitable isomorphous DNA model was chosen for structure solution and refinement, this was the berenil-$d(CGCGAATTCGCG)_2$ complex (PDB entry 2DBE or NDB entry GDL009).[23] Only the DNA part of this model was used for structure solution and refinement. Structure solution involved rigid-body refinement using the CNS program package[24] starting at 4.0 Å resolution, and this was gradually

(21) Wiseman, T.; Willston, S.; Brandts, J. F.; Lin, L.-N. *Anal. Biochem.* **1989**, *179*, 131−137.
(22) Otwinowski, Z. M.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307−326.
(23) Brown, D. G.; Sanderson, M. R.; Skelly, J. V.; Jenkins, T. C.; Brown, T.; Garman, E.; Stuart, D. I.; Neidle, S. *EMBO J.* **1990**, *9*, 1329−1334.
(24) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr. D* **1998**, *54* (Pt 5), 905−21.

increased to the highest resolution at 1.77 Å. The *R*-factor at this point was 40.5% and R_{free} was 41.5%. The Fourier map was generated using CNS and the region around the minor groove of DNA was observed using CHAIN[25] to check whether there is a continuous electron density region. This density was observed in the minor groove. At this stage, repeated simulated annealing and positional refinement with gradual increase in resolution was performed again using CNS. The *R*-factor was 34.0% and the R_{free} was 30.0% before the ligand was included. The model of ligand was added to the structure at 2.4 Å resolution. The hexa-coordinated magnesium-water complex was included into the structure at 2.0 Å resolution; the *R*-factor and R_{free} at this point were 31.9% and 30.8%, respectively.[26] The refinement was continued to 1.77 Å resolution; the *R*-factor and R_{free} were 33.4% and 32.2%, respectively, at this stage of the refinement. The remaining part of the refinement was transferred to SHELX-97 for atomic refinement.[27] The refinement began at 4.0 Å until it reached 1.77 Å. The Fourier map was checked during refinement using Xfit portion of the XtalView program.[28] Water refinement was carried out after this stage. The automated solvent searching program SHELXWAT was then used to locate molecules, as well as manual peak-picking to locate water molecules. The water molecules were accepted according to B-factor value (<50 Å^2), strong electron density, standard distance and geometry. The SHELX-97 refinement was continued to give the final *R*-factor of 21.5% and R_{free} of 29.1%. The atomic coordinates and structure factors have been deposited to the RCSB Protein and Nucleic Acid Data Banks with accession number 1VZK. Crystallographic data for the structure are given in Table 1.

Results

Thermal Melting. Relative Affinity Ranking. We have found for a wide array of different diamidines that determination of melting temperature increases (T_m of complex − T_m of free DNA, where T_m is the maximum in a derivative plot of dA_{260}/dT) provides a rapid and accurate method for ranking compounds according to their binding affinity.[1] Increases in T_m values for compound-DNA oligomer complexes are shown in Figure 1 and example T_m curves are in Figure 2. Large increases in melting temperature are observed for all of the compounds with the oligomer containing an −AATT− binding site. Under the same conditions, no significant increase in the T_m of a DNA duplex oligomer with an alternating CG sequence was observed with these compounds.

With the AT DNA sequences the furan to thiophene conversion for DB75 to DB351 makes little change in the ΔT_m values but the same conversion for DB293 to DB818 results in a surprisingly large T_m increase for a single atom change. In the same manner conversion of a phenyl to a benzimidazole for DB75 to DB293 results in a modest T_m increase but the same change in the thiophene series (DB351−DB818) results in a significantly larger ΔT_m. In additional T_m determinations with AT containing DNA polymers, DB818 again had abnormally large ΔT_m values (not shown) compared to the other compounds in Figure 1. Clearly the amidine-benzimidazole-thiophene system provides a very favorable unit for strong molecular recognition of AT sequences in DNA.

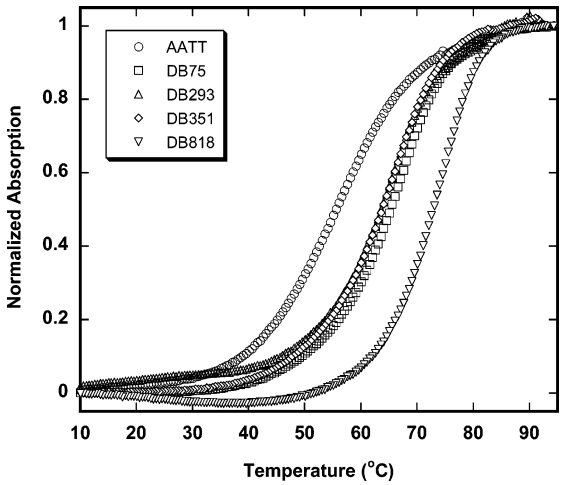

Figure 2. Thermal melting profiles of duplex d(CGCGAATTCGCG)$_2$ in the absence (circles) and in the presence of DB75 (squares), DB293 (up triangles), DB351 (diamonds) and DB818 (down triangles). Experiments were conducted in MES10 buffer.

DNase I Footprinting. Relative affinity and Sequence Specificity. Footprinting is the method of choice for defining sequence specificity in a broad sequence context for compounds that bind to DNA. Footprinting results have been published for both furan derivatives in Figure 1, DB75[16c,f] and DB293,[16d,e] but not for the two thiophene analogues. The DNA sequence recognition specificities of these four compounds are directly compared with three DNA fragments of 117, 160 and 265-bp in length that provide a diversity of potential binding sites (Figure 3). As expected for diamidines with a curvature that closely matches that of the DNA minor groove, both of the furans have strong footprints and the densitometric analysis of the gels (Supporting Information, Figure S1) indicates that the binding sites essentially correspond to AT sequence regions.

The footprinting profiles obtained with DB818 are distinct from those obtained with DB293 but are comparable to those seen with DB75 and DB351. Densitometric results for the gels for DB818 and DB293 are shown in Figure S1 (those for DB75 and DB351 were omitted for clarity) and all results are summarized in Table 2. All results show that DB818 binds to AT sites, as is the case with the other three compounds, but there is a key difference, the footprints with DB818 are much stronger than those at the same concentrations for the other compounds (compare the 1 μM lanes in Figure 3). The footprints with DB818 are easily seen at 0.5 μM whereas a concentration of about 5 μM is required to detect similar effects with DB75 and DB351. In contrast, DB293 can recognize additional types of sequences, mainly −ATGA− sites, and this has been linked to its ability to form a stacked antiparallel dimer in the minor groove at ATGA (Wang et al., 2000, 2001, 2002). DB293 can bind to AT sequences in a monomer complex but it binds more strongly to −ATGA− sequences as a dimer.

Spectroscopy. Relative Affinities and Binding Modes. Absorption spectral titrations of the compounds in Figure 1 with AT DNAs generally result in large decreases in absorbance with slight red shifts when the compound to DNA ratio is high but larger red shifts and an absorbance increase or no significant

(25) Sack, J. S. *J. Mol. Graphics* **1988**, *6*, 224−225.
(26) Tsunoda, M.; Karino, N.; Ueno, Y.; Matsuda, A.; Takenaka, A. *Acta Crystallogr. D* **2001**, *57*, 345−348.
(27) Sheldrick, G. M.; Schneider, T. R. *Methods Enzymol.* **1997**, *276*, 319−343.
(28) McRee, D. E. *J. Struct. Biol.* **1999**, *125*, 156−65.

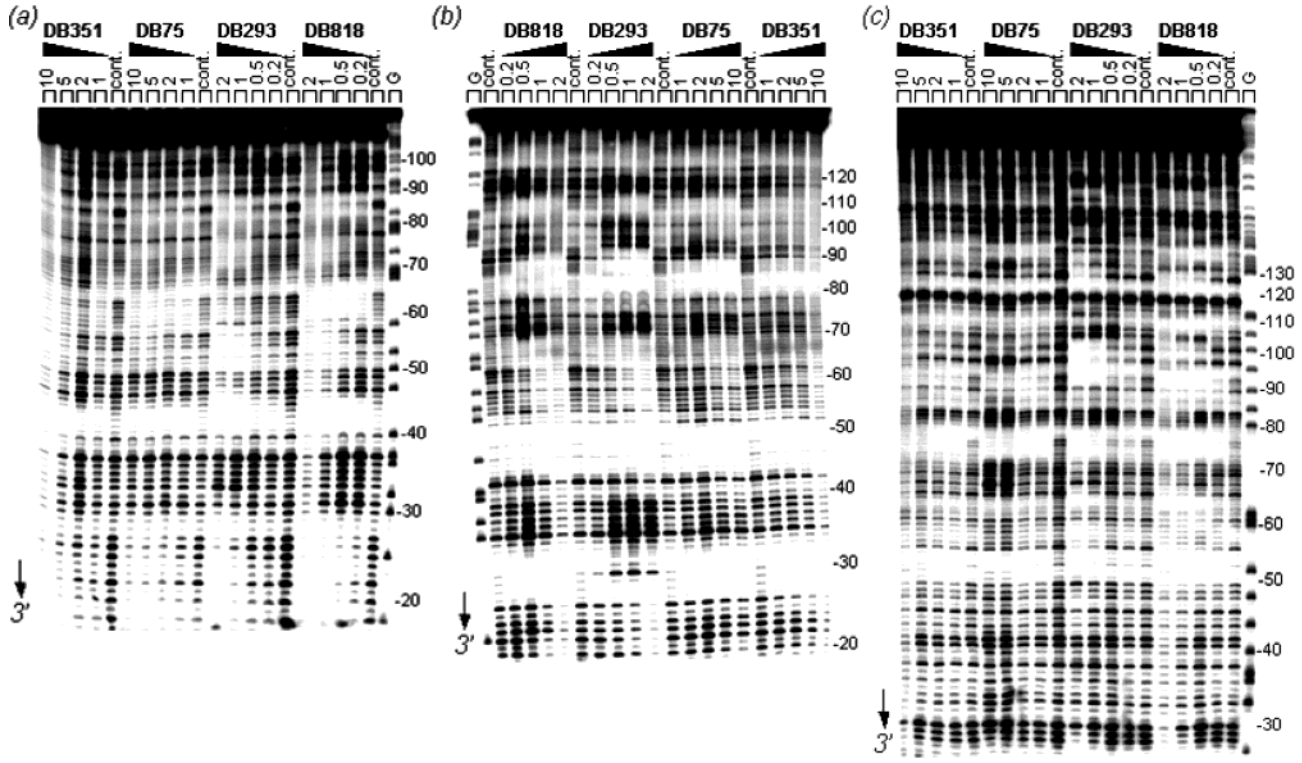

Figure 3. DNase I footprinting with the (a) 117-bp, (b) 160-bp and (c) 265-bp DNA fragments at graded concentrations of the four DB compounds. The concentration (μM) of the drug tested is shown at the top of the appropriate gel lanes. Tracks labeled "cont" contained no drug. Tracks labeled "G" represent dimethyl sulfate-piperidine markers specific for guanine. Numbers at the right side of the gel refer to the sequence and nucleotide position of the different fragments.

Table 2. Sequences of the DNA Fragments Showing DNase I Protection in the Presence of DB818 and DB293, Inferred from Differential Cleavage Plots.

DB818		DB293	
	117-mer		
5′-ATTGTAATA (20−28)[a]	+++[b]	5′-TGTAATA (20−26)	+
5′-TTTT (63−66)	+		
	160-mer		
5′-TTT (29−31)	+	5′-ATCCGT (21−26)	++
5′-TAAAGTGTTACGTT (56−69)	++	5′-GTGT (62−65)	+
5′-TCATATCA (85−92)	++	5′-ATCATATCA (85−93)	++
	265-mer		
5′-TAAAGG (61−66)	+		
5′-ATT (76−78)	+		
5′-ATTACG (90−95)	++	5′-ATTA (90−93)	+
5′-ATGA (101−104)	+	5′-ATGACCATGA (95−104)	++

[a] The position of the sequence is indicated in parentheses. [b] +++, ++, and + indicate strong, medium and weak protection.

change as more DNA is added and only the strongest binding sites are populated. Example spectra for DB818 and DB351 with the AATT oligomer (Figure 1) are shown in Figure 4. This type of behavior does not prove a binding mode but it is consistent with results for other similar compounds that bind in the DNA minor groove in AT sequences. The initial decrease is characteristic of stacking at high compound-DNA ratios followed by partitioning of the compound to the strong binding sites in the minor groove as the ratio of compound to DNA is decreased. At high ratios of DNA to compound in the titrations only the strongest binding sites are populated.

Titrations of DB351 and DB818 with the AT polymer, poly-(dA-T)$_2$, give qualitatively similar results to those with the AATT oligomer (Supplementary Figure S2). The results with the GC polymer, poly(dG-C)$_2$ (Supplementary Figure S3) also illustrate that the compounds can bind to GC sequences but only at higher concentrations. The absorbance decreases on addition of the GC DNA, as expected, but there was no significant increase in absorbance as more DNA was added. Similar behavior, which has been observed for other minor groove binding compounds, is characteristic of stacking interactions but no changes characteristic of a minor groove complex is observed even at high ratios of GC DNA to compound.

Circular dichroism spectral changes provide very useful information for defining the binding mode of unfused aromatic cations, such as those in Figure 1, with DNA.[29,30] For minor groove complexes, strong positive induced CD signals in the compound absorption region are expected on complex formation with DNA. The CD spectral titrations of DB351 and DB818

(29) Lyng, R.; Rodger, A.; Norden, B. *Biopolymers* **1992**, *32*, 1201−1214.
(30) Rodger, A.; Norden, B. In *Circular Dichroism and Linear Dichroism*; Oxford University Press: New York, 1997.

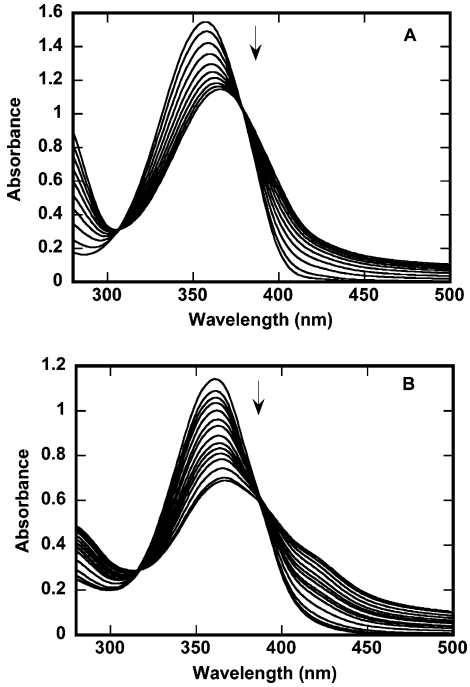

Figure 4. Spectrophotometric titration of DB351 (A) and DB818 (B) with AATT hairpin DNA. The titrations were conducted in MES10 buffer at 25 °C. The concentration of the compound in each case is 2.5×10^{-5} M. In DB351, the ratios of the compound to AATT hairpin DNA from the top to the bottom are free compound, 4.24, 2.12, 1.42, 1.06, 0.85, 0.71, 0.61, 0.53, and 0.47 (A). In DB818, the ratios of the compound to AATT hairpin DNA from the top to the bottom are free compound, 42.44, 21.26, 14.20, 8.50, 6.08, 4.74, 3.88, 3.29, 2.85, 2.52, 2.25, 2.04, 1.79, 1.72, 1.59, 1.48, and 1.39 (B).

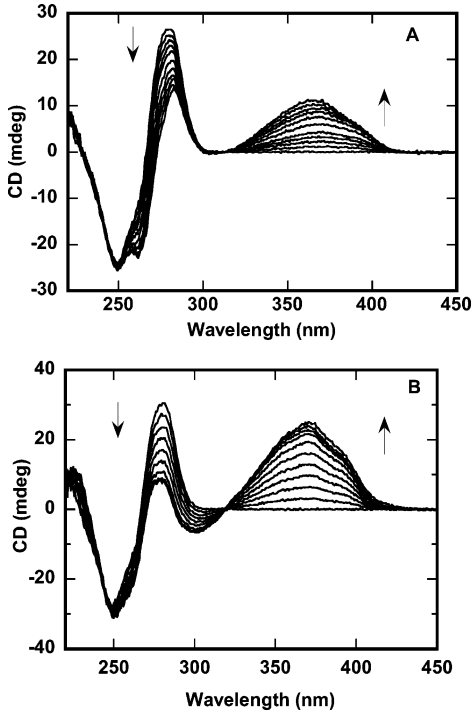

Figure 5. CD spectra of DB351 (A) and DB818 (B) with the AATT hairpin DNA (10 μM bp). The experiments were conducted in MES10 buffer at 25 °C. In DB351 the ratios of compound to DNA from the top to the bottom at 375 nm are 0, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0, 1.2, 1.5, and 2.0. In DB818 the ratios of compound to DNA from the top to the bottom at 375 nm are 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, and 2.0.

with AATT hairpin DNA have large positive induced CD signals above 300 nm where DNA does not absorb (Figure 5). This large induced CD is characteristic of a minor groove binding mode. CD spectra for the complexes of DB818 and DB351 with the alternating AT polymer also display large induced CD spectral changes as expected for minor groove complexes (Supporting Information, Figure S4) and similar changes have been observed for DB75 and DB293.[15,16] With both DNAs the CD changes are characteristic for minor groove complex formation. It should be noted that with both AT DNAs the CD changes are significantly larger for DB818 than for DB351 in agreement with the T_m and footprinting results that indicate much stronger AT DNA interactions for DB818. Over the same range of ratios of compound to DNA, much smaller (DB818) or essentially no (DB351) CD spectral changes are seen in the compound absorption region on addition of GC polymer DNA (Figure S4), as expected for weak nonspecific binding interactions.

Quantitative Comparison of DNA Interactions. Biosensor-SPR. To quantitatively compare the interactions of the compounds of Figure 1 with specific DNA sequences, biosensor SPR experiments were carried out. 5′-biotin labeled hairpin DNA duplexes with alternating AT and GC sequences as well as an AATT sequence (see Materials) were immobilized on three flow cells of a BIAcore sensor chip and the fourth flow cell was left blank for reference subtraction. Sensorgrams for the interaction of DB818 and DB351 with the AATT sequence are compared in Figure 6 and, for two such closely related compounds, the sensorgrams are strikingly different. Both the rates of association and dissociation of DB351 are too fast to

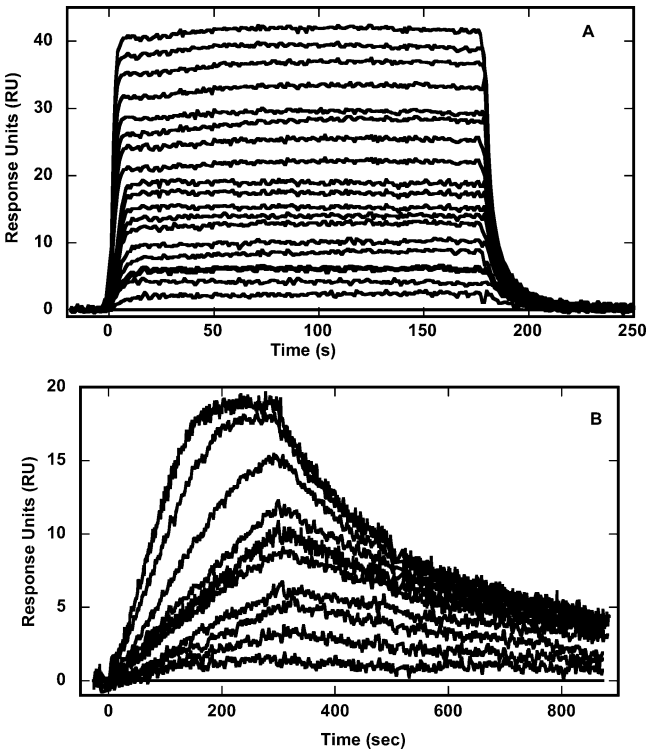

Figure 6. SPR sensorgrams of DB351 (A) and DB818 (B) with the AATT hairpin DNA. The compound concentrations for DB351 from bottom to top are 0 to 0.8 μM (A). The compound concentrations for DB818 from bottom to top are 0 to 10 nM (B).

be determined at the limit of resolution in SPR experiments. A steady-state SPR response is reached for DB351 in a few seconds after initiation of compound or buffer flow. With

Table 3. Binding Constants and Experimental Thermodynamic Values at 298 K for Binding of Furan and Thiophene Compounds to DNA[a]

	AATT				ATAT		CGCG	
compd	K (M^{-1})	ΔG (kcal/mol)	ΔH (kcal/mol)	$T\cdot\Delta S$ (kcal/mol)	K (M^{-1})	ΔG (kcal/mol)	K (M^{-1})	ΔG (kcal/mol)
DB75	1.4×10^7	−9.7	−2.2	7.5	6.0×10^7	−10.6	0.8×10^6	−8.0
DB351	7.5×10^6	−9.4	−3.0	6.4	2.8×10^6	−8.8	3.0×10^3	−4.7
DB293[b]	1.1×10^7	−9.6	−3.6	6.0	- -	- -	- -	- -
DB818	2.8×10^8	−11.5	−6.7	4.8	2.7×10^8	−11.5	- -	- -

[a] Experiments were carried out in MES10 buffer at 25 °C; The K and ΔG values are from biosensor-SPR experiments. Experimental error limits for DB75, DB351, and DB293 are approximately ±10%. The slow kinetics and very strong binding of DB818 increases the error to approximately 20%. The ΔH values are from ITC experiments and have an experimental error of approximately ±10%.

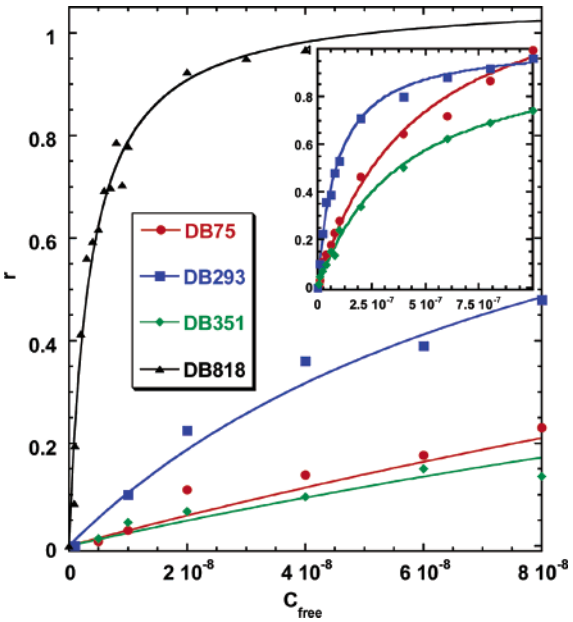

Figure 7. RU values from the steady-state region of SPR sensorgrams were converted to r by r = RU/RUmax and are plotted against the unbound compound concentration (flow solution) for DB75, DB293, DB351, and DB818 binding to the AATT hairpin DNA. The lines in the Figure were obtained by nonlinear least-squares fits of data by using eq 1. Experiments were conducted in MES buffer with 0.1 M NaCl added at 25 °C. The inset shows the same binding curves at higher concentration.

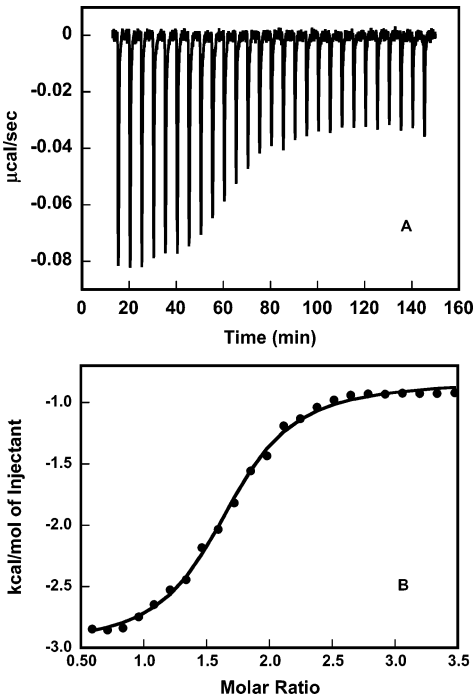

Figure 8. ITC curves for the binding of DB351 (150 μM) to the duplex d(GCGAATTCGC)$_2$ (10 μM base pairs) in MES10 at 25 °C (A). Results were converted to molar heats and plotted against the compound/DNA molar ratio (B). The line shows the fit to the results and gives best fit ΔH (Table 3) values for binding.

DB818 a steady state is only reached after over 100 seconds at concentrations near the saturation level for the compound. The responses for DB75 and DB293 are more similar to those for DB351. The differences in interaction strength for all four compounds can be more easily visualized in the plot of r (moles of compound bound/mole of hairpin DNA) versus C_{free}, the free compound concentration (Figure 7). As can be seen, the concentration for DB818 to reach saturation is at least 10 times lower than for the other compounds.

The compounds all have a strong binding constant ($K > 10^7$ M^{-1}) as well as weaker binding that can be seen at high concentrations as a gradual increase in the r versus C_{free} plots. We thus found better fits to all SPR results with a nonidentical two-site binding model. At concentrations above 1 microM with the compounds in Figure 1 we begin to have baseline subtraction problems, and it is difficult to quantitate the weak secondary binding. It is clear that the K_2 values are near 1×10^5 M^{-1} with DB75 and DB351, 100 times less than K_1, and are even smaller with the benzimidazole derivatives. The strong binding constants (K_1) determined by fitting the curves in Figure 7 with eq 1 (two-site model−see Materials) are collected in Table 3. The binding constants for DB75, DB293, and DB351 with the AATT sequence are between approximately $(1-3) \times 10^7$ M^{-1}, whereas that for DB818 is at least 10 times higher and may be underestimated because of its slow association rate at low concentration. Similar behavior is seen for the AT hairpin oligomer for all the compounds. All four compounds bind to the GC hairpin 100 or more times more weakly than with the AT sequences in agreement with, spectral changes, their lack of footprints and low Tm with GC sequences. The GC interaction may account for part of the secondary binding interactions described above for the AATT oligomer sequence.

Detailed Thermodynamics of the Interactions. ITC Experiments. To understand the energetic basis of the interaction differences between DB818 and the other compounds in Figure 1 in more detail, ITC experiments were conducted to determine the observed binding enthalpy, ΔH (Figure 8) and to allow calculation of ΔS (from $\Delta G = \Delta H - T\cdot\Delta S$). Integration with respect to time of the heats produced on serial injection of ligand into the −AATT− duplex yields the corresponding binding isotherm. When the K value is not too large (with respect to 1/[compound]) these curves can yield a K value. Figure 8B shows the binding isotherms obtained by titrating DB351 into DNA, after subtracting the heat of dilution for addition of the compound into buffer. The observed ΔH value for binding was

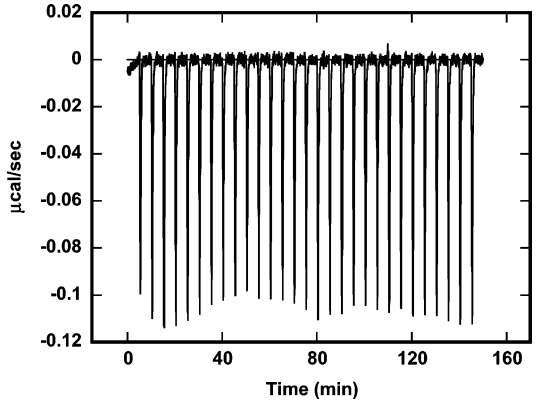

Figure 9. Representative ITC results for addition of DB818 to the AATT hairpin DNA at 35 °C. Experiments were conducted in MES10 buffer. Essentially all DB818 added in this titration range is bound to the AATT site and the average value of the integrated peaks was used to calculate ΔH in Table 3.

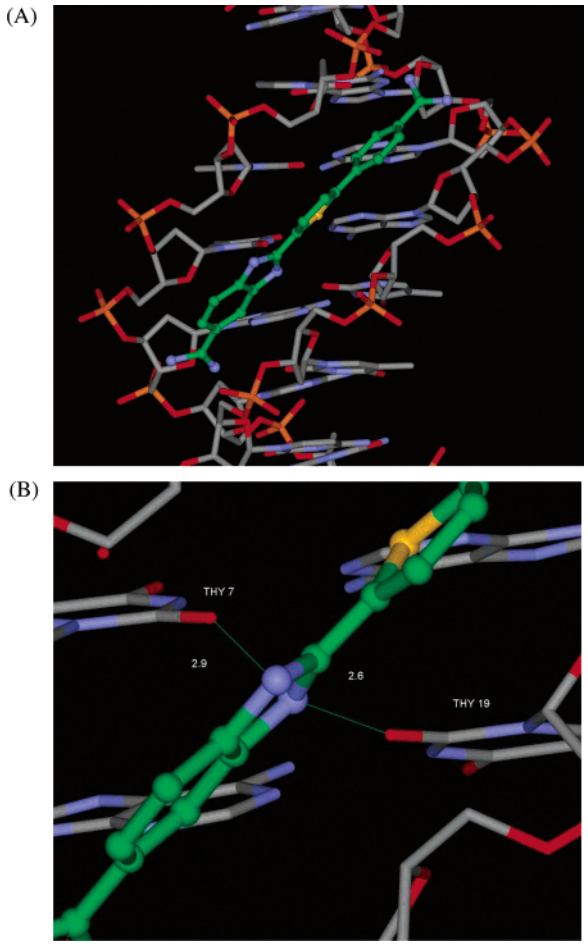

Figure 10. (A) Crystal structure of DB818 bound to the duplex sequence d(CGCGAATTCGCG)$_2$. The view is into the minor groove with DB818 C in green, N in blue and S in yellow; DNA C in gray, N in blue, O in red and P in yellow. The benzimidazole is at the bottom of this view and the two amidines are clearly seen at the ends of the bound compound. (B) Close-up view of the bifurcated pair of hydrogen bonds (2.7 and 2.9 Å), between the benzimidazole inner-facing ring nitrogen atom and O2 atoms of the central two thymines in the crystal structure of the complex between DB818 and d(CGCGAATTCGCG)$_2$.

obtained from fitting the curve with a two binding site model as with the SPR results described above (see Materials and Methods). Only the ΔH values for binding of the compounds to the strong binding site and ΔH values determined for the

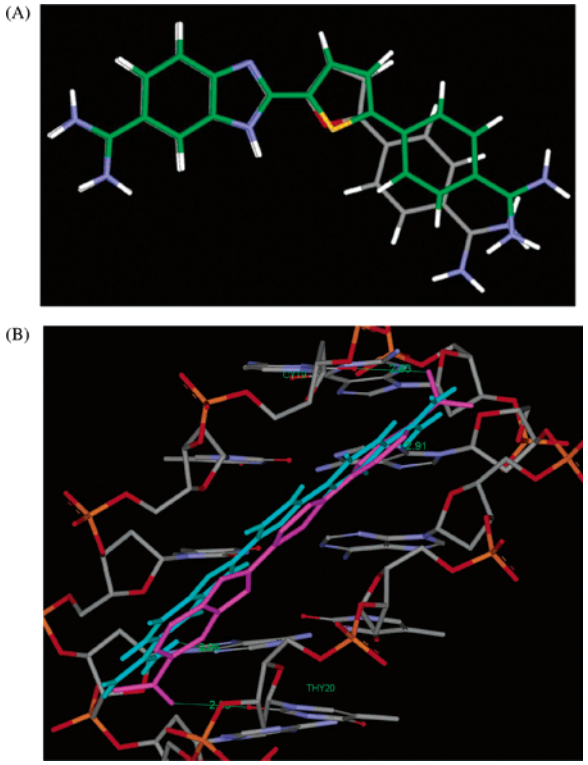

Figure 11. (A) Comparison of the optimized structures of DB818 and DB293: calculations were carried out using the MM2 force field parameter set. The thiophene, DB818 (shown as green) has a significantly less concave inner surface, as shown by the larger N...N amidinium distance of 14.4 Å compared to 12.9 Å in the oxygen analogue, DB293 (shown as gray). (B) Overlay of the optimized structure of DB818 and DB293 into the AATT binding site of d(CGCGAATTCGCG)$_2$.

strong site are collected in Table 3. The K_1 value can be obtained from the DB351 ITC titration curve and the value is in good agreement with that from SPR experiments. Even after subtraction of the heat of dilution for DB351, the ITC curve does not go to zero at ratios above one due to the weak secondary interactions that were also observed in SPR experiments at high ratios. A less extensive titration curve (Figure 9) was obtained for DB818 due to precipitation of the complex as saturation of the AATT site was approached at ITC concentrations. Only the ΔH value for binding can be obtained from these results because of the limited range and large K value. With the K from SPR experiments the ΔG and ΔS values can be also be calculated (Table 3). As can be seen from Table 3, replacing O by S in DB75 yielded only small changes in ΔG and ΔH values for binding to DNA, whereas replacing O by S in DB293 yielded much more favorable (more negative) ΔG and ΔH values. Thus the enhanced binding of DB818 to AATT is largely due to a more favorable enthalpy contribution and the entropy of binding for DB818 is actually less favorable than with the other compounds.

Crystal Structure of the DNA-DB818 Complex. The crystal structure of the DB818-d(CGCGAATTCGCG)$_2$ complex shows that the ligand is bound mostly within the AATT minor groove region of the B-type structure of the duplex DNA. The ligand is twisted along its length, with the four torsion angles defining its conformation having values of 42° (amidine-phenyl), 16° (phenyl-thiophene), 11° (thiophene-benzimidazole) and 24° (benzimidazole-amidine), on going from the phenyl end to the benzimidazole end of the molecule (Figure 10A). The net effect

of this is that the terminal amidinium groups are quite twisted with respect to each other.

There are a number of hydrogen-bond contacts between inner-facing atoms of the ligand and base edges that partially account for the affinity and specificity of binding. The inner nitrogen atom of the benzimidazole group is hydrogen-bonded in a bifurcated manner to atoms O2 of thymines 7 and 19 (2.90 and 2.70 Å respectively, Figure 10B). Such an arrangement has been observed in a number of benzimidazole-containing minor-groove complexes, including those of Hoechst 33258 and analogues. The amidinium group attached to the benzimidazole is hydrogen-bonded to O2 of T 20 (2.90 Å); and it is also the starting-point for an incompletely observed network of water−water contacts that extend along the groove toward the wider G/C region. The amidinium group at the other end of the ligand is also weakly hydrogen-bonded to a base O2 atom (3.20 Å), of cytosine 9.

The ligand−DNA hydrogen bonding observed for DB818 is in contrast to that predicted from computer modeling for the furan analogue, DB293 (Figure 11 A, there is at present no crystal structure for the DB293 complex). This ligand has a significantly less concave inner radius than DB818 (Figure 11 B), due to the smaller van der Waals radius of oxygen compared to sulfur. Thus DB293 cannot simultaneously make base-benzimidazole hydrogen bonds and at the same time be in position for base hydrogen bonding to both terminal amidinium groups. DB818, by virtue of its larger concave radius, can do so. This also enables DB818 to cover a five base-pair site compared to the four of DB293 and to have stronger compound-base interactions.

Discussion

Footprinting, T_m and spectroscopic results (Figures 3−6) show that all four diamidines in Figure 1 are AT specific DNA minor groove binding agents as predicted by their structure and from studies on the furan derivatives, DB75 and DB293.[1,16] These results also qualitatively illustrate that DB818 binds to AT sequences much more strongly than any of the other three diamidines. Conversion of the furan of DB75 to a thiophene does not enhance DNA interactions and conversion of one phenyl in DB75 to a benzimidazole does not cause a major change in the AT minor groove binding affinity. It was thus initially surprising that conversion of the diphenyl-thiophene, DB351 or the furan-benzimidazole, DB293, to DB818 (Figure 1) yields a compound that binds significantly better to AT sequences in DNA than the other diamidines. Both DB818 and DB293 form monomer complexes in −AATT− sequences and since DB818 differs by only a single atom from DB293, the significantly stronger binding was unexpected. The increased DNA interaction of DB818 cannot be due, for example, to simple differences in solution properties of S and O since DB351 actually binds slightly more weakly than DB75.

Biosensor-SPR and ITC experiments were used to quantitatively characterize the binding differences among these compounds.[31,32] The SPR method is particularly well-suited for determination of the strong interactions of compounds, such as the diamidines in Figure 1, with nucleic acids where very low concentrations must be used in order to obtain reliable values for the large equilibrium constants obtained with these systems. The compounds, particularly DB818, do not have strong enough spectroscopic signals to allow accurate determination of binding constants of the magnitude in Table 3. The only method that allowed determination of equilibrium constants for all compounds was the biosensor-SPR system. The interaction differences between DB818 and the other compounds of Figure 1 can be visualized in the very different footprints in Figure 3 and in the sensorgrams in Figure 7. At 0.1 M NaCl DB75, DB351, and DB293 all have equilibrium constants near 1×10^7 M^{-1}, a value typical of dications in this size range under these conditions, while DB818 has a K of over 1×10^8 M^{-1} (Table 3). Conversion of the furan of DB293 to a thiophene or conversion of a phenyl of DB351 to a benzimidazole yields the large increase in binding constant. These combined results indicate that both parts of the benzimidazole-thiophene module of DB818 are required for the large increase in DNA binding constant.

ITC experiments provide a direct method to obtain the binding enthalpy and with SPR K values allow a full thermodynamic characterization of DNA complexes. Such a full thermodynamic picture is essential to complement the structural perspective from crystallography presented here. ITC comparison of the interactions of the four diamidines from Figure 1 with an AATT DNA sequence indicates that the better binding of DB818 to the DNA minor groove is enthalpic in origin (Table 3). The $T{\cdot}\Delta S$ term for DB818 binding to the minor groove of the AATT sequence is actually less favorable than that for the other three diamidines while the ΔH for binding is more favorable by 3−4 kcal/mol. It should be noted that the net binding entropy of all of the diamidines is favorable at 25 °C.[33] Such entropy contributions to binding are common for minor groove interactions and it has been shown that the binding of the benzimidazole derivative, Hoechst 33 258, is almost entirely driven by a favorable binding ΔS term.[34] It is interesting that although Hoechst 33 258 and DB818 have similar binding constants for AT sites in DNA, the molecular basis for their large binding constants is quite different, entropy dominated for Hoechst 33 258[34] and enthalpy dominated for DB818.

The results presented here suggest that strong, favorable interactions of the molecular groups of DB818 with the DNA minor groove provide a very favorable energetic component and account for the larger binding constant of the compound. Such strong interactions explain the larger negative enthalpy on complex formation and they are a welcome example of a very favorable design strategy to increase affinity and binding specificity for monomer minor groove binding agents. The strong interactions, however, also apparently make the DB818 complex less dynamic than with the other three compounds and this more rigid structure could account for the lower entropy of binding for DB818. Such enthalpy−entropy compensations are well-known in biochemical interactions and have been documented for several DNA−protein complexes.[35]

It is quite unlikely that any direct interactions with DNA of the sulfur atom in the thiophene of DB818 could account for

(31) Wilson, W. D. *Science* **2002**, *295*, 2103−2105.
(32) Wilson, W. D.; Wang, L.; Tanious, F.; Kumar, A.; Boykin, D. W.; Carrascol, C. *Biacore, J.* **2001**, *1*, 15−20.
(33) Bailly, C.; Chessari, G.; Carrasco, C.; Joubert, A.; Mann, J.; Wilson, W. D.; Neidle, S. *Nucleic Acids Res.* **2003**, *31*, 1514.
(34) Haq, I.; Ladbury, J. E.; Chowdhry, B. Z.; Jenkins, T. C.; Chaires, J. B. *J. Mol. Biol.* **1997**, *271*, 244.
(35) Dignam, J. D.; Nadda, S.; Chaires, J. B. *Biochemistry* **2003**, *42*, 5333−5340.

the increase in binding, for example, relative to the oxygen atom in the furan of DB293.[16b] DB75 and DB351, for example, have the same O to S change but DB351 binds slightly more weakly than DB75. The explanation for the more favorable interactions and ΔG and ΔH for DB818 binding thus seems likely to be due to structural differences in the complex of this compound with AT minor groove sequences relative to complexes of the other three diamidines of Figure 1. Fortunately, crystallographic results for DB818 bound to the self-complementary DNA duplex, $d(CGCGAATTCGCG)_2$, which has been used in structural analysis of several minor groove-binding compounds including DB75, provide a molecular rationale for the thermodynamic observations. Both DB818 and DB75 are largely bound in the AATT sequence of the duplex with both amidine groups facing the floor of the groove for interactions with DNA base pair edges (no satisfactory crystals of the DB351 and DB293 AATT complexes have yet been obtained). The curvature of DB818, however, appears optimized for interactions of the aromatic system with the walls and floor of the minor groove (Figure 10A). These results suggests that the optimized van der Waals interactions and H-bonding capability of the benzimidazole and amidines of DB818 account for its stronger binding to AATT relative to DB75 and DB351 (Figure 10B). The less favorable binding entropy of DB818 is probably due to its very tight interaction with the bases at the floor of the groove and to the twist in the bound compound that effectively locks it in place in the groove (Figure 10).

To better understand the binding differences between the benzimidazole derivatives of similar size, DB818 and DB293, a molecular model of DB293 was calculated for structural comparison (Methods Section). There is a small bond angle difference between the C−S−C angle of the thiophene of DB818 and the C−O−C angle of the furan of DB293 that is largely caused by the size difference between S and O. This angle change results in a small difference in the positions of the substituents on the thiophene and furan systems in these two compounds that can best be seen in an overlay diagram (Figure 11A). When amplified out to the terminal amidines of the two compounds, the small angle difference at the five-membered ring system yields very significant differences in the positions of the amidines on the two compounds. The DNA binding consequences of this molecular structural difference are apparent in an overlay diagram of the DB818 and DB293 complexes with the $d(CGCGAATTCGCG)_2$ sequence (Figure 11B). The DB818 structure is experimental while the DB293 structure is an optimized docked model. The smaller radius of curvature of DB293 is not well suited to match the DNA minor groove topology as can be seen in the overlay view. Although both amidines of DB293 can form H-bonding interactions with base edges at the floor of the minor groove, the curvature of the molecular system pushes the central region of the compound away from the floor of the groove and prevents the benzimidazole from forming strong H-bonding interactions of the type seen in the DB818 complex. The interactions of DB293 with the AATT sequence are thus similar to the interactions for the DB75 and DB351 complexes but are weaker than those for the closely related benzimidazole derivative, DB818.

In summary, subtle compound structural differences provide the molecular basis for the significant thermodynamic differences observed for formation of the AATT complexes by the four diamidines in Figure 1. Even with their significantly different substituents, the two diphenyl derivatives and DB293 have similar DNA interactions that yield similar ΔG, ΔH, and ΔS values for formation of their AATT complexes. Hydrophobic effects provide a major driving force for formation of the DNA complexes of these compounds, while interactions of the two amidines with AT base pairs provide additional specificity and affinity that probably arises primarily through an enthalpy contribution. With the DB818 complex, however, the compound and DNA minor groove topology are better matched and the DB818 interactions with the AT base pairs and groove walls at AATT are better optimized. Both the amidine and benzimidazole groups can form H-bonds with the base pairs in the DB818 complex and more base pairs are contacted than with either of the diphenyl compounds or DB293. As a result, the binding enthalpy is more favorable for complex formation with DB818. The optimized interactions result in a tighter and somewhat less dynamic complex with enthalpy−entropy compensation and a less favorable ΔS for complex formation with DB818. The net result is that an improved ΔG for binding of DB818 is obtained because the ΔH improvement is larger than the ΔS decrease. Addition of other substituents, which could interact with DNA, to DB818 could yield favorable contacts and molecular interactions at a lower entropy cost, and could result in a more dramatic improvement in ΔG for binding of derivatives of DB818. On a larger scale, by choosing an appropriate central ring or an adjacent benzo-ring (benzothiophene, benzimidazole, benzofuran) it is possible to "tune" the shape of the molecule to match DNA structurally and chemically for enhanced affinity and sequence recognition.

Acknowledgment. Supported by NIH Grant No. GM61587, by the Bill and Melinda Gates Foundation, and by equipment funds from the Georgia Research Alliance (W.D.W. and D.W.B.), the Ligue Nationale Contre le Cancer (C.B.), and the Cancer Research UK (S.N.).

Note Added after ASAP Posting. After this paper was posted ASAP on September 29, 2004, three authors' affiliations and typographical errors in some of the units in the "Crystallization and Data Collection" section were corrected. The corrected version was posted October 4, 2004.

Supporting Information Available: Supplementary Figures S1−S4 and legends. This material is available free of charge via the Internet at http://pubs.acs.org.

JA048175M