J|A|C|S
ARTICLES
Published on Web 12/16/2003

DNA Sequence Dependent Monomer−Dimer Binding Modulation of Asymmetric Benzimidazole Derivatives

Farial A. Tanious,[†] Donald Hamelberg,[†] Christian Bailly,*[,‡] Agnieska Czarny,[†] David W. Boykin,*[,†] and W. David Wilson*[,†]

Contribution from the Department of Chemistry, Georgia State University, Atlanta, GA, 30303, USA, and INSERM U-524 et Laboratoire de Pharmacologie Antitumorale du Centre Oscar Lambret, IRCL, Place de Verdun, 59045 Lille, France

Received July 7, 2003; E-mail: chewdw@panther.gsu.edu

Abstract: A number of studies indicate that DNA sequences such as AATT and TTAA have significantly different physical and interaction properties. To probe these interaction differences in detail and determine the influence of charge, we have synthesized three bisbenzimidazole derivatives, a diamidine, DB185, and monoamidines, DB183 and DB210, that are related to the well-known minor groove agent, Hoechst 33258. Footprinting studies with several natural and designed DNA fragments indicate that the synthetic compounds bind at AT sequences in the minor groove and interact more weakly at sites with TpA steps relative to sites without such steps. Circular dichroism spectroscopy also indicates that the compounds bind in the DNA minor groove. Surprisingly, Tm studies as a function of ratio indicate that the monoamidines bind to TTAA sequences as dimers, whereas the diamidine binds as a monomer. Biosensor-surface plasmon resonance (SPR) studies allowed us to quantitate the interaction differences in more detail. SPR results clearly show that the monoamidine compounds bind to the TTAA sequence in a cooperative 2:1 complex but bind as monomers to AATT. The dication binds to both sequences in monomer complexes but the binding to AATT is significantly stronger than binding to TTAA. Molecular dynamics simulations indicate that the AATT sequence has a narrow time-average minor groove width that is a very good receptor site for the bisbenzimidazole compounds. The groove in TTAA sequences is wider and the width must be reduced to form a favorable monomer complex. The monocations thus form cooperative dimers that stack in an antiparallel orientation and closely fit the structure of the TTAA minor groove. The amidine groups in the dimer are oriented in the 5′ direction of the strand to which they are closest. Charge repulsion in the dication apparently keeps it from forming the dimer. It instead reduces the TTAA groove width, in an induced fit process, sufficiently to form a minor groove complex. The dimer-binding mode of DB183 and DB210 is a new DNA recognition motif and offers novel design concepts for selective targeting of DNA sequences with a wider minor groove, including those with TpA steps.

Introduction

It is now well established that benzimidazole-amidine systems, such as in compounds DB183, DB185, and DB293 (Figure 1), provide a very favorable and flexible DNA recognition module.[1−3] Benzimidazoles, for example, form the core of Hoechst 33258 (H258), one of the most extensively studied minor groove binding agents, as well as of a variety of other related DNA binding compounds.[4] Converting the H258 bulky cationic group into a planar amidine or other similar group significantly enhances its DNA interaction affinity.[3a,b;4f] We have recently shown that DB293, a benzimidazole-amidine of different structure, has unusual DNA recognition properties. Whereas the diphenyl analogue of DB293, DB75 (Figure 1), binds only to AT-rich sequences as a monomer,[5] DB293 prefers formation of minor groove dimers at sites that include GC base pairs.[3c−f] Recent developments in genomics and molecular

[†] Department of Chemistry, Georgia State University.
[‡] INSERM U-524 et Laboratoire de Pharmacologie Antitumorale du Centre Oscar Lambret.

(1) (a) Tidwell, R. R.; Boykin, W. D. 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. (b) 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.; Neidle, S. *J. Med. Chem.* **1995**, *38*, 912−916. (c) Edwards, K. J.; Jenkins, T. C.; Neidle, S. *Biochemistry* **1992**, *31*, 7104−7109. (d) Nunn, C. M.; Jenkins, T. C.; Neidle, S. *Biochemistry* **1993**, *32*, 13 838−13 845.

(2) (a) Lansiaux, A.; Dassonnville, L.; Facompre, M.; Kumar, A.; Stephens, C. E.; Bajic, M.; Tanious, A.; Wilson, W. D.; Boykin, D. W.; Bailly, C. *J. Med. Chem.* **2002**, *45*, 1994−2002.(b) Lansiaux, A.; 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.

(3) (a) Czarny, A.; Boykin, D. W.; Wood, A. A.; Nunn, C. M.; Neidle, S.; Zhao, M.; Wilson, W. D. *J. Am. Chem. Soc.* **1995**, *117*, 4716−4717. (b) Clark, G. R.; Boykin, D. W.; Czarny, A.; Neidle, S. *Nucleic Acids Res.* **1997**, *25*, 1510−1515. (c)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. (d) Wang, L.; Carrasco, C.; Kumar, A.; Stephens, C. E.; Bailly, C.; Boykin, D. W.; Wilson, W. D. *Biochemistry* **2001**, *40*, 2511−2521. (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) Tanious, F.; Wilson, W. D.; Wang, Lei.; Kumar, A.; Boykin, D. W.; Marty, C.; Baldeyrou, B.; Bailly, C. *Biochemistry* (**2003**).

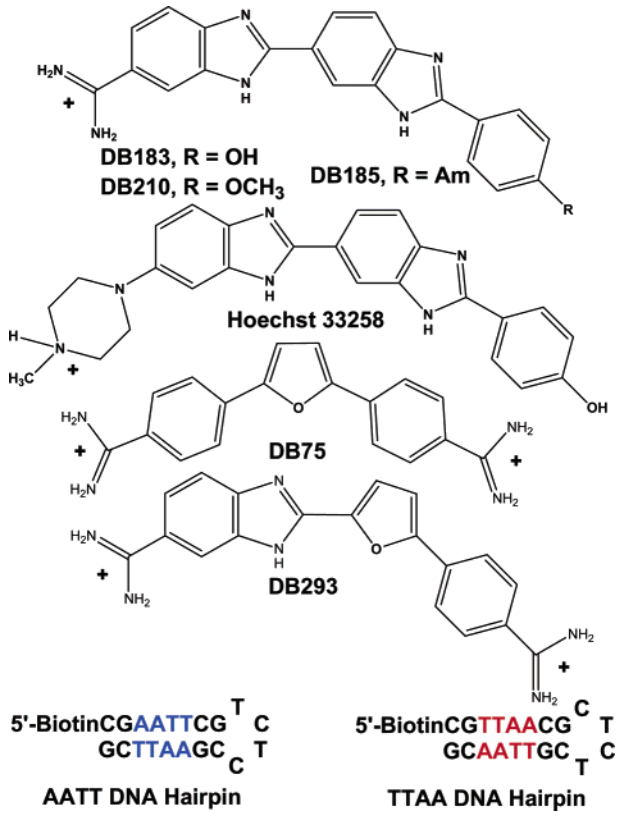

Figure 1. Structure of the compounds and the DNA sequences used in this study.

biology have provided a wealth of new information on important genes from cancer cells to parasitic microorganisms. Despite this new information on gene sequences, DNA-targeted-chemotherapy of both cancer and parasitic microorganisms is still based largely on the application of drugs that have been in use for some time (e.g., nitrogen mustards, anthracyclines, pentamidine). Given that an orally active prodrug of DB75 is currently in Phase II clinical trials and is scheduled to enter Phase III trials within the next year against parasitic diseases,[1a] it is important to study additional amidine-aromatic systems related to DB75 to extend and develop this successful therapeutic discovery. Such development of DNA targeted therapeutics is also essential if we are to begin to develop direct therapeutic benefits from the genome sequence information now available.

Recognition of the DNA minor groove with small molecule dimers is an obvious way to increase drug specificity as well as affinity since dimers can recognize both strands of the double helix.[3,6] To this time, however, very few DNA-binding small molecules have been found to form stacked minor groove dimers. Apart from distamycin[6,7] and related polyamide/lexitropsin compounds[6,8] only two nonamide extended molecular classes have been found to bind to the minor groove of DNA via the formation of stacked dimers. One class is based on the furan diamidine compound, DB293 (Figure 1), described above.[3c−f] The antiparallel stacking of two DB293 molecules is highly cooperative, strongly dependent on the compound structure and occurs preferentially at 5′-ATGA sites.[3] The DB293 analogues with either two phenyls (DB75, Figure 1) or two benzimidazoles (DB270) do not show any tendency to form the stacked dimer.[3c,d]

Besides DB293, another heterocyclic class thought to form minor groove side-by-side dimer structures is related to H258 (Figure 1).[4,9] On the basis of fluorescence measurements and modeling studies, Bruice and co-workers have reported that an −O-phenyl derivative of H258, Hoechst 33377, binds to an AATT sequence in an antiparallel dimer arrangement.[4i−k] The H258 parent compound, however, was reported to bind to the same sequence as a monomer. Thus, some benzimidazole-containing DNA minor groove binders such as DB293 and Hoechst analogues are excellent candidates for the design of dimer-forming gene regulatory molecules capable of reading the genetic information from the minor groove of DNA. To design compounds to achieve this goal, it is essential to define the rules and compound structural elements necessary for formation of 1:1 vs 2:1 complexes. It is clear from the results described above that dimer formation is highly sensitive to compound structure. It should be noted that with the exception of furamidine and its prodrug, DB289, and related structures, all of the agents listed above have limited therapeutic potential due to limited oral bioavailability and/or excessive toxicity[1a]. It is thus desirable to extend the furamidine type structure in our search for additional therapeutic agents that target the DNA minor groove.

(4) (a) Abu-Daya, A.; Brown, P. M.; Fox, K. R. *Nucleic Acids Res.* **1995**, *23*, 3385−3392. (b) Abu-Daya, A.; Fox, K. R. *Nucleic Acids Res.* **1997**, *25*, 4962−4969. (c) Bailly, C. In *DNA and RNA Binders: from Small Molecules to Drugs*; Demeunynck, M., Bailly, C., Wilson, W. D., Eds.; Wiley-VCH: 2003; Vol. 2, Chapter 20, pp 538−575. (d) Bailly, C.; Chessari, G.; Carrasco, C.; Joubert, A.; Mann, J.; Wilson, W. D.; Neidle, S. *Nucleic Acids Res.* **2003**, *31*, 1514−1524. (e) Breusegem, S. Y.; Sadat-Ebrahimi, S. E.; Douglas, K. T.; Bichenkova, E. V.; Clegg, R. M.; Loontiens, F. G. *J. Med. Chem.* **2001**, *44*, 2503−2506. (f) Bostock-Smith, C. E.; Harris, S. A.; Laughton, C. A.; Searle, M. S. *Nucleic Acids Res.* **2001**, *29*, 693−702. (g) Carrondo, M. A. A. de C. T.; Coll, M.; Aymami, J.; Wang, A. H. J.; Van der Marel, G. A.; Van Boom, J. H.; Rich, A. *Biochemistry* **1989**, *28*, 7849−7859. (h) Haq, I.; Ladbury, J. E.; Chowdhry, B. Z.; Jenkins, T. C.; Chaires, J. B. *J. Mol. Biol.* **1997**, *271*, 244−257. (i) Satz, A. L.; Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 86−95. (j) Satz, A. L.; White, C. M.; Beerman, T. A.; Bruice, T. C. *Biochemistry* **2001**, *40*, 6465−6474. (k) Satz, A. L.; Bruice, T. C. *J. Am. Chem. Soc.* **2001**, *123*, 2469−2477. (l) Teng, M.-K.; Usman, N.; Frederick, C. A.; Wang, A. H.-J. *Nucleic Acids Res.* **1988**, *16*, 2671−2690. (m) Loontiens, F. G.; Regenfuss, P.; Zechel, A.; Dumortier, L.; Clegg, R. M. *Biochemistry* **1990**, *29*, 9029−9039. (n) Wood, A. A.; Nunn, C. M.; Czarny, A.; Boykin, D. W.; Neidle, S. *Nucleic Acids Res.* **1995**, *23*, 3678−3684. (o) Rosu, F.; Gabelica, V.; Houssier, C.; Pauw, E. D. *Nucleic Acids Res.* **2002**, *30*, e82. (p) Ebrahimi, S. E.; Bibby, M. C.; Fox, K. R.; Douglas, K. T. *Anticancer Drug Des.* **1995**, *10*, 463−479.

(5) (a) Mazur, S.; Tanious, F. A.; Ding, D.; Kumar, A.; Boykin, D. W.; Simpson, I. J.; Neidle, S.; Wilson, W. D. *J. Mol. Biol.* **2000**, *300*, 321−337. (b) 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.

(6) (a) Wemmer, D. E. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 439−461. (b) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215−2235.

(7) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723−5727.

(8) (a) Moravek, Z.; Neidle, S.; Schneider, B. *Nucleic Acids Res.* **2002**, *30*, 1182−1191.(b) Bailly, C. In *Advances in DNA Sequence Specific Agents*; Palumbo, M., Ed; JAI Press Inc.: London, 1988; Vol. 3 pp 97−156. (c) Chaires, J. B. *Curr. Opin. Struct. Biol.* **1998**, *8*, 314−320. (d) Sondhi, S. M.; Reddy, B.; Lown, J. W. *Curr. Med. Chem.* **1997**, *4*, 313−358. (e) Kopka, M. L.; Larsen, T. A. In *Nucleic Acid Targeted Drug Design*; Propst, C. L., Perun, T. J., Eds.; Marcel Dekker Inc.: New York, 1992; pp 303−374.

(9) (a) Drobyshev, A. L.; Zasedatelev, A. S.; Yershov, G. M.; Mirzabekov, A. D. *Nucleic Acids Res.* **1999**, *27*, 4100−4105. (b) Fede, A.; Billeter, M.; Leupin, W.; Wüthrich, K. *Structure* **1993**, *1*, 177−186. (c) Parkinson, J. A.; Barber, J.; Douglas, K. T.; Rosamond, J.; Sharples, D. *Biochemistry* **1990**, *29*, 10 181−10 190. (d) Pjura, P. E.; Grzeskoniak, K.; Dickerson, R. E. *J. Mol. Biol.* **1987**, *197*, 257−271. (e) Quintana, J. R.; Lipanov, A. A.; Dickerson, R. E. *Biochemistry* **1991**, *30*, 10 294−10 306. (f) Spink, N.; Brown, D. G.; Skelly, J. V.; Neidle, S. *Nucleic Acids Res.* **1994**, *22*, 1607−1612. (g) Sriram, M.; van der Marel, G. A.; Roelen, H. L. P. F.; van Boom, J. H.; Wang, A. H.-J. *EMBO J.* **1992**, *11*, 225−232. (h) Vega, M. C.; Garcia Saez, I.; Aymami, J.; Eritja, R.; Van der Marel, G. A.; van Boom, J. H.; Rich, A.; Coll, M. *Eur. J. Biochem.* **1994**, *222*, 721−726. (i) Breusegem, S. Y.; Clegg, R. M.; Loontiens, F. G. *J. Mol. Biol.* **2002**, *315*, 1049−1061.

A large number of DNA minor groove binders containing one or more benzimidazole heterocycles have been reported to date, including head-to-head bis-benzimidazoles and tris-benzimidazoles endowed with promising antitumor and anti parasitic activities.[1,10−12] In the search for minor groove binding agents, other than polyamides, that can form stacked dimers for enhanced recognition specificity and therapeutic potential, compounds with benzimidazole units are perhaps the most promising direction. As a first step in the search for alternative compounds that can form dimers at selected DNA sequences, we have evaluated the binding of the monocations DB183, DB210, and the dication, DB185 (Figure 1), to different AT containing DNA sequences. DB185 not only contains the same bisbenzimidazole-phenyl core as the Hoechst compounds, for example, but it also retains the diamidine functions of DB293. Given the very different properties of sequences such as AnTn and TnAn,[13] we asked the question of whether these different properties could lead to differences in monocation and dication complexes. Initial studies with these compounds by Tm and X-ray structural methods showed that they could bind to AATT sequences very strongly in a 1:1 complex.[3a,b] Here, we report a complementary biochemical and biophysical analysis of the DNA binding properties of the benzimidazole derivatives with DNaseI footprinting, thermal melting, circular dichroism spectroscopy, biosensor-surface plasmon resonance (SPR) measurements, and theoretical calculations to evaluate the binding strength, stoichiometry, cooperativity and sequence selectivity of these compounds to TTAA. The specific questions that we wish to address are as follows: (i) How does binding of the monocation compare to that of the dication in terms of sequence selectivity and affinity? (ii) Do either the monocation or dication show significant cooperative dimer formation? The results from all techniques show strong binding of the compounds to AT sequences. Surprisingly, however, the monocations, but not the dication, were found to bind to the TTAA, but not the AATT, sequence as dimers. This is a clear example of dimerization driven by both DNA and compound molecular recognition and structural complementary in complex formation.

Materials and Methods

Purification of DNA Restriction Fragments and Radiolabeling. The different 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 198-bp fragment was obtained from plasmid pMS1 (kindly provided by Dr K. R. Fox, University of Southampton) after digestion with the restriction enzymes *Hin*dIII and *Xba*I.[14a] The 131-bp and 116-bp fragments were obtained from plasmids pTayB and pTayD (provided by Dr NE Møllegaard, The Panum Institute, Copenhagen, Denmark) after digestion with *Eco*RI and *Hin*dIII. 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 by Storage Phosphor Imaging. 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 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 previously described.[14b] 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 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. DNA cleavage products were resolved by polyacrylamide gel electrophoresis under denaturating 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, BRL sequencer model S2), gels were soaked in 10% acetic acid for 10 min, transferred to Whatman 3MM 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 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 dimethylsulfate followed by piperidine-induced cleavage at the modified guanine bases in DNA (G-track).

Compounds and Buffers. The Benzimidazole derivatives DB183, DB185, DB210, and DB293 and the furan compound

(10) (a) Baily, C.; Chaires, J. B. *Bioconjugate Chem.* **1998**, *9*, 513−538. (b) Kim, J. S.; Gatto, B.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. *J. Med. Chem.* **1996**, *39*, 992−998. (c) Lombardi, R. L.; Tanious, F. A.; Ramachandran, K.; Tidwell, R. R.; Wilson, W. D. *J. Med. Chem.* **1996**, *39*, 1452−1462. (d) Minehan, T. G.; Gottwald, K.; Dervan, P. B. *Helv. Chim. Acta* **2000**, *83*, 2197−2213. (e) Wang, H.; Gupta, R.; Lown, J. W. *Anti-Cancer Drug Design* **1994**, *9*, 153−180.

(11) (a) Joubert, A.; Sun, X.-W.; Johansson, E.; Bailly, C.; Mann, J.; Neidle, S. *Biochemistry* **2003**, *42*, 5984−5992. (b) Mann, J.; Baron, A.; Opoku-Boahen, Y.; Johansson, E.; Parkinson, G.; Kelland, L. R.; Neidle, S. *J. Med. Chem.* **2001**, *44*, 138−144.

(12) (a) Aymami, J.; Nunn, C. M.; Neidle, S. *Nucleic Acids Res.* **1999**, *27*, 2691−2698. (b) Goldman, G. H.; Yu, C.; Wu, H.-Y.; Sanders, M. M.; La Voie, E. J.; Liu, L. F. *Biochemistry* **1997**, *36*, 6488−6494. (c) Ji, Y. H.; Bur, D.; Hasler, W.; Bailly, C.; Waring, M. J.; Hochstrasser, R.; Leupin, W. *BioOrg. Med. Chem.* **2001**, *9*, 2905−2919. (d) Kim, J. S.; Sun, Q.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. *Bioorg. Med. Chem.* **1998**, *6*, 163−172. (e) Pilch, D. S.; Xu, Z.; Sun, Q.; LaVoie, E.; Liu, L. F.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13 565−13 570. (f) Pilch, D. S.; Xu, Z.; Sun, Q.; LaVoie, E. J.; Liu, L. F.; Geacintov, N. E.; Breslauer, K. J. *Drug. Des. Discovery* **1996**, *13*, 115−133. (g) Xu, Z.; Li, T.-K.; Kim, J. S.; LaVoie, E.; Breslauer, K. J.; Liu, L. F.; Pilch, D. S. *Biochemistry* **1998**, *37*, 3558−3566.

(13) Price, M. A.; Tullius, T. D. *Biochemistry* **1993**, *32*, 127−136.

(14) (a) Lavesa, M.; Fox, K. R. *Anal. Biochem.* **2001**, *293*, 246−250. (b) Bailly, C.; Waring, M. J. *J. Biomol. Struct. Dyn.* **1995**, *12*, 869−898.

DB75 were synthesized as previously described.[1a,3d,15] The Beer's low extinction coefficients for DB183, DB210, and DB185 in MES buffer are as follows: $E_{334} = 36\,120\ M^{-1}cm^{-1}$ for DB183, $E_{336} = 36\,560\ M^{-1}cm^{-1}$ for DB210, and $E_{342} = 35\,160\ M^{-1}cm^{-1}$ for DB185. HBS-EP buffer (BIA Certified) from BIACORE contains 0.01M HEPES (*N*-[2-Hydroxyethyl]-piperazine-*N*′-[2-ethanesulfonic acid]), 0.15M NaCl, 3mM EDTA, and 0.005% polysorbate 20 (v/v), pH 7.4. MES00 buffer was prepared with 0.01 M MES (2-(N-morpholino)ethanesulfonic acid), 0.001 M ethylenediaminetetraacetic acid (EDTA), and the pH was adjusted to 6.25 with NaOH solution, MES10 is the same as MES00 but with 0.1 M NaCl, and MES20 has 0.2M NaCl: MES00, no added NaCl; MES10, 0.1M NaCl; MES20; 0.2MNaCl.

Oligomers. The 5′-biotin labeled hairpin duplexes used in these studies (Midland Certified Reagent Co.-HPLC purified and desalted) are (Biotin-dCGAATTCG*TCTC*CGAATTCG) (AATT hairpin), (Biotin-dCGTTAACG*CTCT*CGTTAACG) (TTAA hairpin) with the hairpin loop sequences in italics. The oligomer concentrations were determined optically using extinction coefficients per mole of strand at 260 nm determined by the nearest neighbor procedure:[16] AATT hairpin (units of M(strand) $^{-1}cm^{-1}$):, $E_{260} = 185\,700$; and TTAA hairpin (units of M(strand) $^{-1}cm^{-1}$) $E_{260} = 184\,700$.

Determination of Binding Constants by Surface Plasmon Resonance. Surface plasmon resonance (SPR) measurements were performed with BIAcore 2000 or 3000 systems and streptavidin coated sensor chips (SA from BIACORE). Briefly, the chips were prepared for use by conditioning with three to five consecutive 1 min injections of 1 M NaCl in 50 mM NaOH followed by extensive washing with buffer. 5′-Biotinylated DNA samples (25nM) in HBS buffer were immobilized on the flow cell surface by noncovalent capture. Three flow cells were used to immobilize DNA oligomer samples and the fourth cell was left blank as a control. Samples of the compounds were prepared in filtered and degassed MES 10 buffer by serial dilution from stock solutions. Drug samples were injected from 7 mm plastic vials with pierceable plastic crimp caps at a flow rate of 7 or 10 μL/min using the KINJECT command. For the benzimidazole-DNA complexes a solution of 10 mM glycine pH 2 was use to dissociate all of the benzimidazole from DNA to regenerate the surface. An array of different benzimidazole concentrations was used in each experiment and the results were analyzed as described below. The injection of the compound (association) was followed by injection of running buffer (compound dissociation). To reduce the probability of nonspecific binding to the chip surface 50 μL/L of surfactant P20 was added to the MES buffers. The amount of DNA immobilized was approximately 350 RUs in each of the three flow cells. SPR experiments were performed at 25 °C in MES10. With the SPR technique the change in refractive index occurring at the surface of the sensor chip is monitored. The change in refractive index in terms of response units is proportional to the amount of compound bound to the DNA.

To obtain the affinity constants the data generated were fitted to different interaction models using Kaleidagraph for nonlinear least squares optimization of the binding parameters using the following equation

$$RU = RUmax \times (K_1 \times C_{free} + 2 \times K_1 \times K_2 \times C_{free}^2)/(1 + K_1 \times C_{free} + K_1 \times K_2 \times C_{free}^2)$$

where K_1 and K_2 are macroscopic equilibrium constants for two types of binding sites, RU is the SPR response at the steady-state level, RU_{max} is the maximum SPR response for binding one molecule per binding site, and C_{free} is the concentration of the compound in solution. For a single-site model ($K_2 = 0$). RUmax can be predicted using the following equation

$$RUmax = (RU_{DNA}/MW_{DNA}) \times MW_{compound} \times RII$$

where RU_{DNA} is the amount of DNA immobilized in response units (RU), MW is molecular weight of compound and DNA, and RII is the refractive index increment ratio of compound to refractive index of DNA.[17]

Thermal Melting (T_m) Experiments. Experiments were conducted in MES00. The concentration of the DNA was about 3×10^{-6} to 5×10^{-6} M in hairpin and the experiments were done using a Cary 3E or 4E spectrophotometer.

Circular Dichroism. CD spectra were obtained on Jasco J-710 spectrometer. The software supplied by Jasco provided instrument control, data acquisition and manipulation. DNA solutions in MES buffer were scanned in 1 cm quartz cuvettes and solutions of the compounds were added to DNA at the desired ratio and the complexes were rescanned.

Molecular Dynamics Simulations and Free Energy Calculations. Each molecular dynamics simulation was carried out with the AMBER 5.1 suites of programs.[18] The solutes were solvated with approximately 4000 TIP3P[19] water molecules to fill a box size of approximately $50 \times 60 \times 50$ Å. The systems were neutralized with Na^+ ions to obtain electronic neutrality, and an additional 10 Na^+ and 10 Cl^- ions were added so that the NaCl concentration was approximately 0.15 M. The all atom Cornell et al.[20] force field was used to model the DNA molecules. The atomic charges and force field parameters of the drug molecules were determined using the RESP[21] methodology at the HF/6-31G* level of theory and comparable standard parameters of the Cornell et al. force field, respectively. The systems were equilibrated by using a standard multistage equilibration protocol.[22] After equilibration, the final equilibrated structures were used to carryout seven 3 ns MD simulations in the NPT ensemble with periodic boundary conditions at a constant temperature of 300 K with the Berendsen temperature algorithm,[23] and at a pressure of 1 bar. The SHAKE[24] algorithm

(15) Czarny, A.; Wilson, W. D.; Boykin, D. W. *J. Heterocyclic Chem*. **1996**, *33*, 1393−1397.
(16) Fasman, G. D. In *Handbook of Biochemistry and Molecular Biology*: *Nucleic Acids*; CRC Press: Cleveland, OH, 1975; Vol. 1 p 589.
(17) Davis, T. M.; Wilson, W. D. *Anal. Biochem*. **2000**, *284*, 348−353.
(18) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III, Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K. and Kollman, P. A. 1997. 5.1 Edition; University of California, San Francisco.
(19) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(20) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Jr., K. M. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(21) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10 269−10 280.
(22) (a) Hamelberg, D.; McFail-Isom, L.; Williams, L. D.; Wilson, W. D. *J. Am. Chem. Soc*. **2000**, *122*, 10 513−10 520. (b) Hamelberg, D.; Williams, L. D.; Wilson, W. D. *J. Am. Chem. Soc.* **2001**, *123*, 7745−7755.
(23) Berendsen, H. J. C.; van Gunsteren, W. F.; Postma, J. P. M.; A. D. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

was applied to all bonds involving hydrogen atoms and an integration time step of 2.0 fs was used in solving Newton's equation of motion. Lennard−Jones interactions were subjected to a 9 Å cutoff and the nonbonded pairlists were updated at every 10 steps. The MD trajectories were sampled at one picosecond intervals.

The energy analysis was done on the last 1 ns (1000 snapshots) of each simulation. To calculated the free energy changes of binding, ΔG, the free energies of the free DNA, free drug, and DNA-drug complexes were calculated according to eq 1

$$G = E_{\text{conf}} + G_{\text{solvation}} - TS_{\text{conf}} \quad (1)$$

where E_{conf} is the internal energy of the solute, $G_{\text{solvation}}$ is the free energy of solvation of the solute, T is the temperature, and S_{conf} is configurational entropy of the solute. The internal energies of the solutes, E_{conf}, were calculated for each snapshot with the force field used in the simulations with no solvent or counterions and no cutoff for nonbonded interaction. The free energies of solvation, $G_{\text{solvation}}$, are the sum of an electrostatic term and a nonpolar term. The electrostatic contributions were calculated using the finite difference solution of the Poisson−Boltzmann equation as implemented in Delphi II program.[25] The solutes were assigned a dielectric constant of 1, to be consistent with the simulation force field, and solvent dielectric constant was set to 80. A grid spacing of 0.5 was used in the calculation. The nonpolar contributions were modeled as a term related to the solvent-accessible surface area by the expression aSASA + b,[26] where SASA is the surface area of the solute, a is the 0.00542 kcal/Å^2, and b is 0.92 kcal/mol. The SASA was calculated using the algorithm of Scanner et al.[27] The free energies of solvation for each snapshot were calculated and averaged over the all the snapshots. Estimation of the configurational entropy, S_{conf}, was done by carrying out a normal-mode analysis, as implemented in AMBER, on the each solute structure. To perform normal mode calculations, each snapshot had to be minimized. The structures were minimized until the rms (root-mean-square) was approximately 10^{-6} kcal/mol Å.

Results

DNaseI Footprinting Studies. Because DB185 is a diamidine, we first compared its sequence recognition properties with two other diamidines, DB293 and DB75 (Figure 1), using conventional DNaseI footprinting methodology appropriate for small molecule-DNA complexes.[14] We chose DB75 and DB293 because they give quite different footprints with DNA: DB75 is an AT specific compound while DB293 gives generally stronger footprints in sites that contain GC base pairs. The electrophoresis gel presented in Figure S1 indicates that DB185 binds extremely tightly to specific DNA sequences to produce pronounced footprints at several sites within the 265-bp restriction fragment used for this experiment. The footprinting pattern obtained with DB185 is closer to that of DB75 than that of DB293, and both DB75 and DB185 produce footprints in AT sequence regions. To get a more precise analysis of the footprints, the intensities of the bands in Figure S1 were determined and converted to numerical probability of cleavage as previously described.[14] Several important conclusions arise from this analysis (Figure S2): (i) the footprints seen with DB185 always coincide with the position of AT-rich tracts; (ii) the footprints have the 3′-offset expected for footprinting of a minor groove complex, in accord with the model for asymmetric cleavage by DNase I across the minor groove of the B-form helix;[28] (iii) DB185 protects AT-rich sequences from cleavage by DNase I much more intensely than DB75 at low concentration indicating that it could exhibit a considerably higher affinity for AT DNA sequences; and (iv) the ATGA-containing sequence near position 100, specific to a DB293 dimer complex, is not a good receptor site for DB185.

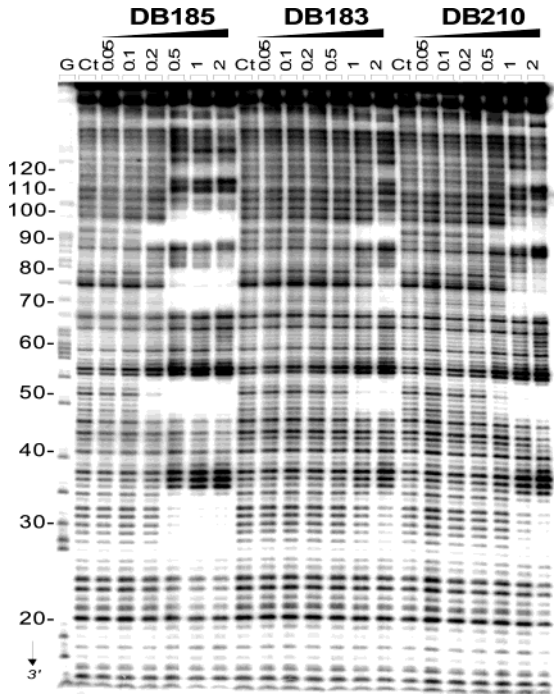

Figure 2. DNase I footprinting of the diamidine DB185 and the monoamidine analogues DB183 and DB210 on the 3′-end radiolabeled 198-bp restriction fragment. The cleavage products of the DNase I digestion were resolved on an 8% polyacrylamide gel containing 8M urea. The concentrations (microM) of the drugs are shown at the top of the appropriate gel lanes. Control tracks labeled "Ct" contained no drug. The track labeled G represents dimethylsulfate−piperidine markers specific for guanine. Numbers at the left side of the gel refer to the numbering scheme used in Figure 3.

The footprinting study was extended to two additional DNA sequences with fragments of 117-bp and 198-bp that offer various combinations of AT arrangements. To provide information on the influence of compound structure and the importance of total charge, the experiments were performed in parallel with the diamidine as well as two monoamidines, DB183 and DB210, that have a hydroxy and a methoxy group, respectively, on the phenyl ring in place of the amidine in DB185 (Figure 1). The representative gel shown in Figure 2 indicates that the two monocations bind to the same AT DNA sequences as the dication but with a lower affinity. Significantly higher drug concentrations are required to detect the same level of footprint

(24) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(25) (a) Rocchia, W.; Alexov, E.; Honig, B. *J. Phys. Chem. B.* **2001**, *105*(28), 6507−6514. (b) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435−445.
(26) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*(7), 1978−1988.
(27) Sanner, M.; Olson, A. J.; Spehner, J. C. *Biopolymers* **1996**, *38*(3), 305−320.
(28) Weston, S. A.; Lahm, A.; Suck, D. *J. Mol. Biol.* **1992**, *226*, 1237−1256.

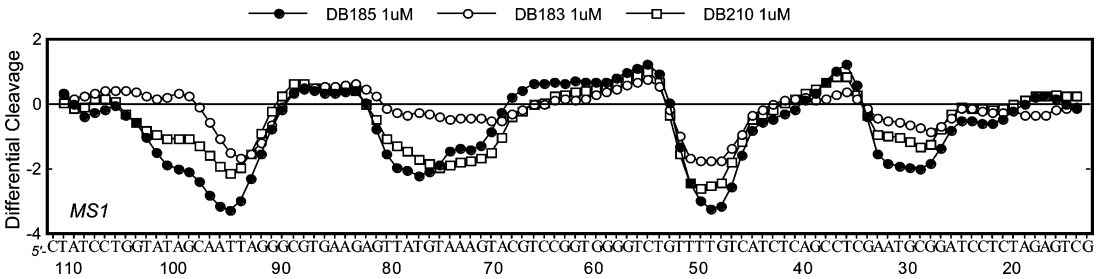

Figure 3. Differential cleavage plots comparing the DNase I cleavage of the 198-mer fragment in the presence of the mono (DB183, DB210) and diamidine (DB185) compounds (1 μM each). In this representation, negative values correspond to a ligand-protected site and positive values represent enhanced cleavage. Vertical scales are in units of ln(fa) − ln(fc), where fa is the fractional cleavage at any bond in the presence of the drug, and fc is the fractional cleavage of the same bond in the control, given closely similar extents of overall digestion. Each line drawn represents a 3-bond running average of individual data points, calculated by averaging the value of ln(fa) − ln(fc) at any bond with those of its two nearest neighbors. Only the region of the restriction fragment analyzed by densitometry is shown.

with the monoamidines. At a concentration of 0.2 μM, the phenyl-amidine compound DB185 strongly inhibits DNaseI cleavage at specific sites, whereas at this concentration, the phenol and methoxy derivatives show no effect. We can estimate from the footprinting experiments that the loss of one amidine function reduces the binding strength by a factor of approximately 10. It should be emphasized, however, that relative to other similar minor groove binding agents such as H258 and DB75 the monoamidines are quite strong minor groove binding agents, whereas the diamidine is an exceptionally strong binding compound. The differential cleavage plots in Figure 3 confirm that the compounds bind preferentially to AT sequences. Four binding sites can be identified in the 198-bp fragment. The large footprints centered on nucleotide positions 102 and 83 correspond to compound binding to two juxtaposed AT tracts: 5′-TATAGCAATTA and 5′-TTATGTAAA. The third site around position 51 refers to the sequence 5′-TTTT and interestingly, the fourth site around position 30 contains only three consecutive A·T bp, 5′-AATGC. Similar results were obtained with the 117-bp fragment. In this case, the following binding sequences were identified: 5′-ATTAA, 5′-TTTT, 5′-TAAAA, 5′-AATTG-TAATA, and 5′-GTAAC (data not shown).

The last series of footprinting experiments was performed with two designed DNA fragments of 116 and 131 bp containing the 16 possible $[A{\cdot}T]_4$ blocks.[29,4a] The interaction enhancement of the two amidine functions appears most clearly when comparing the footprinting patterns obtained with DB185 and DB183 (Figure S3). It takes much lower concentrations of the dication to produce intense footprints and there is complete disappearance of the DNA band in many places. The differential cleavage plots (Figure S4) also show that DB185 does not bind equally to all types of $[A{\cdot}T]_4$ blocks. For example, with the 116- and 131-bp fragments much stronger footprints were detected at 5′-AATT and 5′-ATTT·5′-AAAT, whereas higher concentrations of DB185 were required to observe significant binding at TpA containing sequences such as 5′-TTAA and 5′-TAAA·5′-TTTA. A concentration-dependent analysis was performed to better compare the magnitude of binding of DB185 at distinct AT-binding sites. The drug concentration required for half-maximal footprinting (the C_{50} value) is about 0.5 μM for the 5′-AATT site under these conditions, whereas the C_{50} value is practically doubled when an A is substituted for a T (5′-AATA and 5′-AAAT). Some other AT sites, such as the sequence 5′-ATAA, bind DB185 even more weakly. It is important to note that the presence of a TpA step within an $[A{\cdot}T]_4$ block reduces the extent of binding. This is commonly observed with minor groove binders,[4a] including diamidines.[30]

Thermal Melting. To obtain more information on the different complexes of the compounds with ApT and TpA containing steps, thermal melting experiments were conducted with oligomers containing AATT and TTAA binding sites. The DNAs have monophasic melting curves in the unbound state but exhibit biphasic curves below the saturation ratio for the benzimidazole compounds. The biphasic melting is characteristic of compounds that bind very strongly to DNA. As the free oligomer melts at low temperature, the compound eventually saturates the remaining sites to create a high-melting complex. As the amount of compound is increased, the fraction in the low-melting band decreases and the fraction at higher temperature increases up to complete saturation of the sites. Results for DB210 are shown as examples in Figure 4 and ΔT_m values for all complexes are in the figure legend. The saturation point varies with both DNA sequence and compound structure. The two transitions in the biphasic curves occur at approximately the same temperature as a function of ratio but the intensity shifts to the higher temperature transition as the ratio is increased. With the AATT sequence, all three benzimidazoles (Figure 1) have biphasic melting curves up to a ratio of 1:1 (compound/ hairpin duplex) and the curves are monophasic above that ratio. The dication DB185 stabilizes the DNA significantly more than the two monocations in agreement with previous results.[3a]

With the TTAA sequence some of the results are quite different. DB185 still exhibits biphasic melting below the 1:1 ratio and with monophasic melting behavior at the 1:1 ratio and above (Figure S5). DB183 and DB210, however, have biphasic melting curves up to a ratio of 2:1 but have monophasic curves above the 2:1 ratio. At the 1:1 ratio, the two phases have approximately equal intensity as would be expected for a 2:1 stoichiometry (Figure 4). As with the footprinting results, the diamidine DB185 stabilizes the DNA more than the two monocations. In summary, these results suggest the surprising observation that the two monocationic benzimidazole-amidine compounds bind to the TTAA sequence as dimers but bind as monomer complexes to the AATT site. The dication DB185

(29) Nielsen, P. E.; Møllegaard, N. E.; Jeppesen, C. *Nucleic Acids Res.* **1990**, *18*, 3847−3851.

(30) Nguyen, B.; Tardy, C.; Bailly, C.; Colson, P.; Houssier, C.; Kumar, A.; Boykin, D. W.; Wilson, W. D. *Biopolymers* **2002**, *63*, 281−297.

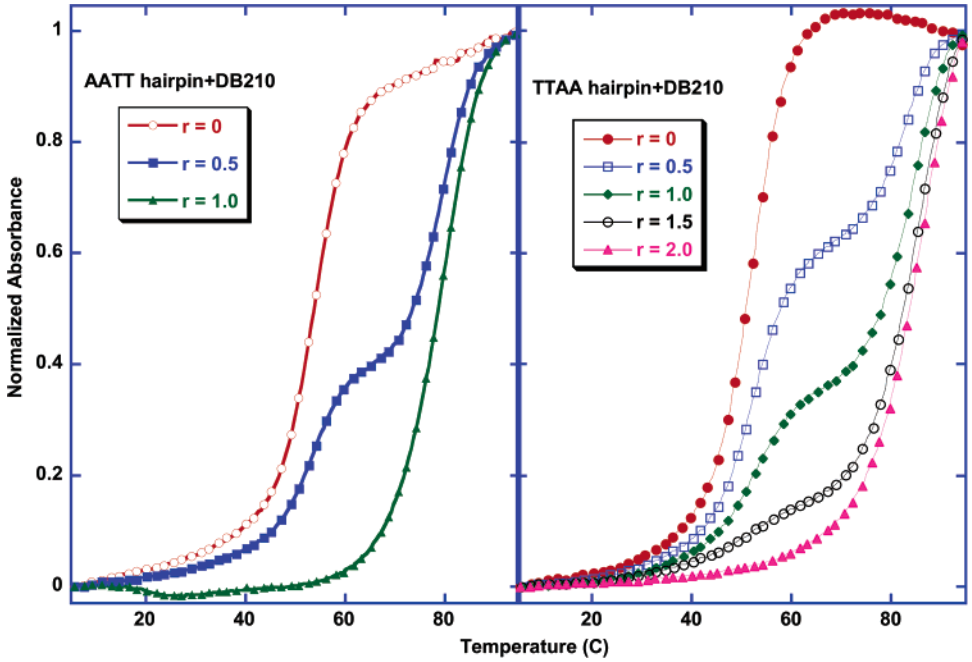

Figure 4. UV melting curves of DNA hairpin duplexes containing the AATT and TTAA sites in the absence and in the presence of DB210 at the indicated ratio of compound to DNA hairpin duplex. DB183 gave essentially identical curves with these sequences. The experiments were conducted in MES 00 and the DNA concentration was approximately 4×10^{-6}M in hairpin strand. The T_m for the AATT hairpin duplex is 53.6 °C and the ΔT_m's are 25.1 °C, 26.4 °C, and >41.5 °C for DB183, DB210, and DB185, respectively, when they bind to this sequence. The T_m for the TTAA hairpin duplex is 52.5 °C, and the ΔT_m's are 30.3 °C, 31.0 °C, and >41.5 °C for DB183, DB210, and DB185 complexes.

does not exhibit dimer binding characteristics with either sequence (Figure S5).

Circular Dichroism. Binding of the benzimidazole cations to the AATT and TTAA sites was characterized by CD spectroscopy in the wavelength range between 220 and 420 nm. The CD spectra monitor the asymmetric environment of the compounds when bound to DNA and therefore can be used to obtain information on the binding mode. The free benzimidazoles do not exhibit CD signals, however, upon addition of the compounds to DNA, substantial positive CD signals arise between 300 and 400 nm, where the compounds, but not DNA, absorb. This large induced CD represents a characteristic pattern for a minor groove binding mode of unfused aromatic cations.[31] The CD results are, however, just a pattern and do not prove a binding mode because some intercalators also have substantial positive induced CD signals. All three benzimidazoles have similar positive induced CD signals in this wavelength range with both the AATT and TTAA binding sites. Example CD spectra for complexes of DB183 and DB185 are shown in the Supporting Information (Figures S6 and S7). Isoelliptic points are observed with the AATT sequence and DB183 at 377, 295 and 248 nm and with TTAA at very similar wavelengths. In conclusion, induced CD spectra on binding of the benzimidazoles to both the AATT and TTAA sites are consistent with a minor groove binding mode for these compounds in agreement with footprinting results.

SPR−Biosensor Binding Determinations. We have found that biosensor−SPR methods are excellent for determination of stoichiometry, cooperativity, and affinity in DNA oligomer-small molecule complexes. Sensorgrams (DNA flow cell − blank) for binding of DB185 to the immobilized AATT and TTAA sequences are compared in Figure 5A. The flow cells for AATT and TTAA oligomers in this experiment have essentially the same amount of DNA immobilized so that the sensorgram saturation levels can be compared directly for stoichiometry differences. In both cases, saturation occurs at RU values of 20−25 RU, and this corresponds to binding of one DB185 per DNA duplex.[17] Saturation of the binding sites is obtained at concentrations of DB185 below 50 nM. As can be seen in Figure 5A, the binding kinetics are slow up to the saturation level, and it is difficult to reach a steady-state plateau at the lower concentrations. Using a single-site binding equation, we estimate binding constants of 2×10^9 and 3×10^{11} for DB185 with the TTAA and AATT sequences, respectively (Table 1). The stronger binding to AATT agrees with the footprinting and Tm results.

DB183 and DB210 have similar sensorgrams and results for these compounds are illustrated with DB183 in Figure 5B. The sensorgrams for these monocations differ in several important aspects from the results with DB185. With the same DNA sequence both DB183 and DB210 associate and dissociate faster than DB185 (Figure 5). With the AATT sequence all three compounds reach similar RU values at saturation and the value corresponds to a single bound molecule. With the TTAA sequence, both DB183 and DB210 reach saturation RU values that are approximately twice as large as with the AATT sequence and that are twice as large as with DB185 and TTAA. As suggested by the Tm results, the sensorgrams indicate that two molecules of DB183 or DB210 bind to the TTAA oligomer, while only one DB185 binds per site. SPR detection is thus a major advantage in small molecule systems of this type where quantitative determination of stoichiometry is of major importance.

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

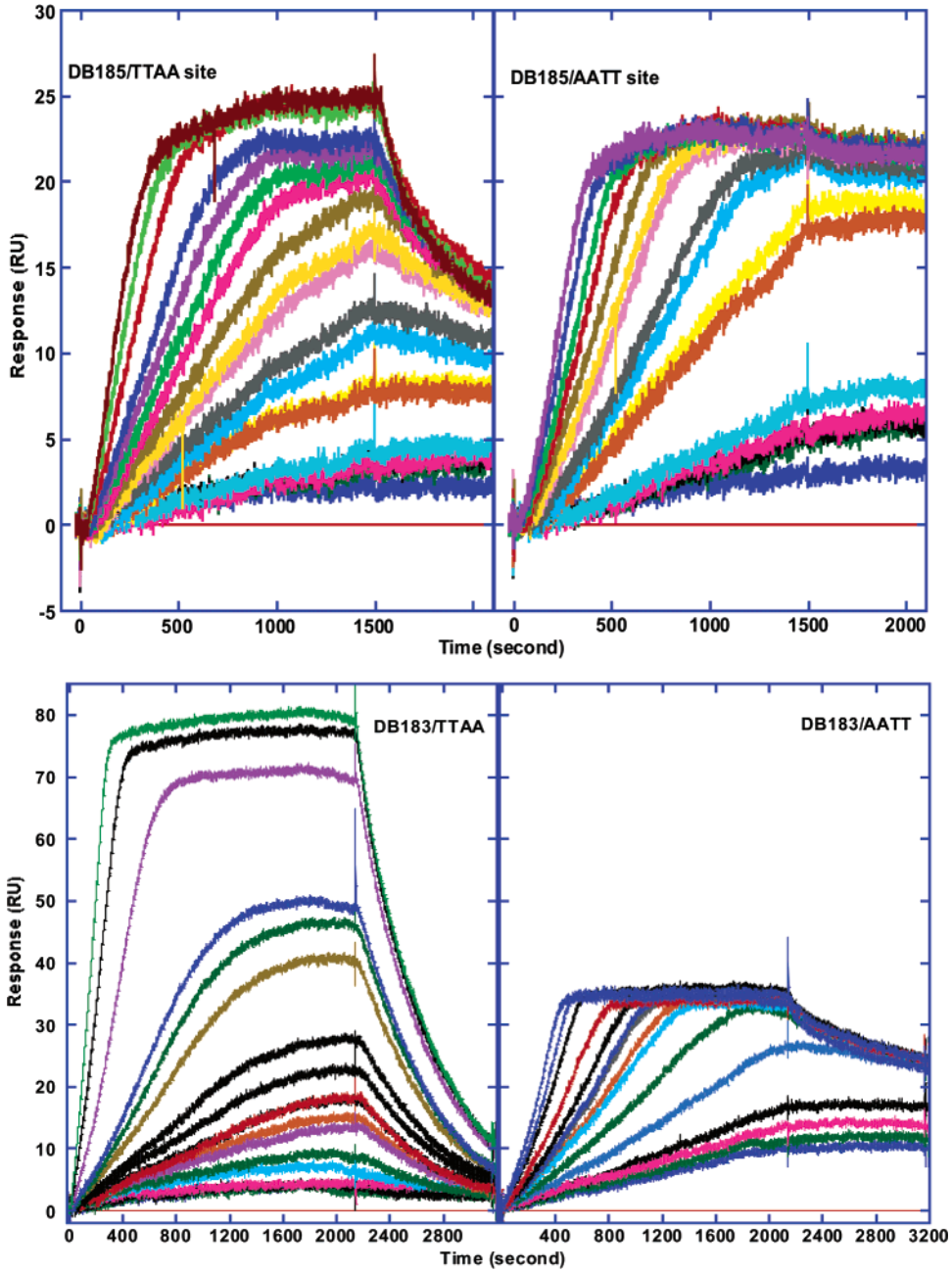

Figure 5. A. SPR sensorgrams for the interaction of DB185 with AATT and TTAA hairpin DNA. The DB185 concentrations are from 1×10^{-11}M (lower curve) to 4×10^{-8}M (highest curve) for AATT binding and from 1×10^{-11}M (lower curve) to 8×10^{-8}M (highest curve) for the TTAA binding. The experiments were conducted at 25 °C in MES 10 with a flow rate of 10 μL/min. B. SPR sensorgrams for the interaction of DB183 with AATT and TTAA hairpin DNA. The DB183 concentrations are from 1×10^{-10}M (lower curve) to 1×10^{-7}M (highest curve) for AATT and TTAA binding. The experiments were conducted at 25 °C in MES 10 with a flow rate of 7μL/min.

Table 1. Binding Constants for the Interaction of DB183, DB210, and DB185 with AATT and TTAA Sites

	DB183			DB210			DB185
sequence	K_1	K_2	$(K_1K_2)^{1/2}$	K_1	K_2	$(K_1K_2)^{1/2}$	K
CGAATTCG		1.6×10^{10}			2.6×10^{11}		$>3 \times 10^{11}$
CGTTAACG	7.2×10^{7}	1.0×10^{9}	2.7×10^{8}	3.0×10^{7}	3.4×10^{9}	3.2×10^{8}	1.8×10^{9}

Experiments were conducted in MES10 at 25 °C. See Figure 1 for the DNA hairpin sequences.

Direct binding plots for DB183 binding to the AATT and TTAA sequences are shown in Figure S8 and plots of very similar shape were obtained for DB210. As with DB185, the results for DB183 binding to the AATT sequence were fit with a single site binding equation with a K near 1.6×10^{10}, over 20 times lower than with DB185 and the AATT sequence (Table 1). The results for TTAA with both DB183 and DB210, however, required fitting with an equation with two DNA binding sites for the monocations to obtain acceptable errors and fitting residuals. The $K1$ values in both cases are smaller than the $K2$ values indicating significant positive cooperativity in binding (Table 1). A Scatchard plot in the Supporting

Table 2. Results for DB183 and DB185 Interacting with [d(CGCGTTAACGCG)$_2$] and [d(CGCGAATTCGCG)$_2$]

	E_{conf}[a,b]	$G_{solvation}$[a]	$-TS_{conf}$[a]	G[a,c]
[d(CGCGTTAACGCG)$_2$]	1140.5 ± 1.1	−5958.9 ± 0.9	−623.7 ± 0.4	−5442.1 ± 2.4
DB183	−54.5 ± 0.2	−64.4 ± 0.0	−47.1 ± 0.1	−166.0 ± 0.3
1:1 complex	530.8 ± 1.2	−5518.0 ± 1.1	−645.3 ± 0.1	−5632.5 ± 2.4
2:1 complex	−135.8 ± 1.1	−5012.2 ± 1.1	−676.1 ± 0.1	−5824.1 ± 2.3
2:1 complex[d]	−85.9 ± 1.1	−5056.6 ± 0.8	−672.5 ± 0.1	−5815.0 ± 2.0
[d(CGCGAATTCGCG)$_2$]	1215.6 ± 1.7	−6032.3 ± 1.6	−623.0 ± 0.1	−5439.7 ± 3.4
1:1 complex	541.7 ± 1.3	−5535.9 ± 1.2	−648.4 ± 0.1	−5642.6 ± 2.6
[d(CGCGTTAACGCG)$_2$]	1140.5 ±1.1	−5958.9 ± 0.9	−623.7 ± 0.4	−5442.1 ± 2.4
DB185	−53.7 ± 0.1	−138.4 ± 0.0	−50.1 ± 0.1	−242.2 ± 0.2
1:1 complex	−22.5 ± 1.0	−5054.8 ± 0.8	−650.4 ± 0.1	−5727.7 ± 1.9
2:1 complex	−1035.0 ± 1.2	−4251.2 ± 1.0	−674.2 ± 0.2	−5960.4 ± 2.4
[d(CGCGAATTCGCG)$_2$]	1215.6 ± 1.7	−6032.3 ± 1.6	−623.0 ± 0.1	−5439.7 ± 3.4
1:1 complex	−8.5 ± 1.1	−5071.2 ± 0.9	−649.7 ± 0.2	−5729.4 ± 2.2

[a] Average energy ± standard error in kcal/mol. [b] $E_{conf} = E_{bond} + E_{angle} + E_{dihedral} + E_{vdw} + E_{elec}$. [c] $G = E_{conf} + G_{solvation} - TS_{conf}$. $T = 300$ K. [d] Two possible ways that DB183 can bind to TTAA; see Figure S10.

Information clearly demonstrates the pronounced positive cooperativity in binding of DB183 to the TTAA site (Figure S9). The square root of the product of the two K values is also reported in Table 1 to allow more direct comparison of the binding on a per molecule basis. As with the AATT sequence, the per molecule K for DB183 binding to TTAA is approximately 10 times lower than the K for DB185 with the same sequence. In summary, the SPR results clearly show that the monoamidine bisbenzimidazoles form cooperative 2:1 complexes with TTAA DNA sequences.

Molecular Dynamics Simulation and Free Energy Calculation Results. To obtain a better understanding of the DB183 and DB185 complexes with AT sites, molecular dynamics simulations were conducted on their complexes with [d(CGCGTTAACGCG)$_2$] and [d(CGCGAATTCGCG)$_2$]. On the basis of the root-mean-square deviation of MD structures with time, the complexes relaxed quickly and remain stable over the entire simulations. The simulations of the 1:1 complex of DB183 with TTAA and AATT were carried out by docking DB183 into the minor groove of TTAA and AATT duplexes 3 Å away from the floor of the grooves with no specific hydrogen bond contacts and with the curvature of compounds complementary to that of the minor groove. The 2:1 complex of DB183 with TTAA can form in two ways with the two DB183 molecules oriented in an antiparallel fashion in the minor groove to allow the charges to point in opposite directions (Figure S10). Both complexes were simulated by putting the two DB183 molecules together in an antiparallel stack and docking them into the minor groove of the DNA duplex as with the monomer complexes. The 2:1 complex with the amidines pointed toward the 5′ end of the adjacent strand had the lower free energy indicating that it is the correct binding mode. As can be seen from Table 2, it is 9.1 kcal/mol lower in energy than the complex with the amidines oriented to the 3′ end of the adjacent DNA chains. Figure S11 shows a schematic representation the 2:1 complex with the lower energy and the hydrogen bonding pattern formed between the two DB183 molecules and the floor of the minor groove of TTAA. Figure 6 shows models for the 1:1 and 2:1 complexes of DB183 with the TTAA site. The dramatic difference in minor groove width in the two complexes is easily seen in this figure.

The calculated free energy of formation of the 1:1 complex with TTAA, −24.4 kcal/mol, is less favorable than for formation of the 2:1 complex, −25.6 kcal/mol, from the 1:1 complex. This result therefore suggests positive cooperativity in binding the second molecule of DB183 as observed experimentally. The formation of the 1:1 complex with the AATT duplex was found to have a free energy of −36.9 kcal/mol; 12.5 kcal/mol more favorable than that of the 1:1 complex of DB183 with TTAA. To compare the different complexes, the average minor groove widths of the duplexes of the 1:1 complexes of DB183 with TTAA and AATT, the 2:1 complex of DB183 with TTAA, free TTAA duplex, and free AATT duplex are shown in Figure S12. The minor groove width of the 2:1 complex of DB183 with TTAA is approximately 10 Å at the central region and is similar to that of the free TTAA duplex. On the other hand, the minor groove of the 1:1 complex of DB183 with TTAA is narrow, relative to that of B-form DNA, and similar to the groove of free AATT and the 1:1 complex of DB183 with AATT.

The results of the molecular dynamics simulations of DB185 with the TTAA and AATT oligomers are summarized in Table 2. The simulations of the 1:1 and 2:1 complexes of DB185 with TTAA and AATT were performed as with DB183. The free energy of formation of the 1:1 complex with TTAA is −43.4 kcal/mol as can be seen from Figure S13. However, the free energy of formation of the 2:1 complex from the 1:1 complex with TTAA is positive, +9.5 kcal/mol, signifying that DB185 does not form a 2:1 complex with TTAA (Figure S13). Figure S14 shows a schematic representation of the 1:1 complex of DB185 with TTAA and the hydrogen bond pattern formed between DB185 and the floor of the minor groove. The free energy of formation of the 1:1 complex of DB185 with AATT is −47.5 kcal/mol and is more favorable than that of the formation of the 1:1 complex of DB185 with TTAA.

Discussion

DNA has been a primary target for anticancer and antiparasitic drugs for over 30 years and remains one of the most promising biological receptors for the development of therapeutic agents against these diseases.[1a,32,33] The recent success of an orally active prodrug of the diamidine, furamidine (Figure 1), against a number of parasitic organisms has greatly expanded the possible therapeutic uses of minor groove binding agents.[1] An exciting extension of the minor groove binding motifs was provided by the discovery that a benzimidazole, DB293 (Figure 1), can bind to specific, mixed base pair, sequences of DNA as a cooperative dimer. This is the first dication to form a dimer in the minor groove, and it expands the sequence recognition

(32) Hurley, L. H. *J. Med. Chem.***1989**, *32*, 2027−2033.
(33) Hurley, L. H. *Nat. Rev. Cancer* **2002**, *2*, 188−200.

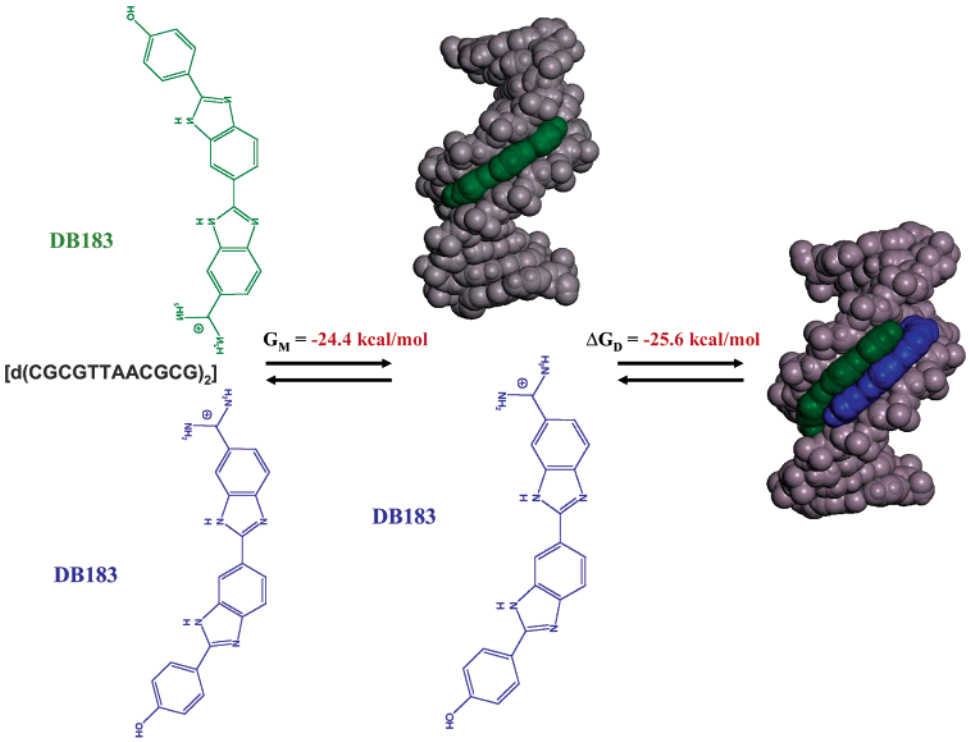

Figure 6. Models and energetics for the formation of 1:1 and 2:1 complexes of DB183 with TTAA: space filling models of the average structure of 1:1 and 2:1 complexes as well as the ΔG values for complex formation.

possibilities for therapeutic targeting of DNA. Given the cell uptake properties and low toxicity of the lead compound in this series, the dimer discovery presents a unique opportunity to develop new therapeutics. To explore and expand this discovery, we are systematically synthesizing and evaluating the DNA interaction properties of benzimidazole-amidine and related derivatives in a multilaboratory collaboration.

H258 has many useful properties as a DNA interactive reagent but it has not achieved any significant therapeutic use. As part of our development program for targeting the minor groove, both monoamidine (DB183 and DB210) and diamidine (DB185) analogues of H258 have been prepared (Figure 1). DNase I footprinting and CD results demonstrate that the compounds bind in the minor groove of DNA with significant binding selectivity for AT base pairs sequences. The footprinting results show that there is significant variation in affinity among different AT sequences that can be targeted by the benzimidazole-amidines. Sites with a central TA step, for example, display lower affinity than similar AT sequences, as has been observed with other diamidines as well as H258.[4,9] One idea for the reduced affinity is that the minor groove is wider in the TA containing sequences than is optimum for monomer interactions.[13,35] It seemed to us, however, that a wider minor groove might be an excellent way to induce a DNA sequence and compound structure dependent stacked dimer complex. We show here that the idea is correct for some sequences containing TA steps.

The complexes of the mono and diamidines were evaluated in detail at AT and TA containing sites with a range of biophysical studies and DNA oligomers. Tm analysis of the amidine complexes with DNA sequences containing AATT and TTAA sites indicated a clear difference in stoichiometry. The dication DB185 has the expected 1:1 maximum stoichiometry with both the AATT and TTAA sequences. The monocations also display 1:1 stoichiometry with AATT but the Tm curves suggest a 2:1 saturation binding ratio with TTAA. These results are confirmed and extended in quantitative detail by biosensor−SPR experiments with immobilized oligomers containing the AATT and TTAA target sites.

The SPR results also indicate a 1:1 stoichiometry of DB185 with both sequences and exceptionally strong binding (Table 1). On the other hand, twice the SPR saturation response for binding of the monocations to the TTAA site, relative to a 1:1 complex at AATT, is observed. These differences in binding suggest that there is a complementary interplay of the compound structure and charge with the surface conformation, chemistry and charge of the minor groove to define the complex stoichiometry and affinity. The SPR results also clearly demonstrate that the 2:1 complex of the monocation forms in TA containing sequences with significant positive cooperativity. The binding energetics are clearly better for binding both mono and diamidines at the AATT site in a 1:1 complex than binding to the TTAA site. At the TTAA sequence the diamidine is able to achieve the most stable complex at a 1:1 ratio, whereas the monoamidine forms the lowest energy species when two monomers form a stacked dimer at the TTAA sequence.

To probe the molecular basis for these unexpected compound structure and DNA sequence dependent binding differences, we conducted MD studies on the complexes of DB183 and DB185 with the AATT and TTAA sites. It is known from a large

(34) Liepinsh, E.; Leupin, W.; Otting, G. *Nucleic Acids Res.* **1994**, *22*, 2249−2254.
(35) (a) Berman, H. M.; Schneider, B. In *Oxford Handbook of Nucleic Acid Structure;* Neidle. S., Ed.; Oxford University Press: New York, 1999, Chapter 9, 295−312. (b) Crothers, D. M.; Shakked, Z. In *Oxford Handbook of Nucleic Acid Structure;* Neidle. S., Ed.; Oxford University Press: New York, 1999, chapter 14, 455−470. (C) Balendiran, K.; Sundaralingam, *J Biomol Struct Dyn.* **1991**, *9*, 511−516.

number of experimental studies that the minor groove in AATT sequences is narrow on average with a well-ordered hydration matrix, whereas with TTAA the groove is wider with a more disordered hydration structure.[13,34] Differences in group positioning and stacking energetics lead to these structural and hydration differences.[35] The groove width and hydration differences are mirrored by our MD calculations (Figure S12) and the MD results provide a unifying molecular explanation for all of the observations in this report. The results with AATT are the same as for many other minor groove binding agents: the groove provides a docking conformation that closely matches the width of the compound. H-bond acceptors help with alignment of the compound in the complex as well as binding specificity and energetics. Release of bound water from the minor groove as well as van der Waals and electrostatic interactions provide the remaining energetics for complex formation. Formation of a 2:1 complex with this sequence would require significant DNA conformational changes that are apparently not favorable for conversion of the 1:1 to a 2:1 complex. The stacking, H-bonding, and electrostatic contributions of the phenyl-amidine group in DB185 yield a very strong interaction of this compound with the AATT sequence as observed experimentally. DB183 is also a strong minor groove binding agent when compared to known compounds, such as H258, but it binds less strongly than DB185. DB210, with an $-OCH_3$ on the phenyl, binds similarly to DB183 indicating that the −OH of DB183 is not playing a strong direct role in the complex.

The results for the TTAA sequence and complexes in Figure S12 are particularly informative. The predicted average groove width of this sequence is significantly wider than that for AATT. To form a 1:1 complex with DB183 at this site requires that the width of the groove be reduced to near that of the 1:1 complex of DB183 with the AATT site (Figure S12). This gives a favorable free energy of binding but it is less favorable than with the AATT receptor site (Table 2). Formation of a 2:1 complex requires a much wider groove but the required groove width is actually quite close to that for the free TTAA sequence (Figure S12). As a result, the 2:1 complex of DB183 with TTAA forms with significant positive cooperativity (Table 2) in agreement with the experimental observations. With the DB185 dication, however, stacking of two molecules in the groove brings two amidine groups quite close together at each end of the complex with strong electrostatic repulsion (Figure S13). As a result, formation of the 2:1 complex with DB185 is disfavored (Figure S13), and the 1:1 complex is the only observed species in our experiments.

Although formation of 2:1 drug:DNA complexes is not restricted to monocations (the dication DB293 forms dimers at ATGA sites [3c−f]), it is obvious that the monocationic nature of DB183 must facilitate its dimerization at TTAA. From our modeling results it is likely that DB183 molecules can arrange cooperatively into head-to-tail complexes to produce a higher affinity DNA binding complex than for monomer interactions. Once formed, the drug dimers can sense the surface of the DNA minor groove to find suitable sites in which they can penetrate. Alternatively, one can envision that a DB183 molecule binds weakly in the minor groove at sites such as TTAA and the bound molecule provides a docking surface to promote cooperative assembly with a second DB183. Drug−DNA recognition is based on the formation of molecular contacts (direct readout) as well as on structural changes of the DNA receptor and solvent interactions (indirect readout). As the second drug molecule (or the preformed drug dimer) penetrates the minor groove, the groove must be wide enough to accommodate the stack and this is only energetically possible at sites with steps such as TpA which display a strong conformational variability and tendency to a wider groove.[13,35,36]

The phosphodiester backbone of TA-containing sequences, in particular the TATA element recognized by the TATA box-binding protein (TBP), is generally flexible and can easily deform to accommodate a ligand or a protein.[37] TBP also recognizes its target sites (TATA boxes) by indirectly reading the DNA sequence through its conformation effects.[38] The factors that combine to produce favorable elements for sequence recognition by TBP include overall flexibility; minor groove widening and relatively low maximal water densities around the DNA.[39,40] DB183, like TBP, can mould the minor groove around the TA sites and optimize the surface complementarity of the complex. The analogy between TBP- and DB183-DNA binding processes also suggests a possible antagonism. Several proteins (such as the nuclear receptors RevErb, NGFI-B, SF-1, and ROR) rely on the intrinsic geometry and flexibility of the TA site to make the required fit. This raises the possibility that compounds designed along the DB183 model could function to control the activity of TA-like elements in promoter regions of specific genes.

Acknowledgment. This work was supported by grants (to W.D.W and D.W.B.) from the National Institutes of Health (GM 61587) and (to C.B.) from the Ligue Nationale Française Contre le Cancer (Equipe labellisée LA LIGUE). The BIAcore instrumentation was purchased through funds from the Georgia Research Alliance. W.D.W. acknowledges an INSERM "Poste Orange" fellowship for research in Lille at INSERM U-524. C.B. thanks Alexandra Joubert for the outstanding footprinting experiments.

Supporting Information Available: DNase I footprinting, and differential cleavage for DB183 and DB185 on 116 and 131 bp HindIII-EcoRI fragments; CD spectra for DB185 and DB183; Scatchard plots for DB183 binding to AATT and TTAA; and MD simulation for DB183 and DB185. This material is available free of charge via the Internet at http://pubs.acs.org.

JA030403+

(36) Tereshko, V.; Urpi, L.; Malinina, L.; Huynh-Dinh, T.; Subirana, J. A. *Biochemistry* **1996**, *35*, 11 589−11 595.
(37) Laughton, C.; Luisi, B. *J Mol Biol.* **1999**, *288*, 953−963.
(38) (a) Bareket-Samish, A.; Cohen, I.; Haran, T. E. *J. Mol. Biol.* **2000**, *299*, 965−977. (b) Moravek, Z.;, Neidle, S.; Schneider, B. *Nucleic Acids Res.* **2002**, *30*, 1182−1191.
(39) Qian, X.; Strahs, D.; Schlick, T. *J. Mol. Biol.* **2001**, *308*, 681−703.
(40) Umezawa, Y.; Nishio, M. *Bioorg. Med. Chem.* **2000**, *8*, 2643−2650.