# Unusually Strong Binding to the DNA Minor Groove by a Highly Twisted Benzimidazole Diphenylether: Induced Fit and Bound Water†

Farial A. Tanious,‡ William Laine,§ Paul Peixoto,§ Christian Bailly,§,|| Kristie D. Goodwin,⊥ Mark A. Lewis,@ Eric C. Long,@ Millie M. Georgiadis,⊥ Richard R. Tidwell,# and W. David Wilson*,‡

*Department of Chemistry, Georgia State University, P.O. Box 4098, Atlanta, Georgia 30302-4098, INSERM U-837, Institute for Cancer Research, Lille 59045, France, Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, and Department of Chemistry and Chemical Biology, Purdue School of Science, Indiana University-Purdue University Indianapolis (IUPUI), Indianapolis, Indiana 46202, and Department of Pathology, University of North Carolina, Chapel Hill, North Carolina 27599*

*Received February 9, 2007; Revised Manuscript Received March 29, 2007*

ABSTRACT: RT29 is a dicationic diamidine derivative that does not obey the classical "rules" for shape and functional group placement that are expected to result in strong binding and specific recognition of the DNA minor groove. The compound contains a benzimidazole diphenyl ether core that is flanked by the amidine cations. The diphenyl ether is highly twisted and gives the entire compound too much curvature to fit well to the shape of the minor groove. DNase I footprinting, fluorescence intercalator displacement studies, and circular dichroism spectra, however, indicate that the compound is an AT specific minor groove binding agent. Even more surprisingly, quantitative biosensor-surface plasmon resonance and isothermal titration calorimetric results indicate that the compound binds with exceptional strength to certain AT sequences in DNA with a large negative enthalpy of binding. Crystallographic results for the DNA complex of RT29 compared to calculated results for the free compound show that the compound undergoes significant conformational changes to enhance its minor groove interactions. In addition, a water molecule is incorporated directly into the complex to complete the compound−DNA interface, and it forms an essential link between the compound and base pair edges at the floor of the minor groove. The calculated $\Delta C_p$ value for complex formation is substantially less than the experimentally observed value, which supports the idea of water being an intrinsic part of the complex with a major contribution to the $\Delta C_p$ value. Both the induced fit conformational changes of the compound and the bound water are essential for strong binding to DNA by RT29.

Heterocyclic cations that bind strongly and specifically to the DNA minor groove are of interest for the fundamental information that they provide about DNA recognition mechanisms as well as for their therapeutic and biotechnology uses (*1*−*14*). Heterocyclic diamidines, for example, have shown excellent activity against several eukaryotic parasitic diseases, and a prodrug of DB75 (Figure 1) is in Phase III clinical trials against sleeping sickness (*1*−*3*). Parasitic diseases, such as sleeping sickness, are caused by vector-transmitted kinetoplastid parasites, and there is currently an alarming worldwide spread of these diseases (*1*−*3*). The seriousness of the problem is emphasized by the fact that, in the absence of treatment, the outcome of sleeping sickness is death of the patient. Fluorescence microscopy of kinetoplastid parasites after treatment with fluorescent diamidines clearly shows that the compounds selectively target the mitochondrial kinetoplast, with subsequent destruction of the kinetoplast and cell death (*2*, *3*). The extensive AT sequences of the kinetoplast minicircle DNA are a favorable target for the diamidines, and the interaction appears to interfere with replication of the catenated kinetoplast DNA (kDNA) network (*2*, *3*, *15*). The minor groove in AT sequences can more easily assume the narrow width that is required for tight binding of the heterocyclic−amidine system of the dications than the groove in G·C or mixed base pair sequences (*3*, *4*, *8*−*10*). The extra H-bond of the G·C base pairs typically leads to a wider minor groove and presents a steric block to deep penetration of compounds into the groove (*4*, *11*).

Structural studies of the DNA interactions of minor groove binding agents over many years have led to derivation of a set of rules concerning shape and properties of the compounds that are required for effective minor groove complex formation in AT sequences (*3*−*5*, *10*, *12*−*14*). In particular, fairly rigid criteria for compound curvature and functional group positioning in the compound have been defined for optimum DNA fit and interactions (*9*, *16*, *17*). Important molecular features of minor groove binders include a crescent

† This work was supported by National Institutes of Health Grants AI064200 (to W.D.W.), GM 62831 (to E.C.L.), and GM055026 (to M.M.G.), by the Bill and Melinda Gates Foundation (to W.D.W. and R.R.T.), and by Ligue Nationale Francaise Contre le Cancer (to C.B.). SPR and ITC instruments were funded in part by the Georgia Research Alliance.

* To whom correspondence should be addressed. Telephone: (404) 651-3903. Fax: (404) 651-1416. E-mail: wdw@gsu.edu.

‡ Georgia State University.

§ Institute for Cancer Research.

|| Present address: Institut de Recherche Pierre Fabre, 3 rue des satellites, 31400 Toulouse, France.

⊥ Indiana University School of Medicine, IUPUI.

@ Purdue School of Science, IUPUI.

# University of North Carolina.

10.1021/bi700288g CCC: $37.00 
Published on Web 05/17/2007

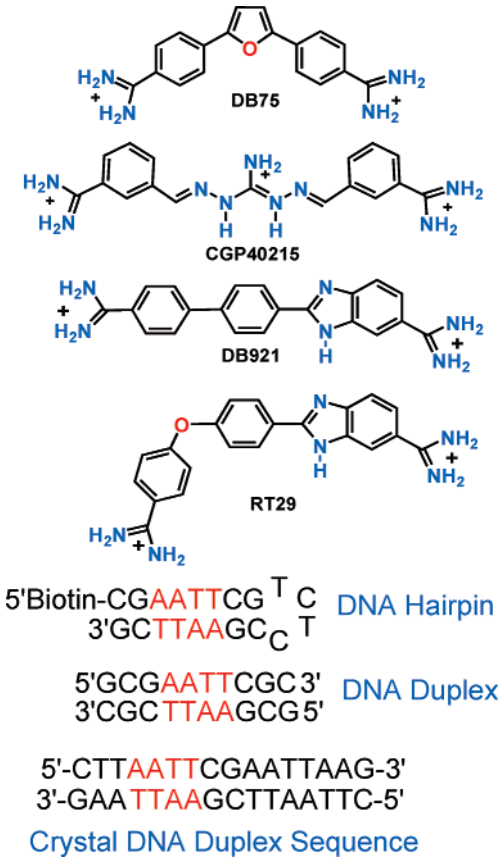


FIGURE 1: Structures for RT29 with other diamidines and DNA hairpin and self-complementary duplex sequences.

shape that complements the helical minor groove, positive charges to enhance electrostatic interactions, appropriately placed inward-facing H-bonding groups for sequence recognition, and an extended unfused heterocyclic structure to allow optimization of the compound for DNA minor groove interactions (*4*, *11*, *17*). These rules have been used successfully to design many minor groove-binding heterocycles, such as DB75, with strong minor groove interactions, significant biological activity, and low general toxicity (*1*−*3*, *17*).

Proteins that recognize DNA, however, use a broad array of different structural units and shapes and frequently involve water molecules in their DNA complexes (*18*, *19*). These results suggest that an array of minor groove interactions by small molecules much broader than that previously observed or envisioned might be possible. Discovery and design of new modes for minor groove recognition could lead to new classes of antiparasitic and anticancer drugs with improved and broader activity as well as different structures for bypassing possible resistance mechanisms if they should develop with classically shaped minor groove compounds. Such nonclassical but relatively simple compounds with well-defined complexes could also improve our fundamental understanding of the recognition of DNA and the role of solvent in DNA molecular recognition (*17*).

In our search for variations on minor groove binding motifs, several criteria have been used to focus the design and search. Diamidine derivatives of different structures have been investigated since such compounds have demonstrated a variety of excellent DNA interaction, cell uptake, and therapeutic properties (*1*−*3*, *20*, *21*). Recently, for example, two relatively linear diamidines, CGP40215 and DB921 (Figure 1), that bind strongly to the minor groove, even though they do not follow the classical rules described above, have been discovered and characterized (*3*, *22*−*25*). Crystallographic analysis of the compounds bound to an -AATT- minor groove site has shown that a water molecule forms a critical part of the complex in both cases. This finding supports the concept that diamidines with varied structure and functional groups may be able to escape the classical constraints on minor groove complexes and yield a wide variety of structures not predicted by the rules described above.

Benzimidazole heterocycles have exhibited excellent DNA interactions in a number of systems (*3*, *4*, *14*, *25*−*29*), and we included benzimidazole derivatives in our design and search exploration for new types of minor groove binding agents. RT29 (Figure 1) is a benzimidazole diamidine with a diphenyl ether group, and because of this linkage, the compound is too highly curved and twisted to bind to the minor groove by the classical mode. In fact, the predicted shape of the compound in solution suggests, on the basis of a comparison to standard minor groove binding agents (*17*), that it should bind only weakly to DNA by nonspecific electrostatic interactions. Surprisingly, biological testing results revealed that the compound has excellent antitrypanosomal activity. In a search for the biological target and mechanism of therapeutic action for RT29, the quite unexpected observation that it displays very strong and selective binding to AT sequences in DNA was made (*30*). The binding is significantly stronger than that of DB75 or related compounds that actually fit the classical rules for formation of minor groove complexes. RT29 presents a striking new motif for DNA recognition and provides new ideas for the design of nonclassical heterocyclic cations for recognizing the DNA minor groove. To fully understand the DNA interactions of unusual minor groove binding compounds, such as RT29, it is essential to characterize the energetics of complex formation in detail. To accomplish this goal, we have used DNase I footprinting and fluorescence intercalator displacement methods to characterize the sequence selectivity of RT29 binding in more detail. Detailed thermodynamic studies have been used to better understand the energetic basis for the strong binding of RT29. All of the results are correlated with the observed X-ray structure of RT29 bound into the minor groove (*30*), and they provide an essential energetic understanding to complement the structural details of the DNA complex.

## MATERIALS AND METHODS

*Compounds, DNAs, and Buffers.* RT29 was synthesized as previously described (*31*). Oligomers and 5′-Biotin-labeled hairpin DNA oligomers were from Midland Certified Reagent Co. with reverse phase HPLC purification and mass spectrometry characterization (Figure 1). PolydA·polydT was from Pharmacia. Tris buffers contained 0.01 M Tris-HCl and 0.001 M EDTA adjusted to pH 7.4. Tris10, Tris20, Tris50, and Tris100 buffers contained those two components and 0.10, 0.20, 0.50, and 1.0 M NaCl, respectively. Cacodylic

[1] Abbreviations: SPR, surface plasmon resonance; ITC, isothermal titration calorimetry; Tris, tris(hydroxymethyl)aminomethane; EtBr, ethidium bromide; FID, fluorescence intercalator displacement; DNase I, bovine pancreatic deoxyribonuclease I.

buffer was used for $T_m$ and ITC experiments; it contained 10 mM cacodylic acid, 1 mM EDTA, and 200 mM NaCl, and the pH was adjusted to 6.25.

*DNase I[1] Footprinting: Purification of DNA Restriction Fragments and Radiolabeling.* The different plasmids were isolated from *Escherichia coli* by a standard sodium dodecyl sulfate−sodium hydroxide lysis procedure and purified by banding in CsCl−ethidium bromide gradients. The 117 bp DNA fragment was 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 same thing was done with the 176 bp fragment (pTUC plasmid). 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) and 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. After being washed with 70% ethanol and after the precipitate had been vacuum-dried, 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 PhosphorImaging.* Bovine pancreatic deoxyribonuclease I (DNase I, Sigma Chemical Co.) was stored as a 7200 units/mL solution in 20 mM NaCl, 2 mM $MgCl_2$, and 2 mM $MnCl_2$ (pH 8.0). The stock solution of DNase I was kept at −20 °C and freshly diluted to the desired concentration immediately prior to use. Footprinting experiments were performed essentially as previously described (*11*, *32*). 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 the appropriate concentration. After incubation for 30 min 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 ∼0.01 unit/mL in the reaction mixture. After 3 min, the reaction was stopped by freeze-drying the mixture. 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 denaturing conditions (0.3 mm thick, 8% acrylamide containing 8 M urea). After electrophoresis (∼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. 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).

*Fluorescence Intercalator Displacement Analyses.* A DNA library of 136 different oligonucleotide hairpins was purchased from Trilink Biotechnologies, Inc., as individual lyophilized solids. Concentrations of the hairpin deoxyoligonucleotides were determined by the method described by Boger (*33*, *34*) using UV at 90 °C and single-strand extinction coefficients to ensure accurate concentration determination. To carry out the assay, each well of a Costar black 96-well plate was loaded with Tris buffer containing ethidium bromide and one hairpin deoxyoligonucleotide of the library. Typical final concentrations in each well were 1.5 μM DNA hairpin, 4.5 μM EtBr, and 0.75−3 μM DNA binding agent. The final buffer consisted of 10 mM Tris (pH 8.0) and 100 mM NaCl. After incubation at 25 °C for 30 min, fluorescence measurements were taken (average of three measurements) on each well with a Varian Cary Eclipse fluorescence plate reader ($\lambda_{ex}$ = 545 nm; $\lambda_{em}$ = 595 nm). Compound assessments were conducted in triplicate (or more) with each well acting as its own control well (no agent, 100% fluorescence, and no DNA, 0% fluorescence). Fluorescence readings are reported as a percent fluorescence decrease relative to the control wells.

*Circular Dichroism (CD).* CD spectra were collected with a Jasco J-810 spectrometer at different ratios of compound to DNA at 25 °C in MES 10 buffer. A DNA solution in a 1 cm quartz cuvette was first scanned over a desired wavelength range. The compounds at increasing ratios were then titrated into the same cuvette and the complexes rescanned under the same conditions.

*Thermal Melting ($T_m$).* $T_m$ experiments were conducted with a Cary 300 UV−visible spectrophotometer in 1 cm quartz cuvettes. The absorbance of the DNA−compound complex was monitored at 260 nm as a function of temperature, and DNA without compound was used as a control. Cuvettes were mounted in a thermal block, and the solution temperatures were monitored by a thermistor in a reference cuvette with a heating rate of 0.5 °C/min. The concentration of DNA was determined by measuring the absorbance at 260 nm. A compound:DNA oligomer duplex ratio of 1:1 was used for the complex or a compound:DNA base pair ratio of 0.6 with the polymer. Higher ratios did not result in any significant $T_m$ increase.

*Calculated Equilibrium Geometry.* The equilibrium geometry of RT29 when not bound to DNA was obtained by full geometry optimization at both the Hartree−Fock (HF) and density functional (B3LYP) levels with a series of basis sets [6-31G(d,p) and higher]. All quantum chemical calculations were carried out with Spartan04.

*Calculated Solvent Accessible Surface Areas.* Solvent accessible surface areas (SASAs) of the RT29−DNA complex (PDB entry 2FJV) were calculated using NACCESS version 2.1.1. This program outlines the accessible surface of a molecule using a path generated by the center of a 1.4 Å radius probe rolled around the van der Waals surface of the molecule based on the method of Lee and Richards (*35*, *36*). The surface area is calculated by slicing the three-dimensional molecular volume to determine the accessible surface of individual atoms. SASAs for unbound DNA or RT29 were calculated using only the coordinates for DNA or RT29, respectively.

$$\Delta SASA_{total} = SASA_{complex} - (SASA_{free\ DNA} + SASA_{free\ RT29})$$

*Biosensor-Surface Plasmon Resonance (SPR).* Biosensor-SPR experiments were conducted with BIAcore 2000 or 3000 (Biacore, Uppsala, Sweden) instruments. 5′-Biotin-labeled DNA hairpins were immobilized on streptavidin-linked sensor chips via noncovalent capture. For kinetics experiments, a Biacore CM4 chip, which has a lower degree of carboxymethylation (relative to the standard CM5 chips), was used. Streptavidin was linked to this chip by standard amine coupling chemistry (1800 RU), and biotinylated DNA was captured on the chip surface. The tight affinity between biotin and streptavidin yields a highly stable surface over time, allowing for regeneration with relatively harsh conditions. After the chip had been washed with 1 M NaCl and 50 mM NaOH and buffer to remove unlinked streptavidin, a 25 nM DNA of a selected sequence was manually injected over each streptavidin-derivatized flow cell at a flow rate of 2 $\mu$L/min until the desired amount of each DNA sequence (RU) was immobilized. Three flow cells contained DNA, and one flow cell was left blank as a reference.

The compounds at different diluted concentrations in degassed and filtered Tris buffer with 0.005% surfactant P-20 were injected over the DNA surface for a selected time. A 10 mM glycine solution at pH 2.0 was used for regeneration. Steady-state binding studies were carried out by averaging the resonance unit values (RU) over a selected time region at different compound concentrations and converting them to $r$ (moles of bound compound per mole of DNA hairpin); $r = RU_{eq}/RU_{max}$, where $RU_{eq}$ is the averaged RU at steady state for each concentration and $RU_{max}$ is the maximum RU value for binding one compound per binding site. The binding constant ($K_a$) was obtained from the best fit of $r$ versus the free compound concentration ($C_{free}$) with a one-site binding model:

$$r = (K_a C_{free})/(1 + K_a C_{free}) \quad (1)$$

Fits with more than one site did not improve the correlation coefficient or residuals significantly for RT29. Analysis of kinetics results was performed by global fitting of the binding data with mass transport-limited binding models (*37*, *38*). Equations 2 and 3 are standard 1:1 kinetics models, while in a mass transport limitation model, eqs 2−5 are used:

$$A + B \leftrightarrow AB$$

$$[A]_{t=0} = 0,\ [B]_{t=0} = cRU_{max},\ [AB]_{t=0} = 0$$

$$K_a = [AB]/([A][B]) \quad (2)$$

$$d[AB]/dt = k_a[A][B] - k_d[AB] \quad (3a)$$

$$d[RU]/dt = k_a[A][RU_{max} - RU] - k_d[RU] \quad (3b)$$

$$d[A]/dt = k_t([A_{bulk}] - [A]) - k_a[A][B] + k_d[AB] \quad (4)$$

$$d[B]/dt = -k_a[A][B] + k_d[AB] \quad (5)$$

where [A] and $[A_{bulk}]$ are the concentrations of the compound at the sensor surface and in the bulk solution, respectively, [B] or $c[RU_{max} - RU]$ is the concentration of the DNA, [AB] or cRU is the concentration of the complex, $cRU_{max}$ is the response at saturation of binding sites, $k_a$ is the association rate constant, $k_d$ is the dissociation rate constant, and $c$ is a proportionality constant that relates RU to concentration. The constant $c$ allows conversion of eq 3a to eq 3b and direct use of the response in determining kinetic constants. Global fitting of entire sensorgrams at all concentrations was used to determine the kinetics constants.

*Isothermal Titration Calorimetry (ITC).* ITC experiments were performed with a MicroCal VP-ITC instrument (MicroCal Inc., Northampton, MA) interfaced with a computer equipped with VP-2000 viewer instrument control software and Origin software for data analysis. In ITC experiments, 10 $\mu$L of compound in cacodylic buffer was injected into the sample cell containing DNA duplexes in the same buffer. The observed heat for each injection was measured by integration of the peak areas with respect to time. Blank titrations with no DNA were conducted by injecting the compound into the sample cell containing only buffer under the same conditions. The corrected interaction heat was determined by subtracting the blank heats from those in the compound/DNA titrations.

## RESULTS

*DNase I Footprinting Studies Establish the Binding Sites in Natural DNA.* DNase I footprinting results for RT29 (Figure 2A,B) indicate that the diamidine binds extremely tightly to specific AT-rich sequences of DNA to produce pronounced footprints within the 117 and 176 bp restriction fragments used for these experiments. For the 117 bp fragment, footprints centered on nucleotide positions 20−30 correspond to the compound binding to juxtaposed AT sequences with a single separating GC base pair (5′-AATTGTAATA). Sites around position 42 (5′-TAAAA), 63 (5′-TTTT), and 85 (5′-ATTAA) have similar, strong footprints. Interestingly, a protected site around position 74 contains only three consecutive A·T bp (5′-TAAC). This finding agrees with ethidium displacement results (below) that indicate significant binding by RT29 to some sites with three consecutive A·T base pairs and a 5′ or 3′ G·C base pair. Footprinting results for the DNA fragment of 176 bp (Figure 2B) confirm that RT29 binds very strongly to AT DNA sequences and particularly strongly to AATT. All of the footprinting results show that RT29 has a strong binding preference for sites with four or more A·T base pairs but with some tolerance for specific sites with three A·T and a terminal G·C base pair.

*Fluorescence Intercalator Displacement (FID): Analysis of RT29 Site Selectivity.* Although DNase I footprinting can scan long DNA sequences, it does not have nucleotide resolving power, and we have thus turned to the FID method to investigate specific sequences in greater detail. Previous analyses of the sequence selectivity of RT29 via ethidium displacement (using an oligonucleotide library containing all possible 4 bp sequences) indicated that RT29 interacted preferentially with DNA sites containing four contiguous A·T base pairs (such as AATT and ATTA) and also with sites that have three A·T base pairs followed by a C or G base pair (*30*) (Figure 3A); the site of this latter type preferred by RT29 was determined to be ATTC. Interestingly, distinct from the behavior of netropsin in this assay (*33*, *34*, *39*), which clearly revealed a preference for four A·T base pair sites only, RT29 revealed that four A·T base pair sites where interspersed among the similarly preferred three A·T base pair sites in the rank order (Figure 3). While the footprinting

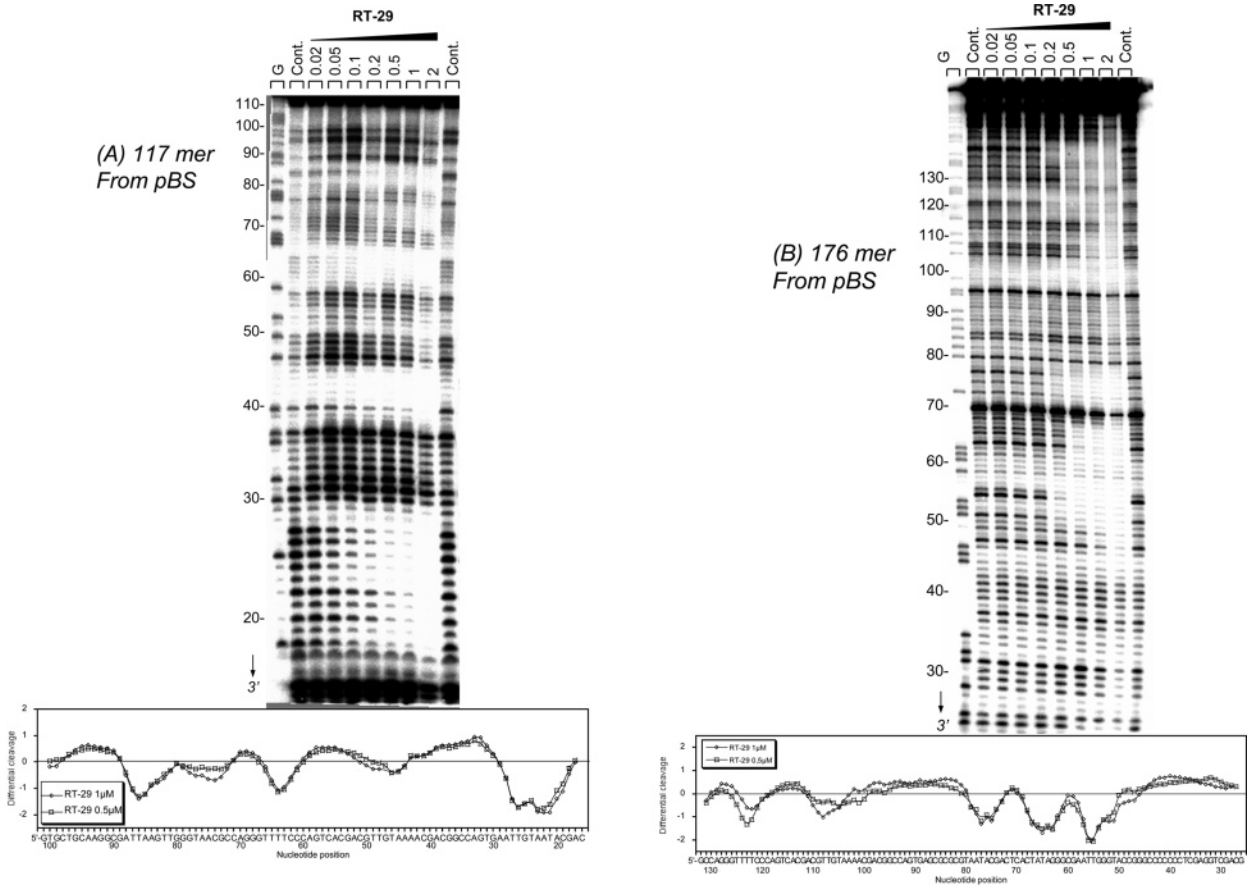


FIGURE 2: DNase I footprinting for RT29 bound to the (A) 3′-end radiolabeled 117 bp restriction fragment from pBS and (B) the 176 bp restriction fragment from pTUC. The cleavage products from the DNase I digestion were resolved on an 8% polyacrylamide gel containing 8 M urea. The concentration (micromolar) of the drug is shown at the top of the appropriate gel lanes. Control tracks labeled "Cont." contained no drug. The track labeled G represents dimethyl sulfate−piperidine markers specific for guanine. The numbering of the sequence is shown at the left. Densitometer traces with sequence numbering are shown below the gel results.

results for RT29 described earlier provide support for this finding, through the footprint observed at the TAAC site in the 117 bp fragment, an ATTC site was absent in the sequences used in the footprinting experiments, and those sequences contain few $(A{\cdot}T)_3$-G·C sites. A histogram of the FID results for analysis of all possible 4 bp $(A{\cdot}T)_3$-G·C sites is shown in Figure 3. As noted previously, the strongest $(A{\cdot}T)_3$-G·C site contained the sequence ATTC, and like the behavior of other AT-rich DNA-targeted molecules, none of the highest-ranked sequences with three AT base pairs preferred by RT29 have a 5′-TA-3′ step; sites containing TpA steps appear only within the lower half of the rank ordered sites among this group of oligonucleotide targets.

Along with our examination of $(A{\cdot}T)_3$-G·C sites in general, we also examined details of the interaction of RT29 with the favored ATTC type of sites. Our X-ray structural analysis of RT29 bound to ATTC determined that RT29 interacts with this particular site, in part, through the formation of a water-mediated hydrogen bond between the RT29 benzamidine moiety and the nucleobase present at the 3′-flank of the ATTC site, i.e., to the N in ATTCN (*30*). While the observation of a water-mediated interaction suggests that the identity of the nucleobase present at N should not strongly influence RT29 binding, direct evidence to support this notion was not available. To probe the ability of RT29 to target the ATTC site in greater detail, the FID method was used further to investigate the influence of ATTC flanking sequences contained in oligonucleotide sequences of the type 5′-ATTCN- (where N is A, T, C, or G in a hairpin duplex of the type previously used). All complexes formed were found to have a 1:1 RT29:duplex stoichiometry (data not shown), and sites for which N was T or A (T > A) exhibited slightly stronger binding to RT29 in comparison to those for which N was G or C within the drug concentration range of 0−0.5 $\mu$M (Figure 3). These results indicate that, as expected on the basis of the observed water-mediated interactions, all ATTC-like binding sites have a similar affinity for RT29 with little flanking sequence influence. These observations further suggest that the overall groove width of ATTC sites, and perhaps 5′-$(A{\cdot}T)_3$-(G·C) sites in general, influences binding to a greater extent than specific intermolecular interactions.

*Circular Dichroism: Probing the Binding Mode in Solution.* CD spectra of the diamidine upon titration into a solution of a DNA oligomer with an -AATT- sequence were monitored from 220 to 400 nm. At the maximum absorption wavelength of the compound, where the DNA CD signals do not interfere, positive induced CD signals were observed around 325 nm (Figure 4). RT29 has no intrinsic CD signals, and these signals are for the bound compound. Such positive induced CD signals are generally a characteristic of binding in the minor groove of DNA (*40*, *41*) as observed for other diamidines. In summary, CD results indicate that RT29 binds in the DNA minor groove in AT sequences in solution in spite of the pronounced molecular twist in the RT29 diphenyl ether system.

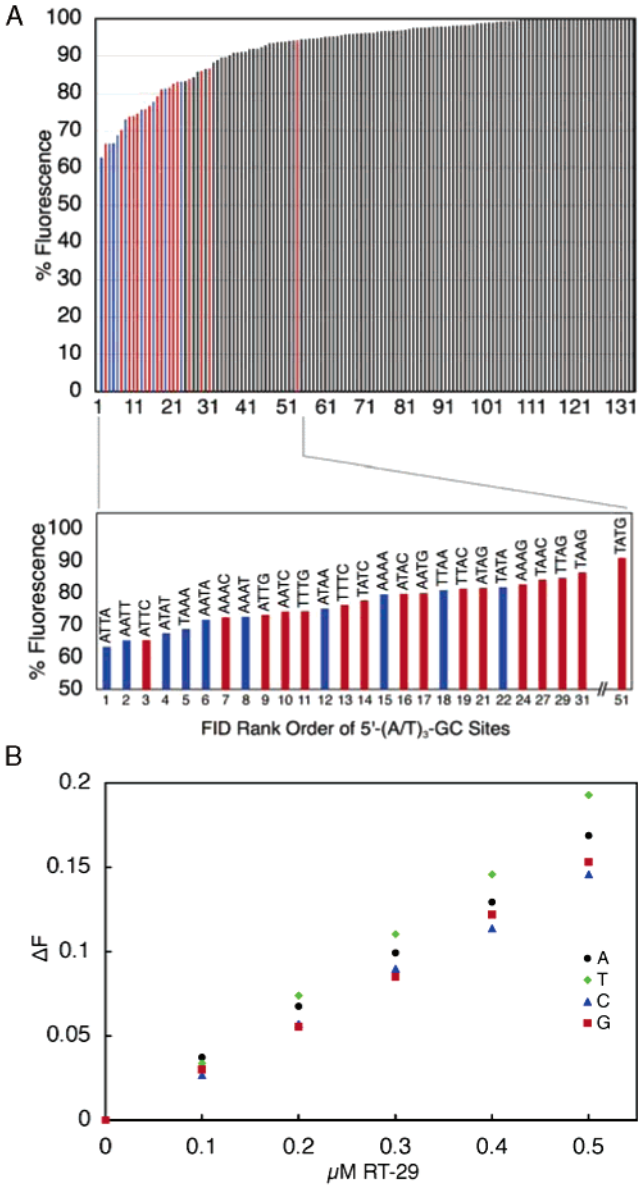


FIGURE 3: (A) FID analysis of binding of RT29 (0.75 $\mu$M) to all possible four base pair DNA sequences emphasizing the overall rank and relative rank order of sites containing three contiguous A·T base pairs flanked by one G·C base pair [5′-(A·T)$_3$-G·C sites]. The merged-bar FID analysis histogram in the top panel is color-coded such that four base pair A·T-only DNA cassettes are blue [5′-(A·T)$_4$ sites] while those containing 5′-(A·T)$_3$-G·C sites are red; all other sites are colored black. (B) The plot illustrates the change in *F* (decrease) that occurs as ethidium bromide is displaced from oligonucleotides containing 5′-ATTC-N sites as a function of increasing RT29 concentration; these data suggest that RT29 exhibits a slight preference for 5′-ATTC-T sites (N=T > A > G > C).

*RT29−DNA Thermal Melting Studies: Relative Binding Affinity.* Thermal melting studies with polydA·polydT serve as a qualitative screen for compound binding affinity for AT sequences (*42*). $\Delta T_m$ values (the $T_m$ of the DNA complex minus the $T_m$ for DNA) for RT29 with polydA·polydT were compared to the value for DB75 ($\Delta T_m$ = 24.8 °C; the error in $\Delta T_m$ values is ±0.5 °C) under the same conditions. Surprisingly for the highly twisted RT29, the binding is so strong that the DNA complex melts above 95 °C and no transition could be observed ($\Delta T_m$ > 28 °C). $T_m$ studies were also conducted with a DNA hairpin duplex containing a single -AATTAA- site, the same sequence as in the crystallographic analysis. The oligomer $T_m$ is 61.2 °C; at a 1:1 ratio of RT29, the $T_m$ increases to 75.4 °C, while at a 2:1 ratio, the $T_m$ increases only to 76.2 °C. This result indicates that RT29 binds to the oligomer duplex in a 1:1 complex as expected from the designed sequence and that it binds to DNA significantly more strongly than the classically shaped minor groove agent, DB75.

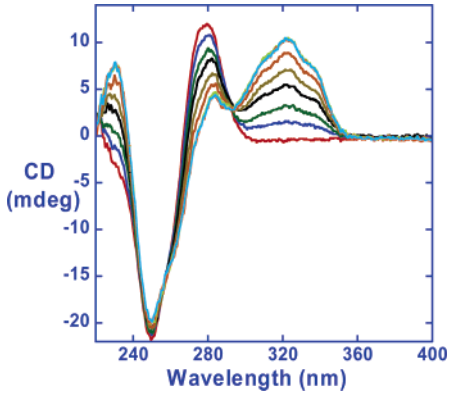


FIGURE 4: Induced CD signals for RT29 with the d(CGCGAATTCGCG)$_2$ duplex in MES 10 buffer at 25 °C. Molar ratios of compound to DNA duplex are 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, and 1.5 from the bottom to the top at the induced CD wavelength maximum near 320 nm.

*Biosensor-Surface Plasmon Resonance (SPR): Quantitative Binding Affinity and Stoichiometry.* To quantitatively evaluate the DNA interactions of RT29, SPR experiments were conducted with hairpin duplexes containing -AATT- sequences, which have been used in studies of many minor groove binding compounds and which have the strongest RT29 affinity in FID experiments, -AATTAA- (the same one that was used in the $T_m$ and crystal structure studies), as well as an alternating G·C base pair sequence as a control. Because the SPR approach responds to mass, it can be used in comparative studies for diamidines that have very large differences in properties and equilibrium binding constants (*K*) as well as weak spectral signals. In many cases, rates for the binding can be observed in real time as the magnitude of the SPR signal increases with formation of the compound−DNA complex. A number of detailed and extensive comparative studies have conclusively shown that carefully conducted biosensor-SPR studies provide results equivalent to those of solution methods (*37*, *38*, *43*−*46*).

In agreement with the DNase I footprinting and FID results, the observed binding of RT29 to the two hairpin AT sequences is quite strong while no significant binding could be detected to the GC sequence at concentrations that saturated the AT binding sites. The binding of RT29 to the AATT (Figure 5A) and AATTAA hairpins is very similar under the experimental conditions (Table 1). Because of the strong binding, low concentrations were required to obtain a mix of free and bound DNA sites, and at the lowest concentrations, a steady-state region was not reached in the 500 s maximum, volume-limited, flow time for the experimental conditions (Figure 5A). Since the observed association rate is a mix of bimolecular complex formation and mass transfer (eq 4), the observed rate increases with an increase in the concentration of RT29 in the flow solution. A steady-state plateau was obtained at the higher concentrations. The observed dissociation rate is also quite slow and is not

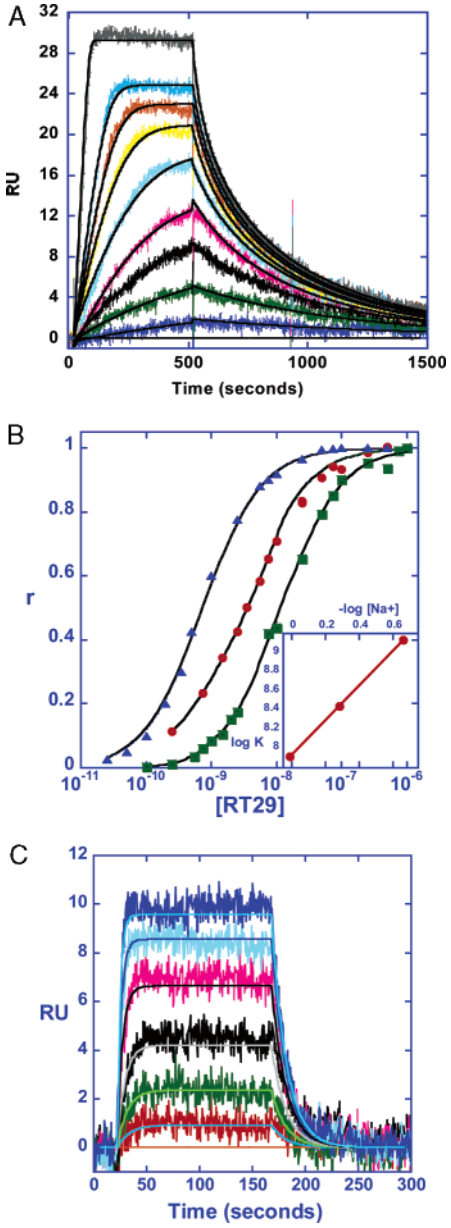


FIGURE 5: (A) SPR sensorgrams for binding of RT29 with AATT in Tris50 buffer at 25 °C. The smooth black lines are best fit curves by global kinetics fitting as described in the text. For RT29, the concentrations from bottom to top are 0, 0.25, 0.75, 1.5, 2.5, 3.5, 5.5, 7.5, 10, and 20 nM. (B) Binding results for RT29 with the AATT hairpin sequence (Figure 1) at different salt concentrations [0.2 (▲), 0.5 (●), and 1.0 (■) M NaCl] at 25 °C. The inset shows the linear relation of log $K$ to $-\log[Na^+]$ which gives a slope of 1.8. (C) SPR sensorgrams for the binding kinetics of RT29 with AATT in Tris100 buffer at 25 °C and a flow rate of 100 $\mu$L/min on a low-density C4 streptavidin chip. The smooth lines are best fit curves by global kinetics fitting. The RT29 concentrations from bottom to top are 0, 2.5, 5.5 10, 20, 30, and 40 nM.

complete in 1000 s, even at a salt concentration of 0.5 M. As one can see in Figure 5B, the compound binding sites are essentially saturated at RT29 concentrations above 25 nM at 0.2 M NaCl. The concentration required for saturation binding increases with an increase in the salt concentration (Figure 5B).

To evaluate the binding equilibrium constants ($K$), the experimental sensorgrams were quantitatively evaluated by fitting with a 1:1 global kinetics analysis model that included a mass transport limitation term (Materials and Methods).

Table 1: Binding and Kinetic Results for Binding of RT29 to AATT and AATTAA Sites at 25 °C

| DNA | $[Na^+]$ (mM) | $k_a$ ($M^{-1}$ $s^{-1}$) | $k_d$ ($s^{-1}$) | $K$ ($M^{-1}$) |
|---|---|---|---|---|
| AATT | 200 | MT[a] | MT[a] | $1.4 \times 10^9$ |
| | 500 | MT[a] | MT[a] | $2.7 \times 10^8$ |
| | 1000 | $6.88 \times 10^6$ | 0.072 | $9.6 \times 10^7$ |
| AATTAA | 200 | MT[a] | MT[a] | $1.6 \times 10^9$ |
| | 500 | MT[a] | MT[a] | $2.5 \times 10^8$ |
| | 1000 | $2.35 \times 10^6$ | 0.026 | $9.2 \times 10^7$ |

[a] Mass transfer-dominated binding rate.

The results clearly show that the rate of binding of RT29 to DNA under these conditions is dominated by the transport of RT29 molecules between the flow solution and the sensor surface with immobilized DNA. Given the rapid association of RT29 with the minor groove in AT sequences and the very low concentrations used in the experiments due to the strong binding of RT29, this is an expected observation (*47*). The apparent kinetics results also increase with an increase in the flow rate, a signature characteristic of mass transport-influenced rates. The calculated $k_aRU_{max}/k_t$ ratios are higher than 5 under these conditions, providing further evidence for mass transport-limited binding. A ratio of less than 5 has been shown to be important for obtaining reliable kinetic data (*38*). True reaction kinetics results could thus not be determined for the RT29−DNA complex under these conditions (0.2 and 0.5 M $Na^+$ at 25 °C); however, the equilibrium binding constant at 0.2 M salt [$K = 1.4 \times 10^9$ $M^{-1}$ (experimental error of ±20%)] obtained from the global fit approach is valid (*38*, *47*), as long as most of the observed association and dissociation rate data can be collected in the experimental time period. The fitting procedure allows the determination of equilibrium constants since mass transfer affects both the association and dissociation reactions and $K$ can thus be determined from mass transfer-limited apparent rate constants. The binding constant obtained from more limited steady-state results at the highest concentrations is in good agreement with the $K$ from kinetics analysis. The binding constant obtained for RT29 is higher than those of other minor groove binding dications of similar size and charge.

To better understand the energetic details of the RT29−DNA complex, binding studies were conducted as a function of salt concentration and temperature. A slope of 1.8 was obtained in a plot of log $K$ versus $-\log[Na^+]$, and this indicates that both cationic amidines of RT29 are electrostatically interacting with the DNA duplex (Figure 5B). $\Delta G$ for binding changes very little over the experimental temperature range (Figure 6), in agreement with observations on other minor groove binding agents (*27*, *29*, *48*). In summary, the SPR results show that RT29, in spite of its unusual structure, binds very tightly to AT sequences of four or more AT base pairs in a 1:1 complex, but that it binds very weakly to sites composed of GC base pairs. For its size, RT29 is one of the strongest DNA binding agents found to date.

As described above, the kinetics of the interaction of RT29 with DNA at salt concentrations of up to 0.5 M NaCl have significant mass transport influence. Kinetics of binding of RT29 to the AATT site at 1 M NaCl with a lower density of DNA on a CM4 chip are not significantly mass transport limited [Table 1 (*38*)]. As the $K$ value decreases at higher

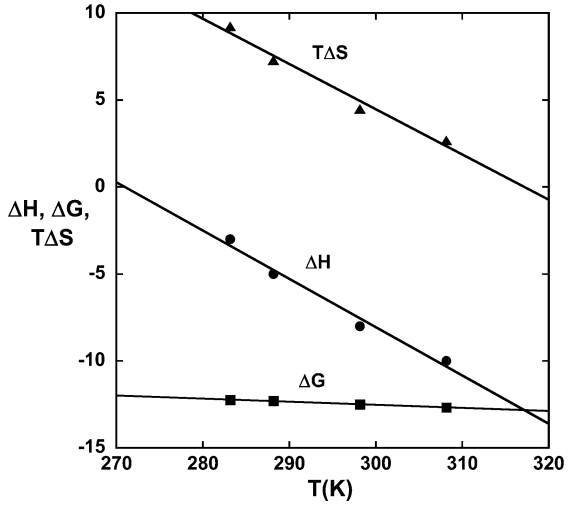


FIGURE 6: Thermodynamic results, $\Delta G$ from SPR, $\Delta H$ from ITC, and $-T\Delta S$ (calculated from $\Delta G = \Delta H - T\Delta S$), for binding of RT29 to the AATT site (Figure 1) at different temperatures.

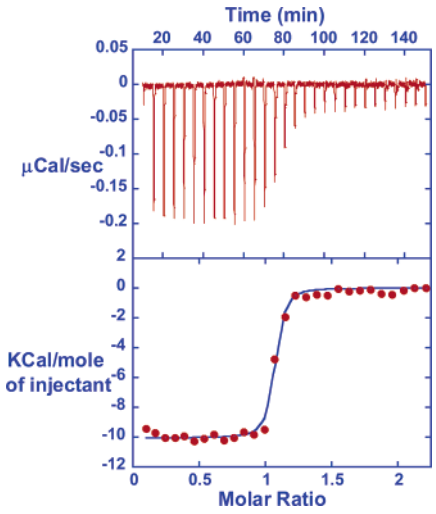


FIGURE 7: ITC curve (top) for titration of RT29 (0.050 mM) into 0.005 mM d(GCGAATTCGC)$_2$ duplex (CAC20 buffer with 200 mM NaCl) at 35 °C. The integrated, blank-subtracted heats are plotted vs molar ratio in the bottom panel.

salt concentrations, the association ($k_a$) rate constant decreases and the mass transfer influence decreases to zero on lower-density DNA surfaces. The $k_a$ and dissociation ($k_d$) rate constants as well as the $K$ value (Table 1 and Figure 5C) obtained from global fits are consistent with non-mass-transport-limited binding kinetics at 1 M salt. Similar kinetics experiments were conducted with the AATTAA sequence (Figure S1 of the Supporting Information and Table 1), and the results are similar to those obtained with the AATT site.

*Isothermal Titration Calorimetry (ITC): Detailed Thermodynamics of the RT29−DNA Interaction.* ITC experiments with an AATT sequence oligomer were conducted to obtain a full thermodynamic understanding of the exceptionally strong binding of RT29 (*27*, *29*, *48*−*50*). Titration of RT29 into buffer and into a DNA solution provided plots of heat per mole versus molar ratio by subtraction of the integrated peak areas for the buffer titration from the DNA interaction titration (an example titration is shown in Figure 7). Due to the large $K$ value for binding of RT29 to DNA, only the binding enthalpy, $\Delta H$, could be obtained in the ITC experiments. For molar ratios of less than approximately 0.9, the observed heat per mole is constant and equal to $\Delta H$ since essentially all of the added compound binds to DNA in the calorimeter under these conditions. The results were fit to a single-site binding model, and the observed $\Delta H$ value for binding is $-8.0 \pm 0.1$ kcal/mol of RT29 at 25 °C. The calculated $-T\Delta S$ values (−4.5 kcal/mol at 25 °C) are based on SPR binding free energies ($\Delta G° = -RT \ln K = -12.5$ kcal/mol at 25 °C) and $\Delta H$ from ITC (Figures 6 and 7). These results clearly show that the strong binding of RT29 at ≥25 °C is dominated by a favorable binding enthalpy. This is in contrast to many diamidine derivatives and other minor groove binding agents for which the binding to AT sequences is entropy-driven (*25*, *48*, *50*). ITC and SPR experiments were also conducted with the -AATTAA- sequence DNA used in crystallization (Figure S2 of the Supporting Information). The observed binding enthalpy is −7.4 kcal/mol, and the calculated $-T\Delta S$ is −5.3 kcal/mol at 25 °C. Thus, the interaction of RT29 with both AT DNAs is principally driven by the interaction enthalpy term.

Footprinting results described above as well as previous fluorescence intercalator displacement (FID) studies (*30*) have shown that RT29 can bind to DNA sequences with three consecutive AT base pairs and a single GC base pair. To compare this type of interaction with that for pure AT sites, ITC experiments with a hairpin oligomer sequence, 5′-CG**ATTC**CGTCTCC**GAAT**CG-3′, with an -ATTC- binding site (in bold with the hairpin loop sequence underlined) which was previously used in SPR studies (*30*), were conducted (Figure S3 of the Supporting Information). These studies provide a thermodynamic understanding of the differences in interactions of RT29 with AT sites containing three or four AT base pairs. The experiments with -ATTC- were conducted as with -AATT-, and the observed $\Delta H$ value for RT29 binding at 25 °C is $-6.2 \pm 0.07$ kcal/mol. The calculated $-T\Delta S$ value at 25 °C for binding to -ATTC-, based on the SPR binding free energies ($\Delta G° = -RT \ln K = -10.4$ kcal/mol) (*2*) and $\Delta H$ from ITC, is −4.2 kcal/mol. These results again show that the binding of RT29 to DNA with three or four AT base pairs is dominated by the enthalpy term at 25 °C. The stronger binding to AATT, relative to that to ATTC, is entirely due to a more favorable binding enthalpy with the AATT site (−8 kcal/mol vs −6.2 kcal/mol).

*Molecular Structural Comparisons and Models: Calculations and Crystallography.* A geometry-optimized structure for RT29 is shown in Figure 8 along with a model of RT29 with the coordinates from the crystal complex with DNA (*30*). An optimized structure for DB75 is also shown to allow comparison with a compound that has a more standard minor groove binding shape (*17*). The structures were calculated by quantum methods at both the HF and B3LYP levels (Materials and Methods) with different basis sets. The basis set level that was used in both cases had a small effect on the final torsional angles in RT29. The structure calculated by DFT methods has predicted torsional angles for all angles joining the benzimidazole, phenyl, and amidine systems slightly smaller than those found in the HF calculations. In the calculated structures for both RT29 and DB75, the amidine−aromatic ring torsional angles are 35−40°. The benzimidazole−amidine angle in the RT29 crystal structure is similar, but the phenyl−amidine angle is reduced to near 0° (Figure 8). The benzimidazole−phenyl angle in both the

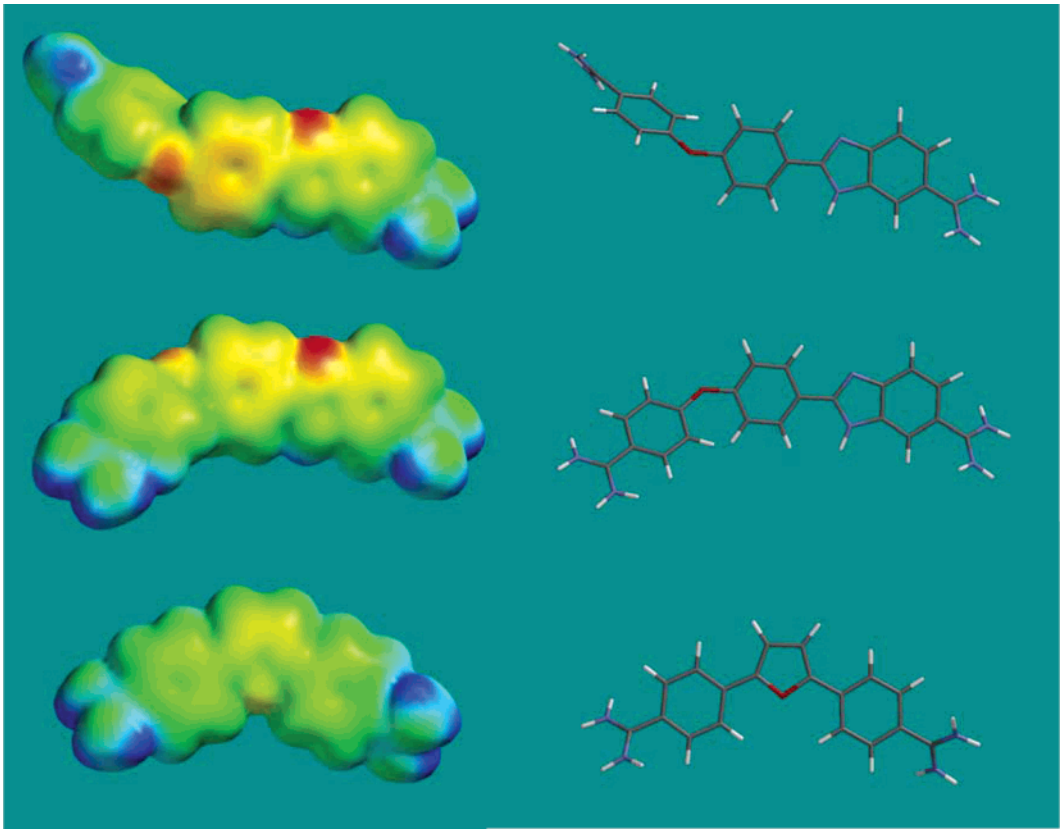

FIGURE 8: Equilibrium geometries of RT29 (top) and DB75 (bottom) were calculated by the density functional method at the 631G(p,d) level. Tube models are shown at the right with space filling electrostatic potential molecular surfaces at the left. The middle structure is RT29 in the conformation observed in the crystal complex. The crystal conformation was frozen in the calculations for comparison to the unconstrained minimum energy conformations of RT29 and DB75. As can be seen, there is a very significant difference among the equilibrium, unbound, and bound conformations of RT29. Note that the bound conformation of RT29 is much closer to that for DB75 and other similar minor groove binding compounds.

calculated and crystal structures of RT29 is near 0°. These results indicate that the amidine−benzimidazole−phenyl end of RT29 does not undergo significant structural changes upon formation of the -AATT- minor groove complex. The amidine−phenyl−O-phenyl end of RT29, on the other hand, undergoes considerable structural change upon complex formation.

There is a dramatic difference in structure between DB75 and RT29 that can also be seen from the comparisons in Figure 8. The aromatic system of DB75 has the classical curvature of DNA minor groove binding compounds with a relatively planar phenyl−furan−phenyl aromatic system, while the diphenylether of RT29 is highly curved and twisted. The twist is especially significant relative to the width of the minor groove, and in the equilibrium conformation, the compound cannot slide deeply into the groove. The X-ray results, however, show that the angle between the phenyl planes of 62° in the calculated, unbound structure is reduced to 35° in the bound form (Figures 8 and 9). It is this decrease in twist that makes it possible for RT29 to slide into the minor groove. The reduction in twist is largely made possible by an increase in the phenyl−O−phenyl bond angle. The bond angle is 126° in the unbound molecule but is 136° in the complex with DNA. This slight opening of the bond angle makes it possible for the reduction in phenyl−O−phenyl twist to occur without extensive atom−atom steric clash.

The induced structure of RT29 that is bound to DNA also has an impressively optimized electrostatic potential molecular surface for DNA interactions (Figure 8). The inner face

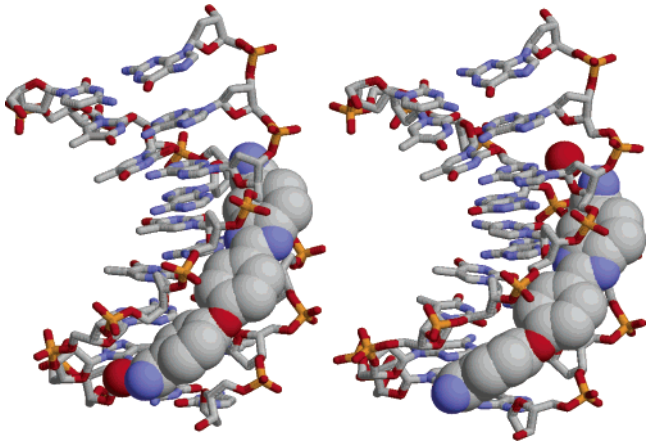

FIGURE 9: Crystal structures for RT29 bound in the minor groove of an AATT (left) and ATTC site (right). The water molecules that complete the complexes are shown as red spheres. The water is at the bottom, phenylamidine group, in the AATT complex and at the top, benzimidazole amidine group, in the ATTC complex.

of the molecule in the bound conformation contains the benzimidazole and amidine NH groups with a significant positive potential. The opposite face with the unprotonated benzimidazole N and the ether oxygen, which faces out into the solvent, is significantly more negatively charged and does not have H-bond donating groups. This is an advantage for both formation of H-bonds with the AT base pairs at the floor of the DNA minor groove and electrostatic interactions. The interaction of RT29 with DNA is thus similar to the

Table 2: Solvent Accessible Surface Areas of d(CTTAATTC)$_2$ and RT29 Calculated Using NACCESS[a]

| | nonpolar area (Å$^2$) | polar area (Å$^2$) | total area (Å$^2$) |
|---|---|---|---|
| complex | 1403.7 | 1715.4 | 3119.1 |
| bound DNA | 1284.4 | 1629.6 | 2914.0 |
| bound RT29 | 119.3 | 85.8 | 205.1 |
| free DNA | 1504.9 | 1734.6 | 3239.5 |
| free RT29 | 416.6 | 246.7 | 663.3 |

[a] $\Delta SASA_{total} = -783.7$ Å$^2$.

folding−binding reaction that many proteins undergo upon formation of a complex with DNA, and the induced conformational change places the functional groups of RT29 in closely optimized positions for interaction with DNA (Figure 9).

Various methods have been proposed to calculate binding heat capacities for nucleic acid complexes based on the amount of polar and nonpolar surface area that is shielded from solvent upon complex formation. From the crystal structure of RT29 bound at the AATT site, the amount of buried polar and nonpolar surface area can be determined, and the values are listed in Table 2. Record and co-workers (*51*), on the basis of protein folding and small molecule transfer to water, and Chaires and co-workers (*52*), on the basis of small molecules binding to DNA, have developed equations for correlating buried surface areas with $\Delta C_p$. The Record equation predicts a heat capacity of $-128$ cal mol$^{-1}$ K$^{-1}$, and the Chaires equation predicts a $\Delta C_p$ of $-165$ cal mol$^{-1}$ K$^{-1}$ for binding of RT29 to the AATT site. Both predictions are substantially below the experimental value of $-278$ cal mol$^{-1}$ deg$^{-1}$ for the AATT complex that is determined from the slope of $\Delta H$ versus temperature in Figure 6.

## DISCUSSION

It is important to study the DNA complexes of compounds of unusual structure, such as RT29, in detail to better understand the molecular basis for their unexpected ability to recognize DNA. This is especially true for RT29 since it has exceptional DNA binding strength, in addition to a nonstandard shape for DNA recognition, and can serve as the basis for design of a new class of nonclassical minor groove binders. DNase I footprinting is the method of choice for defining sequence specificity in DNA binding of compounds with binding sites of different length in a long sequence context (*11*, *32*, *53*). Several important conclusions arise from the footprinting results (Figure 2). (i) The footprints seen with RT29 always coincide with the position of AT-rich sequence 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 (*53*, *54*). (iii) RT29 protects AT-rich sequences from cleavage by DNase I more intensely at AATT sites, indicating that it exhibits a higher affinity for this DNA sequence. (iv) A less intense footprint is seen at a site with three AT base pairs and a terminal GC base pair. These results are complemented and expanded by results from FID experiments which show that AATT is a favored RT29 binding site in four base pair binding sequences (Figure 3). Binding sites that contain TA steps interact with RT29 significantly less strongly than other AT sequences in both footprinting and FID experiments. Similar poor binding to sequences with TA steps has been reported with other minor groove binding agents (*55*−*57*). Both DNase I footprinting and FID as well as SPR studies show that four base pair sites with more than one GC pair bind RT29 too weakly to be detected. DNase I is a large enzyme, and this makes it somewhat difficult to precisely define the exact drug binding site in a footprinting study. In several footprints (Figure 2), a G·C base pair at the 3′-side is protected, and this could be correlated with the FID results that support RT29 binding at (A·T)$_3$-G·C sites. The correlation is complicated, however, by the 3′-offset of minor groove footprints and by the size of DNase I.

X-ray structural results in Figure 9 (*30*) provide a model for RT29 binding in the minor groove, and circular dichroism spectra provide support for a minor groove complex in solution (Figure 4). The strong minor groove binding of RT29 to AATT was unexpected on the basis of the twist and curvature of the compound that is unique for a minor groove binding agent of this size. Similar compounds that are highly curved or twisted bind weakly to DNA (*16*, *58*). To better understand the basis for the strong binding of RT29, crystallographic (*30*) and calculated structures are compared for RT29 in Figure 8. There are significant induced fit changes in the conformation of RT29 upon complex formation. Even with these changes, however, it is not possible to arrange all of the functional groups, the amidine and benzimidazole NH groups, for example, to match the shape of the minor groove and position of DNA receptor groups. To complete the complex requires a bound water molecule that forms an essential part of a ternary complex (Figure 9). Interestingly, the benzimidazole NH group and the amidine group of RT29 form direct H-bonds with DNA in the AATT site, while the phenylamidine interacts with DNA through a bound water molecule. In the crystal structure with the compound bound at the ATTC sequence, the water position shifts to the benzimidazole end of RT29 and the benzimidazole is pushed slightly off of the floor of the minor groove due to the GC base pair and the G-NH$_2$ group in the minor groove in the benzimidazole binding region. As a result, there are no direct H-bonds from RT29 to the DNA bases in the ATTC complex, and only indirect H-bonds are formed through the bound water molecule to the bases edges at the floor of the minor groove. The lack of H-bonding certainly helps explain the weaker binding of RT29 to the ATTC sequence. Bound water molecules thus provide a flexible way of linking RT29, and other diamidines, to the floor of the minor groove with a significant increase in the Gibbs energy of binding. Interestingly, there are also a number of close contacts between phenyl CH groups of RT29 and DNA TO2 keto groups in the minor groove that apparently significantly stabilize the complex (Figure 9).

These interesting and unexpected structural observations raise specific questions that we have addressed in the thermodynamics studies reported here. What is the thermodynamic basis for the exceptionally strong binding of RT29 to AATT sites, and how do the thermodynamic driving components for complex formation change when RT29 is bound to a site with only three AT base pairs (ATTC)? The answer to the first question is summarized in the thermodynamic results as a function of temperature (Figure 6). As with all minor groove dications studied in detail to this time,

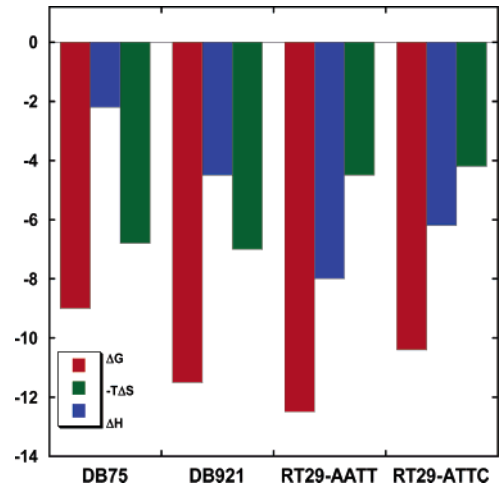


FIGURE 10: Comparison of binding thermodynamics for DB75, DB921, and RT29 with an AATT site and for RT29 with an ATTC site at 25 °C in buffer with 0.2 M NaCl.

the Gibbs energy of binding changes very little with temperature near 25 °C while the calorimetric enthalpy of binding becomes significantly more negative, more favorable for complex formation, as the temperature is increased. The calculated $-T\Delta S$ value for binding, thus, also becomes more negative, less favorable, as the temperature is increased. Formation of a complex of RT29 with the AATT site is strongly enthalpy driven at 25 °C, and the plot in Figure 6 predicts that the binding reaction becomes completely enthalpy driven above 45 °C but completely entropy driven below approximately 0 °C.

The heat capacity difference between the DNA complex of RT29 and the free species is obtained from the slope of the plot of $\Delta H$ versus temperature (Figure 6) and is negative ($-278$ cal mol$^{-1}$ deg$^{-1}$ for the AATT complex). From the crystal structure of RT29 bound at the AATT site, the amount of buried polar and nonpolar surface area can be calculated (Table 2) and used to estimate $\Delta C_p$ (*51*, *52*). Predicted values (Table 2) are substantially below the experimental value, and this suggests that bound water, particularly the water that is an intrinsic part of the complex, adds significantly to the change in heat capacity upon complex formation. Cooper and co-workers (*59*, *60*) have clearly shown that many biopolymer transitions and binding interactions have detailed thermodynamics with the characteristics described above (Figure 6) and that such characteristic thermodynamic values are expected for processes involving numerous weak interactions, particularly those involving changes in bound water, in the overall reaction. They conducted an experimental and theoretical analysis of trisaccharide−protein binding with and without a specifically bound water molecule. Both experiment and theory indicated that for each extra H-bond in the system with water, the binding enthalpy would be approximately $-2$ kcal/mol larger than for the trisaccharide system without water. The binding heat capacity was found to be approximately $-20$ cal mol$^{-1}$ K$^{-1}$ larger for the water system, and this was within the range of theoretical predictions. An RT29 analogue without bound water is not available, but we can compare the results to those of other diamidines that bind without a water.

Figure 10 shows a comparison of the binding thermodynamics for the RT29−AATT complex with results for other dicationic minor groove binding agents with the same sequence. DB75, shown in Figures 1 and 8, is a classically shaped minor groove binding agent in which both amidines form H-bonds with DNA bases (*61*). No water molecule forms an intrinsic part of the interaction between the compound and DNA bases, and complex formation is strongly entropy driven at 25 °C. Although there are many interactions in the RT29 complex that are not possible with the DB75 heterocyclic system, the differences in binding enthalpy and heat capacity with the theoretical results described above (*59*, *60*) suggest that at least two water H-bonds must be present in the RT29 complex that are not found in the DB75 complex.

Thermodynamic results for DB921 (Figure 1), an AT specific minor groove agent that has groups similar to those of RT29 but with a less twisted and more linear shape, are also shown in Figure 10. DB921 also binds very strongly to DNA and requires a bound water molecule at the phenylamidine end of the molecule to H-bond with DNA bases at that end of the molecule (*25*). Interestingly, it has a thermodynamic profile different from that of RT29 with a larger entropy component for binding to AATT sites at 25 °C. These results suggest that RT29 has a more ordered water network system to stabilize its DNA complex than DB921. Such a difference is not apparent in the X-ray structures of the two compound−DNA complexes. What is quite clear is that these two benzimidazole derivatives, with a very nonstandard shape for minor groove complex formation, have some of the largest binding constants for molecules of their size. Results for binding of RT29 to the ATTC sequence are also shown in Figure 10, and the reduction in binding enthalpy and Gibbs energy are clearly seen in comparison to the AATT results. The lower enthalpy is in agreement with the reduced number of H-bonds in the ATTC complex observed in the crystal structure.

For therapeutic reasons, it is essential to design additional compounds that target DNA sequences in different but effective ways. RT29 represents a new paradigm that does not fit the classical model of minor groove interactions. An induced fit molecular structural change in RT29 reduces the twist of the ether group to complement the minor groove and places the functional groups in position to interact with DNA bases at the floor of the minor groove. The interaction is completed by incorporating a water molecule into the complex, and in the AATT binding site, the phenylamidine contacts the bases at the floor of the groove indirectly through a bound molecule of water. The RT29−water pair forms a curved noncovalent minor groove binding module that matches the shape of the groove. With the ATTC site, the water−RT29 interaction again creates a curved minor groove binding adduct that is better able to match the curvature of the groove. These results, along with previous analyses of linear diamidines that bind strongly in the minor groove [Figure 1 (*25*)], clearly suggest that traditional views of the compound shape required for minor groove complex formation must be reevaluated.

## SUPPORTING INFORMATION AVAILABLE

SPR data for binding of RT29 to AATTAA sequence and ITC data for binding of RT29 to AATTAA and ATTC sequences (Figures S1−S3). This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES


1. Tidwell, R. R., and Boykin, D. W. (2003) Dicationic DNA minor-groove binders as antimicrobial agents, in *DNA and RNA Binders: From Small Molecules to Drugs* (Demeunynck, M., Bailly, C., and Wilson, W. D., Eds.) pp 414−460, Wiley-VCH, Weinheim, Germany.
2. Mathis, A. M., Holman, J. L., Sturk, L. M., Ismail, M. A., Boykin, D. W., Tidwell, R. R., and Hall, J. E. (2006) Accumulation and intracellular distribution of antitrypanosomal diamidine compounds DB75 and DB820 in African trypanosomes, *Antimicrob. Agents Chemother. 50*, 2185−2191.
3. Wilson, W. D., Nguyen, B., Tanious, F. A., Mathis, A., Hall, J. E., Stephens, C. E., and Boykin, D. W. (2005) Dications that target the DNA minor groove: Compound design and preparation, DNA interactions, cellular distribution and biological activity, *Curr. Med. Chem. Anti-Cancer Agents 5*, 389−408.
4. Neidle, S. (2001) DNA minor-groove recognition by small molecules, *Nat. Prod. Rep. 18*, 291−309.
5. Dervan, P. B., and Edelson, B. S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides, *Curr. Opin. Struct. Biol. 13*, 284−299.
6. Lacy, E. R., Madsen, E. M., Lee, M., and Wilson, W. D. (2003) Polyamide Dimer Stacking in the DNA Minor Groove and Recognition of T•G Mismatched Base Pairs in DNA, in *DNA and RNA Binders: From Small Molecules to Drugs* (Demeunynck, M., Bailly, C., and Wilson, W. D., Eds.) pp 384−413, Wiley-VCH, Weinheim, Germany.
7. Shapiro, T. A., and Englund, P. T. (1990) Selective cleavage of kinetoplast DNA minicircles promoted by antitrypanosomal drugs, *Proc. Natl. Acad. Sci. U.S.A. 87*, 950−954.
8. Woynarowski, J. M. (2002) Targeting critical regions in genomic DNA with AT-specific anticancer drugs, *Biochim. Biophys. Acta 1587*, 300−308.
9. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., and Dickerson, R. E. (1985) The molecular origin of DNA-drug specificity in netropsin and distamycin, *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376−1380.
10. Goodwin, K. D., Long, E. C., and Georgiadis, M. M. (2005) A host-guest approach for determining drug-DNA interactions: An example using netropsin, *Nucleic Acids Res. 33*, 4106−4116.
11. Bailly, C., and Waring, M. J. (1998) The use of diaminopurine to investigate structural properties of nucleic acids and molecular recognition between ligands and DNA, *Nucleic Acids Res. 26*, 4309−4314.
12. Hannah, K. C., and Armitage, B. A. (2004) DNA-templated assembly of helical cyanine dye aggregates: A supramolecular chain polymerization, *Acc. Chem. Res. 37*, 845−853.
13. Wemmer, D. E. (2000) Designed sequence-specific minor groove ligands, *Annu. Rev. Biophys. Biomol. Struct. 29*, 439−461.
14. Joubert, A., Sun, X. W., Johansson, E., Bailly, C., Mann, J., and Neidle, S. (2003) Sequence-selective targeting of long stretches of the DNA minor groove by a novel dimeric bis-benzimidazole, *Biochemistry 42*, 5984−5992.
15. Shapiro, T. A., and Englund, P. T. (1995) The structure and replication of kinetoplast DNA, *Annu. Rev. Microbiol. 49*, 117−143.
16. Fairley, T. A., Tidwell, R. R., Donkor, I., Naiman, N. A., Ohemeng, K. A., Lombardy, R. J., Bentley, J. A., and Cory, M. (1993) Structure, DNA minor groove binding, and base pair specificity of alkyl- and aryl-linked bis(amidinobenzimidazoles) and bis(amidinoindoles), *J. Med. Chem. 36*, 1746−1753.
17. Nguyen, B., Boykin, D. W., and Wilson, W. D. (2007) DNA Minor Groove Interactions of Antiparasitic Diamidines: Reevaluation of The Crescent-Shape Concept in Groove-Binding, in *Synthetic and Biophysical Studies of DNA Binding Compounds* (Lee, M., and Strekowski, L. Eds.) Research Signpost.
18. Luscombe, N. M., Laskowski, R. A., and Thornton, J. M. (2001) Amino acid-base interactions: A three-dimensional analysis of protein-DNA interactions at an atomic level, *Nucleic Acids Res. 29*, 2860−2874.
19. Bergqvist, S., Williams, M. A., O'Brien, R., and Ladbury, J. E. (2004) Heat capacity effects of water molecules and ions at a protein-DNA interface, *J. Mol. Biol. 336*, 829−842.
20. Ismail, M. A., Batista-Parra, A., Miao, Y., Wilson, W. D., Wenzler, T., Brun, R., and Boykin, D. W. (2005) Dicationic near-linear biphenyl benzimidazole derivatives as DNA-targeted antiprotozoal agents, *Bioorg. Med. Chem. 13*, 6718−6726.
21. Athri, P., Wenzler, T., Ruiz, P., Brun, R., Boykin, D. W., Tidwell, R., and Wilson, W. D. (2006) 3D QSAR on a library of heterocyclic diamidine derivatives with antiparasitic activity, *Bioorg. Med. Chem. 14*, 3144−3152.
22. Nguyen, B., Stanek, J., and Wilson, W. D. (2006) Binding-linked protonation of a DNA minor-groove agent, *Biophys. J. 90*, 1319−1328.
23. Nguyen, B., Lee, M. P., Hamelberg, D., Joubert, A., Bailly, C., Brun, R., Neidle, S., and Wilson, W. D. (2002) Strong binding in the DNA minor groove by an aromatic diamidine with a shape that does not match the curvature of the groove, *J. Am. Chem. Soc. 124*, 13680−13681.
24. Nguyen, B., Hamelberg, D., Bailly, C., Colson, P., Stanek, J., Brun, R., Neidle, S., and Wilson, W. D. (2004) Characterization of a novel DNA minor-groove complex, *Biophys. J. 86*, 1028−1041.
25. Miao, Y., Lee, M. P., Parkinson, G. N., Batista-Parra, A., Ismail, M. A., Neidle, S., Boykin, D. W., and Wilson, W. D. (2005) Out-of-shape DNA minor groove binders: Induced fit interactions of heterocyclic dications with the DNA minor groove, *Biochemistry 44*, 14701−14708.
26. Tanious, F. A., Hamelberg, D., Bailly, C., Czarny, A., Boykin, D. W., and Wilson, W. D. (2004) DNA sequence dependent monomer-dimer binding modulation of asymmetric benzimidazole derivatives, *J. Am. Chem. Soc. 126*, 143−153.
27. Haq, I., Ladbury, J. E., Chowdhry, B. Z., Jenkins, T. C., and Chaires, J. B. (1997) Specific binding of Hoechst 33258 to the $d(CGCAAATTTGCG)_2$ duplex: Calorimetric and spectroscopic studies, *J. Mol. Biol. 271*, 244−257.
28. Satz, A. L., and Bruice, T. C. (2002) Recognition in the minor groove of double-stranded DNA by microgonotropens, *Acc. Chem. Res. 35*, 86−95.
29. Haq, I. (2002) Thermodynamics of drug-DNA interactions, *Arch. Biochem. Biophys. 403*, 1−15.
30. Goodwin, K. D., Lewis, M. A., Tanious, F. A., Tidwell, R. R., Wilson, W. D., Georgiadis, M. M., and Long, E. C. (2006) A high-throughput, high-resolution strategy for the study of site-selective DNA binding agents: Analysis of a "highly twisted" benzimidazole-diamidine, *J. Am. Chem. Soc. 128*, 7846−7854.
31. Tidwell, R. R., Geratz, J. D., Dann, O., Volz, G., Zeh, D., and Loewe, H. (1978) Diarylamidine derivatives with one or both of the aryl moieties consisting of an indole or indole-like ring. Inhibitors of arginine-specific esteroproteases, *J. Med. Chem. 21*, 613−623.
32. Bailly, C., and Waring, M. J. (1995) Comparison of different footprinting methodologies for detecting binding sites for a small ligand on DNA, *J. Biomol. Struct. Dyn. 12*, 869−898.
33. Tse, W. C., and Boger, D. L. (2004) A fluorescent intercalator displacement assay for establishing DNA binding selectivity and affinity, *Acc. Chem. Res. 37*, 61−69.
34. Boger, D. L., Fink, B. E., Brunette, S. R., Tse, W. C., and Hedrick, M. P. (2001) A simple, high-resolution method for establishing DNA binding affinity and sequence selectivity, *J. Am. Chem. Soc. 123*, 5878−5891.
35. Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: Estimation of static accessibility, *J. Mol. Biol. 55*, 379−400.
36. Hubbard, S. J., and Thornton, J. M. (1993) *NACCESS*, Department of Biochemistry and Molecular Biology, University College London, London.
37. Myszka, D. G., He, X., Dembo, M., Morton, T. A., and Goldstein, B. (1998) Extending the range of rate constants available from BIACORE: Interpreting mass transport-influenced binding data, *Biophys. J. 75*, 583−594.
38. Karlsson, R. (1999) Affinity analysis of non-steady-state data obtained under mass transport limited conditions using BIAcore technology, *J. Mol. Recognit. 12*, 285−292.
39. Lewis, M. A., and Long, E. C. (2006) Fluorescent intercalator displacement analyses of DNA binding by the peptide-derived natural products netropsin, actinomycin, and bleomycin, *Bioorg. Med. Chem. 14*, 3481−3490.
40. Rodger, A., and Nordén, B. (1997) *Circular Dichroism and Linear Dichroism*, Oxford University Press, New York.
41. Lyng, R., Rodger, A., and Norden, B. (1992) The CD of ligand-DNA systems. 2. Poly(dA-dT) B-DNA, *Biopolymers 32*, 1201−1214.
42. Wilson, W. D., Tanious, F. A., Fernandez-Saiz, M., and Rigl, C. T. (1997) Evaluation of drug-nucleic acid interactions by thermal melting curves, *Methods Mol. Biol. 90*, 219−240.

43. Myszka, D. G., Abdiche, Y. N., Arisaka, F., Byron, O., Eisenstein, E., Hensley, P., Thomson, J. A., Lombardo, C. R., Schwarz, F., Stafford, W., and Doyle, M. L. (2003) The ABRF-MIRG'02 study: Assembly state, thermodynamic, and kinetic analysis of an enzyme/inhibitor interaction, *J. Biomol. Technol. 14*, 247−269.
44. Day, Y. S., Baird, C. L., Rich, R. L., and Myszka, D. G. (2002) Direct comparison of binding equilibrium, thermodynamic, and rate constants determined by surface- and solution-based biophysical methods, *Protein Sci. 11*, 1017−1025.
45. Day, Y. S., and Myszka, D. G. (2003) Characterizing a drug's primary binding site on albumin, *J. Pharm. Sci. 92*, 333−343.
46. Drake, A. W., Myszka, D. G., and Klakamp, S. L. (2004) Characterizing high-affinity antigen/antibody complexes by kinetic- and equilibrium-based methods, *Anal. Biochem. 328*, 35−43.
47. Karlsson, R., and Larsson, A. (2004) Affinity measurement using surface plasmon resonance, *Methods Mol. Biol. 248*, 389−415.
48. Mazur, S., Tanious, F. A., Ding, D., Kumar, A., Boykin, D. W., Simpson, I. J., Neidle, S., and Wilson, W. D. (2000) A thermodynamic and structural analysis of DNA minor-groove complex formation, *J. Mol. Biol. 300*, 321−337.
49. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter, *Anal. Biochem. 179*, 131−137.
50. Wang, L., Kumar, A., Boykin, D. W., Bailly, C., and Wilson, W. D. (2002) Comparative thermodynamics for monomer and dimer sequence-dependent binding of a heterocyclic dication in the DNA minor groove, *J. Mol. Biol. 317*, 361−374.
51. Spolar, R. S., Livingstone, J. R., and Record, M. T., Jr. (1992) Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water, *Biochemistry 31*, 3947−3955.
52. Ren, J., Jenkins, T. C., and Chaires, J. B. (2000) Energetics of DNA intercalation reactions, *Biochemistry 39*, 8439−8447.
53. Bailly, C., Kluza, J., Martin, C., Ellis, T., and Waring, M. J. (2004) DNase I Footprinting of Small Molecule Binding Sites on DNA, in *Oligonucleotide Synthesis: Methods and Applications* (Herdewijn, P., Ed.) pp 319−342, Humana Press, Totowa, NJ.
54. Weston, S. A., Lahm, A., and Suck, D. (1992) X-ray structure of the DNase I-d(GGTATACC)$_2$ complex at 2.3 Å resolution, *J. Mol. Biol. 226*, 1237−1256.
55. Abu-Daya, A., Brown, P. M., and Fox, K. R. (1995) DNA sequence preferences of several AT-selective minor groove binding ligands, *Nucleic Acids Res. 23*, 3385−3392.
56. Abu-Daya, A., and Fox, K. R. (1997) Interaction of minor groove binding ligands with long AT tracts, *Nucleic Acids Res. 25*, 4962−4969.
57. Fox, K. R., Yan, Y., and Gong, B. (1999) DNA sequence recognition by a novel series of minor groove-binding ligands, *Anticancer Drug Des. 14*, 219−230.
58. Nguyen, B., Tardy, C., Bailly, C., Colson, P., Houssier, C., Kumar, A., Boykin, D. W., and Wilson, W. D. (2002) Influence of compound structure on affinity, sequence selectivity, and mode of binding to DNA for unfused aromatic dications related to furamidine, *Biopolymers 63*, 281−297.
59. Cooper, A., Johnson, C. M., Lakey, J. H., and Nollmann, M. (2001) Heat does not come in different colours: Entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions, *Biophys. Chem. 93*, 215−230.
60. Cooper, A. (2005) Heat capacity effects in protein folding and ligand binding: A re-evaluation of the role of water in biomolecular thermodynamics, *Biophys. Chem. 115*, 89−97.
61. Laughton, C. A., Tanious, F., Nunn, C. M., Boykin, D. W., Wilson, W. D., and Neidle, S. (1996) A crystallographic and spectroscopic study of the complex between d(CGCGAATTCGCG)$_2$ and 2,5-bis(4-guanylphenyl)furan, an analogue of berenil. Structural origins of enhanced DNA-binding affinity, *Biochemistry 35*, 5655−5661.

BI700288G