

Article

Thermodynamic Factors That Drive Sequence-Specific DNA Binding of Designed, Synthetic Minor Groove Binding Agents

Ananya Paul ¹, Abdelbasset A. Farahat ^{1,2}, David W. Boykin ¹ and W. David Wilson ^{1,*}

¹ Department of Chemistry and Center for Diagnostics and Therapeutics, Georgia State University, Atlanta, GA 30303, USA; apaul@gsu.edu (A.P.); a4f2002@mans.edu.eg (A.A.F.); dboykin@gsu.edu (D.W.B.)

² Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Mansoura University, Mansoura 35516, Egypt

* Correspondence: wdw@gsu.edu; Tel.: +1-404-413-5503; Fax: +1-404-413-5505

Abstract: Ken Breslauer began studies on the thermodynamics of small cationic molecules binding in the DNA minor groove over 30 years ago, and the studies reported here are an extension of those ground-breaking reports. The goals of this report are to develop a detailed understanding of the binding thermodynamics of pyridine-based sequence-specific minor groove binders that have different terminal cationic groups. We apply biosensor-surface plasmon resonance and ITC methods to extend the understanding of minor groove binders in two directions: (i) by using designed, heterocyclic dicationic minor groove binders that can incorporate a G•C base pair (bp), with flanking AT base pairs, into their DNA recognition site, and bind to DNA sequences specifically; and (ii) by using a range of flanking AT sequences to better define molecular recognition of the minor groove. A G•C bp in the DNA recognition site causes a generally more negative binding enthalpy than with most previously used pure AT binding sites. The binding is enthalpy-driven at 25 °C and above. The flanking AT sequences also have a large effect on the binding energetics with the -AAAGTTT- site having the strongest affinity. As a result of these studies, we now have a much better understanding of the effects of the DNA sequence and compound structure on the molecular recognition and thermodynamics of minor groove complexes.

Keywords: DNA minor groove binder; mixed base-pair DNA sequences; sequence selectivity; ligand–DNA complex thermodynamic; molecular curvature; heterocyclic diamidine; biosensor; calorimetry



Citation: Paul, A.; Farahat, A.A.; Boykin, D.W.; Wilson, W.D. Thermodynamic Factors That Drive Sequence-Specific DNA Binding of Designed, Synthetic Minor Groove Binding Agents. *Life* **2022**, *12*, 681. <https://doi.org/10.3390/life12050681>

Academic Editors: Tigran Chalikian and Jens Völker

Received: 30 March 2022

Accepted: 29 April 2022

Published: 4 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the 1980s, Ken Breslauer and some excellent coworkers initiated a series of fundamental studies on the thermodynamics of small-molecule, minor groove agents, and intercalators, binding to different DNA sequences [1–7]. Among other techniques, they used batch calorimetry in ground-breaking investigations of the DNA complexes of these compounds. They established methods for these types of studies that have continued to influence thermodynamic analysis of small-molecule-DNA binding to this day, including the work reported here. They introduced concepts and methods such as entropy-enthalpy compensation that masked driving forces when looking at free energy alone [5]. They used structural studies to bring an understanding of the structural and solution properties that influence the thermodynamics of DNA complex formation [4]. In these studies, they began to develop a microscopic understanding of the experimental macroscopic thermodynamic results. They started the research with state-of-the-art batch calorimeters, which, from personal experience with one of my early colleagues, Harry Hopkins, require a thermodynamics artist's touch to bring forth beautiful thermodynamics pictures. They later moved to much-improved titration calorimeters from MicroCal but continued with detailed thermodynamics studies of DNA interactions that now are going on around the world. It is a pleasure to write this paper in honor of Ken's 75 birthday.

The Breslauer and other laboratories began studies with available, classical minor groove binders from polyamides, such as netropsin, to other types of cationic heterocycles, such as berenil and Hoechst dyes, which were all specific for binding to AT DNA sequences (Figure 1A) [1–8]. These are all uniformly concave-shaped compounds that fit snugly into the minor groove in A-tract sequences and have groups that can H-bond with the N3 of dA or O2 of dT at the floor of the groove.

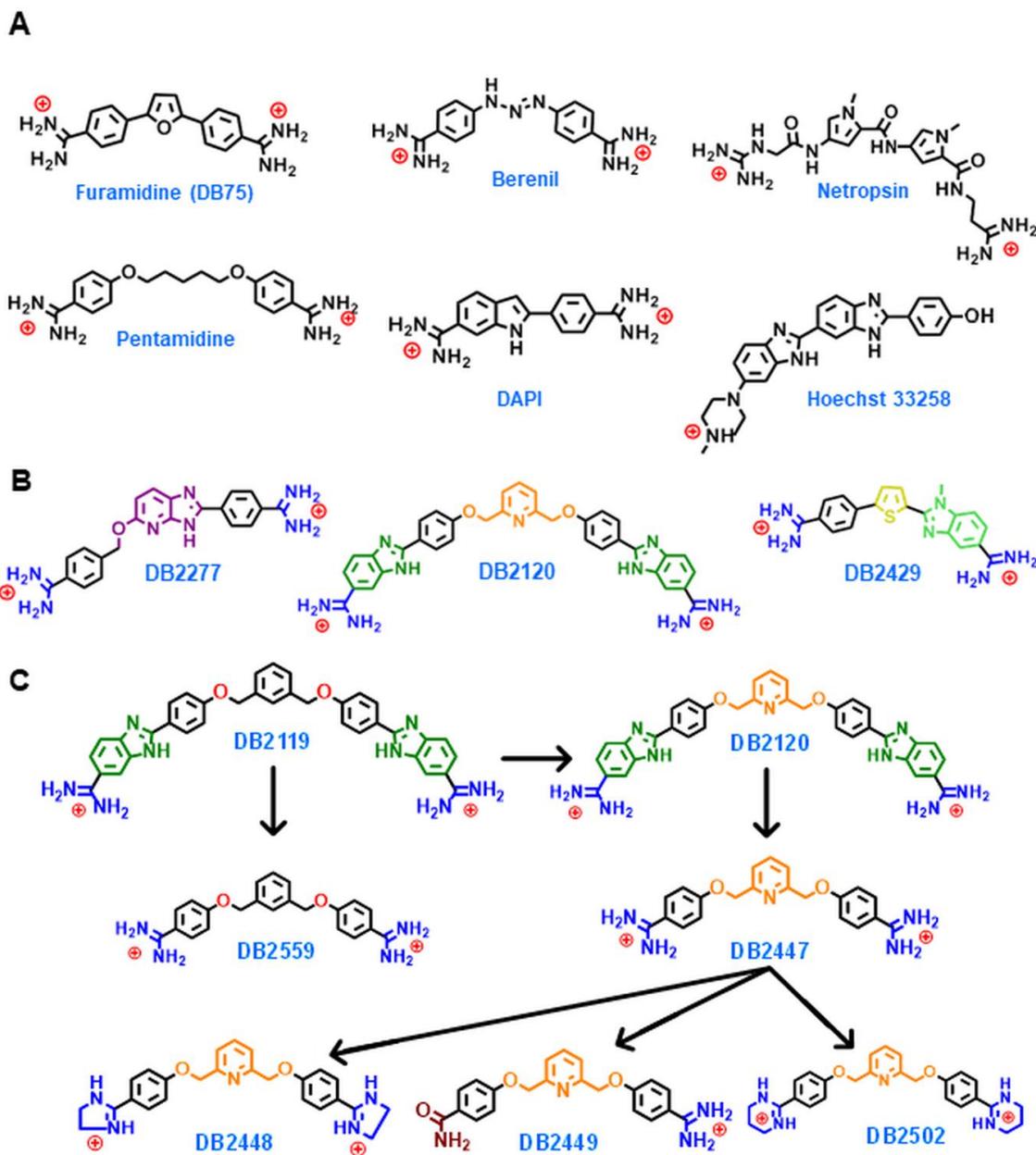


Figure 1. (A) Structures of classical AT-specific DNA minor groove binders; (B) structures of three very different designed mixed-sequence DNA minor groove binders; (C) schematic representation of the development of AT-specific DNA binding compounds from DB2119 to DB2559 and from mixed DNA sequence-specific compounds DB2120 to DB2447 and the analogs for this study. Color schemes used in this figure denote different functional groups.

In addition to forming an H-bond with the cytosine C=O in the minor groove, the 2-amino group of G projects an -N-H into the groove and presents a steric block to classical AT-specific minor groove binders [9]. In spite of the many successes with these early minor groove binders, the lack of a broad selection of sequence-specific compounds was a

limitation for expanded applications. Crystal structures of DNA and complexes suggested that compounds with H-bond acceptor groups could bind to the free G-NH in the minor groove to give GC binding specificity [9–14]. This concept was successful with polyamides but has not been broadly applied to the other minor groove binders such as those listed above (Figure 1).

As heterocyclic amidine minor groove binders such as 4',6-diamidino-2-phenylindole (DAPI) are widely used as cellular nuclear stains and others such as pentamidine, furamidine, and berenil (Figure 1A) have been used or tested in humans and/or animals for therapeutic applications [15–19], we have used the heterocyclic amidine template as the platform for incorporation of single hydrogen bond-accepting groups in generating the first set of compounds with increased specificity. The goal with these first compounds was to recognize a GC base pair in an AT sequence. In this way, it would be possible to create modules that could be combined to give broad DNA sequence recognition. Successful compounds with azabenzimidazole, N-alkylbenzimidazole, and pyridine H-bond acceptor groups for G recognition were created (Figure 1B) [20–25]. Extensive elaborations of the N-alkylbenzimidazole module have significantly improved the affinity and selectivity of that module. With a thiophene adjacent to the imidazole of N-alkylbenzimidazole, a preorganized structure for minor groove recognition was created [25–28].

The original pyridyl-linked amidine-benzimidazole-phenyl compound, DB2120, binds strongly with the single G•C bp-containing -A4GT4- target sequence with a $K_D < 0.1$ nM. DB2120, however, has poor solution properties and aggregation difficulties even at low concentrations under standard experimental conditions for DNA complexes. Because of its size and AT binding benzimidazole units, it also has less selectivity than optimum for use in most applications. To increase the possible uses of the pyridine series of compounds, it is, thus, essential to develop smaller molecules that will have better solution and sequence recognition characteristics. Hence, in the studies reported here, in-depth experiments were conducted with 12 primary DNA hairpin duplex sequences (Figure 2) four of which contain pure AT sequences in their recognition sites (AAAAAA, AAATTT, AATAAT, and ATATAT), four have mixed AT and single GC bp DNA sequences (AAAGAAA, AAAGTTT, AATGAAT, and ATAGTAT), and four have AT sequences with two G•C bps (AAAGCAAA, AAAGCTTT, AATGCAAT, and ATAGCTAT). Three pyridine compounds that are smaller than DB2120, as well as a phenyl compound, were synthesized with the smaller compound design ideas (Figure 1C). The smaller compounds are easier to synthesize, have improved solution properties, and have strong and specific binding to DNA sequences with a single G•C bp. The results with these new pyridine and phenyl compounds are reported here.

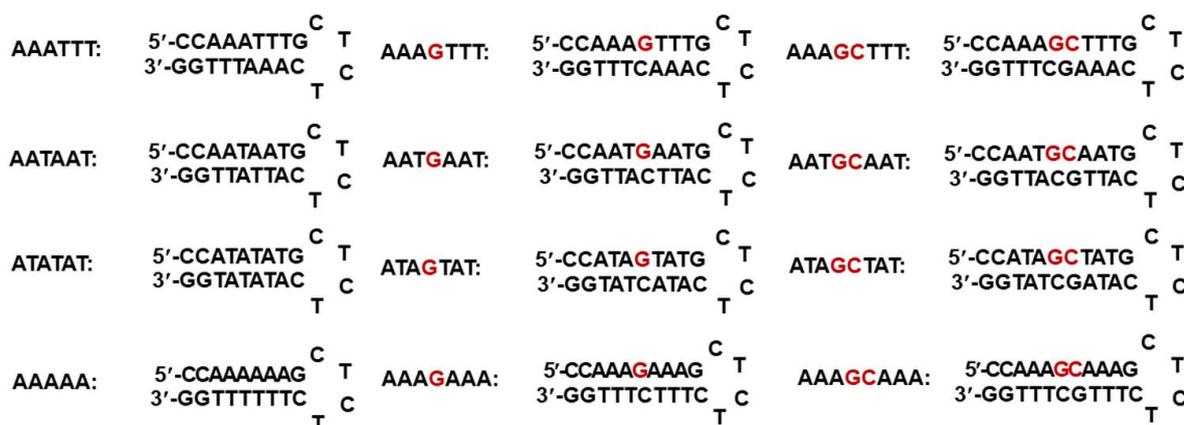


Figure 2. The DNA sequences used in this study; DNA sequences used for Surface Plasmon Resonance (SPR) studies were labeled with 5'-biotin.

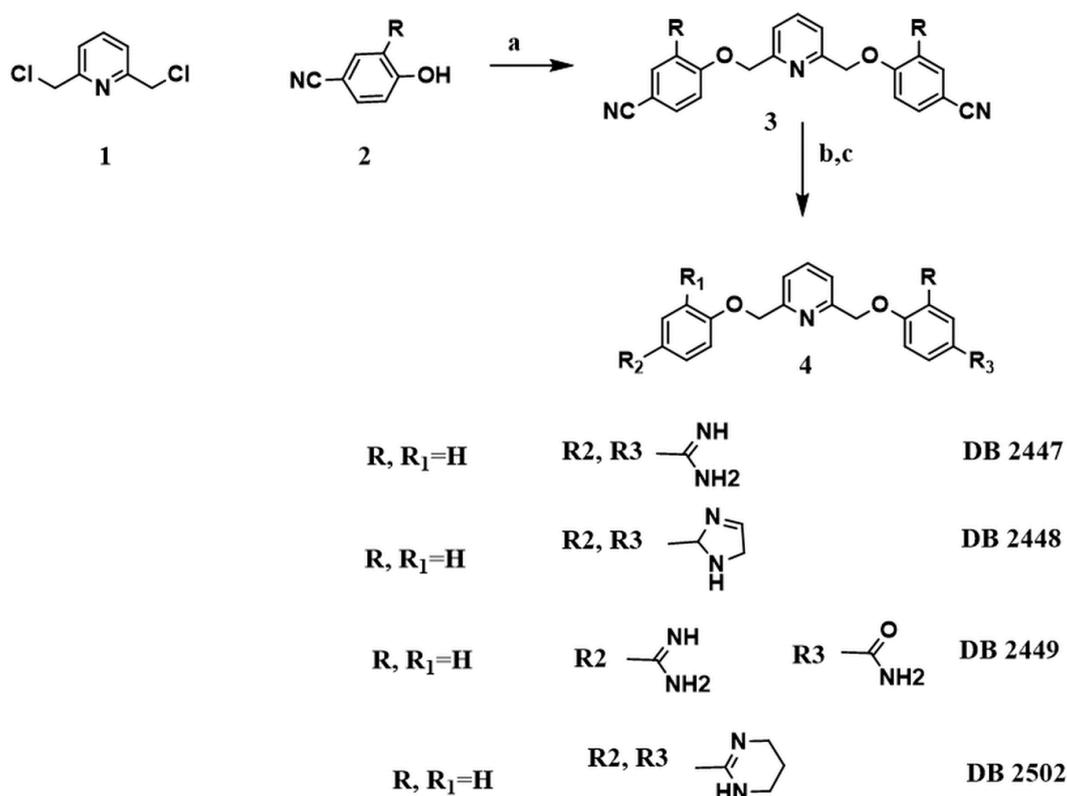
2. Compound Design

The goal with these compounds was to make a series of DB2120 analogs of reduced size but better solution properties and with strong and selective binding to target single G•C bp-containing sequences [23,24]. The compounds would retain the DB2120 core structure but with different terminal cationic groups. Dications with amidine, imidazoline, and tetrahydropyrimidine were successfully prepared along with a monocation with terminal amidine and amide groups (Figure 1C). As a control, an amidine analog with a central phenyl ring that cannot form the H-bond required for GC specificity was also prepared. By using our relative curvature determination method [28] (Results), the compounds were found to have an optimum curvature for the minor groove. Results with these five compounds and the target and control DNA sequences described above (Figure 2) are reported here. For complete studies, target sequences have a single G•C bp in an AT sequence context, while the control sequences have no G•C bp or a two G•C bp insert.

3. Results

3.1. Compound Synthesis

Scheme 1 describes the synthesis of the final amidines **4**. Cyanophenol derivatives **2** were allowed to react with bis(chloromethyl)pyridine **1** in anhydrous dimethylformamide in the presence of potassium carbonate as a base. The formed bisnitriles were converted to the final amidines by applying Pinner reaction conditions [29,30], where the bisnitriles were converted to the intermediate imidate ester hydrochloride by stirring in dry ethanolic HCl. The formed imidate ester was converted to the corresponding amidine by stirring in dry ethanolic ammonia. Finally, the amidines were purified by conversion to the free base using sodium hydroxide and then formation of the hydrochloride salt by stirring in ethanolic HCl. The characterization of the final compounds have described in Experimental Methods.



Scheme 1. Reagents and conditions: (a) K_2CO_3 , DMF, 45 °C, 4 h; (b) ethanol-HCl, 0 °C—r.t.; (c) ethanol, ammonia or amine, r.t or reflux.

3.2. DNA Thermal Melting (ΔT_m): Screening for Relative Affinity and Sequence Selectivity for Target and Control DNA Binding

Changes in DNA thermal melting temperature (ΔT_m) provide an initial ranking of compounds for binding affinity and relative sequence specificity with the target and control DNA sequences [31,32]. The binding affinities of the pyridyl-centered heterocyclic cations were tested with pure AT sequences, which are the primary sites of most known minor groove binders from netropsin to furamidine (Figure 1A) [33–35]. The target DNAs of primary interest for the compounds in this research have a single G•C bp with flanking AT sequences. The T_m experiments were carried out in the presence of nonalternating and alternating AT, and mixed flanking sequences, AAA-TTT, ATA-TAT, AAT-AAT, and AAA-AAA, Figure 2, where the dashes indicate zero, one, or two G•C bps. These are useful test sequences as each flanking AT sequence has different properties, including variations in the minor groove positions of the H-bond acceptor groups on the A•T bp. The number of G•C bp also varies in these selected mixed DNA sequences, which also gives significant differences in minor groove width (Figure 3).

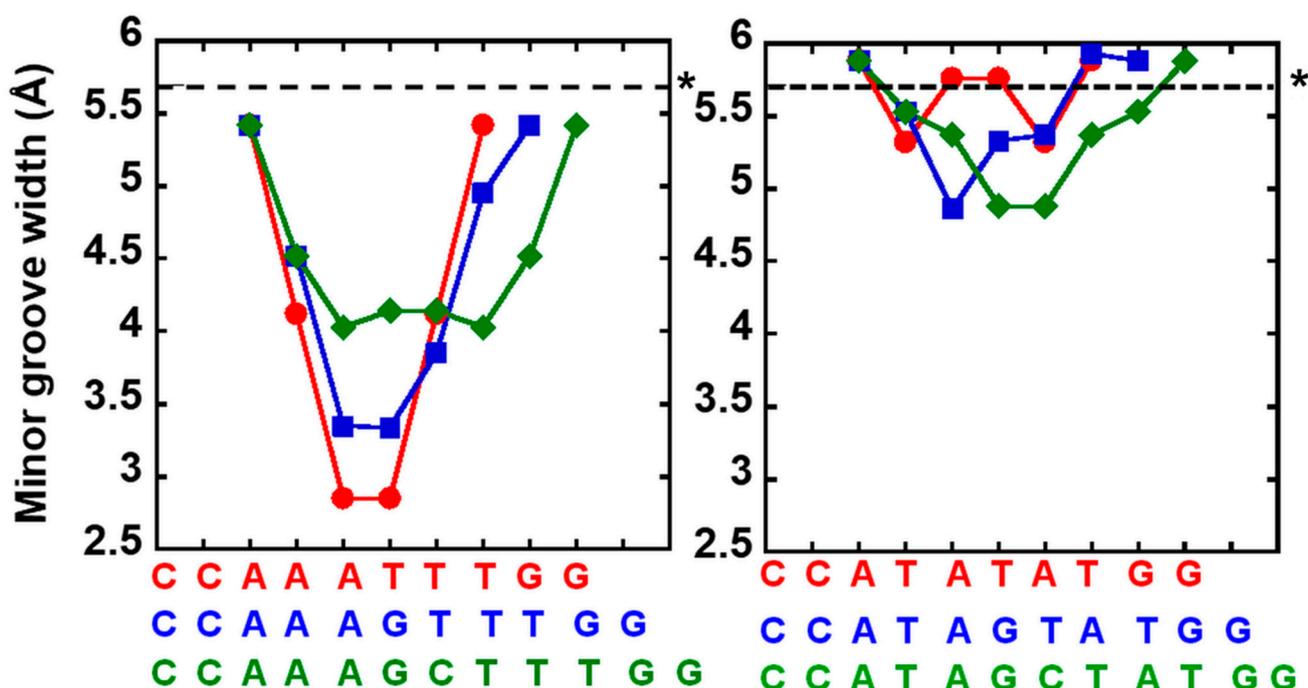


Figure 3. Minor groove width vs. target DNA sequences calculated from the online algorithm of Rohs and coworkers (*Nucleic Acids Res.* 2013, 4, W56–62). * indicates the minor groove width of standard B-form DNA. Groove width gives the perpendicular separation of helix strands drawn through phosphate groups, diminished by 5.8 Å to account for van der Waals radii of phosphate groups.

With AAA-TTT, a type of nonalternating AT sequence, DB2447, the direct truncated analog of the original pyridine, DB2120, resulted in an encouraging increase in the thermal stability of the single G•C bp-containing target site sequences (Figure 4 and Table S1). DB2447 also showed sequence selectivity, as expected from previous results and the compound design approach. The compound showed significantly lower thermal stability increases for the two G•C bps and all AT-containing sequences (Figure 4 and Table S1).

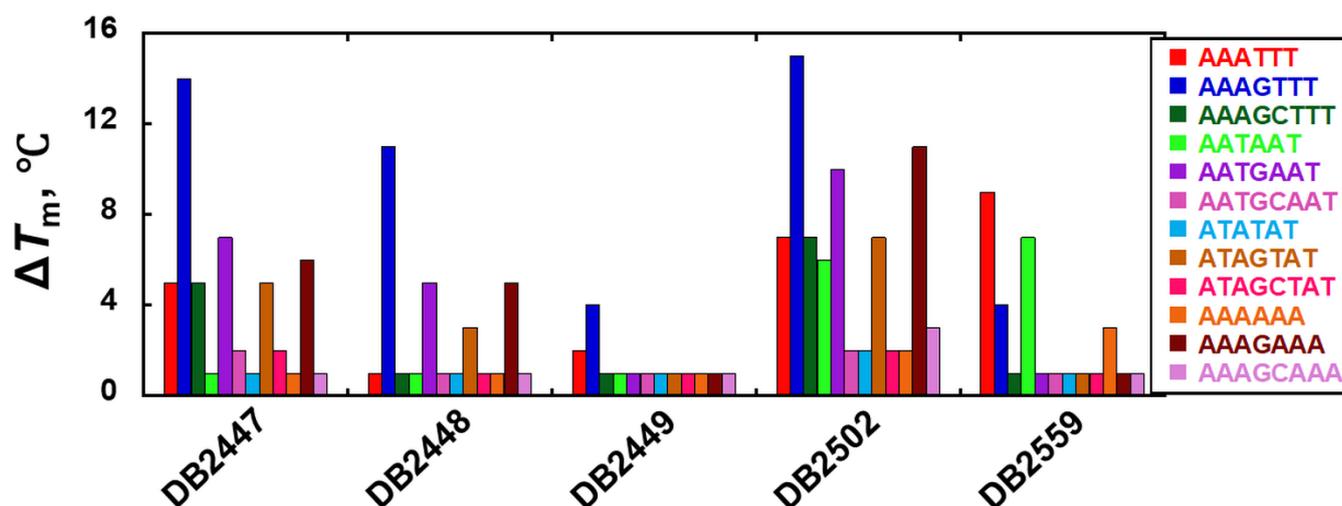


Figure 4. Comparison of relative binding affinities, ΔT_m , °C ($\Delta T_m = T_m(\text{the complex}) - T_m(\text{the free DNA})$) by thermal melting experiments of the designed heterocyclic amidine compounds with pure AT base-pair and mixed single G•C base-pair-containing DNA sequences. The listed values are for the 2:1 [ligand]/[DNA] ratio and an average of two independent experiments with a reproducibility of ± 0.5 °C. Full DNA sequences are as described in Figure 2.

Two analogs of DB2447 with terminal 4,5-dihydro-1-H-imidazole (*Im*) (DB2448) and 1,4,5,6-tetrahydro pyrimidine (*THP*) (DB2502) were synthesized and tested to determine how the molecular size and chemistry of terminal dications affect the sequence binding affinity and selectivity. The *Im* compound binds to a single G•C bp DNA somewhat more weakly than DB2447 but with significantly improved selectivity (Figure 4 and Table S1). The *THP* compound DB2502 binds to the single G•C bp target slightly more strongly than DB2447 and with similar selectivity. These results indicate that the terminal *Im* and *THP* analogs are useful additions to expand the chemical space of the parent, DB2447. The three compounds can be especially useful in biological test applications where their uptake into different cell types can be very different. In cellular applications where strong specificity is needed, the *Im* analog may be preferred. The compound DB2449, with one amidine replaced with an amide, has quite low ΔT_m values with all of the tested DNAs (Figure 4 and Table S1).

In addition to the compound structure and properties, the DNA sequence has a major effect on minor groove binding. The alternating DNAs ATA-ATA have a significant drop in ΔT_m relative to the target, AAA-TTT sequences with all compounds. This is expected based on the wider Minor Groove (MG) of the alternating sequence (Figures 3 and 4 and Table S1) [36,37]. The sequence with a pure A-tract AAA-AAA also has a surprisingly lower ΔT_m relative to the target, AAA-TTT sequences. The AAT-AAT sequences have an intermediate ΔT_m relative to the target, AAA-TTT sequences. In summary, the -AAAGTTT-sequence has the best ΔT_m of all of the flanking AT sequence variations. This can easily be seen in Figure 4, where the three pyridyl-diamidines with each DNA sequence group have the highest peaks in the histogram set. With DB2559, the -AAATTT- and -AATAAT-sequences have the highest ΔT_m s as expected for a central phenyl. Analysis of the minor groove widths of the different single G•C bp sequences showed that of all sequences, -AAAGTTT- has the most narrow minor groove and is most appropriate for binding compounds such as DB2447 with a connected aromatic system of approximately 3.4 Å in width (Figure 3). The weakest binding is seen with the -ATAGTAT- sequence and it has the widest DNA minor groove, basically the same as a standard B-form minor groove. The groove widths for the -AAAGTTT- and -ATAGTAT are shown in Figure 3 for reference.

3.3. Biosensor-Surface Plasmon Resonance (SPR): High-Resolution Evaluation of Binding Affinity, Kinetics, Stoichiometry, and Cooperativity

Biosensor-SPR methods provide a high-resolution, label-free way to quantitatively evaluate the binding affinity, selectivity, and stoichiometry of a set of compounds with a spectrum of immobilized DNAs [38–40]. The parent pyridyl-linked-phenyl-amidine, DB2447, binds strongly with AAAGTTT, and global kinetics fitting defined a single binding site with $K_A = 5.5 \times 10^8 \text{ M}^{-1}$ ($K_D = 1.8 \text{ nM}$) at 0.1 M NaCl (Figures 5 and 6, Table S2). The strong binding of DB2447 is the result of the rapid association that is at the instrumental limitation ($k_a = \sim 4.4 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$) and a comparatively slow dissociation rate constant ($k_d = 7.8 \times 10^{-2} \text{ s}^{-1}$). DB2447 binds to the pure AT, -AAATTT-, sequence as a monomer complex with rapid dissociation and a 200-fold lower affinity compared to AAAGTTT. This result indicates that DB2447 maintains surprisingly high sequence selectivity for the single G•C bp sequence. The sensorgram of AAATTT shows an off-rate that is much faster, and complete dissociation from the complex occurs within the first few seconds of the dissociation phase (Figure S1). With the -AAAGCTTT- binding site, DB2447 shows a 30-fold weaker binding affinity than with -AAAGTTT- under the same experimental conditions (Figures 5 and S1, Table S2). It appears that DB2447 can induce a somewhat favorable minor groove site in -AAAGCTTT- for binding, but the extra GC is a mismatch that reduces the binding affinity. The phenyl derivative, DB2559, with a simple -N- to -CH conversion, has a stronger binding for pure AT sequences, over ten times stronger than with the single G•C sequence, as with most classical minor grooves binders (Figures 5 and S2, Table S2). This result is as expected for compounds without an H-bond acceptor group in a position to bind to the G-NH that faces into the minor groove.

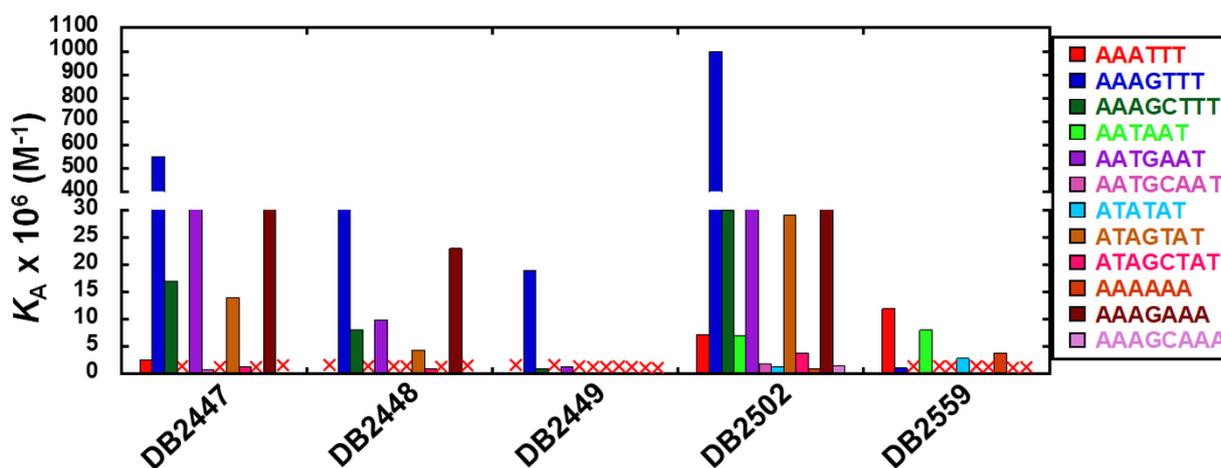


Figure 5. Comparison of equilibrium binding constants (K_A , M^{-1}) of phenyl and pyridine analogs with pure AT and mixed single G•C base-pair-containing DNA sequences. “X” represents no measurable K_A under our experimental conditions. The listed binding affinities are an average of two independent experiments carried out with two different sensor chips, and the values are reproducible within a 10% experimental error. Full DNA sequences are as described in Figure 2.

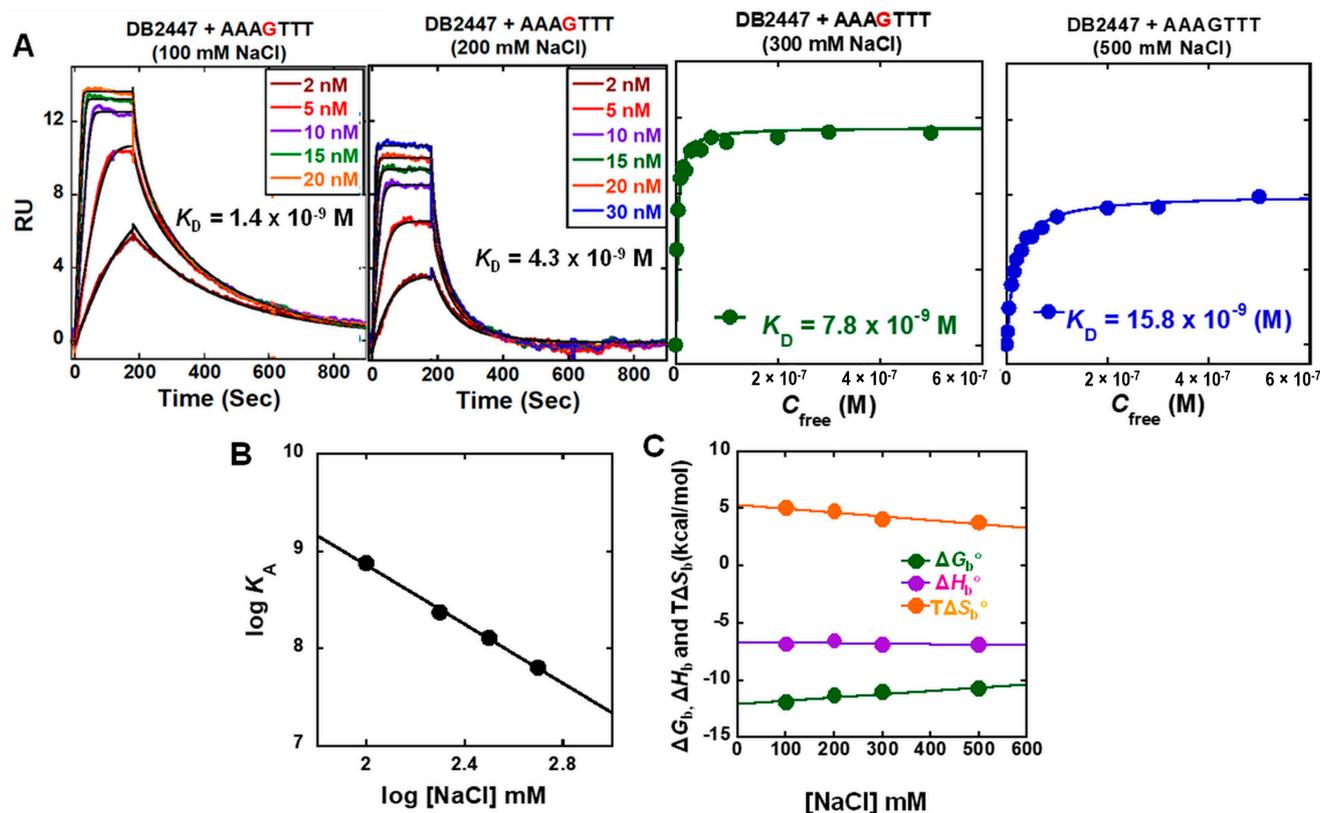


Figure 6. (A) SPR sensorgrams with kinetics fits and steady-state binding plots for DB2447 with the AAAGTTT DNA sequence at different salt concentrations; (B) salt dependence of K_A for DB2447 binding as determined by SPR. The K_A values were obtained by global kinetics (at two lower salt concentrations) and steady-state fits (at two higher salt concentrations); (C) plot of ΔG_b° , ΔH_b° , and $T\Delta S_b^\circ$ versus salt concentrations for DB2447 with AAAGTTT sequence at 25 °C. The listed binding affinities are standard thermodynamic values and are an average of two independent experiments carried out with two different sensor chips. The values are reproducible within a 10% experimental error; RU = Response Unit based on amounts of bound compounds on the immobilized DNA.

The derivative, DB2448, with an imidazoline terminal group has a K_D value about half as strong as DB2447 for the single G•C bp sequence, but it has negligible binding to AAATTT under our experimental conditions, an impressive improvement in sequence selectivity (Figures 5 and S3, Table S2). DB2502, on the other hand, with six-atom terminal cations has a K_D value that is about twice as strong as that of DB2447. Unfortunately, its K_D value for the pure AT sequence is about three times stronger than for DB2447. With the AAAGCTTT sequence, DB2448 has a K_D value over twice that of the DB2447 constant but the DB2502 value is below half of the DB2447 K_D value. The *Im* substitution thus has a promising increase in overall selectivity while the results with DB2502 are a disappointment in selectivity for this series (Figures 5 and S4, Table S2). As with the T_m experiments, DB2449, the monocation, has relatively weak binding to all tested DNA sequences except -AAAGTTT with a K_D of 52 nM (Figures 5 and S5, Table S2). This compound has very different solution properties relative to all other compounds and may be an advantage in cell studies.

With the other three sets of flanking sequences, the single G•C bp sequences always have the strongest binding (Figures 5 and S1–S5, Table S2), as with -AAAGTTT-, but the binding is weaker than with -AAAGTTT-. With -AAAGAAA- and AATGAAT, the binding with DB2447 is about a factor of ten weaker than with -AAAGTTT-. With the fully alternating sequence -ATAGTAT-, the binding is reduced by close to a factor of 40. Similar reductions in affinity are seen with the other compounds and these DNA sequences

(Figure 5, Table S2). With the Figure 5 histograms, the single G•C set of sequences have the highest plot in each set of DNAs. It should also be noted, however, that the GC sequences have a stronger-than-expected binding.

3.4. Molecular Curvature Determination

Molecular curvature plays a crucial role in sequence-selective DNA minor groove recognition in conjunction with DNA minor groove binders' molecular functionality and stacking surface. Correct curvature is important for strong H-bonding interaction, charge interactions, and van der Waals stacking in the groove of the DNA. Our previous report [28] stated a graphical approach method to determine relative molecular curvature values for minor groove binding compounds. In this approach, the diamidines compounds are energy-minimized in the SPARTAN software package using the DFT/B3LYP theory with the 6-31+G* basis set. The compounds are then matched up in a PowerPoint graphics package. A reference circle (black circle) is drawn through both amidine carbons, the center of the compound where the circle's periphery passes through the center point of the individual molecular unit of the entire molecule, illustrated in DB2447, DB2448, and DB2502 (Figure 7). Two straight lines (orange) are drawn from the circle point at the center of the molecule to the amidine carbons. The midpoint of these two lines defines the comparative curvature value for the diamidines compound. The curvature values are 147° for DB2447, 147° for DB2448, and 143° for DB2502. The curvature analysis of a library of strong DNA minor groove binding compounds by this method offers a standardization value of around 140–145° curvature angle. DB2447 and its analogs also show a similar ° of curvature angle, which supports the strong binding affinity of these diamidines compounds.

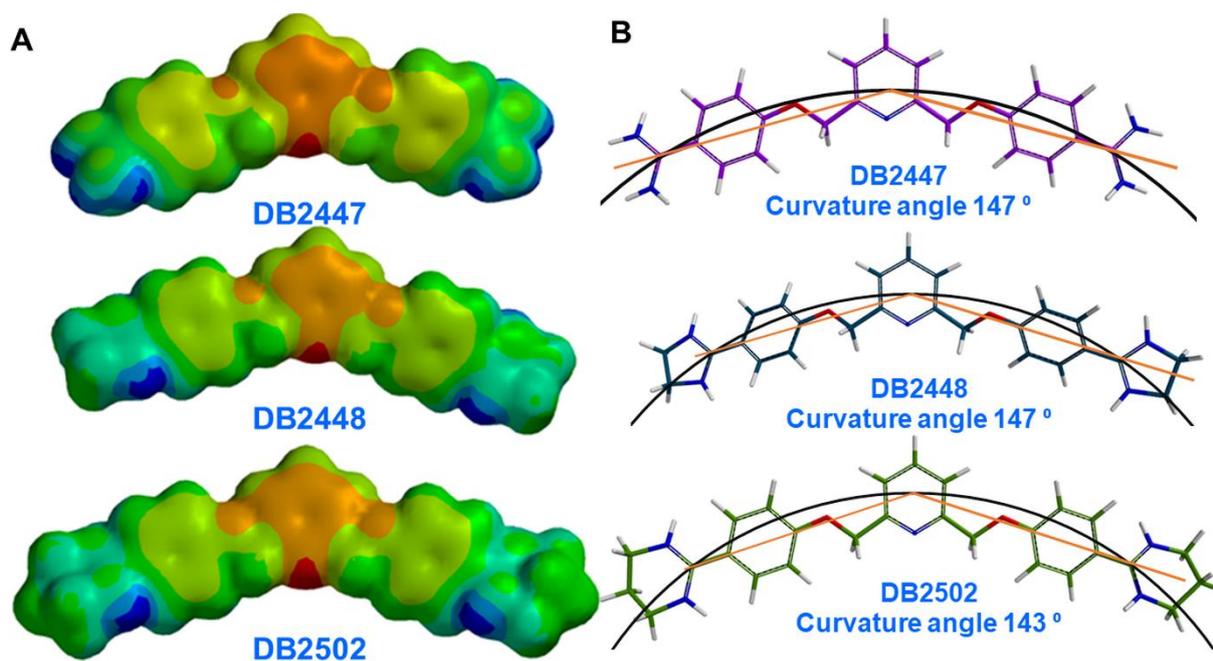


Figure 7. (A) Energy-minimized structures of DB2447, DB2448, and DB2502 at B3LYP/6-31G* (p,d) level of theory. In the electrostatic potential maps, red indicates high electronegativity and blue indicates electron-deficient/positively charged regions. (B) Molecular curvatures for DB2447, DB2448, and DB2502. Black circles represent reference circles and orange lines represent two straight lines drawn from the circle point at the center of the molecule to the amidine carbons.

3.5. Effect of Salt Concentration and Temperature on DB2447 Binding to Single G•C bp DNA Sequences

The T_m , structure evaluation, and SPR binding results indicate that DB2447 has an optimized length, curvature, and flexibility for effective and selective recognition of a single G•C bp in an AT minor groove sequence. The SPR results also indicate that DB2447

has improved solution properties relative to DB2120, which allows us to explore the thermodynamic behavior of this compound with single G•C bp and other sequences in more detail. These results help to provide a fundamental understanding of the molecular basis for specific recognition of the DNA minor groove.

To evaluate the effect of ionic strength on DB2447 binding affinity with the -AAAGTTT-sequence, SPR experiments were carried out from 100 to 500 mM NaCl concentrations at 25 °C (Figure 6, Table 1). The equilibrium binding constants (K_A) obtained either by global kinetics fits at low salt concentration or by steady-state fits at higher salt concentration are collected in Table 1 and Figure 6. Both theory and experiment suggest that the logarithm of the equilibrium binding constant K_A is a linear function of the logarithm of NaCl concentration for many organic cations binding to DNA [41,42]. For a typical DNA–cation complex, the equilibrium binding constant values decrease as the salt concentration increases with a slope that depends on the compound charge [42–46]. As seen in Figure 6, the $\log(K_A)$ versus $\log[\text{Na}^+]$ plot for DB2447 is linear with a slope of 1.6. The number of phosphate contacts (Z) between DB2447 and the AAAGTTT DNA sequence is predicted to be two and can be obtained in experiments from the slope/0.88, where 0.88 is the fraction of phosphate charge shielded by the total associated counterions. For the 11-base-pair synthetic oligomer AAAGTTT, the obtained Z is 1.8. These results indicate that the dicationic DB2447 releases 2 Na^+ ions when binding to the DNA minor groove. The enthalpy change, ΔH_b° , for binding is essentially constant with salt concentration, while for both ΔG_b° and $T\Delta S_b^\circ$, values decrease by one kilocalorie with the change in salt concentration from 100 to 500 mM NaCl.

Table 1. Kinetics ^a and steady-state ^b analysis of DB2447 with the AAAGTTT sequence at different salt concentrations in 50 mM Tris-HCl, 1 mM EDTA buffer, pH 7.4 at 25 °C.

NaCl (mM)	K_D (nM)	ΔG_b° (kcal/mol)	ΔH_b° (kcal/mol)	$T\Delta S_b^\circ$ (kcal/mol)
100	1.4	−12.0	−6.8	5.2
200	4.3	−11.4	−6.5	4.9
300	7.8	−11.1	−6.9	4.2
500	15.8	−10.7	−6.9	3.8

^a Kinetics analysis was performed by global fitting with a 1:1 binding model and ^b steady-state fits were carried out by using 1:1 binding equation model. ΔH_b° was determined in ITC experiments; $T\Delta S_b^\circ = -\Delta G_b^\circ + \Delta H_b^\circ$.

To obtain an additional understanding of the thermodynamic basis for DB2447 interactions with DNA, SPR experiments were conducted from 25 °C to 40 °C at 100 mM salt concentration. The SPR results reveal that temperature significantly affects the DB2447-DNA binding thermodynamics and kinetics, as shown in Figure 8 and Table 2. The ligand-DNA binding affinity decreases (K_A) with experimental temperature. However, the temperature changes have a smaller effect on the ΔG_b° as previously observed for other minor groove binders [42]. As can be seen in Figure 7, the enthalpy and entropy for binding have compensating decreases with increasing temperature. The $T\Delta S_b^\circ$ and ΔH_b° values are similar at 20 °C but $T\Delta S_b^\circ$ approaches zero at 45 °C, and ΔH_b° completely accounts for ΔG_b° .

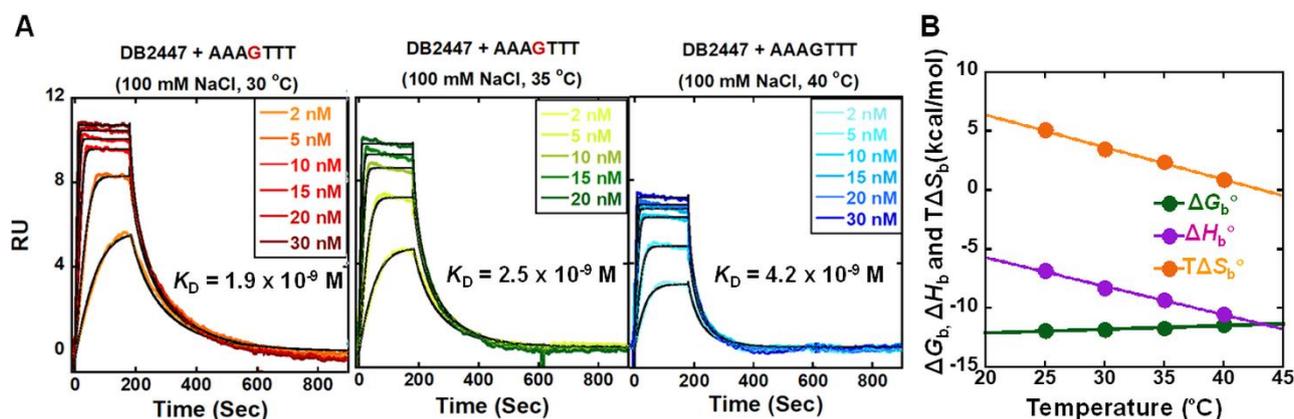


Figure 8. (A) SPR sensorgrams with kinetics fit for DB2447 with the AAAGTTT DNA sequence at different temperatures at 100 mM NaCl; (B) thermodynamic results, from Table 2 for binding of DB2447 to the AAAGTTT site at different temperatures. The listed binding affinities are standard thermodynamic values and binding affinities are an average of two independent experiments carried out with two different sensor chips, and the values are reproducible within a 10% experimental error; RU = Response Unit based on amounts of bound compounds on the immobilized DNA.

Table 2. Kinetics and thermodynamic results for DB2447 with the AAAGTTT sequence at different experimental temperatures at 50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA buffer at pH 7.4.

Temperature (°C)	K_D (nM)	ΔG_b° (kcal/mol)	ΔH_b° (kcal/mol)	$T\Delta S_b^\circ$ (kcal/mol)
25	1.4	−12.0	−6.8	5.2
30	2.0	−11.8	−8.3	3.5
35	2.5	−11.7	−9.3	2.4
40	4.2	−11.4	−10.5	0.9

3.6. Isothermal Titration Calorimetry (ITC) of Complex Formation: Effects of Salt Concentration and Temperature

Characterization of the full thermodynamics of cationic compound–DNA interactions is an essential component of any detailed analysis of DNA molecular recognition and rational drug design. Isothermal titration calorimetry (ITC) is the method of choice to analyze the energetic basis for the strong and selective binding of compounds with the DNA minor groove. It has been used with minor groove binders since the early studies of Ken Breslauer and colleagues and it is applied here to DB2447–DNA complexes. ITC provides a key component of the thermodynamic profile of ligand–DNA interactions by direct determination of the enthalpy (ΔH_b°) of binding. With the enthalpy and ΔG_b° from biosensor experiments, it is possible to calculate the entropy of binding (ΔS_b) from the known thermodynamic relationships ($\Delta G_b^\circ = -RT\ln K_A$) and ($\Delta G_b^\circ = \Delta H_b^\circ - T\Delta S_b$). Thermodynamic profiles are valuable in drug design because they provide quantitative data on drug–DNA interactions that cannot be obtained directly by structural or computational methods [47–51]. They also provide valuable ideas about the compound–DNA–water components of complex formation [52].

In the experiments reported here, ITC was used to monitor the heat released upon the binding of DB2447 to the target binding sites. For strong binding compounds such as DB2447, ITC experiments require considerably higher concentrations than the compound K_D for DNA interactions. In cases such as this, the enthalpy of binding can be determined quite accurately in the presence of excess DNA in the calorimetry cell such that the compound is fully bound to the DNA. In this model-free approach, the ΔH_b° is simply determined from the average of ΔH_b° versus the binding ratio below saturation binding. The binding constant must then be determined at lower concentrations by an alternative method such as SPR or fluorescence methods.

The ITC curves were fitted using Origin software to obtain the enthalpy at each titration point. The data (Figure 9) indicate an exothermic interaction after adding DB2447 to the solution containing -AAAGTTT-DNA at each NaCl concentration. The subtraction of the integrated peak areas for ligand/buffer titration from the ligand/DNA titration gives a direct determination of ΔH_b° at each temperature. Figure 9 shows the titration of DB2447 into AAAGTTT with the blank buffer correction, and the ΔH_b° value is -6.8 ± 0.2 kcal/mol at 25 °C at a 100 mM salt concentration. The favorable negative enthalpy change suggests strong H-bonding, electrostatic, and van der Waals interactions between DB2447 and AAAGTTT DNA. Interestingly, the complex formation of DB2447-AAAGTTT has an enthalpy at 25 °C that is larger than most of the A-tract minor groove binders reported in the literature [44,47–49]. This is expected from the known sequence-dependent differences in minor groove structure and hydration between the single G•C bp and all AT DNA sequences (Table 1).

To evaluate the relationship between thermodynamic measurements of the DB2447-AAAGTTT complex and experimental salt concentrations, the ITC experiments were carried out at 100–500 mM NaCl concentrations. The SPR results show that the binding constant decreases by almost ten times (Table 1) with increasing salt concentration. However, the ITC experiments show that salt concentrations have a much smaller effect on ΔH_b° than on ΔG_b° (Figure 9 and Table 1). This phenomenon reveals the enthalpy of complex formation for an energetic component, which is the sum of interactions such as hydrogen bond formation and van der Waals interactions and is essentially independent of salt concentration.

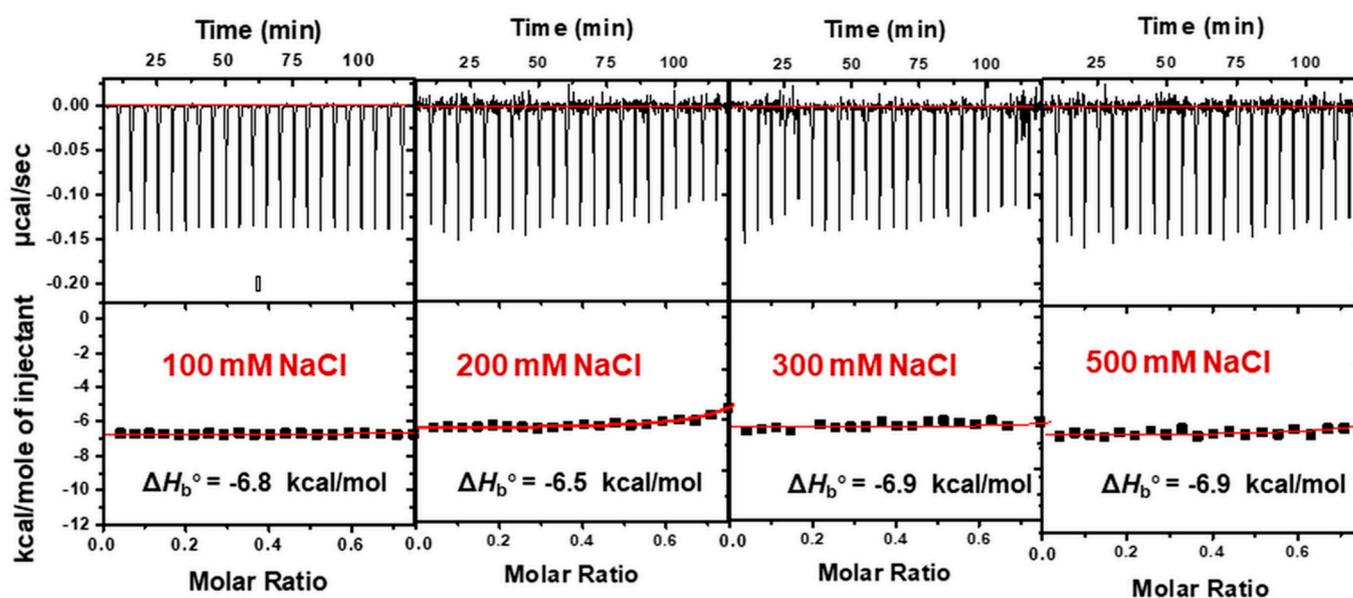


Figure 9. ITC data for the titration of DB2447 with AAAGTTT DNA at different salt concentrations at 25 °C. The listed binding enthalpies are an average of two independent experiments, and the values are reproducible within a 10% experimental error.

3.7. Determination of the Heat Capacity of the DB2447–AAAGTTT Complex

The ITC experiments of DB2447 with AAAGTTT were also carried out at different temperatures (15–40 °C) with a constant 100 mM NaCl concentration (Figure 10). The titration profiles indicate that the enthalpy of DB2447–DNA complex formation strongly depends on the experimental temperature and becomes more negative with increases in temperature (Table 2). The temperature-dependent differences in the binding enthalpy of DB2447–DNA complexes were used to calculate the heat capacity (ΔC_p) for binding from the slope of a linear least-squares fit of the plot of ΔH_b° versus temperature, $\Delta C_p = -285$ cal/mol K (Figure 10). The temperature-dependent ITC results also show the effect of temperature on the entropy term. As the ΔG_b° of binding is essentially constant with temperature (Figure 8,

Table 2), the subtraction of ΔH_b° from ΔG_b° yields results that show that the entropy of binding decreases as the temperature is increased (Table 2) and, as previously noted, $T\Delta S_b^\circ$ approaches zero at 40–45 °C.

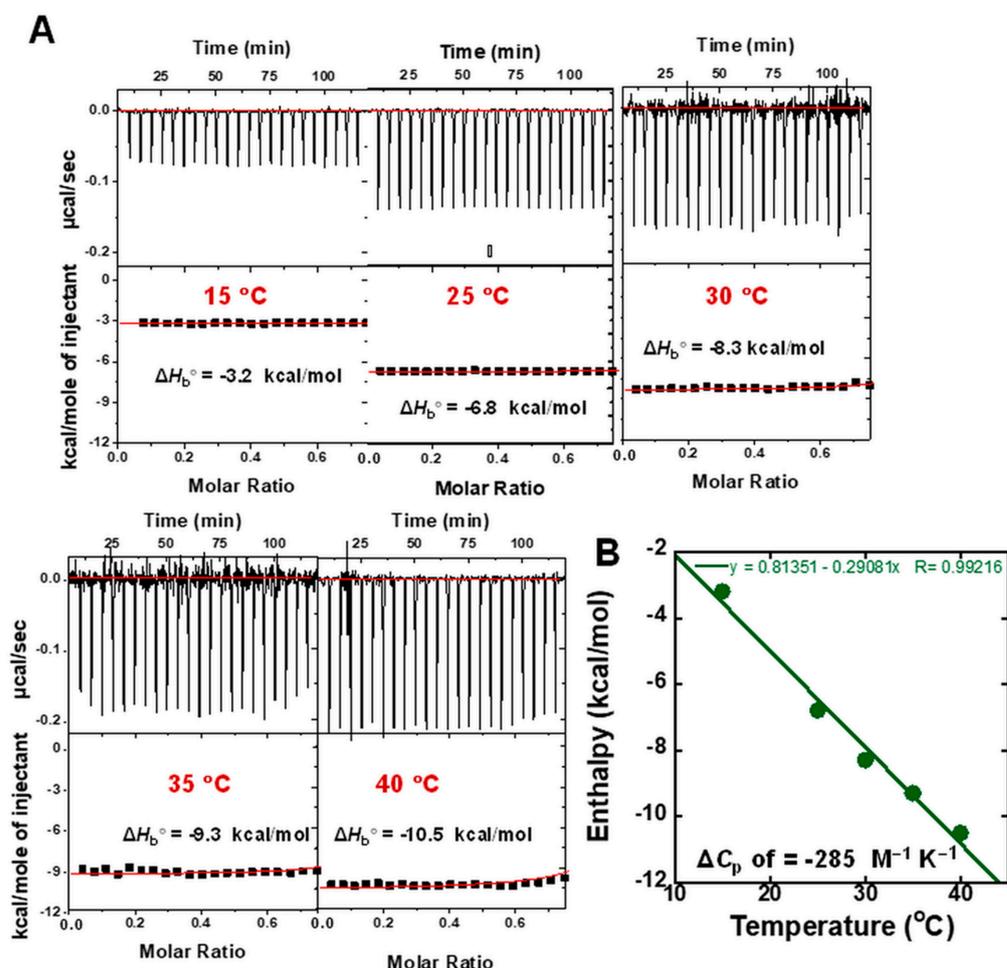


Figure 10. (A) ITC data for the titration of DB2447 with AAAGTTT DNA at different experimental temperatures by using 100 mM NaCl. (B) Plot of ΔH_b° versus temperature for DB2447 with AAAGTTT DNA, and the linear fit yields a ΔC_p of $-285 \text{ M}^{-1} \text{ K}^{-1}$. The listed binding enthalpies are an average of two independent experiments, and the values are reproducible within a 10% experimental error.

3.8. The Effects of AT Flanking Sequence Variations on DB2447 Binding Thermodynamics

In Figure 11, the effects of the four sequences with different AT sequences flanking the single G•C bp binding site are shown. Determination of the ITC ΔH_b° and the SPR ΔG_b° (Figure 4) allows calculation of the ΔS_b° for each sequence (Figure 11). The -AAAGTTT-sequence has the most favorable ΔG_b° and with a substantial ΔS_b° , which is why it is the best binding sequence. The AATGAAT sequence has a similar ΔH_b° but a smaller ΔS_b° with a decreased ΔG_b° . With AAAGTTT and ATAGTAT, the binding ΔH_b° is decreased but the ΔS_b° is higher than with AATGAAT. With AAAGTTT, the binding becomes entropy-driven, and binding to AAATTT is shown as a strongly entropy-driven reference (Figure 11).

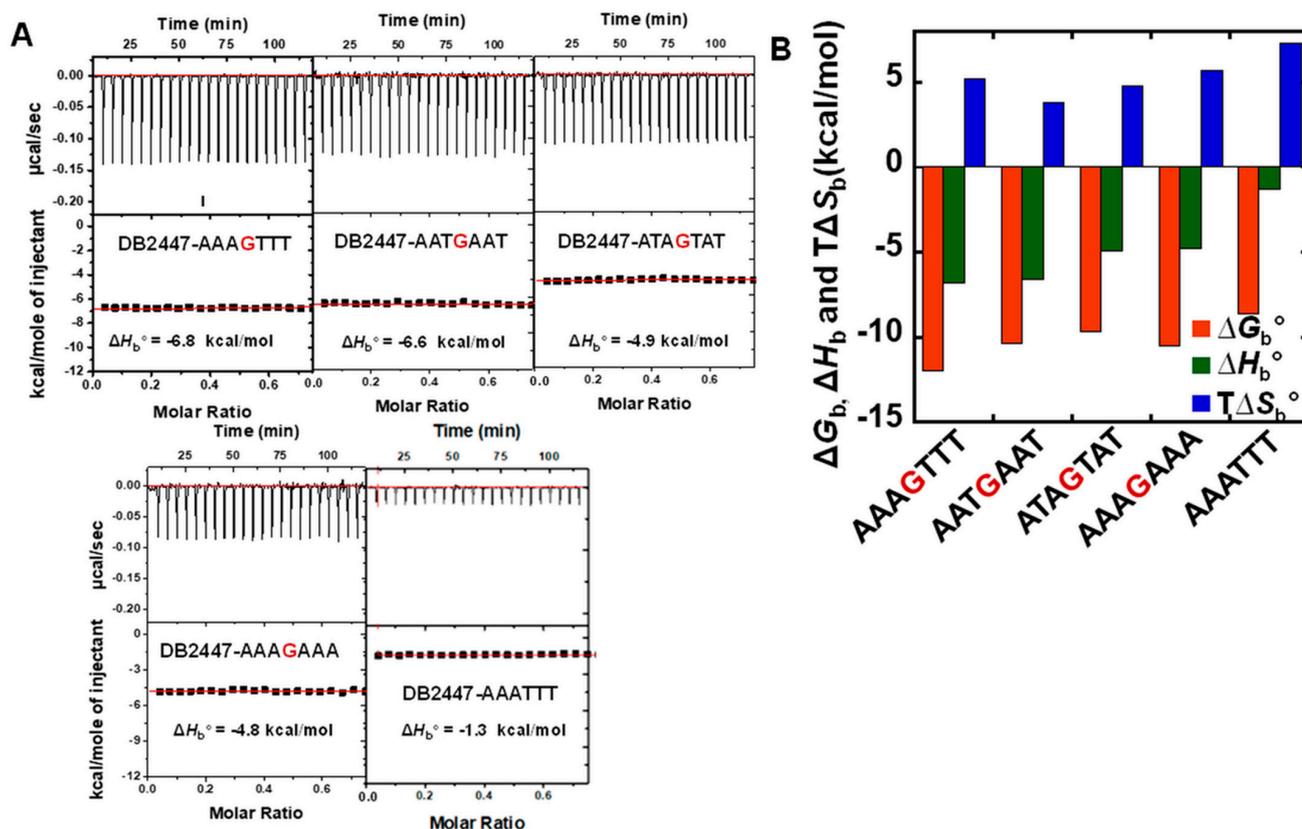


Figure 11. (A) ITC data for the titration of DB2447 with different single G•C base-pair-containing mixed DNA sequences in 50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA buffer, pH 7.4 at 25 °C. (B) Comparison of thermodynamic parameters, ΔG_b° , ΔH_b° , and $T\Delta S_b^\circ$, of DB2447 with different single G•C base-pair-containing mixed DNA sequences. The plotted thermodynamic parameters are an average of two independent experiments, and the values are reproducible within a 10% experimental error. Full DNA sequences as described in Figure 2.

3.9. Thermodynamic Effects in the Binding of the Pyridyl Diamidine Compounds

The three compounds have similar thermodynamics with DB2502 having a slightly larger ΔG_b° , while the value is smallest with DB2448 (Figure 12). DB2447 has the most favorable ΔH_b° , suggesting that the unsubstituted amidines form the most favorable H-bonds. DB2502 has the most favorable ΔS_b° , indicating that the large tetrahydropyrimidine displaces the most water from the minor groove on binding.

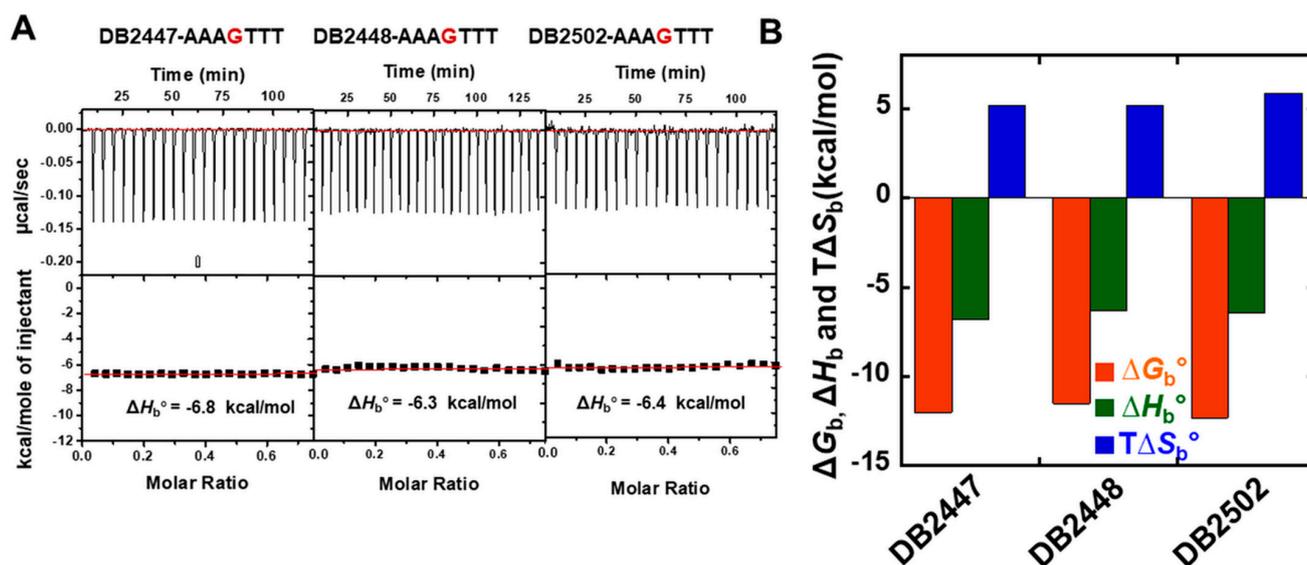


Figure 12. (A) ITC data for the titration of DB2447, DB2448, and DB2502 with AAAGTTT DNA sequence in 50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA buffer, pH 7.4 at 25 °C. (B) Comparison of thermodynamic parameters, ΔG_b° , ΔH_b° , and $T\Delta S_b^\circ$, of DB2447, DB2448, and DB2502 with AAAGTTT. The plotted thermodynamic parameters are an average of two independent experiments, and the values are reproducible within a 10% experimental error.

3.10. Competition Electrospray Ionization Mass Spectrometry (ESI-MS) of DB2447

Competition MS allows high-throughput screening for the comparison of binding of compounds to a panel of DNA sequences for the evaluation of relative affinity and selectivity [53,54]. The dashes represent zero, one, or two G•C bps with the AAA-TTT sequences used. In Figure 13A, the free DNA peaks are shown for AAATTT (m/z 6684), AAAGTTT (m/z 7302), and AAAGCTTT (m/z 7921). After the addition of DB2447, the intensity of the peak for AAAGTTT (m/z 7302) is reduced with the appearance of a new peak at m/z 7672, which is the characteristic of a 1:1 AAAGTTT-DB2447 complex (Figure 13B). This is in agreement with the ΔT_m and SPR results that show stronger binding to the single G•C bp sequence. At the 1:1 ratio of compound-to-DNA in this experiment, only binding to -AAAGTTT- is seen. There is no appearance of other DNA-ligand complex peaks at this compound-to-DNA ratio. As the ratio is increased beyond that in Figure 13, binding to other, less favored sequences begins to be observed. The observed ESI-MS spectra strongly indicate the high sequence specificity and affinity of DB2447 for the single G•C bp sequence.

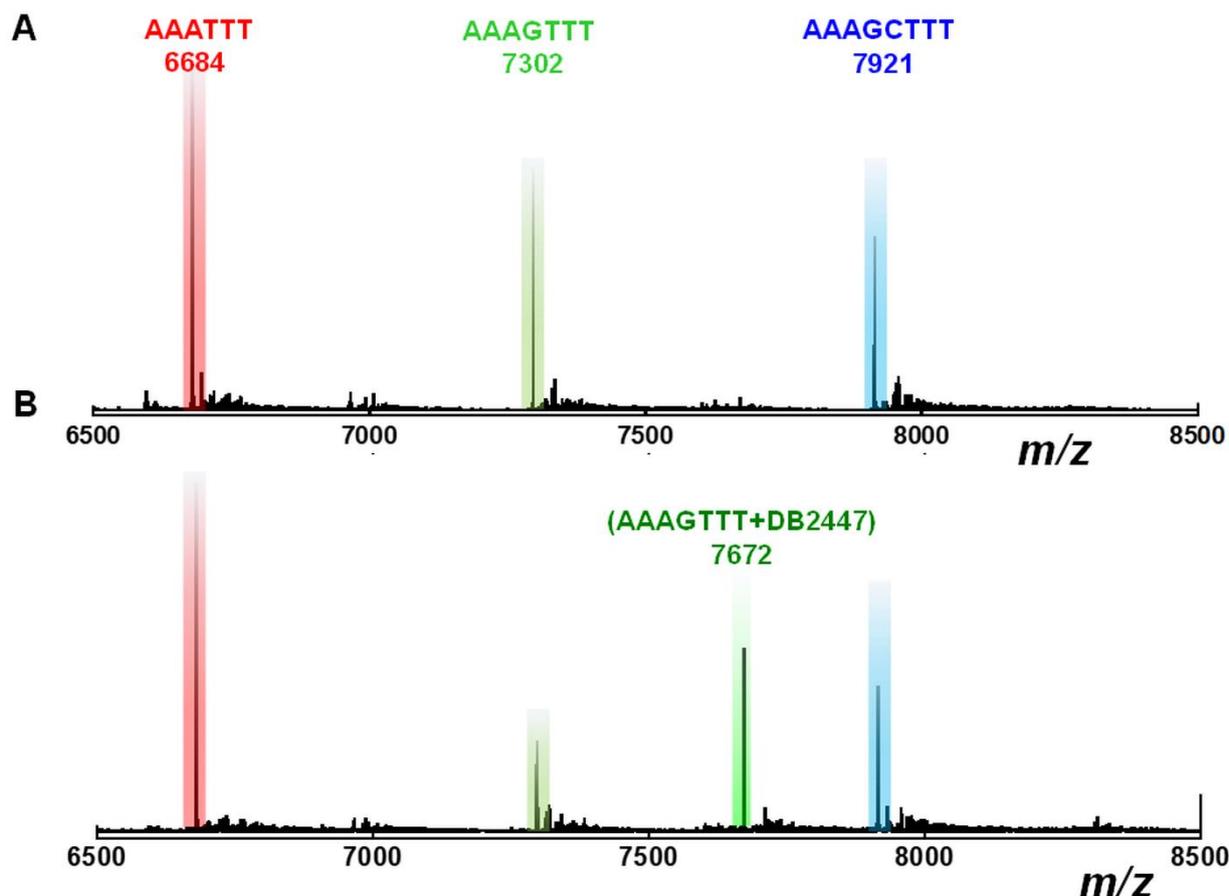


Figure 13. ESI-MS negative mode spectra of the competition binding of sequences AAATTT, AAAGTTT, and AAAGCTTT (10 μ M each), with 30 μ M DB2447 in buffer (100 mM ammonium acetate with 10% methanol (v/v), pH 6.8). (A) The ESI-MS spectra of free DNA mixtures; (B) the ESI-MS spectra of DNA mixtures with DB2447. The ESI-MS results are deconvoluted spectra, and molecular weights are shown with each peak.

4. Discussion

Over the last 50 years, there have been extensive studies on the interaction of a broad range of minor groove binders with DNA. Most of these compounds, especially those in the initial studies such as netropsin, DAPI, and Hoechst 33258 (Figure 1A), have been specific for binding to pure AT bp sequences, especially with A-tract type sequences. In the project described in this paper, the studies have been broadened to include new compounds from our laboratories that were designed to include a pyridine group with the goal to add a G•C bp to the traditional AT recognition sequence (Scheme 1, Figure 1C). The DNA sequences were also designed to include variations in the AT bp sequences that flank the G•C bp (Figure 3) to probe the effects of flanking sequences on affinity and specificity in binding. The pyridine group has been incorporated in the minor groove binders to complex with the G-NH that protrudes into the minor groove. One control compound has a phenyl in place of the pyridine to quantitatively evaluate the pyridine effects. The terminal cationic groups on the compounds (Figure 1) were varied to determine their role in DNA recognition. The compounds have either amidines, imidazolines, tetrahydropyridines, or, in one case, a combination of one amidine and one amide for a total of five compounds. The DNA samples included in the studies have zero, one, or two central GC base pairs with various flanking AT sequences (Figure 2). The combination gives five compounds and 12 DNA sequences for broad comparison with thermal melting, SPR, ITC, and combination MS methods under a variety of salt concentration and temperature conditions.

The ΔT_m results present a complete but low-resolution picture of all five compounds with the 12 DNA samples. For all of the pyridine derivatives, the single GC sequences have the highest ΔT_m values, and the highest of all is obtained with the -AAAGTTT- sequence (Figure 4, Table S1). It clearly has the optimum combination of groove width, curvature, and placement of H-bond acceptor groups in the minor groove to interact with the pyridine-substituted compounds (Figure 1C). With DB2447, for example, the ΔT_m with -AAAGTTT- is 14 °C, while the ΔT_m values for all other single G DNAs are between 5 and 7, about one-half the -AAAGTTT- ΔT_m . In all cases, the alternating AT sequence -ATAGTAT- has the lowest ΔT_m values with all DNA sequences. This sequence has a significantly wider minor groove width than -AATGTTT- and is a less favorable binding site for these types of minor groove binders (Figure 3). DB2448, with imidazoline cationic groups, binds a little weaker than the amidine, but it has excellent selectivity and only has a significant ΔT_m with single G•C sequences. DB2502 with tetrahydropyrimidine terminal cations binds more strongly than the other two compounds but with lower specificity. Ab initio calculations on the three pyridine dications indicate that DB2448 is the most planar structure while DB2502 is the least planar. DB2559 with the pyridine replaced with a phenyl binds best to pure AT sequences as expected. It has the best binding with -AAATTT-, while all other pure AT sequences only have about one-half of the ΔT_m of -AAATTT-.

Biosensor-SPR methods provide more quantitative binding results than the ΔT_m values but they are in qualitative agreement (Figures 4 and 5). The K_D for DB2447 with -AAAGTTT- is 1.8 nM, and results with -AAAGAAA- and AATGAAT are ten-fold higher, in agreement with their lower ΔT_m values. As with ΔT_m , the weakest binding is seen with -ATAGTAT- with a K_D of 68 nM. The same trend is observed with the other pyridine compounds with the ΔT_m results: DB2448 binds slightly weaker but with greater selectivity, while DB2502 binds slightly more strongly but with lower selectivity. DB2559 binds best to pure AT sequences and, as expected, it has the strongest binding to -AAATTT-.

Given that -AAAGTTT- is the best binding sequence, more detailed thermodynamic studies were conducted with it and DB2447 to determine what components are most important for complex formation. As is typical of many biological complexes that are formed from numerous relatively weak interactions, the pyridine DB2447 complex in this set has a large negative heat capacity for complex formation, -248 cal/mol deg. With DB2447, the ΔH_b° and $-T\Delta S_b$ values are -6.8 and -5.1 kcal/mol at 25 °C, respectively. As expected from the negative heat capacity, the values at 40 °C are -11.4 and -0.90 kcal/mol and the compensating changes maintain an almost constant ΔG_b° . This agrees with the thermodynamics for other minor groove binders that interact with a G•C bp and indicates an enthalpy-driven complex stabilized by an array of H-bonding, van der Waals, and electrostatic interactions [22,55]. With compounds such as DB75 that recognize only AT bp sequences, the binding entropy is the dominant component of the complex formation due to the release of water from the minor groove AT sites [49]. With DB2559, for example, the ΔH_b° and $-T\Delta S_b^\circ$ values are approximately -4 (Figure S6) and -6 kcal/mol at 25 °C, respectively, and the difference is even larger with DB75, approximately -2 and -7 kcal/mol, respectively [49]. The clear conclusion from the available results then is that adding a G•C bp to a minor groove binder recognition sequence significantly increases the binding enthalpy and reduces the binding entropy. A similar thermodynamic shift was seen with DB293 and DB2277 [22,55]. With an -AATT- sequence, however, DB293 binds as an entropy-driven complex [56]. In all of these systems, the binding energetics shift more to energetic emphasis on H-bond formation through the G-NH to compound acceptor group and more to water release and entropy in a pure AT minor groove sequence.

Strong support for the highly selective binding of DB2447 to -AAAGTTT- is seen in competition mass spectroscopy experiments (Figure 13). This is in agreement with the ΔT_m and SPR results that show stronger binding to the single G•C sequence. To help better understand the structural basis of molecular recognition of DNA sequences with a GC bp in an AT flanking sequence context, a molecular dynamics (MD) simulation for a complex of the pyridine compound, DB2447, with the DNA sequence *ds*[5'-CCAAAGTTTGG-

3') (5'CCAAACTTTGG-3') was conducted (Figure 14). Force constants for DB2447 were determined as described previously and added to the force field for the simulation [57,58]. The MD simulation was performed by using Amber 16 in the presence of 0.15 M NaCl as previously described. The DB2447 complex can dynamically orient to provide favorable curvature to the DNA complex and interactions between the compound and DNA. The pyridine N and amidine -NH groups are positioned for strong H-bonds with the -G-NH (3.1 Å) and -two T=O (2.9–3.0 Å) groups at the floor of the minor groove (Figure 14). The H-bonding ability, stacking with the minor groove walls, and dynamics of the bound system help provide the high binding affinity of DB2447 to the -AAAGTTT- binding site. The strong G-NH to pyridine N H-bond provides high binding selectivity of DB2447 toward the AAAGTTT sequence, in agreement with the ESI-MS results. Additional selectivity in binding is provided by -CH groups of the two phenyls that point into the minor groove (Figure 14B). The -CH groups form a dynamic weak interaction with -dT=O that are adjacent to the central G-C bp. The combination of weak to strong interactions in the complex gives the large negative heat capacity and strong binding of DB2447.

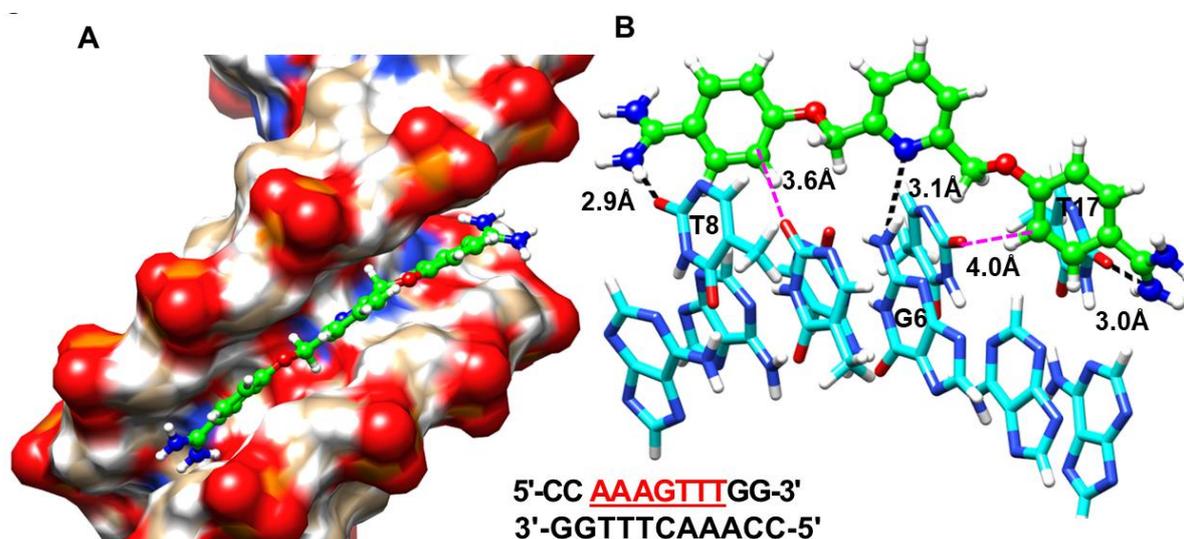


Figure 14. (A) A Snapshot of minor groove view of the DB2447-d[(5'-CCAAAGTTTGG-3')(5'-CCAAACTTTGG-3')] complex from MD simulations. The sphere model in a green-white-blue-red-yellow (C-H-N-O-S) color scheme represents DB2447. The DNA bases are represented in a ball and stick and ribbon model with a cyan-white-red-blue-orange (C-H-O-N-P) color scheme. (B) The important interactions between different sections of the DB2447–DNA complex are illustrated. DB2447 forms three direct H-bonds (black dashed lines) with DNA bases. Red lines indicate compound aromatic -CH interactions with -T=O groups.

For drug design, it is essential to understand the effects of solution conditions on minor groove binder–DNA complexes. The equilibrium constant, K_A , decreases as the salt concentration is increased as with other minor groove binders. As expected, the slope of a $\log K_A$ versus $\log[\text{Na}^+]$ plot is linear with a slope of 1.6 (Figure 6). The enthalpy change, ΔH_b° , is affected by a very small amount with changes in salt concentration. The effects of salt concentration on ΔG_b° and $T\Delta S_b^\circ$ are complementary and amount to about a one-kilocalorie decrease as the salt concentration increases from 100 to 500 mM NaCl. As the temperature increases, however, ΔH_b° becomes much more negative and the heat capacity for binding is a large negative value for binding a small molecule.

ITC experiments were also used to evaluate the effects of the DNA sequence on the binding of DB2447. The results show (Figure 11) that the -AAAGTTT- sequence has substantial ΔH_b° and ΔS_b° values and the most favorable ΔG_b° , which is why it is the best binding sequence. The -ATAGATA- sequence has a relatively small ΔH_b° and ΔS_b° that sum to give it the lowest ΔG_b° value of all the single G sequences. Evaluation of the three pyrimidine

diamidine compounds shows that they all have similar binding thermodynamics. DB2502 has the most favorable ΔS_b , indicating that the cyclic tetrahydropyrimidine group displaces the most water from the minor groove on binding. DB2448 has the lowest ΔS_b that sums with its ΔH_b° to give it the lowest ΔG_b° of the three compounds.

In summary, the results presented here show that minor groove binding thermodynamics depends on both compound structure and DNA sequence.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/life12050681/s1>. References [21,22,24,28] are cited in the supplementary materials. Experimental details of synthesis, DNA melting experimental data table, SPR sensorgrams, and affinity binding curves for DB2447, DB2448, DB2449, DB2502, and DB2559; SPR equilibrium dissociation constants (K_D , nM) data table for DB2447 and analogs with pure A·T and mixed DNA sequences; ITC of DB2559 and DB2448 with AAATTT sequence; ^1H NMR spectra of final products. Figure S1: Comparison of equilibrium binding constants (K_D , M) of DB2447 with pure AT and mixed single/two G•C base pair(s) containing DNA sequences; Figure S2: Comparison of equilibrium binding constants (K_D , M) of DB2559 with pure AT and mixed single G•C base-pair containing DNA sequences; Figure S3: Comparison of equilibrium binding constants (K_D , M) of DB2448 with mixed single/two G•C base pair(s) containing DNA sequences; Figure S4: Comparison of equilibrium binding constants (K_D , M) of DB2502 with pure AT and single/two G•C base pair(s) containing DNA sequences; Figure S5: Comparison of equilibrium binding constants (K_D , M) of DB2449 with mixed single/two G•C base pair(s) containing DNA sequences; Table S1: Thermal Melting Studies (ΔT_m , $^\circ\text{C}$) of the designed heterocyclic amidine compounds with pure A·T and mixed DNA sequences; Table S2: Biosensor-SPR equilibrium dissociation constants (K_D , nM) of DB2447 and analogues with pure A·T and mixed DNA sequences; Figure S6: ITC data for the titration of DB559 and DB2448 with AAATTT DNA at 100 mM NaCl at 25 $^\circ\text{C}$.

Author Contributions: W.D.W., D.W.B. and A.P. designed the study and experiments. A.A.F. and D.W.B. designed the novel compound syntheses. W.D.W., A.P. and A.A.F. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: National Institutes of Health Grant GM111749 (W.D.W. and D.W.B.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank the National Institutes of Health Grant GM111749 (W.D.W. and D.W.B.) for financial support and Arvind Kumar for providing some compounds.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Marky, L.A.; Blumenfeld, K.S.; Breslauer, K.J. Calorimetric and spectroscopic investigation of drug-DNA interactions. I. The binding of netropsin to poly d(AT). *Nucleic Acids Res.* **1983**, *11*, 2857–2870. [[CrossRef](#)] [[PubMed](#)]
2. Marky, L.A.; Snyder, J.G.; Breslauer, K.J. Calorimetric and spectroscopic investigation of drug-DNA interactions: II. Dipyridium binding to poly d(AT). *Nucleic Acids Res.* **1983**, *11*, 5701–5715. [[CrossRef](#)] [[PubMed](#)]
3. Marky, L.A.; Snyder, J.G.; Remeta, D.P.; Breslauer, K.J. Thermodynamics of drug-DNA interactions. *J. Biomol. Struct. Dyn.* **1983**, *1*, 487–507. [[CrossRef](#)] [[PubMed](#)]
4. Marky, L.A.; Breslauer, K.J. Origins of netropsin binding affinity and specificity: Correlations of thermodynamic and structural data. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 4359–4363. [[CrossRef](#)]
5. Breslauer, K.J.; Remeta, D.P.; Chou, W.Y.; Ferrante, R.; Curry, J.; Zaunczkowski, D.; Snyder, J.G.; Marky, L.A. Enthalpy-entropy compensations in drug-DNA binding studies. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8922–8926. [[CrossRef](#)]
6. Jin, R.; Breslauer, K.J. Characterization of the minor groove environment in a drug-DNA complex: Bisbenzimidazole bound to the poly[d(AT)].poly[d(AT)]duplex. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 8939–8942. [[CrossRef](#)]
7. Marky, L.A.; Patel, D.; Breslauer, K.J. Effect of tetramethylammonium ion on the helix-to-coil transition of poly(deoxyadenylthymidine): A nuclear magnetic resonance and calorimetric investigation. *Biochemistry* **1981**, *20*, 1427–1431. [[CrossRef](#)]
8. Pilch, D.S.; Kirolos, M.A.; Liu, X.; Plum, G.E.; Breslauer, K.J. Berenil [1,3-bis(4'-amidinophenyl)triazene] binding to DNA duplexes and to a RNA duplex: Evidence for both intercalative and minor groove binding properties. *Biochemistry* **1995**, *34*, 9962–9976. [[CrossRef](#)]

9. Kopka, M.L.; Goodsell, D.S.; Han, G.W.; Chiu, T.K.; Lown, J.W.; Dickerson, R.E. Defining GC-specificity in the minor groove: Side-by-side binding of the di-imidazole lexitropsin to C-A-T-G-G-C-C-A-T-G. *Structure* **1997**, *5*, 1033–1046. [[CrossRef](#)]
10. Reddy, B.S.; Sharma, S.K.; Lown, J.W. Recent developments in sequence selective minor groove DNA effectors. *Curr. Med. Chem.* **2001**, *8*, 475–508. [[CrossRef](#)]
11. O'Hare, C.C.; Mack, D.; Tandon, M.; Sharma, S.K.; Lown, J.W.; Kopka, M.L.; Dickerson, R.E.; Hartley, J.A. DNA sequence recognition in the minor groove by crosslinked polyamides: The effect of N-terminal head group and linker length on binding affinity and specificity. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 72–77. [[CrossRef](#)] [[PubMed](#)]
12. Zhang, Q.; Dwyer, T.J.; Tsui, V.; Case, D.A.; Cho, J.; Dervan, P.B.; Wemmer, D.E. NMR Structure of a cyclic polyamide-DNA complex. *J. Am. Chem. Soc.* **2004**, *126*, 7958–7966. [[CrossRef](#)] [[PubMed](#)]
13. Goodsell, D.S.; Ng, H.L.; Kopka, M.L.; Lown, J.W.; Dickerson, R.E. Structure of a dicationic monoimidazole lexitropsin bound to DNA. *Biochemistry* **1995**, *34*, 16654–16661. [[CrossRef](#)] [[PubMed](#)]
14. Quintana, J.R.; Lipanov, A.A.; Dickerson, R.E. Low-temperature crystallographic analyses of the binding of Hoechst 33258 to the double-helical DNA dodecamer C-G-C-G-A-A-T-T-C-G-C-G. *Biochemistry* **1991**, *30*, 10294–10306. [[CrossRef](#)]
15. Ming, X.; Ju, W.; Wu, H.; Tidwell, R.R.; Hall, J.E.; Thakker, D.R. Transport of dicationic drugs pentamidine and furamidine by human organic cation transporters. *Drug. Metab. Dispos.* **2009**, *37*, 424–430. [[CrossRef](#)]
16. Paine, M.F.; Wang, M.; Generaux, C.N.; Boykin, D.W.; Wilson, W.D.; Koning, H.P.; Olson, C.A.; Pohlig, G.; Burri, C.; Brun, R.; et al. Diamidines for human African trypanosomiasis. *Curr. Opin. Investig. Drugs* **2010**, *11*, 876–883.
17. Peregrine, A.S.; Mamman, M. Pharmacology of diminazene: A review. *Acta. Trop.* **1993**, *54*, 185–203. [[CrossRef](#)]
18. Elamin, E.A.; Homeida, A.M.; Adam, S.E.; Mahmoud, M.M. The efficacy of berenil (diminazene aceturate) against *Trypanosoma evansi* infection in mice. *J. Vet. Pharmacol. Ther.* **1982**, *5*, 259–265. [[CrossRef](#)]
19. Soeiro, M.N.; Werbovets, K.; Boykin, D.W.; Wilson, W.D.; Wang, M.Z.; Hemphill, A. Novel amidines and analogues as promising agents against intracellular parasites: A systematic review. *Parasitology* **2013**, *40*, 929–951. [[CrossRef](#)]
20. Paul, A.; Guo, P.; Boykin, D.W.; Wilson, W.D. A new generation of minor-groove-binding-heterocyclic diamidines that recognize G-C base pairs in an AT sequence context. *Molecules* **2019**, *24*, 946. [[CrossRef](#)]
21. Chai, Y.; Paul, A.; Rettig, M.; Wilson, W.D.; Boykin, D.W. Design and synthesis of heterocyclic cations for specific DNA recognition: From AT-rich to mixed-base-pair DNA sequences. *J. Org. Chem.* **2014**, *79*, 852–866. [[CrossRef](#)] [[PubMed](#)]
22. Paul, A.; Chai, Y.; Boykin, D.W.; Wilson, W.D. Understanding mixed sequence DNA recognition by novel designed compounds: The kinetic and thermodynamic behavior of azabenzimidazole diamidines. *Biochemistry* **2015**, *54*, 577–587. [[CrossRef](#)] [[PubMed](#)]
23. Paul, A.; Nanjunda, R.; Kumar, A.; Laughlin, S.; Nhili, R.; Depauw, S.; Deuser, S.S.; Chai, Y.; Chaudhary, A.S.; David-Cordonnier, M.H.; et al. Mixed up minor groove binders: Convincing A-T specific compounds to recognize a G-C base pair. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 4927–4932. [[CrossRef](#)]
24. Paul, A.; Kumar, A.; Nanjunda, R.; Farahat, A.A.; Boykin, D.W.; Wilson, W.D. Systematic synthetic and biophysical development of mixed sequence DNA binding agents. *Org. Biomol. Chem.* **2017**, *15*, 827–835. [[CrossRef](#)] [[PubMed](#)]
25. Guo, P.; Paul, A.; Kumar, A.; Farahat, A.A.; Kumar, D.; Wang, S.; Boykin, D.W.; Wilson, W.D. The Thiophene “Sigma-Hole” as a concept for preorganized, specific recognition of G-C base pairs in the DNA minor groove. *Chemistry* **2016**, *22*, 15404–15412. [[CrossRef](#)] [[PubMed](#)]
26. Guo, P.; Farahat, A.A.; Paul, A.; Harika, N.K.; Boykin, D.W.; Wilson, W.D. Compound shape Effects in minor groove binding affinity and specificity for mixed sequence DNA. *J. Am. Chem. Soc.* **2018**, *140*, 14761–14769. [[CrossRef](#)]
27. Guo, P.; Paul, A.; Kumar, A.; Harika, N.K.; Wang, S.; Farahat, A.A.; Boykin, D.W.; Wilson, W.D. A modular design for minor groove binding and recognition of mixed base pair sequences of DNA. *Chem. Commun.* **2017**, *53*, 10406–10409. [[CrossRef](#)]
28. Guo, P.; Farahat, A.A.; Paul, A.; Boykin, D.W.; Wilson, W.D. Engineered modular heterocyclic-diamidines for sequence-specific recognition of mixed AT/GC base pairs at the DNA minor groove. *Chem. Sci.* **2021**, *12*, 15849–15861. [[CrossRef](#)]
29. Boykin, D.W.; Kumar, A.; Xiao, G.; Wilson, W.D.; Bender, B.C.; McCurdy, D.R.; Hall, J.E.; Tidwell, R.R. 2,5-bis[4-(N-alkylamidino)phenyl]furans as anti-Pneumocystis carinii agents. *J. Med. Chem.* **1998**, *41*, 124–129. [[CrossRef](#)]
30. Das, B.P.; Boykin, D.W. Synthesis and antiprotozoal activity of 2,5-bis(4-guanylphenyl) furans. *J. Med. Chem.* **1977**, *20*, 531–536. [[CrossRef](#)]
31. Wilson, W.D.; Tanius, F.A.; Fernandez-Saiz, M.; Rigl, C.T. Evaluation of drug-nucleic acid interactions by thermal melting curves. *Methods Mol. Biol.* **1997**, *90*, 219–240. [[PubMed](#)]
32. Shi, X.C.; Chaires, J.B. Sequence- and structural-selective nucleic acid binding revealed by the melting of mixtures. *Nucleic Acids Res.* **2006**, *34*, e14. [[CrossRef](#)] [[PubMed](#)]
33. Nguyen, B.; Neidle, S.; Wilson, W.D. A role for water molecules in DNA-ligand minor groove recognition. *Acc. Chem. Res.* **2009**, *42*, 11–21. [[CrossRef](#)] [[PubMed](#)]
34. Wilson, W.D.; Nguyen, B.; Tanius, F.A.; Mathis, A.; Hall, J.E.; Stephens, C.E.; Boykin, D.W. Dications that target the DNA minor groove: Compound design and preparation, DNA interactions, cellular distribution and biological activity. *Curr. Med. Chem. Anticancer Agents* **2005**, *5*, 389–408. [[CrossRef](#)]
35. Munde, M.; Wang, S.; Kumar, A.; Stephens, C.E.; Farahat, A.A.; Boykin, D.W.; Wilson, W.D.; Poon, G.M. Structure-dependent inhibition of the ETS-family transcription factor PU.1 by novel heterocyclic diamidines. *Nucleic Acids Res.* **2014**, *42*, 1379–1390. [[CrossRef](#)]

36. Rohs, R.; West, S.M.; Sosinsky, A.; Liu, P.; Mann, R.S.; Honig, B. The role of DNA shape in protein-DNA recognition. *Nature* **2009**, *461*, 1248–1253. [[CrossRef](#)]
37. Zhou, T.; Yang, L.; Lu, Y.; Dror, I.; Dantas-Machado, A.C.; Ghane, T.; Di Felice, R.; Rohs, R. DNAshape: A method for the high-throughput prediction of DNA structural features on a genomic scale. *Nucleic Acids Res.* **2013**, *4*, W56–W62. [[CrossRef](#)]
38. Liu, Y.; Chai, Y.; Kumar, A.; Tidwell, R.R.; Boykin, D.W.; Wilson, W.D. Designed compounds for recognition of 10 base pairs of DNA with two AT binding sites. *J. Am. Chem. Soc.* **2012**, *134*, 5290–5299. [[CrossRef](#)]
39. Nguyen, B.; Tanius, F.A.; Wilson, W.D. Biosensor-surface plasmon resonance: Quantitative analysis of small molecule–nucleic acid interactions. *Methods* **2007**, *42*, 150–161. [[CrossRef](#)]
40. Nanjunda, R.; Munde, M.; Liu, Y.; Wilson, W.D. *Methods for Studying DNA/Drug Interactions*; Chapter 4; Wanunu, M., Tor, Y., Eds.; CRC Press-Taylor & Francis Group: Boca Raton, FL, USA, 2011.
41. DeHaseth, P.L.; Lohman, T.M.; Record, M.T., Jr. Nonspecific interaction of iac repressor with DNA: An association reaction driven by counterion release. *Biochemistry* **1977**, *16*, 4783–4790. [[CrossRef](#)]
42. Wang, S.; Kumar, A.; Aston, K.; Nguyen, B.; Bashkin, J.K.; Boykin, D.W.; Wilson, W.D. Different thermodynamic signatures for DNA minor groove binding with changes in salt concentration and temperature. *Chem. Commun.* **2013**, *49*, 8543–8545. [[CrossRef](#)] [[PubMed](#)]
43. Manning, G.S. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **1978**, *11*, 179–246. [[CrossRef](#)] [[PubMed](#)]
44. Chaires, J.B. A thermodynamic signature for drug-DNA binding mode. *Arch. Biochem. Biophys.* **2006**, *453*, 26–31. [[CrossRef](#)] [[PubMed](#)]
45. Privalov, P.L.; Dragan, A.I.; Crane-Robinson, C. Interpreting protein/DNA interactions: Distinguishing specific from non-specific and electrostatic from non-electrostatic components. *Nucleic Acids Res.* **2011**, *39*, 2483–2491. [[CrossRef](#)]
46. Record, M.T., Jr.; Anderson, C.F.; Lohman, T.M. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: The roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* **1978**, *11*, 103–178. [[CrossRef](#)]
47. Munde, M.; Lee, M.; Neidle, S.; Arafa, R.; Boykin, D.W.; Liu, Y.; Bailly, C.; Wilson, W.D. Induced fit conformational changes of a “reversed amidine” heterocycle: Optimized interactions in a DNA minor groove complex. *J. Am. Chem. Soc.* **2007**, *129*, 5688–5698. [[CrossRef](#)]
48. Wang, S.; Aston, K.; Koeller, K.J.; Harris, G.D., Jr.; Rath, N.P.; Bashkin, J.K.; Wilson, W.D. Modulation of DNA-polyamide interaction by β -alanine substitutions: A study of positional effects on binding affinity, kinetics and thermodynamics. *Org. Biomol. Chem.* **2014**, *12*, 7523–7536. [[CrossRef](#)]
49. Liu, Y.; Collar, C.J.; Kumar, A.; Stephens, C.E.; Boykin, D.W.; Wilson, W.D. Heterocyclic diamidine interactions at AT base pairs in the DNA minor groove: Effects of heterocycle differences, DNA AT sequence and length. *J. Phys. Chem. B.* **2008**, *112*, 11809–11818. [[CrossRef](#)]
50. Buurma, N.J.; Haq, I. Calorimetric and spectroscopic studies of Hoechst 33258: Self-association and binding to non-cognate DNA. *J. Mol. Biol.* **2008**, *381*, 607–621. [[CrossRef](#)]
51. Ferreira, J.M.; Sheardy, R.D. Linking temperature, cation concentration and water activity for the B to Z conformational transition in DNA. *Molecules* **2018**, *23*, 1806. [[CrossRef](#)]
52. Kumar, S.; Xue, L.; Arya, D.P. Neomycin-neomycin dimer: An all-carbohydrate scaffold with high affinity for AT-rich DNA duplexes. *J. Am. Chem. Soc.* **2011**, *133*, 7361–7375. [[CrossRef](#)] [[PubMed](#)]
53. Laughlin, S.; Wilson, W.D. May the best molecule win: Competition ESI mass spectrometry. *Int. J. Mol. Sci.* **2015**, *16*, 24506–24531. [[CrossRef](#)] [[PubMed](#)]
54. Laughlin, S.; Wang, S.; Kumar, A.; Farahat, A.A.; Boykin, D.W.; Wilson, W.D. Resolution of mixed site DNA complexes with dimer-forming minor-groove binders by using electrospray ionization mass spectrometry: Compound structure and DNA sequence effects. *Chemistry* **2015**, *21*, 5528–5539. [[CrossRef](#)] [[PubMed](#)]
55. Tanius, F.; Wilson, W.D.; Wang, L.; Kumar, A.; Boykin, D.W.; Marty, C.; Baldeyrou, B.; Bailly, C. Cooperative dimerization of a heterocyclic diamidine determines sequence-specific DNA recognition. *Biochemistry* **2003**, *42*, 13576–13586. [[CrossRef](#)] [[PubMed](#)]
56. Wang, L.; Kumar, A.; Boykin, D.W.; Bailly, C.; Wilson, W.D. Comparative thermodynamics for monomer and dimer sequence-dependent binding of a heterocyclic dication in the DNA minor groove. *J. Mol. Biol.* **2002**, *317*, 361–374. [[CrossRef](#)] [[PubMed](#)]
57. Harika, N.K.; Wilson, W.D. Bound compound, interfacial water, and phenyl ring rotation dynamics of a compound in the DNA minor groove. *Biochemistry* **2018**, *57*, 5050–5057. [[CrossRef](#)]
58. Athri, P.; Wilson, W.D. Molecular dynamics of water-mediated interactions of a linear benzimidazole-biphenyl diamidine with the DNA minor groove. *J. Am. Chem. Soc.* **2009**, *131*, 7618–7625. [[CrossRef](#)]