Published in final edited form as:
J Phys Chem B. 2008 September 18; 112(37): 11809–11818. doi:10.1021/jp804048c.

Heterocyclic Diamidine Interactions at AT Base Pairs in the DNA Minor Groove: Effects of Heterocycle Differences, DNA AT Sequence and Length

Yang Liu, **Catharine J. Collar**, **Arvind Kumar**, **Chad E. Stephens**, **David W. Boykin**, and **W. David Wilson***
Department of Chemistry, Georgia State University, Atlanta, GA 30302, USA

Abstract

Given the increasing significance of diamidines as DNA–targeted therapeutics and biotechnology reagents, it is important to establish the variations in thermodynamic quantities that characterize the interactions of closely related compounds to different sequence AT binding sites. In this study, an array of methods including biosensor–surface plasmon resonance (SPR), isothermal titration microcalorimetry (ITC), circular dichroism (CD), thermal melting (Tm) and molecular modeling have been used to characterize the binding of dicationic diamidines related to DB75 (amidine-phenyl-furan-phenyl-amidine) with alternating and nonalternating AT sequences. Conversion of the central furan of DB75 to other similar groups, such as thiophene or selenophene, can yield compounds with increased affinity and sequence binding selectivity for the minor groove. Calorimetric measurements revealed that the thermodynamic parameters (ΔG, ΔH, ΔS) that drive diamidine binding to alternating and nonalternating oligomers can be quite different and depend on both DNA sequence and length. Small changes in a compound can have major effects on DNA interactions. By choosing an appropriate central group it is possible to "tune" the shape of the molecule to match DNA for enhanced affinity and sequence recognition.

Keywords

DB75; alternating and nonalternating AT DNAs; thermodynamics; isothermal titration microcalorimetry; biosensor–surface plasmon resonance

Introduction

Synthetic organic dications that bind reversibly to the DNA minor groove provide important examples of clinically useful therapeutic agents as well as specific nucleic acid florescence stains. A prodrug of the diphenylfuran diamidine, furamidine, DB75 in Figure 1, has progressed to phase III clinical trials against human African trypanosomiasis, sleeping sickness.[1,2] The prodrug concept provides new opportunities for development of orally effective agents against lethal parasitic diseases.[3,4] Such compounds also have the potential to provide new types of cellular biotechnology regents and new approaches to probe the molecular basis of DNA recognition and control of gene expression.[3] Other dications that are based on the DB75 lead compound are being tested against trypanosomes as well as other diseases in humans. Fluorescence microscopy analysis of trypanosomes from animals treated with the dications shows that furamidine and related compounds target the parasite kinetoplast DNA with

***Corresponding author:** W. David Wilson, Department of Chemistry, Georgia State University, Atlanta, GA 30302-4098, USA, Tel: +1-404-413-5503, Fax: +1-404-413-5505, Email: wdw@gsu.edu.

subsequent, time-dependent, destruction of the kinetoplast.[5,6] Although DNA binding is clearly involved, not all DNA binding diamidines are active and it is important to understand the effects on DNA of different compound features and how the differences lead to variations in biological activity.

In an x-ray structure of the furamidine complex with d(CGCGAATTCGCG)$_2$, the compound is deep in the minor groove in the –AATT– site and in close contact with the edges of AT base pairs at the floor of the groove.[7] The amidine groups form hydrogen bonds with the AT base pairs but such close contact is sterically prevented by the 2-NH_2 group of G in GC containing sequences. For compounds such as DB75, a site of four or more AT base pairs in length is required for strong binding. Recent results show that the affinity of DB75 for different AT sites can vary by a surprisingly large amount.[7,8] The diversity of affinities for minor groove compounds with AT sequences, however, has not been extensively investigated. Well-characterized diamidine derivatives such as berenil and DAPI have significant differences in their affinities for AT sites of different sequence.[9]

For detailed understanding of the biological mechanism of the diamidines as well as for drug design and development progress, it is important to understand how small changes in compound structure and properties affect DNA interactions. Such information can provide new insights into how binding of the compounds to the parasite kinetoplast DNA can yield biological activity. There are multiple AT sequences in kDNA that are targeted by the diamidines and we wish to model these interactions. The kinetoplast of *T. brucei*, the cause of sleeping sickness, for example, has approximately 10,000 DNA minicircles of 1000 base pairs that are interlinked with DNA maxicircles.[10] The minicircles are AT rich and contain phased A–tracts that can lead to significant bending of the DNA helix. The AT-rich minicircle DNAs forms an attractive target for AT specific DNA binding diamidines. Interference with transcription and replication of minicircles by the compounds is a possible mechanism for destruction of the complex interlocked kinetoplast structure that leads to death of the trypanosome. A–tract bent DNA was discovered in early studies of minicircle DNAs from both trypanosomes and Leishmania organisms.[11–13] Clearly, to better understand the biological action of the diamidines in parasites and to design a broader range of improved drugs, it is essential to investigate the interaction of diamidines related to DB75 with DNAs of different AT sequences.

The approach that we describe here involves making single atom changes in the furan of DB75 by substitution of the O for S and Se (Figure 1) and analyzing the interaction of the compounds with different AT DNA sequences. Our goal is to understand the molecular recognition of DNA and use this understanding to develop a more detailed model to better understand the antiparasitic biological action of this class of dicationic heterocycles. We have used DNAs with alternating and nonalternating AT base pairs as our model systems and the interactions of the compound set with AT polymers and oligomers have been investigated. The results for binding to the alternating and nonalternating AT sequences are quite different at the fundamental thermodynamic level.

Materials and methods

Materials

DB1213, DB351 and DB75 were synthesized as previously described.[1,14] Their purity was verified by NMR and elemental analysis. The double stranded polymers poly(dA)·poly(dT), poly(dA-dT)·poly(dA-dT) were purchased from Pharmacia (U. S. A.). The hairpin oligomers used in this study were A_{15} [d (CCAAAAAAAAAAAAAAAGCCCCCGCTTTTTTTTTTTTTTTGG)], A_8 [d (CAAAAAAAACCCCCTTTTTTTTG)], A_4 [d(CCAAAAGCCCCCGCTTTTGG)], $(AT)_7$ [d (CCATATATATATATATAGCCCCCGCTATATATATATATATGG], $(AT)_4$ [d

(CATATATATCCCCATATATATG)], and ATAT [d(CCATATGCCCCCGCTATAGG] with the hairpin loop sequences underlined. The above hairpin oligomers are all from Integrated DNA Technologies, Inc. with reverse phased HPLC purification and mass spectrometry characterization. In SPR experiments, 5'-Biotin labeled hairpin DNA oligomers were used. The MES10 buffers used in these experiments contained 0.01 M [2-(N-morpholino) ethanesulfonic acid] (MES), 0.001 M EDTA, 0.1 M NaCl, pH 6.25.

UV melting studies

Ultraviolet melting curves were determined in 1 cm path length quartz cells using a Cary 300 UV–Visible spectrophotometer (Varian Inc., Palo Alto, CA), equipped with a thermoelectric temperature controller. Absorbance versus temperature profiles were measured at 260 nm with a heating rate of 0.5 °C/min. Experiments were generally conducted at a concentration of 1.5×10^{-6} M oligomers for hairpin DNA. Thermal melting experiments for the DNA–compound complexes were conducted as a function of different ratios. Melting temperature (*T*m) was taken as the temperature of half–dissociation of the DNA duplex or hairpin and was obtained from the maximum of the first derivative d*A*/d*T* plots (where *A* is absorbance and *T* is temperature).

CD titration studies

CD spectra were recorded using a Jasco J−810 instrument with a 1cm cell and a scan speed of 50 nm/min with a response time of 1 s. The spectra from 500 to 220 nm were averaged over five scans. A buffer baseline scan was collected in the same cuvette and subtracted from the average scan for each sample. The titration experiments were performed at 25 °C. The desired ratios of compound to DNA were obtained by adding compound to the cell containing a constant amount of DNA. Data processing and plotting were performed with Kaleidagraph software.

ITC thermodynamic studies

The ITC experiments were performed with a MicroCal VP–ITC (MicroCal Inc., Northampton, MA, USA) interfaced with a computer for instrument control and data collection with Origin 7.0 software. Binding enthalpies for nonalternating and alternating oligmers [A_{15}, A_8, A_4 hairpins and $(AT)_7$, $(AT)_4$, ATAT hairpins] and polymers [poly(dA)·poly(dT) and poly(dA-dT)·poly(dA-dT)] binding with different compounds were determined using a protocol with a high DNA concentration in the calorimeter to ensure that all added compound is effectively bound after each addition in the low ratio region.[15,16] Specifically, for oligomeric DNA, a 0.03 mM DNA solution was loaded into the sample cell and a 0.15 mM compound solution in MES10 buffer was added. For polymeric DNA, a 0.24 mM base pair DNA solution was loaded into the sample cell and a 0.08 mM compound solution was added. Usually 20 injections of 7 μl were done with 300 s between injections to ensure equilibration. For the additions where all compound is bound the heat of reaction (Δ*H*) is obtained by integration of the peaks after each injection. The dilution heats, determined by injecting drug solution into the same sample cell loaded with buffer alone, were subtracted from the Δ*H* value determined for addition into DNA solution to render a corrected value for the binding–induced enthalpy change. An average overall low ratio values gives the Δ*H* of binding.

Biosensor–SPR studies

SPR measurements were performed with a four–channel BIAcore 2000 optical biosensor system (BIAcore Inc.). 5'-biotin labeled DNA samples [A_{15}, A_8, A_4 hairpins and $(AT)_7$, $(AT)_4$, ATAT hairpins] were immobilized onto streptavidin–coated sensor chips (BIAcore SA) as previously described.[17] Three flow cells were used to immobilize the DNA oligomer samples, while a fourth cell was left blank as a control. The SPR experiments were performed

in filtered, degassed MES10 buffer with 5×10^{-3}% v/v Surfactant P20. Steady state binding analysis was performed with multiple injections of different compound concentrations over the immobilized DNA surface at a flow rate of 25 μl/min and 25 °C. Solutions of known ligand concentration were injected through the flow cells until a constant steady–state response was obtained. Compound solution flow was then replaced by buffer flow resulting in dissociation of the complex. The reference response from the blank cell was subtracted from the response in each cell containing DNA to give a signal (*RU*, response units) that is directly proportional to the amount of bound compound. The predicted maximum response per bound compound in the steady-state region (RU_{max}) was determined from the DNA molecular weight, the amount of DNA on the flow cell, the compound molecular weight, and the refractive index gradient ratio of the compound and DNA, as previously described.[17,18] The number of binding sites and the equilibrium constant were obtained from fitting plots of *RU* versus C_{free}. Binding results from the SPR experiments were fit with two– (K_3, $K_4 = 0$), three– ($K_4 = 0$) or four–sites interaction models:

$$r=(K_1{}^*C_{free}+2{}^*K_1{}^*K_2{}^*C_{free}{}^2+3{}^*K_1{}^*K_2{}^*K_3{}^*C_{free}{}^3+4{}^*K_1{}^*K_2{}^*K_3{}^*K_4{}^*C_{free}{}^4)$$
$$/(1+K_1{}^*C_{free}+K_1{}^*K_2{}^*C_{free}{}^2+K_1{}^*K_2{}^*K_3{}^*C_{free}{}^3+K_1{}^*K_2{}^*K_3{}^*K_4{}^*C_{free}{}^4)$$

where *r* represents the moles of bound compound per mole of DNA hairpin duplex, K_1, K_2, K_3, K_4 are macroscopic binding constants, and C_{free} is the free compound concentration in equilibrium with the complex. K_{mean} is calculated as the average of the binding constants obtained.

Molecular modeling

Geometry optimized structures for DB75, DB351 and DB1213 were calculated at the Hartree-Fock 631G** level with the Spartan 04 software package.[19,20] Docking studies were performed with SYBYL 7.3 software package on a Fedora Core 5 Linux Workstation.[20,21] DNA duplexes d(CCAAAAGC)·d(GCTTTTGG) and d(CCATATGC)·d(GCATATGG) with the recognition sites –AAAA– and –ATAT– were constructed in the Biopolymer module. A crystal structure, 227D from the protein data bank, of $d(CGCGAATTCGCG)_2$ with bound DB75 was selected as a template. The designed DNA was aligned to the DNA of the crystal structure and the single atom changes were made to the docked structure to create DB351 and DB1213.

Atoms of the ligand to be docked were assigned Gasteiger-Marsili charges and then minimized using the Tripos force field until a terminating conjugate gradient of 0.01 kcal/mol Å was reached. The DNA from 227D was deleted leaving the designed DNA in complex with the respective ligand. Before each docking, the ligand of interest (DB75, DB351, or DB1213) was moved to a second memory location, so that the ligand could move independently of the DNA. For each docking, the genetic algorithm of the Flexidock module was employed implementing five different random numbers and a large number of generations.[19]

To ensure that low-energy complexes were obtained, 348,000 generations, 3000 × 116 rotatable bonds, were used for each docking. The suggested minimum number of generations to be used in Flexidock studies should be the number of rotatable bonds plus 6 times 500. To ensure that we received low-energy complexes, we recalculated one compound from each prior docking by implementing five different random numbers for 61,000 generations. Both the DNA and the bound compounds were permitted torsional flexibility in the docking process. Atomic charges were calculated using the Kollman All-Atom protocol for the DNA, and all of the hydrogen bond sites were marked for the DNA and compounds. Each docking generated 20 low-energy structures for a total of 1200 structures, 2 × 6 DNA-ligand complexes × 5 random number starting points × 20 compounds produced for docking, were obtained.

Results

Biosensor–SPR: Quantitative binding affinity and stoichiometry

To quantitatively compare the interactions of the compounds of Figure 1 with alternating and nonalternating DNA sequences of different lengths, biosensor-SPR experiments were carried out as previously described.[22] Representative SPR sensorgrams for the binding of DB1213 to the immobilized alternating and nonalternating DNA are compared in Figure 2 (a). Essentially, the same moles of DNA oligomer were immobilized on the surface of sensor chips so that the sensorgram saturation levels can be compared directly for stoichiometry differences. The differences in interaction strength and binding stoichiometry for all three compounds can be visualized in a plot of *r* (moles of compound bound/mole of hairpin DNA) versus C_{free}, the free compound concentration [Figure 2 (b)]. The data for compound binding were best fit with a one–site binding model for A_4 and ATAT, a two–site binding model for A_8 and $(AT)_4$ hairpin and a four–site binding model for A_{15} and $(AT)_7$ hairpin.

The K_{mean} values for these compounds binding with alternating and nonalternating DNA sequences at different lengths are shown in Table 1. As can be seen, the binding constants for DB1213 with nonalternating DNA oligomers (A_4, A_8, A_{15}) increased significantly (almost 10 times higher) with increasing DNA length. Similar behavior is seen for DB1213 binding to the alternating DNA oligomers [ATAT, $(AT)_4$, $(AT)_7$] with an approximately five fold increase with increasing length. For all lengths the K for binding of DB1213 to the nonalternating sequence is approximately three fold higher than with the alternating sequence. When the central atom of DB1213 is converted from C–Se–C to C–S–C (DB351) or to C–O–C (DB75), decreased affinity and sequence binding selectivity are obtained and the effect is larger for the nonalternating sequence. For example, the binding constant for DB75 with the A_{15} hairpin is almost 10 times lower than that for DB1213. Thus, the binding affinity with AT sequences for compounds in this series increases with DNA length, central atom size and from alternating to nonalternating sequences.

Circular Dichroism: Probing the binding mode and binding ratio

CD titration experiments as a function of compound concentration were utilized to monitor the binding mode and saturation limit for compound binding with alternating and nonalternating DNA sequences of different lengths (Figure 3). CD spectra monitor the asymmetric environment of the compounds when bound to DNA and therefore can be used to obtain information on the binding mode. The free compounds do not exhibit CD signals, however, upon addition of the compounds to DNA, substantial positive induced CD signals arise between 300 and 400 nm. These large positive induced CD signals for the complexes are characteristic patterns for a minor grove binding mode in AT sequences.[23,24] The induced CD signals were also plotted against compound–DNA ratio (inset in Figure 3). The saturated maximum compound to DNA ratios obtained from the inset plot are 4 compounds / hairpin for A_{15} and $(AT)_7$, 2 compounds / hairpin for A_8 and $(AT)_4$ and 1 compound / hairpin for A_4 and ATAT respectively, and these values are consistent with results from SPR experiments.

Thermal Melting: Relative affinity ranking

DNA thermal melting studies provide a rapid, qualitative method for ranking compounds according to their binding affinity. To obtain more information on the different complexes of the compounds with long nonalternating and alternating sequences [A_{15} and $(AT)_7$], relative stabilities of the compound–DNA complexes at saturation binding ratio 4:1 were evaluated by temperature–dependent UV absorbance measurements at 260 nm (Figure 4). Large increases in melting temperature are observed for all of the compounds with A_{15} and $(AT)_7$. For the A_{15} sequence converting from C–O–C (DB75) to C–S–C (DB351) makes small change in the ΔTm values but the conversion to C–Se–C (DB1213) results in a larger *T*m increase. In

additional *T*m determinations with poly(dA) poly(dT), DB1213 again had a significantly larger Δ*T*m value (Figure S1, Supplementary material) than the other compounds. For the $(AT)_7$ sequence, the central atom conversion results in smaller changes of Δ*T*m values. These observations are in agreement with the SPR results.

Isothermal Titration Calorimetry (ITC): Detailed thermodynamics of the interaction

To probe the detailed energetic basis of the interaction differences between DB1213 and the other two compounds in more detail, ITC experiments were conducted. Binding enthalpies (Δ*H*) for different length DNA hairpins with these compounds were determined at low binding ratios to directally obtain the Δ*H* for binding.[16] Figure 5 shows representative primary data from the addition of nonalternating (A-I) and alternating (J-R) DNA sequences with different compounds at 25 °C, respectively. Integration of the peaks followed by normalization for the number of moles of added compound and subtraction of the blank heat provides a direct estimate of the binding enthalpy. The average binding enthalpy values for the alternating and nonalternating DNA oligomers binding with these compounds are shown in Table 1.

Binding enthalpies for nonalternating and alternating polymers [poly(dA)·poly(dT) and poly (dA-dT)·poly(dA-dT)] with these compounds were also determined and a comparison of the ITC results for nonalternating and alternating polymers and oligomers [A_{15} and $(AT)_7$] with different compounds is shown in Figure 6. The values with the oligomer duplexes closely match the values for poly(dA) poly(dT) and poly(dA-dT) poly(dA-dT). The binding enthalpies for the shorter DNA hairpins, however, decrease significantly relative to the polymers and 15mer hairpin sequences (Figure 5). The average binding enthalpies are substantially more negative than with the longer sequences. These results suggest that the A_{15} and $(AT)_7$ hairpins are good models to mimic results with polymer DNAs. From the Gibbs free energy (obtained from SPR experiments) and enthalpy values determined above, the binding entropy (Δ*S*) of complex formation could be calculated from $\Delta G = \Delta H - T\Delta S$. The complete thermodynamics parameters for all compounds and DNAs are summarized in Table 1. To visualize the different thermodynamic quantities more easily, the overall thermodynamic profiles for the interactions of alternating and nonalternating DNA sequences of different length with these compounds are compared in Figure 6. The thermodynamics driving these compounds to bind to alternating and nonalternating AT–tracts at 25 °C are quite different and are DNA sequence and length dependent.

Molecular modeling: providing possible explanations for binding mode

The compound *ab initio* calculations produced structures for comparison to the models below and are similar to those found in the docking study. The effects of the central atom changes on the compound structures can be seen in the molecular models in Figure S2, Supplementary material. The interactions observed in these DNA duplexes in complex with DB1213, DB351, and DB75 were analyzed for hydrogen bonds, conformational angles and fit, as well as, lowest energy conformation values. For d(CCAAAAGC)·d(GCTTTTGG), the most stable complex consists of the DB1213, which is locked down into essentially a single conformation. This complex displays hydrogen bonding from –NH groups of the DB1213 amidines to adenosine –N and thymine –O atoms in the minor grove of the DNA. The conformation of DB1213 when bound to DNA appears to be determined by the selenium atom, which causes more rotation in the phenyl rings than the S and O. This DB1213 compound fits well into the minor groove of DNA and the complex formed has a low energy value (Table S1, Supplementary material). DB351 is more mobile in the binding site and the sulfur atom causes less rotation of the phenyl substituates than that seen in DB1213 (Figure 8). This complex shows very similar binding to that of DNA-DB1213, although DB351 does not conform as well to the DNA. In the lowest energy conformation its energy value is 92% of the DB1213 value. The complex with DB75, which is essentially a planar compound, provides several conformations of interaction with

only two hydrogen bonds from the –NH groups of the amidines at each end of the compound. In this dataset, DNA-DB75 has the most positive value for the lowest energy conformation, 89% of the DB1213 value. The energy values from the docking can be compaired to the ΔG of binding data in Table 1, where it is shown that the ΔG of A_4-DB351 is 94% and A_4-DB351 is 91% of the value A_4-DB1213. The experemental binding data is in good agreement with the modeling values. These three compounds did not bind as well to d(CCATATGC)·d (GCATATGG) (Figure S3, Supplementary material). The rotation of the phenyl groups is more pronounced in the alternating sequence and this changes the binding ability in terms of shape and bond length for – NH amidine hydrogen bonds. The curvature and rotational change differences of DB1213, DB351 and DB75 lead to slightly different binding conformations with different hydrogen bond lengths and binding affinity. All of the calculated energy values are summarized in Table S1 (Supplementary material).

Discussion

Because of the interest in development of synthetic compounds that have therapeutic effects and that can modify gene expression, there have been many studies on the interaction of small organic compounds with DNA. Detailed quantitative studies on the interaction of series of closely related compounds, however, are more rare. We have found large thermodynamic differences for the interaction of DB75 with AT sequences that have different length or that have either alternating or nonalternating base pairs.[22,25] For optimized antiparasitic drug design as well as improved understanding of the energetic basis for the ability of diamidines to distinguish among different sequences in AT binding sites such as those in kDNA, we are extending these quantitative studies to additional AT specific minor groove compounds that are related to DB75.

Multiple AT sites in parasitic mitochondrial kinetoplast DNA have been shown to be the biological target of DB75 and other diamidines,[2] and the studies described here were designed to determine how small variations in compound structure at the central five-member ring of DB75 (Figure 1) affect the interactions of the compounds at different AT sequences. The diamidine AT binding sites in kinetoplast DNA range from four base pairs to longer sequences and they can extend to approximately 20 base pairs. The sequences range from A-tracts to mixed A and T sequences.[26–28] To model these interactions and understand their differences, we have used sequences of A-tracts and alternating AT base pairs of length from 4 to 15 base pairs. The 15 base pair sequence behaves in a similar manner to polymer DNA and it is not necessary to go to longer oligomer sequences.

DB75 has been crystallized with $d(CGCGAATTCGCG)_2$ [7,29] and it binds in the minor groove at the –AATT– site with stabilization of the complex by hydrogen bonding, van der Waals and electrostatic interactions. The curvature of the compound allows it to slide to the bottom of the minor groove and the twist of the diamidines allows them to hydrogen bond with terminal T bases in the –AATT– sequences on opposite strands of the duplex.[7,29] In agreement with this binding mode, CD studies indicate that DB75, DB351 and DB1213 bind in a similar manner to the minor groove of the alternating and nonalternating AT sequences of different length (Figure 3).[24]

The single atom changes, O to S to Se, cause significant quantitative differences in the interaction of the compounds with AT sequences. In general, as the bond angles and amidine positions change from O to Se as shown in Figure S2 (Supplementary material), the affinity of the compound for all AT sequences increases (Figure 7 and Table 1). DB75 has similar affinities for the alternating and nonalternating AT sites, while the Se compound, DB1213, binds to all sequences more strongly than DB75, and its affinity for nonalternating AT sequences is greater than with the alternating AT.

ITC results show much larger differences in the interactions of the compounds with the alternating and nonalternating AT sites than are seen in the Gibbs energy differences (Table 1 and Figure 7). DB1213, for example, has a $\Delta H = -1.3$ kcal/mole at 25 °C with the A15 sequence but a more exothermic $\Delta H = -7.2$ kcal/mole with $(AT)_7$. The ΔG difference for these two sequences is less than one kcal/mole. Interestingly, DB75 has more endothermic binding enthalpies of 3.0 and −4.4 kcal/mole, respectively, for the same two sequences, while DB351 has values between those for the O and Se compounds. Two major trends are observed in the binding enthalpy variations for the compounds and DNA sequences. The enthalpies become more negative in going from DB75 to DB1213 and they become more negative in going from the longer to the shorter sequences. Binding of all compounds is almost completely driven by entropy with A15 but is more balanced between enthalpy and entropy with $(AT)_7$ and with the shorter A4 sequence. Binding of all compounds to –ATAT– is almost completely driven by enthalpy at 25 °C.

We believe that both of these observed trends in binding thermodynamics have a common explanation. The narrow minor groove in nonalternating AT sequences has an array of tightly bound water molecules and release of at least some of this high entropy water can account for the entropy differences on binding. Release of less tightly bound water from the wider minor groove in alternating AT sequences does not provide as large an entropy component to the Gibbs energy. The binding also has a more favorable binding enthalpy for the shorter sequences and this agrees with a major water release component to the entropy. The hydration structure will become more stable in longer sequences and will provide a larger favorable entropy term for the binding energetics. The release of the high entropy water comes at the cost of significant enthalpy for hydrogen bond breaking and as a result, the binding enthalpy will be lower for these systems. The binding of the diamidines to the non–alternating and alternating AT–tracts thus results in distinctive, sequence dependent, thermodynamic profiles.

Single heteroatom changes at the central five-atom ring (Figure 1) are particularly informative because they can cause a significant change in the relative positions of the amidines with little change in the properties of the phenyl-amidine group.[30] Significant differences in amidine positions from DB75 to DB1213 can be seen in Figure S2 (Supplementary materials). As the heteroatom size increases, the bond angles between the five and six member aromatic systems increase, the overall curvature of the compounds is decreased and the twist between the five and six member rings increases.

In the docking investigation the optimal models show that central atom size plays an important role in sequence recognition for the binding of compounds to DNA. As with the *ab initio* results, when the furan oxygen of DB75 is changed to the larger selenium atom to form DB1213, compound curvature decreases and the rotation of the phenyl groups increases. It is possible that hydrophobicity is playing a key role in the binding of these compounds. The larger Se is also more hydrophobic and this could improve the interaction of the relatively nonpolar environment of DB1213 with AT rich sequences of DNA.[31] Groove widths and depths also play an important role in these compounds ability to form hydrogen bonds to the DNA. For the DNA sequence with recognition site –AAAA– the groove is slightly narrower and deeper than the DNA groove of the DNA with recognition site –ATAT–. This allows DB1213 to fit well within the minor grove of the DNA with the recognition site –AAAA–. When the same compound is observed in complex with the DNA site –ATAT–, the compound is closer to the sugars of the bases and has much more room to hydrogen bond nonspecifically and more weakly to sugars or within the groove. Nonspecific hydrogen bonding may be one of the causes for loss of affinity to the DNA with recognition site –ATAT–.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

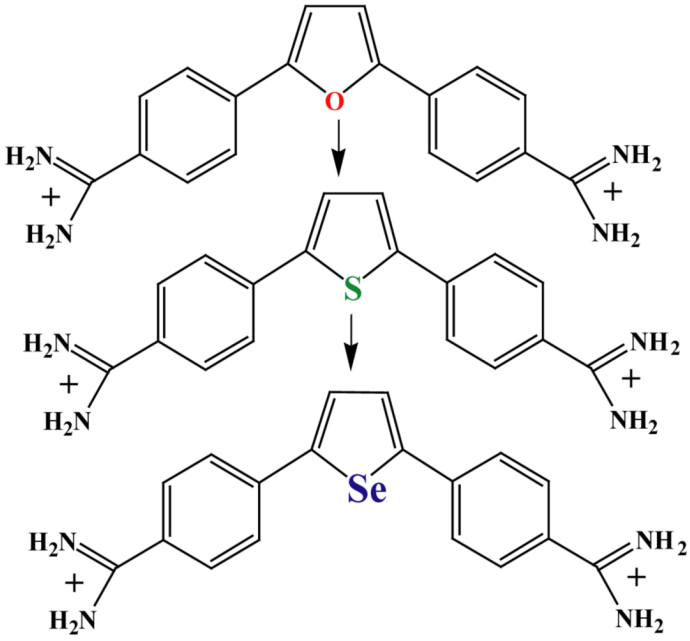
This research was supported by the National Institutes of Health Grant AI064200 (to WDW and DWB) and by a Georgia State University Molecular Basis of Disease Fellowship (to CJC). Equipment purchases were partially funded by the Georgia Research Alliance.

References

1. Tidwell, RRB.; DW. DNA and RNA Binders: from Small Molecules to Drugs; Demeunynck. Bailly, M.; Wilson, C.; WD, editors. Wiley-VCH; 2003. p. 414
2. Wilson WD, Nguyen B, Tanious FA, Mathis A, Hall JE, Stephens CE, Boykin DW. Curr. Med. Chem. Anticancer Agents 2005;5:389. [PubMed: 16101490]
3. Wilson WD, Tanious FA, Mathis A, Tevis D, Hall JE, Boykin DW. Biochimie 2008;90:999. [PubMed: 18343228]
4. Delespaux V, de Koning HP. Drug Resist Updat 2007;10:30. [PubMed: 17409013]
5. Mathis AM, Bridges AS, Ismail MA, Kumar A, Francesconi I, Anbazhagan M, Hu Q, Tanious FA, Wenzler T, Saulter J, Wilson WD, Brun R, Boykin DW, Tidwell RR, Hall JE. Antimicrob Agents Chemother 2007;51:2801. [PubMed: 17517831]
6. Mathis AM, Holman JL, Sturk LM, Ismail MA, Boykin DW, Tidwell RR, Hall JE. Antimicrob Agents Chemother 2006;50:2185. [PubMed: 16723581]
7. Laughton CA, Tanious F, Nunn CM, Boykin DW, Wilson WD, Neidle S. Biochemistry 1996;35:5655. [PubMed: 8639524]
8. Nguyen B, Tardy C, Bailly C, Colson P, Houssier C, Kumar A, Boykin DW, Wilson WD. Biopolymers 2002;63:281. [PubMed: 11877739]
9. Fox KR. Nucleic Acids Res 1992;20:6487. [PubMed: 1336176]
10. Hong M, Simpson L. Protist 2003;154:265. [PubMed: 13677453]
11. Liu B, Liu Y, Motyka SA, Agbo EE, Englund PT. Trends Parasitol 2005;21:363. [PubMed: 15967722]
12. Lukes J, Guilbride DL, Votypka J, Zikova A, Benne R, Englund PT. Eukaryot. Cell 2002;1:495. [PubMed: 12455998]
13. Shapiro TA, Englund PT. Annu. Rev. Microbiol 1995;49:117. [PubMed: 8561456]
14. Das BP, Boykin DW. J. Med. Chem 1977;20:531. [PubMed: 321783]
15. Barcelo F, Capo D, Portugal J. Nucleic Acids Res 2002;30:4567. [PubMed: 12384604]
16. Ren J, Jenkins TC, Chaires JB. Biochemistry 2000;39:8439. [PubMed: 10913249]
17. Nguyen B, Tanious FA, Wilson WD. Methods 2007;42:150. [PubMed: 17472897]
18. Davis TM, Wilson WD. Anal. Biochem 2000;284:348. [PubMed: 10964419]
19. Spartan '04. Irvine, CA: Wavefunction Inc.; 2004.
20. Munde M, Ismail MA, Arafa R, Peixoto P, Collar CJ, Liu Y, Hu L, David-Cordonnier MH, Lansiaux A, Bailly C, Boykin DW, Wilson WD. J Am Chem Soc 2007;129:13732. [PubMed: 17935330]
21. SYBYL Molecular Modeling Software. 7.3 ed.. St. Louis, MO: Tripos Inc.; 2006.
22. Liu Y, Kumar A, Boykin DW, Wilson WD. Biophys Chem 2007;131:1. [PubMed: 17889984]
23. Garbett NC, Ragazzon PA, Chaires JB. Nat Protoc 2007;2:3166. [PubMed: 18079716]
24. Malin, ABN.; Tomas, K. Circular Dichroism. Nina, B.; Koji, N.; Robert, W., editors. Wiley-VCH; 2000. p. 741
25. Mazur S, Tanious FA, Ding D, Kumar A, Boykin DW, Simpson IJ, Neidle S, Wilson WD. J. Mol. Biol 2000;300:321. [PubMed: 10873468]
26. Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE, Haas B, Bohme U, Hannick L, Aslett MA, Shallom J, Marcello L, Hou L, Wickstead B, Alsmark UC, Arrowsmith C, Atkin RJ, Barron AJ, Bringaud F, Brooks K, Carrington M, Cherevach I, Chillingworth TJ, Churcher C, Clark LN, Corton CH, Cronin A, Davies RM, Doggett

J, Djikeng A, Feldblyum T, Field MC, Fraser A, Goodhead I, Hance Z, Harper D, Harris BR, Hauser H, Hostetler J, Ivens A, Jagels K, Johnson D, Johnson J, Jones K, Kerhornou AX, Koo H, Larke N, Landfear S, Larkin C, Leech V, Line A, Lord A, Macleod A, Mooney PJ, Moule S, Martin DM, Morgan GW, Mungall K, Norbertczak H, Ormond D, Pai G, Peacock CS, Peterson J, Quail MA, Rabbinowitsch E, Rajandream MA, Reitter C, Salzberg SL, Sanders M, Schobel S, Sharp S, Simmonds M, Simpson AJ, Tallon L, Turner CM, Tait A, Tivey AR, Van Aken S, Walker D, Wanless D, Wang S, White B, White O, Whitehead S, Woodward J, Wortman J, Adams MD, Embley TM, Gull K, Ullu E, Barry JD, Fairlamb AH, Opperdoes F, Barrell BG, Donelson JE, Hall N, Fraser CM, Melville SE, El-Sayed NM. Science 2005;309:416. [PubMed: 16020726]

27. Levene SD, Wu HM, Crothers DM. Biochemistry 1986;25:3988. [PubMed: 3017412]
28. Marini JC, Levene SD, Crothers DM, Englund PT. Proc Natl Acad Sci U S A 1982;79:7664. [PubMed: 16593261]
29. Nunn CM, Neidle S. J. Med. Chem 1995;38:2317. [PubMed: 7608897]
30. Mallena S, Lee MP, Bailly C, Neidle S, Kumar A, Boykin DW, Wilson WD. J Am Chem Soc 2004;126:13659. [PubMed: 15493923]
31. Jin R, Breslauer KJ. Proc Natl Acad Sci U S A 1988;85:8939. [PubMed: 2461559]



poly(dA)·poly(dT): $dA_n \cdot dT_n$

poly(dA-dT)·poly(dA-dT): $d(AT)_n \cdot d(AT)_n$

A_{15} Hairpin: 5'– $CCA_{15}GC$ / $GGT_{15}CG$

$(AT)_7$ Hairpin: 5'– $CC(AT)_7AGC$ / $GG(TA)_7TCG$

A_8 Hairpin: 5'– CAAAAAAAA / GTTTTTTTT

$(AT)_4$ Hairpin: 5'– CATATATAT / GTATATATA

A_4 Hairpin: 5'– CCAAAAGC / GGTTTTCG

ATAT Hairpin: 5'– CCATATGC / GGTATACG

Fig. 1.

Structure of the compounds and the DNA sequences used in this study. The DNA sequences were also obtained with 5' biotin for SPR experiments. Additional GC base pairs were added to the ends of the short and long DNA sequences so that they could be studied at high temperature.

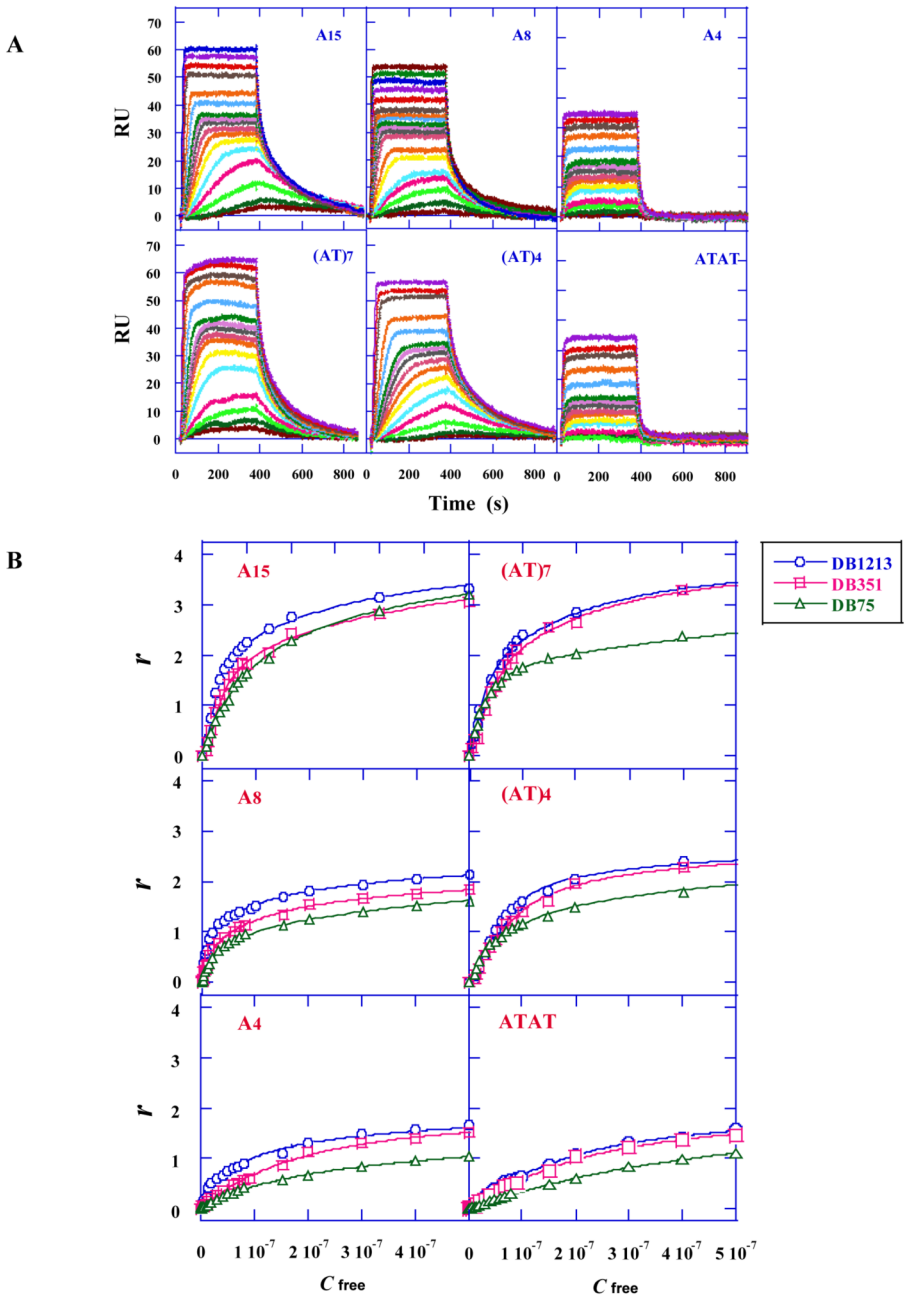


Fig. 2.

SPR binding affinity: (A) Representative SPR sensorgrams for the interaction of DB1213 with alternating and nonalternating DNA sequences at different lengths. The compound concentrations for DB1213 from bottom to top are 0 to 0.5 μM. (B) Comparison of the SPR binding affinity for nonalternating and alternating DNA sequences with different compounds. *RU* values from the steady–state region of SPR sensorgrams were converted to *r* (*r* = *RU*/RU_{max}) and are plotted against the unbound compound concentration (flow solution) for DB1213 (circles), DB351 (squares) and DB75 (triangle) binding with alternating and nonalternating DNA sequences at different lengths. The lines are the best fit values using appropriate binding models as described in the text.

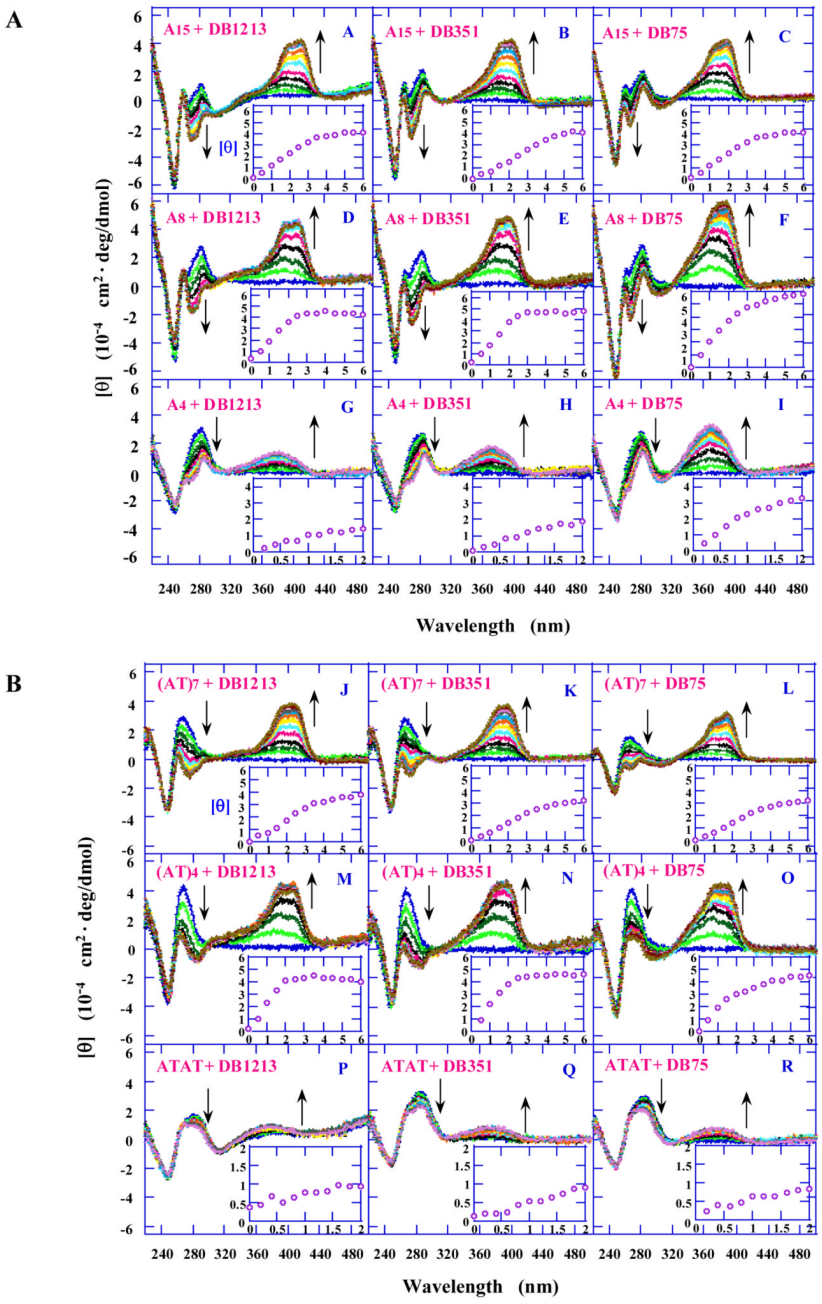


Fig. 3.

Comparison of CD spectra of nonalternating (A-I) and alternating (J-R) DNA sequences of different lengths in complex with different compounds at various mixing ratios. Insert: increase in induced CD magnitude plotted versus the binding ratio of compound per oligomer.

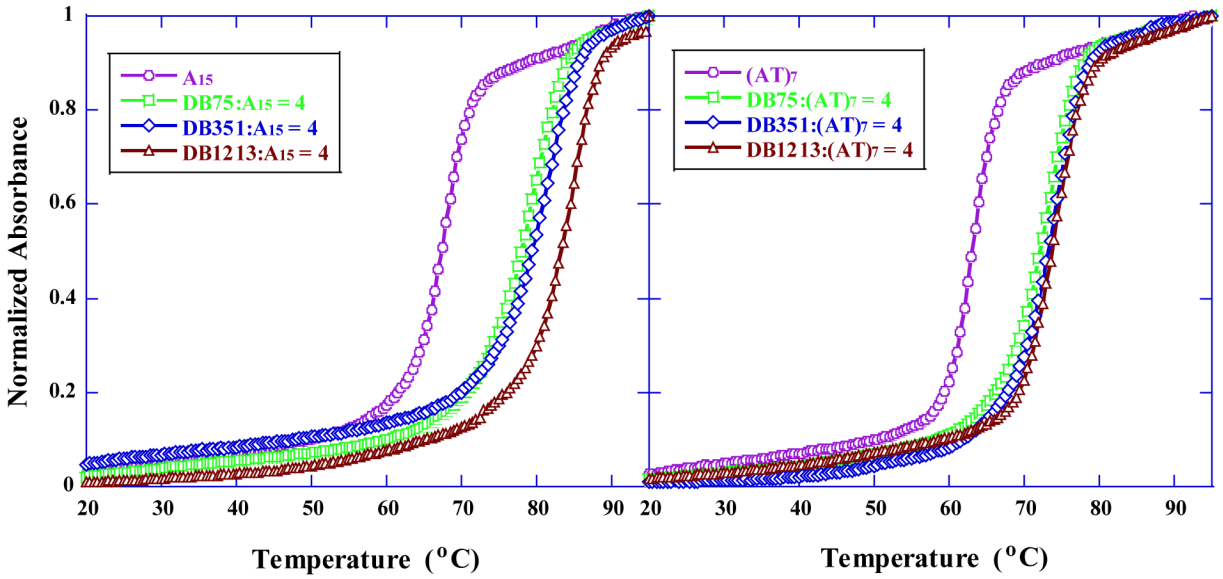


Fig. 4.

UV melting profiles at 260nm of the DNA hairpins in the absence and presence of compounds. (A) Thermal melting curves of A_{15} hairpin at the indicated ratio of compound per oligomer. (B) Thermal melting curves of $(AT)_7$ hairpin at the indicated ratio of compound per oligomer.

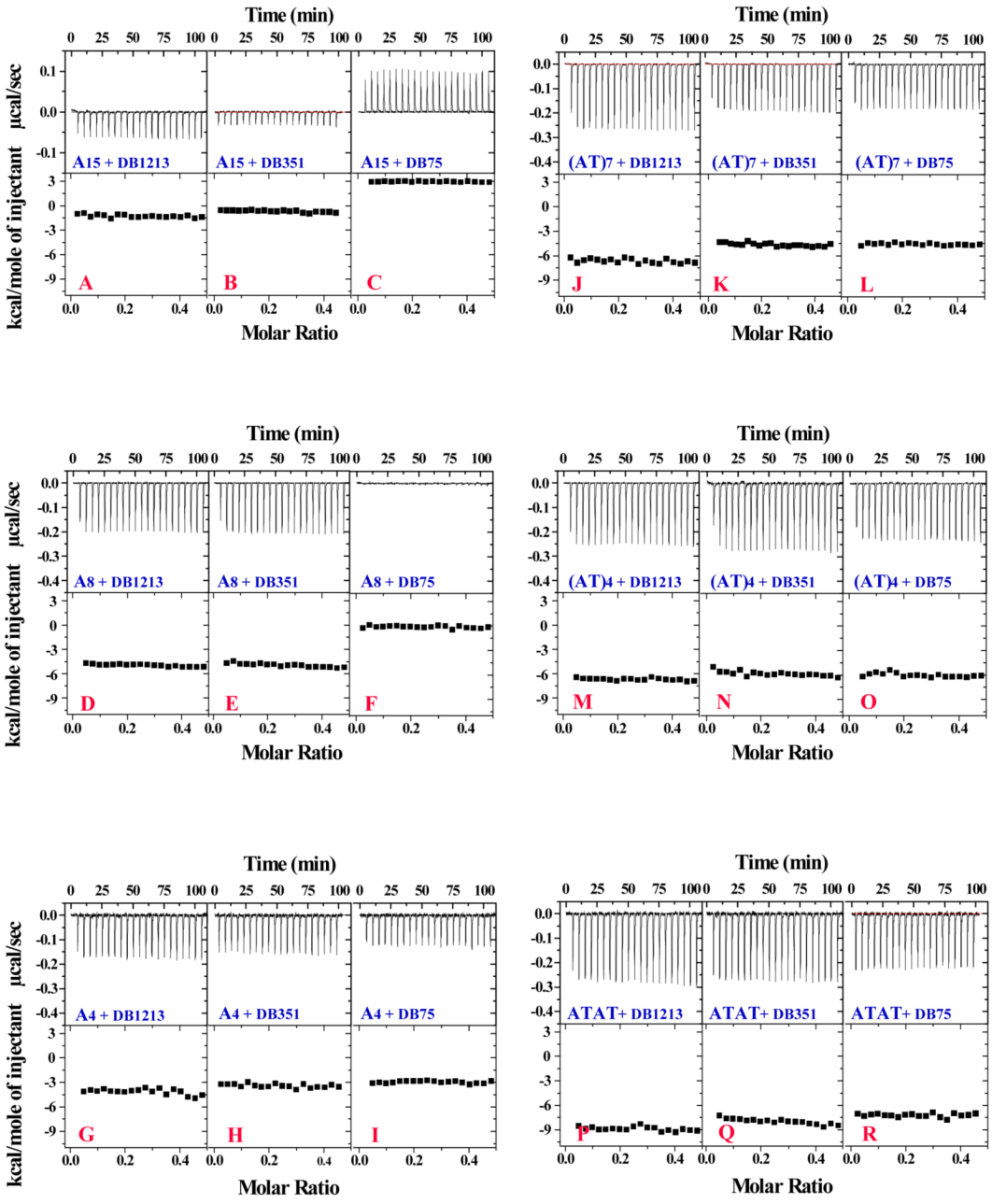


Fig. 5.

Comparison of the ITC results for nonalternating (A-I) and alternating (J-R) DNA sequences binding with different compounds in MES10 at 25 °C. Every peak represents the instrument response for injection of compound into DNA (top). The area of each peak is determined and corrected by subtracting the heat of dilution (bottom).

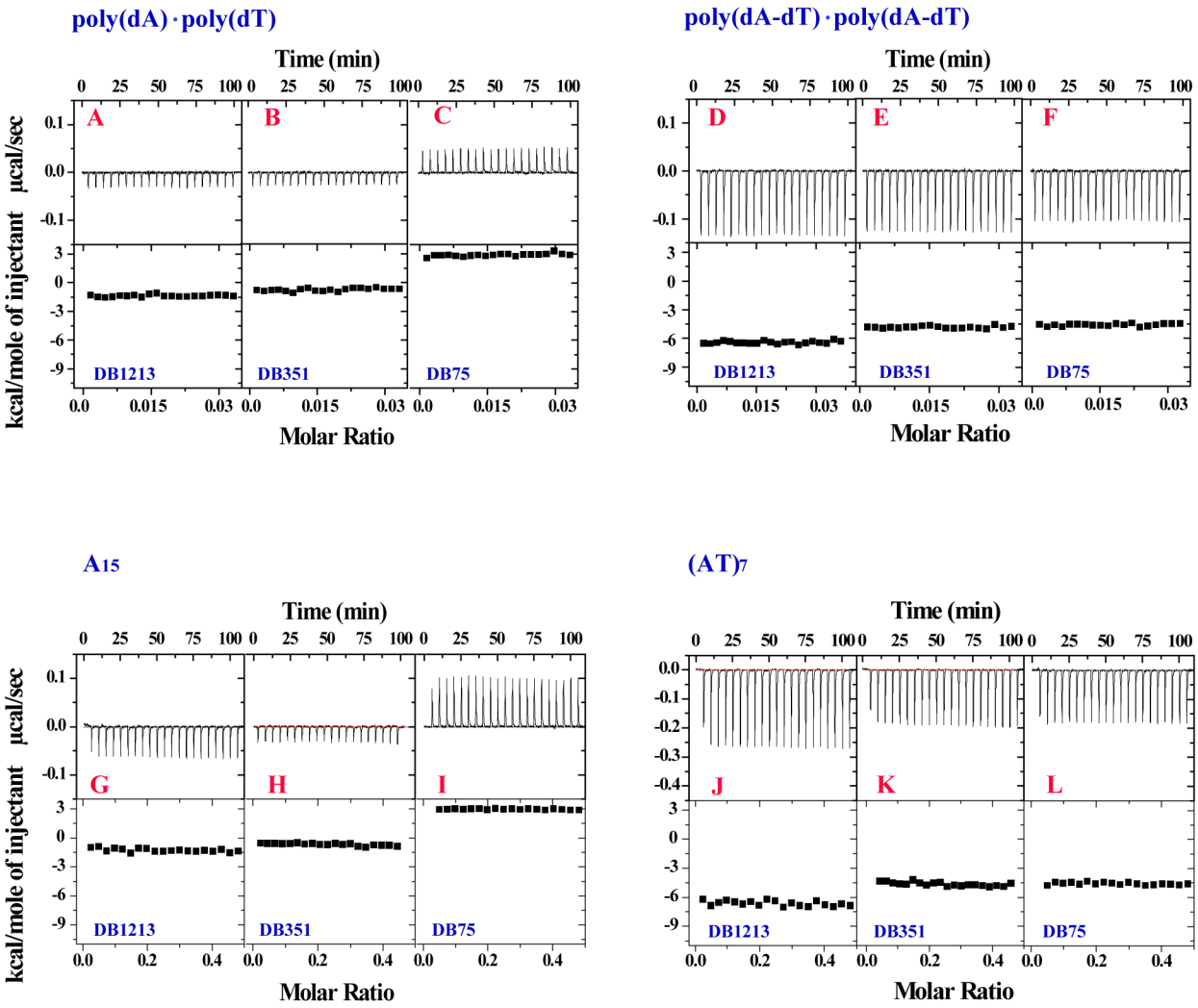


Fig. 6.

Comparison of the ITC results for nonalternating and alternating polymers [poly(dA)·poly(dT) and poly(dA-dT)·poly(dA-dT)] and oligmers [A_{15} and $(AT)_7$] binding with different compounds in MES10 at 25 °C.

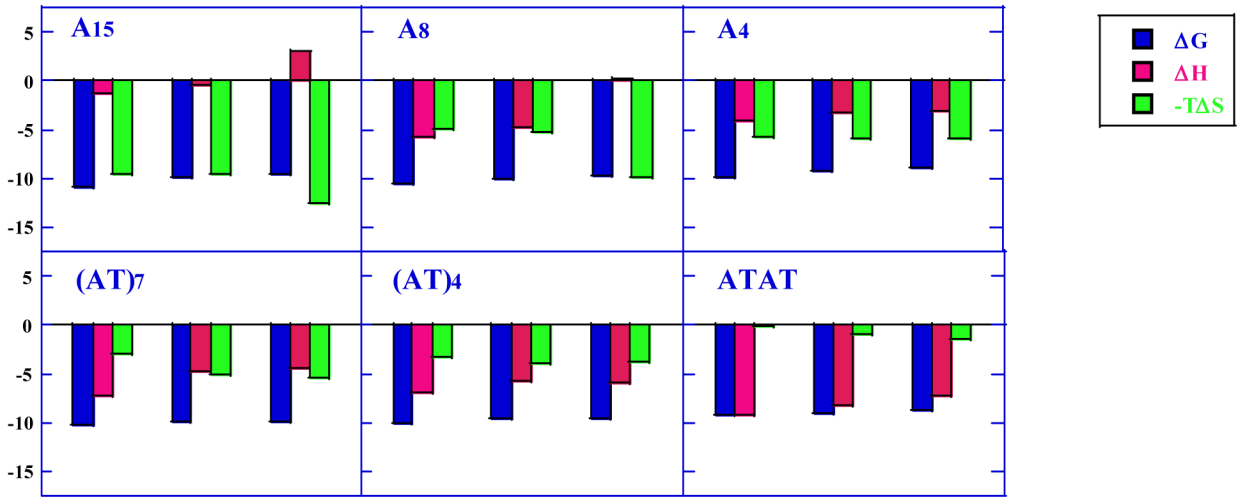


Fig. 7.
Comparison of thermodynamic parameters for nonalternating and alternating DNA sequences binding with different compounds at 25 °C.

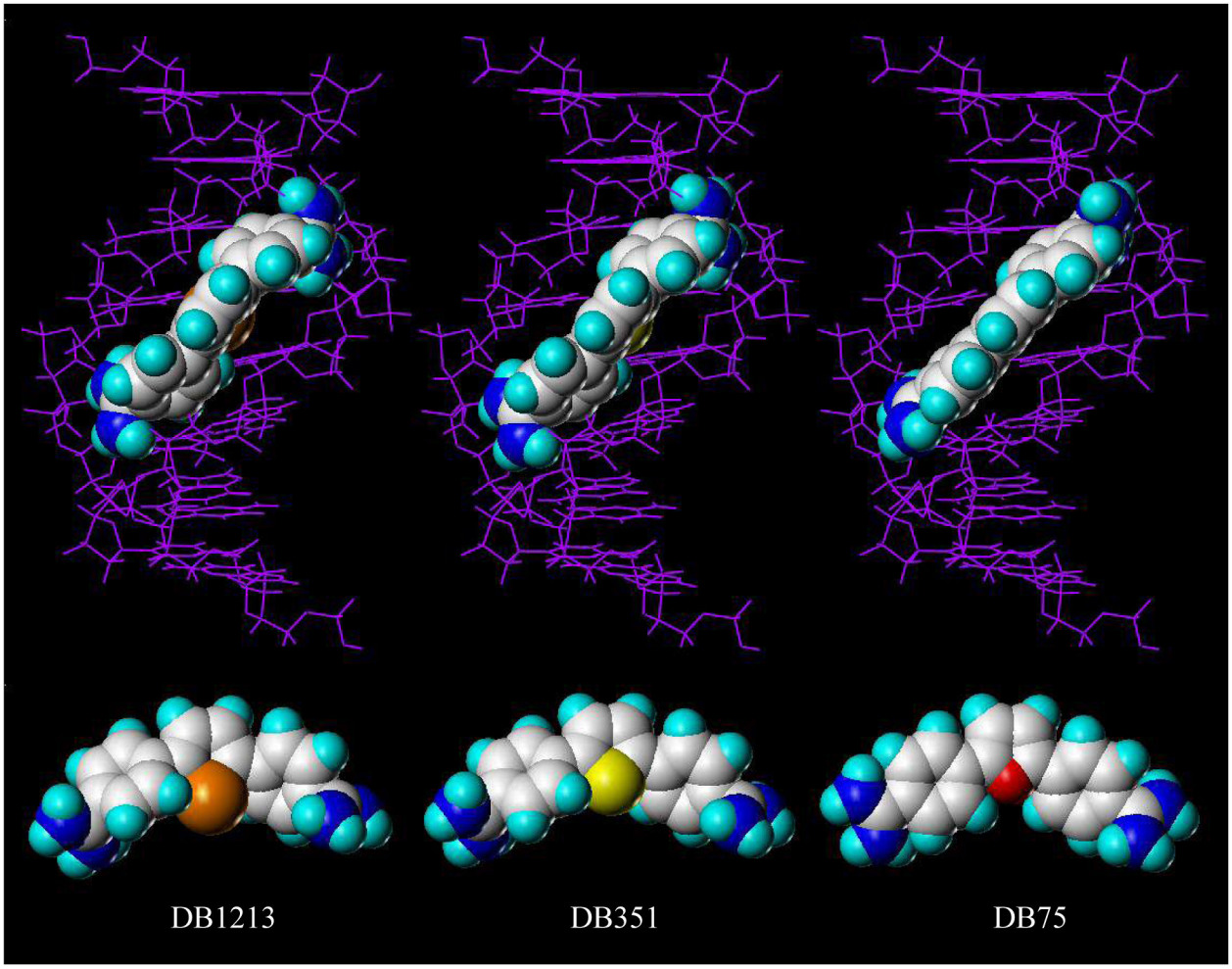


Fig. 8.

From left to right are docked compounds DB1213, DB351 and DB75 in complex with d(CCAAAAGC)·d(GCTTTTCG). Note that DB1213 displays more pronounced rotation of its phenyl groups away from the selenium (orange), DB351 displays a distinct yet smaller rotation of the phenyl groups away from the sulfur (yellow), and DB75 is fairly planar (red).

Table 1

Thermodynamic parameters for the interactions of the compounds in Figure 1 binding to nonalternating and alternating sequences at different lengths.

DNA sequence	Compound	K_{mean} [a] (M^{-1})	ΔG (kcal/mole)	ΔH [b] (kcal/mole)	$-T\Delta S$ (kcal/mole)
A_{15}	DB1213	1.01×10^8	−10.91	−1.34	−9.57
	DB351	1.90×10^7	−9.92	−0.37	−9.55
	DB75	1.02×10^7	−9.56	3.02	−12.58
A_8	DB1213	5.57×10^7	−10.56	−5.66	−4.90
	DB351	2.03×10^7	−9.96	−4.75	−5.21
	DB75	1.37×10^7	−9.73	0.22	−9.95
A_4	DB1213	1.71×10^7	−9.86	−4.11	−5.75
	DB351	6.06×10^6	−9.25	−3.27	−5.98
	DB75	3.59×10^6	−8.94	−3.04	−5.9
$(AT)_7$	DB1213	2.72×10^7	−10.14	−7.17	−2.97
	DB351	1.98×10^7	−9.95	−4.80	−5.15
	DB75	1.82×10^7	−9.90	−4.42	−5.48
$(AT)_4$	DB1213	2.29×10^7	−10.03	−6.87	−3.26
	DB351	1.11×10^7	−9.61	−5.74	−3.87
	DB75	1.13×10^7	−9.62	−5.92	−3.7
ATAT	DB1213	5.62×10^6	−9.22	−9.15	−0.07
	DB351	4.80×10^6	−9.11	−8.25	−0.86
	DB75	2.52×10^9	−8.73	−7.31	−1.42

[a] K_{mean} values are from SPR experiments as described in Materials and methods and ΔG is obtained from the equation: $\Delta G = -RT\ln K_{mean}$.

[b] ΔH binding values are from ITC as described in Materials and methods Experimental error limits for SPR K values and ITC ΔH values are both approximately ±10%.