

CHEMISTRY

A European Journal

A Journal of



Accepted Article

Title: Small Size Sequence-sensitive Compounds for Specific Recognition of G·C Base Pair in DNA Minor Groove.

Authors: Abdelbasset Ahmed Farahat, Pu Guo, Hadir Shoeib, Ananya Paul, David Boykin, and W.David Wilson

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *Chem. Eur. J.* 10.1002/chem.201904396

Link to VoR: <http://dx.doi.org/10.1002/chem.201904396>

Supported by
ACES

WILEY-VCH

Small Size Sequence-sensitive Compounds for Specific Recognition of G·C Base Pair in DNA Minor Groove.

Abdelbasset A. Farahat^{1,2,*}, Pu Guo¹, Hadir Shoeib¹, Ananya Paul¹, David W. Boykin¹, W. David Wilson^{1*}

1. Department of Chemistry and Center for Diagnostics and Therapeutics Georgia State University, 50 Decatur St SE, Atlanta, GA 30303, USA.

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

* Corresponding authors emails. A. A. Farahat (afarahat@gsu.edu), W.D. Wilson (wdw@gsu.edu)

Abstract.

A novel series of small size diamidines with thiophene and modified *N*-alkylbenzimidazole σ -hole module represent specific binding to single G·C base pair (bp) DNA sequence. The variation of *N*-alkyl or aromatic rings were sensitive to microstructure of DNA minor groove. Thirteen new compounds were synthesized to test their binding affinity and selectivity. The dicyanobenzimidazoles needed to synthesize the target diamidines were made via condensation/cyclization reactions of different aldehydes with different 3-amino-4-(alkyl- or phenyl-amino) benzonitriles. The final diamidines were synthesized using lithium bis-trimethylsilylamide (LiN[Si(CH₃)₃]₂) or Pinner methods. The newly synthesized compounds showed strong binding and selectivity to AAAGTTT compared to similar sequences AAATTT and AAAGCTTT investigated by several biophysical methods including biosensor-SPR, fluorescence spectroscopy, DNA thermal melting, ESI-MS spectrometry, Circular dichroism, and Molecular Dynamics. The binding affinity results determined by fluorescence spectroscopy are in accordance

with those obtained by biosensor-SPR. These small size single G·C bp highly specific binders extend the compound database for future biological applications.

Keywords.

Biosensor-SPR, Diamidines, Minor Groove Binders, Molecular Dynamic Simulation, *N*-alkylbenzimidazole G·C Base Pair.

1. Introduction.

Minor groove binders that can recognize mixed base pairs (bps) DNA with sequence selectivity are of interest for a variety of research areas ^[1]. The finding that minor groove binding compounds can be very effective inhibitors of major groove binding proteins is particularly important for the use of these compounds as therapeutic agents as well as biological probes ^[2-5]. The heterocyclic diamidine minor groove binders that have been designed and prepared in our group have been demonstrated to get into a variety of cell types and to have clinically useful biological activity with low toxicity ^[6-8]. In addition, diamidines have been shown to be excellent inhibitors of transcription factors (TFs) both in vitro ^[9,10] and in vivo ^[11,12]. The diamidine minor groove binders synthesized previously have been essentially pure AT interacting compounds. To recognize and inhibit a broad range of TFs, however, it is essential to develop compounds that bind to the DNA sequence that contain G·C as well as A·T bps. To do this, we have initiated a project to incorporate to G·C bp recognizing modules into heterocyclic diamidines.

Ligands that can recognize mixed base pair sequences of DNA containing A·T and G·C bps are an important development area for minor groove binders. At first, the development took a major step forward with modification of the minor groove binders, netropsin (Nt.) and distamycin (Dst.) ^[13-15]

The proposal was to replace one or more pyrrole groups in Nt. and Dst. with *N*-methylimidazole groups^[13-15]. The extra N in *N*-methylimidazole relative to pyrrole should allow that group to H-bond with the G-NH₂ that protrudes into the minor groove. Although several groups have worked with synthetic polyamides, solution difficulties have limited their applications^[16-18].

Our goal was to prepare entirely new types of compounds without the amide group that could bind to mixed A·T and G·C bps containing sequences but would also have varied chemical and structural properties that would allow them to specifically bind to target sequences and effectively target a variety of different cell types. The initial research in the area has produced quite different compounds that are undergoing additional development to increase affinity and selectivity for the target sequence type with a single G·C flanked by A·T bps^[19-21]. Among them, the *N*-MeBI-Thiophene containing compounds were found to be effectively modified to improve their binding affinity and specificity^[22].

Incorporating the *N*-MeBI thiophene unit into a heterocyclic cation base structure, as in **Figure 1**, resulted in DB2429 which can recognize a target G·C bp in an AT sequence^[21]. Exploration of the importance of the *N*-MeBI-thiophene module to binding and specificity shows that both the *N*-MeBI and thiophene are essential for the strong, specific interactions with the target sequence. Replacing the thiophene or *N*-MeBI with other similar structures decreased the binding affinity and specificity. Compounds with the *N*-MeBI replaced by benzimidazole (BI), for example, bind much more strongly to pure AT DNA sequences than to any sequence with a G·C bp^[21, 22]. For the design of strong binders in DNA minor groove, compounds should have, for example, the appropriate curvature to fit the convex shape of the minor groove.^[23-27]

One or two cationic groups are necessary for new compounds to aid solubility and increase the electrostatic interaction with the negative DNA backbone [23, 28, 29]. For example, amidine cationic groups, which have been quite successful for targeting DNA, provide a hydrogen-bond-donating planar surface for optimizing interactions with A·T bps at flanking sites of G·C bps. Also, the twist dihedral between amidine and phenyl ring would help the compound structure to fit the helical curvature of the minor groove [30,31]. Moreover, the groove binding ligand should include a H-bond acceptor (*N*-MeBI) to recognize G-NH₂. In order to improve binding specificity, it is important to attempt to take advantage differences in groove-width of targeted sequences. Based on Rohs [25] and our work [22], the wider DNA minor groove at GC versus AT sequences can be targeted. So, strategic structural modification of the groove-binders could lead to optimization of binding specificity. The significant improvement of binding specificity with such small changes in the overall compound structure [21] encouraged us to design new agents to specifically recognize mixed sequence DNA.

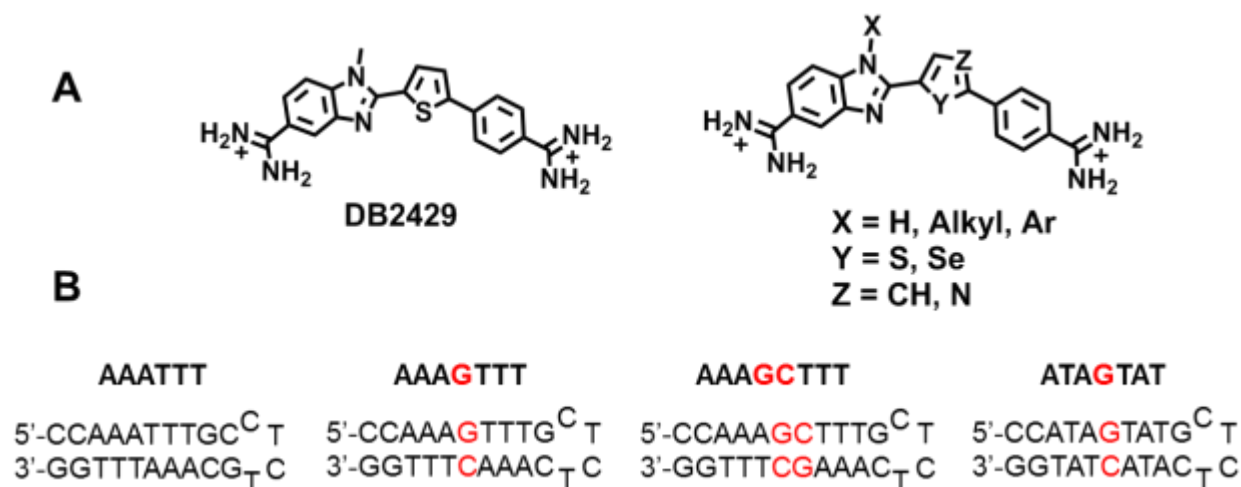


Figure 1. A) Chemical structure of a reported single G·C bp binder, DB2429, and the modified compounds from DB2429 used in this study. B) The DNA sequences used in this study; DNA sequences used for SPR studies were labeled with 5'-biotin.

We have recently shown that larger compounds with the *N*-MeBI-thiophene module can be modified to have significant improvements in binding specificity for a target sequence with a single G·C bp^[22]. This is a very significant step forward in preparation of biologically useful compounds. For drug design, however, there is a significant improvement in synthetic ease and cost by using smaller compounds such as DB2429 and its analogs. In making these small molecules we used readily available starting materials, avoided using the Palladium catalyzed expensive organometallic reactions and the final products were obtained after few synthetic steps. There is also a general improvement in cell uptake with smaller compounds^[30, 16-18]. A problem with DB2429, however, is that its ratio of binding to a target single G·C bp sequence over an equivalent pure AT sequence is only about a factor of 10. A concept for improving this ratio is that by designing DB2429 analogs with a variety of substituents, local variations in DNA microstructure could be recognized better with higher binding selectivity^[32].

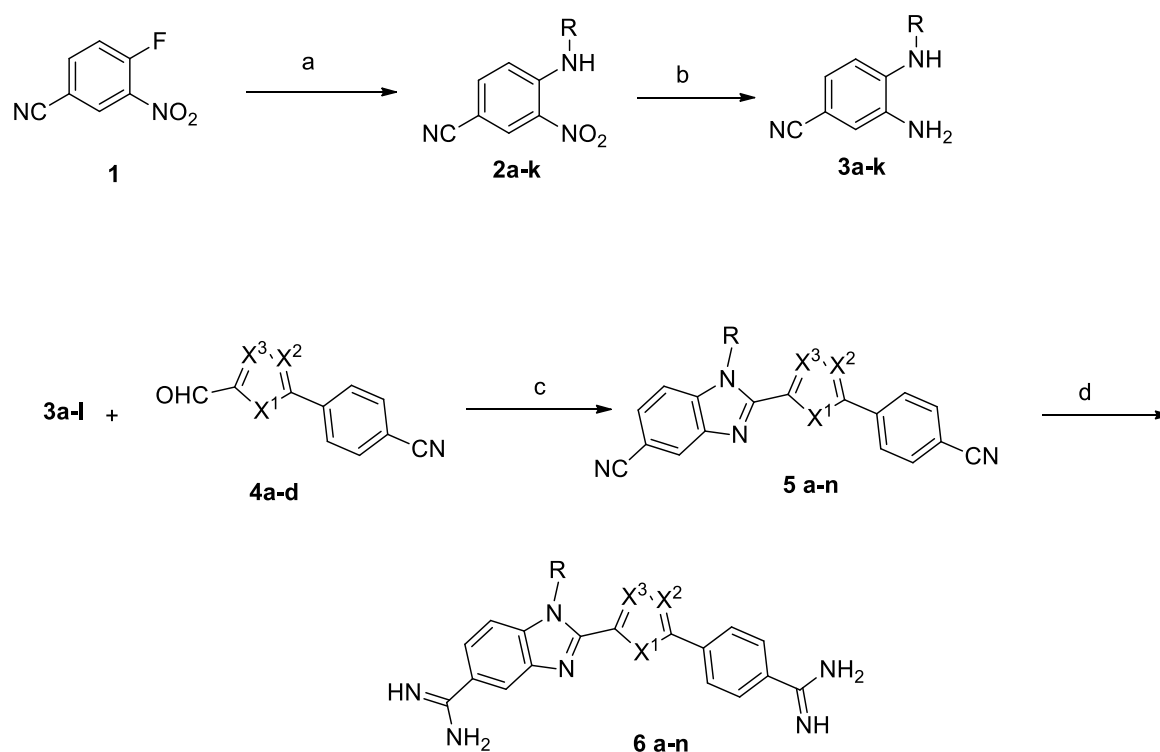
The first set of compounds that were prepared had modifications at the *N*-BI position of DB2429 (Scheme 1, Table 1). We wanted to increase the size of the alkyl group to reduce binding to narrow DNA sequences like AAATTT, and bigger size substituents may fit the wider DNA minor groove such as AAAGTTT, for example, methyl < ethyl < isopropyl < isobutyl < neopentyl substituents. We also employed the *N*-trifluoroethyl DB2716 to test the different influence between the trifluoroethyl and ethyl substitutions. The linear group with a more polar substituent as in the *N*-propylmethoxy DB2722 was prepared by the same methods (Scheme 1). To dramatically change the shape of the alkyl group, compounds with *N*-cyclobutyl (DB2720), *N*-cyclopentyl (DB2713) and *N*-cyclohexyl (DB2721) substituents were prepared. Finally, the substituent was significantly changed with a phenyl (Ph) substituent on the BI group, *N*-Ph (DB2739). Additional compounds were prepared with substitutions designed to test different σ -hole systems. The σ -hole system

preorganizes the thiophene-*N*-XBI module for G·C bp interaction in the minor groove. The C–S single bonds of thiophene ring presents a relatively positive electrostatic potential that can form an interaction with electron-donating atoms such as the unsubstituted N of *N*-XBI. The interaction is based on the presence of low lying thiophene C–S σ^* orbitals on S that give rise to the positive electrostatic potential or a σ -hole^[33]. One compound, with an *N*-MeBI was prepared with the thiophene S replaced by Se to evaluate the effect of size and changes in polarizability at that position,^[34, 35] as well as, DB2591 with a thiazole structure replacing the original thiophene ring. A control compound for the *N*-Me DB2591, the *NH* DB2652, without an *N*-BI substituent, was also prepared. Brief synthetic details are given here for the synthesis of the new compounds along with Scheme 1.

2. Results and Discussion

2. 1. Chemistry

Scheme 1 outlines the synthesis of the final thiophene-benzimidazole diamidines **6a-n**. Different alkyl and arylamines underwent nucleophilic substitution reaction with 4-fluoro-3-nitrobenzonitrile **1** to afford the alkyl and aryl substituted nitro benzonitrile **2a-k**. The previous nitro compound was reduced using stannous chloride in ethanol to produce the diamine **3a-k**. The diamine **3a-k** underwent oxidative condensation and cyclization with the thiophene, selenophene or thiazole aldehydes **4a-d** using sodium metabisulphite in DMSO to afford the prefinal bisnitriles **5a-n**. The final diamidines **6a-n** were produced by allowing the bisnitriles to react with either lithium bistrimethylsilylamide or ethanolic/HCl (Pinner methodology).



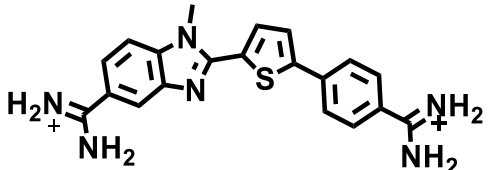
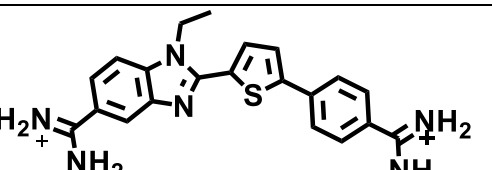
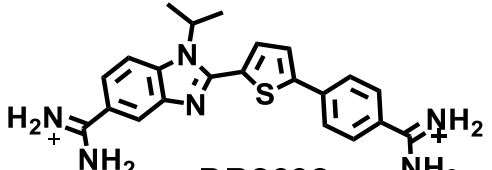
Scheme 1. Reagents and conditions: a) RNH_2/EtOH , rt; b) $\text{SnCl}_2/\text{EtOH}$, reflux; c) $\text{Na}_2\text{S}_2\text{O}_3/\text{DMSO}$, 130 °C; d) $i\text{-LiN}(\text{TMS})_2/\text{THF}$, ii- HCl/EtOH or $i\text{-HCl}/\text{EtOH}$, ii- NH_3 gas.

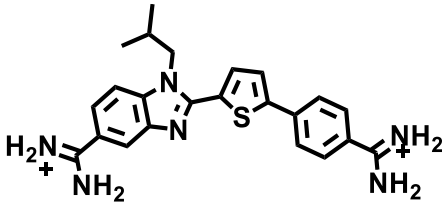
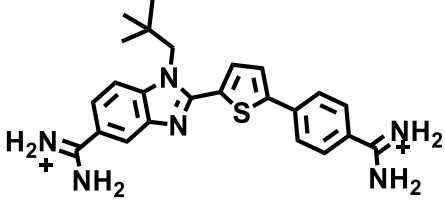
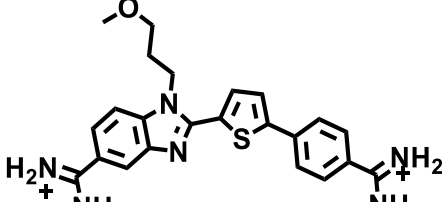
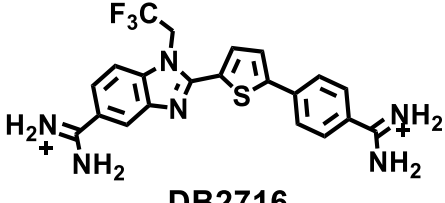
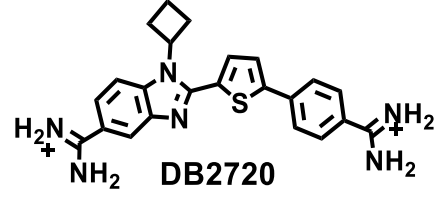
2.2. DNA Binding.

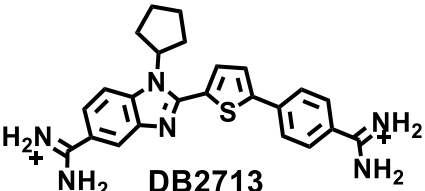
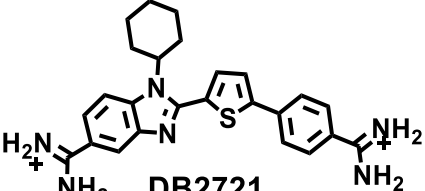
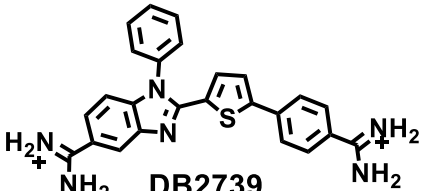
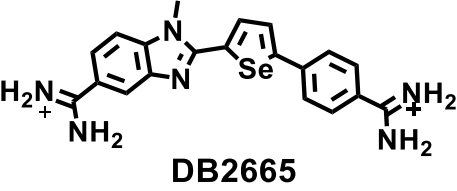
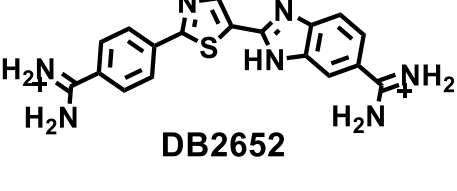
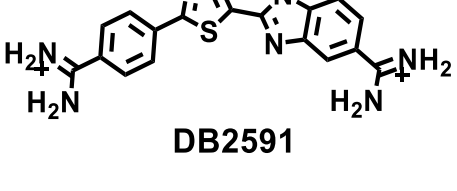
2.2.1. DNA Thermal Melting: Screening for mixed sequence binding.

We have used changes in DNA thermal melting temperature (T_m) to provide an effective way for initial testing of compounds for binding affinity with different DNA sequences. We have used the three sequences shown in Figure 1 for comparative experiments. Most of the dicationic amidines reported so far have shown pure AT sequence binding activity and a DNA with the binding site -AAATTT- was used to test for AT binding affinity. Our primary interest was studying DNA that has a single G·C bp with the binding sequence -AAAGTTT-. The third sequence with two G·C bps was also used to test for binding affinity and selectivity versus AAAGTTT.

Table 1. Thermal Melting (ΔT_m , °C)^a, Fluorescence Titration (K_D , nM)^b, and SPR (K_D , nM)^c of Studied Compounds with Pure AT and Mixed DNA Sequences.

Compounds	Experiments Sequences	ΔT_m (°C) ^a			FL (K_D , nM) ^b	SPR AAAGTTT (K_D , nM) / AAATTT (K_D , nM) / Selectivity Ratio vs AAATTT
		AAA TTT 64.5 (°C)	AAA G TTT 66.5 (°C)	AAA GC TTT 70.5 (°C)		
 DB2429 6a		5	11	5	45 ± 2	51 / 522 / 10
 DB2689 6b		3	10	2	44 ± 2	49 / NB / NB
 DB2692 6c		2	10	2	21 ± 2	20 / NB / NB

 DB2710 6d	2	9	2	33 ± 3	36 / NB / NB
 DB2717 6e	2	9	2	30 ± 2	--
 DB2722 6f	2	10	2	19 ± 2	--
 DB2716 6g	3	10	4	29 ± 2	--
 DB2720 6h	2	10	2	21 ± 1	--

 DB2713 6i	2	9	2	15 ± 2	21 / NB / NB
 DB2721 6j	2	10	2	17 ± 1	--
 DB2739 6k	3	11	3	10 ± 4	8 / 1330 / 166
 DB2665 6l	5	8	3	59 ± 2	--
 DB2652 6m	13	7	4	82 ± 4	--
 DB2591 6n	3	9	3	83 ± 2	--

a. $\Delta T_m = T_m$ (the complex) - T_m (the free DNA). 3 μM DNA sequences were tested in TNE100 buffer with the ratio of 2:1 [ligand] / [DNA]. An average of two independent experiments with a reproducibility of 0.5 $^{\circ}\text{C}$.

b. Binding affinity (K_D , nM) calculated from fluorescence experiments for 20 nM compound titrated with sequence AAAGTTT in TNE100 buffer at 25 $^{\circ}\text{C}$. The excitation wavelength is 337 nm. Two independent experiments were studied. The errors represent the standard errors of the mean.

c. SPR binding affinity (K_D , nM) were investigated in Tris-HCl buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 0.05% P20, pH 7.4) at a 100 $\mu\text{L}\cdot\text{min}^{-1}$ flow rate. “--” means not done, “NB” means binding between ligands and sequence AAATTT is no detectible binding.

The lead compound, DB2429, with an *N*-Me substituent was previously reported ^[21] and has strong binding to the single G·C bp sequence, the ΔT_m is 11 $^{\circ}\text{C}$, with lower ΔT_m and weaker binding to the AT and two G·C bps sequences, 5 and 5 $^{\circ}\text{C}$ respectively. Except for the *N*-MeBi-selenophene DB2665 and the NH DB2652, the rest of the tested compounds with different bulky alkyl and aromatic groups showed similar binding to DB2429 with AAAGTTT and much weaker binding to both the AAATTT and AAAGCTTT sequences. Most of the *N*-substitutions modified compounds showed improved selectivity to AAAGTTT as planned (Table 1). DB2665 with selenophene as a spacer will be helpful to future cell uptake studies, however, showed less selectivity towards AAAGTTT. Since the bigger size of selenium over sulfur, the entire structure of DB2665 is more linear than DB2429 and does not fit the minor groove curvature well. DB2652 with thiazole as a spacer and unsubstituted NH of benzimidazole showed strong binding to AAATTT sequence than other tested sequences and this agrees with other previously tested NH-benzimidazole compounds prepared in our lab ^[36]. DB2591 with thiazole as a spacer and *N*-Me substituted benzimidazole still shows strong binding and

selectivity towards AAAGTTT sequence compared to other tested sequences and this indicates that changing the spacer from thiophene to thiazole preserves but weakens the target DNA binding activity and selectivity. We conclude that changing the alkyl substituents on the benzimidazole moiety has no large effect on the AAAGTTT binding activity but has a great effect on increasing the selectivity towards that sequence. The presence of these alkyl substituents is crucial for binding and selectivity towards AAAGTTT and its absence leads to compounds with strong binding to pure AT sequences.

To test the sequence sensitivity of the compounds, the wider flanking sequence with single G·C bp ATAGTAT was used in DNA melting experiments (Table S4). Comparing to ATAGTAT and AAAGCTTT, the sequence AAAGTTT is still the strongest binding target for this type of compound. Since the G·C bp is present in the ATAGTAT sequence, the binding affinity is stronger than two G·C bps sequence. Depending on the different microstructure on flanking part between AAAGTTT and ATAGTAT, the bigger group modification of *N*-MeBI only improves the G-binding part but not the AT-flanking part.

2.2.2. Fluorescence Emission Spectroscopy: Binding Affinity

The aim of binding assays is to measure quantitatively the interactions between two molecules, such as a small molecule binding with DNA. If a change in the fluorescence intensity accompanies the binding of two species, this can be used to quantitatively monitor the binding interaction and determine the stoichiometry of binding using equilibrium titration methods. Such methods have been used extensively to examine numerous drug–DNA interactions in solution including dicationic amidines–DNA interactions. Generally, if the molecule changes its fluorescence, then this species is titrated with the non-fluorescent molecule (e.g.,

DNA) and the change in fluorescence intensity is monitored as a function of total DNA concentration [12,30].

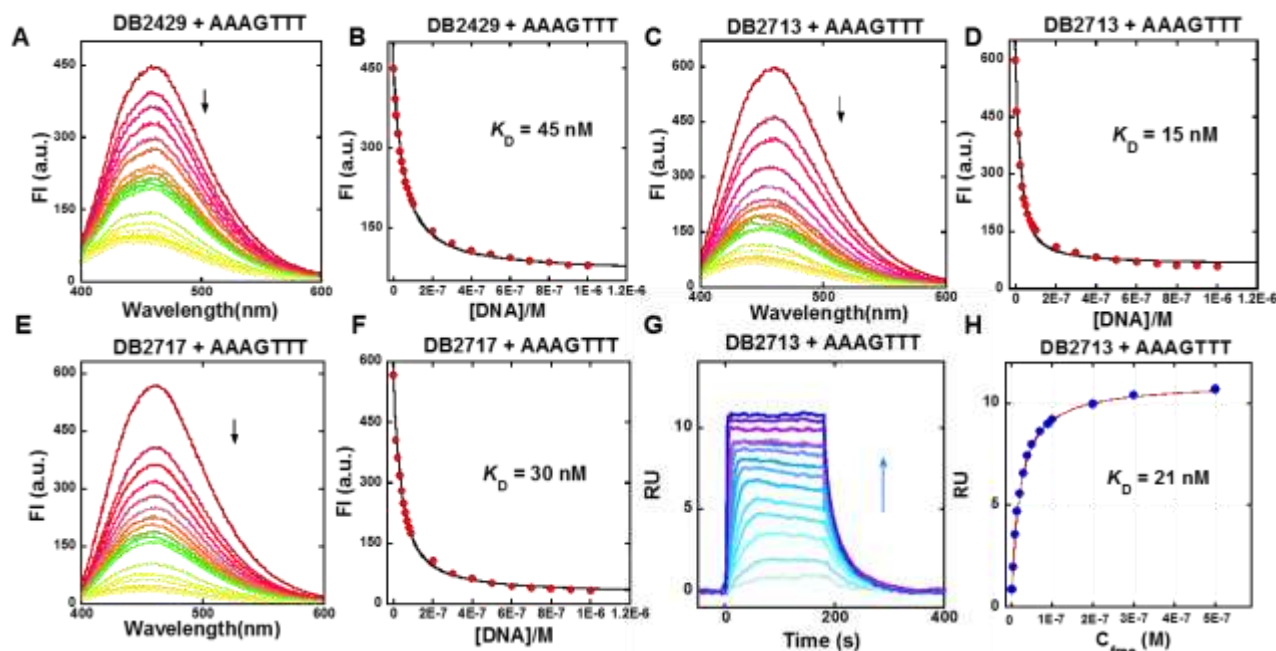


Figure 2. A, C, E) Fluorescence Emission spectra for 20 nM DB2429, DB2713 and DB2717 titrated with sequence AAAGTTT in TNE100 buffer at 25 °C. The excitation wavelength is 337 nm. The slit width is [10, 5 nm]; B, D, F) Fluorescence binding curve between 20 nM DB2429, DB2713 and DB2717 and sequence AAAGTTT in TNE100 buffer to determine equilibrium constant; G) Representative SPR sensorgrams for DB2713 in the presence of AAAGTTT sequence, concentrations of DB2713 from bottom to top are 2-500 nM.; H) SPR steady-state binding plots for DB2713 with AAAGTTT sequence. The data are fitted to a steady-state binding function using a 1:1 model to determine equilibrium binding constants.

The fluorescence intensity of our tested compounds decreased on titration with increasing concentrations of the sequence AAAGTTT. The binding affinities (K_D values) were determined by fluorescence titration experiments are shown in Table 1. Generally, the *N*-alkyl-substituted benzimidazole diamidines bind strongly to the DNA sequence AAAGTTT as it appears from the K_D . The *N*-phenyl substituted benzimidazole diamidine DB2739 is the strongest binder to DNA with K_D of 10 nM. Compounds with alkyl substitutions on the benzimidazole bind strongly to DNA such as the *N*-isopropyl DB2692, *N*-isobutyl DB2710, and *N*-

methyl with selenophene DB2665 except for the methyl and ethyl substituted DB2429, DB2689 that showed slightly weaker binding. In the meantime, compounds with cycloaliphatic substitutions such as *N*-cyclopentyl (DB2713) and *N*-cyclohexyl (DB2721) showed very good binding. To test the effect of increasing the length and bulkiness of the *N*-alkyl substituent, we prepared the *N*-isobutyl (DB2710), *N*-trifluoroethyl (DB2716), *N*-*tert*-butyl (DB2717), and *N*-propylmethoxy (DB2722) derivatives, these new derivatives still show high DNA affinity as seen from the K_D (19-35 nM). For selected new compounds, the K_D obtained from the SPR binding affinity studies compared closely with the values obtained from the fluorescence titrations. Compounds containing either thiophene or selenophene as spacer binds strongly to DNA while DB2652 and DB2591 that contains thiazole as a spacer has weaker binding to DNA (K_D = 82 nM and 141 nM respectively).

We further tested the fluorescence activity of six selected compounds against two other DNA sequences namely, ATAGTAT and AAAGCTTT (Table S1). The tested compounds showed moderate binding with the sequence ATAGTAT, compound *N*-propylmethoxy DB2722 is the strongest binder with K_D of 58 nM while the same compound showed K_D of 19 nM on binding to AAAGTTT, this demonstrates that DB2722 is more specific towards the sequence AAAGTTT than ATAGTAT. After studying the K_D values of the tested compounds against ATAGTAT, we realized that in general these six compounds are more specific binders with AAAGTTT than ATAGTAT. These six tested compounds showed no binding to AAAGCTTT. Although sequence AAAGTTT and ATAGTAT look similar, there is a large difference (almost 2 Å) of minor groove width between AAAGTTT and ATAGTAT according to high-throughput DNA shape prediction^[32]. Our *N*-substituents compounds are so sensitive to the variations of the

microstructure of DNA that they bind to AAAGTTT and ATAGTAT with different binding affinities.

2.2.3. Biosensor-SPR: Methods for Quantitative Binding

Biosensor-SPR methods provide an excellent way to quantitatively evaluate the interaction of small organic molecules with immobilized biomolecules^[38]. SPR provides sensitive, real-time progress of the binding reaction as well as the equilibrium binding affinity, kinetics, and stoichiometry of complex formation^[23, 38]. Based on the DNA thermal melting and fluorescence spectroscopy results, the interactions of selected compounds, DB2429, DB2689, DB2692, DB2710, and DB2713, with AAAGTTT were evaluated by SPR (Table S2). In accordance with the fluorescence spectroscopy experiments, SPR results indicate the similar binding affinity to sequence AAAGTTT (Table 1, Table S2). Representative SPR sensorgrams for DB2713 with *N*-cyclopentyl is shown in Figure 2G. Moreover, the data are fitted to a steady-state binding function using a 1:1 model to achieve the binding affinity ($K_D = 19$ nM) with the sequence AAAGTTT. Consistent with the fluorescence quenching experiment of sequence AAAGTTT titrated into DB2713 (Figure 2C and D), the binding affinity is nearly the same ($K_D = 21$ nM). As shown in Table S2, we studied these selected compounds on sequence specificity with sequences AAATTT, AAAGTTT, and AAAGCTTT. The parent compound, DB2429, binds to AAAGTTT as 51 nM K_D . The selectivity ratios for AAATTT and AAAGCTTT are 10 and 20 respectively. When we change the *N*-substitution to ethyl (DB2689), isopropyl (DB2692), isobutyl (DB2710) and cyclopentyl (DB2713), the sequence specificity was improved. DB2689 and DB2710 show similar binding to the sequence AAAGTTT but decreases binding to no detectable binding to the pure AT sequence and two G·C bps sequences. However, DB2692 and DB2713 even improve the binding two times to our target single G sequence with extremely high sequence selectivity. All the *N*-alkyl except *N*-Me compounds in SPR studies show

no detectible binding to the narrowest minor groove of pure AT sequence because of their bulky substitution. Moreover, the minor groove of two G·C bps sequence is wider than the single G sequence. The only one G-NH₂ recognition unit (N of *N*-substitution benzimidazole) and the appropriate size of entire structure lead to much weaker or no detectible binding to AAAGCTTT. The binding affinity with target sequence AAAGTTT and the selectivity ratio compared with a similar sequence AAATTT are shown in Table 1 to indicate the selectivity of each tested compound. The original compound DB2429 with small Me group can fit into narrow minor groove like AAATTT and results in poorer selectivity (10 fold) than the compounds with bigger bulk substituents such as ethyl (DB2689), isopropyl (DB2692), isobutyl (DB2710), and cyclopentyl (DB2713). The bigger bulk would allow ligands to fit into the wider minor groove microenvironments as in AAAGTTT but blocks them from binding to narrow minor groove sequences as in AAATTT^[32]. DB2739 with *N*-phenyl substituent is the strongest binder in this series but with worse selectivity since the phenyl ring could rotate freely around the C-N bond to adjust the structure to fit different widths of the minor groove.

2.2.4. Circular dichroism (CD): Binding Mode

CD titration experiments are a reliable method for monitoring the saturation limit for compounds binding with DNA sequences and the binding mode. Spectra of CD monitor the asymmetric environment of the compounds that bind to DNA and therefore can be used to give information regarding the binding mode^[39, 40]. Free diamidines have no CD signals but on the titration of the diamidines into DNA, substantial positive CD signals (ICD) shown in the absorption region between 240 and 480 nm. These positive ICD signals confirm a minor-groove binding mode by these molecules, as previously reported with similar structures. The induced positive CD, consistent with binding in the minor groove, can be explained by nondegenerate exciton coupling with transitions in the nucleobases^[41]. ICD signals at the range of

320 to 440 nm were observed due to the gradual increase of *N*-isopropyl DB2692, *N*-cyclopentyl DB2713, and *N*-cyclobutyl DB2710 to the AAAGTTT sequence. As shown from Figure 3A, 3B and 3C of DB2692, DB2710 and DB2713 bind in the minor groove of the AAAGTTT sequences, and this agrees with fluorescence and SPR results. Finally, the CD titration experiments prove a minor-groove binding mode for the compounds in Figure 3.

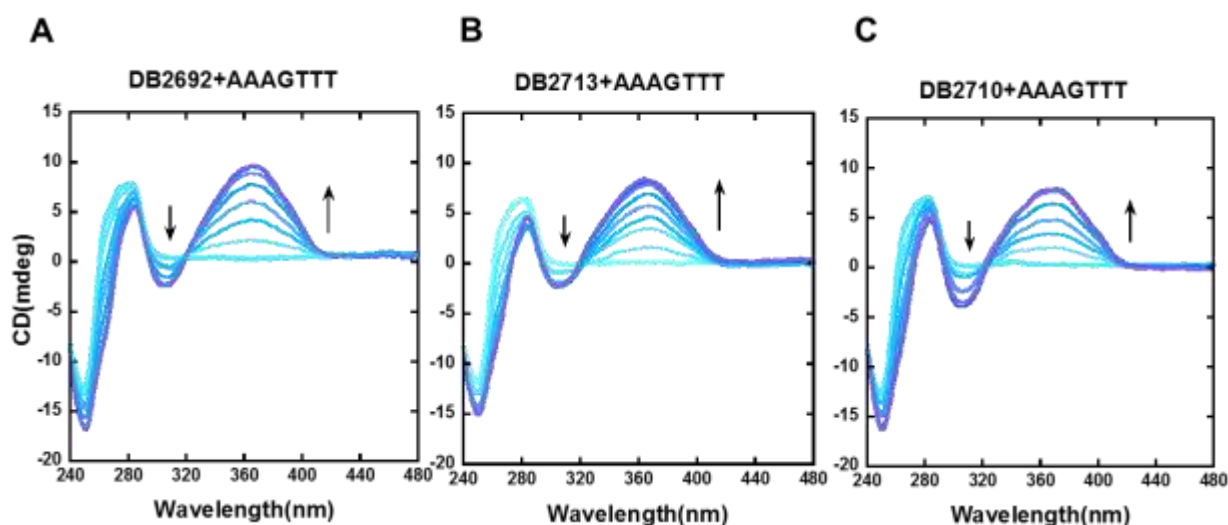


Figure 3. Circular dichroism spectra of the titration of representative compounds, A) DB2692, B) DB2713, and C) DB2710 with a 5 μ M AAAGTTT sequence in the TNE100 buffer at 25 $^{\circ}$ C . Arrows indicate the changes.

2.2.5. Competition electrospray ionization mass spectrometry (ESI-MS): Binding Stoichiometry and Relative Binding Affinity

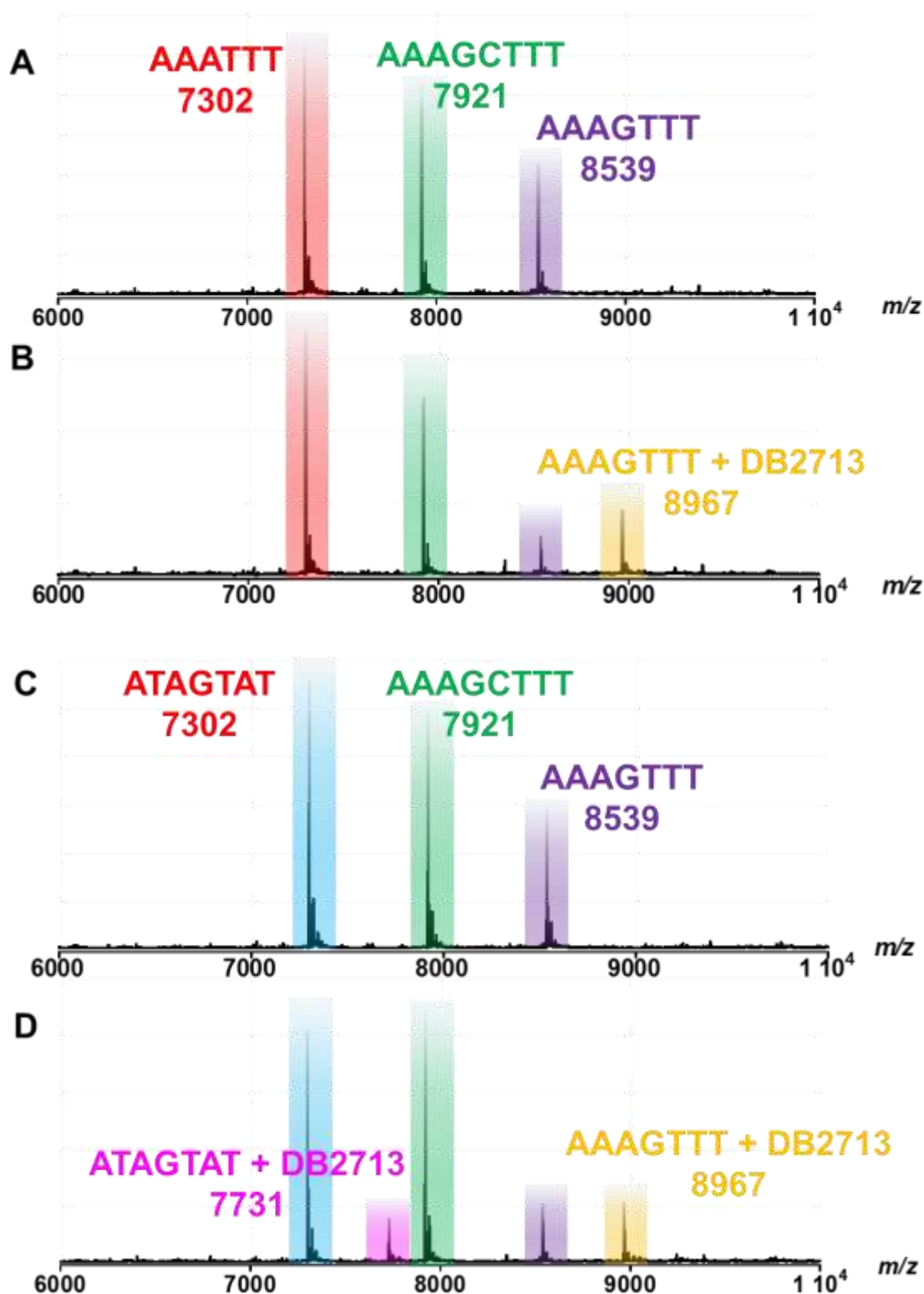


Figure 4. ESI-MS negative mode spectra of the competition binding of sequences (A and B) AAATTT (5'-CCAAATTTGCCTCTCGAAATTGG-3'), AAAGTTT (5'-GCCAAAGTTTGCCTCTGCAAAGCTTTGGC-3') and AAAGCTTT (5'-CCAAAGCTTTGCTCTCAAAGCTTTGG-3') (10 μ M each); (C and D) ATAGTAT (5'-

CCATAGTATGCTCTCATACTATGG-3'), AAAGTTT (5'-GCCAAAGTTTGCCTCTGCAAAGTTTGGC-3'), and AAAGCTTT (5'-CCAAAGCTTTGCTCTCAAAGCTTTGG-3') (10 μ M each); with 20 μ M DB2713 in buffer (50 mM ammonium acetate with 10% methanol (v/v), pH 6.8). (A and C) The ESI-MS spectra of free DNA mixture. (B and D) The ESI-MS spectra of DNA mixture with DB2713. The ESI-MS results shown here are deconvoluted spectra, and molecular weights are shown with each peak. Full DNA sequences are in Table S3.

Competition ESI-mass spectrometry (MS) provides a direct analysis of relative binding affinity, stoichiometry, and specificity for small molecule DNA complexes [42]. The use of some DNA sequences simultaneously mixed with a ligand creates a competitive binding environment for direct comparison of DNA interactions. The competition ESI-MS analysis results of the *N*-cyclopentyl DB2713 with DNA sequences AAATTT, AAAGTTT, ATAGTAT, and AAAGCTTT, are shown in Figure 4. Since the same molecular weight between sequence AAATTT and ATAGTAT which we used for MS experiments, we divide the experiments into two parts. One is shown in Figure 4 A and 4 B with sequences AAATTT, AAAGTTT, and AAAGCTTT. The other one is shown as Figure 4 C and 4 D with sequences ATAGTAT, AAAGTTT, and AAAGCTTT. The A and C plots show three peaks for the three DNA sequences as shown in the figure. On addition of the *N*-cyclopentyl DB2713, the peak for AAAGTTT (8539) decreases with the simultaneous appearance of a new peak at $m/z = 8967$ that is characteristic of a 1:1 AAAGTTT–DB2713 complex (Figure 4 B). There is no appearance of any complex peak with the other DNA sequences. In Figure 4D, with the addition of DB2713, the peak for AAAGTTT (8539) decreases a lot compared to the other two sequences with a new peak of AAAGTTT – DB2713 complex (8967) appearing. However, another peak at $m/z = 7731$ also appears as 1:1 complex of ATAGTAT and DB2713. Referring to the binding affinity results in Table S1, DB2713 binds to ATAGTAT ($K_D = 95$ nM) weaker than AAAGTTT ($K_D = 15$ nM) but much stronger than AAAGCTTT. In agreement to binding results from other experiments, results for competition ESI-

MS analysis of DB2713 with sequences AAATTT, ATAGTAT, AAAGTTT, and AAAGCTTT are done to test the sequence specificity.

2.2.6. Molecular Dynamic Simulation (MD).

To help better understand the structural basis of molecular recognition of DNA sequences with a single G·C bp in an AT context, molecular dynamics (MD) simulations for the representative *N*-isopropyl compound, DB2692, were conducted. The DB2692-A3GT3 bound complex showed that DB2692 is selectively bound to the minor groove of the AAAGTTT binding site and makes excellent van der Waals interactions with the walls of the minor groove (Figure 5 A). There are three optimum H-bonds in the complex. Both amidine groups form –N–H to T=O H-bonds (Figure 5 B) that are an average of 2.03 Å in length. The amidines also form numerous highly dynamic H-bonds to water molecules. These water molecules frequently also form H-bonds with A·T bps at the floor of the minor groove and help link the compound to the specific binding site in the groove and stabilize the complex. The third strong H-bond with 1.97 Å bond length is from the central G-NH that projects into the minor groove to the unsubstituted imidazole N in the BI group of DB2692, Figure 5 C, to account for binding selectivity of DB2692. Additional selectivity in binding is provided by the –CH group of the six-member ring of BI that points into the minor groove. This –CH forms dynamic close interactions with the N3 of dA at the floor of the minor groove (Figure 5 C). Additional direct stabilizing interactions are formed by –CH of the phenyl group that point to the floor of the minor groove with dA-N3 groups on the bases at the floor of the groove. Extensive interactions are formed by the conjugated aromatic-system of DB2692 with the sugar–phosphate back bone of the minor groove. In addition to H-bonds with the A·T bps, the amidines of DB2692 also form electrostatic interactions with the backbone phosphates.

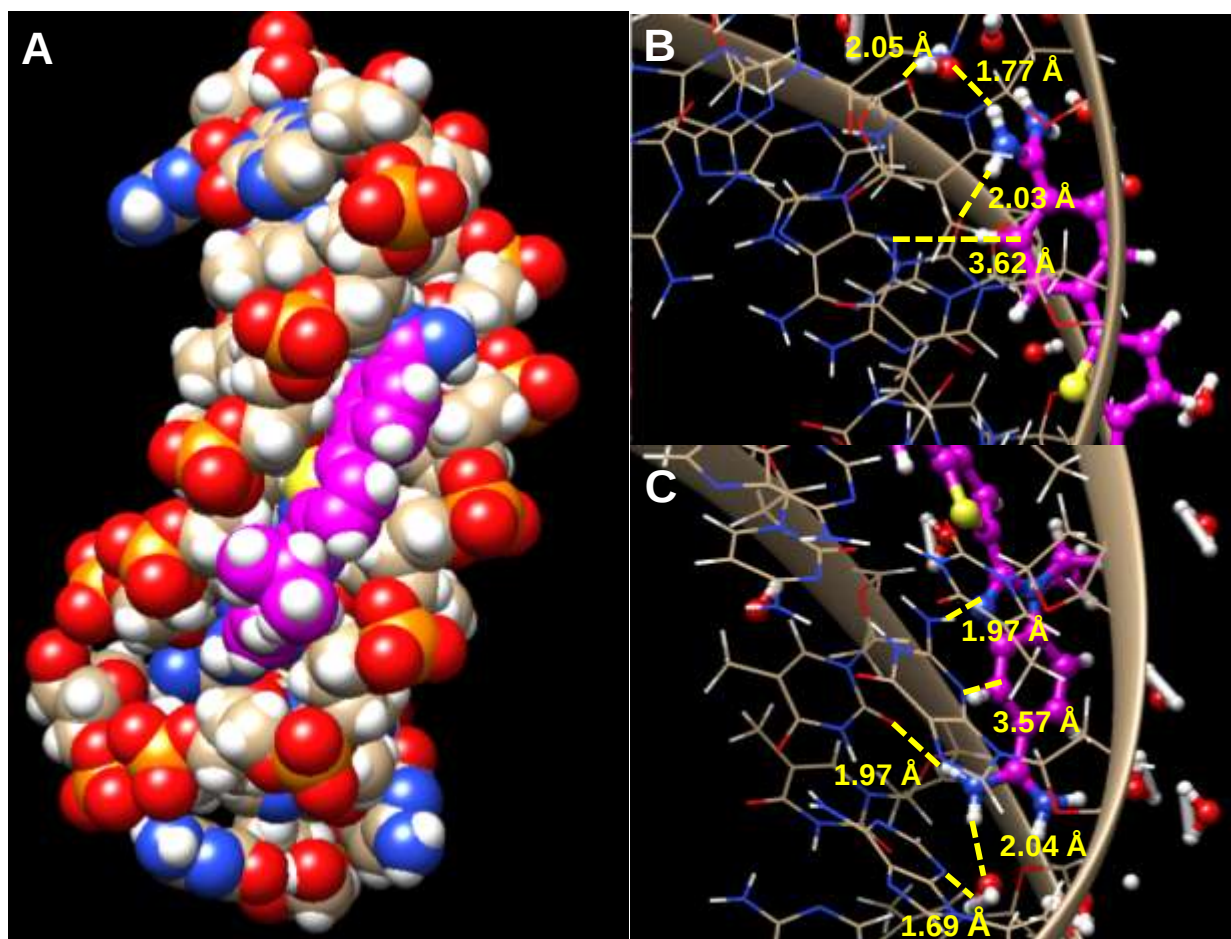


Figure 5. A, B, C. Molecular Dynamics (MD) model of DB2692 bound to an AAAGTTT site: (A) A space-filling model viewed into the minor groove of the AAAGTTT binding site with bound DB2692. The DNA bases are represented in tan-white-red-blue-orange (C-H-O-N-P) color scheme and DB2692 is magenta-white-blue-yellow (C-H-N-S) color scheme. The important interactions between different sections of the DB2692-DNA complex are illustrated in (B) and (C). The terminal amidine group forms a strong hydrogen bond with the carbonyl group of a dT (yellow dashed lines), and another direct stabilizing interaction is observed from a Ph-CH group of DB2692 which points to N3 of a dA at the floor of the minor groove (B). In (C), the imidazole-N makes a strong hydrogen bond interaction with the exocyclic NH of dG. The BI-C-H that points to the floor of the minor groove forms strong interactions with N3 of a dA.

To understand more about the sequence specificity of DB2692, we did detailed MD simulation studies with both pure AT and two GC bps containing sequences from the experimental studies. The DB2692-A3T3 bound complex showed that due to the absence of H-bond donor group at the minor groove of A3T3 sequence and also partial hydrophobic nature of the minor groove in AT sequence the unsubstituted imidazole N in the BI group of DB2692 rotates to the outside of

the minor groove. The hydrophobic bulky isopropyl group of DB2692 partly fits near the minor groove floor. This unusual orientation of the BI group places the adjacent amidine group out of the minor groove which leads to loss of the favorable amidine binding interactions.

DB2692-A3GCT3 MD simulated result showed that DB2692 does not fit tightly at the wider minor groove of A3GCT3 sequence. Due to the poor stacking interaction, the $-CHs$ from *N*-Me BI and Ph make weak bonding interactions with DNA bases and loses several favorable binding interactions. The single G·C bp mismatch at the minor groove also caused unfavorable weak amidine $NH-O=dC$ interaction and again the complex loss some favorable binding enthalpy. The MD simulation results with three sequences with different minor groove widths not only validate the experimental results on the sequence-specific binding of DB2692 at the groove of single GC bp containing sequence but also provides very important information about rationally designed organic small molecules for sequence-specific DNA recognition.

3. Conclusion.

In the current investigation, we have prepared thirteen new small size compounds of DNA sequence-specific recognition diamidines based on the structure of DB2429 to function as specific mixed DNA sequence binders. The new ligands contain *N*-alkyl-BI-thiophenes which are preorganized by both a σ -hole interaction to fit the shape of the DNA minor groove and H-bond to the $-NH$ of G·C bps that project into the minor groove. We prepared a series of diamidines with different size alkyl groups on the benzimidazole nitrogen to determine if we could take advantage of the difference in groove width created by incorporating G·C bps to the AAATTT sequences with the goal of increasing the sequence specificity of minor groove binders. All the alkyl substituents on the *N*-alkyl BI produced a recognizable

increase of selectivity for the single G·C bp sequence (Table 1) with the *N*-phenyl substituted benzimidazole diamidine DB2739 being the strongest binder to DNA with K_D of 10 nM with high selectivity to single G·C bp sequence. These results were confirmed by using different DNA binding techniques namely, Fluorescence emission spectroscopy, ΔT_m , SPR, CD, mass spectrometry, and molecular dynamics.

4. Experimental.

4.1. DNA Melting experiments

DNA thermal melting experiments were performed on a Cary 300 Bio UV-vis spectrophotometer (Varian). The concentration of each hairpin DNA sequence was 3 μ M in Tris-HCl buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 7.4) using 1 cm quartz cuvettes. The mixture solutions of DNA and ligands were tested with the ratio of 2:1 [ligand/DNA]. All samples were annealed prior to each experiment. The spectrophotometer was set at 260 nm with a 0.5 °C/min increase beginning at 25 °C, which is below the DNA melting temperature and ending above it at 95 °C. The absorbance of the buffer was subtracted, and a graph of normalized absorbance vs. temperature was created using KaleidaGraph 4.0 software. The ΔT_m values were calculated using a combination of the derivative function and estimation from the normalized graphs.

4.2. Fluorescence Titrations

Fluorescence spectra were recorded on a Cary Eclipse Spectrophotometer, with excitation and emission slit width fixed at [10, 5 nm] depending on the concentrations of ligands. The free compound solutions at different concentrations were prepared in an appropriate buffer in Tris-HCl buffer (50 mM Tris-HCl, 100

mM NaCl, 1 mM EDTA, pH 7.4), and DNA sequence (AAAGTTT) aliquots were added from a concentrated stock. All titration spectra were collected after allowing an incubation time of 10 min. All compounds were excited at 350 nm based on molecular absorbance from UV-vis spectroscopy. Emission spectra of these compounds were monitored from 400 to 620 nm wavelength. All the fluorescence titrations were performed at 25 °C. Then the Fluorescence Titration equation was made in *Kaleidagraph 4.0* to determine the K_D value.

4.3. Biosensor-Surface Plasmon Resonance (SPR).

SPR measurements were performed with a four-channel Biacore T200 optical biosensor system (GE Healthcare, Inc., Piscataway, NJ). A streptavidin-derivatized (SA) CM5 sensor chip was prepared for use by conditioning with a series of 180 s injections of 1 M NaCl in 50 mM NaOH (activation buffer) followed by extensive washing with HBS buffer (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, and 0.05% P20, pH 7.4). Biotinylated-DNA samples (AAAATTTT, AAAAGTTT, AAAGAAGTTT, AAAGAACTTT, AAAGAAACTTT, and AAGAAAACTT hairpins, Table S1) of 25-30 nM were prepared in HBS buffer and immobilized on the flow cell surface by noncovalent capture as previously described.^[43] Flow cell 1 was left blank as a reference, while flow cells 2–4 were immobilized separately by manual injection of biotinylated-DNA stock solutions (flow rate of 1 μ L/min) until the desired amount of DNA response units (RU) was obtained (250–300 RU). Ligand solutions were prepared with degassed and filtered TNE 100 with 0.05% (v/v) surfactant P20 by serial dilutions from a concentrated stock solution. Typically, a series of different ligand concentrations (2 nM to 500 nM) were injected over the DNA sensor chip at a flow rate of 100 μ L/min for 180 s, followed by buffer flow for ligand dissociation (600–1800 s). After each cycle, the sensor chip surface was regenerated with a 10 mM glycine solution (pH 2.5) for 30 s followed by multiple

buffer injections to yield a stable baseline for the following cycles. The reference response from the blank cell was subtracted from the response in each flow 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. RU was plotted as a function of free ligand concentration (C_{free}), and the equilibrium binding constants were determined with a one-site binding model ($K_2 = 0$).

$$r = (K_1 \cdot C_{\text{free}} + 2K_1 \cdot K_2 \cdot C_{\text{free}}^2) / (1 + K_1 \cdot C_{\text{free}} + K_1 \cdot K_2 \cdot C_{\text{free}}^2) \dots \dots \dots (1)$$

where r represents the moles of bound compound per mole of DNA hairpin duplex, K_1 and K_2 are macroscopic binding constants (for a single-site or 1:1 model $K_2 = 0$), and C_{free} is the free compound concentration in equilibrium with the complex. $RU_{\max}$ in the equation was used as a fitting parameter, and the obtained value was compared to the predicted maximal response per bound ligand to evaluate the stoichiometry. Kinetic analysis was performed by globally fitting the binding results for the entire concentration series using a standard 1:1 kinetic model with integrated mass transport-limited binding parameters as described previously^[44].

4.4. Circular dichroism.

CD titration experiments were performed on a Jasco J-1500 CD spectrometer in 1 cm quartz cuvette at 25 °C. Tris-HCl buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 7.4) was used. First, this buffer was added in 1 cm cuvette then buffer scan was collected before the titration experiments and then the hairpin DNA sequence AAAGTTT (5 μ M) in the same cuvette and then the ligand was added to the DNA solution and incubated for 10 min to reach equilibrium binding for the DNA-ligand complex formation. With each addition of the ligand, scan was

collected. For each titration point, three spectra were averaged from 480 to 240 nm wavelength with scan speed 50 nm/min, with a response time of 1s. Buffer scan was used as baseline, subtracted from the scan of following samples. Baseline subtracted graphs were done using the KaleidaGraph 4.0 software

4.5. Competition Electrospray Ionization Mass Spectrometry (ESI-MS).

Electrospray Ionization Mass Spectrometry (ESI-MS) analyses were performed on a Waters Q-TOF micro Mass Spectrometer (Waters Corporate, Milford, MA) equipped with an electrospray ionization source (ESI) in a negative ion mode. DNA sequences AAATTT, ATAGTAT, AAAGTTT and AAAGCTTT, Table S1, for ESI-MS experiments were purified by dialyzing it in 50 mM ammonium acetate buffer (pH 6.7) at 4 °C with 3x buffer exchange. Test samples were prepared in 50 mM ammonium acetate with 10% v/v methanol at pH 6.7 and introduced into the ion source through direct infusion at 5 µl/min flow rate. The competitive experiments were done by mixing a ligand and DNAs with different sequences at different ratios. The instrument parameters were typically as follows: capillary voltage of 2800 V, sample cone voltage of 30 V, extraction cone voltage of 1.0 V, desolvation temperature of 70 °C, and source temperature of 100 °C. Nitrogen was used as nebulizing and drying gas. A multiply charged spectra were acquired through a full scan analysis at mass range from 300-2500 Da and then deconvoluted to the spectra presented. MassLynx 4.1 software was used for data acquisition and deconvolution.

4.6. Molecular Dynamic (MD) Simulation.

Structure optimization of DB2692 was performed by using DFT/B3LYP theory with the 6-31+G* basis set in Gaussian 09 (Gaussian, Inc., 2009, Wallingford, CT) with

Gauss-view 5.09^[45]. Partial charges were derived using the RESP fitting method (Restrained Electrostatic potential)^[46, 47]. AMBER 14 (Assisted Model Building with Energy Refinement) software suite was used to perform molecular dynamic (MD) simulations^[48]. Canonical B-form ds[(5'-CCAAAGTTTGG-3')(5'-CCAAACTTTGG-3')] DNA was built in Nucleic Acid Builder (NAB) tool in AMBER. AMBER preparation and force field parameter files required to run molecular dynamic simulations for DB2692 molecule were produced using ANTECHAMBER^[49]. Specific atom types assigned for DB2692 molecule were adapted from the ff99 force field. Most of the force field parameters for DB2692 molecule were derived from the existing set of bonds, angles and dihedrals for the similar atom types in parm99 and GAFF force fields^[50]. Some dihedral angle parameters were obtained from previously reported parametrized data^[51,52]. Molecular structure with specific atom types used for the DB2692 molecule is shown in Figure S1. Parameters of DB2692 in frcmod file are listed at the Table S5.

AutoDock Vina program was used to dock the DB2692 in the minor groove of DNA to obtain the initial structure for DB2692-DNA complex^[53]. MD simulations were performed in explicit solvation conditions where the DNA-DB2692 complex was placed in a truncated octahedron box filled with TIP3P water using xleap program in AMBER. Sodium ions were used to neutralize the system. A 10 Å cutoff was applied on all van der Waals interactions. The MD simulation was carried out using the Sander module with SHAKE algorithm applied to constrain all bonds. Initially, the system was relaxed with 500 steps of steepest-descent energy minimization. The temperature of the system was then increased from 0 K to 310 K for over 10 ps under constant-volume conditions. In the final step, the production run on the system was subsequently performed for 500 ns under NPT (constant-pressure) conditions.

4.7. Chemistry.

All commercial reagents were used without further purification. All melting points were determined on a Mel-Temp 3.0 melting point instrument, and are uncorrected. TLC analysis was carried out on silica gel 60 F254 precoated aluminum sheets using UV light for detection. ^1H NMR and ^{13}C NMR spectra were recorded on a Bruker 400 MHz and 100 MHz spectrometer respectively using the indicated solvents. Mass spectra were obtained from the Georgia State University Mass Spectrometry Laboratory, Atlanta, GA. Elemental analysis were performed by Atlantic Microlab Inc., Norcross, GA. Compounds **5a** and **6a** (DB 2429) was prepared according to the procedure published in reference 21.

Synthesis of 4-(alkylamino)-3-nitrobenzonitrile (2a-k)

Amines (40 mmol) were added to 4-fluoro-3-nitrobenzonitrile (3.32 gm, 20 mmol) suspension in ethanol (20 ml) and stirred at room temperature for 24 h. The precipitated yellow solid was filtered and washed with cooled ethanol and dried.

In case of aniline, the reaction mixture was heated at 80 °C for 24h.

4-(Methylamino)-3-nitrobenzonitrile (2a)

Orange solid (3.39 gm 89 %), mp 168-169 °C (as reported, ^[54]). ^1H NMR (DMSO- d_6): δ 8.64 (br s, 1 H), 8.49 (s, 1 H), 7.83 (d, J = 9 Hz, 1 H), 7.09 (d, J = 9 Hz, 1 H), 2.99 (s, 3 H); ESI-HRMS: m/z calculated for $\text{C}_8\text{H}_8\text{N}_3\text{O}_2$: 177.0533, found: 177.0524 (M^+ + 1).

4-(Ethylamino)-3-nitrobenzonitrile (2b).

Orange solid (3.39 gm 89 %), mp 134-135 °C (as reported, ^[54]). ^1H NMR (DMSO- d_6): δ 8.56 (br s, 1 H), 8.49 (d, J = 1.2 Hz, 1 H), 7.80 (d, J = 9.2 Hz, 1 H), 7.16 (d, J = 9.2 Hz, 1 H), 3.45 (q, J = 6.4 Hz, 2 H), 3.45 (t, J = 7.2 Hz, 3 H); ESI-HRMS: m/z calculated for $\text{C}_9\text{H}_{10}\text{N}_3\text{O}_2$: 192.0768, found: 192.0761 (M^+ + 1).

4-(Isopropylamino)-3-nitrobenzonitrile (2c).

Orange solid (3.36 gm 82 %), mp 111-112 °C (Lit. mp 110-113 °C, ^[55]). ¹HNMR (DMSO-d₆): δ 8.49 (s, 1 H), 8.20 (d, *J* = 7.2 Hz, 1 H), 7.81 (d, *J* = 9.2 Hz, 1 H), 7.21 (d, *J* = 9.2 Hz, 1 H), 4.01 (m, 1 H), 1.27 (d, *J* = 6 Hz, 6 H); ¹³CNMR (DMSO-d₆): 145.5, 137.1, 131.3, 130.3, 117.5, 115.4, 96, 43.7, 21.4; ESI-HRMS: *m/z* calculated for C₁₀H₁₁N₃O₂: 228.0749, found: 228.0759 (M⁺ + Na).

4-(Isobutylamino)-3-nitrobenzonitrile (2d).

Orange solid (3.47 gm 79 %), mp 77-78 °C. ¹HNMR (DMSO-d₆): δ 8.61 (br s, 1 H), 8.49 (s, 1 H), 7.79 (d, *J* = 8.8 Hz, 1 H), 7.19 (d, *J* = 8.8 Hz, 1 H), 3.25 (t, *J* = 6.4 Hz, 2 H), 1.94 (m, 1 H), 0.94 (d, *J* = 6.4 Hz, 6 H); ESI-HRMS: *m/z* calculated for C₁₁H₁₄N₃O₂: 220.1081, found: 220.1070 (M⁺ + 1).

4-(Neopentylamino)-3-nitrobenzonitrile (2e).

Orange solid (3.3 gm 71 %), mp 126-127 °C. ¹HNMR (DMSO-d₆): δ 8.52 (m, 2 H), 7.81 (dd, *J* = 1.2, 9 Hz, 1 H), 7.93 (d, *J* = 9 Hz, 1 H), 3.27 (d, *J* = 6 Hz, 2 H), 0.98 (s, 9H); ¹³CNMR (DMSO-d₆): 147.1, 137.1, 131.3, 130.2, 117.5, 115.4, 96.1, 53, 31.4, 26.5; ESI-HRMS: *m/z* calculated for C₁₂H₁₆N₃O₂: 234.1237, found: 234.1226 (M⁺ + 1).

3-nitro-4-((2,2,2-trifluoroethyl)amino)benzonitrile (2f).

Orange solid (3.82 gm 79 %), mp 87-88 °C. ¹HNMR (DMSO-d₆): δ 8.72(m, 1 H), 8.56 (s, 1 H), 7.93 (d, *J* = 9.2 Hz, 1 H), 7.45 (d, *J* = 9.2 Hz, 1 H), 4.43 (m, 2H); ESI-HRMS: *m/z* calculated for C₉H₇FN₃O₂: 246.0458, found: 246.0476 (M⁺ + 1).

4-((3-Methoxypropyl)amino)-3-nitrobenzonitrile (2g).

Orange solid (3.52 gm 75 %), mp 68-69 °C. ¹HNMR (DMSO-d₆): δ 8.76 (br s, 1 H), 8.49 (s, 1 H), 7.82 (d, *J* = 8.8 Hz, 1 H), 7.15 (d, *J* = 8.8 Hz, 1 H), 3.45 (m, 4 H), 3.26 (s, 3 H), 1.86 (t, *J* = 5.6 Hz, 2 H); ¹³CNMR (DMSO-d₆): 146.3, 137.1, 131.3, 130.2, 117.7, 115, 95.7, 69.5, 57.5, 40.2, 27.5; ESI-HRMS: *m/z* calculated for C₁₁H₁₄N₃O₃: 236.1030, found: 236.1019 (M⁺ + 1).

4-(Cyclobutylamino)-3-nitrobenzonitrile (2h).

Orange solid (3.55 gm 82 %), mp 107-108 °C. ¹HNMR (DMSO-d₆): δ 8.50 (d, *J* = 1.6 Hz, 1 H), 8.46 (d, *J* = 5.2 Hz, 1 H), 7.82 (dd, *J* = 1.6, 9.2 Hz, 1 H), 7.03 (d, *J* = 9.2 Hz, 1 H), 4.20 (m, 1H), 2.44 (m, 2H), 2.05 (m, 2H), 1.80 (m, 2H); ¹³CNMR (DMSO-d₆): 145, 137.2, 131.2, 130.3, 117.6, 115.6, 96.4, 46.9, 29.2, 14.3; ESI-HRMS: *m/z* calculated for C₁₁H₁₂N₃O₂: 218.0924, found: 218.0914 (M⁺ + 1).

4-(Cyclopentylamino)-3-nitrobenzonitrile (2i).

Orange solid (3.68 gm 80 %), mp 105-106 °C. ¹HNMR (DMSO-d₆): δ 8.50 (br s, 1 H), 8.28 (d, *J* = 5.6 Hz, 1 H), 7.83 (d, *J* = 9 Hz, 1 H), 7.21 (d, *J* = 9 Hz, 1 H), 4.11 (m, 1H), 2.06 (m, 2H), 1.70 (m, 2H), 1.62 (m, 4H); ¹³CNMR (DMSO-d₆): 145.8, 137, 131, 130, 117.6, 115.8, 96, 53.4, 23.1; ESI-HRMS: *m/z* calculated for C₁₂H₁₃N₃O₂: 231.1002, found: 231.0991 (M⁺).

4-(Cyclohexylamino)-3-nitrobenzonitrile (2j).

Orange solid (3.77 gm 77 %), mp 113-114 °C, ¹HNMR (DMSO-d₆): δ 8.50 (br s, 1 H), 8.28 (d, *J* = 7.4 Hz, 1 H), 7.80 (d, *J* = 9.2 Hz, 1 H), 7.26 (d, *J* = 9.2 Hz, 1 H), 3.72 (m, 1H), 1.93 (m, 2H), 1.69 (m, 2H), 1.59 (m, 1H), 1.42 (m, 4H), 1.23 (m, 1H); ¹³CNMR (DMSO-d₆): 145.4, 137.1, 131.4, 130, 117.6, 115.6, 96, 50.1, 31.1, 24.4, 23.4; ESI-HRMS: *m/z* calculated for C₁₃H₁₆N₃O₂: 246.1237, found: 246.1227 (M⁺ + 1).

3-Nitro-4-(phenylamino)benzonitrile (2k).

Orange solid (2.54 gm 53 %), mp 128-129 °C, ¹HNMR (DMSO-d₆): δ 9.91 (br s, 1 H), 8.58 (d, *J* = 1.6 Hz, 1 H), 7.76 (dd, *J* = 1.6, 9.2 Hz, 1 H), 7.49 (m, 2 H), 7.34 (m, 3 H), 7.06 (d, *J* = 8.8 Hz, 1 H); ¹³CNMR (DMSO-d₆): 144.6, 137.3, 137, 131.9, 131.2, 129.2, 126, 124.9, 117.4, 116.7, 98.1; ESI-HRMS: *m/z* calculated for C₁₃H₁₀N₃O₂: 240.0768, found: 240.0765 (M⁺ + 1).

Synthesis of 3-amino-4-(alkyl- or phenyl-amino)benzonitrile (3a-k)

SnCl₂ · 2H₂O (4.5 gm, 20 mmol) was added to a suspension of the nitro compound **2** (5mmol) in ethanol (30 ml). The reaction mixture was refluxed for 12 h and concentrated under reduced pressure. The formed residue was neutralized by sodium hydroxide solution in an ice bath. The formed precipitate was filtered, dried under vacuum at room temperature and then dissolved in acetone (50 ml) and filtered. The filtrate was evaporated under reduced pressure and dried under vacuum at room temperature and used in the next step without further characterization.

Synthesis of the bisnitrile compounds (**5 a-n**)^[56,57].

Sodium metabisulphite (1.14 g, 6 mmol) was added to a solution of the diamine **3a-k** (3 mmol) and the aldehyde **4a-d** (3 mmol) in DMSO (10 mL) and the mixture was heated at 140 °C for 30 min. The reaction mixture was poured into water, filtered and dried. Purification was by crystallization from acetone.

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-ethyl-1H-benzo[d]imidazole-5-carbonitrile (5b).

Yellow solid (0.73 gm, 69 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.23 (br s, 1 H), 7.99 (d, *J* = 7.6 Hz, 2 H), 7.94 (br s, 1 H), 7.91 (m, 2 H), 7.87(d, *J* = 7.6 Hz, 1H), 7.74 (d, *J* = 8.4 Hz, 1 H), 4.62 (q, *J* = 6.8 Hz, 2H), 1.42 (t, *J* = 6.8 Hz, 3H); ESI-HRMS: *m/z* calculated for C₂₁H₁₅N₄S: 355.1012, found: 355.0995 (M⁺ + 1).

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-isopropyl-1H-benzo[d]imidazole-5-carbonitrile (5c).

Yellow solid (0.81 gm, 74 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.23 (br s, 1 H), 8.06 (d, *J* = 9 Hz, 1 H), 7.98 (d, *J* = 8.4 Hz, 2 H), 7.92 (d, *J* = 8.4 Hz, 2 H), 7.88 (d, *J* = 4 Hz, 1H), 7.71 (d, *J* = 4 Hz, 1H), 7.64 (d, *J* = 9 Hz, 1 H), 5.19 (m, 1H), 1.68 (m, 6H); ESI-HRMS: *m/z* calculated for C₂₂H₁₇N₄S: 369.1168, found: 369.1159 (M⁺ + 1).

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-isobutyl-1H-benzo[d]imidazole-5-carbonitrile (5d).

Yellow solid (0.75 gm, 66 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.21 (br s, 1 H), 7.95 (m, 2 H), 7.91 (m, 3 H), 7.86 (br s, 2 H), 7.67 (d, *J* = 8 Hz, 1H), 4.44 (d, *J* = 6 Hz, 2 H), 2.10 (m, 1H), 0.86 (d, *J* = 5.2 Hz, 6 H); ESI-HRMS: *m/z* calculated for C₂₃H₁₉N₄S: 383.1325, found: 383.1306 (*M*⁺ + 1).

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-neopentyl-1H-benzo[d]imidazole-5-carbonitrile (5e).

Yellow solid (0.79 gm, 67 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.23 (br s, 1 H), 7.98 (m, 4 H), 7.92 (d, *J* = 8.4 Hz, 2 H), 7.87 (d, *J* = 4 Hz, 1H), 7.68 (dd, *J* = 1, 8.6 Hz, 1H), 7.64 (d, *J* = 9 Hz, 1 H), 4.56 (s, 2H), 0.79 (s, 9H); ESI-HRMS: *m/z* calculated for C₂₄H₂₁N₄S: 397.1481, found: 397.1462 (*M*⁺ + 1).

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-(3-methoxypropyl)-1H-benzo[d]imidazole-5-carbonitrile (5f).

Yellow solid (0.70 gm, 58 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.18 (br s, 1 H), 7.92 (d, *J* = 8.4 Hz, 2H), 7.88 (d, *J* = 8.4 Hz, 2H), 7.84 (br s, 2H), 7.78 (d, *J* = 8.2 Hz, 1H), 7.66 (d, *J* = 8.2 Hz, 1H), 4.58 (m, 2H), 3.33 (m, 2H), 3.19 (s, 3H), 2.02 (m, 2H); ESI-HRMS: *m/z* calculated for C₂₃H₁₉N₄OS: 399.1274, found: 399.1262 (*M*⁺ + 1).

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-(2,2,2-trifluoroethyl)-1H-benzo[d]imidazole-5-carbonitrile (5g).

Yellow solid (0.70 gm, 58 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.30 (br s, 1 H), 8.01 (m, 4 H), 7.93 (m, 3 H), 7.80 (d, *J* = 8.4 Hz, 1H), 5.71 (m, 2H); ESI-HRMS: *m/z* calculated for C₂₁H₁₂F₃N₄S: 409.0729, found: 409.0710 (*M*⁺ + 1).

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-cyclobutyl-1H-benzo[d]imidazole-5-carbonitrile (5h).

Yellow solid (0.75 gm, 66 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.24 (br s, 1 H), 8.06 (d, *J* = 8.6 Hz, 1H), 7.97 (d, *J* = 8.2 Hz, 2H), 7.92 (d, *J* = 8.2 Hz, 2H), 7.88 (d, *J* = 3.8 Hz, 1H), 7.75 (d, *J* = 3.8 Hz, 1H), 7.67 (d, *J* = 8.6 Hz, 1H), 5.44 (m, 1H), 2.73 (m, 2H), 2.57 (m, 2H), 1.91 (m, 2H); ESI-HRMS: *m/z* calculated for C₂₃H₁₇N₄S: 381.1168, found: 381.1155 (M⁺ + 1).

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-cyclopentyl-1*H*-benzo[d]imidazole-5-carbonitrile (5i).

Yellow solid (0.81 gm, 69 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.23 (m, 1 H), 7.98 (d, *J* = 7.6 Hz, 2H), 7.93 (d, *J* = 7.6 Hz, 2H), 7.88 (br s, 1H), 7.83 (d, *J* = 8.8 Hz, 1H), 7.77 (d, *J* = 4 Hz, 1H), 7.65 (d, *J* = 8.8 Hz, 1H), 5.30 (m, 1H), 2.21 (m, 4H), 2.05 (m, 2H), 1.76 (m, 2H); ESI-HRMS: *m/z* calculated for C₂₄H₁₉N₄S: 395.1325, found: 395.1315 (M⁺ + 1).

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-cyclohexyl-1*H*-benzo[d]imidazole-5-carbonitrile (5j).

Yellow solid (0.74 gm, 61 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.24 (br s, 1 H), 8.12 (br s, 1H), 7.95 (m, 5H), 7.66 (m, 1H), 4.71 (m, 1H), 2.31 (m, 2H), 2.00 (m, 5H), 1.69 (br s, 1H), 1.45 (m, 2H); ESI-HRMS: *m/z* calculated for C₂₅H₂₁N₄S: 409.1481, found: 409.1465 (M⁺ + 1).

2-(5-(4-Cyanophenyl)thiophen-2-yl)-1-phenyl-1*H*-benzo[d]imidazole-5-carbonitrile (5k).

Yellow solid (0.70 gm, 59 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.31 (br s, 1 H), 7.87 (m, 4H), 7.74 (m, 5H), 7.61 (m, 2H), 7.19 (br s, 1 H), 6.71 (br s, 1 H); ESI-HRMS: *m/z* calculated for C₂₅H₁₅N₄S: 403.1012, found: 403.0992 (M⁺ + 1).

2-(5-(4-Cyanophenyl)selenophen-2-yl)-1-methyl-1*H*-benzo[d]imidazole-5-carbonitrile (5l).

Yellow solid (0.75 gm 65 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 8.17 (br s, 1 H), 8.04 (m, 2 H), 7.88 (m, 4 H), 7.83 (d, *J* = 6.8 Hz, 1 H), 7.66 (d, *J* = 6.8 Hz, 1 H),

4.08 (s, 3 H); ESI-HRMS: m/z calculated for $C_{20}H_{13}N_4Se$: 389.0300, found: 389.0303 ($M^+ + 1$).

2-(2-(4-Cyanophenyl)thiazol-5-yl)-1H-benzo[d]imidazole-5-carbonitrile (5m).

Yellow solid (0.50 gm, 51 %), mp > 300 °C. 1H NMR (DMSO- d_6): δ 8.71 (br s, 1 H), 8.19 (m, 3H), 8.01 (d, J = 7.6 Hz, 2H), 7.71 (m, 1H), 7.64 (d, J = 7.2 Hz, 1H); ESI-HRMS: m/z calculated for $C_{18}H_{10}N_5S$: 328.0651, found: 328.0639 ($M^+ + 1$).

2-(2-(4-Cyanophenyl)thiazol-5-yl)-1-methyl-1H-benzo[d]imidazole-5-carbonitrile (5n).

Yellow solid (0.70 gm 69 %), mp > 300 °C. 1H NMR (DMSO- d_6): δ 8.73 (br s, 1 H), 8.23 (m, 3 H), 8.02 (d, J = 6.8 Hz, 2 H), 7.91 (d, J = 8.4 Hz, 1 H), 7.72 (d, J = 8.4 Hz, 1 H), 4.14 (s, 3 H); ESI-HRMS: m/z calculated for $C_{19}H_{12}N_5S$: 342.0808, found: 342.0798 ($M^+ + 1$).

Synthesis of the diamidines (6a-n) ^[58-60].

Method 1.

The above bis-nitriles (1 mmol) were suspended in freshly distilled THF (5 mL), and treated with a 1M LiN(TMS) $_2$ in THF solution (6 mL, 6.0 mmol) and the mixture was stirred for 24 h at room temperature. The reaction mixture was cooled to 0 °C and HCl saturated ethanol (3 mL) was carefully added. The mixture was stirred for 2 h, ether was added and the resultant solid was collected by filtration. The diamidine was purified by neutralization with 1M sodium hydroxide solution followed by filtration of the resultant solid, washed with water and dried. The free base was stirred with ethanolic HCl for 24 h, acetone was added, and the solid that formed was filtered and dried under vacuum at 100 °C for 24 h. This method was used to prepare all final compounds except **6a-m**.

Method 2.

The bis-nitrile (1 mmol) was suspended in dry ethanol (20 ml) and cooled in ice bath, HCl gas was bubbled through the reaction mixture for 30 min and the flask was tightly sealed and stirred at room temperature for 7 days. The formed yellow precipitate of imidate ester hydrochloride was filtered and washed with dry ethanol and anhydrous ether and dried under vacuum at room temperature for 3 h. The imidate ester was suspended in dry ethanol and cooled in ice bath, ammonia gas was bubbled through the reaction mixture for 30 min and the flask was tightly sealed and stirred at room temperature for 2 days. The formed yellow precipitate was filtered and washed with dry ethanol and acetone and dried under vacuum at 100 °C for 12 h. This method was used to prepare compounds **6g**, **6m**.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-ethyl-1*H*-benzo[d]imidazole-5-carboximidamide (6b).

Yellow solid (0.26 gm, 48 %), mp > 300 °C. ¹HNMR (DMSO-*d*₆): δ 9.57 (s, 2 H), 9.48 (s, 2 H), 9.35 (s, 2 H), 9.27 (s, 2 H), 8.27 (s, 1 H), 8.05 (d, *J* = 8.2 Hz, 2H), 7.98 (d, *J* = 8.2 Hz, 2H), 7.95 (m, 2H), 7.92 (d, *J* = 3.6 Hz, 1H), 7.83 (d, *J* = 8.4 Hz, 1H), 4.66 (m, 2H), 1.44 (t, *J* = 6.8 Hz, 3H); ¹³CNMR (DMSO-*d*₆): 165.4, 164.4, 147.8, 144.2, 141.2, 138.8, 137.1, 131.8, 129.3, 128.7, 126.8, 126.6, 125.3, 122, 121.3, 119, 110.7, 39.6, 14.3; ESI-HRMS: *m/z* calculated for C₂₁H₂₂N₆S: 195.0808, found: 195.0800 (Amidine base M⁺ + 2, double charged). Anal. Calcd. For C₂₁H₂₀N₆S. 3HCl. 2.25H₂O. 0.35C₃H₆O: C, 47.51; H, 5.35; N, 15.08. Found: C, 47.63; H, 4.98; N, 14.68.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-isopropyl-1*H*-benzo[d]imidazole-5-carboximidamide (6c).

Yellow solid (0.27 gm, 50 %), mp > 300 °C. ¹HNMR (DMSO-*d*₆): δ 9.59 (s, 2 H), 9.52 (s, 2 H), 9.39 (s, 2 H), 9.32 (s, 2 H), 8.31 (s, 1 H), 8.14 (d, *J* = 8.8 Hz, 1H), 8.04 (d, *J* = 8.4 Hz, 2H), 8.00 (d, *J* = 8.4 Hz, 2H), 7.99 (d, *J* = 3.8 Hz, 1H), 7.81 (br s, 1

H), 7.78 (d, $J = 3.8$ Hz, 1H), 5.22 (m, 1H), 1.44 (d, $J = 6.8$ Hz, 6H); ^{13}C NMR (DMSO- d_6): 165.2, 164.4, 147.8, 144.7, 141, 137, 136.6, 130.8, 130.4, 128.7, 126.7, 126.4, 125.3, 122, 121.4, 119, 113.2, 49.2, 20.4; ESI-HRMS: m/z calculated for $\text{C}_{22}\text{H}_{24}\text{N}_6\text{S}$: 202.0886, found: 202.0876 (Amidine base $\text{M}^+ + 2$, double charged). Anal. Calcd. For $\text{C}_{22}\text{H}_{22}\text{N}_6\text{S} \cdot 3\text{HCl} \cdot 1.75\text{H}_2\text{O}$: C, 48.74; H, 5.30; N, 15.51. Found: C, 48.72; H, 5.30; N, 15.21.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-isobutyl-1H-benzo[d]imidazole-5-carboximidamide (6d).

Yellow solid (0.12 gm, 21 %), mp > 300 °C. ^1H NMR (DMSO- d_6): δ 9.53 (s, 2 H), 9.43 (s, 2 H), 9.30 (s, 2 H), 9.21 (s, 2 H), 8.25 (br s, 1 H), 8.05 (d, $J = 7.6$ Hz, 2H), 7.98 (m, 2H), 7.93 (m, 3H), 7.79 (d, $J = 8.4$ Hz, 1H), 4.50 (d, $J = 6$ Hz, 2H), 2.16 (m, 1 H), 8.88 (d, $J = 5.6$ Hz, 6H); ^{13}C NMR (DMSO- d_6): 165.4, 164.4, 148.2, 144, 141, 140, 137.1, 136, 132.3, 129.3, 128.7, 126.7, 125.3, 122, 121.3, 119, 111.4, 50.7, 28.3, 19; ESI-HRMS: m/z calculated for $\text{C}_{23}\text{H}_{26}\text{N}_6\text{S}$: 209.0964, found: 209.0965 (Amidine base $\text{M}^+ + 2$, double charge). Anal. Calcd. For $\text{C}_{23}\text{H}_{24}\text{N}_6\text{S} \cdot 3\text{HCl} \cdot 1.7\text{H}_2\text{O}$: C, 49.75; H, 5.52; N, 15.14. Found: C, 50.13; H, 5.35; N, 14.76.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-neopentyl-1H-benzo[d]imidazole-5-carboximidamide (6e).

Yellow solid (0.12 gm, 20 %), mp > 300 °C. ^1H NMR (DMSO- d_6): δ 9.61 (s, 2 H), 9.52 (s, 2 H), 9.41 (s, 2 H), 9.33 (s, 2 H), 8.30 (br s, 1 H), 8.04 (m, 6H), 7.91 (d, $J = 3.6$ Hz, 1H), 7.83 (d, $J = 8.4$ Hz, 1H), 4.61 (s, 2 H), 0.82 (s, 9 H); ^{13}C NMR (DMSO- d_6): 165.3, 164.3, 148.8, 143.9, 140.6, 140.1, 137.1, 132.6, 130.2, 128.7, 126.5, 126.3, 125.2, 121.8, 121.2, 119, 112.6, 53.8, 34.6, 27.4; ESI-HRMS: m/z calculated for $\text{C}_{24}\text{H}_{28}\text{N}_6\text{S}$: 2016.1043, found: 2016.1033 (Amidine base $\text{M}^+ + 2$, double charge). Anal. Calcd. For $\text{C}_{24}\text{H}_{26}\text{N}_6\text{S} \cdot 3\text{HCl} \cdot 2.75\text{H}_2\text{O}$: C, 49.00; H, 5.91; N, 14.29. Found: C, 48.78; H, 5.59; N, 14.00.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-(3-methoxypropyl)-1H-benzo[d]imidazole-5-carboximidamide (6f).

Yellow solid (0.26 gm, 43 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 9.60 (s, 2 H), 9.51 (s, 2 H), 9.38 (s, 2 H), 9.30 (s, 2 H), 8.28 (br s, 1H), 8.03 (d, *J* = 8.4 Hz, 2H), 8.03 (d, *J* = 8.4 Hz, 2H), 7.94 (m, 2H), 7.90 (d, *J* = 8.4 Hz, 1H), 7.83 (d, *J* = 8.4 Hz, 1H), 4.66 (t, *J* = 6.8 Hz, 2H), 3.35 (t, *J* = 5.6 Hz, 2H), 3.19 (s, 3 H), 2.06 (m, 2 H); ¹³CNMR (DMSO-d₆): 165.3, 164.4, 148, 144.1, 141, 139.3, 137.1, 132, 129.2, 128.7, 126.7, 126.5, 125.2, 122, 121.2, 119, 110.7, 68, 57.5, 41.4, 28.7; ESI-HRMS: *m/z* calculated for C₂₃H₂₆N₆OS: 217.0939, found: 217.0931 (Amidine base M⁺ + 2, double charged). Anal. Calcd. For C₂₃H₂₄N₆OS. 3HCl. 3H₂O: C, 46.45; H, 5.59; N, 14.14. Found: C, 46.18; H, 5.22; N, 13.77.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-(2,2,2-trifluoroethyl)-1H-benzo[d]imidazole-5-carboximidamide (6g).

Yellow solid (0.15. gm, 25 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 9.51 (s, 2 H), 9.45 (s, 2 H), 9.26 (s, 2 H), 9.21 (s, 2 H), 8.28 (d, *J* = 1.6 Hz, 1H), 8.06 (m, 3H), 8.03 (d, *J* = 4 Hz, 1H), 7.97 (br s, 1H), 7.95 (m, 2H), 7.84 (dd, *J* = 1.6, 8.8 Hz, 1H), 5.76 (m, 2 H); ESI-HRMS: *m/z* calculated for C₂₁H₁₉F₃N₆S: 222.0667, found: 222.0658 (Amidine base M⁺ + 2, double charge). Anal. Calcd. For C₂₁H₁₇F₃N₆S. 3HCl. 2.75H₂O: C, 42.02; H, 4.28; N, 14.01. Found: C, 42.03; H, 4.03; N, 13.61.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-cyclobutyl-1H-benzo[d]imidazole-5-carboximidamide (6h).

Yellow solid (0.38 gm, 68 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 9.60 (s, 2 H), 9.56 (s, 2 H), 9.39 (s, 2 H), 9.33 (s, 2 H), 8.31 (br s, 1H), 8.13 (d, *J* = 8.8 Hz, 1H), 8.03 (d, *J* = 8.4 Hz, 2H), 7.99 (d, *J* = 8.4 Hz, 2H), 7.93 (d, *J* = 3.8 Hz, 1H), 7.84 (br s, 1H), 7.81 (d, *J* = 3.8 Hz, 1H), 5.47 (m, 1H), 2.77 (m, 2H), 3.59 (m, 2 H), 1.94 (m, 2H); ¹³CNMR (DMSO-d₆): 165, 164.3, 147.7, 145.1, 139.5, 137.1, 137, 132, 129.3, 128.7, 126.6, 126.3, 125.3, 122.4, 121.8, 118.6, 113, 50.6, 28.5, 14; ESI-HRMS: *m/z*

calculated for $C_{23}H_{24}N_6S$: 208.0886, found: 208.0887 (Amidine base $M^+ + 2$, double charge). Anal. Calcd. For $C_{23}H_{22}N_6S \cdot 3HCl \cdot 1.75H_2O$: C, 49.85; H, 5.18; N, 15.17. Found: C, 49.95; H, 5.01; N, 14.78.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-cyclopentyl-1H-benzo[d]imidazole-5-carboximidamide (6i).

Yellow solid (0.15 gm, 25 %), mp > 300 °C. 1H NMR (DMSO- d_6): δ 9.57 (s, 2 H), 9.50 (s, 2 H), 9.35 (s, 2 H), 9.28 (s, 2 H), 8.03 (d, J = 1.2 Hz, 1H), 8.05 (d, J = 8.4 Hz, 2H), 7.98 (d, J = 8.4 Hz, 2H), 7.94 (d, J = 3.2 Hz, 1H), 7.91 (m, 1H), 7.82 (m, 2H), 5.34 (m, 1H), 2.24 (m, 4H), 2.05 (m, 2 H), 1.79 (m, 2H); ESI-HRMS: m/z calculated for $C_{24}H_{26}N_6S$: 215.0964, found: 215.0959 (Amidine base $M^+ + 2$, double charged). Anal. Calcd. For $C_{24}H_{24}N_6S \cdot 3HCl \cdot 3.5H_2O$: C, 48.06; H, 5.71; N, 14.02. Found: C, 47.95; H, 5.52; N, 13.77.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-cyclohexyl-1H-benzo[d]imidazole-5-carboximidamide (6j).

Yellow solid 0.20 m, 34 %), mp > 300 °C. 1H NMR (DMSO- d_6): δ 9.56 (s, 2 H), 9.50 (s, 2 H), 9.33 (s, 2 H), 9.25 (s, 2 H), 8.29 (br s, 1H), 8.17 (d, J = 8.4 Hz, 1H), 8.04 (d, J = 8.4 Hz, 2H), 7.98 (d, J = 8.4 Hz, 2H), 7.95 (d, J = 4.2 Hz, 1H), 7.77 (d, J = 8.4 Hz, 1H), 7.71 (d, J = 4.2 Hz, 1H), 4.73 (m, 1H), 2.34 (m, 2H), 2.01 (m, 2H), 1.90 (m, 2 H), 1.70 (br s, 1H), 1.48 (m, 3H); ^{13}C NMR (DMSO- d_6): 165.2, 164.4, 148, 144.6, 141.2, 137, 136.9, 131.5, 130.7, 128.7, 126.7, 126.5, 125.3, 121.9, 121.3, 119.1, 113.5, 57, 30, 24.9, 23.7; ESI-HRMS: m/z calculated for $C_{25}H_{28}N_6S$: 222.1043, found: 222.1034 (Amidine base $M^+ + 2$, double charge). Anal. Calcd. For $C_{25}H_{26}N_6S \cdot 3HCl \cdot 1.75H_2O$: C, 51.57; H, 5.63; N, 14.44. Found: C, 51.81; H, 5.61; N, 14.09.

2-(5-(4-Carbamimidoylphenyl)thiophen-2-yl)-1-phenyl-1H-benzo[d]imidazole-5-carboximidamide (6k).

Yellow solid (0.26 gm, 45 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 9.54 (s, 2 H), 9.52 (s, 2 H), 9.33 (s, 2 H), 9.30 (s, 2 H), 8.37 (br s, 1H), 7.94 (d, *J* = 8.6 Hz, 2H), 7.90 (d, *J* = 8.6 Hz, 2H), 7.77 (m, 4H), 7.69 (m, 3H), 7.26 (d, *J* = 8.4 Hz, 1H), 6.74 (d, *J* = 4 Hz, 1H); ¹³CNMR (DMSO-d₆): 165.3, 164.3, 148, 144.2, 141.2, 140.4, 137, 135.8, 134.5, 131.7, 130.1, 129.4, 128.7, 127.7, 126.7, 126.2, 125.2, 122.9, 122, 119.1, 110.3; ESI-HRMS: *m/z* calculated for C₂₅H₂₂N₆S: 219.0808, found: 219.0802 (Amidine base M⁺ + 2, double charged). Anal. Calcd. For C₂₅H₂₀N₆S. 3HCl. 1.25H₂O: C, 52.94; H, 4.53; N, 14.82. Found: C, 52.78; H, 4.80; N, 14.52.

2-(5-(4-Carbamimidoylphenyl)selenophen-2-yl)-1-methyl-1*H*-benzo[d]imidazole-5-carboximidamide (6l).

Yellow solid (0.25 gm, 47 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 9.55 (s, 2 H), 9.45 (s, 2 H), 9.32 (s, 2 H), 9.24 (s, 2 H), 8.25 (s, 1 H), 8.16 (br s, 1 H), 8.11 (br s, 1 H), 7.96 (m, 5H), 7.82 (d, *J* = 8 Hz, 1H), 3.51 (s, 3H); ESI-HRMS: *m/z* calculated for C₂₀H₂₀N₆Se: 212.0452, found: 212.0442 (Amidine base M⁺ + 2, double charged). Anal. Calcd. For C₂₀H₁₈N₆Se. 3HCl. 0.35H₂O: C, 44.72; H, 4.07; N, 15.64. Found: C, 45.07; H, 4.21; N, 15.25.

2-(2-(4-Carbamimidoylphenyl)thiazol-5-yl)-1*H*-benzo[d]imidazole-6-carboximidamide (6m).

Yellow solid (0.16 gm, 32 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 9.60 (s, 2 H), 9.44 (s, 2 H), 9.36 (s, 2 H), 9.17 (s, 2 H), 8.95 (br s, 1H), 8.26 (d, *J* = 7.6 Hz, 2H), 8.19 (br s, 1H), 8.01 (d, *J* = 7.6 Hz, 2H), 7.80 (d, *J* = 8.2 Hz, 1H), 7.72 (d, *J* = 8.2 Hz, 1H); ¹³CNMR (DMSO-d₆): 166.5, 165.5, 164.4, 146.5, 144, 143.8, 136.2, 129.6, 129, 128.7, 128, 126, 125.7, 122, 121.1, 114.4; ESI-HRMS: *m/z* calculated for C₁₈H₁₆N₇S: 362.1182, found: 362.1179 (Amidine base M⁺ + 1). Anal. Calcd. For C₁₈H₁₅N₇S. 3HCl. 2H₂O. 0.1C₃H₆O: C, 42.87; H, 4.44; N, 19.12. Found: C, 43.24; H, 4.33; N, 18.81.

2-(2-(4-Carbamimidoylphenyl)thiazol-5-yl)-1-methyl-1H-benzo[d]imidazole-5-carboximidamide (6n).

Yellow solid (0.19 gm, 37 %), mp > 300 °C. ¹HNMR (DMSO-d₆): δ 9.66 (s, 2 H), 9.50 (s, 2 H), 9.45 (s, 2 H), 9.29 (s, 2 H), 8.74 (br s, 1 H), 8.29 (br s, 1 H), 8.25 (d, *J* = 7.8 Hz, 2H), 8.04 (d, *J* = 7.8 Hz, 2H), 7.95 (d, *J* = 8.4 Hz, 1H), 7.84 (d, *J* = 8.4 Hz, 1H), 4.16 (s, 3H); ¹³CNMR (DMSO-d₆): 166.9, 165.3, 164.4, 146.7, 144, 141, 139.5, 136, 129.2, 128.8, 128.6, 126.1, 122.2, 121.3, 119, 111, 31.8; ESI-HRMS: *m/z* calculated for C₁₉H₁₉N₇S: 188.5706, found: 188.5695 (Amidine base M⁺ + 2, double charged). Anal. Calcd. For C₁₉H₁₇N₇S. 3HCl. 0.75H₂O. 0.25C₃H₆O: C, 46.37; H, 4.53; N, 19.17. Found: C, 46.46; H, 4.68; N, 19.27.

Acknowledgment

We thank the National Institutes of Health Grant GM111749 (W.D.W. and D.W.B.) for financial support. We also thank the Leukemia & Lymphoma Society (LLS) award: Leukemia & Lymphoma Society 6504-17 for support. We thank Dr. Yanyi Chen for his help with our CD experiments.

References.

1. Bhaduri, S; Ranjan, N; Arya, D. P. An overview of recent advances in duplex DNA recognition by small molecules. *Beilestein J. Org. Chem.* 2018, 1051-1086.
2. Cossar, P.J.; Lewis, P.J.; McCluskey, A. Protein–protein interactions as antibiotic targets: A medicinal chemistry perspective. *Med. Res. Rev.* 2018. 1-26.
3. Freilich, R.; Arhar, T.; Abrams, J.L.; Gestwicki, J.E. Protein–protein interactions in the molecular chaperone network. *Acc. Chem. Res.* 2018, 51, 940–949.
4. Nero, T.L.; Parker, M.W.; Morton, C.J. Protein structure and computational drug discovery. *Biochem. Soc. Trans.* 2018, 46, 1367–1379.

5. Sánchez, M. I; Vázquez, O.; Martínez-Costas, J.; Vázquez M. E.; Mascareñas, J. L. Straightforward access to bisbenzamidine DNA binders and their use as versatile adaptors for DNA-promoted processes. *Chem. Sci.* 2012, 2383-2387.
6. 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.
7. Giordani, F.; Munde, M.; Wilson, W.D.; Ismail, M.A.; Kumar, A.; Boykin, D.W.; Barrett, M.P. Green fluorescent diamidines as diagnostic probes for trypanosomes. *Antimicrob. Agents Chemother.* 2014, 58, 1793–1796.
8. Wilson, W.D.; Nguyen, B.; Tanious, 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. Anti-Cancer Agents* 2005, 5, 389–408.
9. Lambert, M.; Jambon, S.; Depauw, S.; David-Cordonnier, M.-H. Targeting transcription factors for cancer treatment. *Molecules* 2018, 23, 1479-1529.
10. Koehler, A. N. A complex task? Direct modulation of transcription factors with small molecules. *Curr. Opin. Chem. Biol.* 2010, 14, 331-340.
11. Antony-Debré, I.; Paul, A.; Leite, J.; Mitchell, K.; Kim, H. M.; Carvajal, L. A.; Todorova, T. I.; Huang, K.; Kumar, A.; Farahat, A. A.; Bartholdy, B.; Narayanagari, S.-R.; Chen, J.; Ambesi-Impiombato, A.; Ferrando, A. A.; Mantzaris, I.; Gavathiotis, E.; Verma, A.; Will, B.; Boykin, D. W.; Wilson, W. D.; Poon, G. M. K.; Steidl, U. Pharmacological inhibition of the transcription factor PU.1 in leukemia. *J. Clin. Invest.* **2017**, 127, 4297-4313.
12. Munde, M.; Wang, S.; Kumar, A.; Stephens, C. E.; Farahat, A. A.; Boykin, D. W.; Wilson, W. D.; Poon, G. M. K. Structure-dependent inhibition of the ETS-

family transcription factor PU.1 by novel heterocyclic diamidines. *Nucleic Acids Res.* **2014**, 42, 1379-1390.

13. Kopka, M.L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R.E. Binding of an antitumor drug to DNA: Netropsin and C-G-C-G-A-A-T-T-BrC-G-C-G. *J. Mol. Biol.* **1985**, 183, 553–563.

14. Lown, J.W.; Krowicki, K.; Bhat, U.G.; Skorobogaty, A.; Ward, B.; Dabrowiak, J.C. Molecular recognition between oligopeptides and nucleic acids: Novel imidazole-containing oligopeptides related to netropsin that exhibits altered DNA sequence specificity. *Biochemistry.* **1986**, 25, 7408–7416.

15. Kielkopf, C.L.; Baird, E.E.; Dervan, P.B.; Rees, D.C. Structural basis for G.C recognition in the DNA minor groove. *Nat. Struct. Biol.* **1998**, 5, 104–109.

16. Nozeret, K.; Loll, F.; Cardoso, G.M.; Escude, C.; Boutorine, A.S. Interaction of fluorescently labeled pyrrole-imidazole polyamide probes with fixed and living murine and human cells. *Biochimie.* **2018**, 149, 122–134.

17. Hargrove, A.E.; Raskatov, J.A.; Meier, J.L.; Montgomery, D.C.; Dervan, P.B. Characterization and solubilization of Pyrrole–Imidazole polyamide aggregates. *J. Med. Chem.* **2012**, 55, 5425–5432. 18. Nishijima, S.; Shinohara, K.I.; Bando, T.; Minoshima, M.; Kashiwazaki, G.; Sugiyama, H. Cell permeability of Py-Im-polyamide-fluorescein conjugates: Influence of molecular size and Py/Im content. *Bioorg. Med. Chem.* **2010**, 18, 978–983.

19. 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.

20. 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.

21. 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. *Chem. Eur. J.* **2016**, 22, 15404–15412.).
22. 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.
23. M. Munde, M. Lee, S. Neidle, R. Arafa, D. W. Boykin, Y. Liu, C. Bailly, W. D. Wilson. 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.
24. S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan,. Recognition of the four Watson–Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*. **1998**, 391, 468 –471.
25. R. Rohs, S. M. West, A. Sosinsky, P. Liu, R. S. Mann, B. Honig, The role of DNA shape in protein–DNA recognition. *Nature*. **2009**, 461, 1248 – 1253.
26. S. Mallena, M. P. H. Lee, C. Bailly, S. Neidle, A. Kumar, D. W. Boykin, W. D. Wilson, Thiophene-based diamidine forms a “super” at binding minor groove agent. *J. Am. Chem. Soc.* **2004**, 126, 13659-13669.
27. Vázquez M.; Vázquez, E.; Blanco, J. B.; Castedo, L.; Mascareñas, J. L. Specific DNA Recognition by a Synthetic, Monomeric Cys2His2 Zinc-Finger Peptide Conjugated to a Minor-Groove Binder *Angew. Chem. Int. Ed.* **2007**, 46, 6886 – 6890.
28. Rodríguez, J.; Mosquera, J.; Couceiro, J. R.; Vázquez, M. E.; Mascareñas, J. L. The AT-Hook motif as a versatile minor groove anchor for promoting DNA binding of transcription factor fragments. *Chem. Sci.* **2015**, 6, 4767 –4771.

29. Hahn, L.; Buurma, N. J.; Gade, L. H. A Water-Soluble Tetraazaperopyrene Dye as Strong G-Quadruplex DNA Binder. *Chem. Eur. J.* **2016**, 22, 6314–6322.)
30. Wilson, W. D.; Tanious, F. A.; Mathis, A.; Tevis, D.; Hall, J. E.; Boykin, D. W. Antiparasitic compounds that target DNA. *Biochimie.* **2008**, 999–1014.
31. Nanjunda, R.; Wilson, W. D. Current Protocols in Nucleic Acid Chemistry, **2012**, Chapter 8, Unit 8.8.
32. 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** 41, W56-W62.
33. Beno, B. R.; Yeung, K.-S.; Bartberger, M. D.; Pennington, L. D.; Meanwell, N. A. A survey of the role of noncovalent sulfur interactions in drug design. *J. Med. Chem.* **2015**, 58, 4383-4438.
34. Steinmann, D.; Nauser, T.; Koppenol, W. H. Selenium and sulfur in exchange reactions: a comparative study. *J. Org. Chem.* **2010**, 75, 6696.
35. Reich, H. J.; Hondal, R. J. Why nature chose selenium. *ACS Chem. Biol.* **2016**, 11, 821.
36. Farahat, A. A.; Kumar, A.; Paliakova, E.; Barghash, A. M.; Goda, F. E. Eisa, H. M.; Wenzler T.; Brun, R.; Liu, Y.; Wilson, W. D.; Boykin, D. W. Exploration of Larger Central Ring Linkers in Furamidine Analogues: Synthesis and Evaluation of their DNA Binding, Antiparasitic and Fluorescence Properties. *Bioorg. Med. Chem.* **2011**, 19, 2156-2167.
37. Nanjunda, R.; Munde, M.; Liu, Y.; Wilson, W. D. Methods for Studying Nucleic Acid/Drug Interactions (Eds.: Y. Wanunu), CRC Press, **2011**, pp. 91–119.

38. Y. Liu, Y. Chai, A. Kumar, R. R. Tidwell, D. W. Boykin, W. D. Wilson. Designed Compounds for Recognition of 10 Base Pairs of DNA with Two AT Binding Sites. *J. Am. Chem. Soc.* **2012**, 134, 5290–5299.
39. M. Eriksson, B. Nordén, Drug-Nucleic Acid Interactions, Vol. 340, Elsevier, Amsterdam, **2001**, pp. 68–98.
40. Fornander, L. H.; Wu, L.; Billeter, M.; Lincoln, P.; Nordén, B. Minor-groove binding drugs: where is the second Hoechst 33258 molecule?. *J. Phys. Chem. B* **2013**, 117, 5820–5830.
41. Kubista, M.; Akerman, B.; Norden, B. Induced Circular Dichroism in Nonintercalative DNA–Drug Complexes: Sector Rules for Structural Applications. *J. Phys. Chem.* **1988**, 92, 2352–2356.
42. 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. *Chem. Eur. J.* **2015**, 21, 5528.
43. Liu, Y.; Kumar, A.; Depauw, S.; Nhili, R.; David-Cordonnier, M.-H.; Lee, M. P.; Ismail, M. A.; Farahat, A. A.; Say, M.; Chackal-Catoen, S.; Batista-Parra, A.; Neidle, S.; Boykin, D. W.; Wilson, W. D. Water-mediated binding of agents that target the DNA minor groove. *J. Am. Chem. Soc.* **2011**, 133, 10171.
44. Munde, M.; Kumar, A.; Nhili, R.; Depauw, S.; David-Cordonnier, M.-H.; Ismail, M. A.; Stephens, C. E.; Farahat, A. A.; Batista-Parra, A.; Boykin, D. W.; Wilson, W. D. DNA minor groove induced dimerization of heterocyclic cations: compound structure, binding affinity, and specificity for a TTAA site. *J. Mol. Biol.* **2010**, 402, 847-864.
45. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; GE Scuseria- Inc.; Wallingford; CT; **2010**. Gaussian 09, revision b. 01, Gaussian.

46. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, 97, 10269-10280.
47. Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, 129-145.
48. Case, D. A.; Babin, V.; Berryman, J.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. AMBER 14, **2014**.
49. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, 25, 247-260.
50. Harika, N. K.; Germann, M. W.; Wilson, W. D. First Structure of a Designed Minor Groove Binding Heterocyclic Cation that Specifically Recognizes Mixed DNA Base Pair Sequences. *Chem. Eur. J.* **2017**, 23, 17612-17620.
51. 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.
52. Špačková, N.; Cheatham, T. E.; Ryjáček, F.; Lankaš, F.; van Meervelt, L.; Hobza, P.; Šponer, J. Molecular Dynamics Simulations and Thermodynamics Analysis of DNA–Drug Complexes. Minor Groove Binding between 4',6-Diamidino-2-phenylindole and DNA Duplexes in Solution. *J. Am. Chem. Soc.* **2003**, 125, 1759-1769.

53. Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **2010**, 65, 455-461.
54. Farahat, A. A.; Bennett-Vaughn, C.; Mineva, E. M.; Kumar, A.; Wenzler, T.; Brun, R.; Liu, Y.; Wilson, W. D.; Boykin, D. W. Synthesis, DNA binding and antitrypanosomal activity of benzimidazole analogues of DAPI. *Bioorg. Med. Chem. Lett.* **2016**, 26, 5907-5910.
55. Hakan Göker, Canan Kuş, David W. Boykin, Sulhiye Yildiz, Nurten Altanlar. Synthesis of some new 2-substituted-phenyl-1H-benzimidazole-5-carbonitriles and their potent activity against candida species. *Bioorg. Med. Chem.*, **2002**, 10, 2589-2596.
56. Farahat, A.A.; Ismail, M.A.; Kumar, A.; Wenzler, T.; Brun, R.; Paul, A.; Wilson, W.D.; Boykin, D.W. Indole and benzimidazole bichalcophenes: Synthesis, DNA binding and antiparasitic activity. *Eur. J. Med. Chem.* **2018**, 143, 1590–1596.
57. Depauw, S.; Lambert, M.; Jambon, S.; Paul, A.; Peixoto, P.; Nhili, R.; Marongiu, L.; Figeac, M.; Dassi, C.; Paul-Constant, C.; et al. Heterocyclic Diamidine DNA ligands as HOXA9 Transcription Factor Inhibitors: Design, Molecular Evaluation and Cellular Consequences in HOXA9-Dependant Leukemia Cell Model. *J. Med. Chem.* **2019**, 62, 1306–1329.
58. Farahat, A. A.; Kumar, A.; Say, M.; Wenzler, T.; Brun, R.; Paul, A.; Wilson, W. D.; Boykin, D. W. Exploration of DAPI analogues: Synthesis, antitrypanosomal activity, DNA binding and fluorescence properties. *Eur. J. Med. Chem.* **2017**, 128, 70–78.
59. Farahat, A. A.; Kumar, A.; Say, M.; Barghash, A. E.-D. M.; Goda, F. E.; Eisa, H. M.; Wenzler, T.; Brun, R.; Liu, Y.; Mickelson, L.; Wilson, W. D.; Boykin, D. W. Synthesis, DNA binding, fluorescence measurements and antiparasitic activity of DAPI related diamidines *Bioorg. Med. Chem.* **2010**, 18, 557-566.

60- Branowska, D.; Farahat, A. A.; Kumar, A.; Wenzler, T.; Brun, R.; Liu, Y.; Wilson, W. D.; Boykin, D. W. Synthesis and antiprotozoal activity of 2,5-bis[amidinoaryl]thiazoles. *Bioorg. Med. Chem.* **2010**, 18, 3551-3558.

Graphical abstract

