

Review

A New Generation of Minor-Groove-Binding—Heterocyclic Diamidines That Recognize G·C Base Pairs in an AT Sequence Context

Ananya Paul , Pu Guo , David W. Boykin and W. David Wilson *

Department of Chemistry and Center for Diagnostics and Therapeutics, Georgia State University, 50 Decatur St SE, Atlanta, GA 30303, USA; apaul@gsu.edu (A.P.); pguo@gsu.edu (P.G.); dboykin@gsu.edu (D.W.B.)

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

Received: 24 January 2019; Accepted: 1 March 2019; Published: 7 March 2019



Abstract: We review the preparation of new compounds with good solution and cell uptake properties that can selectively recognize mixed A·T and G·C bp sequences of DNA. Our underlying aim is to show that these new compounds provide important new biotechnology reagents as well as a new class of therapeutic candidates with better properties and development potential than other currently available agents. In this review, entirely different ways to recognize mixed sequences of DNA by modifying AT selective heterocyclic cations are described. To selectively recognize a G·C base pair an H-bond acceptor must be incorporated with AT recognizing groups as with netropsin. We have used pyridine, azabenzimidazole and thiophene-*N*-methylbenzimidazole GC recognition units in modules crafted with both rational design and empirical optimization. These modules can selectively and strongly recognize a single G·C base pair in an AT sequence context. In some cases, a relatively simple change in substituents can convert a heterocyclic module from AT to GC recognition selectivity. Synthesis and DNA interaction results for initial example lead modules are described for single G·C base pair recognition compounds. The review concludes with a description of the initial efforts to prepare larger compounds to recognize sequences of DNA with more than one G·C base pairs. The challenges and initial successes are described along with future directions.

Keywords: DNA G·C base pair recognition; mixed base pair DNA sequences; sequence selectivity; heterocyclic diamidine; σ -hole; thiophene-*N*-methylbenzimidazole; benzimidazole; aza-benzimidazole; biosensor

1. Introduction

Overview of DNA and RNA Targeting Challenges

The interaction of designed, synthetic, small organic molecules with macromolecules is a fascinating and important current research area for both biophysical studies and for the design of new reagents for selective binding applications as well as for therapeutic development. Compounds that bind to proteins have become important molecular probes but also account for many of our most active current drugs [1–3]. The design of agents that selectively target sequences or structures of DNA or RNA has progressed more slowly due to their more repetitive structures and properties [4–8]. The nucleic acids have only four basic components and a high negative charge compared to the complexity of components and charged groups in proteins. The generally repetitive duplex structure of DNA as well as duplex sections of RNA makes selective targeting more of a challenge. Compounds which bind with

some selectively to nucleic acids have, however, produced important cellular probes, chromosomal dyes, and a number of important therapeutic agents [9–14]. Among the earliest discovered agents that interact with nucleic acids are compounds that selectively bind to A·T base pair (bp) sequences in the DNA minor groove [15–20]. These agents generally have some flexibility in AT sequence binding but strongly exclude G·C bps. In duplex regions G·C bps have a third hydrogen bond, compared to A·T bps and the extra group protrudes into the minor groove. The extra group sterically hinders most of the original minor groove agents from sliding deeply into the groove to interact with the edges of bps in the groove [15–21]. The G·C bps are also more electropositive than A·T bps and this also hinders binding of cationic minor groove agents [22,23] such as those shown in Figure 1, for example. Compounds that bind in the nucleic acid grooves have been especially attractive for therapeutic development since they are generally less mutagenic and less toxic than intercalators [24,25].

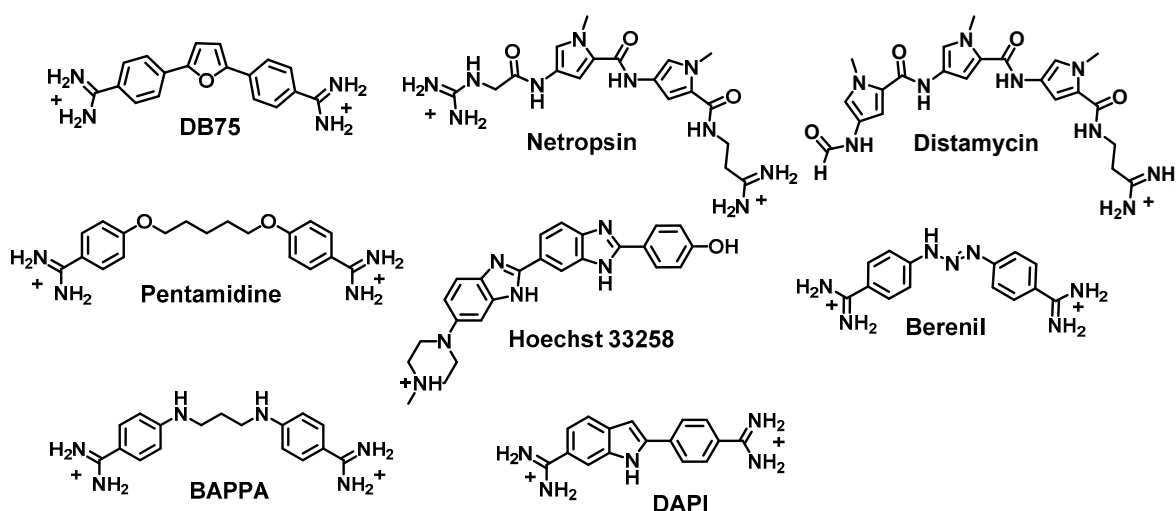


Figure 1. Structures of classical AT specific DNA minor groove binders.

As our knowledge of RNA structures and functional complexity has expanded to RNAs with important cellular functions such as control of gene expression and translation, the number of research groups involved with design and testing of agents for selectively binding to specific RNA base pair groups and structures has rapidly increased [26–33]. Because of the structural complexity of RNA, relative to DNA, designing agents to bind to RNA is quite a different and difficult challenge compared with DNA or protein targeting. To date there are few designed agents that selectively target RNA but it is an area of great potential.

2. AT Specific Minor Groove Drugs: Antiparasitic Agents

AT-specific heterocyclic cations such as Hoechst 33258 and DAPI have been successful in staining components in cells, particularly chromatin. Other AT specific cations such as the diamidines furamide, pentamidine, berenil and related diamidines (Figure 1) have had significant success as antiparasitic agents and pentamidine has been used in humans for over 50 years [25,34–39]. Interestingly, however, no systematic effort to redesign these agents into mixed base pairs recognition compounds has been described previously. Mascareñas and coworkers have made several novel and important designed bis-benzamide compounds such as BAPPA (Figure 1) [40] which can recognize pure AT bps with very high affinity. BAPPA and other pentamidine analogues, where the pentamidine oxygens are replaced by NH, have much improved spectroscopic and solubility properties [12,41]. One synthetic bis-benzamide has a central pyridine group that gave enhanced GC bp recognition in an AT sequence context [42]. New ideas to redesign successful A·T bp specific diamidines, so that they have significantly broadened sequence recognition capability and potentially expanded uses, are reported in this review.

Starting in the mid-1980s, design of minor groove agents to recognize GC bps took a step forward with ideas from the Dickerson and Lown laboratories for modifications of the minor groove binding polyamides netropsin (Nt) and distamycin (Dst), which specifically recognizing A·T bp. The idea was to make new synthetic compounds that could also bind to a G·C bp [43,44]. They proposed replacing one or more pyrrole groups in Nt and Dst with an imidazole group so that the extra N in imidazole could form an H-bond with the G-NH₂ in the minor groove of a G·C bp [43–45]. The idea attracted attention in several additional laboratories and significant success in GC recognition was achieved by the Dervan laboratory by connecting two recognition units to make a hairpin polyamide [46–48]. The synthetic polyamides are very attractive reagents, but solution difficulties, aggregation, and poor cell uptake have limited their applications [49–51]. Several collaborating laboratories at the University of Strathclyde have designed, synthesized and tested polyamides with successfully modified heterocycles to improve their solution properties, DNA recognition and therapeutic potential [52].

Our interest in nucleic acid binding compounds began with a desire to design new types of agents to treat parasitic diseases [36,53]. For most parasitic diseases there are limited small molecule therapeutic options, and there are few or no satisfactory vaccines. Unfortunately, drug design for many parasitic diseases has received little recent attention and funding, since these are largely diseases of economically disadvantaged countries. Trypanosome parasites, for example, cause diseases like sleeping sickness and Chagas diseases, particularly in Africa, Central, and South America [54–57]. There are very few effective drugs to treat these deadly diseases. Many of the available compounds that are used have been available for many years and have significant toxicity [58].

As described above, most classical minor groove binding agents have AT interaction specificity. The malaria genome, for example, is quite AT-rich and many AT specific DNA binding compounds have excellent activity against malaria [59–61]. The mitochondrial genome of trypanosomes that cause human African sleeping sickness is also very AT-rich and presents an excellent AT selective DNA binding site array [62]. The fact that pentamidine (Figure 1), an AT specific DNA minor groove binder, is an active agent for sleeping sickness has provided a valuable lead into this system [63]. Our collaborative synthetic and biophysical groups have prepared and tested a wide variety of heterocyclic amidine derivatives, and we have found a number of compounds that are highly active against sleeping sickness [25,36,53,64,65]. Development of these compounds produced a prodrug of furamidine (DB75 in Figure 1) that has moved to Phase III human clinical trials [25,66]. Because these heterocyclic diamidine compounds are intrinsically fluorescent, they can be directly observed in cells [64]. Fluorescence microscopy showed that the compounds selectively target the AT-rich mitochondrial kinetoplast DNA [64,67,68]. After addition of the compounds to a trypanosome infected mouse, the kinetoplast became highly fluorescent, then disintegrated followed by disappearances and death of the trypanosomes. These initial, relatively simple compounds provided a proof of concept that diamidines could be used effectively in humans.

Over a century ago Paul Ehrlich raised the question “— do the trypanosomes possess in their cells definite groupings which govern the captivation of definite chemical substances?” Cellular fluorescence microscopy of trypanosomes treated with furamidine and related compounds provides a definite “Yes” answer to his question.

3. Design of Compounds that Expand Past Pure AT Sequence

We are also continuing with the design of new compounds that will be more specific in targeting kinetoplast sequences with greater affinity, selectivity, and potentially reduced toxicity. The kinetoplast genome contains thousands of small circular AT-rich DNAs that code for guide RNAs that control mRNA editing in trypanosomes. Since the small DNAs are AT-rich, they are ideal targets for heterocyclic diamidines [36,64,68]. The AT sequences in the kinetoplast typically have many single G·C bps flanked on both sides by AT sequences. We reasoned that compounds that could target the single G·C bp, as well as the flanking AT sequences, would have much greater selectivity in binding to the kinetoplast with reduced toxicity.

Agents that can recognize mixed base pair sequences of DNA containing combinations of A·T and G·C bps are an important development area for minor groove binders. As described above, the development took a major step forward with modification of the minor groove binders, netropsin (Nt) and distamycin (Dst) [43–45]. The proposal was to replace one or more pyrrole groups in Nt. and Dst. with *N*-methylimidazole groups [43–45]. As has been described, 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. Success in mixed sequence binding with both A·T and G·C bp recognition with polyamides required the finding by Wemmer and coworkers that it took two stacked or linked recognition units to make an effective minor groove polyamide for GC targeting [69,70]. Although several groups have worked with synthetic polyamides, solution difficulties have limited their applications [49–51]. Polyamide work on G·C bp targeting has continued to focus on replacement of pyrrole by *N*-methylimidazole or other groups that would give an H-bond acceptor that could bind to the G-NH₂ group that projects into the minor groove [71–74]. Research in several laboratories has expanded past *N*-methylimidazole to new GC recognition agents with the goal of preparing polyamides with broader therapeutic potential [75–77].

Our goal was to prepare entirely new types of non-amide compounds 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 effectively target a variety of different cell types (Figure 2). The initial research in the area has produced quite different compounds that are also quite different from the polyamides and are without pyrrole and imidazole groups [78–80]. Our three most successful compounds in targeting a single G·C bp in an AT sequence are shown in Figure 2 [78–80]. All three of these compounds are undergoing additional development to increase affinity and selectivity for the target sequence type with a single G·C flanked by A·T bps. Brief synthetic schemes for the compounds of Figure 2 that recognize a single G·C bp in an AT sequence are given below.

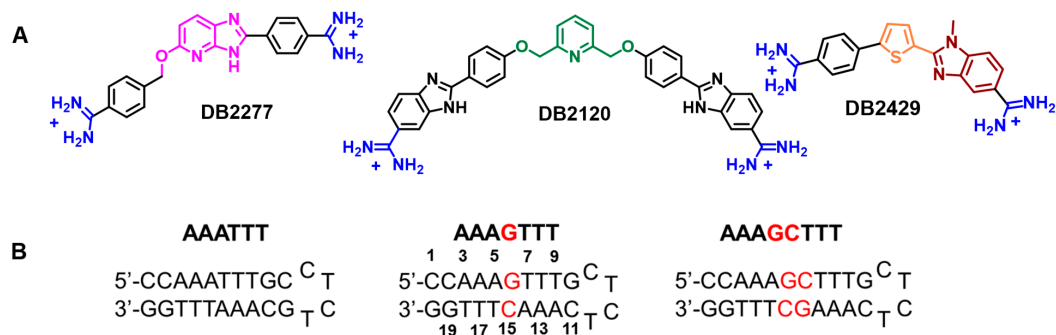
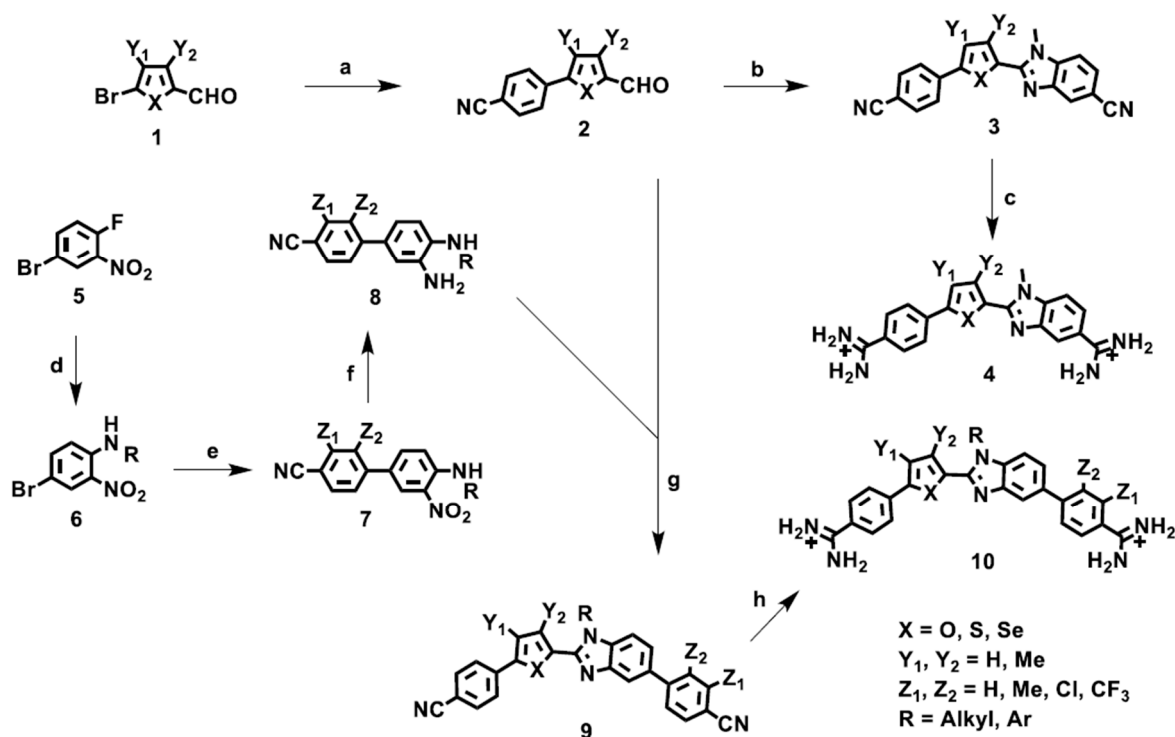


Figure 2. (A) Structures of three very different designed mixed-sequence DNA minor groove binders; (B) DNA hairpin oligomer duplexes used in their study. DNA sequences with 5'-biotin labels were used for SPR studies.

3.1. Synthetic Procedures for DB2429, DB2457 and Analogues

The syntheses of the *N*-substituted benzimidazole heterocycles are outlined in Scheme 1 [80,81]. In this series the H-bond acceptor unit is the σ -hole array 1,4-N-S or 1,4-N-Se provided by *N*-substituted benzimidazoles and either a thiophene or a selenophene. Reaction of the readily available bromoheteroaryl aldehydes **1**, when X = S or Se provides a key component for σ -hole formation, with 4-cyanophenylboronic acid under standard Suzuki coupling conditions conveniently provides the 5-(4-cyanophenyl)-2-formyl 5-ring heterocycles **2**, which can be used to prepare both the tri- and tetra-heteroaryl target molecules **4** and **10**. Condensation, cyclization, and oxidation, mediated by sodium metabisulfite, of the aldehydes **2** with 3-amino-4-(alkyl or aryl amino) benzonitrile, which provides the second component needed for σ -hole formation, gives the bis-nitriles **3**. These can be readily converted into the desired diamidines **4** using LiN(TMS)₂ in THF. In a similar sequence,

the aldehydes **2** are allowed to react with the arylphenylenediamines **8**, again mediated by sodium metabisulfite, to produce the bis-nitriles **9**.



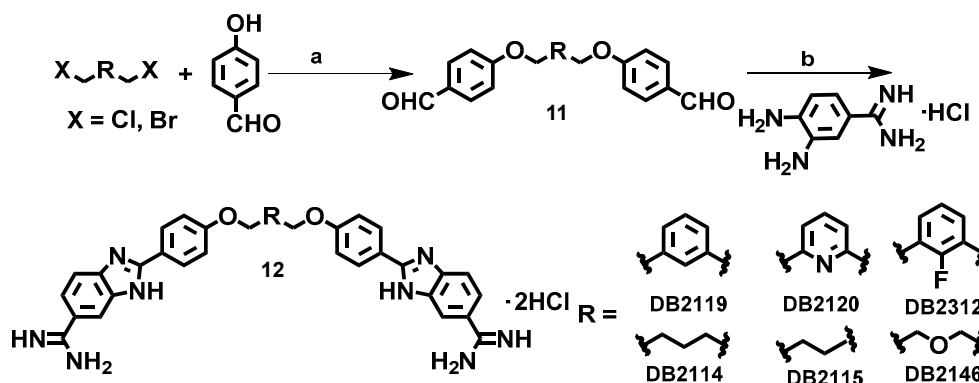
Scheme 1. Reagents and conditions: (a) 4-Cyanophenylboronic acid, Pd(PPh₃)₄, Na₂CO₃/H₂O, dioxane, reflux. (b) 3-Amino-4-alkyl or aryl aminobenzonitrile, sodium metabisulfite, DMF, reflux. (c) i. LiN(TMS)₂/THF, rt ii. HCl/ethanol. (d) RNH₂/EtOH or Cs₂CO₃/DMA. (e) 4-cyanophenylboronic acid, Pd(PPh₃)₄, Na₂CO₃/H₂O, dioxane, reflux. (f) SnCl₂/EtOH, reflux. (g) Na₂S₂O₅/DMSO, 130 °C (h) i. LiN(TMS)₂/THF, ii-HCl/EtOH.

These are then converted to the diamidines **10** as previously described. In order to obtain the arylphenylenediamines **8**, nucleophilic aromatic substitution reactions between 4-bromo-1-fluoro-2-nitrobenzene (**5**) and various aliphatic and aromatic amines are employed which yields the 4-bromo-*N*-alkyl and *N*-aryl-2-nitrobenzenes **6**. The nucleophilic aromatic substitution reaction with aliphatic amines was achieved at room temperature in ethanol, whereas, the less basic aryl amines require heating in dimethylacetamide in the presence of Cs₂CO₃ at 160 °C. Suzuki–Miyaura coupling of substituted 4-cyanophenylboronic acids with the various 4-bromo-*N*-alkyl and *N*-aryl-2-nitrobenzenes under standard conditions provides the nitro biphenyl analogs **7**. The arylphenylenediamines **8** required for formation of the *N*-substituted benzimidazoles are obtained by stannous chloride reduction of the nitro groups of **7** and are subsequently allowed to react directly without characterization. Including the tri- and tetra-heteroaryl target molecules available for single G·C bp binding studies, we have designed and synthesized almost 70 target compounds containing σ -hole modules.

3.2. Synthetic Procedures for DB2120 and Its Analogues

In the Scheme 2 examples a library of new molecules is prepared by including the H-bond accepting group in the linker connecting two terminal benzimidazole-diamidine units **12** in Scheme 2 [78]. The latter are well studied in the field of AT sequence-specific DNA minor groove binders which have been shown to exhibit excellent nuclear uptake and transcription factor inhibitory activity [82]. In this series we have altered the linker to provide H-bond accepting units capable of G·C bp recognition but to retain some conformational flexibility. The preparation of the symmetrical

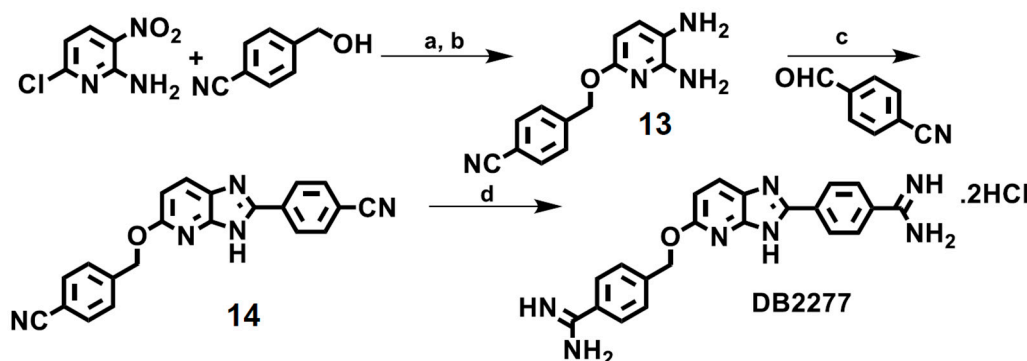
dialdehydes **11** needed to construct the terminal benzimidazole-diamidine units, employs coupling *p*-hydroxybenzaldehyde with the various di-halo substituted linkers in the presence of K_2CO_3/Cs_2CO_3 in dimethylformamide. The compounds **11** are allowed to react with 4-amidino-1,2-phenylene-diamine hydrochloride hydrate in anhydrous ethanol, mediated by 1,4-benzoquinone, to yield DB2120 and various analogues.



Scheme 2. Reagents and conditions: (a) K_2CO_3 , anhydrous DMF, 70 °C. (b) (i) benzoquinone, ethanol, 80 °C. (ii) ethanol-HCl.

3.3. Synthetic Procedure for DB2277

In this series, an azabenzimidazole is employed as the H-bond accepting group, an essential unit for recognition of a G-C bp [79]. The synthesis of the most selective GC binding compound of this type, DB2277, an azabenzimidazole with a flexible -O-CH₂ linker, begins with coupling of commercially available 2-amino-6-chloro-3-nitropyridine and 4-hydroxymethylbenzonitrile in the presence of sodium hydride in dimethylformamide (Scheme 3). The intermediate nitro compound is reduced by iron dust in a 2-propanol/water mixture at reflux which provides the diamine compound **13**. In the following step, 4-cyanobenzaldehyde is allowed to react with diamine **13** in the presence of sodium bisulfite in DMF at reflux to provide the *bis*-nitrile substituted azabenzimidazole intermediate **14**. The desired hydrochloride salt of the diamidine DB2277 is obtained by stirring the intermediate **14** in ethanolic HCl, applying Pinner methodology where the bis-imidate ester hydrochloride is ultimately converted to the hydrochloride salt of the diamidine in the presence of ethanol saturated with ammonia gas [71].



Scheme 3. Reagents and conditions: (a) NaH, DMF, rt. (b) iron dust, ammonium chloride, isopropanol-water, reflux. (c) sodium bisulfite, DMF, 140 °C. (d) (i) HCl gas/EtOH, rt. (ii) NH_3 gas / EtOH, rt.

4. Increasing the Design Diversity of Compounds to Specifically Recognize the Minor Groove

Until recently recognizing mixed base pair sequence in the minor groove of DNA has largely continued to follow the original ideas with pyrrole and imidazole and related groups with amide

linkers in polyamides. Such compounds have an advantage of design simplicity, but have been plagued by solution difficulties and poor cell uptake [49–51,83]. The compounds in Figure 2 have good solution properties, and cell uptake and they were designed with a structural and chemical diversity that gives them an advantage in targeting different cell types [82]. The three different GC recognition groups, pyridine [78,84], azabenzimidazole [79,85,86] thiophene-*N*-methylbenzimidazole (*N*-MeBI) [80,81], in specific minor groove recognition modules offer a much-needed breakthrough in mixed sequence binding with a range of solution properties. Each of these compound types represents the results of extensive development work to enhance sequence selectivity and affinity. The pyridine compounds were found to require the adjacent $-\text{CH}_2\text{-O}-$ linker to give acceptable affinity and selectivity [78]. In a similar manner, the azabenzimidazole (DB2277) was found to require the flexible linker $-\text{O-CH}_2-$ to function well [86]. Compounds without that group bind with much lower selectivity to target AT sequences with a single G·C bp. The same extensive development yielded the initial thiophene-*N*-MeBI compound in that series (Figures 2 and 3, DB2429). We cannot present further development of all three compounds in the space available and will, therefore, focus on the thiophene-*N*-MeBI module in the remainder of this review as an example for all three compounds.

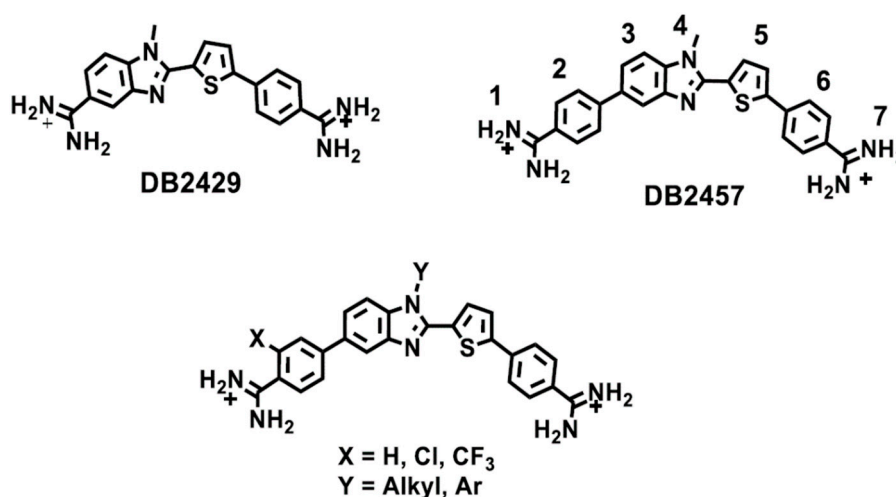


Figure 3. Structure of reported σ -hole single G·C bp binders and the modified compounds from DB2457.

4.1. Initial Concept and Development

The initial concept for development of the thiophene-*N*-MeBI module was that the thiophene S and *N*-alkyl benzimidazole unprotonated *N* could form a σ -hole based interaction that would preorganize the module appropriately for recognition of the shape of the minor groove [80]. This shape also points the unprotonated *N* of *N*-alkylBI into the minor groove to H-bond with the G-NH₂ group in that groove. The σ -hole bonding is a highly directional noncovalent interaction between the covalently bonded S atom in thiophene and the lone pair on the nucleophilic *N* in BI. When one of the half-filled p orbitals in the thiophene S atom is involved in a covalent bond, results in electron deficiency in the outer (noninvolved) lobe of the orbital that is a σ -hole. If the remainder of the molecule is sufficiently electron-withdrawing relative to the polarizable S atom, then there is a positive electrostatic potential associated with the σ -hole that is responsible for the outer-lobe electron deficiency. This positive potential can interact attractively with the negative lone pair on the BI-N; The resulting noncovalent interaction, directed roughly along the extension of the C-S covalent bond in thiophene, is a σ -hole bond. The S in thiophene can have two positive σ -holes, approximately along the extensions of the C-S bonds but there is only one interaction in DB2429 with the one *N*-MeBI. Thiophene C-S single bonds in general provide a relatively positive electrostatic potential that can form an interaction with electron donating atoms such as the unsubstituted *N* in *N*-MeBI, a 1,4 N \cdots S interaction (Figures 2 and 3) [87].

able to successfully synthesize compounds with a variety of substituents at the *N*-BI position and other compounds with substituent modifications on the aromatic groups of the thiophene-*N*-substitution BI module [81]. Representative examples with the modified compounds along with results for their DNA interactions are given here.

4.2. Biosensor-SPR Binding Affinity Results for Modified Thiophene-*N*-RBI Modules

In Table 1, results for seven derivatives of the extended compound, DB2457 are shown. Converting the *N*-Me to *N*-Et gives similar binding to the single G-C bp target sequence with improved selectivity over the pure AT and the GC sequence (not shown). The *N*-*i*-Pr compound, DB2708, is even more improved with a selectivity ratio of 255 for AT and 56 for GC. Enlarging the substituent to isobutyl, DB2711, yields a new potential therapeutic candidate with slightly reduced binding to the single G-C bp sequence, but remarkably no detectable binding to either the pure AT or the GC sequence (Table 1).

Table 1. Summary of binding affinity (K_D , nM) for the interaction of representative compounds with biotin-labeled DNA sequences using biosensor-SPR method ^a.

Compounds	X	Y	K_D (nM)/Selectivity Ratio AAATTT	K_D (nM) AAAGTTT	K_D (nM)/Selectivity Ratio AAAGCTTT
DB2457	H	Me	222/56	4	192/48
DB2708	H	<i>i</i> -Pr	1020/255	4	223/56
DB2711	H		NB/-	13	NB/-
DB2714	H		NB/-	9	1330/148
DB2727	H		NB/-	14	388/28
DB2740	H		NB/-	3	107/36
DB2759	Cl	<i>i</i> -Pr	NB/-	14	NB/-
DB2762	CF ₃	<i>i</i> -Pr	NB/-	94	NB/-

^a All the results in this table were obtained in Tris-HCl buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 0.05% P20, pH 7.4) at a 100 μ L/min flow rate. NB means no binding. The listed binding affinities are an average of two independent experiments carried out with two different sensor chips and the values are reproducible within 10% experimental errors.

SPR sensorgrams for DB2711 with the three DNAs are shown in Figure 5 to illustrate how NB (not binding) is defined experimentally. Results in all three sensorgrams in Figure 5, were collected up to the same maximum concentration. As can be seen under these conditions, neither the sensorgrams for the AAATTT or the two G-C bps sequence moved significantly off of the baseline at the maximum concentration and this is our operational definition of “NB”. The sensorgram results for the single G-C sequence show quite significant binding with slow dissociation kinetics. The black lines in the figure are for a single site kinetics fit to the SPR results. The K_D values are given in Table 1 and the kinetics values are discussed below.

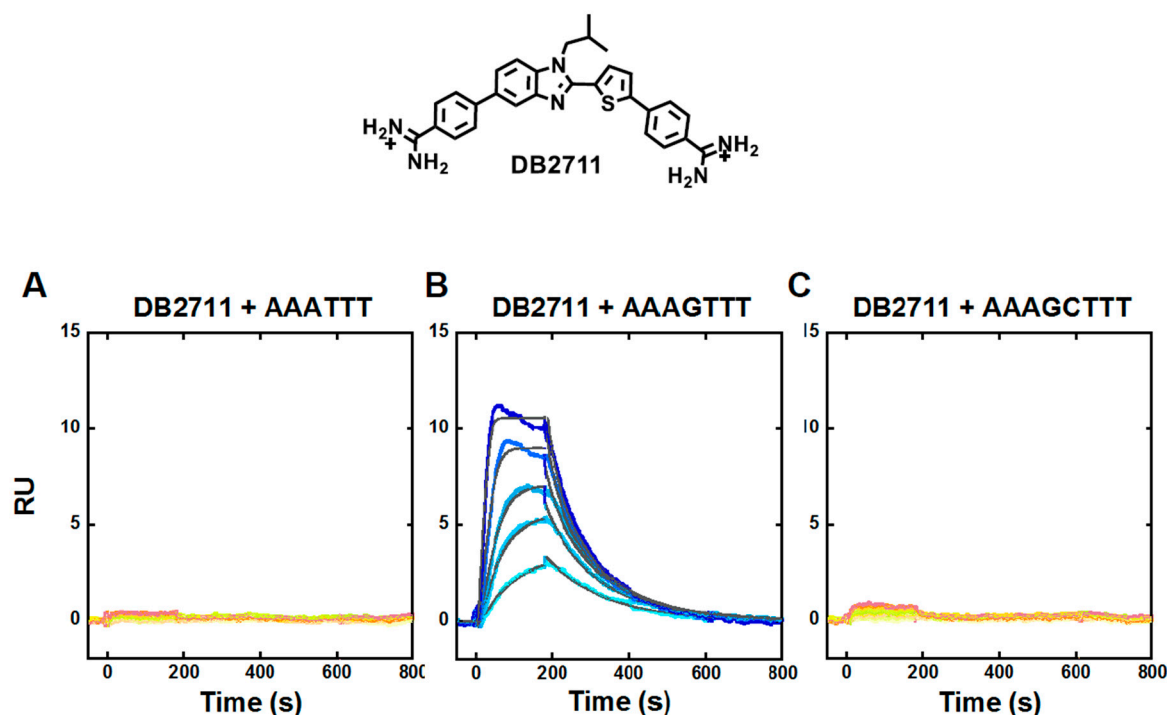


Figure 5. Representative SPR sensorgrams for DB2711 in the presence of (A) AAATTT, (B) AAAGTTT and (C) AAAGCTTT hairpin DNAs. In A, C, the concentrations in the sensorgrams are 2–100 nM of each compound from bottom to top. In B, the concentrations of DB2711 from bottom to top are 15, 30, 40, 70, and 100 nM; In B, the solid black lines are best-fit values for global kinetic fitting of the results with a single site function.

Cyclic *N*-substituents were also explored for binding affinity and specificity. The cyclopentyl derivative, DB2714, has good affinity for the single G-C bp and high selectivity but is not quite as good as the isobutyl derivative, DB2711. The cyclohexyl derivative, DB2727, has weaker binding to the single G-C bp sequence and poor selectivity with GC. Clearly, the more bulky cyclohexyl has stronger binding to the wider GC minor groove than the cyclopentyl compound but does not fit as well in the AAAGTTT sequence.

To explore *N*-aromatic groups, an *N*-phenyl derivative, DB2740, was prepared. It was found to have good binding to the single GC sequence, no detectible binding to AT but poor selectivity for GC. These results suggest that the twist of the phenyl substituent gives a good fit to the target single GC sequence but poor fit to the narrow minor groove of pure AT. The binding to GC is better, probably because the twist of the phenyl with the lowest energy gives a good fit to the wider GC. A number of other derivatives have been prepared but with no improvement over the isobutyl derivative, DB2711 [81].

In order to change the overall shape of the core aromatic-amidine system, several derivatives were prepared. DB2762 with a $-CF_3$ group *ortho* to an amidine gave no detectible binding to AT and GC but relatively weak binding to the single GC sequence. Clearly the large twist of the phenyl-amidine system caused by the $-CF_3$ group does not fit any of the sequence-dependent minor groove shapes. DB2759 with an *ortho* $-Cl$ substituent is a much improved compound. It has strong binding to the single GC but no detectible binding to AT and GC. DB2759 joins DB2711, the *N*-isobutyl compound, as our two best therapeutic leads. Other derivatives were not as good as these compounds and are not shown.

In order to understand the difference among the $-H$, $-Cl$, and $-CF_3$ substituents at the same phenyl-amidine substitution position, torsional angle maps were constructed in the SPARTAN 16 software (Figure 6) [92]. The lowest relative energy represents the most stable twist structure and also shows the energy needed to change the shape into a more planar conformation to fit the DNA minor

groove. Around 10 kJ/mol is needed to twist the 30° structure of –H ortho to a planar conformation of the phenyl-amidine group. For –Cl, 12 kJ/mol is needed to give a similar result. The –CF₃ modification with its larger bulk, however, needs more energy to twist its larger torsional angle into a planar conformation. All of the above results show that, when lower energy is needed to reach a shape that can bind to the minor groove, stronger binding to the target sequence is obtained. Different final twist angles than 0° can be used but the general conclusions are the same.

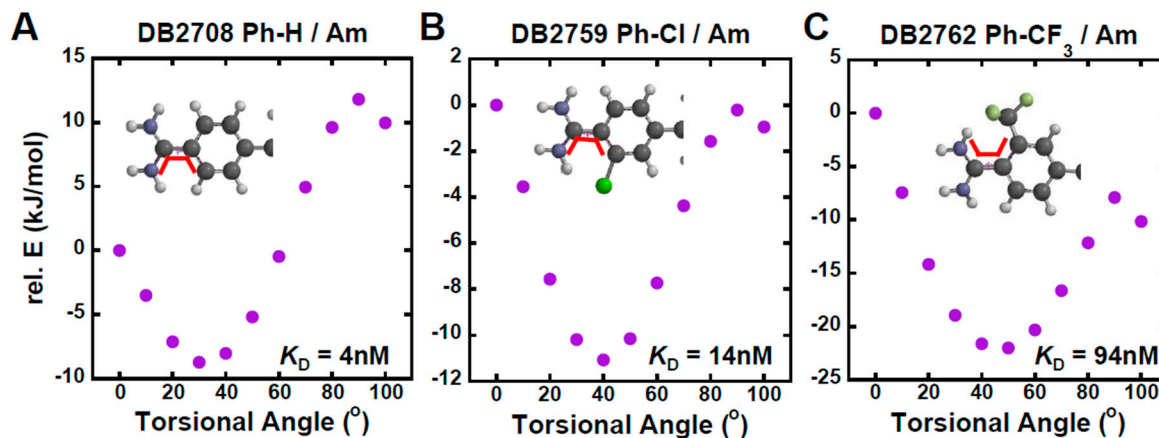


Figure 6. Torsional angle maps of a Ph-Am bond for (A) DB2708, (B) DB2759 and (C) DB2762. All calculations are performed at the B3LYP/6-31G* level of theory. The range of dihedral angle is from 0°–100°. The scanned dihedral is shown as the red bold line at 0°.

4.3. Biosensor-SPR Binding Kinetics Results for Modified Thiophene-N-RBI Modules

Five compounds from Table 1 with a variety of *N*-substituents were used to compare the kinetics of association, k_a , and dissociation, k_d , for the AAAGTTT sequence. The binding association and dissociation kinetics were quite rapid with the AAATTT and AAAGCTTT sequences and could not be determined by SPR methods. For comparison of the kinetics results for the five test compounds with AAAGTTT, $\log k_a$ and $\log k_d$ values are plotted in Figure 7 with results in Table 1. The blue dashed lines in Figure 7 compare compound binding affinities on a logarithmic scale. The *N*-Me reference, DB2457, and strong binding compounds, DB2740 and DB2708, are located on a K_D line with a slope of 3–4 nM. The other two compounds, which bind somewhat more weakly, are located near the 10 nM K_D blue dashed line in Figure 7. The on-rates and off-rates for the compounds change in a correlated manner to give similar K_D values.

For the *N*-substituted compounds in Figure 7, the K_D values with AAAGTTT cover a narrow range, from around 4 to 13 nM. As can be seen in Figure 7, the kinetics results for these compounds also cover a relatively narrow range. The *N*-Me (DB2457) and *N*-Ph (DB2740) compounds have the lowest K_D values in this set and this arises from their similar k_d values. DB2711 with an isobutyl substituent has the fastest disassociation rates with the weakest binding (9.6 nM) to AAAGTTT among the compounds in Figure 7. These results suggest that the very dynamic structure of the butyl substituent interferes with binding in the minor groove giving compounds fast association and dissociation with the highest K_D values in this set. These observations indicate that the faster association results from a lack of optimum penetration into the minor groove and, as a result, the compounds also dissociate more rapidly. It should be noted, however, that even the weakest binding of the compounds in Figure 7 bind very strongly, in the 10 nM K_D range. It should also be noted that the relatively small variation in K_D for the compounds in this set arises primarily from small variations in k_d values (Figure 7). The k_a values have even less variation and are quite large.

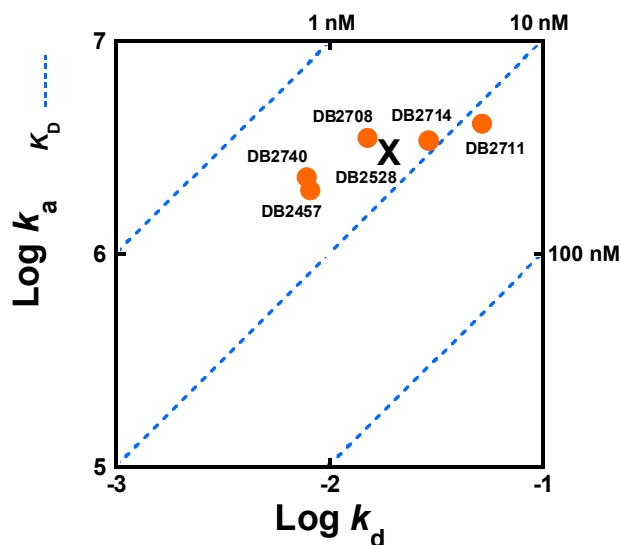


Figure 7. Kinetic plot showing relationships between on-rates (k_a), off-rates (k_d) and binding affinities (K_D) (Blue dash line) represented on the diagonal axis for representative compounds, and determined by SPR. On-rates and off-rates are plotted as logarithmic values (Circles) and for DB2528 (X).

4.4. A DB2708 Model from Molecular Dynamics Simulations

In order to get a better understanding of how the larger substituents on the thiophene-*N*-RBI fit into the minor groove, a molecular dynamics (MD) simulation was conducted on the DB2708 *N*-isopropyl compound DNA complex by using methods that we have previously described [93,94]. DB2708 was docked into our DNA sequence with a single G·C bp and flanking A·T bps. DB2708 force constants for the MD simulation were determined as previously described [94,95]. The MD simulation was conducted for 500 ns, and describe structure for the complex of DB2708 in the AAAGTTT binding site is shown in Figure 8. The MD results were evaluated in detail to determine what major features of the DB2708-DNA complex are responsible for the excellent stability and specificity ratios of the *N*-modified compounds.

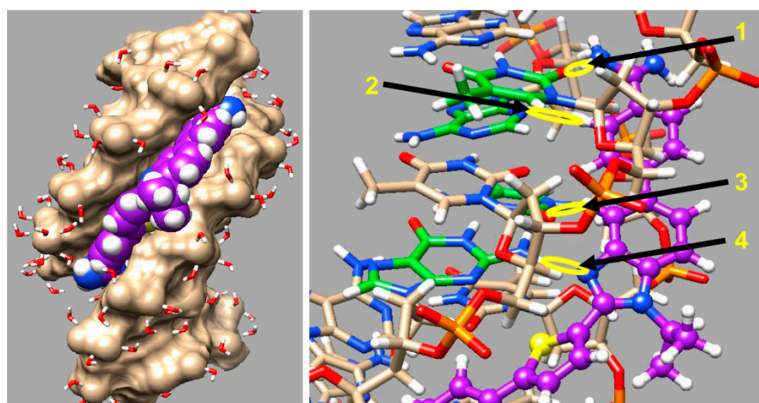


Figure 8. Molecular Dynamics (MD) model of DB2708 bound to an AAAGTTT site. A space filling model with DNA tan color is shown on the left where the isopropyl group and amidines of DB2708 (violet) are clearly seen. On the right side, the DNA bases are represented in tan-white-red-blue-orange (C–H–O–N–P) color scheme and DB2708 is violet -white-blue-yellow (C–H–N–S) color scheme. The detailed important interactions, H-bonds and other interactions described in the text, between the amidine-phenyl-*N*-isopropyl-BI and thiophene of DB2708 with DNA bases are shown with yellow color circle and arrows; arrow 1: Am-NH—O=dT, arrow 2: Ph-C-H—N3-dA, arrow 3: BI-C-H—O=dC, arrow 4: BI-N—H-dG.

There are three critical H-bonds in the complex, two from amidine -NHs to T=O groups that are an average of 2.8 Å in length. The third strong H-bond with 2.2 Å bond length is from the G·C bp G-NH in the minor groove to the unsubstituted imidazole N in the BI group of DB2708 and accounts for a significant amount of the binding selectivity for DB2708 and other thiophene-*N*-RBI compounds length. The complex is additionally stabilized by water molecules that link DB2708 to the DNA binding site. The amidines, for example, form numerous dynamic H-bonds to water molecules that move in and out of the minor groove. Additional H-bonds with DB2708 are frequently formed by water molecules with A·T bp groups at the floor of the minor groove, and these interactions help link the compound to the specific binding site in the groove and stabilize the complex. Selectivity in binding is also enhanced by the -CH group of the six-member ring of BI that points into the minor groove. This -CH forms a dynamic close interaction with the -C=O of the dC base of the central G·C bp as well as to a dT=O on an adjacent A·T bp (Figure 8). This interaction enhances the binding selectivity as well as the binding affinity.

As seen with other minor groove binders, additional stabilizing interactions are formed by phenyl -CH groups of DB2708 that point to the floor of the minor groove. These -CH groups form significant dynamic interactions with dA-N3 groups on the bases at the floor of the groove (see Figure 8 for an example). DB2708 tracks well along the minor groove with a twist and shape to match the minor groove curvature. As a result of this shape match, the aromatic-system of DB2708 forms strong interactions with the sugar-phosphate walls of the minor groove. The cationic amidines of DB2708 also form electrostatic interactions with the backbone phosphates that result in Na⁺ release and an entropy increase that is a stabilizing component, for complex formation.

The release of strongly bound water molecules from the minor groove as DB2708 binds also provide a favorable entropy term to the binding energetics. The sum of these numerous weak, stabilizing interactions results in the observed strong binding of DB2708, a ΔG of binding of 11.5 kcal/mol. The isopropyl group of DB2708 forms a tight interaction with the DNA backbone in the G·C bp region and this is not possible with the narrower minor groove of a pure AT sequence (Figure 4). The K_D for binding to the AAATTT sequence, for example, jumps from 200 nM with DB2457, the *N*-Me compound, to around 1000 nM with DB2708 due to the *N*-isopropyl group (Table 1) and this provides a significant increase in selectivity. It seems that many of the *N*-RBI derivatives can match the minor groove microstructure in AAAGTTT better than with AAATTT.

5. Extending the Capability to Recognize Mixed BP Sequences of DNA to More Than a Single G·C BP

While the three types of compounds described above for converting pure AT DNA minor groove compounds to recognize mixed DNA sequences with a G·C bp flanked by A·T bps was a major step forward, it is clearly only the first step. To make this discovery widely useful, recognition of different DNA sequences, such as, for example, TF promoters, with several G·C bps is essential [96–99]. An attractive way to start on the extended sequence compounds is by combining modules from the three single G·C bp recognition compounds. In this section we will describe our initial steps in this direction with the thiophene-*N*-MeBI module.

The critical question is how to combine single GC recognition modules, such as the thiophene-*N*-MeBI in DB2429, to selectively recognize longer, more complex DNA sequences with additional G·C bp. Although progress in targeting DNA with proteins has been excellent, there is limited progress in the variety of compounds designed to target mixed base pair (bp) DNA sequences. Compounds are needed that can selectively recognize two (or more) G·C bps in DNA sequences such as (A/T)_x-(G/C)-(A/T)_n-(G/C)-(A/T)_y where “n” can be 0–5 bps (Figure 9B) and x and y are variable.

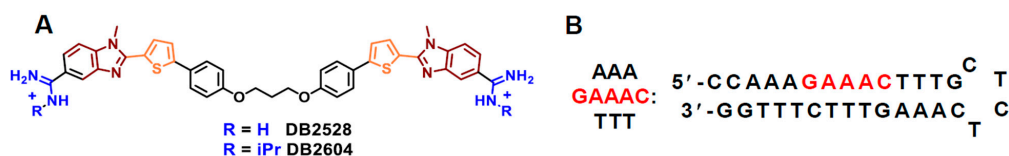
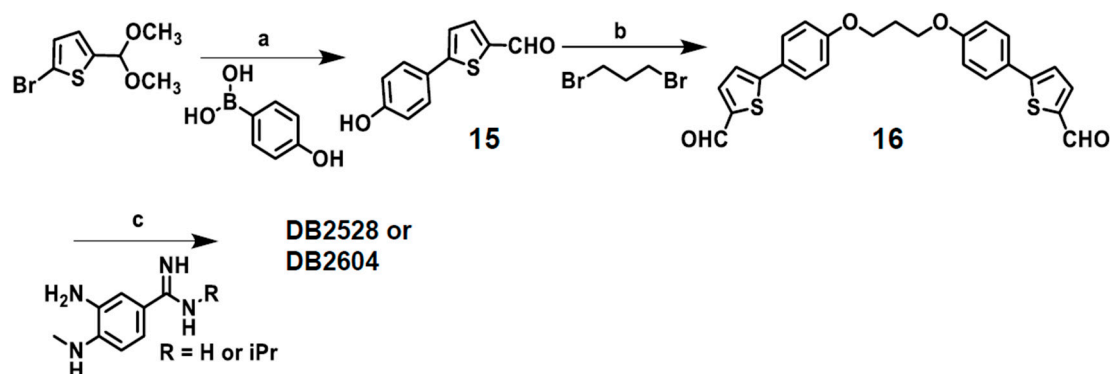


Figure 9. (A) Structures of minor groove binders for two G·C bps containing DNA; (B) DNA hairpin oligomer duplex (B) employed in this study.

Because of the considerable difficulty of preparing longer, complex compounds that have the correct curvature for the minor groove and proper indexing of H-bonding groups to match the array of DNA H-bonding groups at the floor of the minor groove, we started with relatively simple, flexible combinations. For the initial test of this approach, a flexible linker, $-\text{O}-(\text{CH}_2)_3-\text{O}-$, was used to connect GC recognition molecular modules based on DB2429 (Figure 9A). The design strategy provides flexibility to match the shape and curvature of the DNA minor groove with appropriate length so that modules are spaced to H-bond and interact strongly with a GAAAC DNA sequence. Such modules are reasonable to synthesize and can recognize a full turn or more of the DNA double helix.

Synthetic Procedures for DB2528 and DB2604

The syntheses of the two GC recognition compounds are outlined in Scheme 4. Protected 5-(4-hydroxyphenyl)thiophene-2-carboxaldehyde (**15**) is readily prepared through a standard Suzuki coupling reaction between the acetal of 5-bromo-2-thiophene carboxaldehyde and 4-hydroxyphenylboronic acid. Standard Williamson ether synthesis methodology was employed to link 1,3-dibromopropane with 5-(4-hydroxyphenyl) thiophene-2-carboxaldehyde to produce the bisaldehyde **16**. The target products bis-amidino-*N*-methylbenzimidazolethiophenes are prepared from amidino-*N*-methylphenylenediamines by coupling with the symmetrical bis-thiophenecarboxaldehyde in the presence of benzoquinone (Scheme 4) [100].



Scheme 4. Reagents and conditions: (a) (i) $\text{Pd}(\text{PPh}_3)_4$, 2 M aqueous K_2CO_3 , EtOH, reflux, (ii) HCl, H_2O . (b) K_2CO_3 , DMF, 80°C . (c) 1,4-benzoquinone, DMF, EtOH, reflux.

DNA thermal melting experiments (T_m) offer a rapid, screening method for compounds binding to different DNA sequences [101,102]. Binding to DNA gives an increase in T_m that is related to binding affinity. Hairpin DNA oligomers, which have monophasic melting curves, were used in the assay. DB2528 and analogs (Figure 9A) are the first linked compounds to test this concept for two G·C bps binding. DB2528 does not bind well to the single GC test sequence (AAAAGTTTT, $\Delta T_m = 2^\circ\text{C}$), but in an exciting result it binds with a ΔT_m of 9°C with a target two G·C bps sequence, GAAAC. As desired, it has lower T_m values with longer, GAAAAC and GAAAAC sequences (6°C) and shows no significant binding with an all AT sequence or with related GC sequences (GC, GAC, GAAC), all with $\Delta T_m < 2^\circ\text{C}$. As with the simpler compounds above, replacement of *N*-MeBI in DB2528 with benzimidazole (BI) gives a minor groove binder that binds very strongly to extended AT sequences with weaker binding to G sequences (not shown). Circular dichroism (CD) studies with GAAAC and

DB2528 show a strong positive induced CD peak at the compound absorption maximum that indicates minor groove binding. There is little CD change on addition of DB2528 to AT sequences in agreement with weak binding [91].

As above, the interaction affinities of DB2528 were quantitatively evaluated by biosensor-SPR methods with the mixed base pair sequences. DB2528 binds strongly with GAAAC, in agreement with T_m results, and global kinetics fitting yielded a single binding site with a K_D of 5 nM, a rapid on-rate ($k_a = 2.8 \pm 0.5 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$) and a slow off-rate ($k_d = 1.7 \pm 0.4 \times 10^{-2} \text{ s}^{-1}$) (Figure 7). The dissociation kinetics of DB2528 with the GAAG sequence are much faster than with GAAAC and the $K_D = 149 \text{ nM}$ with GAAG. Pure AT and single G sequences were also studied by SPR, and DB2528 has no detectable binding with either of them. These results indicate excellent selectivity and strong binding for a two G sequence that has a specific distance between the two G-C bps. In order to broaden the solution properties and enhance potential cell uptake in future therapeutic applications, we also synthesized DB2604 with isopropyl modified amidine groups (Figure 9A). For SPR results, the K_D value (12 nM) of DB2604 binding to target GAAAC sequence is slightly higher than the DB2528 K_D value (5 nM). To test the SPR binding results with these large complex compounds, the K_D of DB2604 was also evaluated by a fluorescence anisotropy titration of the compound with the GAAAC DNA sequence, and the K_D value is in good agreement with the results from SPR ($K_D = 4 \pm 2 \text{ nM}$).

An additional test of the binding selectivity of DB2528 was conducted with our competition mass spectroscopy method (Figure 10). In the first experiment binding of DB2528 to a mixture of DNAs are with no GC, A_4T_4 ; a DNA with a single GC, A_4GT_4 ; and a DNA with two G-C bps, GAAAC (left side of Figure 10). As can be seen in Figure 10, with these three DNAs, DB2528 binds only to GAAAC. In the second experiment DB2528 was mixed with a combination of two G-C bps containing sequences, GAAC, GAAAC, GAAAAC, and GAAAAAC. In another very exciting result, the compound only showed complex formation with the GAAAC DNA sequence, in excellent agreement with SPR results. This level of selectivity with such closely related DNAs is a strong validation of the module combination approach for recognition of more complex mixed sequences of DNA.

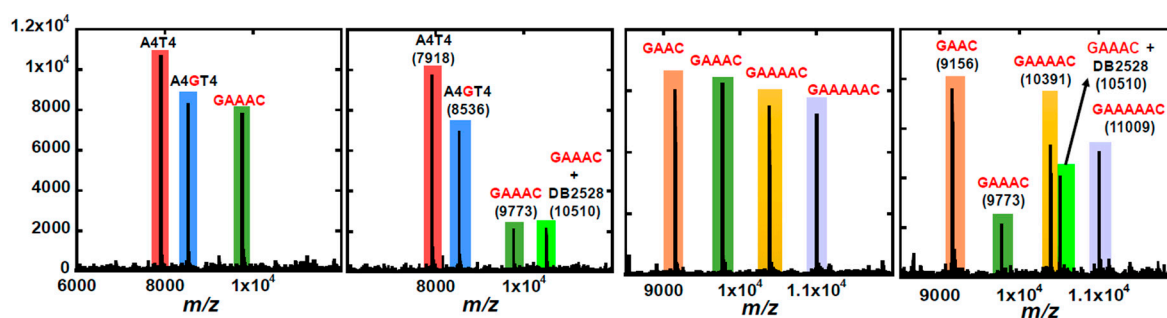


Figure 10. ESI-MS negative mode spectra of the competition binding of sequences (10 μM each) with 40 μM DB2528 in buffer (50 mM ammonium acetate with 10% methanol (v/v), pH 6.8). The ESI-MS results shown here are deconvoluted spectra and molecular weights are shown with each peak.

The results of this work, to link single G-C bp binding modules with a flexible linker for recognition of two G-C bps in the core sequence AGAACT, has functional significance. With optimized design, synthesis and evaluation the linked heterocyclic-diamidine, DB2528 and several derivatives, have been obtained with an excellent match to the DNA minor groove shape and indexing of the compound with DNA functional groups at the floor of the groove. These compounds have syntheses that are reasonable in cost and time (Figure 9A) and provide a promising approach to develop a broad array of linked modular agents for competition with protein-DNA complexes and control of gene expression. These results illustrate what can be accomplished with modular compounds of this type in selective targeting of complex DNA. A number of additional compounds to recognize multiple G-C bps are being tested.

To obtain molecular insight of the very sequence-specific two G·C bps binding by the symmetrical dimer, DB2528, (Figure 9A) with the two G·C bps containing sequence, a 300 ns MD simulation of DB2528 with the GAAAC sequence (Figure 9B) was conducted as previously described [93,95]. The space-filling full view in Figure 11 at left side shows that DB2528 is able to match the curvature of the DNA minor groove and can cover a full turn of the double helix. The MD simulations showed the excellent contacts and van der Waals interactions that the compound makes with the minor groove molecular walls (Figure 11, left). A view of the complex structure at the upper part of the model (yellow circle) is shown in Figure 11 at the right side. This view shows a strong H-bond at a 2.02 Å distance between the *N* of *N*-MeBI and the dG-NH that points out into the groove. This interaction is a critical component of the G·C bp binding selectivity of the compound. There is also an H-bond between the amidine-NH that points to the floor of the groove and a dT=O (dT = O—H—N distance is 2.1 Å). The complex is additionally stabilized by water molecules that link DB2528 to the DNA binding site. The amidines form dynamic H-bonds to water molecules that enhance the stability of the H-bonds from the amidine and DNA (Am-N-H—O—H—N3-dA, distances are 2.1 Å and 2.0 Å respectively), right side of Figure 11.

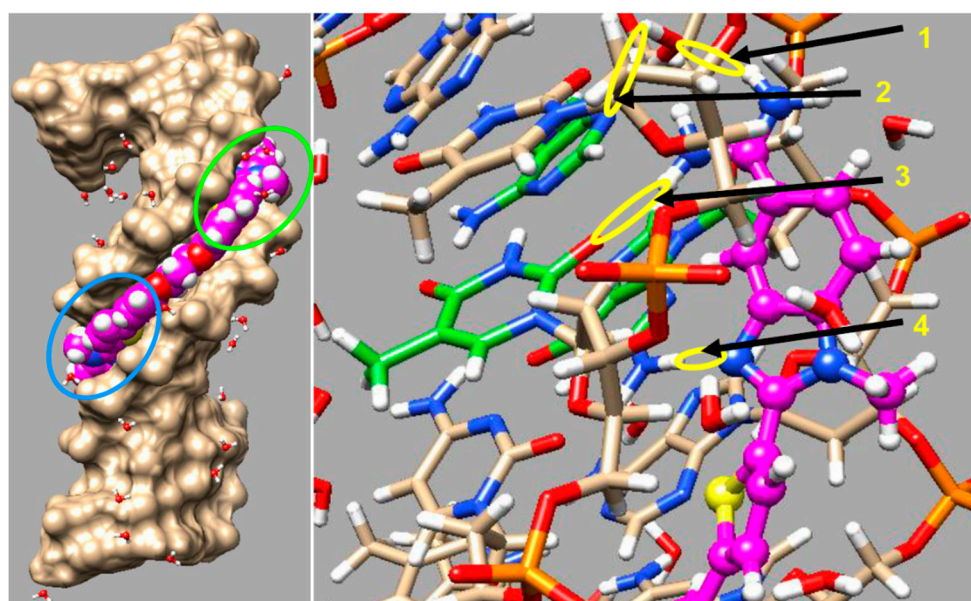


Figure 11. Molecular Dynamics (MD) model of DB2528 bound to a two G·C bps containing sequence AAAGAACTTT site. A space filling model with DNA tan color is shown on the left where the compound DB2528 (magenta) covers a full helical turn of 10 bps in the minor groove. On the right side, the DNA bases are represented in tan-white-red-blue-orange (C—H—O—N—P) color scheme and DB2528 is magenta-white-blue-yellow (C—H—N—S) color scheme. The detailed important interactions, H-bonds and other interactions described in the text, between the symmetrical bis amidine-phenyl-*N*-methyl-BI-thiophene of DB2258 with DNA bases are shown with yellow color circle and arrows; arrow 1: Am-N-H—O-H₂, arrow 2: H₂-O—N₃-dA, arrow 3: Am-NH—O=dT, arrow 4: BI-N—H-dG.

At the other end of the complex (blue circle in Figure 11, left) the same type of H-bond and interaction pattern is obtained (not shown). The flexible -O-(CH₂)₃-O- linker is an appropriate length to effectively cover the AAA sequence between the two G·C bps. The compound excellent curvature, orientation, linker length and the specific interactions observed between DB2528 and the GAAAC sequence at the minor groove provide a rational explanation for the observed two G·C bps DNA recognition.

6. Conclusions

Description of the design, preparation and biophysical testing of entirely new types of heterocyclic amidine compounds with good solution and cell uptake properties that selectively recognize mixed A·T and G·C bp sequences of DNA is described in this report. These compounds are the first example of a series of rationally-designed agents for mixed sequence DNA recognition and they provide very different biotechnology reagents and new therapeutic candidates compared with those previously reported. Synthetic schemes for compounds that bind mixed base pair DNA sequences are given. To selectively recognize a G·C bp an H-bond acceptor must be incorporated into a compound. Pyridine, azabenzimidazole and thiophene-*N*-MeBI GC recognition groups have been appropriately incorporated in modules prepared by using both rational design and empirical optimization of the units. These modules can selectively and strongly recognize a single G·C bp in an AT sequence context. It is shown that in many cases relatively simple changes in substituents can convert a heterocyclic module from AT to GC recognition. Developments of lead compounds for recognition of G·C bps in DNA are presented. These compounds represent the first successful step for a new generation of diverse minor groove binders for mixed DNA sequence binding. Addition of more bulky groups in place of the *N*-methyl on the thiophene-*N*-MeBI GC recognition module shows that, like transcription factors, small molecules can be engineered for selective recognition of both sequence and shape in the minor groove. This dual recognition capability is a very useful step in our compound design efforts and provides compounds with much enhanced recognition selectivity. Modification of our initial sequence selective DNA binding agents has now moved to preparation and testing of larger compounds for recognition of more complex sequences of DNA with more than one G·C bps. Current research is focused on the preparation and linking of different modules to recognize these more complex mixed sequences of DNA.

Author Contributions: A.P., P.G. and W.D.W. designed the study and experiments. A.P., P.G. and W.D.W. designed the biosensor-SPR experiments. A.P., P.G. and W.D.W. designed the DNA Uv-vis thermal melting and ESI-MS and other biophysical experiments. A.P. and W.D.W. designed and executed the molecular dynamic simulations. D.W.B. designed the novel compound syntheses. A.P., and P.G. performed biophysical experiments. A.P., P.G. and W.D.W. analyzed and interpreted the experimental data. A.P., P.G., D.W.B., and W.D.W. wrote the manuscript.

Funding: This work was supported by the US National Institutes of Health (NIH) [Grant No. GM111749 to W.D.W and D.W.B.

Acknowledgments: The authors are thankful to Arvind Kumar, Abdelbasset A. Farahat, and Yun Chai, for synthesizing small organic molecules used in this study. The authors thank Carol Wilson for manuscript assistance.

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

References

1. Cossar, P.J.; Lewis, P.J.; McCluskey, A. Protein–protein interactions as antibiotic targets: A medicinal chemistry perspective. *Med. Res. Rev.* **2018**. [[CrossRef](#)] [[PubMed](#)]
2. 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. [[CrossRef](#)] [[PubMed](#)]
3. Nero, T.L.; Parker, M.W.; Morton, C.J. Protein structure and computational drug discovery. *Biochem. Soc. Trans.* **2018**, *46*, 1367–1379. [[CrossRef](#)] [[PubMed](#)]
4. Vajjayanthi, T.; Pandian, G.N.; Sugiyama, H. Chemical control system of epigenetics. *Chem. Rec.* **2018**, *18*, 1833–1853. [[CrossRef](#)] [[PubMed](#)]
5. Chen, Y.; Yang, D. Sequence, stability, and structure of G-Quadruplexes and their interactions with drugs. *Curr. Protoc. Nucleic Acid Chem.* **2012**. [[CrossRef](#)]
6. Bouhleh, M.; Lambert, M.; David-Cordonnier, M.-H. Targeting transcription factor binding to DNA by competing with DNA binders as an approach for controlling gene expression. *Curr. Top. Med. Chem.* **2015**, *15*, 1323–1358. [[CrossRef](#)] [[PubMed](#)]
7. Neidle, S.; Thurston, D.E. Chemical approaches to the discovery and development of cancer therapies. *Nat. Rev. Cancer* **2005**, *5*, 285–296. [[CrossRef](#)] [[PubMed](#)]

8. Rahman, A.; O'Sullivan, P.; Rozas, I. Recent developments in compounds acting in the DNA minor groove. *MedChemComm* **2019**, *10*, 26–40. [[CrossRef](#)] [[PubMed](#)]
9. Bhaduri, S.; Ranjan, N.; Arya, D.P. An overview of recent advances in duplex DNA recognition by small molecules. *Beilstein J. Org. Chem.* **2018**, *14*, 1051–1086. [[CrossRef](#)] [[PubMed](#)]
10. Evison, B.J.; Sleebs, B.E.; Watson, K.G.; Phillips, D.R.; Cutts, S.M. Mitoxantrone, more than just another Topoisomerase II poison. *Med. Res. Rev.* **2016**, *36*, 248–299. [[CrossRef](#)] [[PubMed](#)]
11. Holt, P.A.; Buscaglia, R.; Trent, J.O.; Chaires, J.B. A discovery funnel for nucleic acid binding drug candidates. *Drug Dev. Res.* **2011**, *72*, 178–186. [[CrossRef](#)] [[PubMed](#)]
12. Sánchez, M.I.; Mosquera, J.; Vázquez, E.M.; Mascareñas, J.L. Reversible supramolecular assembly at specific DNA sites: Nickel-promoted bivalent DNA binding with designed peptide and bipyridyl-bis(benzamidine) components. *Angew. Chem. Int. Ed.* **2014**, *53*, 9917–9921. [[CrossRef](#)] [[PubMed](#)]
13. Mosquera, J.; Jiménez-Balsa, A.; Doderó, V.I.; Vázquez, E.M.; Mascareñas, J.L. Stimuli-responsive selection of target DNA sequences by synthetic bZIP peptides. *Nat. Commun.* **2013**, *4*, 1874. [[CrossRef](#)] [[PubMed](#)]
14. Dilek, I.; Madrid, M.; Singh, R.; Urrea, C.P.; Armitage, B.A. Effect of PNA backbone modifications on cyanine dye binding to PNA–DNA duplexes investigated by optical spectroscopy and molecular dynamics simulations. *J. Am. Chem. Soc.* **2005**, *127*, 3339–3345. [[CrossRef](#)] [[PubMed](#)]
15. Waring, M. Binding of antibiotics to DNA. *Ciba Found. Symp.* **1991**, *158*, 128–142, discussion 142–146, 204–212. [[PubMed](#)]
16. Portugal, J.; Waring, M.J. Comparison of binding sites in DNA for berenil, netropsin and distamycin. *Eur. J. Biochem.* **1987**, *167*, 281–289. [[CrossRef](#)] [[PubMed](#)]
17. Luck, G.; Triebel, H.; Waring, M.; Zimmer, C. Conformation dependent binding of netropsin and distamycin to DNA and DNA model polymers. *Nucleic Acids Res.* **1974**, *1*, 503–530. [[CrossRef](#)] [[PubMed](#)]
18. Pjura, P.E.; Grzeskowiak, K.; Dickerson, R.E. Binding of Hoechst 33258 to the minor groove of B-DNA. *J. Mol. Biol.* **1987**, *197*, 257–271. [[CrossRef](#)]
19. Kubista, M.; Aakerman, B.; Norden, B. Characterization of interaction between DNA and 4',6-diamidino-2-phenylindole by optical spectroscopy. *Biochemistry* **1987**, *26*, 4545–4553. [[CrossRef](#)] [[PubMed](#)]
20. Bailly, C.; Pommery, N.; Houssin, R.; Henichart, J.-P. Design, synthesis, DNA binding, and biological activity of a series of DNA minor-groove-binding intercalating drugs. *J. Pharm. Sci.* **1989**, *78*, 910–917. [[CrossRef](#)] [[PubMed](#)]
21. Nanjunda, R.; Wilson, W.D. Binding to the DNA minor groove by heterocyclic dications: From AT-specific monomers to GC recognition with dimers. *Curr. Protoc. Nucleic Acid Chem.* **2012**. [[CrossRef](#)] [[PubMed](#)]
22. Lavery, R.; Pullman, B.; Zakrzewska, K. Intrinsic electrostatic properties and base sequence effects in the structure of oligonucleotides. *Biophys. Chem.* **1982**, *15*, 343–351. [[CrossRef](#)]
23. Auffinger, P.; Westhof, E. Water and ion binding around r(UpA)₁₂ and d(TpA)₁₂ Oligomers—Comparison with RNA and DNA (CpG)₁₂ duplexes. *J. Mol. Biol.* **2001**, *305*, 1057–1072. [[CrossRef](#)] [[PubMed](#)]
24. Suckling, C. From multiply active natural product to candidate drug? Antibacterial (and other) minor groove binders for DNA. *Future Med. Chem.* **2012**, *4*, 971–989. [[CrossRef](#)] [[PubMed](#)]
25. 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. [[PubMed](#)]
26. Donlic, A.; Hargrove, A.E. Targeting RNA in mammalian systems with small molecules. *Wiley Interdiscip. Rev. RNA* **2018**, *9*, e1477. [[CrossRef](#)] [[PubMed](#)]
27. Disney, M.D.; Yildirim, I.; Childs-Disney, J.L. Methods to enable the design of bioactive small molecules targeting RNA. *Org. Biomol. Chem.* **2013**, *12*, 1029–1039. [[CrossRef](#)] [[PubMed](#)]
28. Abulwerdi, F.A.; Schneekloth, J.S. Microarray-based technologies for the discovery of selective, RNA-binding molecules. *Methods* **2016**, *103*, 188–195. [[CrossRef](#)] [[PubMed](#)]
29. Hermann, T. Small molecules targeting viral RNA. *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 726–743. [[CrossRef](#)] [[PubMed](#)]
30. Wong, C.-H.; Nguyen, L.; Peh, J.; Luu, L.M.; Sanchez, J.S.; Richardson, S.L.; Tuccinardi, T.; Tsoi, H.; Chan, W.; Chan, Y.E.H.; et al. Targeting toxic RNAs that cause Myotonic Dystrophy Type 1 (DM1) with a bisamidinium inhibitor. *J. Am. Chem. Soc.* **2014**, *136*, 6355–6361. [[CrossRef](#)] [[PubMed](#)]

31. Hilimire, T.A.; Chamberlain, J.M.; Anokhina, V.; Bennett, R.P.; Swart, O.; Myers, J.R.; Ashton, J.M.; Stewart, R.A.; Featherston, A.L.; Gates, K.; et al. HIV-1 frameshift RNA-Targeted triazoles inhibit propagation of replication-competent and multi drug-resistant HIV in human cells. *ACS Chem. Biol.* **2017**, *12*, 1674–1682. [[CrossRef](#)] [[PubMed](#)]
32. Marcheschi, R.J.; Tonelli, M.; Kumar, A.; Butcher, S.E. Structure of the HIV-1 frameshift site RNA bound to a small molecule inhibitor of viral replication. *ACS Chem. Biol.* **2011**, *6*, 857–864. [[CrossRef](#)] [[PubMed](#)]
33. Ghosh, A.; Degyatoreva, N.; Kukielski, C.; Story, S.; Bhaduri, S.; Maiti, K.; Nahar, S.; Ray, A.; Arya, D.P.; Maiti, S. Targeting miRNA by tunable small molecule binders: Peptidic aminosugar mediated interference in miR-21 biogenesis reverts epithelial to mesenchymal transition. *MedChemComm* **2018**, *9*, 1147–1154. [[CrossRef](#)] [[PubMed](#)]
34. Chazotte, B. Labeling nuclear DNA with Hoechst 33342. *Cold Spring Harb. Protoc.* **2011**, 2011. [[CrossRef](#)] [[PubMed](#)]
35. 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. [[CrossRef](#)] [[PubMed](#)]
36. 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. [[CrossRef](#)] [[PubMed](#)]
37. Soeiro, M.N.C.; Werbovetz, K.; Boykin, D.W.; Wilson, W.D.; Wang, M.Z.; Hemphill, A. Novel amidines and analogues as promising agents against intracellular parasites: A systematic review. *Parasitology* **2013**, *140*, 929–951. [[CrossRef](#)] [[PubMed](#)]
38. Ming, X.; Ju, W.; Wu, H.; Tidwell, R.R.; Hall, J.E.; Thakker, D.R. Transport of dicationic drugs pentamidine and furamidine by human organic cation transporters. *Drug Metab. Dispos.* **2009**, *37*, 424–430. [[CrossRef](#)] [[PubMed](#)]
39. Millan, C.R.; Acosta-Reyes, F.J.; Lagartera, L.; Ebiloma, G.U.; Lemgruber, L.; Nué Martínez, J.J.; Saperas, N.; Dardonville, C.; de Koning, H.P.; Campos, J.L. Functional and structural analysis of AT-specific minor groove binders that disrupt DNA-protein interactions and cause disintegration of the *Trypanosoma brucei* kinetoplast. *Nucleic Acids Res.* **2017**, *45*, 8378–8391. [[CrossRef](#)] [[PubMed](#)]
40. Vázquez, O.; Sánchez, M.I.; Martínez-Costas, J.; Vázquez, M.E.; Mascareñas, J.L. Bis-4-aminobenzamidines: Versatile, fluorogenic A/T-selective dsDNA binders. *Org. Lett.* **2010**, *12*, 216–219. [[CrossRef](#)] [[PubMed](#)]
41. Bordello, J.; Sánchez, M.I.; Vázquez, M.E.; Mascareñas, J.L.; Al-Soufi, W.; Novo, M. Fluorescence-labeled bis-benzamidines as fluorogenic DNA minor-groove binders: Photophysics and binding dynamics. *Chem. Eur. J.* **2015**, *21*, 1609–1619. [[CrossRef](#)] [[PubMed](#)]
42. 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**, *3*, 2383–2387. [[CrossRef](#)]
43. 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. [[CrossRef](#)]
44. 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 exhibit altered DNA sequence specificity. *Biochemistry* **1986**, *25*, 7408–7416. [[CrossRef](#)] [[PubMed](#)]
45. 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. [[CrossRef](#)] [[PubMed](#)]
46. Dervan, P.B.; Bürli, R.W. Sequence-specific DNA recognition by polyamides. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688–693. [[CrossRef](#)]
47. Wemmer, D.E.; Dervan, P.B. Targeting the minor groove of DNA. *Curr. Opin. Struct. Biol.* **1997**, *3*, 355–361. [[CrossRef](#)]
48. Dervan, P.B.; Edelson, B.S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, *3*, 284–299. [[CrossRef](#)]
49. 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. [[CrossRef](#)] [[PubMed](#)]

50. 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. [[CrossRef](#)] [[PubMed](#)]
51. 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. [[CrossRef](#)] [[PubMed](#)]
52. Scott, F.J.; Khalaf, A.I.; Giordani, F.; Wong, P.E.; Duffy, S.; Barrett, M.; Avery, V.M.; Suckling, C.J. An evaluation of Minor Groove Binders as anti-*Trypanosoma brucei brucei* therapeutics. *Eur. J. Med. Chem.* **2016**, *116*, 116–125. [[CrossRef](#)] [[PubMed](#)]
53. Ismail, M.A.; Brun, R.; Wenzler, T.; Taniou, F.A.; Wilson, W.D.; Boykin, D.W. Novel dicationic imidazo[1,2-a]pyridines and 5,6,7,8-tetrahydro-imidazo[1,2-a]pyridines as antiprotozoal agents. *J. Med. Chem.* **2004**, *47*, 3658–3664. [[CrossRef](#)] [[PubMed](#)]
54. da Silva, C.F.; Batista, D.D.G.J.; de Araújo, J.; Cunha-Junior, E.; Stephens, C.E.; Banerjee, M.; Farahat, A.A.; Akay, S.; Fisher, M.K.; Boykin, D.W.; et al. Phenotypic evaluation and in silico ADMET properties of novel arylimidamides in acute mouse models of *Trypanosoma cruzi* infection. *Drug Des. Dev. Ther.* **2017**, *11*, 1095–1105. [[CrossRef](#)] [[PubMed](#)]
55. 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. [[CrossRef](#)] [[PubMed](#)]
56. Berninger, M.; Schmidt, I.; Ponte-Sucre, A.; Holzgrabe, U. Novel lead compounds in pre-clinical development against African sleeping sickness. *MedChemComm* **2017**, *8*, 1872–1890. [[CrossRef](#)] [[PubMed](#)]
57. Fonseca-Berzal, C.; Arán, V.J.; Escario, J.A.; Gómez-Barrio, A. Experimental models in Chagas disease: A review of the methodologies applied for screening compounds against *Trypanosoma cruzi*. *Parasitol. Res.* **2018**, *117*, 3367–3380. [[CrossRef](#)] [[PubMed](#)]
58. Merouze, F.; Dago, A.; Haller, L.; Adams, H. Clinical and pathological aspects of Human African Trypanosomiasis (T. B. Gambiense) with particular reference to reactive arsenical encephalopathy. *Am. J. Trop. Med. Hyg.* **1986**, *35*, 94–99.
59. Hyman, R.W.; Fung, E.; Conway, A.; Kurdi, O.; Mao, J.; Miranda, M.; Nakao, B.; Rowley, D.; Tamaki, T.; Wang, F.; et al. Sequence of *Plasmodium falciparum* chromosome 12. *Nature* **2002**, *419*, 534–537. [[CrossRef](#)] [[PubMed](#)]
60. Andersson, J.; Li, M.; Lincoln, P. AT-specific DNA binding of binuclear ruthenium complexes at the border of threading intercalation. *Chem. Eur. J.* **2010**, *16*, 11037–11046. [[CrossRef](#)] [[PubMed](#)]
61. Purfield, A.E.; Tidwell, R.R.; Meshnick, S.R. The diamidine DB75 targets the nucleus of *Plasmodium falciparum*. *Malar. J.* **2009**, *8*, 104. [[CrossRef](#)] [[PubMed](#)]
62. Kitchin, P.A.; Klein, V.A.; Ryan, K.A.; Gann, K.L.; Rauch, C.A.; Kang, D.S.; Wells, R.D.; Englund, P.T. A highly bent fragment of *Crithidia fasciculata* kinetoplast DNA. *J. Biol. Chem.* **1986**, *261*, 11302–11309. [[PubMed](#)]
63. Sands, M.; Kron, M.A.; Brown, R.B. Pentamidine: A review. *Clin. Infect. Dis.* **1985**, *7*, 625–634. [[CrossRef](#)]
64. Wilson, W.D.; Taniou, F.A.; Mathis, A.; Tevis, D.; Hall, J.; Boykin, D.W. Antiparasitic compounds that target DNA. *Biochimie* **2008**, *90*, 999–1014. [[CrossRef](#)] [[PubMed](#)]
65. Werbovets, K. Diamidines as antitrypanosomal, antileishmanial and antimalarial agents. *Curr. Opin. Investig. Drugs* **2006**, *7*, 147–157. [[PubMed](#)]
66. Thuita, J.K.; Wolf, K.K.; Murilla, G.A.; Bridges, A.S.; Boykin, D.W.; Mutuku, J.N.; Liu, Q.; Jones, S.K.; Gem, C.O.; Ching, S.; et al. Chemotherapy of second stage Human African Trypanosomiasis: Comparison between the parenteral diamidine DB829 and its oral prodrug DB868 in Vervet monkeys. *PLoS Negl. Trop. Dis.* **2015**, *9*, e0003409. [[CrossRef](#)] [[PubMed](#)]
67. Mathis, A.M.; Holman, J.L.; Sturk, L.M.; Ismail, M.A.; Boykin, D.W.; Tidwell, R.R.; Hall, J. Accumulation and intracellular distribution of antitrypanosomal diamidine compounds DB75 and DB820 in African Trypanosomes. *Antimicrob. Agents Chemother.* **2006**, *50*, 2185–2191. [[CrossRef](#)] [[PubMed](#)]
68. Mathis, A.M.; Bridges, A.S.; Ismail, M.A.; Kumar, A.; Francesconi, I.; Anbazhagan, M.; Hu, Q.; Taniou, F.A.; Wenzler, T.; Saulter, J.; et al. Diphenyl furans and aza analogs: Effects of structural modification on In vitro activity, DNA binding, and accumulation and distribution in Trypanosomes. *Antimicrob. Agents Chemother.* **2007**, *51*, 2801–2810. [[CrossRef](#)] [[PubMed](#)]
69. Pelton, J.G.; Wemmer, D.E. Structural characterization of a 2:1 distamycin A.d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5723–5727. [[CrossRef](#)] [[PubMed](#)]

70. Mrksich, M.; Wade, W.S.; Dwyer, T.J.; Geierstanger, B.H.; Wemmer, D.E.; Dervan, P.B. Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7586–7590. [[CrossRef](#)] [[PubMed](#)]
71. Crowley, K.S.; Phillion, D.P.; Woodard, S.S.; Schweitzer, B.A.; Singh, M.; Shabany, H.; Burnette, B.; Hippenmeyer, P.; Heitmeier, M.; Bashkin, J.K. Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1565–1570. [[CrossRef](#)]
72. Kawamoto, Y.; Bando, T.; Sugiyama, H. Sequence-specific DNA Binding Pyrrole–Imidazole polyamides and their applications. *Bioorg. Med. Chem.* **2018**, *26*, 1393–1411. [[CrossRef](#)] [[PubMed](#)]
73. Kiakos, K.; Pett, L.; Satam, V.; Patil, P.; Hochhauser, D.; Lee, M.; Hartley, J.A. Nuclear localization and gene expression modulation by a fluorescent sequence-selective p-anisyl-benzimidazolecarboxamido Imidazole–Pyrrole polyamide. *Chem. Biol.* **2015**, *22*, 862–875. [[CrossRef](#)] [[PubMed](#)]
74. Szablowski, J.O.; Raskatov, J.A.; Dervan, P.B. An HRE-Binding Py–Im polyamide impairs hypoxic signaling in tumors. *Mol. Cancer Ther.* **2016**, *15*, 608–617. [[CrossRef](#)] [[PubMed](#)]
75. Khalaf, A.I.; Bourdin, C.; Breen, D.; Donoghue, G.; Scott, F.J.; Suckling, C.J.; MacMillan, D.; Clements, C.; Fox, K.; Sekibo, D. Design, synthesis and antibacterial activity of minor groove binders: The role of non-cationic tail groups. *Eur. J. Med. Chem.* **2012**, *56*, 39–47. [[CrossRef](#)] [[PubMed](#)]
76. James, P.L.; Merkina, E.E.; Khalaf, A.I.; Suckling, C.J.; Waigh, R.D.; Brown, T.; Fox, K.R. DNA sequence recognition by an isopropyl substituted thiazole polyamide. *Nucleic Acids Res.* **2004**, *32*, 3410–3417. [[CrossRef](#)] [[PubMed](#)]
77. Brucoli, F.; Guzman, J.D.; Maitra, A.; James, C.H.; Fox, K.R.; Bhakta, S. Synthesis, anti-mycobacterial activity and DNA sequence-selectivity of a library of biaryl-motifs containing polyamides. *Bioorg. Med. Chem.* **2015**, *23*, 3705–3711. [[CrossRef](#)] [[PubMed](#)]
78. 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.* **2016**, *15*, 827–835. [[CrossRef](#)] [[PubMed](#)]
79. Chai, Y.; Paul, A.; Rettig, M.; Wilson, W.D.; Boykin, D.W. Design and synthesis of heterocyclic cations for specific DNA recognition: From AT-rich to mixed-base-pair DNA sequences. *J. Org. Chem.* **2014**, *79*, 852–866. [[CrossRef](#)] [[PubMed](#)]
80. 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. [[CrossRef](#)] [[PubMed](#)]
81. Guo, P.; Farahat, A.A.; Paul, A.; Harika, N.K.; Boykin, D.W.; Wilson, W.D. Compound shape effects in minor groove binding affinity and specificity for mixed sequence DNA. *J. Am. Chem. Soc.* **2018**, *140*, 14761–14769. [[CrossRef](#)] [[PubMed](#)]
82. Antony-Debré, I.; Paul, A.; Leite, J.; Mitchell, K.; Kim, H.; Carvajal, L.A.; Todorova, T.I.; Huang, K.; Kumar, A.; Farahat, A.A.; et al. Pharmacological inhibition of the transcription factor PU.1 in leukemia. *J. Clin. Investig.* **2017**, *127*, 4297–4313. [[CrossRef](#)] [[PubMed](#)]
83. Wang, S.; Aston, K.; Koeller, K.J.; Harris, G.D., Jr.; Rath, N.P.; Bashkin, J.K.; Wilson, W.D. Modulation of DNA-polyamide interaction by β -alanine substitutions: A study of positional effects on binding affinity, kinetics and thermodynamics. *Org. Biomol. Chem.* **2014**, *12*, 7523–7536. [[CrossRef](#)] [[PubMed](#)]
84. Paul, A.; Nanjunda, R.; Kumar, A.; Laughlin, S.; Nhili, R.; Depauw, S.; Deuser, S.; Chai, Y.; Chaudhary, A.S.; David-Cordonnier, M.-H.; et al. Mixed up minor groove binders: Convincing A–T specific compounds to recognize a G–C base pair. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 4927–4932. [[CrossRef](#)] [[PubMed](#)]
85. Harika, N.K.; Paul, A.; Stroeve, E.; Chai, Y.; Boykin, D.W.; Germann, M.W.; Wilson, W.D. Imino proton NMR guides the reprogramming of A•T specific minor groove binders for mixed base pair recognition. *Nucleic Acids Res.* **2016**, *44*, 4519–4527. [[CrossRef](#)] [[PubMed](#)]
86. Paul, A.; Chai, Y.; Boykin, D.W.; Wilson, W.D. Understanding mixed sequence DNA recognition by novel designed compounds: The kinetic and thermodynamic behavior of azabenzimidazole diamidines. *Biochemistry* **2015**, *54*, 577–587. [[CrossRef](#)] [[PubMed](#)]
87. 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. [[CrossRef](#)] [[PubMed](#)]

88. Zhou, T.; Yang, L.; Lu, Y.; Dror, I.; Machado, A.C.D.; Ghane, T.; Felice, R.D.; 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. [[CrossRef](#)] [[PubMed](#)]
89. Bishop, E.P.; Rohs, R.; Parker, S.C.; West, S.M.; Liu, P.; Mann, R.S.; Honig, B.; Tullius, T.D. A map of minor groove shape and electrostatic potential from hydroxyl radical cleavage patterns of DNA. *ACS Chem. Biol.* **2011**, *6*, 1314–1320. [[CrossRef](#)] [[PubMed](#)]
90. Rohs, R.; West, S.M.; Sosinsky, A.; Liu, P.; Mann, R.S.; Honig, B. The role of DNA shape in protein–DNA recognition. *Nature* **2009**, *461*, 1248–1254. [[CrossRef](#)] [[PubMed](#)]
91. Azad, R.N.; Zafiroopoulos, D.; Ober, D.; Jiang, Y.; Chiu, T.-P.; Sagendorf, J.M.; Rohs, R.; Tullius, T.D. Experimental maps of DNA structure at nucleotide resolution distinguish intrinsic from protein-induced DNA deformations. *Nucleic Acids Res.* **2018**, *46*, 2636–2647. [[CrossRef](#)] [[PubMed](#)]
92. Hehre, W.; Ohlinger, S. *Spartan 16 Tutorial and User's Guide*; Wavefunction, Inc.: Irvine, CA, USA, 2016.
93. 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.; et al. *Amber 14*; University of California: San Francisco, CA, USA, 2014.
94. Athri, P.; Wilson, W.D. Molecular dynamics of water-mediated interactions of a linear benzimidazole–biphenyl diamidine with the DNA minor groove. *J. Am. Chem. Soc.* **2009**, *131*, 7618–7625. [[CrossRef](#)] [[PubMed](#)]
95. Harika, N.K.; Wilson, W.D. Bound compound, interfacial water, and phenyl ring rotation dynamics of a compound in the DNA minor groove. *Biochemistry* **2018**, *57*, 5050–5057. [[CrossRef](#)] [[PubMed](#)]
96. Taniguchi, J.; Pandian, G.N.; Hidaka, T.; Hashiya, K.; Bando, T.; Kim, K.; Sugiyama, H. A synthetic DNA-binding inhibitor of SOX2 guides human induced pluripotent stem cells to differentiate into mesoderm. *Nucleic Acids Res.* **2017**, *45*, 9219–9228. [[CrossRef](#)] [[PubMed](#)]
97. Huang, Y.; Sitwala, K.; Bronstein, J.; Sanders, D.; Dandekar, M.; Collins, C.; Robertson, G.; MacDonald, J.; Cezard, T.; Bilenky, M.; et al. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. *Blood* **2012**, *119*, 388–398. [[CrossRef](#)] [[PubMed](#)]
98. Nakagawa, S.; Gisselbrecht, S.S.; Rogers, J.M.; Hartl, D.L.; Bulyk, M.L. DNA-binding specificity changes in the evolution of forkhead transcription factors. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12349–12354. [[CrossRef](#)] [[PubMed](#)]
99. Munde, M.; Poon, G.M.K.; Wilson, W.D. Probing the electrostatics and pharmacological modulation of sequence-specific binding by the DNA-binding domain of the ETS family transcription factor PU.1: A binding affinity and kinetics investigation. *J. Mol. Biol.* **2013**, *425*, 1655–1669. [[CrossRef](#)] [[PubMed](#)]
100. Guo, P.; Paul, A.; Kumar, A.; Harika, N.K.; Wang, S.; Farahat, A.A.; Boykin, D.W.; Wilson, W.D. A modular design for minor groove binding and recognition of mixed base pair sequences of DNA. *Chem. Commun.* **2017**, *53*, 10406–10409. [[CrossRef](#)] [[PubMed](#)]
101. Shi, X.; Chaires, J.B. Sequence- and structural-selective nucleic acid binding revealed by the melting of mixtures. *Nucleic Acids Res.* **2006**, *34*, e14. [[CrossRef](#)] [[PubMed](#)]
102. Munde, M.; Kumar, A.; Peixoto, P.; Depauw, S.; Ismail, M.A.; Farahat, A.A.; Paul, A.; Say, M.V.; David-Cordonnier, M.-H.; Boykin, D.W.; et al. The unusual monomer recognition of guanine-containing mixed sequence DNA by a dithiophene heterocyclic diamidine. *Biochemistry* **2014**, *53*, 1218–1227. [[CrossRef](#)] [[PubMed](#)]

Sample Availability: Samples of the all diamidines compounds, have mentioned in this review article, are available from the authors.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).