Bioorg. Med. Chem. 68 (2022) 116861

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Drug design and DNA structural research inspired by the Neidle laboratory: DNA minor groove binding and transcription factor inhibition by thiophene diamidines

Edwin N. Ogbonna [1], Ananya Paul [1], J. Ross Terrell, Ziyuan Fang, Cen Chen, Gregory M.K. Poon, David W Boykin, W. David Wilson [*]

Department of Chemistry and Center for Diagnostics and Therapeutics, Georgia State University, Atlanta, GA 30303-3083, USA

ARTICLE INFO

Keywords:
DNA minor groove
X-ray crystallography
Diamidine
Surface plasmon resonance (SPR)
Transcription factor inhibition
Curvature
DNA minor groove parameters

ABSTRACT

The understanding of sequence-specific DNA minor groove interactions has recently made major steps forward and as a result, the goal of development of compounds that target the minor groove is an active research area. In an effort to develop biologically active minor groove agents, we are preparing and exploring the DNA interactions of diverse diamidine derivatives with a 5′-GAATTC-3′ binding site using a powerful array of methods including, biosensor-SPR methods, and X-ray crystallography. The benzimidazole-thiophene module provides an excellent minor groove recognition component. A central thiophene in a benzimidazole-thiophene-phenyl aromatic system provides essentially optimum curvature for matching the shape of the minor groove. Comparison of that structure to one with the benzimidazole replaced with an indole shows that the two structures are very similar, but have some interesting and important differences in electrostatic potential maps, the DNA minor groove binding structure based on x-ray crystallographic analysis, and inhibition of the major groove binding PU.1 transcription factor complex. The binding K_D for both compounds is under 10 nM and both form amidine H-bonds to DNA bases. They both have bifurcated H-bonds from the benzimidazole or indole groups to bases at the center of the -AATT- binding site. Analysis of the comparative results provides an excellent understanding of how thiophene compounds recognize the minor groove and can act as transcription factor inhibitors.

1. Introduction

The results described in this paper follow directly from the ground breaking body of studies on DNA minor groove binders for over a period of 40 years by Professor Stephen Neidle and coworkers. Although Dr. Neidle conducted many different experiments on the structure, function, and ligand interactions of DNA over this period,[1–7] this report follows his extensive studies on double-helical DNA complexes with aromatic diamidines. This class of compounds is of fundamental importance for their major influential impact on both therapeutic and biotechnology areas.[2,3,6–16] The structural results from the Neidle laboratory have laid the foundation for most of the other investigations of DNA minor groove complexes.[17–19] Professor Neidle has more recently performed a similar role in the DNA G-quadruplex fields while also conducting many important research studies on the design and testing of DNA G-quadruplex targeted therapeutics.[20–23] As noted, however, this paper will build on his insights into DNA duplex minor groove complexes.

The work in the Neidle group began in the early days of structural studies of DNA small molecule complexes. Their first X-ray structure on an aromatic diamidine was with the therapeutically important compound, berenil (Fig. 1A).[3,24–26] The DNA sequence d $(CGCGAATTCGCG)_2$, which was used in the original netropsin structure, was also used in the initial studies with diamidines. They found that, as with netropsin, berenil bound in the AATT region of the DNA. It displaced the minor groove hydration structure and formed diamidine-base H-bonds to help stabilize the complex. The berenil complex with two diamidines was also stabilized by charge interactions with the anionic electrostatic field in the DNA minor groove as well as by the van der Waals interactions from stacking with the walls of the groove.[3,24–26]

The study of berenil was soon followed with structures for the clinically useful diamidine anti-parasitic drug, pentamidine, and its structural analog, propamidine. (Fig. 1A).[27–29] These important compounds

* Corresponding author.
E-mail address: wdw@gsu.edu (W.D. Wilson).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.bmc.2022.116861
Received 29 April 2022; Received in revised form 23 May 2022; Accepted 26 May 2022
Available online 31 May 2022

are also bound in the -AATT region of the DNA oligomer duplex with many of the same features as berenil. However, one big difference in these structures is that berenil requires a bound water molecule to complete the interaction of the compound with the bases at the floor of the minor groove.[3] The findings with these compounds illustrated two key points about minor groove complexes for the first time: (i) compound curvature and its match to the minor groove curvature is a critical feature for strong complex formation, and (ii) some compounds without the optimum curvature can capture a water molecule to complete the complex and give much stronger binding.[2,7] The direct involvement of water as a critical part of the minor groove complex of berenil was an important insight from the structural results.

The Neidle, Wilson, and Boykin groups began a very enjoyable and productive collaboration in the early 1990 s on heterocyclic diamidines that were synthesized and their DNA complexes biophysics studied at Georgia State University with numerous crystal structures from the Neidle laboratory. These compounds were designed as potential new antiparasitic therapeutics.[2,6–10,30–32] Around this time, Dr. Wilson did a summer sabbatical in England with Professor Neidle in the early 1990s. After a day's work, many enjoyable evenings were spent in a local pub with productive discussions around the Neidle group. An important compound examined in the collaborative research in this period, the aromatic diamidine, furamidine, DB75 (Fig. 1A),[3,9,10,31,32] has been found to have very promising anti-parasitic clinical activity in humans. The prodrug of furamidine has, for example, progressed to phase III clinical trials against the trypanosome organism that causes sleeping sickness with excellent results.[10] It is a highly fluorescent compound and can be seen in trypanosomes from the blood of infected mice after injection.[32] It quickly leads to the destruction of their kinetoplast DNA as an essential part of its mechanism of action.[9] Furamidine has also shown very promising therapeutic activity against several other diseases[12,13,13,33–34] by targeting DNA or structural RNAs. Several derivatives of DB75 have also shown activity against other diseases.[11,12,14,13,33]

To better understand the mechanism of therapeutic action of furamidine and help to design better analogs, the Neidle group determined the structure of furamidine and several of its alkylamidine derivatives bound to the same DNA as with the other compounds described above.[2,3,6,7] The chemical structure of DB75 is shown for a minor groove binding reference in Fig. 1A and the X-ray structure in Fig. 1B.[3] As expected from the close relationship of the compounds to the diamidines berenil and pentamidine, DB75 binds deeply in the DNA minor groove in the -AATT- sequence. Both amidine groups form H-bonds to AT bp at the floor of the groove and the entire molecule is in van der Waals contact with the walls of the minor groove. Rotations about the bonds connecting the amidine, phenyl, and furan groups in DB75 allow the compound to track along the minor groove curvature. The inner-facing nitrogen of both amidine groups are hydrogen-bonded to the O2 atom of

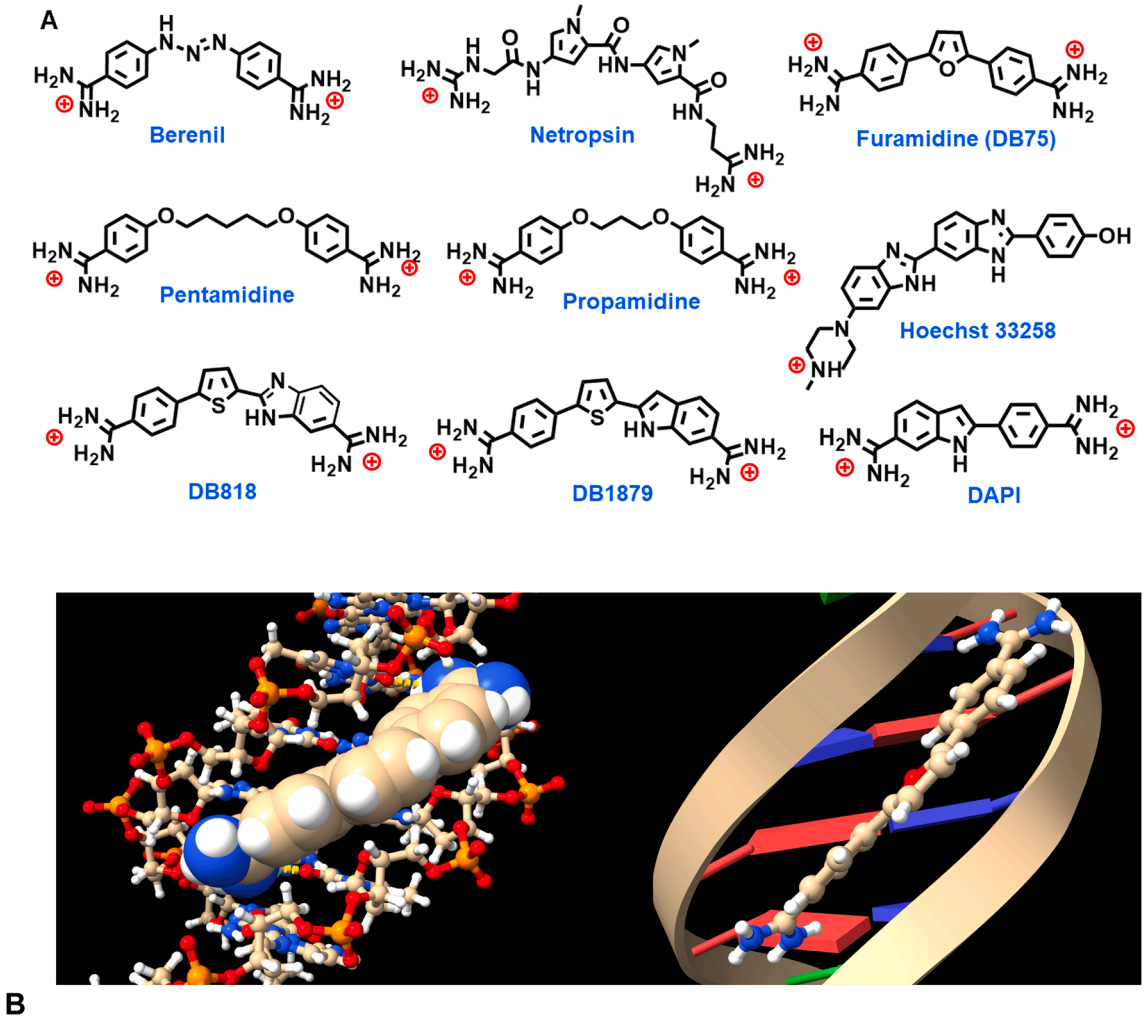

Fig. 1. (A) Chemical structures of classical AT-specific DNA minor groove binders; **(B)** X-ray structure from the Neidle group of the DB75 complex with d (CGCGAATTCGCG) (3), 227d.pdb: Left Side, DNA in ball and spoke, DB75 in space fill model; Right Side, DNA backbone in ribbon, base pairs in slab view and DB75 in ball and spoke. Colors: Left Side - H white, C tan, N blue, O red, P orange; in the Right Side Model: helix ribbons, tan, A red, T blue.

T bases at 3.1–3.2 Å. As can be seen in the model, the T bases are four base pairs apart and on opposite strands. There are close van der Waals contacts between the DB75 furan and phenyls and DNA bases and sugars. The combination of H-bonds, van der Waals contacts, and electrostatic interactions account for the strong interactions of furamidine and analogs with the AATT site.[3] Given the resemblance to berenil and propamidine, the structural similarity and binding strength are reasonable. The alkyl amidine derivatives of furamidine have generally slightly stronger binding, which is probably due to favorable positioning of the alkyl groups in the minor groove, extra water release and van der Waals contacts.[35] All of the compounds have interactions with water molecules at the amidine terminal groups in the minor groove and this is an important stabilizing interaction for these complexes. Both DB75 and berenil produce similar widening of the minor groove of 1–2 A that is most pronounced at the last T of -AATT-.

The Neidle group continued their extensive structural investigation of heterocyclic amidines and established a solid foundation for understanding the interaction of these compounds with the DNA minor groove. In extending these studies to thiophene derivatives, also synthesized by the Boykin group, they determined the structure for the minor groove complex of a very strong binding compound, DB818 (Fig. 1A), that was termed a "super AT binding minor groove agent", based on its similar size to weaker binding compounds.[36]

This thiophene derivative is a single example of the extensive applications of sulfur derivatives in medicinal chemistry.[37] Many sulfur derivatives have been synthesized with activities across a broad array of diseases.[37–41] -The thiophene sulfur-*N*-alkylbenzimidazole nitrogen interaction has been used as the basis for recognition of many mixed sequence DNA samples.[42–44]

In a continued investigation of thiophene compounds, an indole analog was synthesized,[11] and the DNA binding of both benzimidazole and indole were compared as well as their ability to inhibit the binding of the transcription factor PU.1 to a DNA promoter sequence. PU.1 is a member of the ETS family of transcription factors and is involved in the development of acute myeloid leukemia in many people.[45,46] The development of drugs against AML is a high priority but has proven difficult. Our approach is to target the minor groove of PU.1 promoter sequence and cause dissociation of PU.1 from the major groove. Initial results with AT-specific minor-groove binding diamidines are very promising.[14] The results of structural, binding, and PU.1 inhibition studies of the thiophene benzimidazole and indole are presented here as part of our studies to better understand minor groove binding agents and to develop new types of drugs against AML. The fundamental goal of these studies is to provide new design ideas for improved diamidine inhibition of PU.1 and for inhibition of other transcription factors involved in disease development.

1.1. Molecular Curvature:

Molecular curvature is a key feature for minor groove recognition and couples with compound molecular stacking surface, H-bonding to minor groove bases, and charged groups to establish the energetics of binding and recognition.[47] Since there is no systematic method to evaluate the molecular curvature of minor groove binders, a graphical approach for the determination of comparative molecular curvature values has been developed and is applied here to DB75, DB818 and DB1879. In this method, compounds are first energy minimized in the SPARTAN software package. Next, optimization calculations are performed for the compounds using the DFT/B3LYP theory with the 6-31+G* basis set in the SPARTAN software package.[48] The compounds are then compared in the PowerPoint graphics package. The first step is to define a reference circle that passes through both amidine carbons. The reference circle that has a radius to allow it to pass as closely as possible through the center of each molecular unit of the entire molecule and the two amidine carbons is then selected. This is shown with the three compounds in Fig. 2. Two lines are next connected from the amidine carbons to the circle point at the center of the molecule. The angle between these two lines defines a relative curvature value for comparison of each molecule (Fig. 2). Analysis of a range of minor groove binding compounds by this method indicates a value of 140-145° is optimum for compounds binding strongly in the DNA minor groove. As can be seen, the curvature of DB818 and DB1879 falls in this range while DB75 is too curved to make optimum contacts with the groove and it binds approximately 10 times weaker than the thiophenes. The two amidines of DB75 contact the groove surface, but the center of DB75 loses direct contact with the minor groove. Due to its lower flexibility and curved structure, the furan and adjacent atoms of DB75 are pushed away from the floor of the groove. This relative comparison thus provides a useful numerical value to use to evaluate possible new compounds.

2. Experimental

2.1. Materials and methods

2.1.1. Recombinant protein Expression and Purification:

The gene coding for the ETS domain of human PU.1 (residues 165–270) was cloned into the pET-28b plasmid and transformed into BL21 DE3 pLysS cells (Novagen). Cells were grown to an OD_{600nm} of 0.6 at 37 °C in LB media. Expression of protein was induced with the addition of IPTG to a final concentration of 0.5 mM at 22 °C for 16 h and the cells pelleted by centrifugation. The cell pellet was resuspended in buffer containing 10 mM HEPES, 500 mM NaCl, at pH 7.4. Prior to lysis the suspension was treated with 1 mM PMSF and cells lysed by sonication. Cell debris was pelleted by centrifugation and clarified lysate loaded onto a 1 mL HiTrap SP HP column (Cytiva). Protein was eluted against a 2 M NaCl gradient and fractions containing isolated PU.1 were pooled. Buffer exchange and additional purification by gel filtration were carried out on a Superdex 75 HiLoad 16/60 column (GE Life Sciences) equilibrated with buffer containing 10 mM HEPES, 150 mM NaCl at a pH of 7.4. Pooled fractions from gel filtration were concentrated to 500 μM prior to use in experiments.

2.2. Compound Properties:

The free compounds were built, and their equilibrium geometries and electrostatic potentials calculated in the Spartan 20 software package. The calculations were done by the density functional method at the 6-31+G* level. The electrostatic potential maps were compared at the same potential level in the Spartan software.[48]

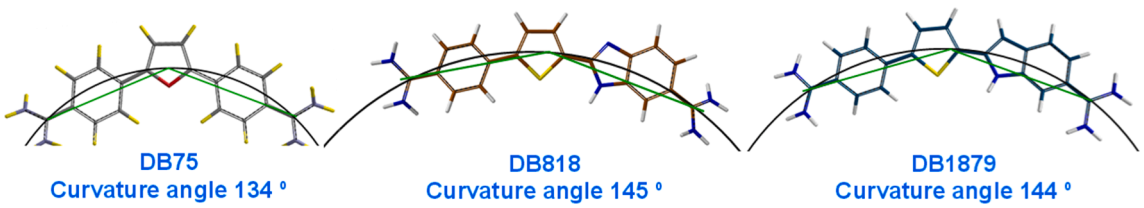

Fig. 2. By using the procedure described in the text, the relative molecular curvature values for these compounds were determined.

2.3. Crystallization and data collection

The oligonucleotide duplex d(5′-CGCGAATTCGCG-3′)$_2$ (HPLC-Purified, IDT) was annealed at 85 °C for 6 mins in 20 mM NaCl with 1.5 stoichiometric equivalents of compound (10 mM stock, H_2O) prior to crystallization. All crystals were grown by vapor diffusion in 3 μL hanging drops (1:1) at 298 K in 24-well VDX plates (Hampton Research) in drops containing the 24-conditions of the Nucleic Acid Mini-screen (Hampton Research) against wells containing 600 μL of a 35% solution of (±)-2-methyl-2,4-pentanediol (MPD). Crystals containing DB818 were grown using condition 7 in a drop comprised of 20 mM $MgCl_2{\cdot}6H_2O$, 80 mM NaCl, 0.5 mM double-stranded DNA, 0.75 mM DB818, 10% v/v (±)-2-methyl-2,4-pentanediol (MPD), 12 mM spermine tetrahydrate, 40 mM sodium cacodylate trihydrate buffer at pH 6.0. Crystals containing DB1879 were grown using condition (1) comprised of 20 mM $MgCl_2{\cdot}6H_2O$, 0.5 mM double-stranded DNA, 0.75 mM DB1879, 10% v/v MPD, 20 mM hexamine cobalt (III) chloride, 40 mM sodium cacodylate trihydrate buffer at pH 5.5. Rod-shaped crystal growth was observed within three weeks with both complexes. The first *d*(5′-CGCGAATTCGCG-3′)$_2$-DB818 crystal was colorless, while under the second set of conditions the *d*(5′-CGCGAATTCGCG-3′)$_2$-DB1879 crystal had a bright-yellow color. All crystals were transferred to appropriately sized cryo-loops and flash-frozen in liquid nitrogen prior to data collection.

X-ray diffraction data sets were collected at the Lawrence Berkeley National Laboratory Advanced Light Source on beamline 5.0.1 with a Dectris Pilatus3S 6 M detector at 100 K, at a wavelength of 0.97 Å. Data sets for both crystals contained 1440 frames collected over 360° with frames exposed for 0.25 s at 0.25 deg per frame. Crystallographic indexing, integration, and scaling were carried out using the HKL2000 software package.[49]

2.4. Structure solution and refinement

Structures of d(5′-CGCGAATTCGCG-3′)$_2$ with both DB818 and DB1879 were solved by molecular replacement using maximum-likelihood search procedures in the Collaborative Computative Project No.4 (CCP4i) software suite.[50] Structure solution and refinement were carried out using an established DNA model (PDB entry 1BNA). Refinement strategies included rigid body and restrained refinements. Final statistics for data collection and refinement can be found in Table S1, SI. The addition of ligands and water in Coot[51] followed by additional refinements to the models brought the final R-values to 22.0% R_{work}, 23.0% R_{free} for the DB818 structure, and 22.1% R_{work}, 26.4% R_{free} for the DB1879 structure. See Table S1, SI for complete crystallographic statistics. The electron density map showed substantial electron density coverage of ligands in the minor groove of both structures ($2mF_o$-DF_c maps shown in Figs. S5, S6, SI). The atomic structure and coordinate factors for both DNA complexes have been deposited to the RCSB Protein and Nucleic Acid Data Bank; DNA-DB818 with an accession ID 7KU4, and DNA-DB1879 with an accession ID 7KWK. All the figures containing crystal structures and models were generated in Chimera X software.[52]

Structure solution and refinement were carried out using an established DNA model (PDB entry 1BNA). Refinement strategies included rigid body and restrained refinements. Final statistics for data collection and refinement can be found in Table S1, SI. The addition of water and further refinements to the models brought the final R-values to 22.0% R_{work}, 23.0% R_{free} for the DB818 structure, and 22.1% R_{work}, 26.4% R_{free} for the DB1879 structure. The electron density map (MTZ) showed substantial electron density coverage in the minor groove of both structures. The atomic structure and coordinate factors for both DNA complexes have been deposited to the RCSB Protein and Nucleic Acid Data Bank; DNA-DB818 with an accession ID 7KU4, and DNA-DB1879 with an accession ID 7KWK. All the figures containing crystal structures and models were generated in Coot[51] and Chimera X software.[52]

2.5. Biosensor-SPR assays for compound binding affinity and PU.1-DNA complex inhibition by compound

SPR experiments were carried out with a four-channel-based Biacore T200 optical biosensor system. A streptavidin-derivatized CM5 sensor chip was created by covalently linked of streptavidin with active ester functionalized carboxymethylated dextran. Carboxymethylated dextran was covalently attached to the gold surface of the chip. To activate the streptavidin-derivatized surface, several 180 *sec* injections of 1 M NaCl in 50 mM NaOH solution mixture (activation buffer) were injected on the surface of the chip, followed by extensive washing with HBS buffer [10 mM HEPES, 150 mM NaCl, 3 mM EDTA, and 0.05% P20 (pH 7.4)]. (54,55) To determine the binding constant of the ligand with -AATT- and λB DNA sequences, 20 nM 5′-biotin-labeled hairpin DNA, 5′-CGAATTCGTCTCCGAATTCG-3′ (-AATT-) and 5′- CCAAA-TAAAAGGAAGTGAAACCAAGCTCTCTTGGTTTCACTTCCTTT-TATTTGG-3′ (λB DNA), sequences were prepared in HBS buffer and immobilized on the flow cell surface.[53,54] To determine the binding constant of the ligand with -AATT- and λB DNA sequences, a series of different compound concentrations (from 1 nM to 1 μM) were injected over the DNA sensor chip at a flow rate of 100 μL/min and the SPR response was followed for 3 min. This was followed by buffer flow to monitor the compound dissociation from the DNA complex and yielded a complete sensorgram for each compound concentration. After each cycle, the sensor chip surface was regenerated with a 10 mM glycine solution at pH 2.5 for 30 s followed by multiple buffer injections to yield a stable baseline for the following cycles. Kinetic analyses were performed by fitting the SPR sensorgram set by using a standard 1:1 kinetic model. The steady-state data analysis was also performed by using previously described method, where response from the blank cell (cell 1) was subtracted from the response in each flow cell containing DNA to give a signal RU_{obs}. RU_{obs} is directly related to the amount of bound ligand on the DNA immobilized surface. The expected maximum response, RU_{max} per bound ligand in the steady-state region, was determined from the molecular weight of the DNA, the ligand molecular weight, and the refractive index gradient ratio of the ligand and DNA. KaleidaGraph 4.0 software was used to plot RU_{obs} versus free ligand concentration (C_{free}). The equilibrium binding constants, K_1, were determined with a one-site binding model. In this model, $r = (RU_{obs}/RU_{max})$ represents the moles of bound compound/mol of DNA hairpin duplex, and K_1 is macroscopic binding constant.

$$r = K1 * C_{free}\%1 + K1 * C_{free} \quad (1)$$

Kinetic analysis was achieved by globally fitting the ligand-binding sensorgrams using a standard 1:1 kinetic model with incorporated mass transport-limited binding parameters.

SPR experiments were performed at 25 °C in filtered and degassed 25 mM Na_2HPO_4 (pH 7.4) containing 0.05% P_{20}, 400 mM NaCl, and 1 mM EDTA. For protein inhibition studies by the ligands, SPR experiments were performed at 25 °C in filtered and degassed 25 mM Na_2HPO_4 (pH 7.4) containing 0.05% P_{20}, 400 mM NaCl, and 1 mM EDTA buffer. A 100 nM constant concentration of PU.1 protein was injected on the immobilized λB DNA surface to acheive the saturation of λB DNA binding site. The increasing concentrations of the ligands were added to the fixed protein solution.[55,56] Then the mixtures of these solutions were injected to the λB DNA surface. The decrease in the protein binding signals in the presence of different concentrations of ligands was plotted against compound concentration to determine the inhibition by each ligand. The midpoint of the normalized inhibition plot was used as the IC_{50} of the ligand.[55,56]

3. Results

3.1. Compound Properties:

Although the compounds are very similar in equilibrium structure

and differ only at a single position, from a -N- in the benzimidazole to a —CH— in the indole, the benzimidazole binds about 4–5 times more strongly (see below). The electrostatic potential maps in Fig. 3 provide an explanation for that difference. At the thiophene-phenyl-amidine ends of the two molecules, the maps are not significantly different. At the benzimidazole/indole-amidine ends, however, there is an important difference in the two compounds. When compared to the indole, DB818 is more positive on the inner curve of the molecule while it is more negative on the outside edge. When considering binding into the negative electrostatic potential of the minor groove, this provides a favorable component to the binding energetics of DB818 compared to DB1879 that can, at least partially, explain the binding difference.

3.2. Biosensor surface plasmon resonance (SPR) determination of binding affinities and Kinetics:

Binding equilibrium constants for DB818 and DB1879 were evaluated to characterize their interaction with an -AATT- site. This was done by SPR methods with chip-immobilized streptavidin capture of biotin-labeled hairpin DNA on a sensor chip surface. The two thiophene compounds were injected over the chip surface in buffer solution at different salt concentrations (from 100 mM to 550 mM NaCl). Examples of sensorgrams are shown for reference in Fig. 4 and equilibrium constants obtained by global fitting to a one-site model are listed in Table S2, SI. Given the close similarity of the two compound structures, it is somewhat surprising that the benzimidazole, DB818, has measurably different binding kinetics and affinities than the indole. DB818 binds strongly with -AATT- and global kinetics fitting defined a single binding site with $K_A = 6.6 \times 10^8\ M^{-1}$ ($K_D = 1.5$ nM) at 0.1 M NaCl (Fig. 5, Table S2). The strong binding of DB818 is the result of the rapid association ($k_a = \sim 5.8 \times 10^6\ M^{-1}s^{-1}$) and a comparatively slow dissociation rate constant ($k_d = 8.9 \times 10^{-3}\ s^{-1}$. However, the indole derivative, DB1879, has a binding affinity ($K_D = 6.5$ nM) about four times weaker than DB818, for -AATT- binding sequence. Surprisingly, it has been observed that the rate of association of the DB1879-AATT binding complex decreased about ten times ($k_a = \sim 4.5 \times 10^5\ M^{-1}s^{-1}$) compared to the DB818-AATT complex, however, the rate of dissociation of DB1879, from DB1879-AATT complex is about 3–4 times slower ($k_d = 2.9 \times 10^{-3}\ s^{-1}$ than DB818-AATT complex. This complementary behavior in binding kinetics of DB1879 leads to four-fold weaker binding affinity for DB1879 compared to DB818.

To evaluate the effect of ionic strength on DB818 and DB1879 binding affinities with the -AATT- sequence, SPR experiments were carried out from 100 to 550 mM NaCl concentrations at 25 °C (Fig. 5, S1-S4, SI). Due to the dicationic nature of DB818 and DB1879, the effects of ionic strength of the solution play a crucial role in DNA – ligand complex formation. The shapes of SPR sensorgrams at different salt concentrations indicate that both DB818 and DB1879 (Fig. S1-S4, SI) vary on and off rates with changing salt concentrations while binding to -AATT- sequence. The equilibrium binding constants (K_A) obtained either by global kinetic fits at low salt concentrations or by steady-state fits at higher salt concentrations are collected in Fig. 4 and Figs. S1-S4, Table S2, SI. Both theoretical and experimental results suggest that the logarithm of the equilibrium binding constant K_A is a linear function of the logarithm of NaCl concentration for many organic cations binding to DNA. For a typical DNA – cation complex, the equilibrium binding constant values decrease as the salt concentration increases with a slope that depends on the compound charge.[57–59] As seen in Fig. 5 the log(K_A) versus log[Na^+] plots for DB818 and DB1879 are linear with a slope of 1.7 and 1.6 respectively. The number of phosphate contacts (Z) between DB818-AATT and DB1879-AATT complexes are predicted to be two for each complex and can be found in experiments from the slope/0.88, where 0.88 is the fraction of phosphate charge shielded by the total associated counterions. For the oligomer -AATT-, the obtained Z is 1.9 for both complexes. These results indicate that dicationic DB818 and DB1879 release $2Na^+$ ions when binding to the AT rich minor groove of DNA.

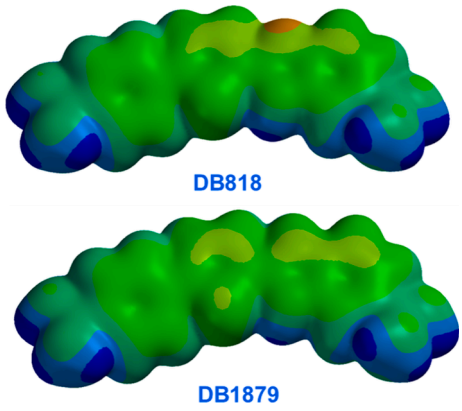

Fig. 3. Energy minimized structures of DB818, and DB1879 at B3LYP/6-31G* (p,d) level of theory. In the electrostatic potential maps, red indicates high electronegativity, and blue indicates electron-deficient/positively charged regions.

3.3. Biosensor-SPR assays for PU.1-DNA complex inhibition by compound:

The inhibition of PU.1 binding in the major groove by minor groove binding heterocyclic diamidines depends on two factors: (i) the binding affinity of the compounds and (ii) secondary allosteric effects that act through DNA from the minor to the major groove. To elucidate the potential of the PU.1 inhibition properties of DB818 and DB1879 at the cognate PU.1 binding site, we immobilized biotin-labeled-λB DNA sequence, a high-affinity cognate sequence for PU.1, on the chip surface. In the absence of ligand, PU.1 formed a 1:1 complex with the immobilized λB motif.[55,56] With λB site immobilized on the chip surface, a 100 nM solution of the PU.1 was injected over the surface with increasing ligand concentrations. Displacement of PU.1 was detected by the reduction in SPR signal as a function of the compound concentrations (Fig. 6). Steady-state signals were used to determine IC_{50} values in Fig. 6. This demonstration indicates that minor-grove binding diamidines DB818 and DB1879 are able to block the ability of the PU.1 protein to bind site specifically in the DNA major groove.

3.4. Compound-DNA Structures:

The conclusion from the biophysical studies described above is that both the benzimidazole, DB818, and the indole derivative, DB1879, bind to AT sequences of four or more base pairs exceptionally well for their size. The results from the comparative analysis presented above are that DB818 binds more strongly and inhibits PU.1 binding to its promoter sequence more strongly than DB1879. The two compounds are similar in equilibrium structure, curvature, and placement of H-bonding groups and this raises the question of whether there are differences in DNA bound structure of the two compounds that could contribute to the binding and structure differences? The structure of DB818 bound to the -AATT- site as with previous minor groove binders has been redetermined and the structure of DB1879 determined at the same site to answer this question.

The re-determined DNA-DB818 structure, 7KU4 (Fig. 7) was compared with the previously published structure by Niedle's laboratory, 1VZK. Comparison of the two structures was reassuring as the DNA-DB818 complex structures determined almost twenty years apart by different laboratories in different countries with different facilities gave virtually identical crystal structures.[36]

All the interactions between the two DB818 samples and the DNA minor groove are the same, and distances between DB818 and the interacting groups on DNA are very similar (Fig. 8). The DB818 model

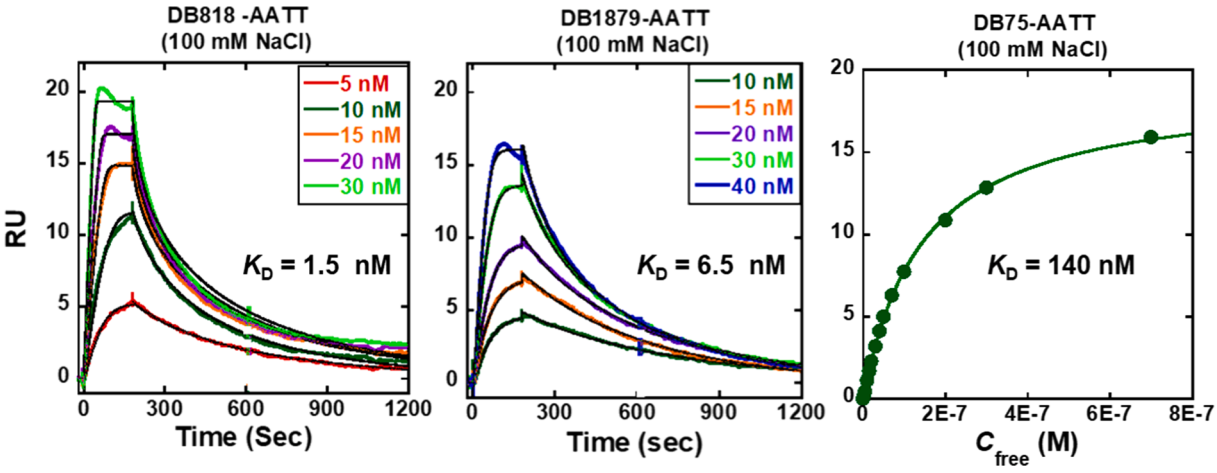

Fig. 4. SPR sensorgrams with kinetics fit for DB818 and DB1879, and steady-state binding plots for DB75 with the -AATT- DNA sequence. The listed binding affinities are an average of two independent experiments carried out with two different sensor chips, and the values are reproducible within 10% experimental errors.

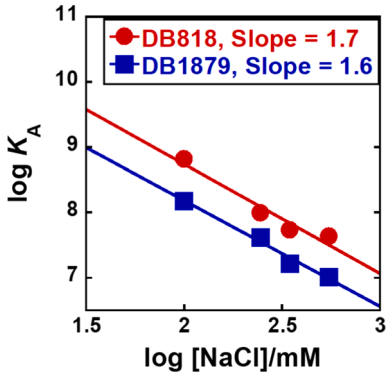

Fig. 5. Salt dependence of K_A for DB818 and DB1879 binding as determined by SPR. The K_A values were obtained by global kinetic (at two lower salt concentrations) and steady-state fits (at two higher salt concentrations).

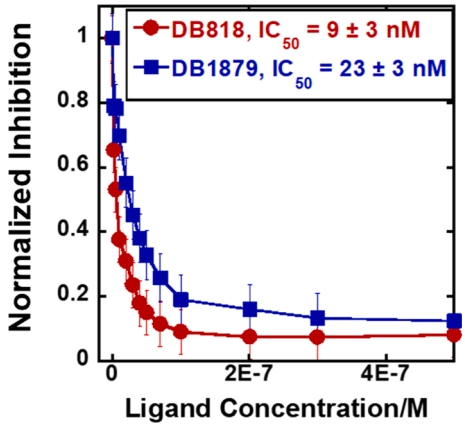

Fig. 6. Normalized PU.1 inhibition resulted from biosensor SPR experiments. The plots represent the amount of PU.1-DNA complex inhibition as a function of the added compound concentration.

for comparison to the new DNA-DB1879 complex, 7KWK (Fig. 9), is thus on firm ground and provides a strong basis for analyzing the structural differences between the DB818 and DB1879 along the -AATT- binding site.

The DNA dodecamer is numbered from 5′ C1 to 3′ G12 in one strand and from 5′ C13 to 3′ G24 in the other.[60] The binding conformation in

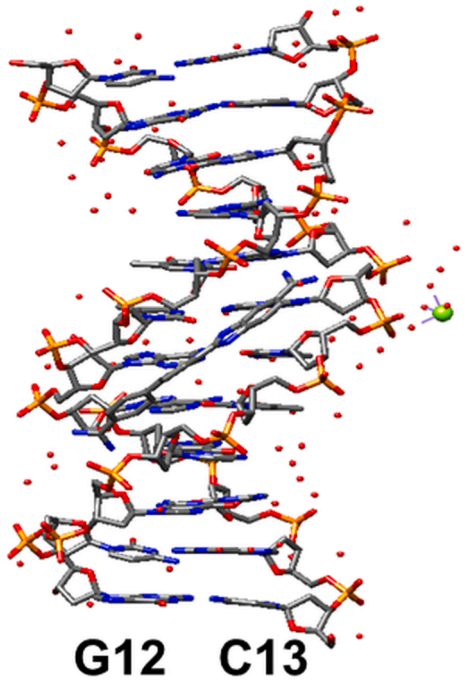

Fig. 7. Structure of 7KU4 at a resolution of 1.6 A. Shown is DB818 bound to the minor groove of the sequence -GAATTC- with surrounding water network of the complex.

the minor groove exerts a twist along the length of DB818 creating torsion angles within the ligand. This results in the amidine groups of the ligand being twisted with respect to each other.[36] The molecular interactions of DB818 in the minor groove involve specific hydrogen bonding in a bifurcated manner between the inner benzimidazole nitrogen atom of the ligand and the base edges in the minor groove (H-bonds of 2.7 Å to O2 of T19 and 2.9 Å to O2 of T7) (Fig. 10). This bifurcated hydrogen bonding by benzimidazole compounds has also been observed in similar AT rich DNA sequences.[36] Also, the amidine groups of the ligand form hydrogen bonds with bases that extended beyond the AATT track (Fig. 10). The amidine group serves as a hydrogen bond donor to the acceptor atoms (O or N) of bases along the minor groove. Amidine group stabilization of DNA complex structures via direct or water-mediated contact have been firmly established.[2,7]

The benzimidazole-amidine formed an H-bond to O2 of T20 while the phenyl-amidine formed a bond with O2 of C9 that also protrudes into

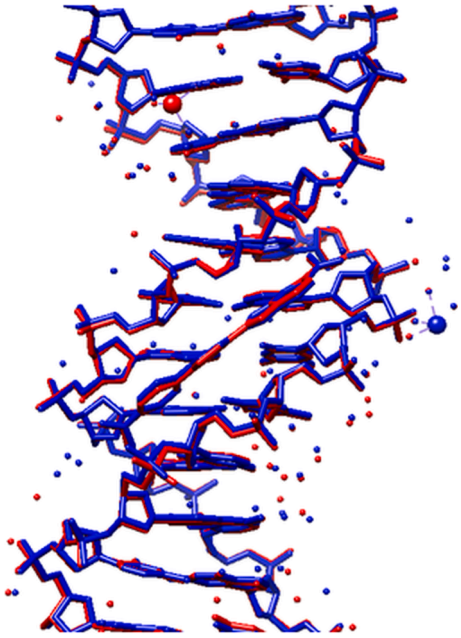

Fig. 8. Overlay of **1VZK** (red) on **7KU4** (blue). **7KU4** matches the structural coordinates of **1VZK** almost exactly. Comparing their intermolecular interactions, benzimidazole nitrogen bonding distances with O2 of T7 and T19 in **7KU4** is 2.8 Å and 2.7 Å respectively (in **1VZK** is 2.9 Å and 2.7 Å respectively).

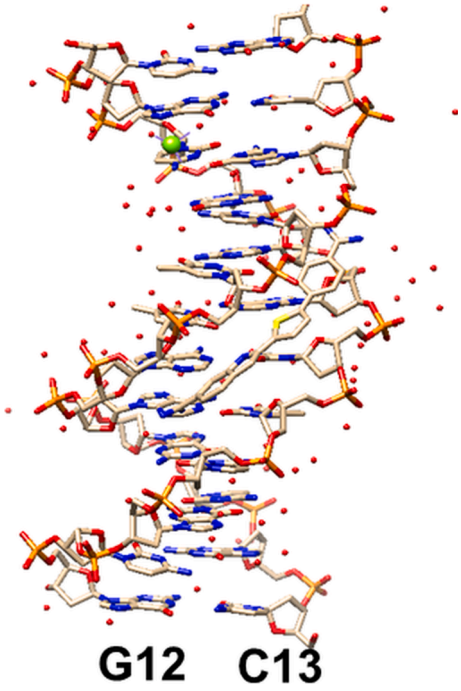

Fig. 9. Structure of 7KWK at a resolution of 1.4 A. Shown is DB1879 bound to the minor groove of -GAATTC- with surrounding water network of the complex. The indole of 7KWK points towards the G12-C13 end of the double-helix.

the minor groove in essentially the same position as O2 of T20. Overall, DB818 binding covers a five-base site in the minor groove. With DB818, the benzimidazole end of the molecule points toward the C1-G24 end of the double helix while the indole end of DB1879 points in the opposite direction. The interactions of the two compounds with DNA binding sites are similar but with different groups due to their opposite orientation (see Figs. 10 and 11).

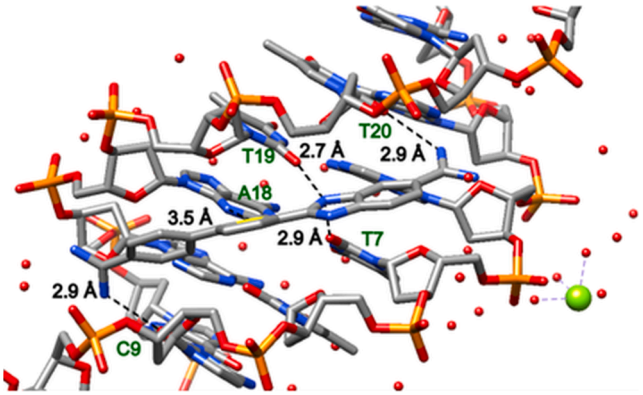

Fig. 10. Significant bonding distances (in black) between DB818 and interacting bases (in green) in the structure of 7KU4. A bifurcated hydrogen bonding between benzimidazole, N with O2 of T7 and T19 is 2.9 Å and 2.7 Å respectively. Thiophene S with N3 of A18 is 3.5 Å. Benzimidazole-amidine N and O2 of T20 is 2.9 Å. Phenyl-amidine N with O2 of C9 is 2.9 Å.

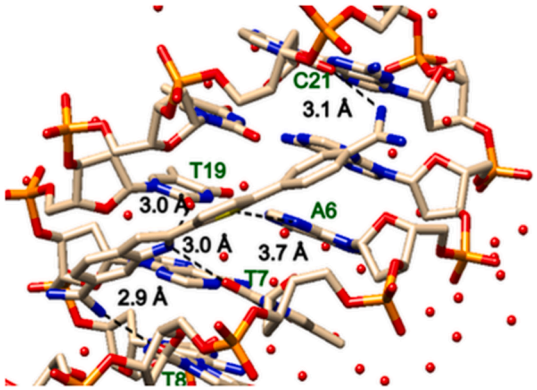

Fig. 11. Significant bonding distances (in black) DB1879 ligand and interacting bases (in green) in the structure of 7KWK. A bifurcated hydrogen bonding between benzimidazole, N with O2s of T7 and T19 is 3.0 Å and 3.0 Å respectively. Thiophene S with N3 of A6 is 3.7 Å. Benzimidazole-amidine N and O2 of T20 is 2.9 Å. Phenyl-amidine N with O2 of C21 is 2.9 Å.

The indole N is 3.0 A from both O2s T19 and T7. The phenyl amidine is 3.1 Å from O2 of C21 while the indole amidine is 2.9 Å from O2 of T8. Both compounds form some important interactions between —CH groups and polar groups on AT base pair edges at the floor of the minor groove (Fig. 11). It is interesting to note that no water-mediated DNA contact was observed for both structures-only direct binding interactions were observed - further confirming how much of an excellent binder both compounds are.[2]

As can be seen from the overlay in Fig. 12 the compounds are slightly offset because, in their opposite orientation, the phenyl-amidine N of DB818 forms an H-bond with O2 of C9 while DB1879 slides to the other end of the binding site to form an H-bond with O2 of C21. The benzimidazole-amidine N of DB818 forms an H-bond with T20 while that of DB1879 forms an H-bond with C21. The two sulfur atoms of the thiophene are displaced by 4.7 Å. The S of DB818 is 3.5 Å from N3 of A6, while the S of DB1879 is 3.7 Å from N3 of A6. As can also be seen in Fig. 12, the two DNA strands and the structured bases form almost a perfect overlay and indicate a similar DNA structure in the two complexes.

The modelled x-ray structures of 7KU4 and 7KWK were further analyzed using the software Chimera[52] to measure the changes in their minor groove width. The minor groove distance comparison between 7KU4, 7KWK, and 5′-GAATTC-3′ is shown (Fig. 13). The minor groove distance measures the phosphate-to-phosphate distance across the DNA

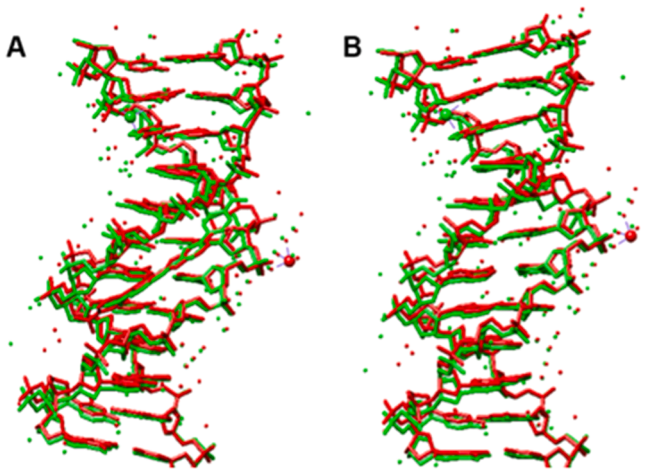

Fig. 12. A) Overlay of **7KU4** (red) on **7KWK** (green). DB818 and DB1879 are only partially aligned. Comparing their intermolecular interactions, benzimidazole nitrogen bonding distances with O2s of T7 and T19 in **7KWK** is 3.0 Å and 3.0 Å respectively (in **7KU4** is 2.8 Å and 2.7 Å respectively). B) Overlay of **7KU4** on **7KWK** without a ligand in the minor groove.

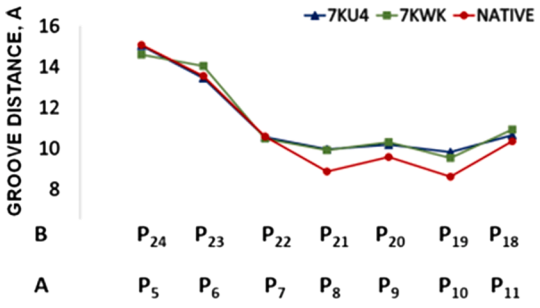

Fig. 13. Minor groove distance comparison of 7KU4 (DNA-DB818) and 7KWK (DNA-DB1879) with native -GAATTC-. The presence of ligand in both complexes slightly widens their respective minor grooves along interacting bases.

strands in a direction perpendicular to the helical axis.[61] The resulting graph shows a distinctive difference between 5′-GAATTC-3′ and the two complexes. The minor groove distance is widest towards the 5′-end of each DNA, and narrowest as the DNA makes a turn.

More detailed analysis of the DNA structure was conducted using Curves+.[62] Curves+ analyzes the conformation of nucleic acid structures, calculating helical parameters (Table S3,S4, SI) and backbone angles (Table S4, S5, SI) of a double helix.[62,63] The calculated helical parameters of each DNA structure were translated into graphical plots to analyze and quantify any structural changes.

Base-pair step parameters measure the translational and rotational relationship within each base pair.[62] Some notable base-pair parameters (shear, stagger, buckle) of all three structures were determined and compared (Fig. 14). Base-step parameters measures the translational and rotational relationship between the two stacked base pairs.[62] Some base-step parameters (shift, rise, roll) of all structures were determined and analyzed as well (Fig. 15).

4. Discussion

The benzimidazole, DB818, and corresponding indole, DB1879, are strong, AT-specific DNA binding agents that have a number of features that make them an ideal pair for understanding and evaluation of DNA minor groove interactions and inhibition of the PU.1 transcription factor. These two compounds have H-bond donor groups pointed into the minor groove at appropriate positions to interact with N3 of A and O2 of T acceptor groups at the floor of the groove. For such minor groove complexes, the compounds must inhibit major groove binding proteins by some type of indirect effects. Such effects could involve induced structural changes in the PU.1-DNA conformation and/or electrostatic effects not conducive to protein binding to DNA. The compounds interact with water molecules at the terminal amidines and these enhance affinity but it is unlikely that they play a direct role in protein inhibition.

To start to answer the question of how these compounds interact with the DNA minor groove, the structure and properties of the free compounds were evaluated with *ab initio* calculations (Fig. 3). With the free compounds, the three linked aromatic systems in both are essentially planar while the two-terminal amidines are twisted 30–40° out of the aromatic plane. The electrostatic potential maps show partial

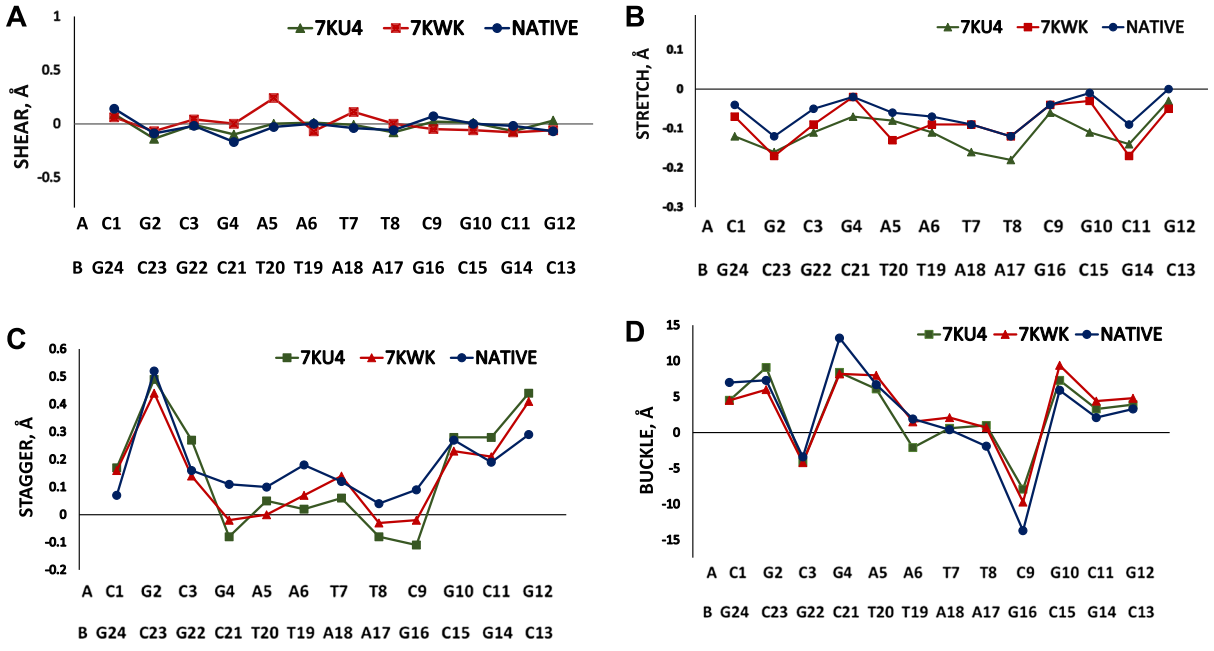

Fig. 14. Base-Pair Parameters. DNA shear, stretch, stagger, and buckle changes at the different DNA base-pair.

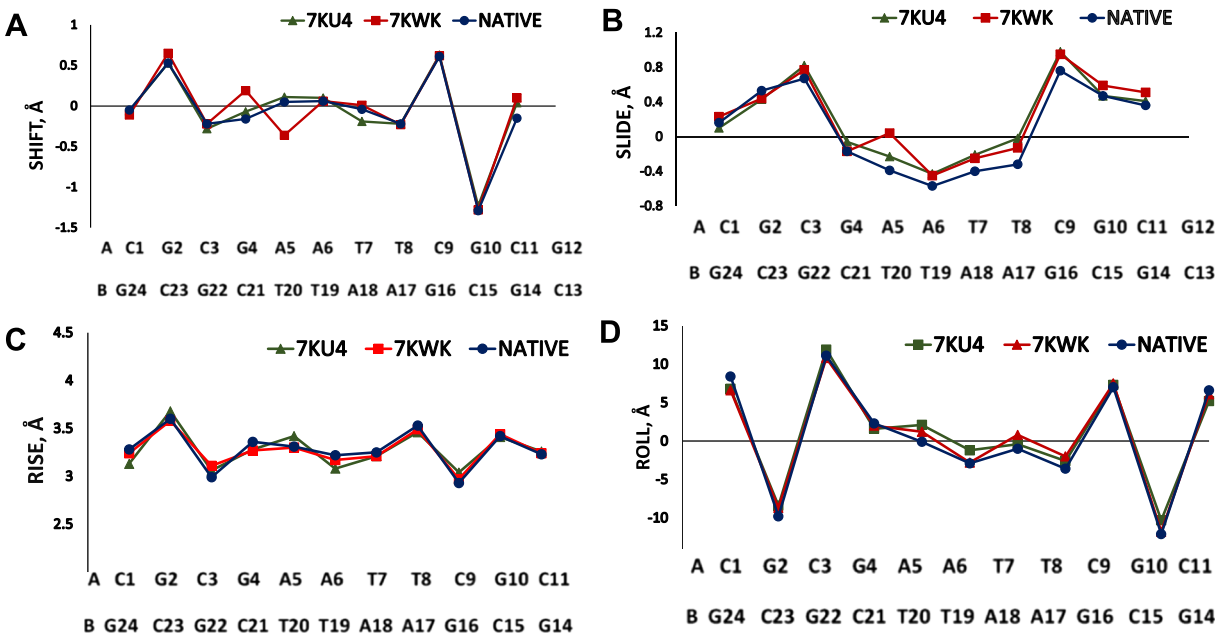

Fig. 15. Base-Step Parameters. DNA shift, slide, rise, and roll changes at the different DNA base-pair stack.

positive charges located on the amidines and the —NH groups of the two compounds. With the indole, the remaining charge is spread fairly evenly over other parts of the compound and is close to zero. With the benzimidazole, there is a partial negative charge on the N at the outer edge of the structure and this causes a more positive charge at the inner edge of the compound and —NH group. Such a distribution is favorable for binding the DNA minor groove and it is thus not surprising that DB818 binds more strongly than DB1879. It should also be noted that the —NH in the benzimidazole is in exchange tautomerization and DB818 could in principle initially bind with the benzimidazole bound with either side of the benzimidazole pointed into the minor groove.

SPR experiments with DNA immobilized on a sensor chip were conducted to evaluate the DNA affinities of the two compounds. The immobilized DNAs had a single, well-studied -AATT- DNA binding site and, as expected, showed strong binding with a one-site binding model. As noted, both compounds bind very strongly and both have K_D values of below 10 nM at 100 mM NaCl. The K_D for DB818 is 4–5 times lower than for DB1879 as expected for stronger binding of the benzimidazole. For strong binding to the minor groove, a compound should have as close a match to the minor groove curvature as possible or should have a more linear shape than will fit the minor groove but have a shape that allows it to capture an interfacial water molecule to complete the minor groove matching curvature. Analysis of the curvature of these compounds using our relative curvature calculation method gave a curvature of 145 +/- 2° which is the ideal value for strong interactions with the minor groove (Fig. 2). DB75 is also shown in Fig. 2 as a much weaker binding compound, $K_D \sim 100$ nM, which has too high a curvature for close fit to the minor groove. Thus DB75 cannot make base contacts and at the same time be in position for base hydrogen bonding to both terminal amidinium groups. DB818, by virtue of its larger concave radius, can form benzimidazole and amidine H-bonds to DNA bases (Figs. 7, 10). This also enables DB818 to cover a five base-pair site compared to the four of DB75 and to have stronger compound-base.

interactions.

Detailed studies of the salt-dependent thermodynamics of DB818 and DB1879 binding can help answer the essential information about compounds that target the minor groove of DNA. Salt-dependent SPR studies show that the rate of association decreases and the rate of dissociation increases with increasing the salt concentration for both DB818-AATT and DB1879-AATT complexes (Fig. 5 and Table S2, SI). A linear dependence has been obtained between the logarithm of the binding constant (log K_A) and the logarithm of salt concentrations (log[Na^+]) (Fig. 5) and it shows that the binding affinity decreases with an increase in the ionic strength or salt concentration which has been expected for dicationic compounds.

To evaluate the DNA binding of these compounds in more detail, x-ray crystal structures of both compounds were determined with the -AATT- DNA and compared to the structure of the unbound DNA. The structure IVZK from the Neidle (Fig. S6, SI) laboratory was used to evaluate the structure of 7KU4. DB818 in 7KU4 makes optimal contacts with the base edges in the minor groove. DB1879 makes similar optimal contacts in the minor groove but with slightly greater hydrogen-bonding distances in all its interactions except for the phenyl-amidine. These slight differences in the hydrogen-bond distances seem in line with the binding properties shown by DB818 and DB1879. Being a better minor groove binder, DB818 shows a higher affinity for the DNA - than its DB1879 counterpart – with its shorter hydrogen-bond distances. The difference in binding orientation of DB1879 does not change the bifurcated hydrogen bonding specificity of the indole with the 5′-GAATTC-3′ sequence. Like benzimidazole, the indole interacts with the same O2 of T7 and T19. However, as previously stated, the change in orientation does alter the specificity of the amidine groups. In the overlay of the 7KU4 and 7KWK, differences can be observed in the conformation of the bound ligands in the minor groove, but little difference is observed in the overall global DNA structure. Thus, these results suggest the interactions of the indole do not induce significant structural changes that vary from its benzimidazole. The similarity of 7KU4 and 7KWK can be seen from the resulting minor groove parameter (Fig. 13), the minor groove width of the two DNA complexes is shown to be about 1.1 Å wider than the native structure along the path of the bound ligand. The increase in minor groove width further confirms the ligand effect on the DNA minor groove. However, there is no significant difference in the minor groove width between 7KU4 and 7KWK. The results were a little surprising considering the difference in PU.1 inhibition between the two complexes. Nonetheless, since the minor groove binders are structurally similar, their groove distances seem probable.

Slight differences in the local DNA structure begin to emerge when the base-pair parameters of 7KU4 and 7KWK are considered. DB818 is bound in the minor groove of 7KU4 between the base's G/4 and T/8, while DB1879 in 7KWK lies between the bases A/5 to C/9. The

differences, albeit small, observed in the base-pair parameters of 7KU4 and 7KWK become more apparent within the base-pair region of the DNA bound by ligand. These differences in the ligand-bound regions are shown in Fig. 10, where for each parameter, the uniformity of the plots to one another distinctively changes between base pairs G/4 and C/9, the region of ligand binding. This suggests the presence of the ligand is exerting some effect along the local axis of the DNA base pairs. For the shear parameter, the indole, 7KWK shows an observable deviation from native -GAATTC- and 7KU4. For the stretch, stagger, and buckle parameter, both 7KU4 and 7KWK show some differences from the native. As minor as the differences in the base-pair parameters between all three structures may appear, these differences are observed along the entire double helix. Nonetheless, the differences in the base-pair parameters do not translate to a significant difference in the global structures of 7KU4 and 7KWK (Fig. 12).

With the base-step parameter, the major differences observed between the different structures only seem to be confined along the region of the bound ligand. Except for the slide parameter, little or no change is observed beyond the ligand-bound region. This suggests the base-pair parameters exert more influence than the base-step parameter over the larger DNA structure beyond the minor groove. Of course, the difference in the base-pair and base-step relationship with the DNA is expected because the base-step parameter primarily involves changes between stacked base pairs, not base-pairing interactions. The analysis of the overall structural parameters of 7KU4 and 7KWK does suggest that the ligand-binding in the minor groove of DNA induces differential structural changes in both complexes but does not seem considerable enough to account for the marked difference shown in the DB818 and DB1879 inhibition profiles.

The final major question of this study was whether these compounds could inhibit the binding of the PU.1 transcription factor to DNA? PU.1 binds in the major groove and inhibition by minor groove binding compounds depends on both the binding affinity of the compounds and indirect effects that act through DNA from the minor to the major groove. Compounds that bind strongly but do not perturb some aspect of the DNA structure from the PU.1 bound conformation will be poor inhibitors while compounds that bind more weakly but have a major effect on DNA conformation can be strong inhibitors. All inhibitors, however, must have sufficient affinity to displace PU.1.

The inhibition of PU.1 binding to the λB sequence was evaluated by an SPR method that was previously developed and used in our laboratory. As shown in Fig. 6, both compounds are strong inhibitors of PU.1 with IC_{50} values below 25 nM. DB818 is a stronger binder than DB1879 and about twice as strong an inhibitor of PU.1. From the electrostatic potential maps in Fig. 1B, DB818 has a higher positive charge along its inner face that can account for its slightly stronger binding. This positive potential could also orient DB818 optimally in solution prior to minor groove binding and account for its more rapid binding to the 5′-GAATTC-3′ sequence. These effects may also account for the slightly better inhibition of PU.1 by DB818. Comparison of PU.1 inhibition by additional closely related compounds should also help to clarify the induced component of PU.1 inhibition by minor groove binding compounds.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the National Institutes of Health Grant GM111749 (W.D. W. and D.W.B.) for financial support. We also thank HL155178 (G.M.K. P.) and NSF 2028902 (G.M.K.P.) for financial support.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bmc.2022.116861.

References

[1] Parkinson GN, Lee MP, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*. 2002;417:876–880.
[2] Nguyen B, Neidle S, Wilson WD. A role for water molecules in DNA-ligand minor groove recognition. *Acc Chem Res*. 2009;42:11–21.
[3] Laughton CA, Tanious F, Nunn CM, Boykin DW, Wilson WD, Neidle SA. Crystallographic and spectroscopic study of the complex between d $(CGCGAATTCGCG)_2$ and 2,5-bis(4-guanylphenyl)furan, an analogue of berenil. Structural origins of enhanced DNA-binding affinity. *Biochemistry*. 1996;35:5655–5661.
[4] Neidle S, Thurston DE. Chemical approaches to the discovery and development of cancer therapies. *Nat RevCancer*. 2005;5:285–296.
[5] Chen YW, Jhan CR, Neidle S, Hou MH. Structural basis for the identification of an i-motif tetraplex core with a parallel-duplex junction as a structural motif in CCG triplet repeats. *Angew Chem Int Ed Engl*. 2014;53:10682–10686.
[6] Wei D, Wilson WD, Neidle S. Small-molecule binding to the DNA minor groove is mediated by a conserved water cluster. *J Am Chem Soc*. 2013;135:1369–1377.
[7] Liu Y, Kumar A, Depauw S, et al. Water-mediated binding of agents that target the DNA minor groove. *J Am Chem Soc*. 2011;133:10171–10183.
[8] Wilson WD, Nguyen B, Tanious FA, et al. Dications that target the DNA minor groove: compound design and preparation, DNA interactions, cellular distribution and biological activity. *Curr Med Chem Anticancer Agents*. 2005;5:389–408.
[9] Mathis AM, Bridges AS, Ismail MA, 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.
[10] Paine MF, Wang MZ, Generaux CN, et al. Diamidines for human African trypanosomiasis. *Curr Opin Investig Drugs*. 2010;11:876–883.
[11] Depauw S, Lambert M, Jambon S, et al. Heterocyclic diamidine DNA ligands as HOXA9 transcription factor inhibitors: Design, molecular evaluation, and cellular consequences in a HOXA9-dependant leukemia cell model. *J Med Chem*. 2019;62:1306–1329.
[12] Jenquin JR, Coonrod LA, Silverglate QA, et al. Furamidine rescues myotonic dystrophy type I associated mis-splicing through multiple mechanisms. *ACS Chem Biol*. 2018;21:2708–2718.
[13] Matthes F, Massari S, Bochicchio A. Reducing mutant huntingtin protein expression in living cells by a newly identified RNA CAG binder. *ACS Chem Neurosci*. 2018;9:1399–1408.
[14] Antony-Debré I, Paul A, Leite J, et al. Pharmacological inhibition of the transcription factor PU.1 in leukemia. *J Clin Invest*. 2017;127:4297–4313.
[15] Vázquez O, Sánchez MI, Martínez-Costas J, Vázquez ME, Mascareñas JL. Bis-4-aminobenzamidines: versatile, fluorogenic A/T-selective dsDNA binders. *Org Lett*. 2010;12:216–219.
[16] Rodriguez J, Mosquera J, Learte-Aymamı S, Vázquez ME, Mascareñas JL. Stimuli-Responsive DNA Binding by Synthetic Systems. *Acc Chem Res*. 2020;53:2286–2298.
[17] Berman HM, Stallings W, Carrell HL, et al. Molecular and crystal structure of an intercalation complex: Proflavine-cytidylyl-(3′,5′)-guanosine. *Biopolymers*. 1979;18:2405–2429.
[18] Wood AA, Nunn CM, Czarny A, Boykin DW, Neidle S. Variability in DNA minor groove width recognised by ligand binding: the crystal structure of a bis-benzimidazole compound bound to the DNA duplex d(CGCGAATTCGCG)2. *Nucleic Acids Res*. 1995;23:3678–3684.
[19] Spink N, Brown DG, Skelly JV, Neidle S. Sequence-dependent effects in drug-DNA interaction: the crystal structure of Hoechst 33258 bound to the d (CGCAAATTTGCG)2 duplex. *Nucleic Acids Res*. 1994;22:1607–1612.
[20] Simone R, Balendra R, Moens TG, et al. G-quadruplex-binding small molecules ameliorate *C9orf72* FTD/ALS pathology *in vitro* and *in vivo*. *EMBO Mol Med*. 2018;10:22–31.
[21] Marchetti C, Zyner KG, Ohnmacht SA, et al. Targeting multiple effector pathways in pancreatic ductal adenocarcinoma with a G-quadruplex-binding small molecule. *J Med Chem*. 2018;61:2500–2517.
[22] Ahmed AA, Angell R, Oxenford S, et al. Asymmetrically substituted quadruplex-binding naphthalene diimide showing potent activity in pancreatic cancer models. *ACS Med Chem Lett*. 2020;11:1634–1644.
[23] Neidle S. A personal history of quadruplex-small molecule targeting. *Chem Rec*. 2015;15:691–710.
[24] Brown DG, Sanderson MR, Skelly JV, Jenkins TC, Brown T, Garman E, Stuart DI, Neidle S. Crystal structure of a berenil-dodecanucleotide complex: The role of water in sequence-specific ligand binding. *EMBO J*. 1990;9:1329–1334.
[25] Trent JO, Clark GE, Kumar A, et al. Targeting the minor groove of DNA: crystal structures of two complexes between furan derivatives of Berenil and the DNA dodecamer d$(CGCGAATTCGCG)_2$. *J Med Chem*. 1996;39:4554–4562.
[26] Brown DG, Sanderson MR, Garman E, Neidle S. Crystal structure of a berenil-d (CGCAAATTTGCG) complex. An example of drug-DNA recognition based on sequence-dependent structural features. *J Mol Biol*. 1992;226:481–490.
[27] Nunn CM, Jenkins TC, Neidle S. Crystal structure of d(CGCGAATTCGCG) complexed with propamidine, a short-chain homologue of the drug pentamidine. *Biochemistry*. 1993;32:13838–13843.

[28] Edwards KJ, Jenkins TC, Neidle S. The structure provides a rationale for the superior DNA-binding properties of propamidine as compared to pentamidine. *Biochemistry*. 1992;31:7104–7109.
[29] Greenidge PA, Jenkins TC, Neidle S. DNA minor groove recognition properties of pentamidine and its analogs: a molecular modeling study. *Mol Pharmacol*. 1993;43:982–988.
[30] Miao Y, Lee MP, Parkinson GN, et al. Out-of-shape DNA minor groove binders: induced fit interactions of heterocyclic dications with the DNA minor groove. *Biochemistry*. 2005;44:14701–14708.
[31] Midgley I, Fitzpatrick K, Taylor LM, et al. Boykin DW Pharmacokinetics and metabolism of the prodrug DB289 (2,5-bis[4-(N-methoxyamidino)phenyl]furan monomaleate) in rat and monkey and its conversion to the antiprotozoal/antifungal drug DB75 (2,5-bis(4-guanylphenyl)furan dihydrochloride). *Drug Metab Dispos*. 2007;35:955–967.
[32] Wilson WD, Tanious FA, Mathis A, Tevis D, Hall JE, Boykin DW. Antiparasitic compounds that target DNA. *Biochimie*. 2008;90:999–1014.
[33] Ming X, Ju W, Wu H, Tidwell RR, Hall JE, Thakker DR. Transport of dicationic drugs pentamidine and furamidine by human organic cation transporters. *Drug Metab Dispos*. 2009;37:424–430.
[34] Donlic A, Zafferani M, Padroni G, Puri M, Hargrove AE. Regulation of MALAT1 triple helix stability and in vitro degradation by diphenylfurans. *Nucleic Acids Res*. 2020;48:7653–7664.
[35] Mazur S, Tanious FA, Ding D, et al. A thermodynamic and structural analysis of DNA minor-groove complex formation. *J Mol Biol*. 2000;300:321–337.
[36] Mallena S, Lee MP, Bailly C, et al. Thiophene-based diamidine forms a "super" AT binding minor groove agent. *J Am Chem Soc*. 2004;126:13659–13669.
[37] Laxmikeshav K, Kumari P, Shankaraiah N. Expedition of sulfur-containing heterocyclic derivatives as cytotoxic agents in medicinal chemistry: A decade update. *Med Res Rev*. 2022;42:513–575.
[38] Kamal A, Sreekanth K, Shankaraiah N, Sathish M, Nekkanti S, Srinivasulu V. Dithiocarbamate/piperazine bridged pyrrolobenzodiazepines as DNA-minor groove binders: synthesis, DNA-binding affinity and cytotoxic activity. *Bioorg Chem*. 2015;59:23–30.
[39] Yadav U, Sakla AP, Tokala R, et al. Design and synthesis of 5-morpholino-thiophene-indole/Oxindole hybrids as cytotoxic agents. *Chem Select*. 2020;5:4356–4363.
[40] Sharma P, Reddy TS, Kumar NP, Senwar KR, Bhargava SK, Shankaraiah N. Conventional and microwave-assisted synthesis of new 1H-benzimidazole-thiazolidinedione derivatives: A potential anticancer scaffold. *Eur J Med Chem*. 2017;138:234–245.
[41] Sharma P, Srinivasa Reddy T, Thummuri D, et al. Synthesis and biological evaluation of new benzimidazole-thiazolidinedione hybrids as potential cytotoxic and apoptosis inducing agents. *Eur J Med Chem*. 2016;124:608–662.
[42] Guo P, Farahat AA, Paul A, Harika NK, Boykin DW, Wilson WD. Compound shape effects in minor groove binding affinity and specificity for mixed sequence DNA. *J Am Chem Soc*. 2018;140:14761–14769.
[43] Guo P, Paul A, Kumar A, et al. The Thiophene "Sigma-Hole" as a Concept for Preorganized, Specific Recognition of G·C Base Pairs in the DNA Minor Groove. *Chemistry*. 2016;22:15404–15412.
[44] Guo P, Paul A, Kumar A, et al. A modular design for minor groove binding and recognition of mixed base pair sequences of DNA. *Chem Commun (Camb)*. 2017;53:10406–10409.
[45] Rosenbauer F, Wagner K, Kutok JL, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet*. 2004;36:624–630.
[46] Mueller BU, Pabst T, Osato M, et al. Heterozygous PU.1 mutations are associated with acute myeloid leukemia. *Blood*. 2002;100:998–1007.
[47] Guo P, Farahat AA, Paul A, Boykin DW, Wilson WD. Engineered modular heterocyclic-diamidines for sequence-specific recognition of mixed AT/GC base pairs at the DNA minor groove. *Chem Sci*. 2021;12:15849–15861.
[48] Shao Y, Molnar LF, Jung Y, et al. Except for molecular mechanics and semi-empirical models, the calculation methods used in Spartan have been documented. *Phys Chem Chem Phys*. 2006;8:3172.
[49] Otwinowski Z, Minor W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol*. 1997;276:307–326.
[50] Winn MD, Ballard CC, Cowtan KD, et al. Overview of the CCP4 suite and current developments. *Acta Cryst D*. 2011;67:235–242.
[51] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D*. 2004;60:2126–2132.
[52] Pettersen EF, Goddard TD, Huang CC, et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*. 2021;30:70–82.
[53] Paul A, Musetti C, Nanjunda R, Wilson WD. Biosensor-Surface Plasmon Resonance: label-free method for investigation of small molecule-quadruplex nucleic aAcid interactions. *Methods Mol Biol*. 2019;2035:63–85.
[54] Nguyen B, Tanious FA, Wilson WD. Biosensor-surface plasmon resonance: quantitative analysis of small molecule-nucleic acid interactions. *Methods*. 2007;42:150–161.
[55] Munde M, Wang S, Kumar A, et al. Structure-dependent inhibition of the ETS-family transcription factor PU.1 by novel heterocyclic diamidines. *Nucleic Acids Res*. 2014;42:1379–1390.
[56] Munde M, Poon GM, Wilson WD. 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.
[57] Record Jr MT, Anderson CF, Lohman TM. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q Rev Biophys*. 1978;11:103–178.
[58] Wang S, Kumar A, Aston K, et al. Different thermodynamic signatures for DNA minor groove binding with changes in salt concentration and temperature. *Chem Commun*. 2013;49:8543–8545.
[59] Paul A, Chai Y, Boykin DW, Wilson WD. Understanding mixed sequence DNA recognition by novel designed compounds: the kinetic and thermodynamic behavior of azabenzimidazole diamidines. *Biochemistry*. 2015;54:577–587.
[60] Drew HR, Wing RM, Takano T, et al. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc Natl Acad Sci*. 1981;78:2179–2183.
[61] Seligman H. DNA replication–current advances. IntechOpen, London. https://doi.org/10.5772/791.
[62] Lavery R, Moakher M, Maddocks JH, Petkeviciute D, Zakrzewska K. Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res*. 2014;37:5917–5929.
[63] Blanchet C, Pasi M, Zakrzewska K, Lavery R. CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Res*. 2011;39(Web Server issue):W68–W73.