BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

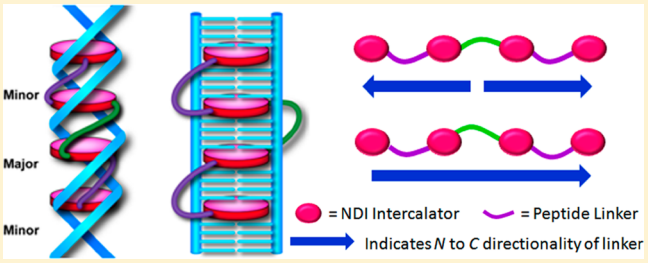
Subtle Recognition of 14-Base Pair DNA Sequences via Threading Polyintercalation

Amy Rhoden Smith, Brian A. Ikkanda, Garen G. Holman, and Brent L. Iverson*

Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712, United States

Supporting Information

ABSTRACT: Small molecules that bind DNA in a sequence-specific manner could act as antibiotic, antiviral, or anticancer agents because of their potential ability to manipulate gene expression. Our laboratory has developed threading polyintercalators based on 1,4,5,8-naphthalene diimide (NDI) units connected in a head-to-tail fashion by flexible peptide linkers. Previously, a threading tetraintercalator composed of alternating minor−major−minor groove-binding modules was shown to bind specifically to a 14 bp DNA sequence with a dissociation half-life of 16 days [Holman, G. G., et al. (2011) *Nat. Chem.* 3, 875−881]. Herein are described new NDI-based tetraintercalators with a different major groove-binding module and a reversed N to C directionality of one of the minor groove-binding modules. DNase I footprinting and kinetic analyses revealed that these new tetraintercalators are able to discriminate, by as much as 30-fold, 14 bp DNA binding sites that differ by 1 or 2 bp. Relative affinities were found to correlate strongly with dissociation rates, while overall C_2 symmetry in the DNA-binding molecule appeared to contribute to enhanced association rates.



Synthetic molecules that bind double-stranded DNA in a sequence-specific manner are of interest because of their potential ability to modulate gene expression. Several modes of DNA recognition have been explored,[1] including triple helix-forming oligonucleotides[2] and peptide nucleic acids (PNAs) that recognize DNA through specific hydrogen bonding between bases.[3,4] Minor groove-binding polyamides represent the most successful non-nucleic acid-based DNA-binding molecules to date and have been used to modulate gene expression both in vitro and in vivo.[5−9]

Intercalation is a mode of DNA binding that involves the insertion of a flat aromatic unit between the base pairs of DNA. While a plethora of mono- and bisintercalators have been studied,[10−16] including several investigated for therapeutic activity, only a handful of intercalators containing three or more intercalating units have been reported.[17−22] In addition, intercalators conjugated to polypeptides have often shown increased sequence specificity.[23−25]

We have developed a class of modular threading polyintercalators by connecting 1,4,5,8-naphthalenetetracarboxylic diimide (NDI) units together in a head-to-tail fashion via flexible peptide linkers.[26] NDI was first discovered as a threading intercalator by Wilson and co-workers,[27,28] and we have extended NDI polyintercalation to include the development of the first octakisintercalator, which was shown to bind DNA with all units intercalated,[29] albeit without demonstrated sequence specificity beyond a preference for GC-rich DNA.

Peptidic linkers specific for 6 bp sequences of double-stranded DNA[30−32] were discovered by using DNase I footprinting to screen NDI bisintercalator libraries. Using NMR structural analysis, an NDI bisintercalator with a Gly_3Lys linker was found to bind in the major groove to the sequence $d(CG|GTAC|CG)_2$ with 4 bp between intercalation sites.[30] A second bisintercalator with a β-Ala_3Lys linker was shown to bind the sequence d(CG|ATAA|GC)·d(GC|TTAT|CG) in the minor groove, again with 4 bp between intercalation sites.[31,32]

With these novel specificities and peptide linkers in mind, tetraintercalator **1** was designed, synthesized, and shown to bind specifically to a 14 bp palindromic sequence, 5′-G|ATAA|GTAC|TTAT|C-3′. The NMR-determined structure of the **1**−DNA complex verified that the molecule bound in a sequence-specific fashion with the linkers residing in the minor−major−minor grooves as predicted[33] (Figure 1). The minor groove-binding portions of **1** were the aforementioned β-Ala_3Lys linkers, but the major groove linker was derived from adipic acid, rather than Gly_3Lys, to add overall C_2 symmetry to the molecule to provide for a palindromic DNA binding site, simplifying the NMR structural analysis. This tetraintercalator was shown to bind to its 14 bp hybrid site with a dissociation half-life of 16 days, the longest reported half-life, to date, for any DNA-binding molecule.[34]

Herein are described the design, synthesis, footprinting, and kinetic analyses of new NDI tetraintercalators **2** and **3** that differ from **1** by targeting the major groove with the Gly_3Lys linker and by reversing the N to C directionality of one of the minor groove-binding β-Ala_3Lys linkers. This change in linker symmetry and directionality is expected to have a correspond-

Received: March 5, 2012
Revised: April 28, 2012
Published: May 3, 2012

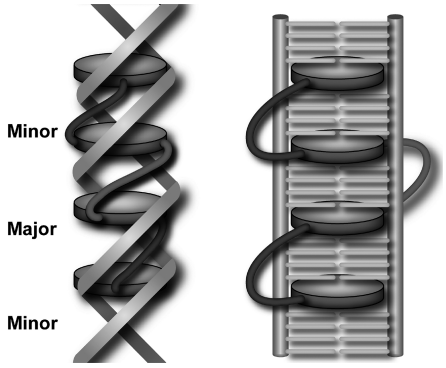


Figure 1. Cartoon depicting threading tetraintercalation.

ing influence on the directionality of the preferred DNA binding sequences of the different molecules. In addition, it is reasonable to propose that C_2 symmetric molecules such as **1** should have a distinct advantage when it comes to association rate constants, because overall C_2 symmetry ensures that all initial encounters with DNA occur with a backbone orientation consistent with formation of a bound complex. Molecules with unidirectional backbones lacking C_2 symmetry, such as **2** and **3**, would be expected to initially interact with DNA in the incorrect orientation approximately half of the time, thus resulting in a corresponding decrease in the productive association rate. Additionally, compounds **2** and **3** differ in the number of positive charges at neutral pH, possessing overall charges of +5 and +4, respectively, allowing for an analysis of any charge dependence of binding affinity and specificity.

EXPERIMENTAL PROCEDURES

Gel Mobility Shift Assays. Association and dissociation kinetics were measured using gel shift mobility assays. The 24-mer DNA (Supporting Information) was radiolabeled with ^{32}P, and the association gel shift assays were performed as described previously.[34] For the dissociation gel shift assays, we incubated the intercalator (1.1 μM) and DNA (0.05 μM radiolabeled, 1.0 μM unlabeled) by heating the DNA to 60 °C for 5 min, adding the intercalator, and cooling the incubation to 25 °C at a rate of 0.5 °C/min. The incubations were then allowed to remain at 25 °C overnight. Full association was confirmed by gel electrophoresis. Dissociation was instigated and quantified as previously described.[34]

DNase I Footprinting. PAGE-purified primers containing the binding sites to be tested were ordered to produce 5′- and 3′-sticky ends for ligation into pMoPac16[35] (Supporting Information). We annealed the primers by heating a 12.5 μM dsDNA solution at 95 °C and cooling the mixture at a rate of 1 °C/min until room temperature was reached. The pMoPac16 vector was digested with SfiI and purified with an agarose gel, and the annealed primers were ligated into the digested plasmid. The 5′-^{32}P-end-labeled DNA fragments were prepared by polymerase chain reaction (Supporting Information), and DNase I footprinting was performed according to a previously reported procedure.[36] The length of DNA used for footprinting was 95 bp. We incubated the tetraintercalators **1**−**3** with radiolabeled DNA by heating the DNA to 80 °C for 5 min, adding the intercalator, cooling the incubation at a rate of 0.5 °C/min until it reached 25 °C, and allowing the incubations to remain at 25 °C overnight. Incubations were digested with 2.5 units/mL DNase I for 4 min. The adenine-specific cleavage reaction was conducted according to the published procedure.[37] DNA fragments were separated on a 6% denaturing polyacrylamide gel. The gels were exposed to a phosphor screen overnight and imaged using Quantity One version 4.6.3 (Bio-Rad).

RESULTS

By changing the major groove-binding module in **1** from adipic acid to the Gly_3Lys linker, we found the putative tetraintercalators **2** and **3** have an overall backbone directionality and no longer exhibit C_2 symmetry. To emphasize this point, the arrows in Figure 2 indicate N to C

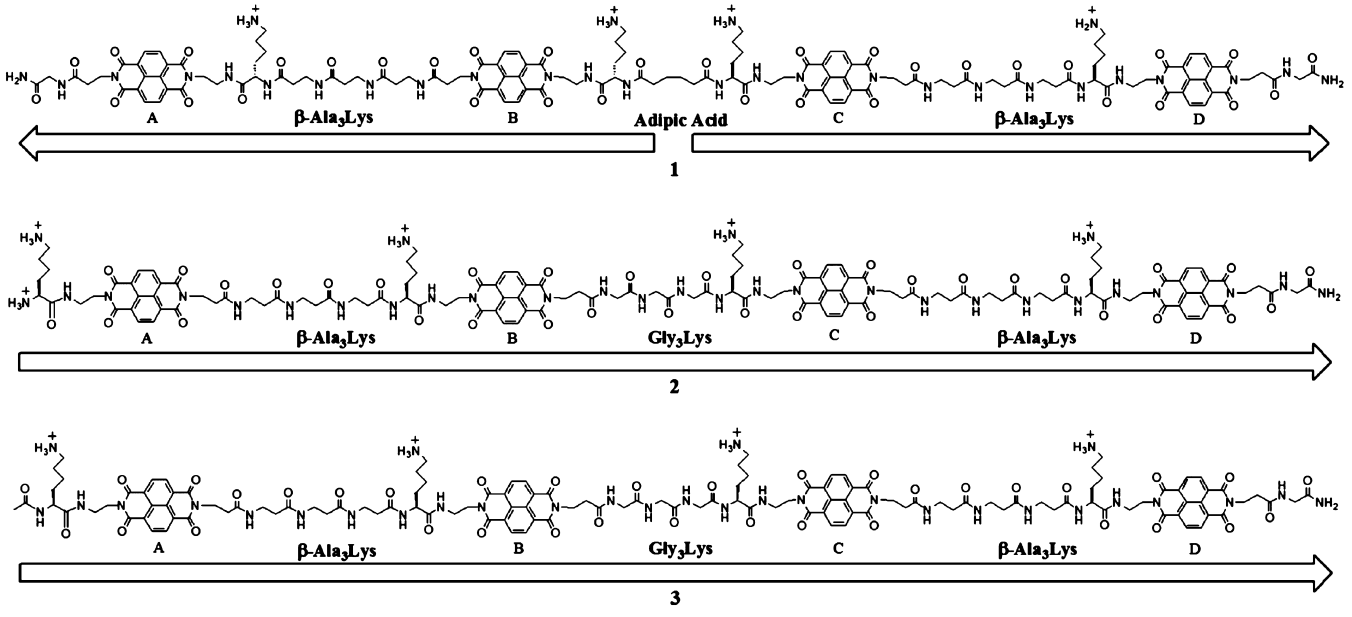


Figure 2. Structures of the adipic acid-containing tetraintercalator **1** and the Gly_3Lys linker-containing tetraintercalators **2** and **3** showing N to C amide bond directionality. Tetraintercalator **2** has an N-terminal free amine, while the N-terminal amine of **3** is capped as the acetamide. NDI intercalating units are denoted with letters for reference.

directionality in the amide bonds of the backbones for **1**–**3**. Although **1** could be synthesized by cross-linking two bisintercalators on the resin with adipic acid, **2** and **3** had to be synthesized in a linear, stepwise fashion.

Synthesis. Tetraintercalators **2** and **3** were synthesized using standard Fmoc solid phase peptide synthesis (SPPS) employing an orthogonal *t*-Boc protection for the lysine side chains as described previously.[38] To minimize coupling steps and maximize yield, Fmoc-$(\beta\text{-Ala})_3$-OH and Fmoc-$(Gly)_3$-OH were synthesized prior to SPPS[39,40] (Supporting Information). While **2** was left with a terminal free amine, the N-terminus of **3** was capped as the acetamide (using acetic anhydride) to produce species with +5 and +4 charges, respectively, at neutral pH.

Binding Site Design. The DNA binding site for **2** and **3** was designed by creating a hybrid 14 bp binding site from the individual bisintercalator binding sites. A combination of the minor–major–minor groove linker binding sites produces **DNA sequence 2** (Figure 3). For reference, the palindromic

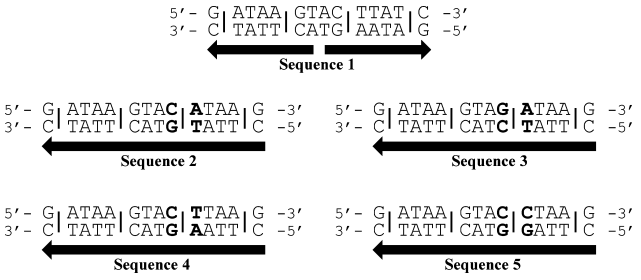


Figure 3. DNA sequence 1 is the binding site for **1**. **DNA sequences 2–5** are the proposed binding site sequences for **2** and **3**. Arrows indicate the overall expected N to C amide bond directionality of the corresponding bound tetraintercalator.

binding site for **1** is also shown (**DNA sequence 1**). Arrows in the figure have been added to emphasize the overall expected N to C directionality of the corresponding bound tetraintercalator. We foresaw a potential problem with the CA intercalation step (highlighted in bold) in the hybrid **DNA sequence 2**. This position is expected to be the site of intercalation for NDI unit **B**, but NDI in the context of our previously reported molecules is known to prefer purine-purine intercalation steps, GG especially.[30,32] For this reason, three other potential binding sites for **2** were designed in an attempt to evaluate this potential problem, listed as **DNA sequences 3–5**. For example, in **DNA sequence 3**, the major groove linker binding site is modified by switching a CG base pair to a GC base pair to allow for a GA intercalation step. In **DNA sequences 4** and **5**, the second minor groove linker binding site is modified. The AT base pair is switched to a TA base pair in **DNA sequence 4** and is exchanged for a CG base pair in **DNA sequence 5**. All modifications produce a purine-purine intercalation site, with only **DNA sequence 5** allowing for a GG intercalation site.

DNase I Footprinting. Concentration-dependent DNase I footprinting was used to evaluate binding of **2** and **3** to **DNA sequences 2–5** (Figure 4). In these studies, the DNA was heated to 80 °C, and then the intercalator was added followed by slow cooling. Note that footprints obtained in this way were identical to those of a 4 day incubation at 37 °C, providing confidence that equilibrium had been reached. To compare rigorously the binding to each sequence, the DNA used for footprinting contained only one of the proposed binding sites with the same flanking sequences on either side (Supporting Information). Intercalators **2** and **3** bind **DNA sequence 3** with the lowest affinity and appear to never occupy the entire site. For **DNA sequence 4**, **2** begins to bind at 150 nM but never fully occupies the sequence in the concentration range tested, and **3** begins to occupy the site at 100 nM. For **DNA sequence 2**, **2** begins to occupy the site at 100 nM, while **3** gives attenuation of digestion bands in the binding site starting at 50 nM. For **DNA sequence 5**, the binding site is mostly bound for both **2** and **3** at 50 nM. Thus, while all the target DNA sequences seem to experience some binding, the footprints seen with **DNA sequence 5** are the most distinct and occur at the lowest concentrations for both **2** and **3**.

DNase I footprinting studies were also performed with **2** and **3** on DNA containing the preferred palindromic binding site for **1**, **DNA sequence 1**, and with **1** on DNA containing the most promising target for **2** and **3**, **DNA sequence 5** (Figure 5). The strong footprint of **1** with **DNA sequence 1** is also shown for reference. Perhaps surprisingly, for **DNA sequence 1**, both **2** and **3** display significant footprints. Compound **2** begins to occupy the site at 100 nM, and **3** begins to bind at 50 nM, ranking these interactions the second strongest seen in the study, behind only those seen with **DNA sequence 5**. For tetraintercalator **1**, some binding was seen with **DNA sequence 5**, with a footprint beginning to appear at 100 nM, although the site is never fully occupied in the concentration range tested. For **2** and **3** at higher concentrations, other sites on the DNA, generally GC-rich regions, also seem to be occupied, implying that **2** and **3** may be less specific in binding compared to **1** in the context of these sequences. This is perhaps not altogether too surprising given that NDI monointercalators prefer binding GC-rich DNA.[28]

Gel Shift Analysis of Dissociation and Association Rates. The dissociation rate constants for **2** and **3** were determined with **DNA sequences 1–5** using gel shift mobility assays as previously reported.[34] Briefly, the tetraintercalator is incubated with a stoichiometric amount of ^{32}P-labeled 24-mer containing a 14 bp binding site. DNA with a bound tetraintercalator displays a retarded gel mobility so that bound and unbound radiolabeled DNA can be quantified using native polyacrylamide gel electrophoresis (PAGE) and autoradiography. For the dissociation rate constant analysis, complete initial binding is confirmed by a gel shift, and then a 100-fold excess of unlabeled DNA 24-mer is added. The large excess of unlabeled DNA duplex will bind any dissociated tetraintercalator and thereby prevent reassociation with radiolabeled DNA. The amount of bound versus unbound radiolabeled DNA is monitored over time (Figure 6), and the data are fit to a monoexponential decay equation to derive dissociation rate constants (Table 1 and Figure S2 of the Supporting Information). Both **2** and **3** display dissociation rates that decrease in the following order: **DNA sequence 3** > **4** > **2** > **1** > **5** (spanning a roughly 30-fold range of values). Dissociation from a control sequence, a 24-mer DNA duplex with no known tetraintercalator binding site (5′-CATTTAA-CAACATGTTGTTGGCTC-3′), was also analyzed. While the control sequence did display a gel shift when incubated with a stoichiometric amount of **2** and **3**, full association was never seen, and both were fully dissociated from the control sequence within 1 h; therefore, reliable values could not be determined. For the purposes of comparison, the dissociation kinetics of compound **1** were also investigated with **DNA sequence 5**. The previously reported extraordinarily slow dissociation rate

dx.doi.org/10.1021/bi300317n | *Biochemistry* 2012, 51, 4445–4452

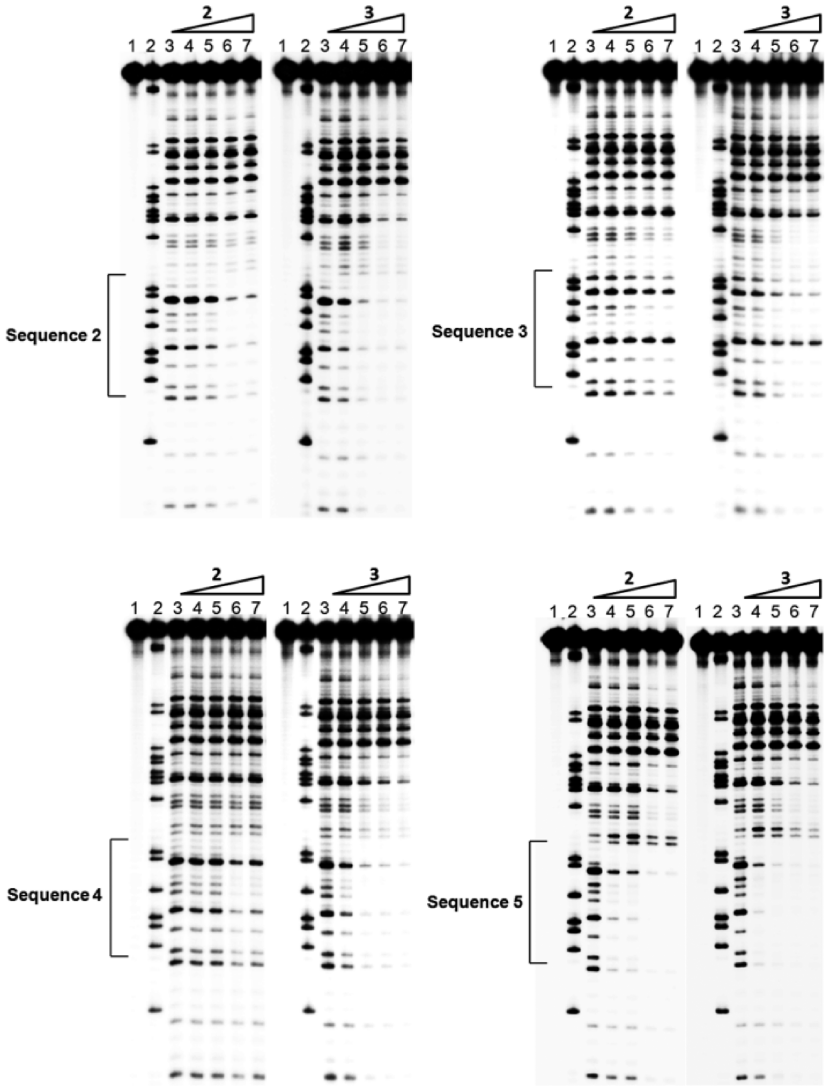


Figure 4. Concentration-dependent DNase I footprints of **2** and **3** with **DNA sequences 2**–**5**. Lane 1 contained no DNase I. Lane 2 contained an adenine-specific cleavage reaction mixture.[37] Lane 3 contained no intercalator. Tetraintercalator concentrations for lanes 4–7 were 50, 100, 150, and 200 nM, respectively.

constant for binding of **1** to **DNA sequence 1** is also shown in Table 1, revealing a >100-fold difference in dissociation rate constants for **1** with **DNA sequences 1** and **5**.

Association kinetics were also monitored by gel shift, yielding approximate k_a values for **2** and **3** with **DNA sequences 1**–**5**, and **1** with **DNA sequences 1** and **5**. For these experiments, a tetraintercalator and a radiolabeled 24-mer were combined in stoichiometric amounts at 500, 1000, and 1500 nM.[34] The integrated rate equation for stoichiometric binding was used to calculate association rate constants[41] (Table 2). Note that this experiment has historically given a relatively poor reproducibility in repeated measurements despite our attempts to establish a systematic protocol, explaining the relatively large errors listed in Table 2. Overall, for the DNA sequences that were analyzed using compound **1** (**DNA sequences 1** and **5**), compound **1** bound faster than either **2** or **3**. In addition, compound **2**, with one more positive charge than either **1** or **3**, bound several-fold faster than **3** but not as fast as **1**. Finally, no compound showed significant association rate differences (larger than error) between the different sequences examined in this study.

DISCUSSION

An important feature of molecules **2** and **3** is the fact that the backbone has an overall N to C orientation and thus a lack of C_2 symmetry as emphasized by the arrows in Figure 2. This contrasts with our original tetraintercalator **1**, in which there is overall C_2 symmetry by virtue of a symmetric middle linker and two external linkers with N to C directionality pointing away from the center. This change in symmetry is expected to have two significant consequences. First, the predicted DNA binding site for **2** and **3** is expected to exhibit a sequence progression that reflects the overall N to C backbone orientation of these molecules, contrasting with the palindromic binding site favored by the C_2 symmetric **1**. Second, the association kinetics of **2** and **3** can also be expected to be slower than those of **1** because of overall symmetry considerations. This is because **1** and its known preferred binding site, **DNA sequence 1**, are both C_2 symmetric, ensuring that either possible orientation of the intercalator initially associated with DNA will be aligned correctly to intercalate fully. For **2** and **3**, the initial electrostatic association of the molecules with their binding site will place approximately half of the molecules in the correct alignment for

dx.doi.org/10.1021/bi300317n | *Biochemistry* 2012, 51, 4445–4452

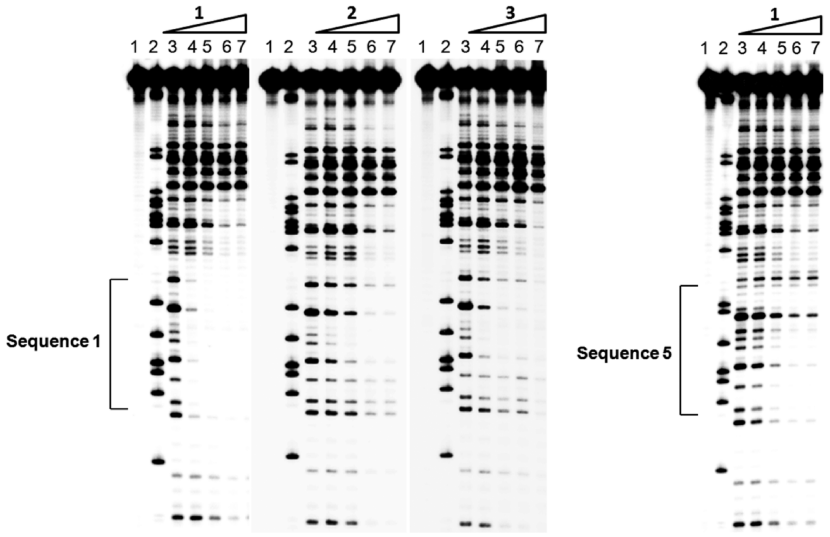


Figure 5. Concentration-dependent DNase I footprints of **1**−**3** with **DNA sequence 1** and of **1** with **DNA sequence 5**. Lane 1 contained no DNase I. Lane 2 contained an adenine-specific cleavage reaction mixture.[37] Lane 3 contained no intercalator. Tetraintercalator concentrations for lanes 4−7 were 50, 100, 150, and 200 nM, respectively.

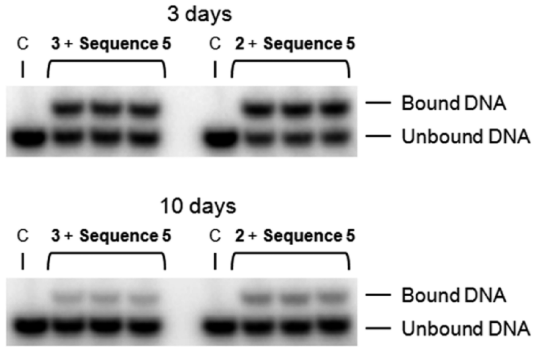


Figure 6. Examples of gel shift mobility assays of the dissociation of **2** and **3** from **DNA sequence 5** at 3 and 10 days. Each dissociation experiment was performed in triplicate.

binding, but the other half will be required to reverse their orientation on the DNA to match the directionality of their binding site prior to productive complex formation. Overall, such a requirement for proper orientation should decrease the rate of productive association of **2** and **3** relative to that of **1**.

Overall, rather subtle sequence discrimination was seen within the family of tetraintercalators **1**−**3**. **DNA sequences 2**−**5**, the sequences tested as predicted binding sites for **2** and **3**, varied by at most 2 bp of 14, yet significantly different footprinting results were obtained for these sequences. The order of binding preference from most to least preferred for **2** and **3** for these sequences is as follows: **DNA sequences 5**, **2**, **4**, and **3**. **DNA sequence 2**, the original hybrid sequence composed of the minor−major−minor groove bisintercalator binding sites, lacked a preferred purine-purine intercalation step. **DNA sequence 5** was modified from **DNA sequence 2** in the second minor groove linker binding site portion to give a more favorable GG intercalation step, so the preference of **2** and **3** for **DNA sequence 5** over **DNA sequence 2** is not altogether surprising. **DNA sequence 4** was also modified in the second minor groove binding site portion to give a GA intercalation step, and even though **DNA sequences 4** and **5** differ by only a single base pair, **2** and **3** display a multifold apparent preference for **DNA sequence 5** based on the footprinting results. **DNA sequence 3** is modified in the major groove binding site portion to allow for a GA intercalation step as well, and this sequence was by far the least preferred. Surprisingly, footprinting revealed the non-C_2 symmetric

Table 2. Association Rate Constants of 1−3 with DNA Sequences 1−5

	k_a ($\times 10^3$ M^{-1} s^{-1})		
	1	**2**	**3**
sequence 1	10 ± 5[a]	5 ± 1	1 ± 0.4
sequence 2	–[b]	4 ± 2	2 ± 1
sequence 3	–[b]	4 ± 1	1 ± 0.1
sequence 4	–[b]	3 ± 1	1 ± 0.5
sequence 5	7 ± 3	5 ± 2	3 ± 2

[a]See ref 34. [b]No measurement was attempted.

Table 1. Dissociation Rate Constants and Corresponding Half-Lives of 1−3 with DNA Sequences 1−5

	k_d ($\times 10^{-6}$ s^{-1})			$t_{1/2}$ (h)		
	1	**2**	**3**	**1**	**2**	**3**
sequence 1	0.50 ± 0.05[a]	2.8 ± 0.4	3.3 ± 0.4	390	69	59
sequence 2	–[b]	7.1 ± 1.2	5.6 ± 0.8	–[b]	27	34
sequence 3	–[b]	67 ± 7	63 ± 10	–[b]	2.9	3.1
sequence 4	–[b]	11 ± 2	9.5 ± 1.9	–[b]	17	20
sequence 5	69 ± 16	2.0 ± 0.3	2.8 ± 0.3	2.8	96	68

[a]See ref 34. [b]No measurement was attempted.

dx.doi.org/10.1021/bi300317n | *Biochemistry* 2012, 51, 4445−4452

compounds **2** and **3** also bound **DNA sequence 1** quite well, only surpassed by binding to their most highly preferred sequence, **DNA sequence 5**. However, the previously reported and C_2 symmetric compound **1** displayed a marked preference for the palindromic **DNA sequence 1** in comparison to **DNA sequence 5**. Thus, while it is clear that significant specificity has been achieved within the family of tetraintercalators **1**−**3** and that overall symmetry is important, it is not yet possible to identify all of the structural features, on the DNA or the tetraintercalators, responsible for the fine specificities observed.

Kinetic analyses revealed that relative affinities for the different DNA sequences correlated almost exclusively with dissociation rate constants, because association rate constants were found to be relatively sequence independent within error. Consistent with this notion, a strong qualitative correlation was seen between the footprinting results and relative dissociation rates, with stronger footprints at lower concentrations being observed for interactions that also displayed slower dissociation rates. For example, the previously reported remarkably slow dissociation rate constant (corresponding to a 16 day half-life) of compound **1** dissociating from its preferred binding site (**DNA sequence 1**) was the slowest measured, and compound **1** binding to **DNA sequence 1** also displayed the strongest footprint at the lowest concentration of any examined. Compounds **2** and **3** also displayed a similarly strong qualitative correlation between the overall strength of the interaction measured by footprinting and decreasing dissociation rates. Particularly interesting were the 5.6- and 3.4-fold differences in dissociation rates seen for binding of **2** and **3** to **DNA sequences 4** and **5**, respectively, which vary by a single base pair.

Upon comparison of **2** and **3**, with +5 and +4 charges, respectively, it is perhaps counterintuitive that the compound with the greater positive charge, **2**, displayed a modestly increased degree of selectivity. The expectation might be that **2** should dissociate slower than **3** from all sequences tested, because it is more highly charged. While compound **2** does dissociate slower than **3** from **DNA sequence 5**, the dissociation rate constants of **2** and **3** for dissociation from **DNA sequences 1**−**4** are approximately the same, within error. The same trend is also seen qualitatively in the footprinting data where **2** appears to bind with somewhat more discrimination to **DNA sequences 1**−**4** compared to **3**.

We have previously reported that **1** binds with high specificity to **DNA sequence 1** in the context of a 467 bp DNA fragment as shown by DNase I footprinting.[34] Here again, this specificity is further demonstrated by comparing the binding of **1** with **DNA sequences 1** and **5**. While **1** did show some occupation of **DNA sequence 5** with DNase I footprinting, gel shift dissociation analysis revealed a 140-fold difference in dissociation rates for binding of **1** to **DNA sequences 1** and **5**, which differ by 3 bp. While **2** and **3** show subtle sequence discrimination by exhibiting as much as a 30-fold difference in dissociation rate constants for the sequences tested, neither of these appears to be as specific as our previously reported **1**.

The subtle sequence discrimination observed here with 14 bp sequences would seem to imply that compounds **2** and **3** have all four NDI units fully intercalated in their preferred DNA complexes. Although a detailed structural analysis is beyond the scope of this study, a preliminary ^{1}H NMR analysis (Figure S3 of the Supporting Information) showed substantial movement of the DNA base pair imide H atoms in the complex formed with compound **3** bound to **DNA sequence 5**, consistent with full intercalation of all four NDI moieties. Note that for the sequences displaying weaker footprints and faster dissociation rates, it is quite possible that only partial intercalation is occurring across a portion of the binding site. We tacitly assume **2** and **3** bind to their preferred 14 bp DNA sequences with the same threading polyintercalation topology seen with compound **1**, although a confirmation of the bound topology for these molecules also must await completion of a detailed structural study.

Interestingly, compounds **2** and **3** displayed consistently slower association rates compared to those measured for compound **1**. For example, compounds **1** and **3** have identical overall charges of +4, yet their association rate constants differ by 1 order of magnitude. It may well be the case that overall C_2 symmetry can aid in the association of a DNA binding molecule (i.e., compound **1**) because such symmetry ensures an initial encounter with DNA will always involve the proper backbone orientation for binding to a preferred site. Binding of molecules without C_2 symmetry, such as **2** and **3**, may be at a relative association rate disadvantage because roughly half of all DNA encounters with such molecules would take place with the incorrect backbone directionality. This discussion assumes that both **2** and **3** bind with a single backbone orientation to preferred sites such as **DNA sequence 5**, an assumption that is strongly supported by the presence of a single bound species in the ^{1}H NMR analysis of **3** bound to **DNA sequence 5** (Figure S3 of the Supporting Information). Of course, the symmetry analysis described above is only one explanation for the observed increased association rate constants seen with **1** compared with those of **2** and **3**. Other structural differences not considered here, such as overall backbone flexibility (expected to be greater in **1**), could also be playing a significant role.

CONCLUSION

Taken together, the data presented here verify that subtle levels of discrimination can be achieved by altering the backbone structure and orientation within our growing family of DNA tetraintercalators. In particular, our kinetic data support the design idea that using C_2 symmetric molecules leads to faster overall association rates. Furthermore, 1 or 2 bp changes in a 14 bp binding site were seen to alter dissociation rates for **2** and **3** by as much as 30-fold. Relative affinities for the different tetraintercalator−DNA binding site sequence combinations used in this study were found to correlate with relative dissociation rates. Nevertheless, much remains to be learned about the details of DNA sequence recognition by our polyintercalators. For example, **DNA sequence 1** differed from **DNA sequence 5** by 3 bp, explaining the more than 100-fold difference in dissociation rates seen with compound **1**. However, compounds **2** and **3** displayed only modest differences in dissociation rates when analyzed using these DNA sequences but showed much larger differences among sequences that would seem more homologous (i.e., **DNA sequences 4** and **5**).

We are currently analyzing the details of binding of **2** and **3** to sequences such as **DNA sequence 5** by NMR to identify the structural features responsible for the subtle recognition observed. We are also attempting to exploit the modularity of our polyintercalating system to prepare different and longer hybrid molecules designed to prefer extended DNA sequences.

■ ASSOCIATED CONTENT

Supporting Information

Monomer synthesis, details of solid phase peptide synthesis, characterization of **2** and **3**, DNA sequence details for gel shift assays and DNase I footprinting, representative dissociation plots, and NMR spectra of **DNA sequence 5** with and without **3** bound. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*The University of Texas at Austin, Department of Chemistry and Biochemistry, Welch Hall 2.204, 105 E. 24th St. Stop A5300, Austin, TX 78712-1224. Telephone: (512) 471-5053. E-mail: iversonb@austin.utexas.edu.

Funding

This work was supported by the Robert A. Welch Foundation (F1188) and the National Institutes of Health (GM-069647).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Steven Sorey for his help with the ^{1}H NMR spectra.

■ ABBREVIATIONS

NDI, 1,4,5,8-naphthalenetetracarboxylic diimide; SPPS, solid phase peptide synthesis; Fmoc, 9-fluorenylmethyloxycarbonyl; *t*-Boc, *tert*-butyloxycarbonyl; PAGE, polyacrylamide gel electrophoresis.

■ REFERENCES

(1) Hannon, M. J. (2006) Supramolecular DNA recognition. *Chem. Soc. Rev. 36*, 280−295.

(2) Moser, H. E., and Dervan, P. B. (1987) Sequence-Specific Cleavage of Double Helical DNA by Triple Helix Formation. *Science 238*, 645−650.

(3) Nielsen, P. E. (2001) Targeting Double Stranded DNA with Peptide Nucleic Acid (PNA). *Curr. Med. Chem. 8*, 545−550.

(4) Hu, J., and Corey, D. R. (2007) Inhibiting Gene Expression with Peptide Nucleic Acid (PNA)−Peptide Conjugates That Target Chromosomal DNA. *Biochemistry 46*, 7581−7589.

(5) Dervan, P. B. (2001) Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem. 9*, 2215−2235.

(6) Mackay, H., Brown, T., Uthe, P. B., Westrate, L., Sielaff, A., Jones, J., Lajiness, J. P., Kluza, J., O'Hare, C., Nguyen, B., Davis, Z., Bruce, C., Wilson, W. D., Hartley, J. A., and Lee, M. (2008) Sequence specific and high affinity recognition of 5′-ACGCGT-3′ by rationally designed pyrrole-imidazole H-pin polyamides: Thermodynamic and structural studies. *Bioorg. Med. Chem. 16*, 9145−9153.

(7) Dervan, P. B., and Edelson, B. S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol. 13*, 284−299.

(8) Olenyuk, B. Z., Zhang, G.-J., Klco, J. M., Nickols, N. G., Kaelin, W. G., and Dervan, P. B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. U.S.A. 101*, 16768−16773.

(9) Muzikar, K. A., Nickols, N. G., and Dervan, P. B. (2009) Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc. Natl. Acad. Sci. U.S.A. 106*, 16598−16603.

(10) Portugal, J., Cashman, D. J., Trent, J. O., Ferrer-Miralles, N., Przewloka, T., Fokt, I., Priebe, W., and Chaires, J. B. (2005) A New Bisintercalating Anthracycline with Picomolar DNA Binding Affinity. *J. Med. Chem. 48*, 8209−8219.

(11) Westerlund, F., Nordell, P., Nordén, B., and Lincoln, P. (2007) Kinetic Characterization of an Extremely Slow DNA Binding Equilibrium. *J. Phys. Chem. B 111*, 9132−9137.

(12) Hampshire, A. J., Rusling, D. A., Bryan, S., Paumier, D., Dawson, S. J., Malkinson, J. P., Searcey, M., and Fox, K. R. (2008) DNA Binding by Analogues of the Bifunctional Intercalator TANDEM. *Biochemistry 47*, 7900−7906.

(13) Kricka, L. J., and Fortina, P. (2009) Analytical Ancestry: "Firsts" in Fluorescent Labeling of Nucleosides, Nucleotides, and Nucleic Acids. *Clin. Chem. 55*, 670−683.

(14) Wheate, N. J., Brodie, C. R., Collins, J. G., Kemp, S., and Aldrich-Wright, J. R. (2007) DNA Intercalators in Cancer Therapy: Organic and Inorganic Drugs and Their Spectroscopic Tools of Analysis. *Mini Rev. Med. Chem. 7*, 627−648.

(15) Boer, D. R., Canals, A., and Coll, M. (2008) DNA-binding drugs caught in action: The latest 3D pictures of drug-DNA complexes. *Dalton Trans.,* 399−414.

(16) Fathi, A. T., and Karp, J. E. (2009) New agents in acute myeloid leukemia: Beyond cytarabine and anthracyclines. *Curr. Oncol. Rep. 11*, 346−352.

(17) Hansen, J. B., Koch, T., Buchardt, O., Nielsen, P. E., Nordén, B., and Wirth, M. (1984) Trisintercalation in DNA by N-[3-(9-acridinylamino)propyl]-N,N-bis[6-(9-acridinylamino)hexyl]amine. *J. Chem. Soc., Chem. Commun.,* 509−511.

(18) Laugaa, P., Markovits, J., Delbarre, A., Le Pecq, J. B., and Roques, B. P. (1985) DNA Tris-intercalation: First Acridine Trimer with DNA Affinity in the Range of DNA Regulatory Proteins. Kinetic Studies. *Biochemistry 24*, 5567−5575.

(19) Wirth, M., Buchardt, O., Koch, T., Nielsen, P. E., and Norden, B. (1988) Interactions between DNA and mono-, bis-, tris-, tetrakis-, and hexakis(aminoacridines). A linear and circular dichroism, electric orientation relaxation, viscometry, and equilibrium study. *J. Am. Chem. Soc. 110*, 932−939.

(20) Takagi, M., Yokoyama, H., Takenaka, S., Yokoyama, M., and Kondo, H. (1983) Poly-intercalators Carrying Threading Intercalator Moieties as Novel DNA Targeting Ligands. *J. Inclusion Phenom. Mol. Recognit. Chem. 32*, 375−383.

(21) Ueyama, H., Takagi, M., and Takenaka, S. (2002) Tetrakis-acridinyl peptide: A novel fluorometric reagent for nucleic acid analysis based on the fluorescence dequenching upon DNA binding. *Analyst 127*, 886−888.

(22) Far, S., Kossanyi, A., Verchère-Béaur, C., Gresh, N., Taillandier, E., and Perrée-Fauvet, M. (2004) Bis- and Tris-DNA Intercalating Porphyrins Designed to Target the Major Groove: Synthesis of Acridylbis-arginyl-porphyrins, Molecular Modeling of Their DNA Complexes, and Experimental Tests. *Eur. J. Org. Chem. 2004*, 1781−1797.

(23) Bailly, C., Helbecque, N., Hénichart, J., Colson, P., Houssier, C., Rao, K. E., Shea, R. G., and Lown, J. W. (1990) Molecular recognition between oligopeptides and nucleic acids. DNA sequence specificity and binding properties of an acridine-linked netropsin hybrid ligand. *J. Mol. Recognit. 3*, 26−35.

(24) Takenaka, S., Iwamasa, K., Takagi, M., Nishino, N., Mihara, H., and Fujimoto, T. (1996) Synthesis of a 9-acridinyl nonapeptide containing the DNA recognizing region of 434 phage repressor protein. *J. Heterocycl. Chem. 33*, 2043−2046.

(25) Thompson, M. (2007) Spectral Properties and DNA Targeting Features of a Thiazole Orange−Peptide Bioconjugate. *Biomacromolecules 8*, 3628−3633.

(26) Lokey, R. S., Kwok, Y., Guelev, V., Pursell, C. J., Hurley, L. H., and Iverson, B. L. (1997) A New Class of Polyintercalating Molecules. *J. Am. Chem. Soc. 119*, 7202−7210.

(27) Yen, S. F., Gabbay, E. J., and Wilson, W. D. (1982) Interaction of aromatic imides with DNA. 1. Spectrophotometric and viscometric studies. *Biochemistry 21*, 2070−2076.

(28) Tanious, F. A., Yen, S. F., and Wilson, W. D. (1991) Kinetic and equilibrium analysis of a threading intercalation mode: DNA sequence and ion effects. *Biochemistry 30*, 1813−1819.

(29) Murr, M. M., Harting, M. T., Guelev, V., Ren, J., Chaires, J. B., and Iverson, B. L. (2001) An octakis-intercalating molecule. *Bioorg. Med. Chem. 9*, 1141−1148.
(30) Guelev, V., Lee, J., Ward, J., Sorey, S., Hoffman, D. W., and Iverson, B. L. (2001) Peptide bis-intercalator binds DNA via threading mode with sequence specific contacts in the major groove. *Chem. Biol. 8*, 415−425.
(31) Guelev, V. M., Harting, M. T., Lokey, R. S., and Iverson, B. L. (2000) Altered sequence specificity identified from a library of DNA-binding small molecules. *Chem. Biol. 7*, 1−8.
(32) Guelev, V., Sorey, S., Hoffman, D. W., and Iverson, B. L. (2002) Changing DNA Grooves: A 1,4,5,8-Naphthalene Tetracarboxylic Diimide Bis-Intercalator with the Linker $(\beta\text{-Ala})_3$-Lys in the Minor Groove. *J. Am. Chem. Soc. 124*, 2864−2865.
(33) Lee, J., Guelev, V., Sorey, S., Hoffman, D. W., and Iverson, B. L. (2004) NMR Structural Analysis of a Modular Threading Tetraintercalator Bound to DNA. *J. Am. Chem. Soc. 126*, 14036−14042.
(34) Holman, G. G., Zewail-Foote, M., Smith, A. R., Johnson, K. A., and Iverson, B. L. (2011) A sequence-specific threading tetra-intercalator with an extremely slow dissociation rate constant. *Nat. Chem. 3*, 875−881.
(35) Hayhurst, A., Happe, S., Mabry, R., Koch, Z., Iverson, B. L., and Georgiou, G. (2003) Isolation and expression of recombinant antibody fragments to the biological warfare pathogen *Brucella melitensis*. *J. Immunol. Methods 276*, 185−196.
(36) Trauger, J. W., and Dervan, P. B. (2001) Footprinting Methods for Analysis of Pyrrole-Imidazole Polyamide/DNA Complexes. *Methods Enzymol. 340*, 450−466.
(37) Iverson, B. L., and Dervan, P. B. (1987) Adenine specific DNA chemical sequencing reaction. *Nucleic Acids Res. 15*, 7823−7830.
(38) Guelev, V. M., Cubberley, M. S., Murr, M. M., Lokey, R. S., and Iverson, B. L. (2001) Design, Synthesis, and Characterization of Polyintercalating Ligands. *Methods Enzymol. 340*, 556−570.
(39) Eggen, I. F., Bakelaar, F. T., Petersen, A., Ten Kortenaar, P. B. W., Ankone, N. H. S., Bijsterveld, H. E. J. M., Bours, G. H. L., Bellaj, F. E., Hartsuiker, M. J., Kuiper, G. J., and Ter Voert, E. J. M. (2005) A novel method for repetitive peptide synthesis in solution without isolation of intermediates. *J. Pept. Sci. 11*, 633−641.
(40) Lapatsanis, L., Milias, G., Froussios, K., and Kolovos, M. (1983) Synthesis of N-2,2,2-(Trichloroethoxycarbonyl)-L-amino Acids and N-(9-Fluorenylmethoxycarbonyl)-L-amino Acids Involving Succinimidoxy Anion as a Leaving Group in Amino Acid Protection. *Synthesis 8*, 671−673.
(41) Anslyn, E. V., and Dougherty, D. A. (2006) *Modern Physical Organic Chemistry*, University Science Books, Sausalito, CA.