Chapter 4

Detection of DNA by Sequence Specific Fluorescent Polyamides

The text of this chapter is taken in part from a manuscript co-authored with Shane Foister and Christian Melander.

Rucker, V. C., Foister, S. F., Melander, C. M., Dervan, P. B. (2003). Sequence Specific Fluorescence Detection of Double Stranded DNA. *J. Am. Chem. Soc.*, **125**, 1195.

Abstract

Methods to sequence specifically localize small, fluorescent molecules to a prechosen sequence of double stranded DNA may be useful for detection purposes in biological applications. A series of hairpin polyamides derivatized with fluorescein or tetramethyl rhodamine (TMR) at an internal pyrrole ring was synthesized, and the fluorescence properties of the polyamide-fluorophore conjugates were examined. We observe weak TMR fluorescence in the absence of DNA. Significantly, addition of ≥ 1:1 match DNA affords a significant fluorescence increase over equimolar mismatch DNA for each polyamide-TMR conjugate. Using a parallel fluorescence assay, we are able to screen the interactions between the residues of the polyamide and the minor groove of DNA even when the binding site contains a non-Watson-Crick DNA base pair. This chapter is divided into two sections. In section 1, we describe our initial inquiries into the phenomenon that allows a polyamide to quench the xanthene fluorophore to which it is attached. In section 2, we rank the specificity of five polyamide residue pairs – Py/Py, □/Py, Im/□, Im/Py, and Im/Im – against all 16 possible base pairs of A, T, G and C in the minor groove. We find that Im/Im is an energetically favorable ring pair for minor groove recognition of the T•G base pair.

Section 1

Introduction

Interest in the detection of specific nucleic acid sequences in homogeneous solution has increased due to major developments in human genetics. Single nucleotide polymorphisms (SNPs) are the most common form of variation in the human genome and can be diagnostic of particular genetic predispositions to disease.[1-7] Most methods of DNA detection involve hybridization of an oligonucleotide probe to its complementary single-strand nucleic acid target leading to signal generation.[8-18] One example is the "molecular beacon" which consists of a hairpin DNA labeled in the stem with a fluorophore and a quencher.[4] Binding to a complementary strand results in opening of the hairpin and separation of the quencher and the fluorophore. Detection by hybridization requires DNA denaturation conditions, and it remains a challenge to develop sequence specific fluorescent probes for DNA in the double strand form.[19-25]

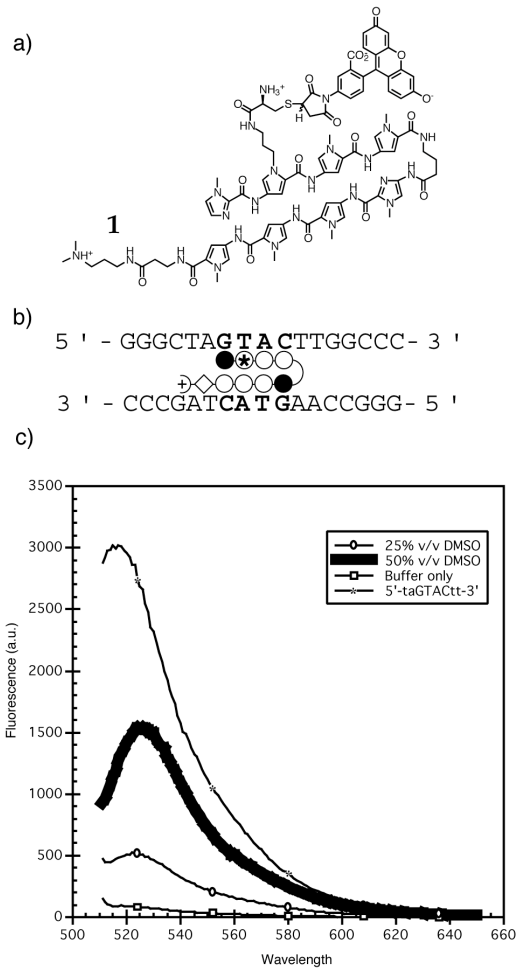


Figure 4.1 A. Polyamide-fluorescein conjugate observed to have weak fluorescence in absence of DMSO or match DNA. B. Ball and stick model of **1** and DNA sequence it is match for. C. Fluorescence behavior of 1 μM **1** in presence of DMSO or 1 equivalent of match oligo.

Within the context of developing non-denaturing stains for DNA, Laemmli

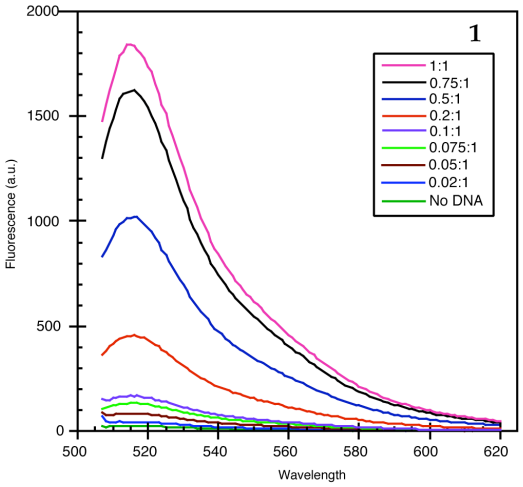


Figure 4.2 Emission increases from a 1 μM solution of **1** as match DNA is titrated to the one equivalency point.

demonstrated that polyamides labeled with Texas Red or fluorescein at the C-terminus are capable of specifically staining 5'-GAGAA-3' repeats in *Drosophila* satellites and 5'-TTAGG-3' and 5'-TTAGGG-3' teleomeric repeats in insect and human chromosomes, respectively.[26,27] Similarly, Trask has demonstrated the use of polyamide-dye conjugates to fluorescently label specific repetitive regions on human chromosomes 9, Y and 1 (5'-TTCCA-3' repeats) for discrimination in cytogenetic preparations and flow cytometry.[28]

We discuss here our inquiries into a phenomenon that allows hairpin polyamide-fluorophore conjugates to fluorescently report specific DNA sequences within short segments of double helical DNA in homogeneous solution. We observe that the fluorescence of polyamide-TMR or fluorescein conjugates is largely quenched in the absence of DNA containing the polyamide's match binding site. Addition of duplex DNA containing the match site for a polyamide-fluorophore conjugate affords a ≥ 10-fold increase in fluorescence upon polyamide binding. Attenuated fluorescence occurs when mismatch DNA is added at the same concentration. In this section we describe our initial inquiries and provide proof of principle data. In the following section we employ this fluorescence detection method to analyze the specificity of hairpin polyamides challenged to bind DNA sequences containing *non-Watson-Crick* paired bases.

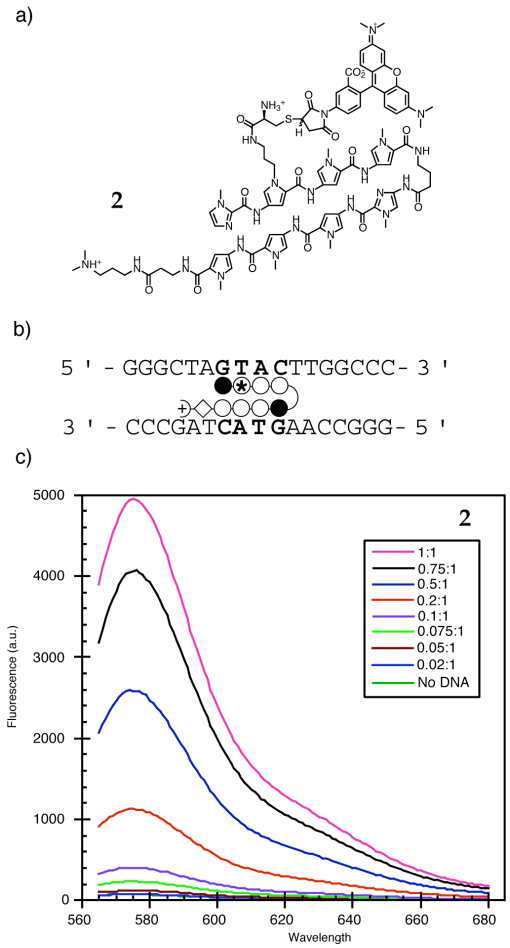


Figure 4.3 A. Structure of **2**, a TMR analog of **1**. B. Match recognition site for **2**. C. Emission profile from 1 μM of **2** when match DNA is titrated in from 20 nM to 1 μM.

Discussion of the principles underlying the fluorescence phenomenon described here is presented in Chapter 5.

Results and discussion

Fluorescein is quenched when conjugated to a hairpin polyamide

Initially, we were interested in designing bifunctional polyamide-fluorophore conjugates that would sequence specifically transport a fluorophore to a prechosen sequence of DNA. Interesting applications such as FRET or chromosome staining, as previously reported,[26-28] are easily imaginable. Our initial efforts to achieve this end began with polyamide-fluorophore conjugate **1**, a match for the sequence 5'-GTAC-3' (Figure 4.1 - synthesis provided at end of section). We were confounded to notice, however, that **1** demonstrated little fluorescence, a surprising observation for fluorescein given its large quantum yield at physiological pH[29] (1X TKMC: pH 7.3, 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$). Presuming that not all of **1** was fully dissolved, fluorescence measurements were made on samples of **1** in water containing dimethyl sulfoxide as co-solvent. Surprisingly, the fluorescence increased as the amount of DMSO (v:v) was increased (Figure 4.1C). More surprising, though, was

the observation that addition of oligonucleotide (1X TKMC buffer) containing the polyamide's match binding site evinced the largest increase in fluorescence (Figure 4.1C).

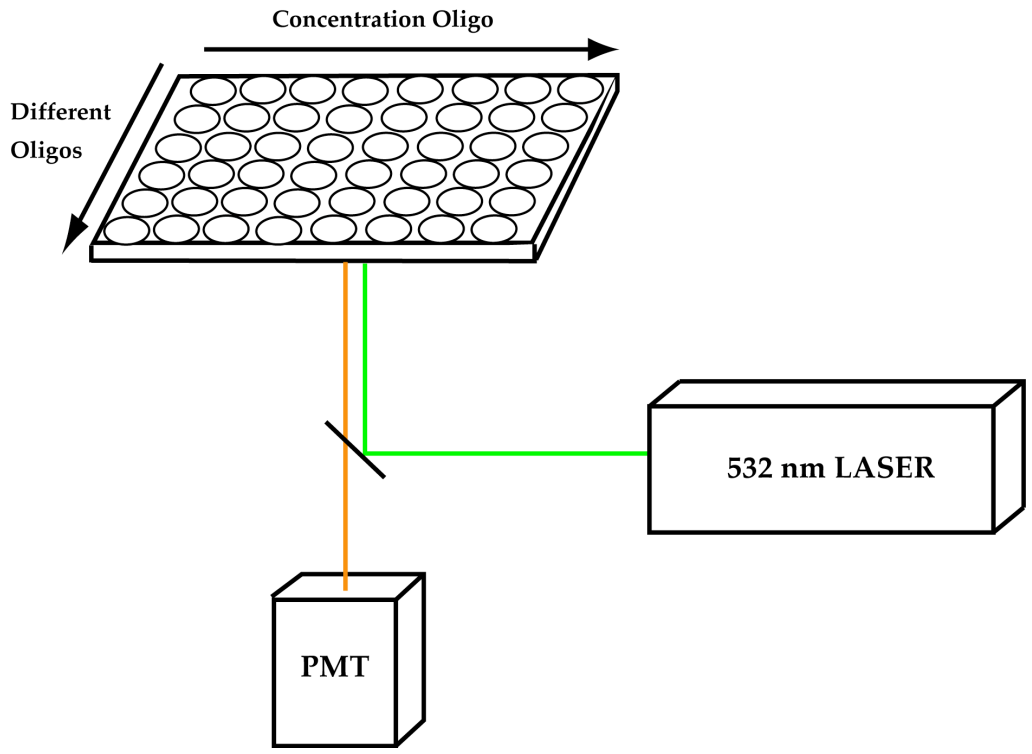


Figure 4.4 Design of parallel plate assay for probing polyamide interaction with a potential recognition site embedded in a short oligo. Plates are scanned through the bottom and emission collected after processing through a correct filter.

Taking this data at face value, it seemed obvious that some remarkable chemistry may be at hand, so we examined the possibility of a concentration dependence for the fluorescence "rescue" that we noted upon addition of 1 equivalent of match site containing DNA. Fixing the concentration of **1** at 1 μM, such that [**1**] $\geq 2K_d$,[23,24] we titrated in a 17-mer oligonucleotide containing the match site for **1** (5'-GTAC-3') over the concentration range 0.02 μM - 1 μM. We observed a concentration-dependent fluorescence increase (Figure 4.2).

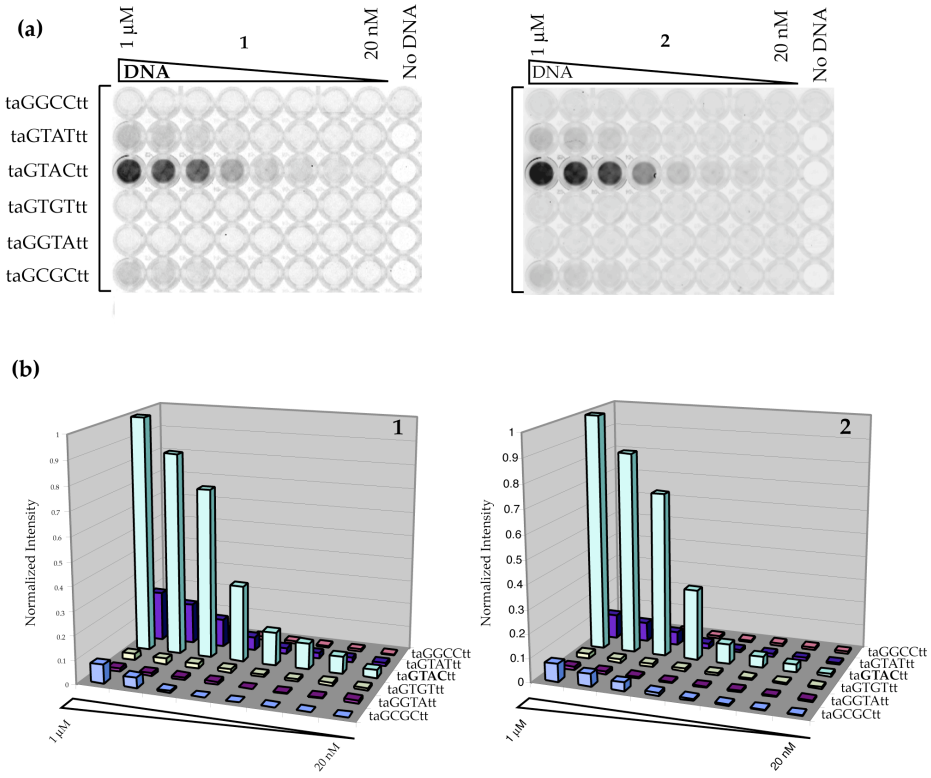


Figure 4.5 Data from parallel screen of **1** and **2** against six oligos, 1 match and six mismatches. **1** is excited at 532 nm with a 526 nm short-pass filter used for collection of emission. **2** is excited at 532 nm and a 580 nm filter is used to collect emission. A. Raw emission data. "No DNA" provides for measure of background fluorescence in absence of DNA. B. Normalized histogram of data in (A).

Tetramethyl rhodamine (TMR) is quenched when conjugated to a hairpin polyamide

Excited by these results, we explored fluorophore generality by synthesizing and investigating the fluorescence properties of conjugates incorporating tetramethyl rhodamine (TMR). We chose TMR because it is more photostable than fluorescein, and, anticipating the parallel screening assays we present below, we needed a fluorophore more compatible with the 532 nm laser of our Molecular Dynamics Typhoon imaging system. As shown in Figure 4.3, we synthesized a TMR containing analog of **1**,

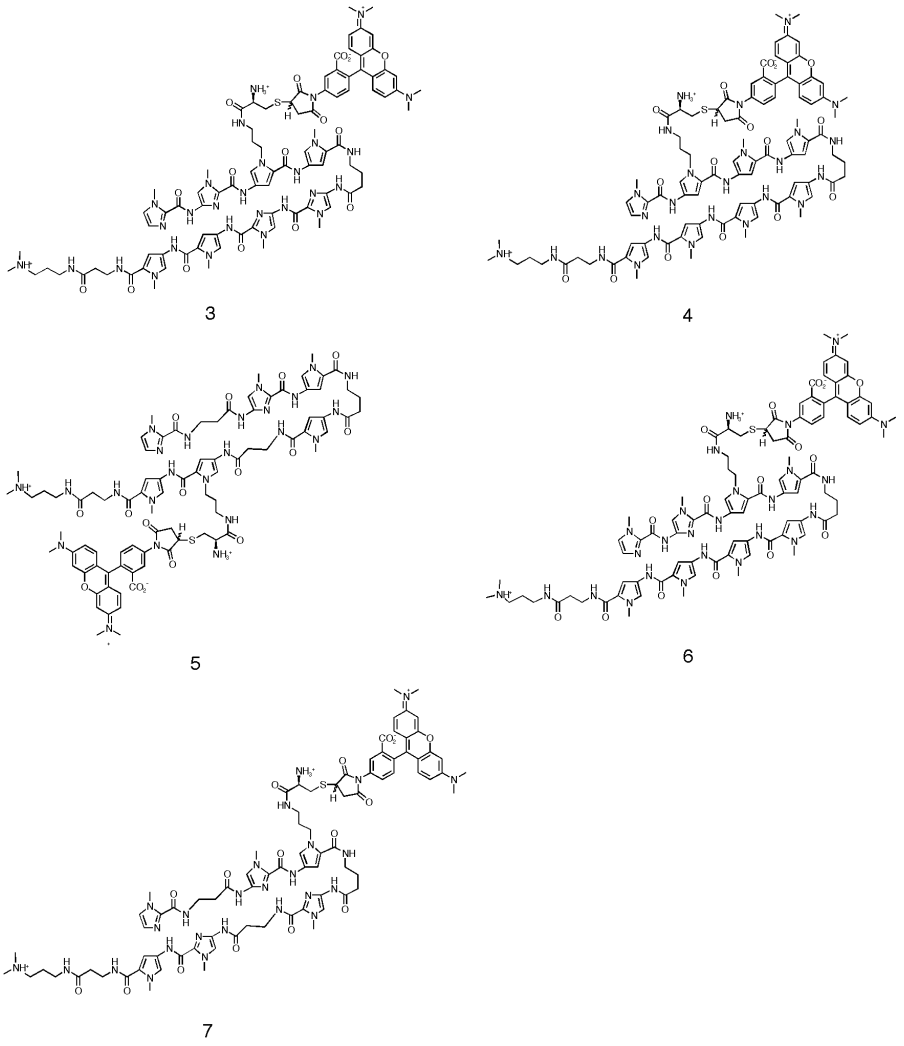


Figure 4.6 Polyamides synthesized to explore polyamide generality.

conjugate **2**. In the presence of match DNA, conjugate **2** exhibited fluorescent properties identical to those of **1** (save for inherent bathochromically shifted rhodamine absorption and emission maxima). Thus emboldened, we set out to explore the interactions of our conjugates with pairing rules *mis*match site containing oligonucleotides.

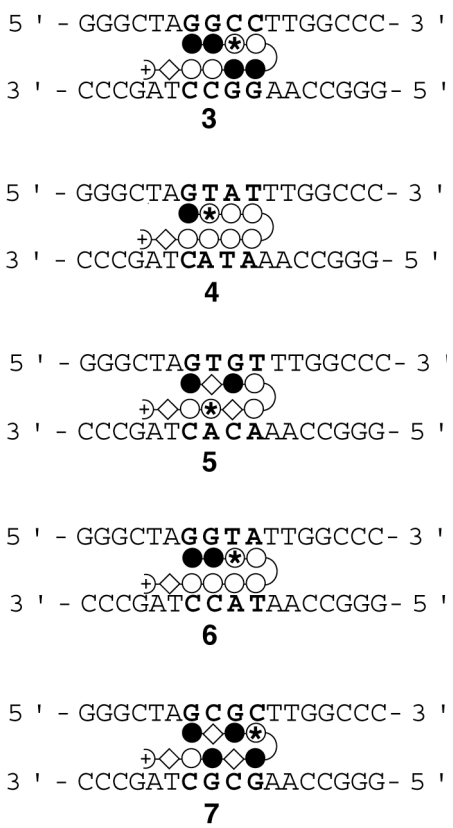


Figure 4.7 Ball-and-stick representation of polyamide-fluorophore conjugates **3** – **7** shown with DNA sequence recognized by each. Asterisk indicates position of fluorophore.

Explorations with mismatch binding sites

Up to this point, no explorations against mismatched DNA had been performed. We characterized the binding of **1** and **2** to match oligonucleotide, but what would be the properties of **1** and **2** in the presence of DNA containing mismatch binding sites? To investigate this, we chose to explore the interactions of **1** and **2** in parallel with 5 other mismatched oligonucleotides. Given the concentration dependence of **1** and **2** at their match (Figures 4.2 and 4.3) we postulated that the if the polyamide is sequence specifically recognizing the floor of the minor groove, and binding the minor groove is necessary for rescue of fluorescence, then less fluorescence should be observed when our conjugates are challenged to bind DNA containing mismatch recognition sites. We chose to characterize the interactions of **1** and **2** interacting with the match oligonucleotide (5'-taGTACtt-3') and 5 separate oligonucleotides containing the mismatch binding sequences 5'-taGGCCtt-3', 5'- taGTATtt-3', 5'-taGTGTtt-3', 5'-taGGTAtt-3', and 5'-taGCGCtt-3'.[30]

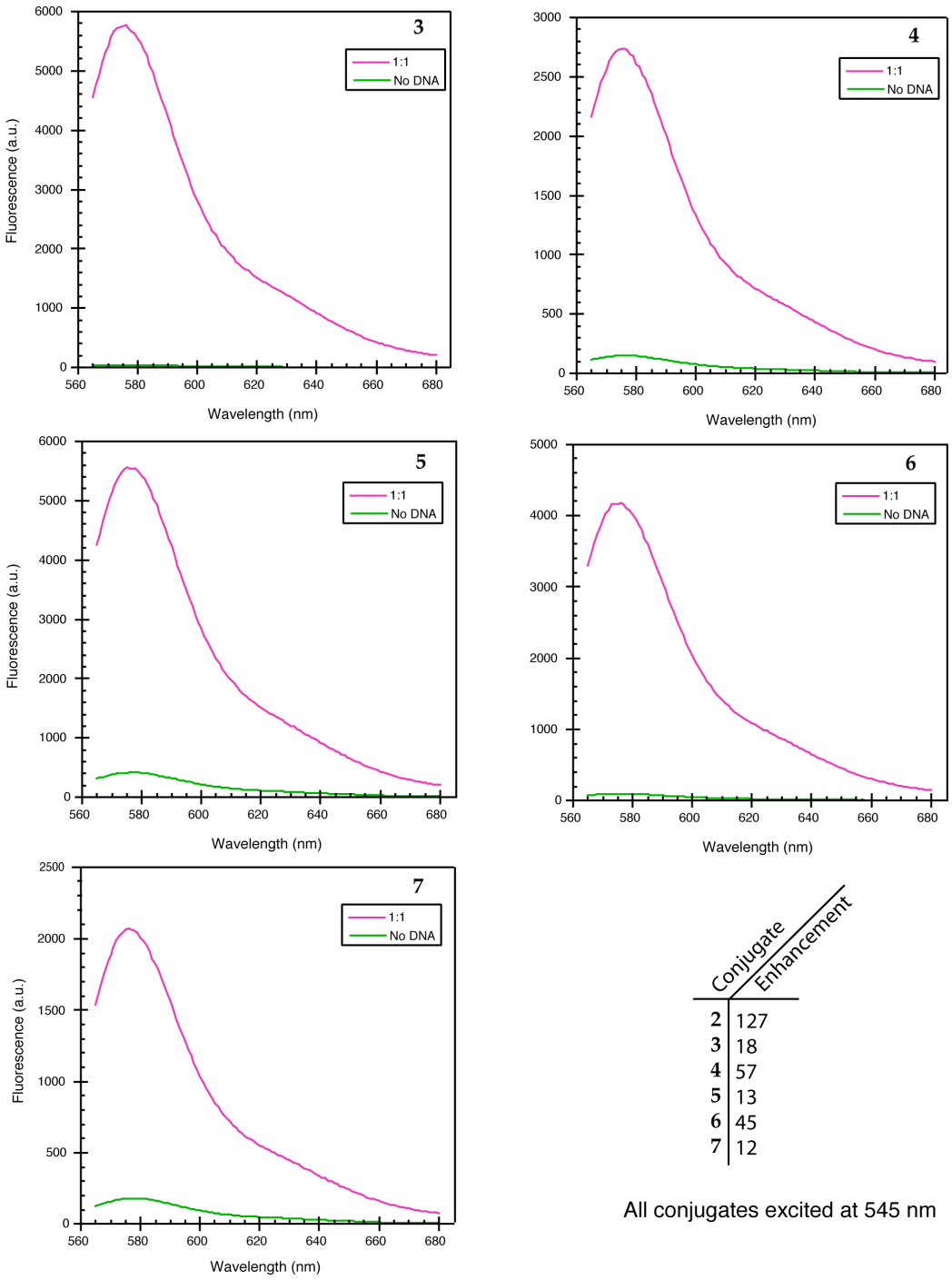


Figure 4.8 Emission observed for 1 μM **3** – **7** in the presence of 1 equivalent of respective match DNA. Table in lower right corner is the enhancement in fluorescence observed under these conditions, and it is simply the ratio of fluorescence observed in the presence of match DNA divided by the fluorescence observed in the absence.

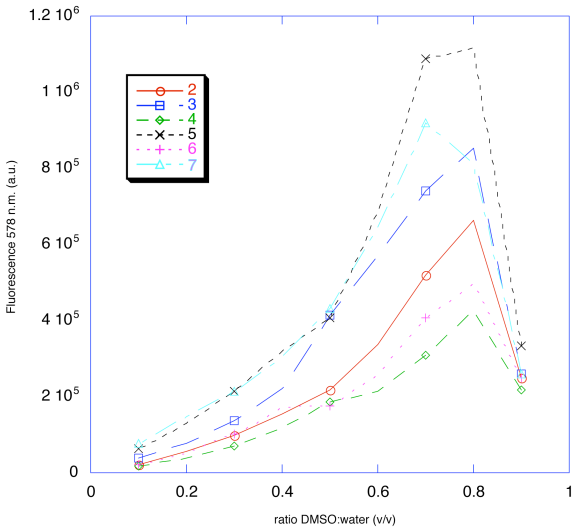


Figure 4.9 Conjugates **2** – **7** exhibit increase in quantum yield as ratio of DMSO:water is increased. This behavior is similar to that observed with prototype **1**. Sudden decrease in fluorescence at ratio of DMSO:water ~ 0.9 may be due to solvent effects on hydroxide pKa.[49]

Direct excitation with a 532 nm laser through the clear bottom of 96-well polystyrene plates was performed on 150 μL the conjugate, held at 1 μM, screened against 6 oligos, 5 mismatches and 1 match, of linearly increasing concentration (Figure 4.4). Fluorescein emission was collected through a 526 nm short-pass filter and TMR emission was collected with a 580 nm (± 15 nm FWHM) filter. As shown in Figure 4.5, we were gratified to note that the polyamide exhibits sequence specific fluorescence increase most greatly in the presence of the match recognition site for the polyamide. The fluorescence increase is much less in the presence of an oligonucleotide containing a mismatch polyamide binding site. Normalizing the fluorescence data from the direct excitation experiment and plotting these values as a histogram allows one to see that the thermodynamic underpinnings of the pairing rules allow **1** and **2**, by means of a fluorescence rescue dependent on extent of polyamide binding, to report on the sequence of DNA present in homogenous solution (Figure 4.5B).

Exploring polyamide generality

Emboldened, we synthesized 5 polyamides (Figure 4.6) that are match compounds for the DNA we employed as mismatches for conjugates **1** and **2**. Each

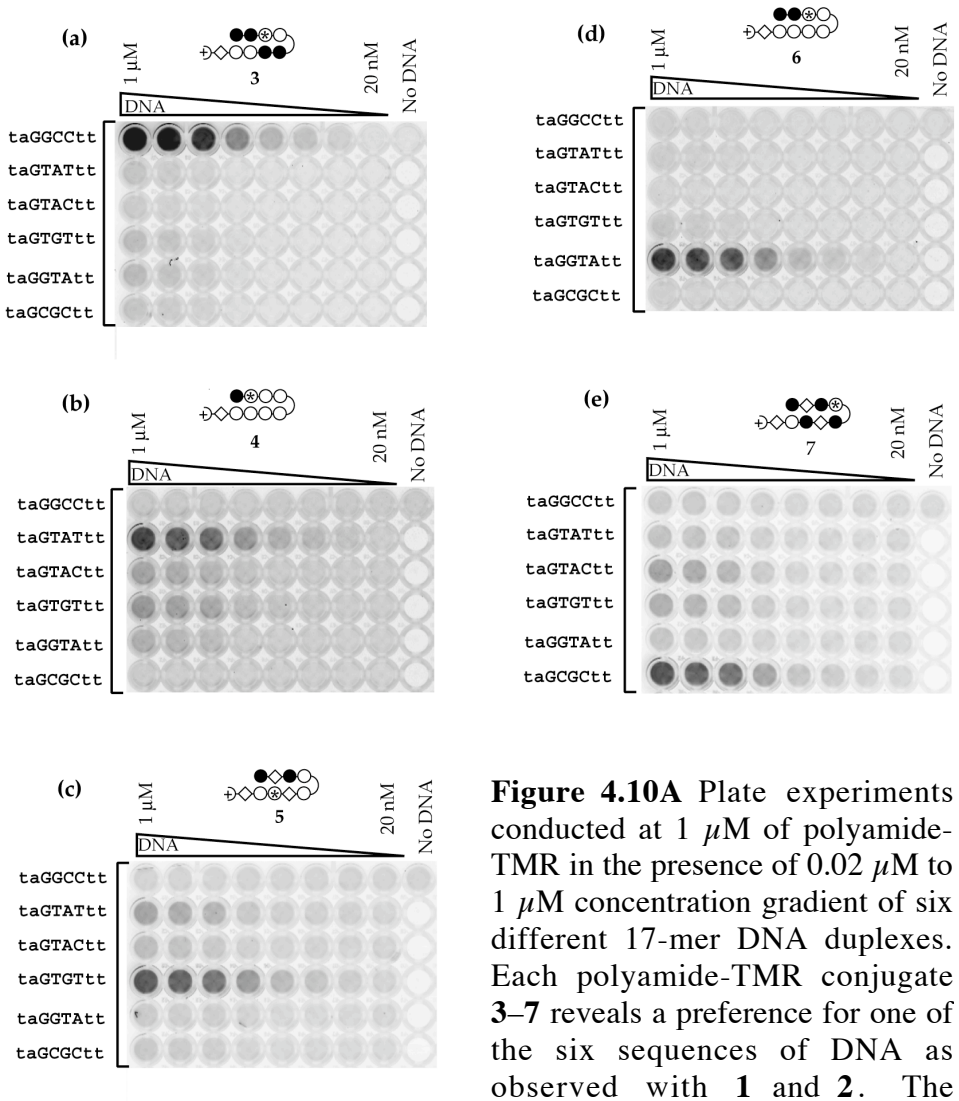


Figure 4.10A Plate experiments conducted at 1 μM of polyamide-TMR in the presence of 0.02 μM to 1 μM concentration gradient of six different 17-mer DNA duplexes. Each polyamide-TMR conjugate **3**–**7** reveals a preference for one of the six sequences of DNA as observed with **1** and **2**. The "minus" DNA well is shown for assessment of background fluorescence. Intensity correlates with the dark color of a well.

polyamide, as with **1** and **2**, is derivatized with TMR through a short cysteine linker and

is programmed to recognize a unique DNA sequence as determined by the pairing rules (Figure 4.7). The advantage of these five compounds, in addition to allowing for the possibility of exploring the sequence recognition properties of these compounds by

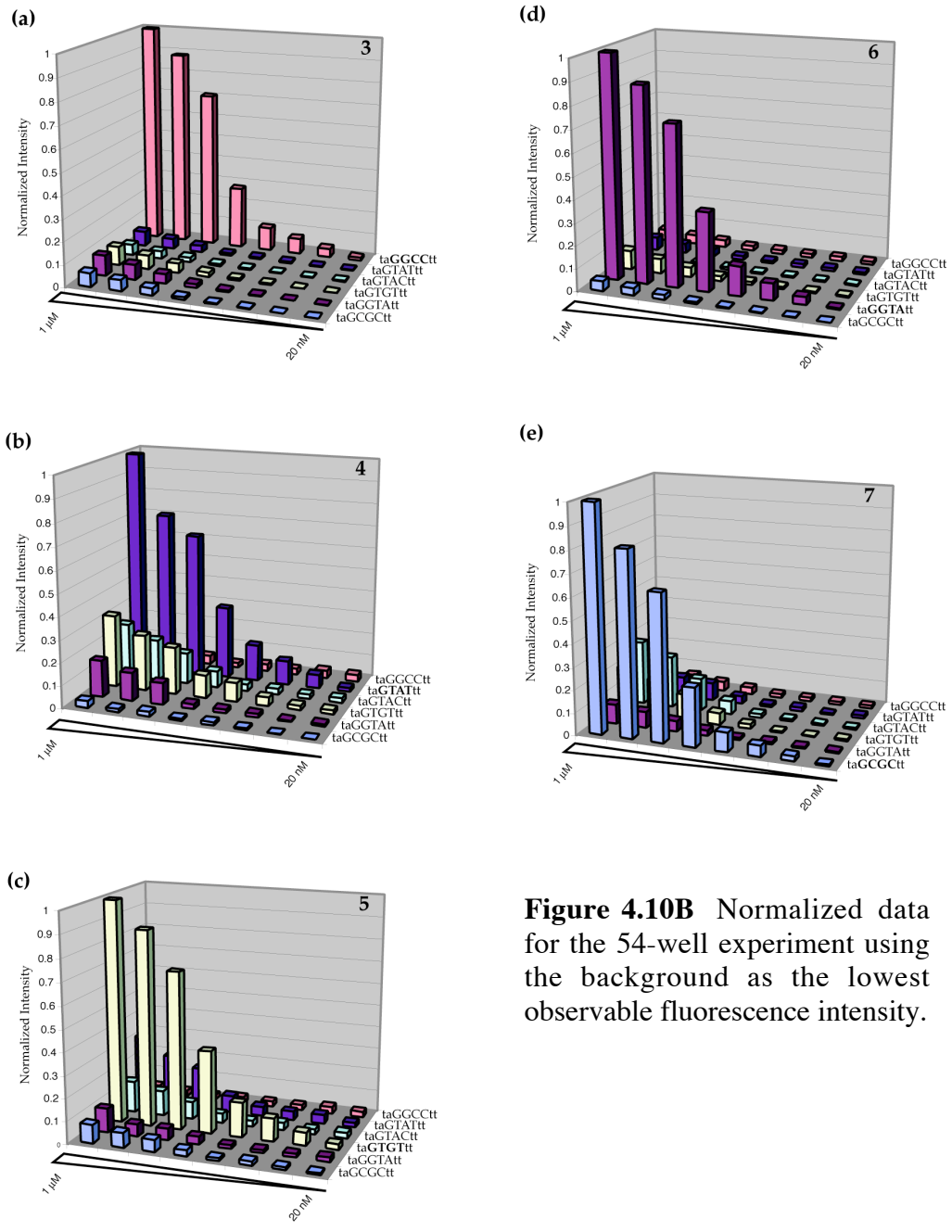


Figure 4.10B Normalized data for the 54-well experiment using the background as the lowest observable fluorescence intensity.

fluorimetry, is that they allow us to examine a diversity of polyamide structure and composition. At this point, given our limited knowledge about this new class of

Table 4.1 Normalized fluorescence data for conjugates **2-7** at [conjugate]=[DNA]=1 μM.[a]

	2	**3**	**4**	**5**	**6**	**7**
taGGCCtt	0.03	**1**	0.04	0.04	0.05	0.08
taGTATtt	0.09	0.04	**1**	0.30	0.20	0.10
taGTACtt	**1**	0.02	0.22	0.12	0.01	0.27
taGTGTtt	0.01	0.04	0.30	**1**	0.08	0.18
taGGTAtt	0.03	0.05	0.13	0.09	**1**	0.07
taGCGCtt	0.05	0.07	0.02	0.07	0.04	**1**

[a]The assays were carried out at 22°C at pH 7.0 in the presence of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

polyamide-fluorophore conjugates, we revisited experiments performed in characterizing **1** and **2**. We observed fluorescence increases over background when 1 μM **3** - **7** are in the presence of 1 μM oligonucleotide containing the match polyamide recognition site (Figure 4.8). We observed (Figure 4.9) that our new conjugates **3** - **7** are fluorescently quenched in the absence of DNA, but that upon adding DMSO there is an increase in quantum yield. Additionally, the fluorescence behavior of **3** - **7** in the presence of match DNA recapitulates that observed for **1** and **2**.

Screening in parallel 1 μM **3** - **7** against 0.02 - 1 μM of 1 match and 5 different mismatch oligos produced results equivalent to those observed previously for the interaction of **1** and **2** with match and mismatch binding sites (Figure 4.5, Figure 4.10A). Again, normalizing the raw fluorescence to that of the conjugate background fluorescence provided histograms as shown in Figure 4.10B. Furthermore, this data can be extracted into tabular form to yield a numeric ranking of the relative thermodynamics of polyamide discrimination between different sequences of DNA (Table 4.1).

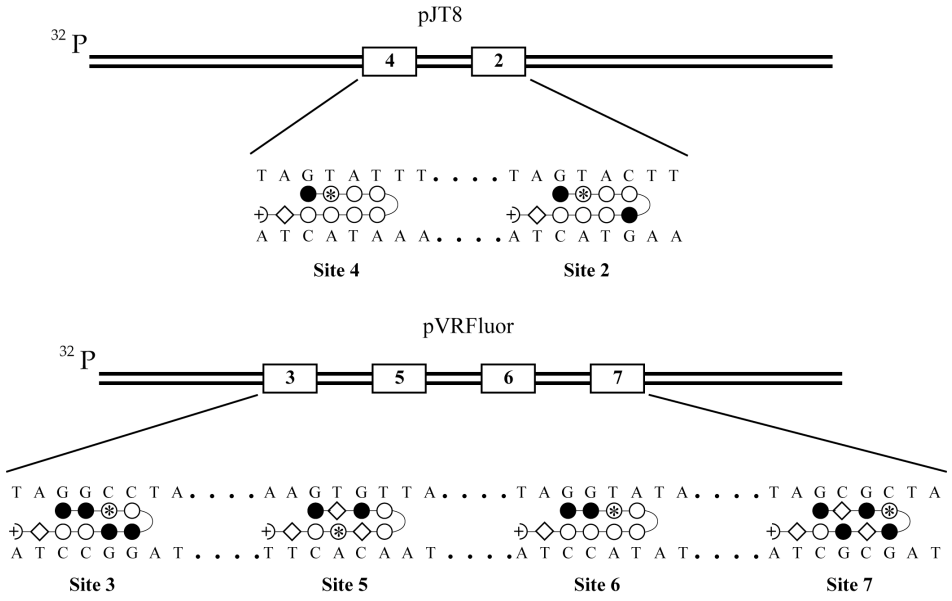


Figure 4.11 Binding sites on plasmids pJT8 and pVRfluor for characterizing **2** – **7**.

Quantitative DNAse I footprint titration assay confirms fluorescence

To determine if our fluorescence data truly represents the thermodynamic sequence preference of the polyamide DNA recognition domain, we turned to quantitative DNAse I footprinting to characterize the interaction of **2** - **7** with radiolabeled DNA containing the binding sites we employed in our parallel screening assays (Figures 4.5 and 4.10). Using a plasmid available from previous studies (pJT8),[31] and constructing a new plasmid (pVRfluor), DNAse I footprinting was conducted on all 6 compounds to characterize the thermodynamics of interaction with match site and mismatch binding sites (Figure 4.11). We find that each compound does preferentially recognize its match site with the highest affinity (Table 4.2A). Tabulating normalized values of the association constants in Table 4.2A, it is clear from Table 4.2B that DNAse I footprint titration confirms that the values reported in Table 4.1 represent the

Table 4.2A Equilibrium association constants for **2 - 7** (M^{-1}) as determined by DNase I footprinting.[a]

	2	**3**	**4**	**5**	**6**	**7**
taGGCCtt	$< 10^7$	**2.1×10^9**	$< 10^7$	2.0×10^7	$\leq 2.8 \times 10^7$	$< 10^7$
taGTATtt	5.2×10^7	$< 10^7$	**2.5×10^8**	$\leq 5.2 \times 10^8$	$< 10^7$	$\leq 5.4 \times 10^7$
taGTACtt	**1.7×10^9**	$< 10^7$	5.0×10^7	$\leq 4.2 \times 10^8$	3.2×10^8	$\leq 1.4 \times 10^8$
taGTGTtt	$< 10^7$	$< 10^7$	1.0×10^8	**1.4×10^9**	$\leq 4.5 \times 10^7$	$\leq 2.0 \times 10^8$
taGGTAtt	$< 10^7$	$< 10^7$	1.0×10^8	1.1×10^8	$\leq$ **9.7×10^8**	$< 10^7$
taGCGCtt	$< 10^7$	$< 10^7$	$< 10^7$	3.0×10^7	$\leq 8.7 \times 10^7$	**$\leq 4.9 \times 10^8$**

[a]The assays were carried out at 22°C at pH 7.0 in the presence of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

Table 4.2B Normalized equilibrium association constants for **2 - 7** as determined by DNase I footprinting.[a]

	2	**3**	**4**	**5**	**6**	**7**
taGGCCtt	< 0.01	**1**	< 0.05	0.01	≤ 0.03	< 0.02
taGTATtt	0.03	< 0.01	**1**	≤ 0.37	< 0.01	≤ 0.11
taGTACtt	**1**	< 0.01	0.20	≤ 0.30	0.33	≤ 0.29
taGTGTtt	< 0.01	< 0.01	0.20	**1**	≤ 0.05	≤ 0.41
taGGTAtt	< 0.01	< 0.01	0.20	0.08	**1**	< 0.02
taGCGCtt	< 0.01	< 0.01	< 0.05	0.02	≤ 0.09	**1**

[a]The assays were carried out at 22°C at pH 7.0 in the presence of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

fluorescent report of thermodynamic DNA discrimination through sequence specific polyamide binding in the minor groove.

Conclusions

The observation that the fluorescence of the fluorophores fluorescein and TMR is significantly diminished when attached by a short tether to an internal pyrrole in hairpin polyamide-TMR conjugates was not expected. We have found that other xanthene fluorophores, such as OregonGreen®,[32] when conjugated to an internal pyrrole of hairpin polyamides behave similarly. Presumably, the linker holds the fluorophore in very close

proximity[§] or in direct contact with the hairpin polyamide. After excitation, the polyamide-fluorophore decays from the excited state via a non-radiative pathway. Incorporation of the aliphatic β-alanine (β) residues lead to a decrease in fluorescence quenching (Figure 4.8). The more flexible aliphatic β may allow increased

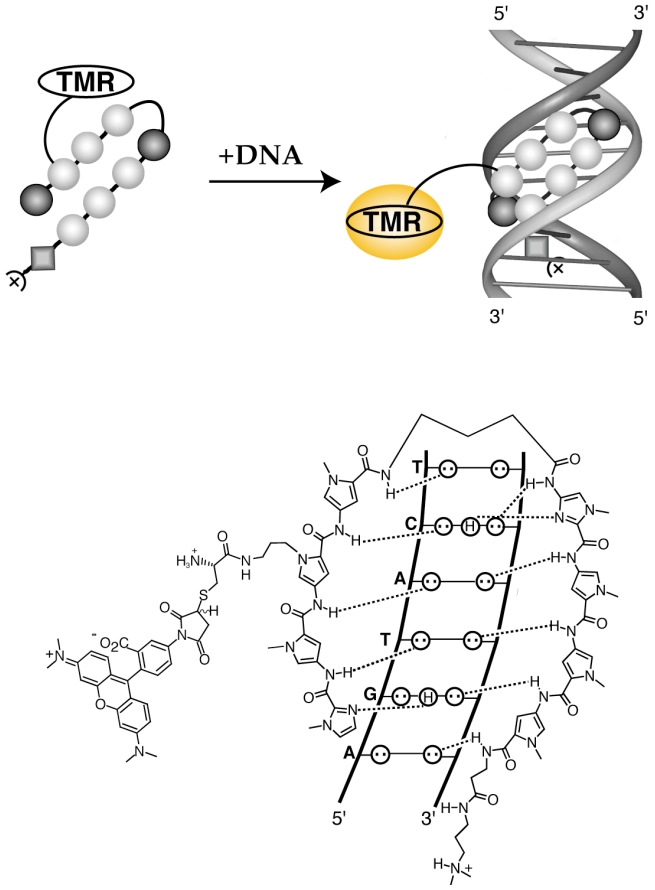


Figure 4.12 TMR-polyamide conjugates are fluorescently quenched in the absence of DNA containing a match recognition site. Upon binding, fluorescence is rescued, and the extent of biniding is dependent on the pairings, thus providing us with small molecule "light up probes" for the detection of DNA through sequence specific polyamide recognition of the minor groove.

[§] Conjugates without the β-amine of the cysteine, e.g., β-mercaptopropionic acid linker, behave in similar fashion to **2** – **7**.

conformational freedom to the conjugates **5** and **7**, possibly disrupting the quenching event between dye and polyamide.[33,34] In terms of the cause of fluorescence quenching, one can imagine a simple model encompassing two conformations possible for the polyamide-dye conjugate (Figure 4.12): polyamide and TMR are in close proximity when free in solution (quenching), and polyamide and TMR are separated in an extended conformation with TMR unable to contact the polyamide sequestered in the minor groove (fluorescing).

Both Laemmli and Trask have demonstrated the utility of polyamide-dye conjugates in the context of *gigabase size* DNA as sequence non-denaturing specific chromosomal repeat "paints".[26-28] This report extends the properties of polyamide-dye molecules for specific DNA sequence detection in solution *within short segments* of DNA. We suggest that an application for this science may be related to "SNP typing" in which the identity of the SNP containing sequence is known, and a proper means for

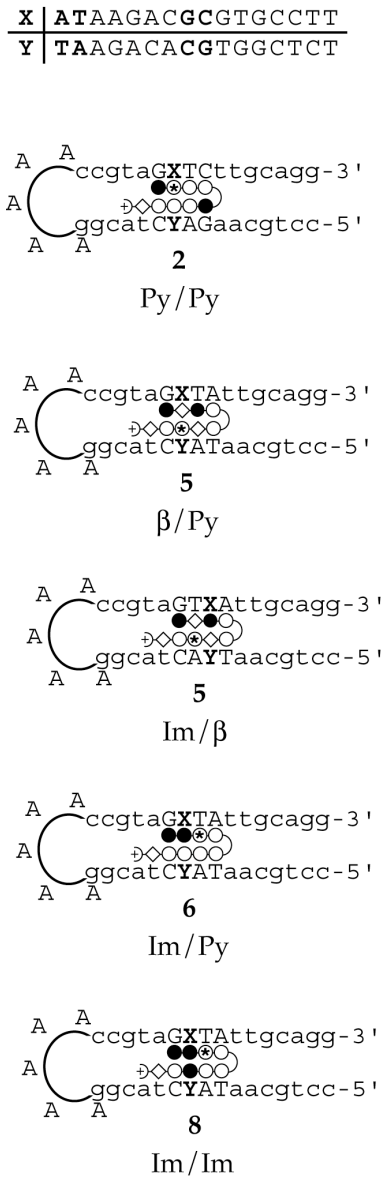
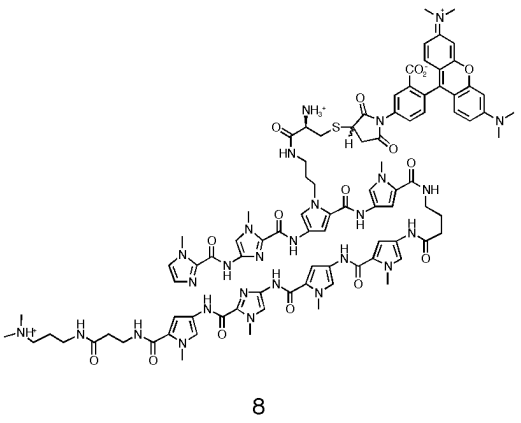


Figure 4.13 (Top) 16 possible X/Y pairings encompassing A, T, G, and C. (Lower) Design of hairpin forming oligos with polyamide binding site and X/Y variable position indicated.

SNP detection is the only obstacle to diagnosis. In a more immediate sense, we are happy to present a set of sequence specific DNA binding molecules in which recognition and detection can be tuned separately, i.e., the polyamide module is programmed for a desired DNA sequence by the pairing rules and the dye is chosen for optimal absorption/emission properties.



8

Figure 4.14 Structure of **8** containing novel Im/Im pairing.

Section 2

Polyamides as Molecular Calipers for Non-Watson-Crick Base Pairs in the DNA Minor Groove

Our fluorescent DNA detection method allows us to screen, in parallel, how a hairpin polyamide will behave when forced to bind over a non-Watson-Crick paired base. To date, the energetics of hairpin polyamide binding over non-Watson-Crick base pairs has not been examined, mainly due to the efforts required for the preparation of special end-radiolabeled DNA fragments containing discrete non-Watson-Crick

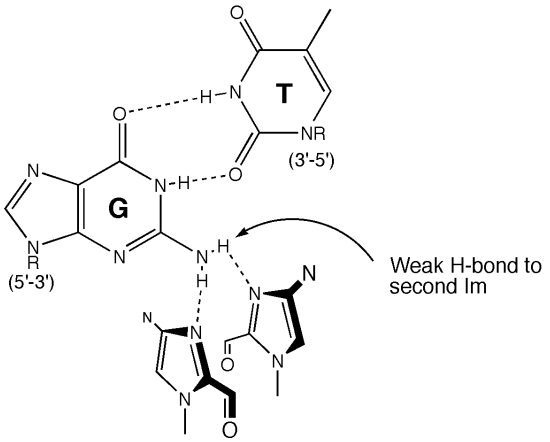


Figure 4.15 Sugiyama-Wang-Lee model for symmetric Im/Im pair recognition of G•T mismatch.

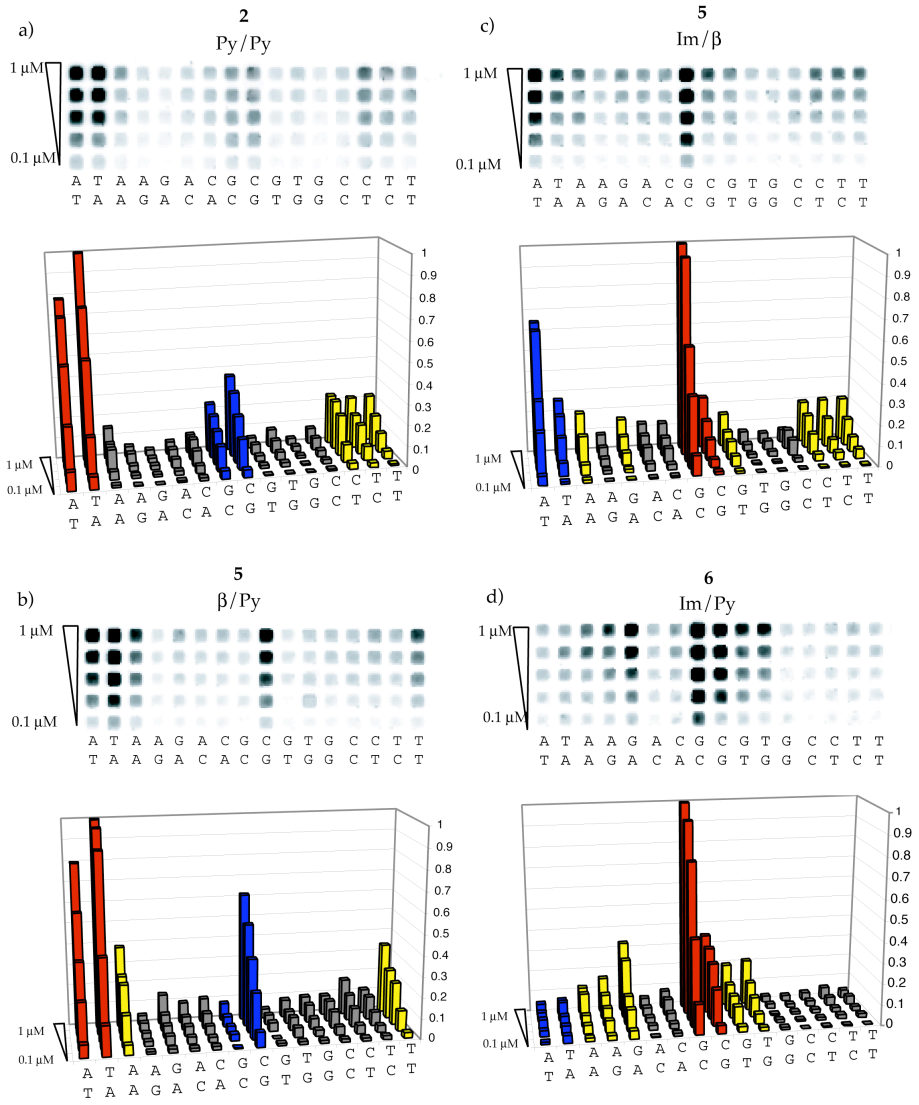


Figure 4.16 Screening of 1 μM **2**, **5**, and **6** against mismatch containing oligo. Oligo is titrated over concentration range 0.1 μM - 1 μM. Red bars are pairing rules matches, blue bars are pairing rules mismatches, yellow indicates binding over a Watson-Crick mismatch that is at least as significant as the binding we observe for placing a polyamide over a pairing rules mismatch.

base pairs for use in footprinting titration analysis. Three of the hairpin polyamide-fluorophore conjugates discussed in Section 1, **2**, **5,** and **6**, and a new conjugate **8** were used to screen the pairings Py/Py, β/Py, Im/β, Im/Py and Im/Im, respectively, against all 16 possible pairing permutations of the nucleic acid bases A, T, G, and C (Figure 4.13).

Note that screening of the □/Py and Im/□ pairing is achieved by screening **5** against 32 oligos, 2 separate sets of 16 under each position. Polyamide **8** (Figure 4.14) contains an unusual Im/Im ring pairing, previously shown by footprinting to be energetically unfavorable for recognition of any of the four Watson-Crick base pairings.[35] With regard to minor groove recognition of non-Watson-Crick paired bases, Wang and coworkers have reported an NMR structure for the polyamide dimer $(ImImImDp)_2$ that indicates binding with a minor groove site containing G•T and T•G base pairs under the Im/Im pair (Figure 4.15).[36] Im/Im recognition of a G•T and T•G wobble mismatch is the first example of ring pairs for specific recognition of non-Watson-Crick paired bases in the minor groove, though concentrations required for NMR preclude thermodynamic characterization of the interaction. There has also been a report of DAPI binding a minor groove recognition site containing a T•T mismatch with the O2 of thymidine accepting a hydrogen bond donated from an indole nitrogen.[37]

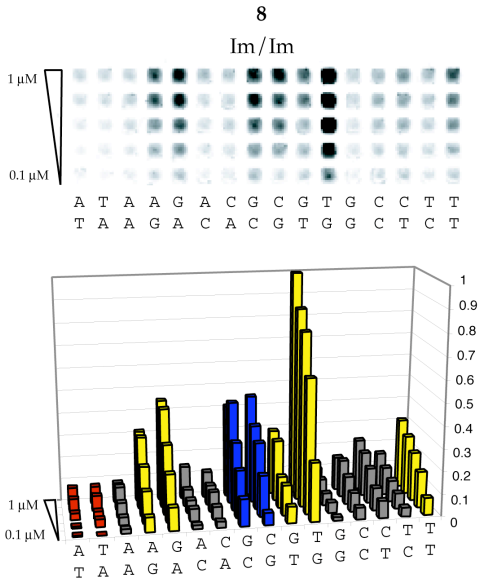


Figure 4.17 Fluorescence data collected for binding of 1 μM **8** over oligos containing all 16 possible mismatches.

Hairpin-forming oligos (37-mer) were synthesized and annealed (pH 7.0, 1X Tris-EDTA buffer). Each duplex was designed to allow for base pair variance at a single unique X/Y position within the context of a 6-base pair hairpin match site. Assays were conducted in 30 μL solutions with the concentration of polyamide conjugate fixed at 1

μM, and the concentration of oligonucleotide varied from 0.1 μM to 1 μM. After 4 hrs incubation (22°C), fluorescence intensities at the 16 variable positions were acquired by direct excitation in parallel on the optical scanner (16 X/Y position x 5 concentrations).

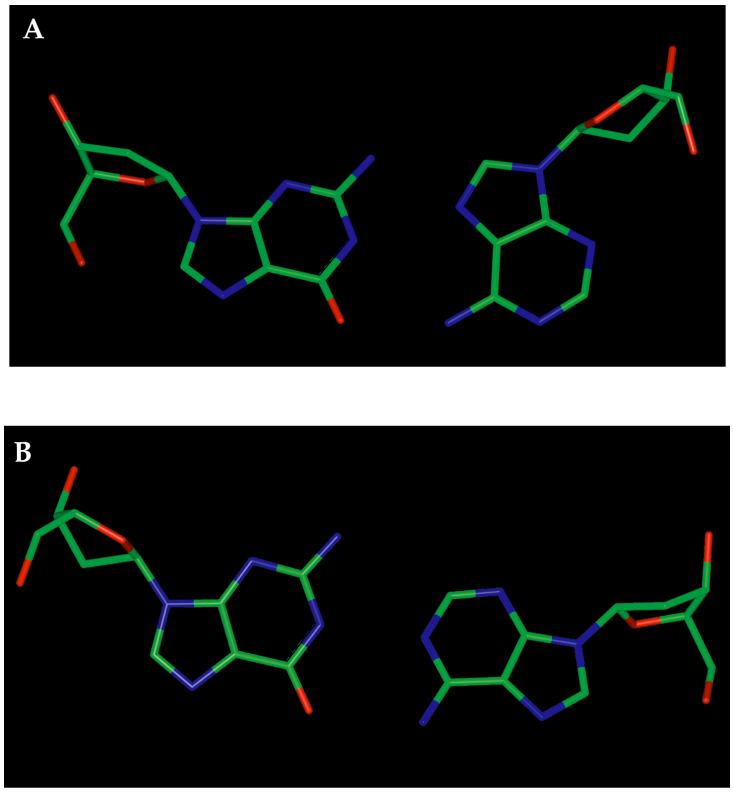


Figure 4.18 X-ray structure of G•A mismatch with A(*syn*)[39] or A(*anti*).[40] Note exocyclic amine of G preserved in minor groove as H-bond donor.

Figure 4.16 shows the collected data and normalized graphical presentation for the interactions of **2**, **5**, **6**, and **7** when forced to bind over a non-Watson-Crick base pair.

Figure 4.17 contains the data for polyamide **8**. Single mismatched bases in a large oligo are reported to induce only minor distortions from canonical B-form DNA.[37-47] Thus, we are not surprised to fluorescently measure polyamide binding over some mismatches. Aside from non-complementarity between hydrogen bond donors and acceptors on the edges of the base pairs and the polyamide, some mismatches are recognized with affinities bordering on those measured for binding over sites containing pairing rules mismatches (Figure 4.16, yellow bars). Take **5** binding to place the A•A mismatch under a /Py pairing. The A•A mismatch is quite stable, because the adenines are intercalated

into the stack of the helix,[50] thus we are observing disfavored binding not because of an instable mismatch (as is the case with C•C or A•C mismatches).[38] Note the binding (K_a ~ 3.9 x 10^8 M^{-1}) by **6** at a site containing G•A under Im/Py, possibly the result of the exocyclic amine of guanine preserved in the minor groove to donate a hydrogen bond to imidazole (Figure 4.18). It is difficult to say, though, if the polyamide is binding over the G(*anti*)•A(*anti*) or G(*anti*)•A(*syn*) conformation as both have been documented.[39,40] As a "control," it is comforting to note that we observe, for all compounds in Figure 4.16, strongest binding at the pairing rules match recognition site. Polyamide **8** with the Im/Im pair shows a slight preference for G•C over A/T (Figure 4.17), but does not discriminate G•C from C•G as expected for a symmetrical ring pair. Remarkably, the Im/Im pair binds T•G with the highest affinity validating the hypothesis that the exocyclic amine of guanine is available to donate a hydrogen bond to an appropriately placed acceptor (Figure 4.19).[36,46,47] In our

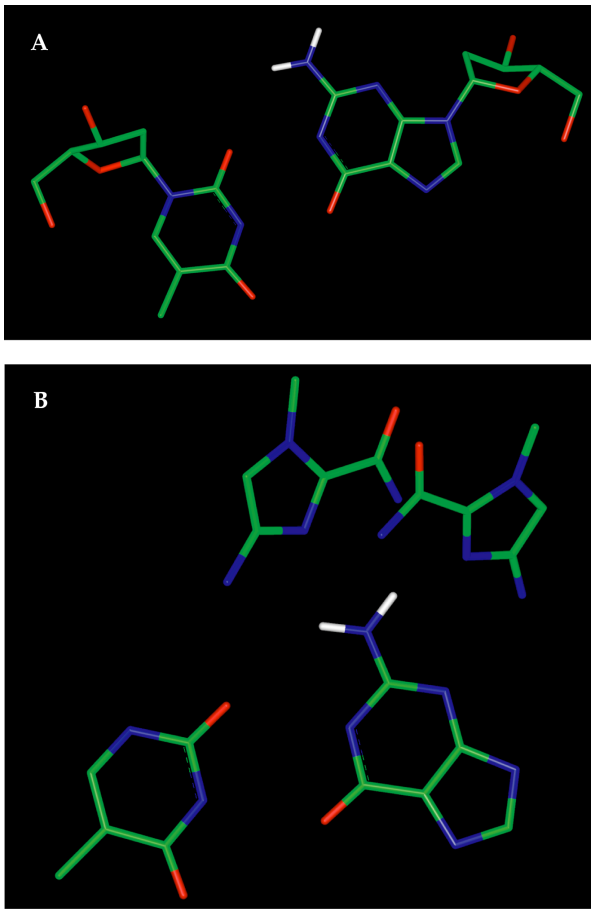


Figure 4.19 A. X-ray structure of unliganded G•T mismatch.[46,47] B. NMR structure of G•T mismatch determined with $(ImImImDp)_2$ ligand bound in minor groove. In both cases, note exocyclic amine is preserved to donate a hydrogen bond to an appropriately placed acceptor.

study, however, Im/Im distinguishes T•G from G•T, a result not anticipated in the NMR structure, which may be due to different sequences used in the two studies. Further studies may address G•T and T•G asymmetry through use of 3-hydroxypyrrole/Im pairings to bind the asymmetric cleft of the G•T mismatch. Regardless, it is noteworthy that a minor groove binding polyamide can target *a single non-Watson-Crick base pair* out of 12 unique possibilities.

In summary, polyamide-fluorophore conjugates have shown promise as stains for sequence repeats in chromosomes for the interrogation of repeat location and estimation of repeat length.[26-28] In this section of Chapter 4, we extend the use of polyamide-fluorophore conjugates to distinguish match and mismatch sequences within short segments of DNA in solution. The correlation of sequence identity with fluorescence enhancement may be useful in other applications such as screening for sequence variation in specific segments of genomic DNA.

Experimental Section

Materials

UV spectra were measured in water on a Beckman model DU 7400 spectrophotometer. MALDI-TOF mass spectrometry data was collected by the Protein and Peptide Microanalytical Facility at the California Institute of Technology. HPLC analysis was performed using a Beckman Gold Nouveau system using a Rainin C_{18}, Micosorb MV, 5 μm, 300 x 4.6 mm reversed phase column using 0.1% (wt/v) TFA with acetonitrile as eluent and a flow rate of 1.0 mL/min, gradient elution 1.25% acetonitrile/min. Preparatory reversed phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 x 100 mm, 100 μm C_{18} column equipped with a guard, 0.1% (wt/v) TFA/ 0.25% acetonitrile/min. Water was from either a Millipore MilliQ water purification system or RNase free water from USB. All buffers were 0.2 μm filtered prior to storage. Oligonucleotides were synthesized and purified by the Caltech Biopolymer Synthesis and Analysis Resource Center and Genbase Inc., San Diego. Enzymes for molecular biology were purchased from either New England Biolabs or Boeringer-Mannheim. [α-^{32}P] adenosine triphosphate and [α-^{32}P] thymidine triphosphate were purchased from New England Nuclear. L-Boc(Trt)-Cys-OH was purchased from Bachem. Tetramethyl rhodamine-5-maleimide was from Molecular Probes. Polystyrene plates were from VWR Scientific. Fluorescence spectrophotometer measurements were made at room temperature on an ISS K2 spectrophotometer employing 5 nm emission and excitation slits in conjunction with a Hg lamp. Polystyrene plate measurements were made on a Molecular Dynamics Typhoon employing a 532 nm 10-20 mW excitation

laser. The emission filter employed is 580±15 nm filter for observation of tetramethyl rhodamine fluorescence.

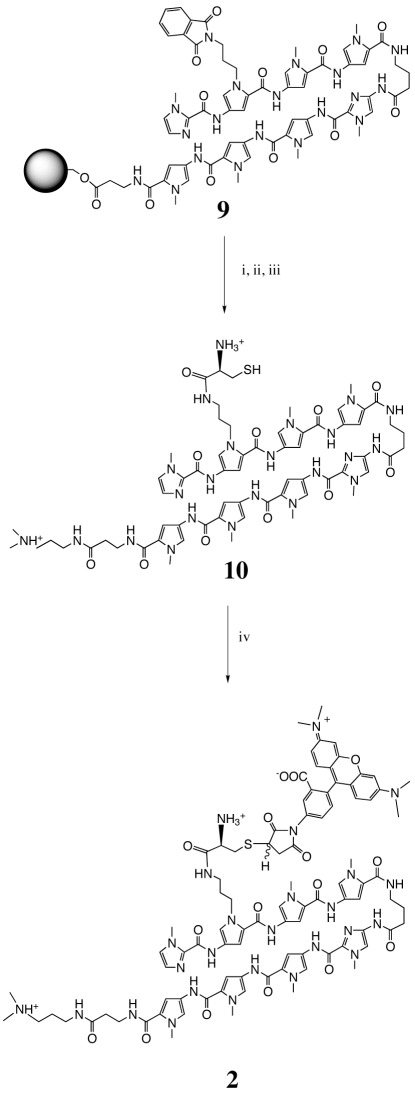


Figure 4.20 Intermediate polyamide **9** is synthesized on Boc-β-Ala PAM resin using standard protocols. i. Nitrogen deprotection and polyamide liberation accomplished by aminolysis with Dp, purification. ii. Boc-Cys-(Trt)-OH, 1 equivalent, purification iii. TFA, triethyl silane, purification. iv. 1 eq. TMR-5-maleimide in neat DMF, purfication.

Synthesis of polyamide-fluorophore conjugates

Polyamides **1**-**8** were synthesized by solid phase methods (Figure 4.20). A new monomer, 4-[(*t*-butoxycarbonyl)amino]-1-(phthalimidopropyl)pyrrole-2-carboxylic acid (prepared by Shane Foister), to provide a primary amine for fluorophore attachment from the pyrrole back of the polyamide, was incorporated in the usual manner by DCC/HOBt activation (DMF, 4 hrs., 37 ºC). Polyamide **9** was simultaneously deprotected and cleaved from the resin by aminolysis with 3-(dimethylamino)-propylamine (60 ºC, 10 hrs.). Purification by reversed phase HPLC afforded a diamine which was allowed to react with one equivalent of L-Boc(Trt)-Cys-OH activated with DCC/HOBt. The polyamide-cysteine conjugate **10** was characterized by analytical HPLC and MALDI-TOF mass

spectrometry. The L-cysteine modified polyamides in DMF were allowed to react with one equivalent of TMR-5-maleimide in a minimal volume of DMF. This reaction was monitored by analytical HPLC and was usually complete in less than 15 mins. The products were purified by reversed phase preparatory HPLC. All products were analyzed by MALDI/TOF MS. For **1** (monoisotopic) [M + H] 1795.7 (1795.69 calc'd for $C_{87}H_{94}N_{24}O_{18}S^{+}$); **2** (monoisotopic) [M + H] 1853.6 (1853.8 calc'd for $C_{89}H_{105}N_{28}O_{16}S$); **3** (monoisotopic) [M + H] 1850.54 (1850.81 calc'd for $C_{92}H_{108}N_{25}O_{16}S$); **4** (monoisotopic) [M + H] 1851.74 (1851.81 calc'd for $C_{91}H_{107}N_{26}O_{16}S$); **5** (monoisotopic) [M + H] 1749.9 (1749.79 calc'd for $C_{85}H_{105}N_{24}O_{16}S$); **6** (monoisotopic) [M + H] 1851.8 (1851.81 calc'd for $C_{91}H_{107}N_{26}O_{16}S$); **7** (monoisotopic) [M + H] 1751.6 (1751.78 calc'd for $C_{83}H_{103}N_{26}O_{16}S$); **8** (monoisotopic) [M + H] 1852.7 (1852.63 calc'd for $C_{90}H_{106}N_{27}O_{16}S$). For **1** - **4**, **6, 8** UV (H_2O) λ_{max}(ε) 312 (68800), **5** and **7**, UV (H_2O) λ_{max}(ε) 300 (51600), **1** UV (H_2O) λ_{max}(ε) 494 (~71000, pH 7.3), **2** - **8**, Vis (H_2O) λ_{max}(ε) 563 (65700).

Optical characterization

All measurements in Section 1were performed in TKMC buffer {10 mM Tris-HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$}. Those in Section 2 in 1X TE buffer as described in the text of Section 2. The concentration of polyamide-TMR conjugate was 1 μM and the volume of solution used for each measurement was 150 μL. In the fluorimeter, polyamide **1** was excited at 475 nm with emission collected over the interval 505 nm to 620 nm. Polyamides **2**-**7** were excited at 545 nm and measured over the interval of 570 nm to 680 nm. Emission maxima are 578 nm for **2**-**8**, 515 nm for **1**. 1 μM solutions of each conjugate were irradiated in the presence and absence of 1 μM of

the appropriate match DNA to generate the fluorescence enhancements reported in Figure 4.8.

96-well plate characterization

In the dark, 150 μL solutions were made by titrating together a fixed concentration of 1 μM conjugate against 20 nM to 1 μM of each DNA 17-mer. The solutions were gently swirled for a few mins and then allowed to sit for four hrs before measurements were made. No changes in fluorescence intensity were observed for longer equilibration times. Plates containing polyamides **1**-**8** were excited at 532 nm, and data were collected at 526 nm with a short-pass filter (**1**) or a 580 (± 15 nm FWHM) nm filter (**2**-**8**). ImageQuant software (Molecular Dynamics) was used to analyze the fluorescence intensity of each titration experiment. From this data, histograms were constructed to assist in determining DNA sequence preference for each conjugate. The normalized data used in each histogram is polyamide specific assuming the background to be the lowest observable fluorescence intensity, which is subtracted in the denominator from the highest observed fluorescence intensity, that of the match sequence at 1:1 conjugate:duplex as given by the relation $(F_{obs}-F_{min})/(F_{max}-F_{min})$.

Footprinting and plasmid construction

DNAse I footprinting titrations were conducted as previously described.[48] Oligos 5' GAT CCG TCC CTT AGG CCT ATG GTC CAC GTT AGT GTT ATG GTC CAC GTT AGG TAT ATG GTC CAC GTT AGC GCT ATG GCCA 3' and 5' AGC TTG GGC CTA TGC GCT AAC GTG GAC CAT ATA CCT AAC GTG GAC CAT AAC

ACT TAC GTG GAC CAT AGG CCT AAG GGA CG 3' were annealed and ligated into *BamHI*/*HindIII* cut sites of pUC19. Blue/white selection protocols and Promega Wizard MidiPreps were used for plasmid purification. A_{260} absorption values of 50 μg/ml were used to quantitate yield.

References

1. Wang, D. G.; Fan, J.-B; Siao, C.-J.; Berno, A.; Young, P.; Sapolsky, R.; Ghandour, G.; Perkins, N.; Winchester, E.; Spencer, J.; Kruglyak, L; Stein, L.; Hsie, L.; Topaloglou, T.; Hubbell, E.; Robinson, E.; Mittmann, M.; Morris, M. S.; Shen, N.; Kilburn, D.; Rioux, J.; Nusbaum, C.; Rozen, S.; Hudson, T. J.; Lipshutz, R.; Chee, M.; Lander, E. S. *Science* **1998**, *280*, 1077-1082.
2. Hacia, J. G. *Nature Genetics Suppl.* **1999**, *21*, 42-47.
3. Ihara, T.; Chikaura, Y.; Tanaka, S.; Jyo, A. *Chem. Commun.* **2002**, 2152-2153.
4. Frutos, A. G.; Pal, S.; Quesada, M.; Lahiri, J. *J. Am. Chem. Soc.* **2002**, *124*, 2396 -2397.
5. Kukita, Y.; Higasa, K.; Baba, S.; Nakamura, M.; Manago, S.; Suzuki, A.; Tahira, T.; Hayashi, K. *Electrophoresis* **2002**, *23*, 2259-2266.
6. Brazill, S. A.; Kuhr, W. G. *Anal. Chem.* **2002**, *74*, 3421-3428.
7. Yamashita, K.; Takagi, M.; Kondo, H., Takenaka, S. *Anal. Biochem.* **2002**, *306*, 188-196.
8. Whitcombe, D.; Theaker, J.; Guy, S. P.; Brown, T.; Little, S. *Nat. Biotechnol.* **1999**, *17*, 804-807.
9. Thelwell, N. Millington, S.; Solinas, A.; Booth, J. A.; Brown, T. *Nucleic Acids Res.* **2000**, *28*, 3752-3761.
10. Holland, P. M.; Abramson, R. D.; Watson, R.; Gelfand, D. H. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 7276-7280.
11. Woo, T. H. S.; Patel, B. K. C.; Smythe, L. D.; Norris, M. A.; Symonds, M. L.; Dohnt, M. F. *Anal. Biochem.* **1998**, *256*, 132-134.

12. Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996**, *14*, 303-308.

13. Tyagi, S.; Bratu, D. P.; Kramer, F. R. *Nat. Biotechnol.* **1998**, *16*, 49-53.

14. Kostrikis, L. G.; Tyagi, S.; Mhlanga, M. M.; Ho, D. D.; Kramer, F. R. *Science* **1998**, *279*, 1228-1229.

15. Jenkins, Y.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 8736-8738.

16. Rajir, S. B.; Robles, J.; Wiederholt, K.; Kiumelis, R. G.; McLaughlin, L. W. *J. Org. Chem.* **1997**, *62*, 523-529.

17. Ranasinghe, R. T.; Brown, L. J.; Brown, T. *Chem. Comm.* **2001**, 1480-1481.

18. Sando, S.; Kool, E. T. *J. Am. Chem. Soc.* **2001**, *124*, 2096-2097.

19. For non-sequence specific fluorescent dyes, such as ethidium, thiazole orange and oxozole yellow dimers, see (19-20): Rye, H. S.; Yue, S.; Wemmer, D. E.; Quesada, M. A.; Haughland, R. P; Mathies, R. A.; Glazer, A. N. *Nucleic Acids Res.* **1992**, *20*, 2803-2812.

20. Cosa, G.; Focsaneanu, K. S.; McLean, J. R. N.; McNamee, J. P.; Scaiano, J. C. *Photochem. & Photobiol.* **2001**, *73*, 585-599.

21. Seifert, J. L.; Connor, R. E.; Kushon, S. A.; Wang, M.; Armitage, B. A. *J. Am. Chem. Soc.* **1999**, *121*, 2987-2995.

22. Wang, M.; Silva, G. L.; Armitage, B. A. *J. Am. Chem. Soc.* **2000**, *122*, 9977-9986.

23. For distamycin-Hoechst hybrids, see (23-25): Satz, A. L.; Bruice, T. C. *Bioorg. & Med. Chem.* **2000**, *8(8)*, 1871-1880.

24. Satz, A. L.; Bruice, T. C. *J. Am. Chem. Soc.* **2001**, *123(11)*, 2469-2477.

25. Satz, A. L.; Bruice, T. C. *Bioorg. & Med. Chem.* **2002**, *10*, 241-252.

26. Janssen, S.; Durussel, T.; Laemmli, U. K. *Mol. Cell* **2000**, *6(5)*, 999-1011.

27. Maeshima, K.; Janssen, S., Laemmli, U. K. *EMBO J*. **2001**, *20*, 3218-3228.

28. Gygi, M. P.; Ferguson, M. D.; Mefford, H. C.; Kevin, P. L.; O'Day, C.; Zhou, P.; Friedman, C.; van den Engh, G.; Stobwitz, M. L.; Trask, B. J. *Nucleic Acids Res*. **2002**, *30*, 2790-2799.

29. Sjöback, R.; Nygren, J.; Kubista, M. *Spectrochim. Acta A* **1995**, *51*, L7-L21.

30. Dervan, P. B. *Bioorg. & Med. Chem*. **2001**, *9*, 2215-2235.

31. Trauger, J.W.; Baird, E.E.; Dervan, P.B. *Nature* **1996**, *382*, 559-561.

32. Chevillet, J. Thesis, *California Institute of Technology*.

33. Johansson, M. K.; Fidder, H.; Dick, D.; Cook, R. M. *J. Am. Chem. Soc*. **2002**, *124*, 6950-6956;

34. Rabinowitch, E.; Epstein, L. F. *J. Am. Chem. Soc*. **1941**, *63*, 69-78.

35. White, S.; Baird, E. E.; Dervan, P. B. *Chem. & Biol*., **1997**, *4*, 569-578.

36. Yang, X.-L; Hubbard IV, R. B.; Lee, M.; Tao, Z.-F.; Sugiyama, H.; Wang, A. H.-J. *Nucleic Acids Res*. **1999**, *27*, 4183-4190.

37. Trotta, E.; Paci, M. *Nucl. Acids Res*. **1998**, *26*, 4706-4713.

38. Peyret, N.; Seneviratne, P. A.; Allawi, H.T.; SantaLucia, J., Jr. *Biochem.* **1999**, *38*, 3468-3477.

39. Prive, G.G.; Heinemann, U.; Chandrasegaran, S.; Kan, L.S.; Kopka, M.L.; Dickerson, R.E. *Science* **1987**, *238*, 498-504.

40. Hunter, W.N.; Brown, T.; Kennard, O. *J. Mol. Struct. Dyn*. **1986**, *4*, 173-191.

41. Kennard, O.; *Biochem. Soc. Trans*. **1986**, *14*, 207-210.

42. Arnold, F.H.; Wolk, S.; Cruz, P.; Tinoco, I., Jr. *Biochem.* **1987**, *26*, 4068-4075.

43. Cornelis, A.G.; Haasnoot, H.J.; den Hartog, J.F.; de Rooij, M.; van Boom, J. H.; Cornelis, A. *Nature* **1979**, *281*, 235-236.

44. Kouchakdjian, M.; Li, B.F.L.; Swann, D.F.; Patel, D.J. *J. Mol. Biol*. **1988**, *202*,

139-155.

45. Hunter, W.N.; Brown, T.; Anand, N.N.; Kennard, O. *Nature* **1986**, *320*, 555-552.

46. Kneale, G.; Brown, T.; Kennard, O. *J. Mol. Biol.* **1985**, *186*, 805-814.

47. Rabinovich, D.; Haran, T.; Eisenstein, M.; Shakked, Z. **1988**, *200*, 151-161.

48. Trauger, J.W.; Dervan, P. B. *Methods Enzymol.* **2001**, *340*, 450-466.

49. Rucker, V.C.; Byers, L.D. *J. Am. Chem. Soc.* **2000**, *112*, 8365-8369.

50. Boon, E.M.; Kisko, J.L.; Barton, J.K. *Methods Enz.* **2002**, *353*, 506-522.