

Targeting Repeat Sequences with DNA-Binding Small Molecules

Thesis by

John Richard Chevillet

In Partial Fulfillment of the Requirements

for the Degree of

Master of Science

California Institute of Technology

Pasadena, California

2002

Abstract

Recognition of repeat sequences in DNA would have applications in molecular biology. One of the most biologically interesting repeat sequences is the telomeric repeat which composes the ends of eukaryotic chromosomes; in vertebrates 5'-TTAGGG-3'. This sequence has been used as a model to study how DNA-binding polyamide molecules composed of pyrrole (Py) and imidazole (Im) residues bind to repeating sequences. DNase I footprinting shows that the polyamide-fluorophore conjugate ImImImPy- γ -PyPy((CH₂)₃N,N',N''trimethylbis (hexamethylene)triamine-OregonGreen488) PyPy- β -Me can bind the sequence 5'-AGGGTT-3' $K_a = 1.8 \times 10^8 \text{ M}^{-1}$. Quantitative fluorescence titrations with varying patterns of telomeric repeat suggest that the molecule can tolerate another polyamide binding contiguously, but not two. Truncation of the tail of the conjugate to yield the molecule ImImImPy- γ -PyPy((CH₂)₃N,N',N''trimethylbis (hexamethylene)triamine-OregonGreen488) PyPy-Me allows the compound to bind three contiguous sites, suggesting that steric polyamide-polyamide interactions control binding in this manner.

Table of Contents

| | |
|--|-------|
| Abstract | iii. |
| Table of Contents | iv. |
| List of Figures and Tables | v-vi. |
| I. Introduction | 1 |
| II. Background and Significance | 3 |
| III. Project Aims | 4 |
| IV. Project Design and Implementation | 5 |
| Experimental | 34 |
| References | 44 |

List of Figures and Tables

| | |
|--|----|
| Figure 1: Watson-Crick base pairs | 1 |
| Figure 2: Polyamide pairing Rules | 3 |
| Figure 3: Polyamide design motifs | 6 |
| Figure 4: Ball and stick models of repeat binding | 7 |
| Figure 5: Synthesis of Compound 1 | 8 |
| Figure 6: Synthesis of Compound 2 | 9 |
| Figure 7: Plasmid pJC01 design | 10 |
| Figure 8: DNase footprint of Compound 3 | 11 |
| Figure 9: Synthesis of Compound 6 | 14 |
| Figure 10: DNase footprint of Compound 6 | 16 |
| Figure 11: Fluorescence titration helices | 17 |
| Figure 12: Synthesis of Compound 9 | 19 |
| Figure 13: Fluorescence spectrum of Compound 6 | 20 |
| Figure 14: Fluorescence titration of Compound 6 | 21 |
| Figure 15: Linear region of Compound 6 fluorescence titration | 21 |
| Figure 16: Synthesis of Compound 13 | 24 |
| Figure 17: Binding isotherms for 13 on pJC01 | 25 |

| | |
|---|----|
| Figure 18: DNase footprint of Compound 13 | 26 |
| Figure 19: MPE footprints of Compounds 6 and 13 | 27 |
| Figure 20: Fluorescence titration helices with H5 | 28 |
| Figure 21: Fluorescence image of 96-well plates | 29 |
| Figure 22: Fluorescence titration of Compound 6 | 30 |
| Figure 23: Fluorescence titration of Compound 13 | 30 |
| Figure 24: Linear region of Compound 6 fluorescence titration | 31 |
| Figure 25: Linear region of Compound 13 fluorescence titration | 31 |
| | |
| Table 1: Least squares regression analysis of Compound 6 titration | 22 |
| Table 2: Least squares regression analysis of Compound 6 titration with H5 | 32 |
| Table 3: Least squares regression analysis of Compound 13 titration | 32 |

I. Introduction

The biologically most relevant form of DNA is B-DNA, which displays a wide, shallow major groove and a deep, narrow minor groove.¹ Sequence specific recognition of duplex DNA by other molecules is governed by the H-bonding patterns that they can form with the bases. The major groove has conventionally been thought to be a better candidate for specific recognition as it is more accessible and presents a characteristic H-bond pattern for all four Watson-Crick base pairs (Figure 1).² However, collaborative studies in the Dervan group illustrated it is possible to recognize all four base pairs in the minor groove via a polyamide molecule composed of pyrrole (Py), imidazole (Im), and hydroxypyrrole amino acids (Figure 2).³

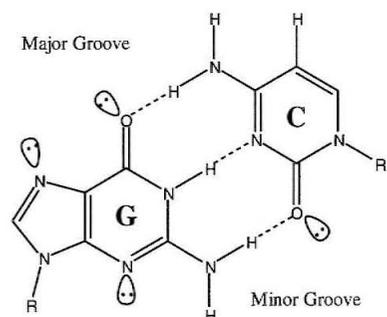
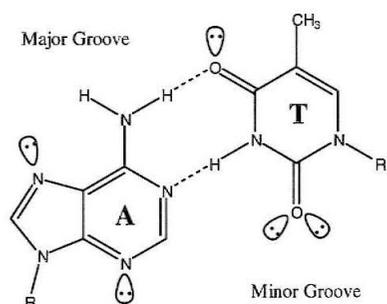


Figure 1: Watson-Crick base pairs showing major and minor groove H-bonding patterns

DNA binding polyamide molecules have been shown to be cell permeable and have high affinity and specificity for DNA (comparable to DNA binding proteins).⁴⁻⁸ The residues pair side by side in the minor groove; an Im•Py pairing specifies for G•C, while a Py•Im pairing is specific for C•G. A Py•Py pairing is degenerate for A•T and T•A. This degeneracy can be removed by substitution of a hydroxypyrrole residue.^{3,9} An Hp•Py pair is specific for T•A and a Py•Hp pair specifies A•T. The

design of polyamides is limited to five contiguous rings, as an increase in this number of residues does not affect binding affinity and decreases specificity.¹⁰ This limitation is overcome by the insertion of a flexible aliphatic linker, β -alanine. This substitution adds conformational flexibility and β - β pairings in the minor groove have been found specify for A•T and T•A.¹¹⁻¹³

The ability to sequence specifically recognize DNA by artificial small molecules gives rise to many potential applications in molecular biology and medicine. Investigation in this area has been ongoing. In collaborative studies, hairpin polyamides have been shown to inhibit RNA Polymerase II transcription.¹⁴ In addition, conjugate molecules utilizing a polyamide as a DNA recognition domain, a linker moiety, and a peptide activation domain have been shown to activate transcription *in vitro*.¹⁵

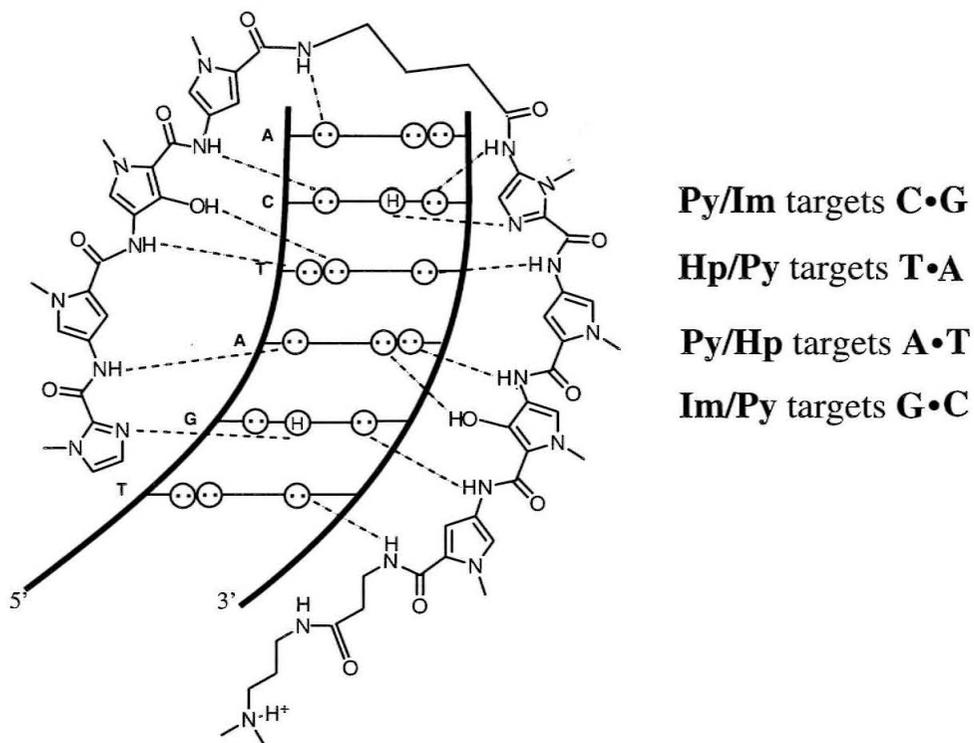


Figure 2: Schematic representation of the polyamide ImPyHpPy- γ -ImPyHpPy- β -Dp recognizing 5'-TGATCA-3'

II. Project Background and Significance

Recognition and detection of repeat sequences in DNA would have applications in chromosome painting and genomic analysis. One of the most biologically interesting repeat sequences is the telomeric repeat which composes the termini of eukaryotic chromosomes.¹⁶ In vertebrates, it follows the hexameric pattern 5'-TTAGGG-3'.¹⁷ The bulk of the region is double stranded, however the extreme 3' terminus displays a 12-16 nucleotide single stranded overhang.^{18,19} The length of the telomere is maintained at

approximately 15kb in germline cells, but is variable in somatic cells and decreases with each cell division,^{18,20-22} but not necessarily with chronological age in a mechanism referred to as the "mitotic clock."²³ The telomeres shorten until they trigger a p53 dependent check point arrest,²⁴ causing the cell to enter a non-dividing state known as senescence. This state is thought to prevent the ends of the chromosomes from being recognized as double strand breaks, leading to repair by fusion, genomic instability, and apoptosis.²⁵

Senescence has been implicated in aging and its associated pathologies.²⁶⁻²⁸ Rapidly dividing cancer cells escape this state by activation of telomerase, the enzyme responsible for telomere synthesis (normally inactive in somatic cells) and thus immortalize.^{29,30} Thus, a probe of telomere length would have significant applications in studies of cancer and aging.³¹

III. Project Aims

The goals of this project are:

1. To design and synthesize polyamides that bind repeating DNA sequences.
2. To discern what factors influence their mode of binding.

IV. Project Design and Implementation

The telomeric repeat will serve as the experimental model to study how polyamides bind repeating sequences. Toward satisfaction of Project Aim 1, two polyamide motifs have been designed to target the telomeric repeat (Figure 3). The first employs the formation of a side-by-side heterodimer couple targeting 11bp of the repeat: 5'-AGGGTTAGGGT-3'. This will be achieved through the pairing rule complimentary polyamides ImImImPy- β -PyImImIm- β -Dp (**1**) and PyPyPyPy- β -PyPyPyPy- β -Dp (**2**) (Dp: dimethylaminopropylamine). The second targets the six base-pair sequence 5'-AGGGTT-3' by way of the eight-ring hairpin polyamide ImImImPy- γ -PyPyPyPy- β -Dp (**3**) (γ : γ -aminobutyric acid) that has been previously shown to bind 5'-AGGGAA-3' with a K_a of $3.7 \times 10^8 \text{ M}^{-1}$ and to have good specificity over mismatch sequences.³²

Hypothesized modes of binding for these polyamides (PAs) are illustrated in Figure 4. It is proposed that each design will adopt a preferred binding state, either saturating the available binding sites or by binding alternating sites. The mode of binding will be dependent upon polyamide-polyamide interactions and polyamide-DNA interactions.

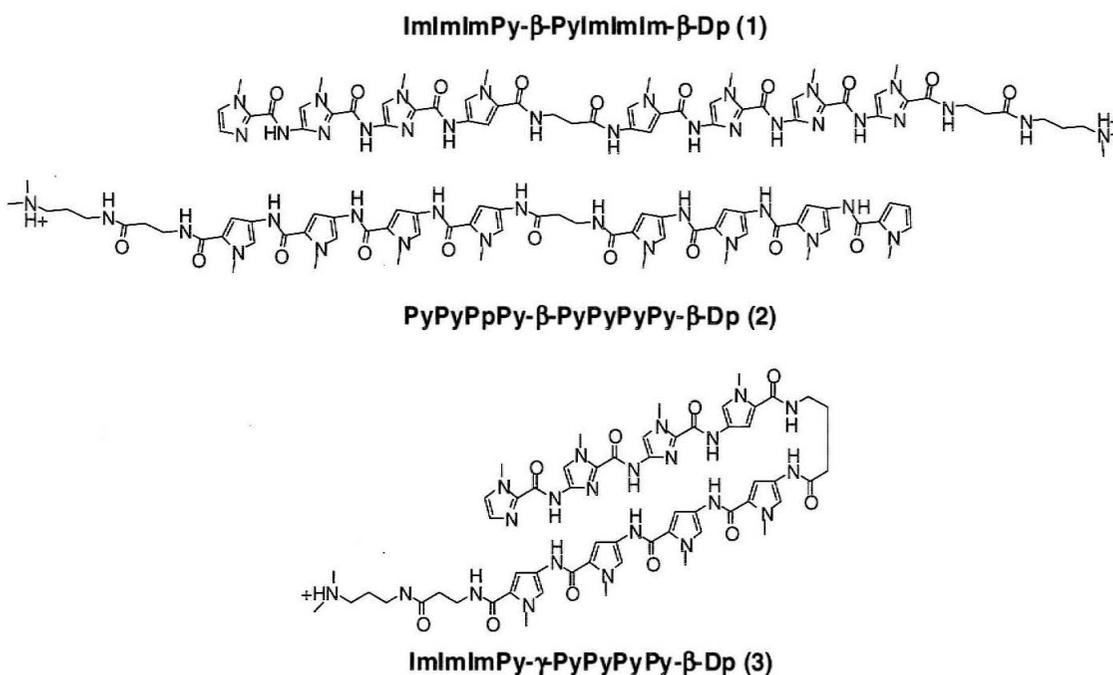


Figure 3: Heterodimeric polyamide couple ImImImPy- β -PyImImIm- β -Dp, PyPyPyPy- β -PyPyPyPy- β -Dp and the hairpin ImImImPy- γ -PyPyPyPy- β -Dp

Examples of PA-PA interactions include the potential cation-cation repulsion in the case of the heterodimer (which may be negated by phosphate contacts) as well as tail-tail steric interactions, and the possible steric clash of the γ -aminobutyric turn residue of one hairpin with the Dp tail of the following hairpin. PA-DNA interactions include the distortion of the helix by polyamide binding which may be beneficial or detrimental to a subsequent binding event.

Following synthesis and characterization for DNA binding affinity, experiments are performed on polyamide-fluorophore conjugates to study their mode of binding contiguous match sites by quantitative fluorescence spectroscopy.

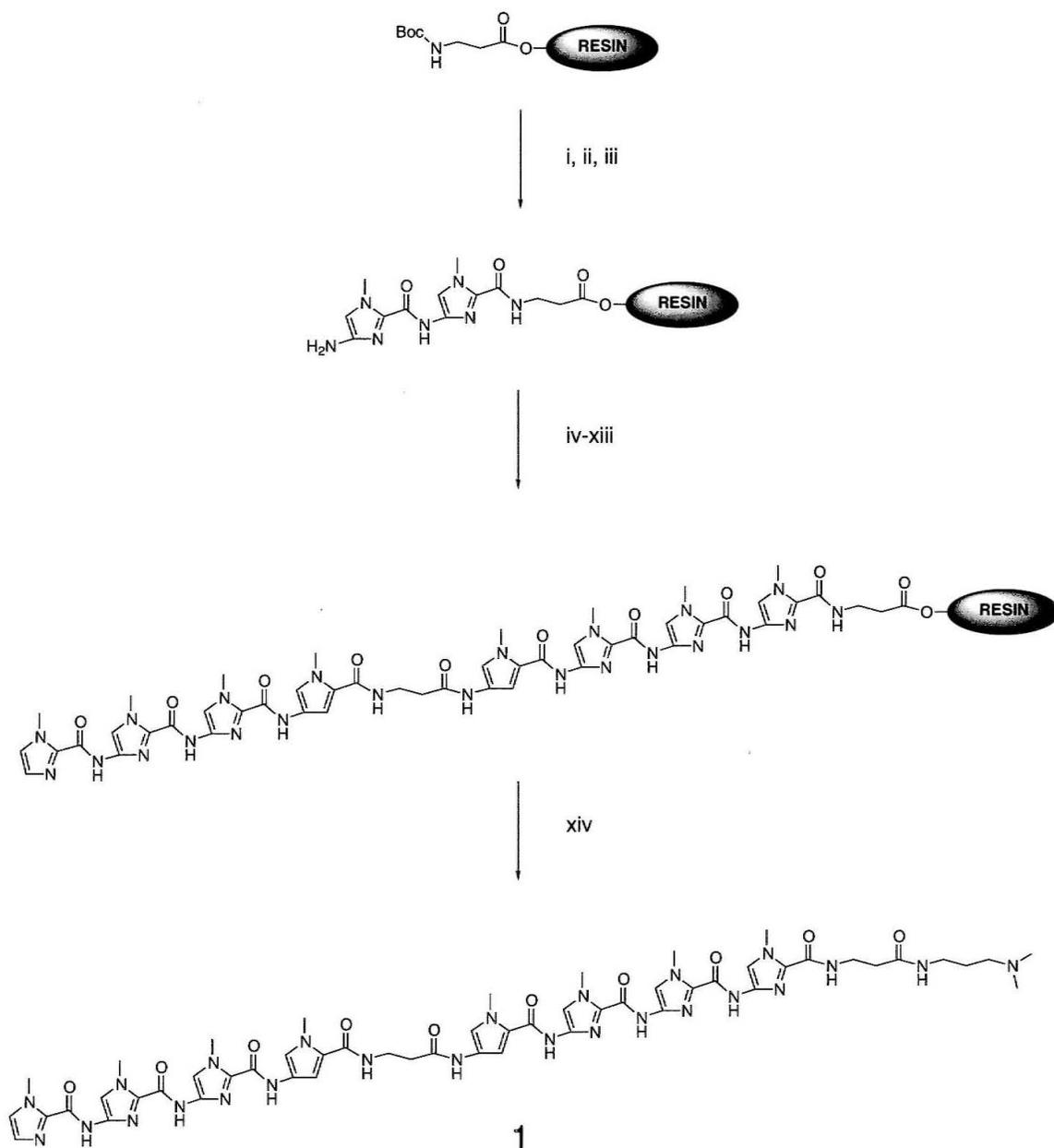


Figure 5: Synthetic outline for ImImImPyβPylmImβDp. (i) 80%TFA/DCM/0.5M PhSH; (ii) BocImImCOOH, DCC, HOBt, DMF, DIEA; (iii) 80%TFA/DCM/0.5M PhSH; (iv) BocPylmCOOH, DCC, HOBt, DMF, DIEA; (v) 80%TFA/DCM/0.5M PhSH; (vi) Boc-β-Ala, DCC, HOBt, DMF, DIEA; (vii) 80%TFA/DCM/0.5M PhSH; (viii) BocPyObt, DMF, DIEA; (ix) 80%TFA/DCM/0.5M PhSH; (x) BocImImCOOH, DCC, HOBt, DMF, DIEA;; (xi) 80%TFA/DCM/0.5M PhSH; (xii) 80%TFA/DCM/0.5M PhSH; (xiii) ImCOOH, DCC, DMF, DIEA; (xiv) dimethylaminopropylamine, 37°C, 18hr

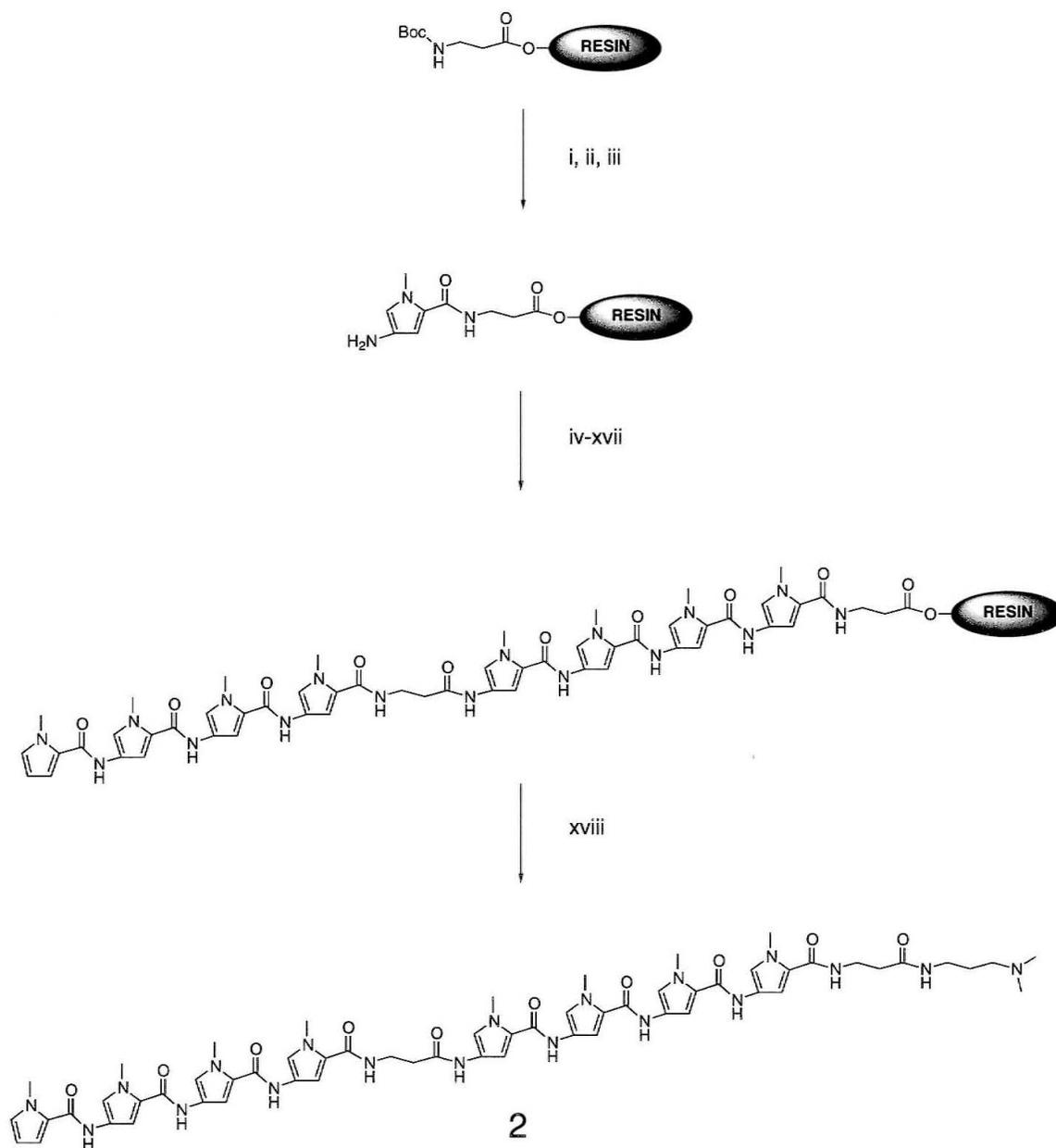


Figure 6: Synthetic outline for PyPyPyPybPyPyPyPybDp. (i) 80%TFA/DCM; (ii) BocPyOBt, DIEA, NMP; (iii) 80%TFA/DCM; (iv) BocPyOBt, DIEA, NMP; (v) 80%TFA/DCM; (vi) BocPyOBt, DIEA, NMP; (vii) 80%TFA/DCM; (viii) BocPyOBt, DIEA, NMP; (ix) Boc- β -Ala, HBTU, DIEA, NMP; (x) 80%TFA/DCM; (xi) BocPyOBt, DIEA, NMP; (xii) 80%TFA/DCM; (xiii) BocPyOBt, DIEA, NMP; (xiv) 80%TFA/DCM; (xv) BocPyOBt, DIEA, NMP; (xvi) 80%TFA/DCM; (xvii) 1-methyl-2-pyrrole carboxylic acid, DCC, HOBT, DIEA, NMP; (xviii) dimethylaminopropylamine, 37°C, 12hr

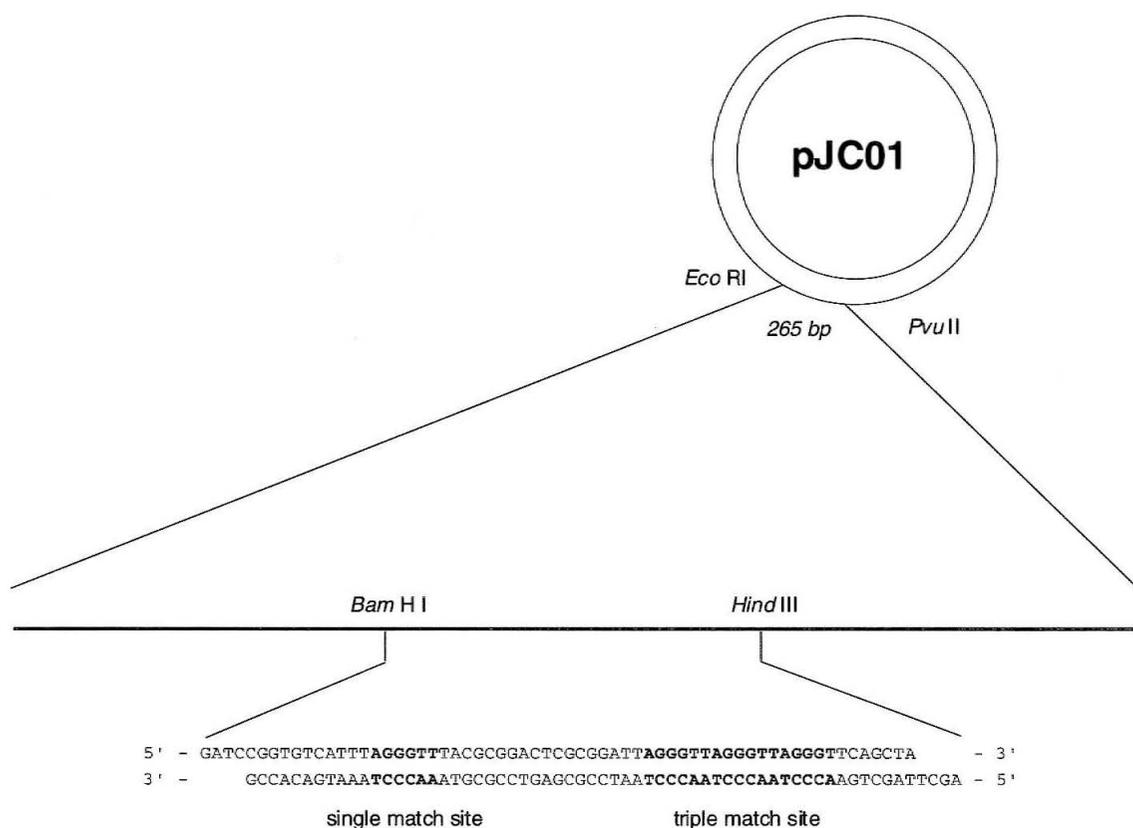


Figure 7: Oligonucleotide insert into pUC19 to create plasmid pJC01 showing single and triple binding sites

Quantitative DNase I Footprinting Titrations

To evaluate equilibrium binding affinity for the designed match site and discern the ability of the polyamide to bind match sites in a contiguous repeating sequence (Project Aim 1), plasmid pJC01 was constructed containing a single match site for (3) and three contiguous match sites (Figure 7). Quantitative DNase I footprinting titrations were performed according to the method of Trauger and Dervan.³⁴ The hairpin polyamide

bound the match site with a K_a of $7.2 \times 10^8 \text{ M}^{-1}$ (Figure 8). A footprint was observed over the whole triple site region. An average quantitative analysis over the triple site yielded a $K_a = 9.2 \times 10^8 \text{ M}^{-1}$ indicating that the polyamides could bind that area with comparable affinity but unknown mode of binding.

Design and Preparation of Fluorescent Conjugate

To maintain binding affinity (as it is usually reduced upon conjugation of a polyamide to another molecule), a rational redesign of the molecule was implemented. Exchanging the positions of the Dp tail and a pyrrole N-methyl group has been shown to yield a 10-fold increase in binding affinity with no loss of specificity,³⁵ thus the cationic charge was relocated to the third pyrrole residue. The dye chosen was commercially available Oregon Green 488 (Molecular Probes), which has an excitation/emission spectrum almost identical to fluorescein, but higher photostability. The linker design reflects the need for a tertiary amine on the alkyl pyrrole residue, a nucleophile for dye conjugation, and for the finished conjugate to have a net cationic charge, aiding in solubility.

The synthesis of **6** (Figure 9) was accomplished beginning with solid phase preparation utilizing the N-functionalized monomer 3-hydroxypropyl-4-[(tert-butoxycarbonyl)amino]-pyrrole-2-carboxylic acid.³⁵ Cleavage from the solid support

resulted in ImImImPy- γ -PyPy((CH₂)₃OH)PyPy- β -Me (**4**). The crude material was extracted, lyophilized, and activated by toluenesulfonyl chloride in pyridine. Precipitation by ether and nucleophilic displacement by N,N',N''-trimethylbis(hexamethylene)triamine at 37°C gave the free-amine compound (**5**), purified by HPLC. Reaction in dimethylformamide/diisopropylethylamine with the succinimidyl ester of Oregon Green 488 resulted in **6**, subsequently purified by HPLC.

Quantitative DNase I Footprinting Titrations of Fluorescent Conjugate

DNase I footprinting studies were performed with **6** as above (Figure 10). The conjugate bound the match sequence with a $K_a = 1.8 \times 10^8 \text{M}^{-1}$, averaged over the triple site with $K_a = 2.2 \times 10^8 \text{M}^{-1}$, and bound the match site proximal to the 5' ³²P label within the triple site $K_a = 2.2 \times 10^8 \text{M}^{-1}$ (others not reliably quantifiable due to 3' shift in footprint as an artifact of DNase cleavage). Unlike **3**, the footprint of this molecule in the triple match region showed cleavage bands in the center site. This evidence contradicts binding in a saturation mode.

Fluorescence Titration Experiment

It was discovered that various designs of polyamide-dye conjugates exhibited fluorescence when in the presence of match DNA, but were quenched over an order of magnitude when alone in solution.³⁶ The fluorescence was found to be dependent on DNA binding as dictated by polyamide pairing rules; fluorescence in the presence of

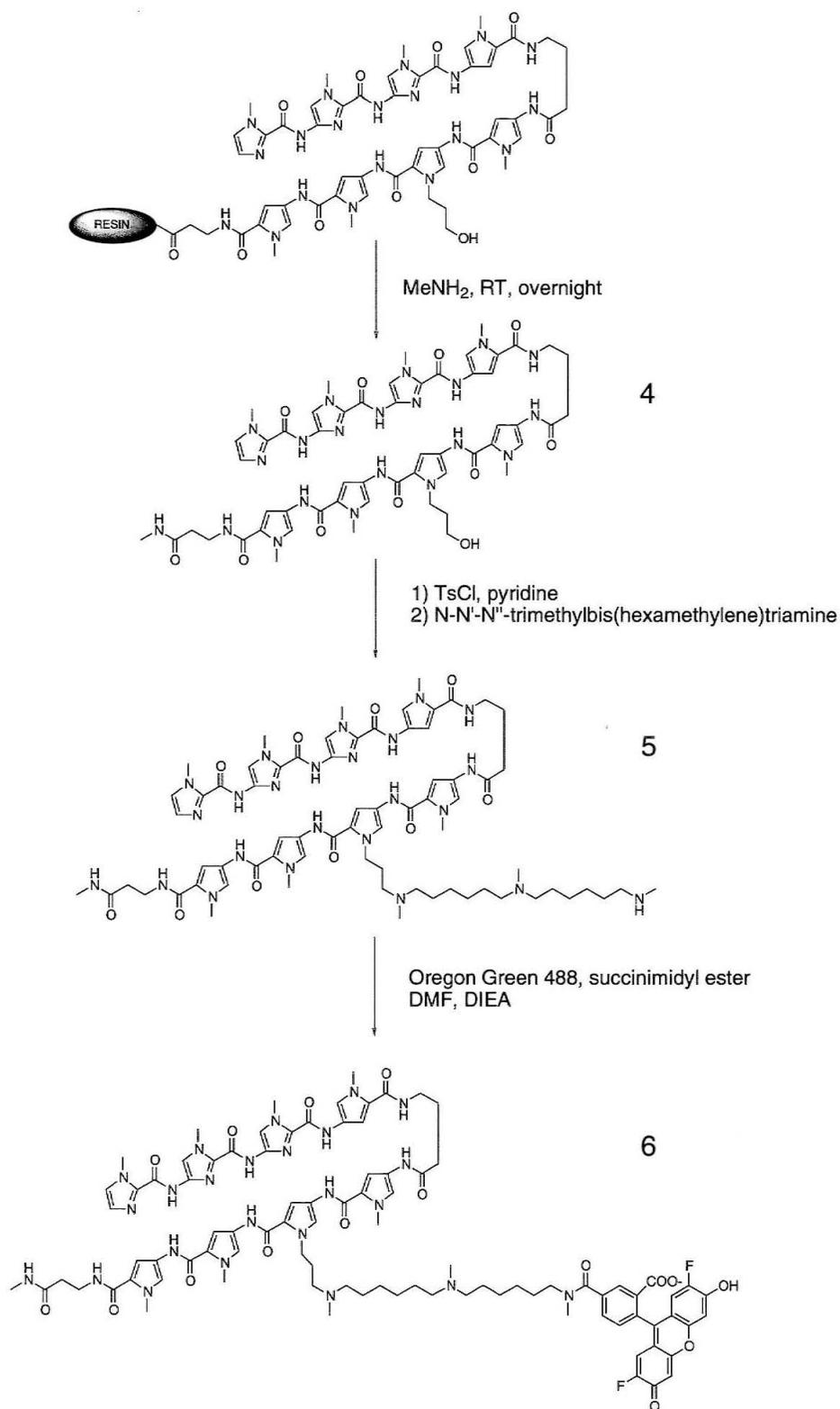


Figure 9: Synthesis of fluorescent conjugate (6)

mismatch DNA was correspondingly lower. Investigation is proceeding into the mechanism of this phenomenon.

This discovery presented the opportunity to determine mode of binding by fluorescence increase. A series of DNA helices were designed containing varying degrees and patterns of match sites (Figure 11). Helix 1 (**H1**) contained a single match site, Helix 2 (**H2**) a contiguous double, Helix 3 (**H3**) a contiguous triple, and Helix 4 (**H4**) two match sites separated by six base pairs that constituted a double mismatch. It was hypothesized that titrating a given helix into a fixed concentration of polyamide conjugate would result in an increase in fluorescence that would indicate mode of binding by rate of fluorescence increase over the titration. Titrating **H1** into a fixed concentration of polyamide-dye conjugate (PA-dye) was expected to result in a linear increase of fluorescence until the concentration of **H1** was equal to that of the PA-dye. The PA-dye could then be considered titrated and there would be maximal fluorescence. Beyond this point, no increase in fluorescence should occur. Fluorescence of the PA-dye when titrated with **H2** was expected to be dependent on how the PA bound the contiguous double site. If the sites could be saturated (two PAs bound) fluorescence should increase over the titration at double the rate of **H1** and reach maximal fluorescence when the concentration of DNA is half that of the PA-dye. If the sites are bound by alternation,

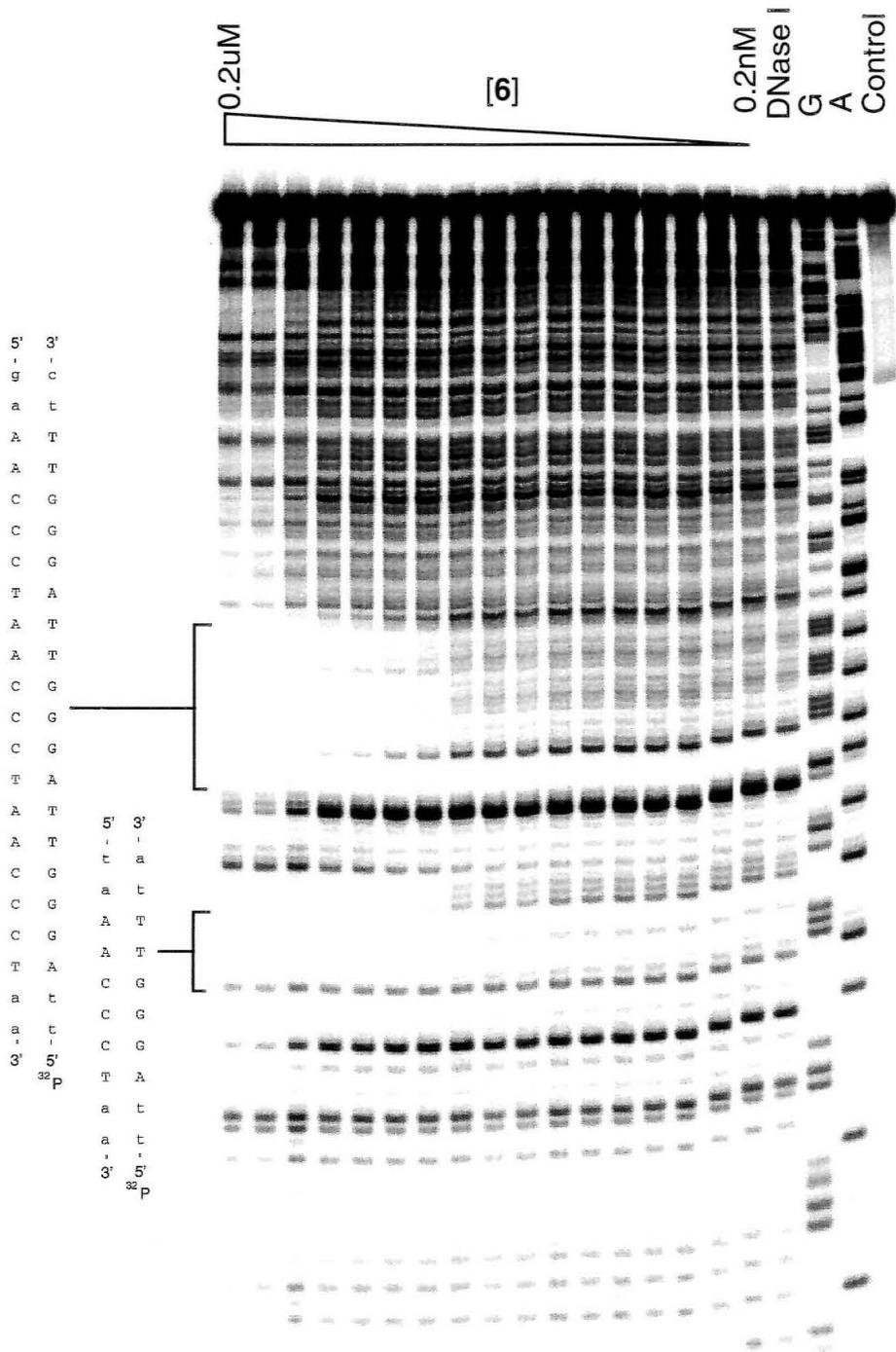


Figure 10: DNase I footprinting gel of 6 on 5' ³²P labeled pJCO1. Control lane represents and intact fragment. A and G represent sequencing lanes for these bases. D represents DNase control lane, no polyamide. Other lanes increase from 0.2nM to 0.2µM as shown.

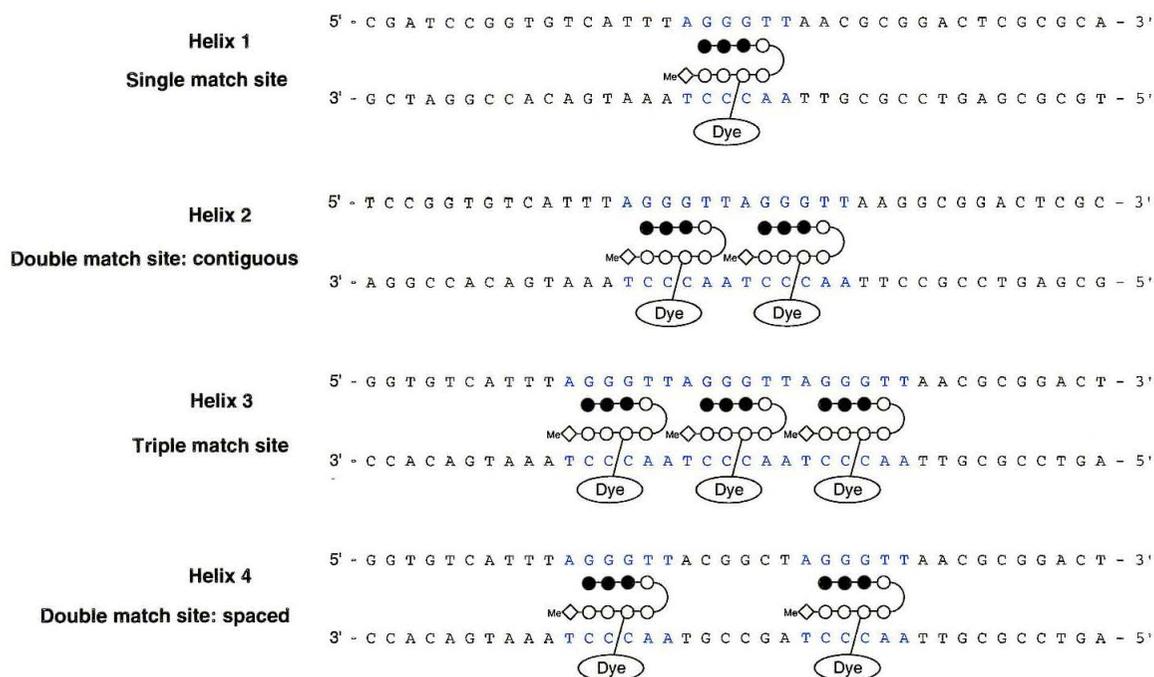


Figure 11: Schematic of quantitative fluorescence experiment. Ball and stick polyamides are shown as they would bind by saturation.

the data should resemble that of **H1** as one polyamide binds the contiguous double site.

In the case of **H3**, data for saturation should show a fluorescence increase over the titration at three times the rate of **H1** and maximal fluorescence when $[DNA] = 1/3[PA-dye]$. If the helix is bound by alternation, the data should reflect two polyamides binding.

H4 was designed as a control comparison and to determine if PA-DNA interactions could effect a PA binding 6bp away (as there should be no PA-PA interactions). If both sites were occupied, then the data should indicate two PAs binding.

In the case of one PA binding, the data should resemble that from the **H1** titration.

An additional PA was designed utilizing a cysteine linker and fluorescein coupled by maleimide conjugation to imitate the structure of molecules that had shown this effect. **9** was prepared by stepwise solid phase synthesis incorporating a phthalimide protected amino-pyrrole in place of the previous $\text{Py}((\text{CH}_2)_3\text{OH})$ residue (Figure 12). Cleavage from the solid support by methyl amine gave the deprotected, methyl tail polyamide **7**. Subsequent reaction with Boc-Cys-Trt-OBt in DMF/DIEA and linker deprotection with trifluoroacetic acid and triethylsilane results in the amine/thiol polyamide **8**, which was purified by HPLC. Reaction of the thiol with fluorescein 5-maleimide in the presence of sodium bicarbonate and an reducing agent yielded **9**, purified by HPLC. However, this molecule displayed poor solubility and thus was not further characterized for its DNA binding or fluorescent properties.

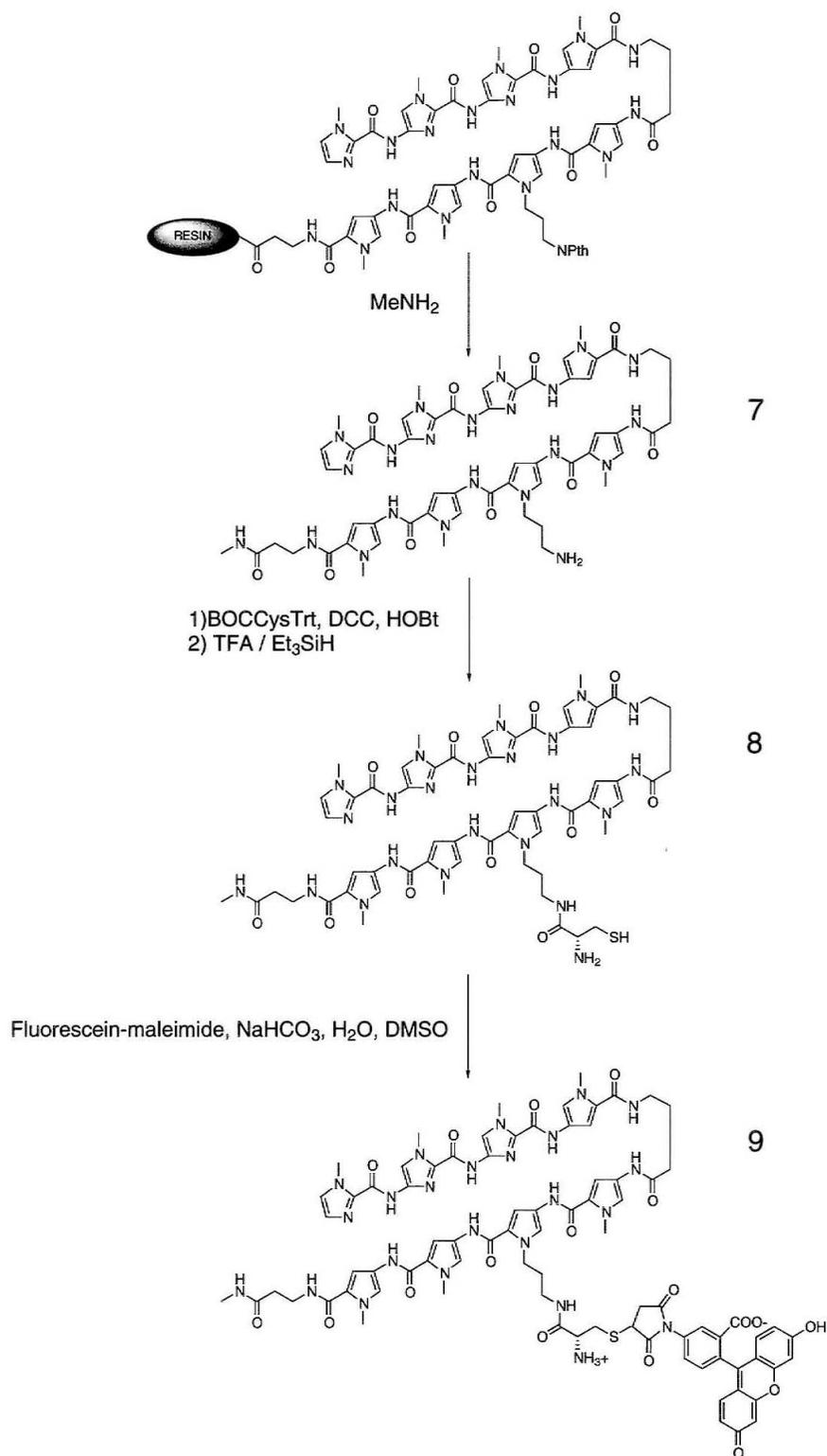


Figure 12: Synthesis of fluorescent conjugate 9

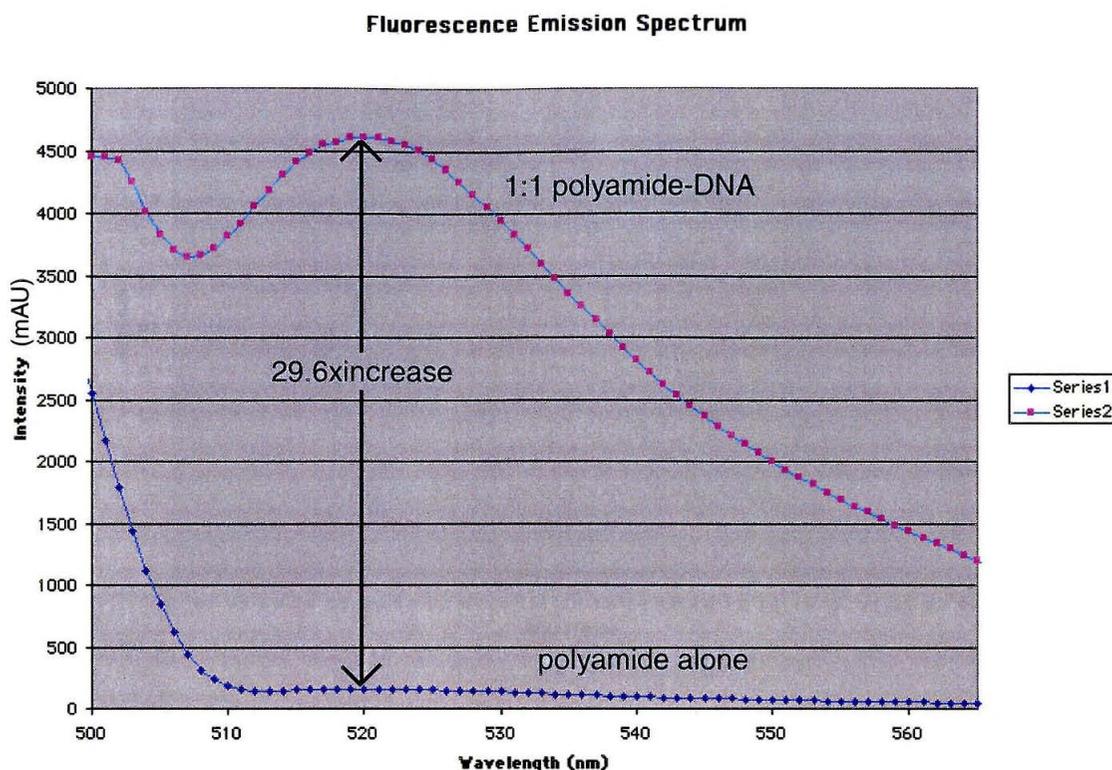
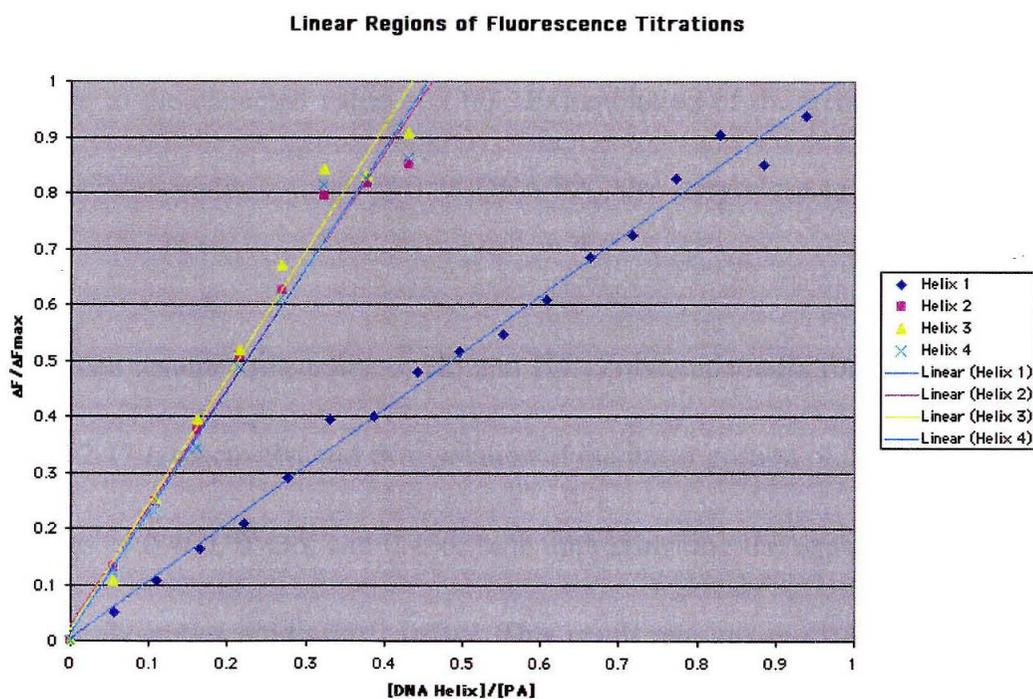
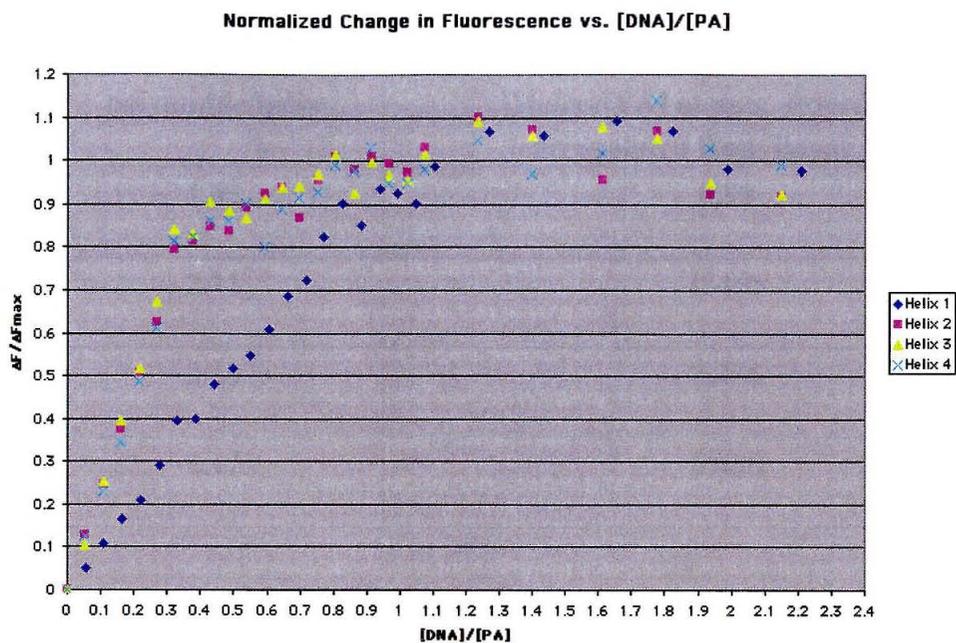


Figure 13: Fluorescence spectrum of compound 6 alone in aqueous solution and in the presence of equimolar match DNA

The previous conjugate (**6**) exhibited a 29.6-fold increase in fluorescence in the 1:1 molar presence of single match site DNA as compared to free in solution (Figure 14) and it was selected for use in the titration experiment.

A 96-well plate assay was developed where the concentration of polyamide conjugate was held constant and titrated against increasing concentration of helix. Buffer conditions and equilibration times were identical to footprinting experiments. The transparent plates were analyzed for fluorescence by scanning on an imaging



| Helix | polyamide/helix | [Helix]/[PA] at max. fluorescence (extrapolated from linear trend) |
|--------------|------------------------|---|
| 1 | 1.02 | 0.98 |
| 2 | 2.11 | 0.460 |
| 3 | 2.25 | 0.436 |
| 4 | 2.17 | 0.456 |

Table 1: Least squares regression analysis of quantitative fluorescence titration. Reported numbers are the average of two experiments.

system, and the individual well images quantified for fluorescent intensity. The results of the experiments are shown in Figures 14 and 15. values listed in Table 1. The data for **H1** supports the predicted 1 PA:1 Helix binding mode for this duplex. Polyamide/helix = 1.02, very close to the expected value of 1.00. Extrapolation of the trend to maximal fluorescence increase yields a value of 0.98 DNA/PA, also very close to the predicted value of 1.00.

H2 yielded results comparable to **H3** and **H4**. Polyamide/helix for H2-H4 were 2.11, 2.25, and 2.17 respectively, and extrapolation of the linear portion of the data shows DNA/PA values of 0.460, 0.436, and 0.456, each approximating the expected result of 0.500 for two polyamides binding per helix. This result requires re-evaluation of the

mode of binding by alternation, as the **H2** data should have more closely resembled the **H1** data and not **H3** and **H4** if this were the case.

Instead the data suggest that it is possible to bind two contiguous polyamides if no other match sites are available, but alternation is preferred when more sites are available. It is hypothesized that this is the result of the ability of the conjugate to tolerate one proximal polyamide binding (as in the case of H2) but not two (as in the case of H3).

Conclusions

The methyl tail polyamide **6** binds the match sequence 5'-AGGGTT-3' with $K_a = 1.8 \times 10^8 \text{ M}^{-1}$. The manner that it binds contiguous repeats of this sequence is dependent on their presentation. Two contiguous match sites are bound by two polyamides, saturating the available sites. Three contiguous match sites are bound by two polyamides occupying either contiguous or alternating sites.

Conjugate Re-design

As **6** was not able to saturate the triple binding site, molecular design was reconsidered. It was hypothesized that omission of the β -alanine residue in the polyamide tail would relieve steric interference between two polyamides attempting to bind in a contiguous manner. Polyamide **10** (Fig. 16) was synthesized in a stepwise manner using reaction conditions suitable to the oxime resin support.³⁵ Upon cleavage with neat methyl

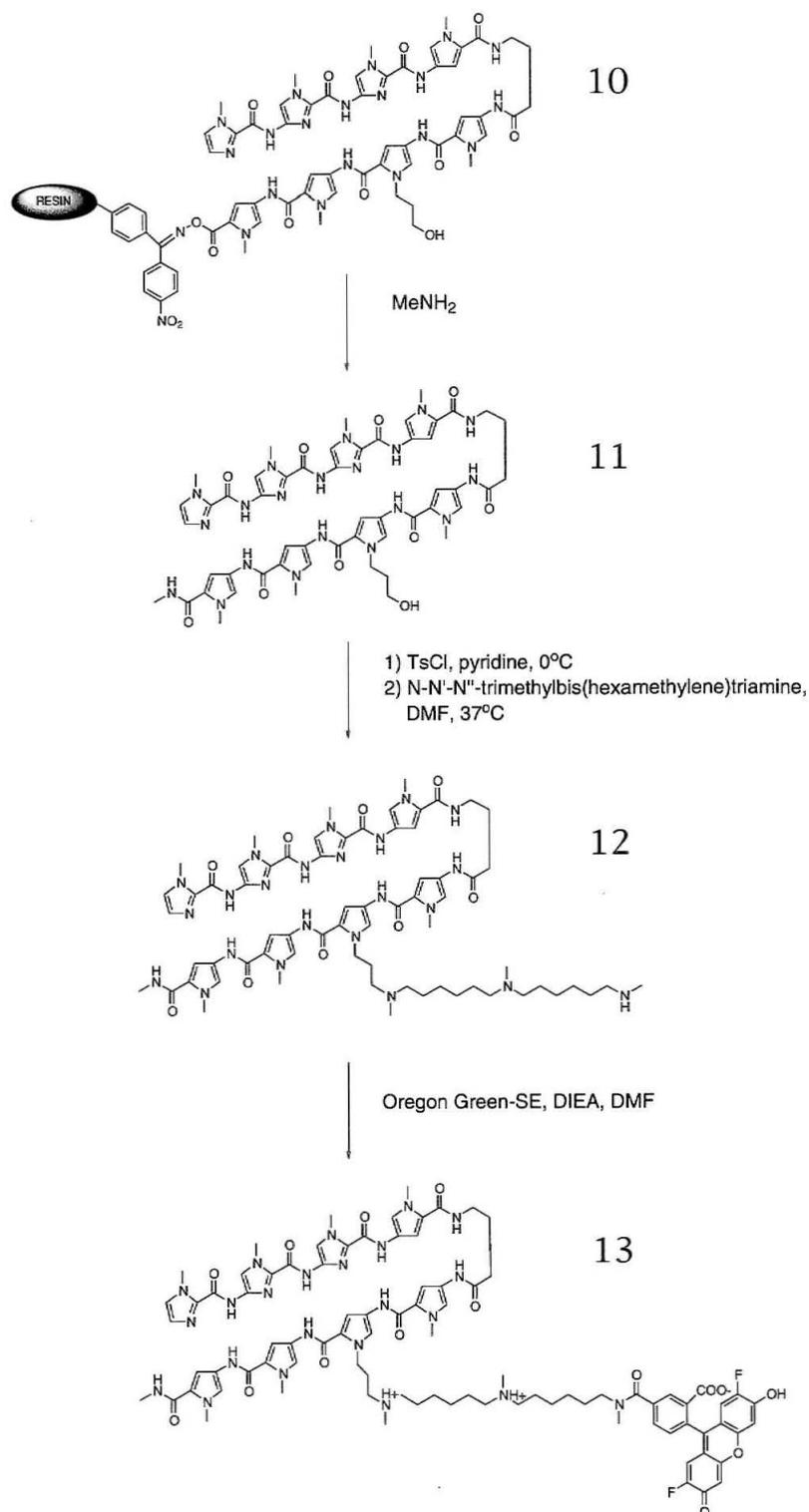


Figure 16: Synthesis of truncated tail conjugate 13

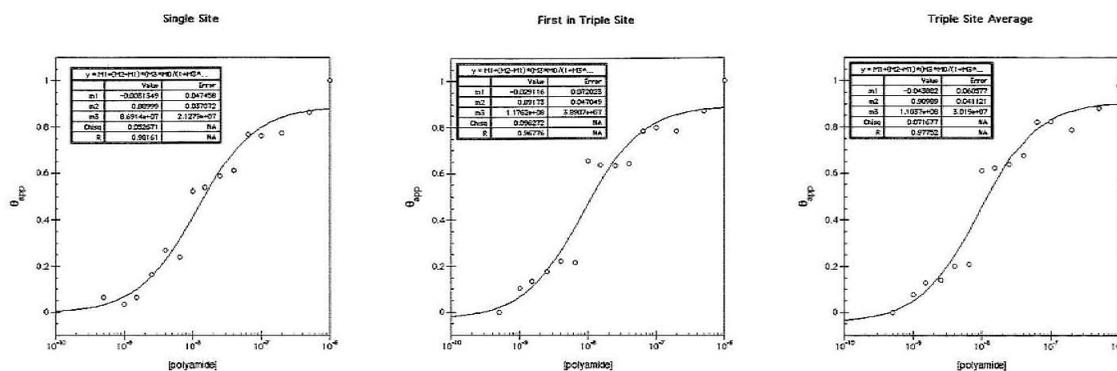


Figure 17: DNA binding isotherms for polyamide conjugate 13 on pJC01

amine, the C-terminal methyl-amide molecule **11** was obtained. Conjugation was then performed as before, yielding **13**.

Quantitative DNase I Footprinting Titrations of Fluorescent Conjugate 13

DNase I footprinting was performed with **13** as before (Fig. 18). The conjugate bound the match sequence with $K_a = 8.7 \times 10^7 M^{-1}$, the first site in the triple match with $K_a = 1.2 \times 10^8 M^{-1}$ and averaged over the triple match site with $K_a = 1.1 \times 10^8 M^{-1}$.

MPE Footprinting of Fluorescent Conjugate 13

To obtain a highly resolved footprint of both conjugates, MPE studies were performed with 4nM-4 μ M conjugate equilibrium concentrations (Fig. 19). The short-tail conjugate showed a footprint over the entire triple site region at the highest concentration performed, the same concentration at which it gave a total footprint over the single site. The β -alanine containing conjugate was unable to completely protect the triple site at 4 μ M, although it was able to footprint over the single site at lower concentrations. This

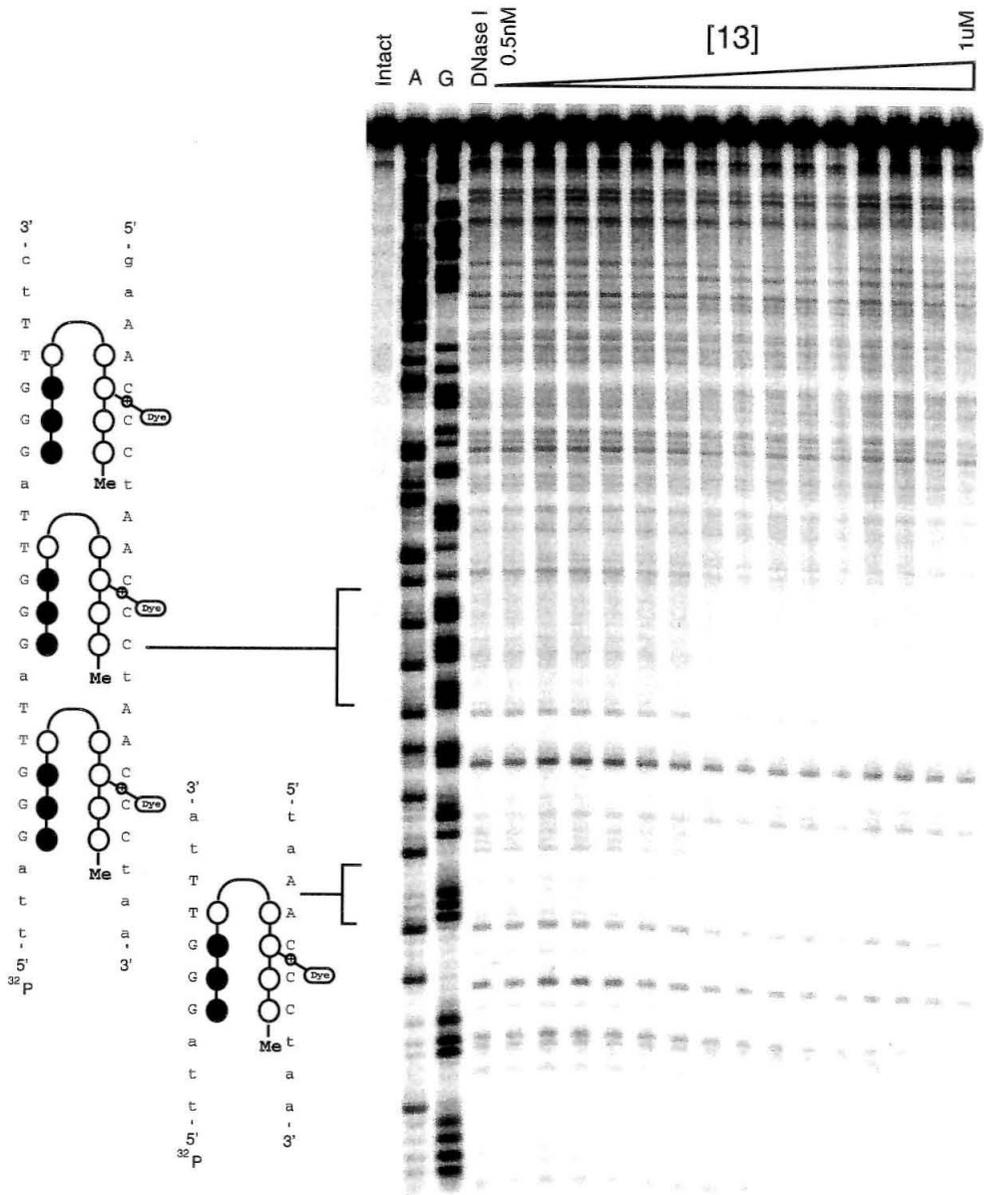


Figure 18: DNase I footprinting of 13 on 5' labeled pJC01

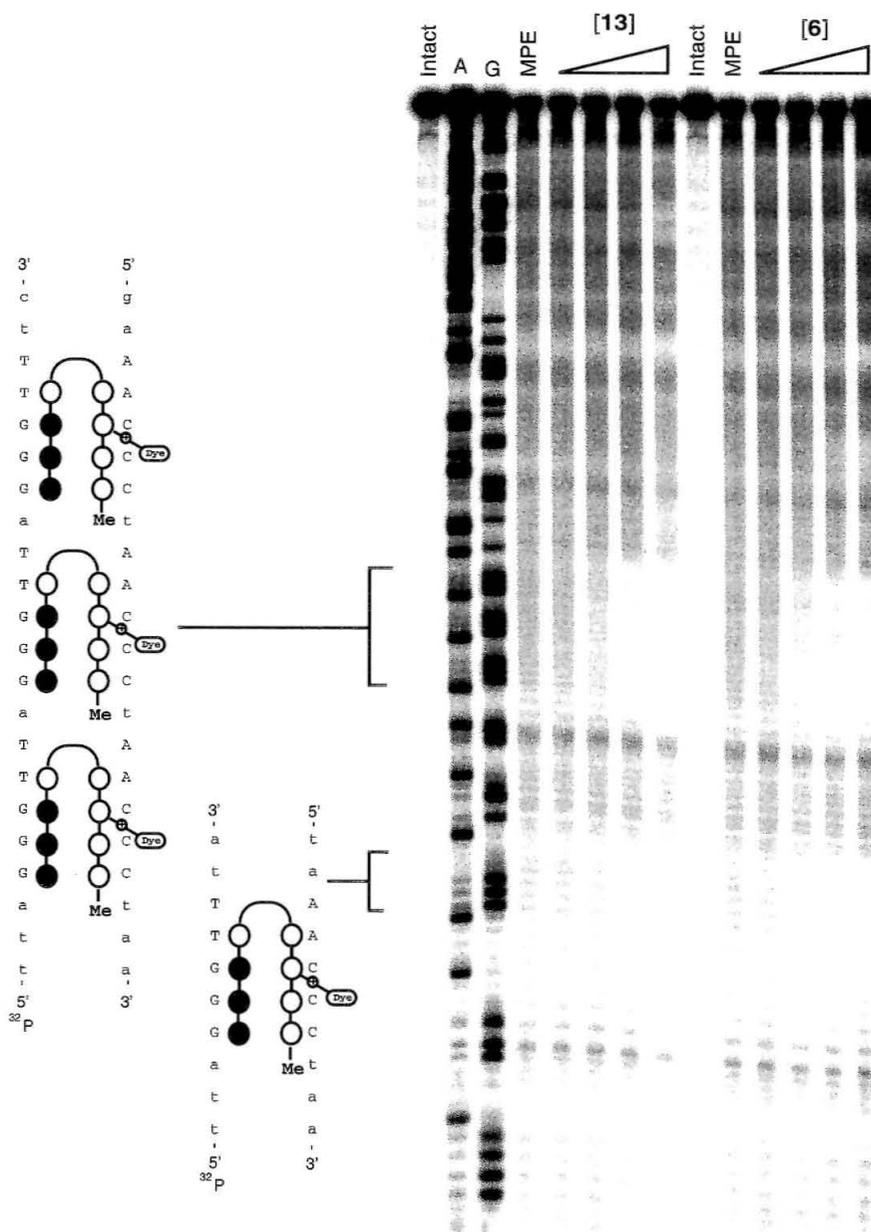


Figure 19: Autoradiogram of MPE footprinting gel for 6 and 13 on 5' labeled pJC01

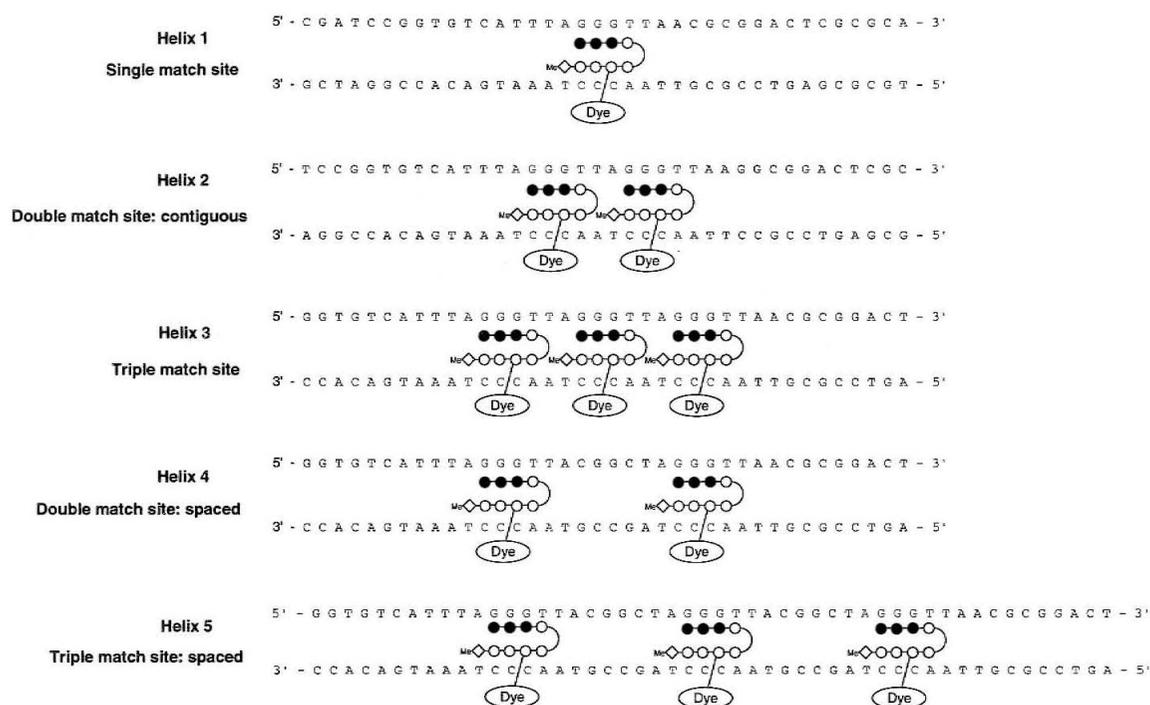


Figure 20: Helices for new series of titrations including spaced triple match H5.

evidence supports the saturation mode of binding for the short-tail compound (**13**) and contradicts it for the long tail compound (**6**).

Fluorescence Titration Experiment with Both Conjugates

The fluorescent titration of both compounds was performed as before with the addition of a new helix (Fig. 20). **H5** contained three binding sites separated by six base-pairs each. This design should completely remove steric effects between polyamides binding 3:1 with the helix and would allow for the observation of the fluorescence increase of this event. The results of these titrations are shown in Figs. 21-23.

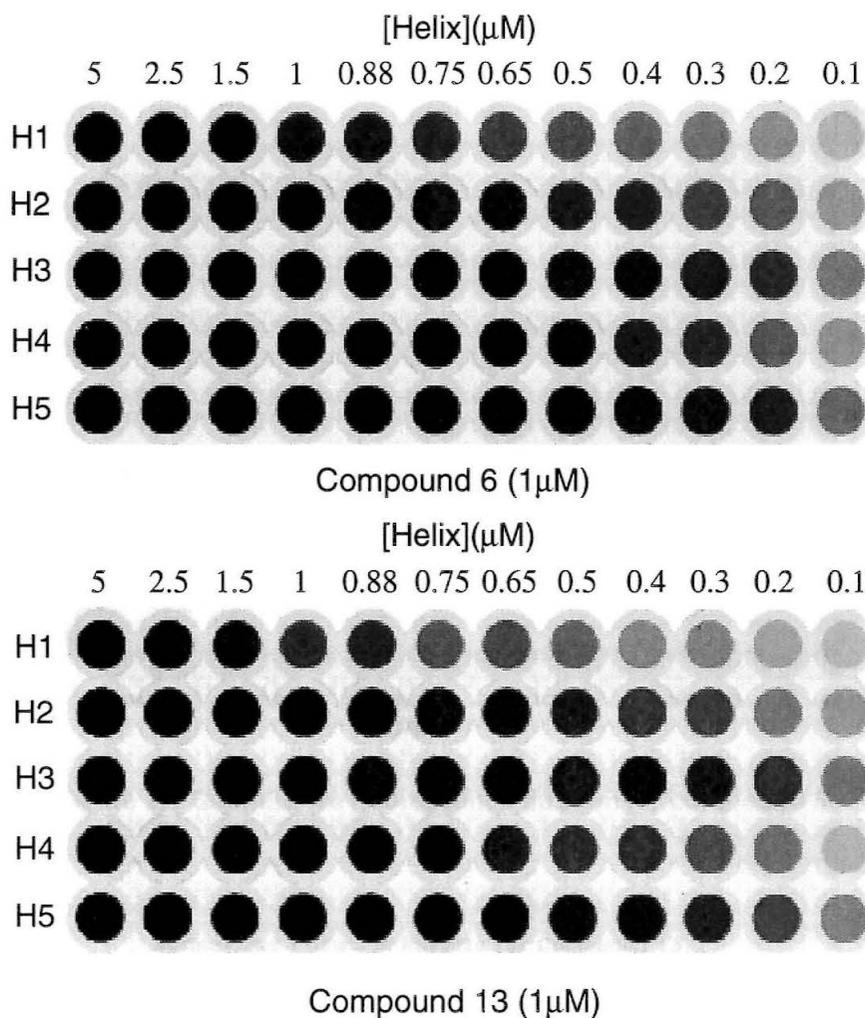


Figure 21: Fluorescence intensity image of 96-well plate titrations

Conjugate **6** yielded results similar to the first series of titrations for **H1-4**. A least squares regression analysis of the linear portion of the normalized data resulted in the values for polyamide per helix in Table 2. **H5** showed a value of 2.74 polyamides per helix, approximating the theoretical value of 3.00.

The titration data for **13** did not maintain linearity as far as the data for **6**. This may be explained by the lower K_a for **13** and may account for the lower than anticipated

Normalized Titration Results for Compound 6

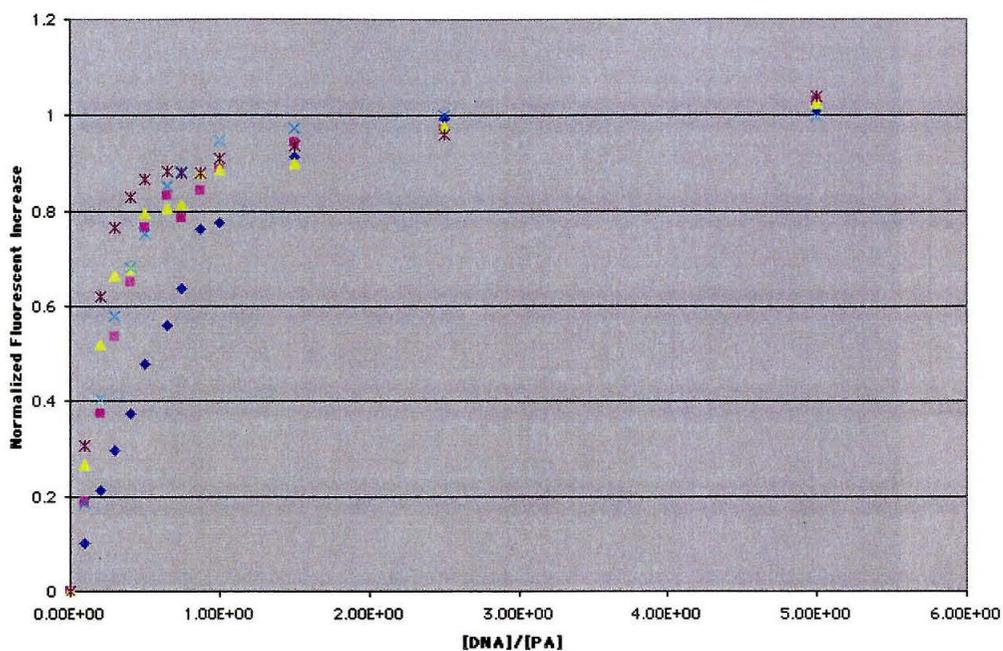


Figure 22: Quantitative fluorescence titration for compound 6

Normalized Titration Results for Compound 13

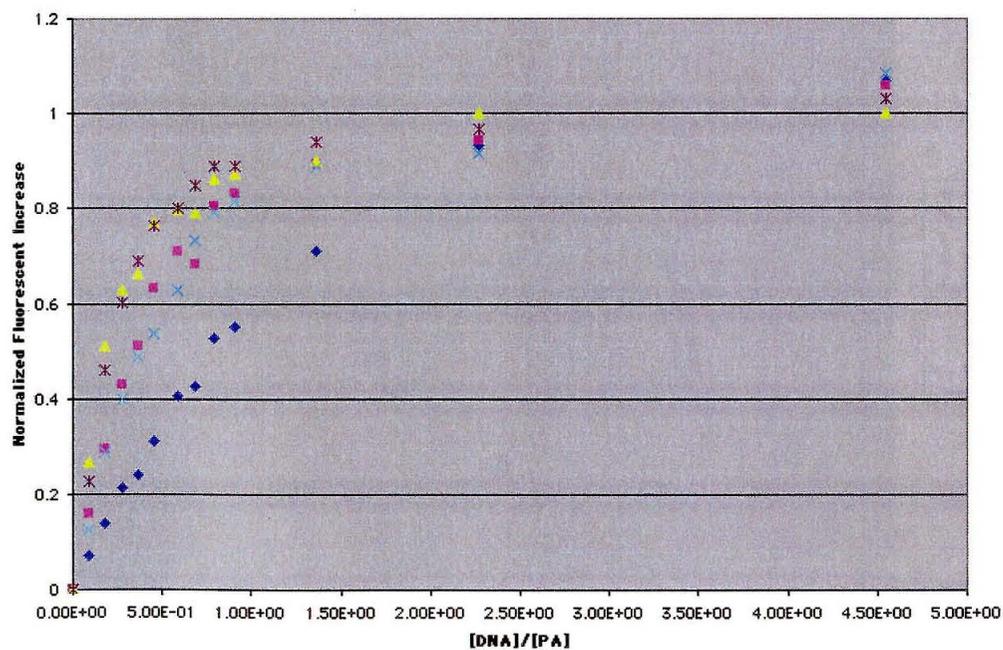


Figure 23: Quantitative fluorescence titration for compound 13

Linear Region of Compound 6 Titration

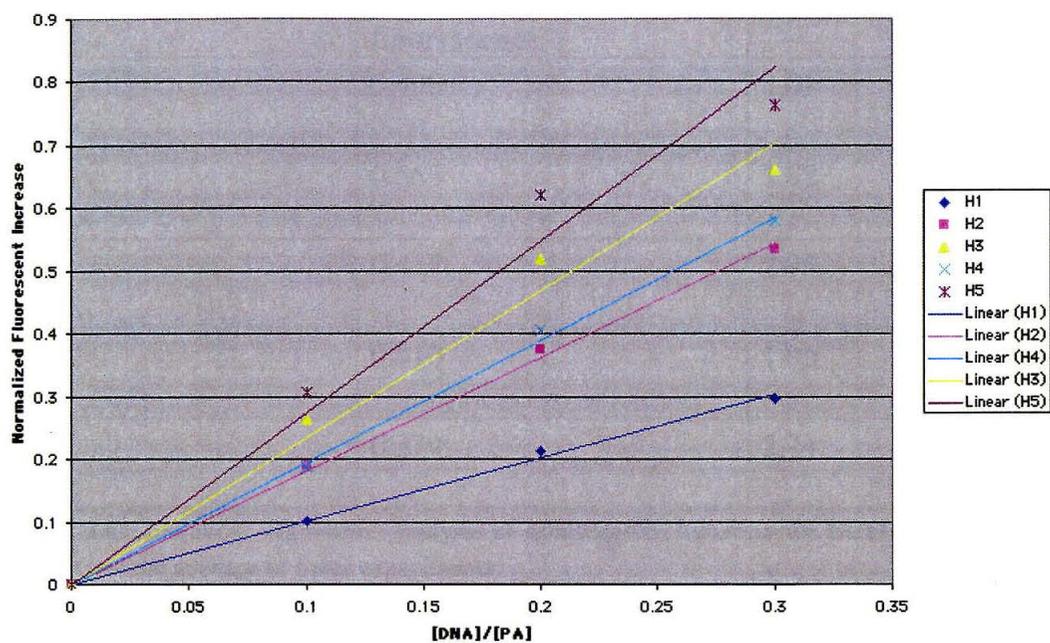


Figure 24: Linear region of quantitative fluorescence titration for compound 6

Linear Region of Compound 13 Titration

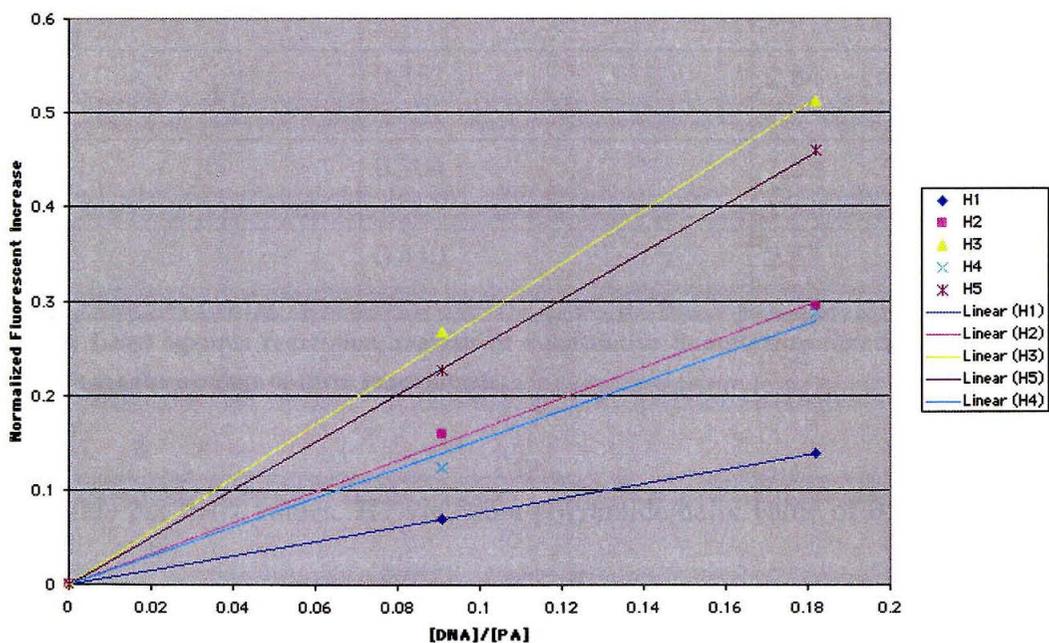


Figure 25: Linear region of quantitative fluorescence titration for compound 13

| Helix | [Helix]/[PA] at maximum fluorescence | polyamides per helix |
|-------|--------------------------------------|----------------------|
| 1 | 0.990 | 1.01 |
| 2 | 0.565 | 1.82 |
| 3 | 0.435 | 2.35 |
| 4 | 0.510 | 1.96 |
| 5 | 0.379 | 2.74 |

Table 2: Least squares regression analysis of quantitative fluorescence titration on compound 6. Numbers are the average of three experiments.

| Helix | [Helix]/[PA] at maximum fluorescence | polyamides per helix |
|-------|--------------------------------------|----------------------|
| 1 | 1.39 | 0.76 |
| 2 | 0.641 | 1.65 |
| 3 | 0.352 | 2.84 |
| 4 | 0.704 | 1.53 |
| 5 | 0.410 | 2.52 |

Table 3: Least squares regression analysis of quantitative fluorescence titration on compound 13. Numbers are the average of three experiments.

polyamide per helix values. **H1** yielded a polyamide/helix value of 0.76, approximating the theoretical value of 1.00. **H2** and **H4** yielded comparable results to each other, 1.65 and 1.53 respectively, approximating the theoretical value of 2.00. **H3** and **H5** showed

values of 2.84 and 2.52 respectively, approximating the theoretical value of 3.00. The observation that **H2** and **H4** follow similar trends and that **H3** and **H5** follow similar trends provides evidence that **13** can bind the repeat sequence contiguously.

Conclusions

The β -methyl tail compound (**6**) binds the match sequence 5'-AGGGTT-3' with $K_a = 1.8 \times 10^8 \text{M}^{-1}$. The truncated methyl tail compound (**13**) binds the match site 5'-GGGTT-3' with $K_a = 8.7 \times 10^7 \text{M}^{-1}$. MPE footprinting illustrates that **6** is unable to bind contiguously three adjacent match sites at the highest concentration tested, but the short-tail conjugate **13** protects all three sites from cleavage. Fluorescence titration data compliment this result and evidence that **13** can bind three 5'-TTAGGG-3' sites contiguously.

Discussion

The data suggest that in the model case of 5'-(TTAGGG)₃-3', a steric clash or other localized polyamide-polyamide interaction occurs when the sequence is targeted using a β -methyl tail eight ring hairpin polyamide. This prevents the molecule from consistently binding the three contiguous sites, as evidenced by MPE footprinting and fluorescence titrations. Omission of the β -alanine residue to form **13** appears to relieve this steric clash, allowing for contiguous binding.

Experimental

Materials

Boc-protected monomers were prepared as described previously.³³ Dicyclohexylcarbodiimide (DCC), Hydroxybenzo-triazole (HOBT) and were purchased from Peptides International. N,N-diisopropylethylamine (DIEA), N,N-dimethylformamide (DMF), dimethylaminopropylamine, N,N',N''-trimethylbis(hexamethylene)triamine, pyridine, dimethylsulfoxide (DMSO), triethylsilane, and toluenesulfonyl chloride were from Aldrich. Trifluoroacetic acid (TFA) was from Halocarbon. UV-spectra were measured on a Beckman-Coulter DU-7400. MALDI-TOF was performed at the Protein and Peptide Microanalytical Facility at Caltech. Analytical HPLC was performed on a Beckman Gold system using a Rainin C18, Microsorb MV, 5 μ m, 300x4.6mm reversed phase column with a 0.1% w/v TFA/acetonitrile gradient. Preparatory HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25x100mm, 100 μ m C18 column with a 0.1% w/v TFA/acetonitrile gradient. Water was obtained from a Millipore MilliQ system, 0.2 μ m filtered.

Synthesis of Polyamides and Conjugates

ImImImPy- β -PyImImIm- β -Dp (1) ImImImPy- β -PyImImIm- β -Pam resin was synthesized in a stepwise fashion by Boc-protected manual solid phase protocols.³³ A

sample of the resin was treated with neat (dimethylamino)-propylamine (1.5mL) and heated (37°C, 18 hours). The reaction mixture was then filtered diluted to 8mL with 0.1% w/v trifluoroacetic acid, and purified by reversed phase HPLC. ImImImPy-β-PyImImIm-β-Dp was recovered as a white powder upon lyophilization of the appropriate fractions (4.0mg, 5.0% recovery). ESI-MS[M+H]⁺ calc. 1212.5 for C₅₃H₆₆N₂₅O₁₀⁺ found 1212.6.

PyPyPyPy-β-PyPyPyPy-β-Dp (2) PyPyPyPy-β-PyPyPyPy-β-Pam resin was synthesized in a stepwise fashion by Boc-protected manual solid phase protocols.³³ A sample of the resin was treated with neat dimethylaminopropylamine (2mL) and heated (37°C, 18 hours). The reaction mixture was then filtered diluted to 8mL with 0.1% w/v trifluoroacetic acid, and purified by reversed phase HPLC. PyPyPyPy-β-PyPyPyPy-β-Dp was recovered as a white powder upon lyophilization of the appropriate fractions (14.5mg, 6.3% recovery). ESI-MS[M+H]⁺ calc. for C₅₉H₇₂N₁₉O₁₀⁺ 1206.6 found 1206.7.

ImImImPy-γ-PyPy((CH₂)₃OH)PyPy-β-Me (4) ImImImPy-γ-PyPy((CH₂)₃OH)Py-β-Pam resin was synthesized in a stepwise fashion by Boc-protected manual solid phase protocols.³³ A sample of resin was placed in a Parr-bomb apparatus cooled in dry ice/isopropanol. Approximately 60mL of condensed methylamine was added, the bomb sealed and allowed to stand at room-temperature overnight with occasional swirling. The bomb was then opened, methylamine allowed to evaporate. Residue dissolved in 1:1

CH₃CN:1M NH₄OH, filtered to remove resin, flash-frozen and lyophilized. Crude material was used without further purification. (42.6mg, 47.9% crude yield). ESI-MS[M+H]⁺ calc. for C₅₉H₇₅N₂₂O₁₁⁺ 1267.6 found 1267.5.

ImImImPy-γ-PyPy((CH₂)₃N,N',N''-trimethylbis(hexamethylene)triamine)PyPy-β-

Me (5) ImImImPy-γ-PyPy(CH₂)₃OH)PyPy-β-Me (33.6μmol crude from above) was dissolved in 1mL anhydrous pyridine and cooled to 0°C. Toluenesulfonyl chloride (144mg) was dissolved in 600μL anhydrous pyridine and cooled to 0°C. The tosyl chloride solution was then added to polyamide solution with vigorous mixing. Reaction allowed to proceed for 1 hour at 0°C, completion indicated by analytical HPLC. Product precipitated by addition of ether, solution decanted, washed in triplicate, and solids dried *in vacuo*. Product was divided into four aliquots, each dissolved in 300μL pyridine. To each was added 800μL N,N',N''-trimethylbis(hexamethylene)triamine and 100μL anhydrous dimethylformamide. Reactions were shaken at 37°C for 2 hours, monitored by analytical HPLC. Upon completion, product was precipitated as before, dried *in vacuo*, dissolved in 2% w/v trifluoroacetic acid and purified by HPLC. (2.5mg, 4.8% recovery). ESI-MS[M+H]⁺ calc. for C₇₀H₉₉N₂₄O₁₀⁺ 1435.8 found 1436.0.

ImImImPy-γ-PyPy((CH₂)₃N,N',N''-trimethylbis(hexamethylene)triamine-

OregonGreen488) PyPy-β -Me (6) ImImImPy-γ-PyPy((CH₂)₃N,N',N''-trimethylbis(hexamethylene)triamine)PyPy-β-Me (1.97μmol) was dissolved in 15μL

anhydrous dimethylformamide. Oregon-Green 488-succinimidyl ester (2.9 μ mol) was dissolved in 170 μ L dimethylformamide and added to the polyamide solution with 200 μ L diisopropylethylamine. Mixture was allowed to react overnight at room temperature. Product precipitated by ether addition, washed, and dried *in vacuo*. Product dissolved in a minimum amount of acetonitrile, diluted to 4mL with 0.1% w/v trifluoroacetic acid and purified by HPLC. (0.54mg, 14.1% recovery). MALDI-TOF-MS[M+H]⁺ calc. for C₉₁H₁₀₇F₂N₂₄O₁₆⁺ 1829.8 found 1829.9.

ImImImPy- γ -PyPy((CH₂)₃NH₂) Py Py - β - Me (7) ImImImPy- γ -

PyPy((CH₂)₃NPhthalimide)PyPy- β -Pam resin was synthesized in a stepwise fashion by Boc-protected manual solid phase protocols.³³ A sample of resin was placed in a Parr-bomb apparatus cooled in dry ice/isopropanol. Approximately 60mL of condensed methylamine was added, the bomb sealed and allowed to stand at room-temperature overnight with occasional swirling. The bomb was then opened, methylamine allowed to evaporate. Residue dissolved in 1:1 CH₃CN:1M NH₄OH, filtered to remove resin, flash-frozen and lyophilized. Crude material was used without further purification. (117.2mg, 66.3% crude yield). MALDI-TOF-MS[M+H]⁺ calc. for C₅₅H₆₆N₂₂O₁₀⁺ 1195.5 found 1195.6.

ImImImPy- γ -PyPy((CH₂)₃NH-Cys)PyPy- β - Me (8) ImImImPy- γ -

PyPy((CH₂)₃NH₂)PyPy- β -Me (32.7 μ mol) was dissolved in 1mL DMF, 100 μ L DIEA.

BocCysTrt (0.142mg) was activated with DCC (60mg) and HOBt (42mg) for 25 minutes, then added to the polyamide solution through a disposable polypropylene filter. Mixture was shaken at room temperature for four hours. DMF was removed *in vacuo*. Compound was deprotected with 1mL trifluoroacetic acid and 200mL triethylsilane for 10 minutes. Mixture was filtered, diluted to 8mL with 0.1% w/v trifluoroacetic acid, and purified by HPLC. (2.8mg, 6.1% recovery). MALDI-TOF-MS[M+H]⁺ calc. for C₅₈H₇₂N₂₃O₁₁S⁺ 1298.6 found 1298.7.

ImImImPy- γ -PyPy((CH₂)₃NH-Cys-Fluorescein-5')PyPy- β -Me (9) ImImImPy- γ -PyPy((CH₂)₃NH-Cys)PyPy- β -Me (1.14 μ mol) was dissolved in 127 μ L of dimethylsulfoxide. 100mM sodium bicarbonate (475 μ L) was added. Tris(carboxyethyl)phosphine (TCEP, 12.7mg) was dissolved in 100mM sodium bicarbonate (180 μ L), and 12.7 μ L of the resulting solution added to the polyamide solution. Fluorescein-5'-maleimide (7.9mg) was dissolved in 100 μ L dimethylsulfoxide and 9 μ L of the resulting solution added to the reaction mixture. Mixture shaken overnight. Diluted to 5mL with 0.1% w/v trifluoroacetic acid, purified by HPLC. (0.35mg, 16.8% recovery). MALDI-TOF-MS[M+H]⁺ calc. for C₈₂H₈₅N₂₄O₁₈S⁺ 1725.6 found 1725.8.

ImImImPy- γ -PyPy((CH₂)₃OH)PyPy-Me (11) ImImImPy- γ -PyPy((CH₂)₃OH)Py-Oxime resin (**10**) was synthesized in a stepwise fashion by Boc-protected manual solid phase

protocols.³⁷ A sample of resin was placed in a Parr-bomb apparatus cooled in dry ice/isopropanol. Approximately 60mL of condensed methylamine was added, the bomb sealed and allowed to stand at room-temperature overnight with occasional swirling. The bomb was then opened, methylamine allowed to evaporate. Residue dissolved in 1:1 CH₃CN:1M NH₄OH, filtered to remove resin, flash-frozen and lyophilized. Crude material was used without further purification. (33.4mg, 28% crude yield). ESI-MS[M+H]⁺ calc. for C₅₂H₆₁N₂₀O₁₀⁺ 1125.5 found 1125.7.

ImImImPy- γ -PyPy((CH₂)₃N,N',N''-trimethylbis(hexamethylene)triamine)PyPy-Me

(12) ImImImPy- γ -PyPy(CH₂)₃OH)PyPy-Me was treated as 4. (1.3mg, 2.3% recovery).

MALDI-MS[M+H]⁺ calc. for C₆₇H₉₄N₂₃O₉⁺ 1364.8 found 1365.0.

ImImImPy- γ -PyPy((CH₂)₃N,N',N''-trimethylbis(hexamethylene)triamine-

OregonGreen488)PyPy-Me (13) ImImImPy- γ -PyPy((CH₂)₃N,N',N''-trimethyl

bis(hexamethylene)triamine)PyPy-Me (0.97 μ mol) was treated as 5. (0.6mg, 36.2% recovery). MALDI-MS[M]⁺ calc. for C₈₉H₁₀₂N₂₃O₁₄⁻ 1754.8 found 1754.9.

DNA Reagents and Materials Enzymes were purchased from Boehringer-Mannheim and used with their supplied buffers. ³²P α - and γ - labeled nucleotides were purchased from Amersham. Calf thymus DNA (sonicated, deproteinized) and DNaseI (7500U/mL, FPLC pure) were obtained from Pharmacia. Tris-HCl, dithiothreitol (DTT), RNase-free

water, and 0.5M EDTA were purchased from US Biochemicals. XGal and IPTG were from ICN Biomedicals. Ampicillin trihydrate was acquired from Sigma. Ethanol (abs.) was purchased from Equistar. Calcium chloride, potassium chloride, and magnesium chloride were from Fluka. Formamide and pre-mixed tris-borate-EDTA (Gel Mate) were from Gibco. Bromophenol blue was from Acros. All reagents were used without further purification.

Construction of Plasmid DNA The plasmid JC01 was constructed by hybridization of the inserts 5'-GATCCGGTGTCATTTAGGGTTTACGCGGACTCGCGGATTAGGGTTAGGGTTAGGGTTCAGCTA-3' and 5'-AGCTTAGCTGAACCCTAACCTAACCTAATCCGCGAGTCCGCGTAAACCCTAAATGACACCG-3'. The hybridized insert was ligated into linear pUC19 *BamHI/HindIII* plasmid using T4 DNA ligase. The resulting construct was transfected into JM109 competent cells from Promega. Ampicillin-resistant white colonies were selected from 25mL Luria-Bertani (LB) medium agar plates (containing 50µg/mL ampicillin and treated with XGAL and IPTG). Plasmid purification was performed with Qiagen Midi-Prep purification kits. Dideoxy sequencing was used to verify the presence of the insert. The concentration of the resulting plasmid was determined at 260nm using the relationship 1OD unit = 50µg/mL.

Preparation of 3' and 5' End Labeled Fragments

End labelled fragments used in DNase I quantitative footprinting titrations were prepared according to previous methods.³⁴

Quantitative DNase I Footprinting Titrations³⁴

All reactions were carried out in a volume of 400 μ L. No carrier DNA was used in these reactions until after DNase I cleavage. A polyamide stock solution (or water for reference and intact lanes) was added to an assay buffer where the final concentrations were 10mM Tris•HCl buffer (pH 7.0), 10mM KCl, 10mM MgCl₂, 5mM CaCl₂, and 25kcpm of ³²P radiolabeled DNA. The solutions were allowed to equilibrate for 12-18 hours at 22°C. Cleavage was initiated by addition of 10 μ L of DNase I stock solution (diluted with 1mM DTT to give a stock concentration of 1.13U/mL) and allowed to proceed for 7 minutes at 22°C. The reactions were stopped by adding 50 μ L of a solution containing 2.25M NaCl, 150mM EDTA, 0.6 mg/mL glycogen, and 30 μ M base pair calf thymus DNA, and then ethanol precipitated (2.1 volumes). The cleavage products were washed with 75% ethanol, resuspended in 16 μ L RNase free water, lyophilized to dryness, and then resuspended in 100mM tris-borate-EDTA/80% formamide loading buffer (with bromophenol blue as a dye), denatured at 90°C for 10 minutes and loaded onto a pre-run 8% denaturing polyacrylamide gel (5% cross-link, 7M urea) at 2000V for 1 hour. The

gels were then dried *in vacuo* at 80°C and exposed to a storage phosphor screen (Molecular Dynamics).

Quantitation and Data Analysis of Footprinting Titrations³⁴

Data from the footprint titration gels were obtained using a Molecular Dynamics Typhoon imaging system followed by quantitation using Image Quant software (Molecular Dynamics). Background-corrected volume integration of rectangles encompassing the footprint sites and a reference site at which DNase I reactivity was invariant across the titration generated values for the site intensities (I_{site}) and the reference intensity (I_{ref}). The apparent fractional occupancy (θ_{app}) of the sites were calculated using the equation:

$$\theta_{\text{app}} = 1 - \frac{I_{\text{site}}/I_{\text{ref}}}{I_{\text{site}}^0/I_{\text{ref}}^0} \quad (1)$$

where I_{site}^0 and I_{ref}^0 are at the site and reference intensities, respectively, from a control lane to which no polyamide was added. The ($[L]_{\text{tot}}$, θ_{app}) data points were fit to a general Hill equation (eq. 2) by minimizing the difference between θ_{app} and θ_{fit} :

$$\theta_{\text{fit}} = \theta_{\text{min}} + (\theta_{\text{max}} - \theta_{\text{min}}) \frac{K_a^n [L]_{\text{tot}}^n}{1 + K_a^n [L]_{\text{tot}}^n} \quad (2)$$

where $[L]_{\text{tot}}$ is the total polyamide concentration, K_a is the equilibrium association constant, and θ_{min} and θ_{max} are the experimentally determined site saturation values when the site is unoccupied or saturated, respectively. The data were fit using a nonlinear

least-squares fitting procedure (using Kaleidagraph software) with K_a , θ_{\min} , and θ_{\max} as the adjustable parameters and a fixed value for n . The binding isotherms were normalized using the following equation:

$$\theta_{\text{norm}} = \frac{\theta_{\text{app}} - \theta_{\text{min}}}{\theta_{\text{max}} - \theta_{\text{min}}} \quad (3)$$

Reported association constants are the average value obtained from at least three independent footprinting experiments.

96-Well Plate Titration Assays

Oligonucleotides were obtained from the Caltech Oligonucleotide Synthesis Facility and used without further purification. Complementary strands were annealed in 1xTE (Tris-EDTA, Aldrich) buffer. All equilibrations were carried out in a volume of 140 μ L with buffer conditions and equilibration times identical to those used in footprinting experiments. Plates were imaged on a Molecular Dynamics Typhoon imaging system and images quantified using Image Quant software. Reported data are the average of two experiments in the case of the first series on compound **6**, three experiments in the case of the second series on compounds **6** and **13**.

References

1. Watson, J. D.; Crick, F. H. C. *Nature* **1953**, 171, 737-738.
2. Branden, C.; Tooze, J. *Introduction to Protein Structure* (Garland, New York), **1991**, 83.
3. Kielkopf, C. L.; White, S.; Szewczyk, J.W.; Turner, J. M.; Baird, E.E.; Dervan, P.B.; Rees, D. C.; *Science* **1998**, 282, 111-115.
4. Trauger, J.W.; Baird, E.E.; Dervan, P.B. *Nature* **1996**, 382, 559-561.
5. Swalley, S.E.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc.* **1997**, 119, 6953-6961.
6. Turner, J.M.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc.* **1997**, 119, 7636-7644.
7. Trauger, J.W.; Baird, E.E.; Dervan, P.B. *Angew. Chem. Int., Ed. Engl.* **1998**, 37, 1421-1423.
8. Turner, J.M.; Swalley, S.E.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc.* **1998**, 120, 6219-6226.
9. Urbach, A.R.; Szewczyk, J.W.; White, S.; Turner, J. M.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc.* **1999**, 121, 11621-11629.
10. Kelly, J.J.; Baird, E.E.; Dervan, P.B. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, 93, 6981-6985
11. Trauger, J.W.; Baird, E.E.; Mrksich, M.; Dervan, P.B. *J. Am. Chem. Soc.* **1996**, 118, 6160.
12. Swalley, S.E.; Baird, E.E.; Dervan, P.B. *Chem, Eur. J.* **1997**, 3, 1600.

13. Trauger, J.W.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc.* **1998**, 120, 3534.
14. Dickinson, L. A.; Gulizia, R. J; Trauger, J.W.; Baird, E.E.; Mosier, D.E.; Gottesfeld, J.M.; Dervan, P.B. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, 95, 12890-12895.
15. Mapp, A.K.; Ansari, A.Z; Ptashne, M.; Dervan, P.B. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97, 3930-3935.
16. Blackburn, E.H. *Cell* **1994**, 77, 621-623.
17. Moyzis, R. K.; Buckinham, J.M.; Cram, L.S.; Dani, M.; Wu, J.R. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, 85, 6622-6626.
18. Muniyappa, K.; Kironmai, K.M *Crit. Rev. Biochem. Mol. Biol.* **1998**, 33, 297-336.
19. Zakian, V.A. *Annu. Rev. Genet.* **1996**, 30, 141-172.
20. Allsopp, R.C.; Vaziri, H.; Harley, C.B. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, 89, 10114-10118.
21. Vaziri, H.; Schachter, F.; Uchida, I.; Wei, L.; Harley, C.B. *Am. J. Hum. Genet.* **1993**, 52, 661-667.
22. Vaziri, H.; Dragowska, W.; Allsopp, R.C.; Landsorp, P.M. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, 91, 9857-9860.
23. Harley, C.B.; Vaziri, H.; Counter, C.M.; Allsopp, R.C. *Exp. Gerontol.* **1992**, 27, 375-382.

24. Shay, J.W.; Wright, W.E. *Science* **2001**, 291, 839-840.
25. Shay, J.W. *Exp. Cell Res.* **1991**, 196, 33-39.
26. Bodnar, A.G.; Ouellette, M.; Frolkis, M.; Holt, S.E., Shay, J.W., Wright, W.E. *Science* **1998**, 279, 349-352.
27. Hayflick, L.; Moorhead, P.S. *Exp. Cell Res.* **1961**, 25, 585.
28. Campisi, J.; *Cell* **1996**, 84, 497.
29. Wright, W.E.; Shay, J.W. *Exp. Gerontol.* **1992**, 27, 383-389.
30. Kim, N.W. et al. *Science* **1994**, 266, 2011-2015.
31. Poon, S.S.; Martens, U.M.; Ward, R.K.; Landsorp, P.M. *Cytometry* **1999**, 36, 267-278.
32. Swalley, S.E., Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc.* **1996**, 118, 8198-8206.
33. Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc.* **1996**, 118, 6141.
34. Trauger, J.W.; Dervan, P.B.; *in press*
35. Bremmer, R.B., Wurtz, N.R.; Dervan, P.B. *in press*
36. Rucker, V.C., Melander, C.; Foister, S.; Dervan, P.B. *manuscript in preparation*
37. Belitsky, J.M.; Nguyen, D.H.; Wurtz, N.R.; Dervan, P.B. *manuscript in preparation*
38. Herman, D.M.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc.* **1998**, 120, 1382-1391.