The Design, Synthesis and Characterization of Bright Fluorescent Probes for Biomolecule Detection

Submitted in Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy

The Department of Chemistry

Elizabeth E. Rastede

Carnegie Mellon University

Pittsburgh, PA

May, 2015

Abstract

The ability to observe and quantify interactions within a cellular environment has allowed scientist to develop a detailed understanding of biomolecular interactions, in addition to identifying and diagnosing disease states. Advancements in fluorescent probes and imaging techniques have proven to be instrumental for peering deeper into these molecular interactions in real time. Although fluorescent small molecules are well-established and versatile tools, fine tuning the spectral properties of either the probe or the small molecule is necessary in order to obtain a clear and accurate picture of the desired molecular environment. In this thesis, we develop and explore new fluorescent probes for adaptable, efficient, and sensitive detection assays.

The first portion of this thesis looks at manipulating the scaffold of a fluorogenic dye in order to develop a versatile fluorophore to stain non-specifically DNA as well as bind to a promiscuous protein. Fluorogenic dyes are a special class of fluorophores that show distinctly different emission intensities depending on their local environment. When free in a solution, the dye can exploit non-radiative twisting pathways about the central methine bridge; but when in a viscous media or when bound by a biomolecule, this pathway is eliminated and fluorescence increases. The fluoromodule spectral range is dependent upon the properties of the chromophore that is bound. In this thesis, we present a series of thiazole orange analogs with varied substituents and electronic properties, and investigate their spectroscopic properties and their fluorogenic behavior in different solvents and biomolecular environments (e.g. DNA and scFv). It was determined that a 40nm spectral shift could be obtained allowing for a brighter image at common laser wavelengths.

The second portion of the thesis focuses on improving hybridization probes targeting nucleic acids. Traditional strategies lack the ability to distinguish bound from unbound probes. In order to minimize background fluorescence from unbound probes we developed a modified binary probe used to label DNA telomeric repeats in fixed cells and tissue. By utilizing γ-modified Peptide Nucleic Acids (γ PNAs), we were able to shorten and simplify the binary probe design with two terminally-labeled γPNA 9-mer probes designed to hybridize in tandem allowing for a Förster Resonance Energy Transfer (FRET) signal to indicate hybridization to the target. First, the γPNA probes were analyzed for affinity, structure, and fluorescence properties. Next, the γPNA probes were applied for the analysis of telomeric DNA in genomic DNA, as well as cellular imaging using fluorescence in-situ hybridization (FISH).

The final portion of the thesis focuses on enhancing the brightness of hybridization probes. Traditional PNA probes are designed to be complementary to the target and a single dye is covalently attached at the terminus. In this strategy, sensitivity of the probe is dependent on the fluorescent properties of a single dye. In order to enhance the brightness of the probe, we proposed an internal labeling strategy to improve the local concentration of dye per probe. We developed PNA probes with multiple internal fluorescent dyes using the Huisgen-Meldal Sharpless “Click” reaction between an azide-modified fluorescent dye and alkynyl-uracil-containing PNA. We have synthesized a dye-modified PNA monomer and have incorporated it into a series of oligomers. It was determined that the PNA:DNA duplex stability was not affected by the presence of the dye and only a marginal decrease in fluorescence due to quenching was observed, even when the two dyes were placed in adjacent positions. We have demonstrated the feasibility of this approach, which suggests that several fluorescent dyes can be introduced into a PNA probe at internal positions in order to enhance the brightness of DNA/RNA-targeting PNA probes.

We have improved the efficacy of fluorescent probes by altering the structure of a fluorogenic dye, and designing new strategies for nucleic acid hybridization probes. Versatile, brighter, and more accurate probes will ultimately give biologist better tools to diagnose, characterize, and ultimately understand biological interactions at a molecular level.

TABLE OF CONTENTS

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Charts

Chapter 2

Chapter 3

Chapter 4

Table of Schemes

Chapter 2

Chapter 3

Chapter 4

Acknowledgments

In my 5 years at Carnegie Mellon there have been many people that have supported me through this journey. First and foremost , I would like to express my upmost gratitude to Dr. Bruce Armitage, he gave me the freedom to explore, opened new doors to areas of chemistry I didn't even know I would enjoy and guided me when the trail was hard to follow. It has been a privilege to work in his lab and to have him as a mentor.

A special thanks to my committee members Dr. Marcel Bruchez, Dr. Danith Ly for their constructive advice and suggestions; and thank you Dr. Jim Schneider for serving as the external member on my committee.

I could not give a big enough thanks to the Amitage Lab for a fun, and supportive work environment. To past members: Dr. Connor Murphy, Dr. Kim Zanotti, Dr. Anish Gupta, Dr. Ha Pham. Dr. Nathaniel Shank and Dr. Wendy Brotherton, thank you for your continued guidance through graduate school and for putting up with my constant questions. And the current members : Karen, Stan, Taylor, Munira, Xiaohong, and April. I have enjoyed battling ideas on the white board, and laughing over dinner at U.G. I am thankful to have the experience of learning to mentor two very talented undergrads Lukas "Pronner" and Michelle Moon, who I believe will go on to do great things.

A very special thanks to Dr. Rea Freeland for her support, guidance, and words of encouragement, words cannot express my gratitude and I hope to continue our friendship in the future. I would also like to thank Dr. Massimilliano Lamberto for taking me under his wing, sparking a passion for organic chemistry and encouraging me to go to graduate school.

I would also like to thank the many talented scientists that have helped me along the way, Dr. Brigitte Schmidt for her synthetic wisdom, guidance and dye materials. Gloria Silva for welcoming me into the group. Roberto Gil and Gayathri Withers for their help with NMR.

Finally, I would like to thank my parents, Don and Meta Rastede, for everything they have done for me; and my siblings, Freddy and Leighann, for their continued support, and encouragement. My friends for their support, especially Vince, who always seems to make me laugh, even on the dreariest of Pittsburgh days.

1. Chapter 1: Fluorescence and the Fluorescent Detection of Biomolecules

Molecular fluorescence technologies have been instrumental in several scientific breakthroughs within the biological community, from decoding the human genome to drug discovery, providing scientists the ability to visualize and understand the molecular interactions within a complex cellular environment. However, despite these breakthroughs there remains a need for physically diverse, accurate, biocompatible fluorescent materials to selectively detect and visualize a dynamic system such as our cellular enviroment. Scientists have not yet reached the full potential of rationally designed biocompatible materials combined with fluorescence techniques. In this thesis we developed and explored new combinations of bio-compatible materials in order to achieve further flexible, efficient and sensitive detection assays.

1.1. Fluorescence

Fluorescence is the emission of light from a molecule in an excited singlet state typically achieved using direct photoexcitation[1,2]. Fluorescence and its techniques have been extensively used and its popularity can be attributed to the high sensitivity, specificity, and responsiveness to the micro environment, all of which can translate well in a living system. Further, fluorescence provides spatial and temporal control, as well as real time information to help visualize the dynamics of a system of interest. However, first we must understand the physical process of fluorescence at a quantum level in order to logically manipulate molecules and constructs for improved detection assays.

Fluorescence is a multi-step process that can be conveniently visualized using a Jablonski Diagram (see Figure 1)[1,3]. The horizontal lines represent different energy levels of a molecule with S_0 representing the singlet ground state and S_1 and S_2 representing the subsequent singlet excited states as higher energy states. The process begins when the molecule in its ground state absorbs a photon, thus exciting the molecule (an electron) to a higher energy state as seen in Figure 1(A). After excitation, the electron relaxes back down to a

lower energy state (B) or the ground state. This return to the ground state can come in many forms. First, if the energy is released back as a photon, this process is called fluorescence (C). Secondly, the energy can be released through internal conversion or non-radiative decay (D). Non-radiative decay includes vibration, bond rotation, molecular collisions, or Förster resonance energy transfer to another molecule. A third path involves an intersystem crossing to the triplet state (E), from which the electron can relax by releasing energy as a photon (phosphorescence, F) or via non radiative decay (G). All of the aforementioned non-radiative processes in addition to phosphorescence compete with fluorescence and reduce the efficiency of the fluorescence process[2]. The fluorescence efficiency can affect the sensitivity of the assay.

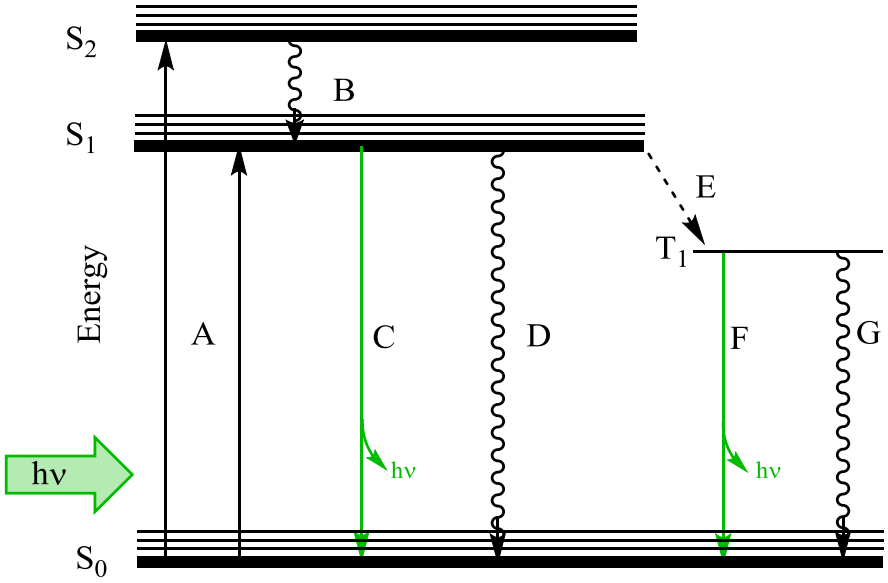


Figure 1: Jabloski diagram, an energy diagram illustrating molecular states and possible transitions between them. Solid lines represent radiative transitions, curvy lines represent non-radiative transitions and the dashed line represents the intersystem crossing to a triplet state. Figure adapted from ref 1 and 2.

1.2. Fluorescence Methods

Fluorescence is typically detected using either a spectrophotometer or a microscope depending on the scale and assay. However, continuing advancements in data analysis, instrumentation, and technology continue to push the limits of fluorescent detection and its applications. Some fluorescence methods include fluorescence anisotropy, flow cytometry, and

single molecule spectroscopy[2]. Furthermore, phenomena such as resonance energy transfer can be utilized to increase the potential of current technologies which will be discussed in further detail in Chapter 3.

1.3. Fluorophore Properties

A fluorophore is a molecule that emits fluorescence when excited. The usefulness of a fluorophore is determined by both its photophysical and chemical properties. Although chemical properties are important, they will not be addressed in detail.

Photophysical properties of a fluorophore can be clarified by observing fluorescence as a quantum based and systematic process as described in Section 1.1. The efficacy of a fluorophore depends on the efficiency of two processes: absorbance of a photon and subsequent release of a photon (fluorescence). All fluorophores have an absorbance spectrum which reflects the transitions between the ground state and the subsequent excited states. The most efficient transition from ground state to excited state is represented as the maxima of the absorbance spectrum and may fluctuate due to changes in the microenvironment of the molecule. The measured absorbance intensity is the product of three variables: the path length, the concentration of the absorbing species, and molar extinction coefficient. The molar extinction coefficient (ε; units = $M^{-1}cm^{-1}$) expresses the efficiency of a species to absorb light at a given wavelength and is an important parameter when selecting a fluorophore[2]. The other important property is the efficiency for a molecule to fluoresce which can be defined by the quantum yield (φ) which is the ratio of the number of emitted photons to the number of absorbed photons. The product of these two important properties, molar extinction coefficient and the quantum yield, is defined as the brightness. Brightness is also the measured fluorescence intensity when a fluorophore is excited at a given wavelength. Brightness is especially useful when comparing the fluorescence properties of different species in the same experimental parameters. Other photophysical properties include spectral maxima, Stokes shifts, extinction

coefficients, quantum yield and fluorescence lifetime[2]. Chapter 2 will look at synthetically modifying a fluorophore and studying the effect on the physical properties.

Chemical properties include reactivity, lipophilicity, pKa, and chemical stability. Fluorophores can be designed to fluoresce only upon a change in physical or chemical environment. This "light up" behavior is termed fluorogenic. Fluorophores have been designed to be sensitive to the pH of the cellular environment, producing a wavelength shift when the lipophilicity changes. Other fluorophores have been designed to light up upon chemical reaction, signaling the presence of the molecule of interest.

Physical and chemical properties can be manipulated by synthetically altering the chemical scaffold of some common fluorescent small molecules. Most fluorophores are not innately specific or responsive to their environment; scientists must modify or functionalize a fluorophore to get desired properties for a particular assay. One method to achieve fluorophore specificity is modular approach, where a specific targeting moiety can be covalently attached to a fluorophore, several examples of this approach will be discussed in Section 2. In order to optimize the physical property of brightness, Chapter 4 will look at attaching multiple fluorophores on a specific targeting moiety.

1.4. Common small molecule fluorophores

Fluorophores can either be endogenous or synthetic, although synthetic fluorophores are more common due to their tune-ability and optimized physical properties. Endogenous fluorescent small molecules include some amino acids (e.g. tryptophan, phenyl alanine, and tyrosine)[4], nicotinamide cofactors[5], and porphyrins. Although naturally occurring, these fluorophores typically have low quantum yields and poor spectral properties. Chemists have become extremely proficient at synthesizing fluorescent molecules that have better properties than naturally occurring small molecules, and that vary in extinction coefficient and quantum

yields all along the UV-Vis spectrum[3,6-8]. Some common examples of fluorescent small molecules can be seen in Figure 2 in addition to their physical properties in Table 1.

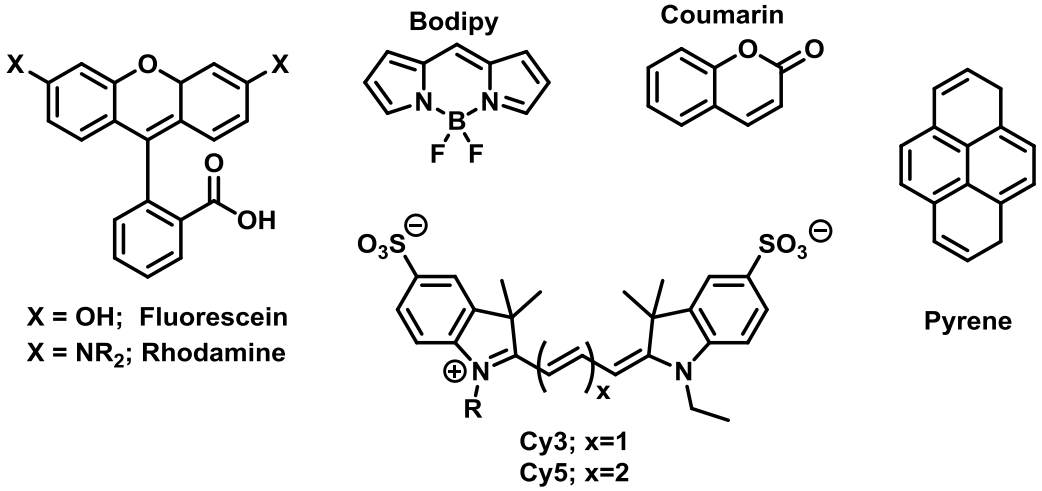


Figure 2: Common Fluorophore Structures and Properties

Table 1: Physical Properties of Common Fluorophores

Fluorophore	λmax / λem (nm)	E ($M^{-1}cm^{-1}$)	φ	ref
Fluorescein	490 / 514	9.3×10^4	0.94	9
Rhodamine	496 / 517	7.4×10^4	0.92	9,10
Pyrene	340 / 376	4.3×10^4	0.75	11,12
Coumarin	360 / 450	1.7×10^3	0.63	13
Bodipy	505 / 511	9.1×10^4	0.94	14
Cy3	560 / 575	1.5×10^5	0.09	15
Cy5	658 / 677	2.5×10^5	0.40	15

Among these fluorophores fluorescein is the most popular and widely used small molecule due to its environmental sensitivity and high quantum yield[9,12,16]. Several derivatives of fluorescein modify the spectral shifts[17,18], photo-activity[19-21] and environmental sensitivity[16,19,22]. Rhodamine has a similar scaffold to fluorescein, substituting the phenol with an N- alkyl amino group results in a red-shifted spectrum with high quantum yields [9,10,23]. Other common small molecule fluorophores include, but are not limited to, pyrene, bodipy, and coumarin. Pyrene is a polycyclic aromatic[12] and has the most useful property of a long

fluorescence lifetime, which allows for an excited state pyrene to associate with a ground state pyrene to form a red shifted "excimer"[24]. Bodipy, another small fluorescent molecule, is also hydrophobic but exhibits favorable fluorescein- like spectral properties[12,25,26]. Coumarins have large Stokes shifts and can be easily tuned for solubility and environmental sensitivity[12,27]. Cyanine dyes (e.g. Cy3, Cy5) are members of another common class of fluorescent small molecules that exhibit favorable tunability by varying the polymethine chain and/or heterocycles to cover a vast range of the electromagnetic spectrum[12,15,28,29]. Cyanine dyes will be discussed further in Chapter 2. All of these common classes of fluorescent dyes can be covalently attached to a biomolecule of interest by strategically synthesizing a chemical handle on the small molecule thus allowing for fluorescent labeling and detection.

1.5. Other fluorophores

Although small molecules have the advantage of size, in order to increase brightness, photo stability and expand biocompatibility, scientists have explored alternative fluorophores such as quantum dots[30-33], nanoparticles[34-36] and auto-fluorescent proteins[37]. Auto-fluorescent proteins such as GFP, will be discussed in further detail in section 2.3. Overall, the ideal fluorophore should have the following properties: high brightness, both chemically and photo chemically inert, and appropriate and selective excitation and emission parameters based on current technology[6,38-40].

2. Fluorescent Detection of Biomolecules

Mapping the human genome gave scientists the blueprint for a human; however, this genetic map lacks dynamic information (e.g. location, movement and function of biomolecules)[41]. In order to fill this gap, scientists have focused on identifying all functional elements within the human genome. In order develop a strong understanding of the functional elements and their interactions, it is essential to be able to track all the molecules involved in the flow of information in vivo from genetic code (DNA) to functional biomolecules (RNA and

proteins). Fluorescence imaging allows scientists the ability to observe and quantify interactions within a cellular environment in order to develop a detailed understanding of biomolecular interactions. In addition, fluorescent tags allow for quantifiable detection and visualization of cellular responses to external stimuli (e.g. drug or environmental changes). In this section we will review the common techniques used to label and detect DNA, RNA and proteins.

2.1. DNA Labeling

DNA is the genetic material that codes for RNA and protein. DNA contains two key structural elements: a negatively-charged phosphate backbone that interacts with the hydrophilic environment of a cell, and nucleobases which contain the code and are protected through hydrophobic interactions[42]. DNA is not naturally fluorescent; therefore, in order to visualize DNA, scientists have developed fluorescent labels. In the following sections we will describe three types of DNA labeling: staining, hybridization probes, and enzymatic labeling,

2.1.1. DNA Staining

Some fluorescent small molecules can be designed to interact non-covalently with DNA[43]. Typically, these small molecules have hydrophobic cores for interactions with the nucleobases and cationic substituents for electrostatic interactions with the negatively charged phosphate backbone. Due to the nature of these interactions, hydrophobic and electrostatic, small molecule and binding to nucleic acids is not typically sequence specific.

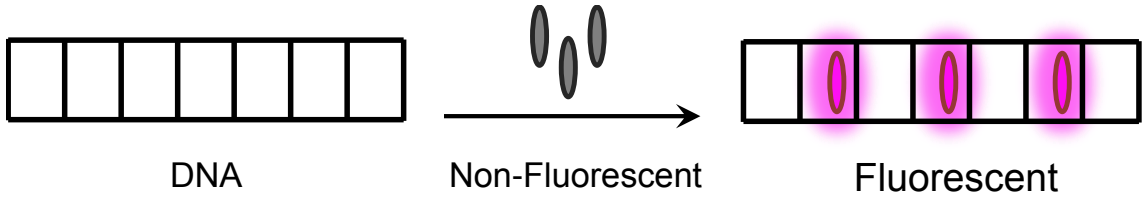


Figure 3: Non-covalent small molecule interaction with DNA using a fluorogenic dye

A common binding mode for small molecules is intercalation which requires a planar aromatic compound that can fit between adjacent base pairs[44] (Figure 3). Although these DNA binding small molecules are not specific, they are commonly used as DNA stains in cellular assays[45], agarose gels, capillary electrophoresis[46,47], among other applications. A key feature of these small molecules is their fluorogenic behavior, eliciting a light up response upon interaction with DNA[48]. When the dye is free in solution, minimal fluorescence is observed, indicating a relaxation pathway though non-radiative decay (e.g. torsional motion within the molecule). However, once this torsional motion is restricted, via intercalation with DNA, the molecule uses a fluorescence pathway to return to the ground state, emitting a 1000 fold increase in fluorescence[49,50]. This striking feature allows for these dyes to have low background fluorescence from unbound dye, which is ideal for DNA detection[51]. Extensive optimization of dye has produced a broad color assortment of commercially-available nucleic acid stains. Some common examples of DNA intercalators are ethidium bromide[52], acridine orange, and thiazole orange[45]. Derivatives of thiazole orange will be highlighted in Chapter 2.

2.1.2. Hybridization Probes

Although DNA stains have the advantage of low background fluorescence and fast kinetics for DNA detection, they lack sequence specificity. In order to distinguish a specific nucleic acid sequence, Watson-Crick base pairing can be used to design a nucleic acid analog probe to hybridize to the target. As a result a detectable marker is positioned on the probe for easy visualization. As seen in Figure 4A, the covalent attachment of a fluorescent tag to a probe permits fluorescence methods to detect and visualize nucleic acids[53-55]. However, it is difficult to distinguish the emission from unbound from bound probe, thus, requiring additional steps to remove the unbound probe. This limits the types of assays these probes can be used in.

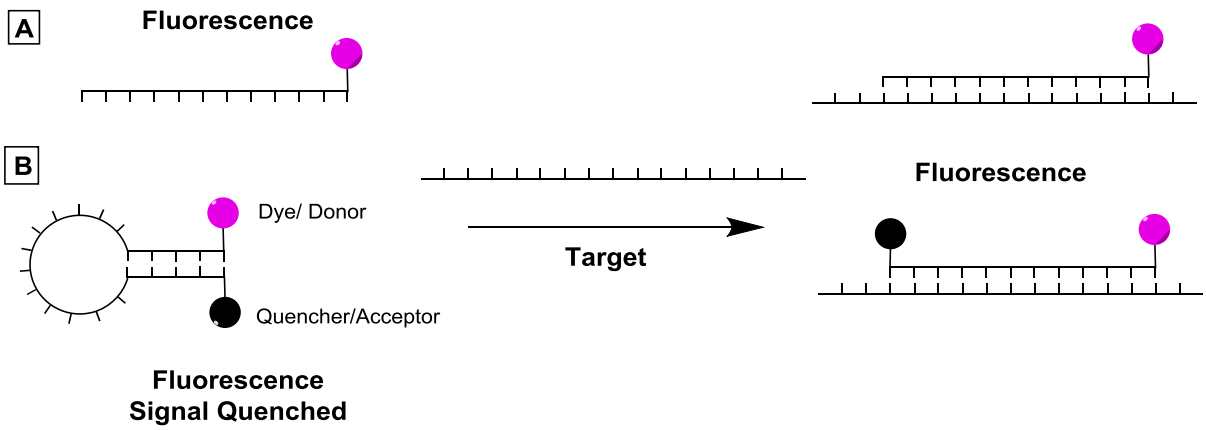


Figure 4: Hybridization probe strategies (A) fluorescently-labeled oligonucleotide with target; (B) fluorescently-labeled oligonucleotide with quencher/acceptor in secondary strand in order to reduce signal from unbound probe and limit off target effects.

In Figure 4B, two general strategies have been utilized to reduce background signal: Energy transfer via contact (Quencher); or energy transfer to another dye (FRET)[56]. In general both strategies reduce the fluorescence of the signaling dye by either collisional quenching or energy transfer to another dye (acceptor), once bound to the target the quencher/acceptor dye is no longer effective and the dye can fluoresce[57]. This strategy allows assays to be performed in excess probe and limits the background noise. Different hybridization probe strategies will be discussed in more detail in Chapter 3.

Most oligonucleotides for probes are generated using solid phase synthesis which has the advantage of incorporating site specific modifications, as well as terminally labeling sequences. Site specific modifications can include nucleotides that have a covalently attached a fluorescent dye to the nucleobase. These internal modifications allow multiple fluorescent dyes to be incorporated in a single probe resulting in a brighter probe[58,59]. Nucleobase modifications leading to internal labeling will be further discussed in Chapter 4.

2.1.3. Enzymatic labeling of Nucleic Acids

Although solid phase synthesis of nucleic acids is a robust and scalable process, the incorporation of nucleobase-modified monomers can be limiting by both oligomer length and chemistry. Solid phase requires that base-modified phophoramidite protecting groups and modifications must be stable to oxidation and have limited reactivity with nucleophiles. Another approach is to let biology do the work through enzymatic synthesis using a DNA polymerase with fluorescently modified 2'-deoxyribonucleoside triphosphates (dN^{x}TPs).

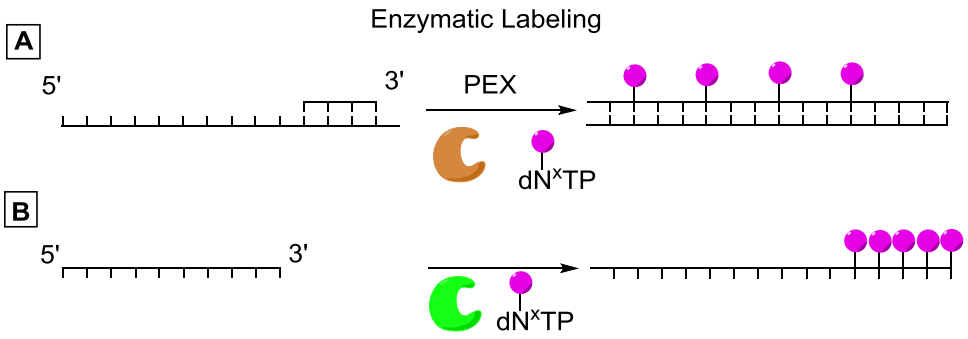


Figure 5: Enzymatic labeling of DNA using fluorescently modified dNTP (A) Primer extension (PEX) (B) terminal labeling using transferase enzyme

Several methods have been developed for synthesis of nucleobase modified DNA using polymerases. One of the simplest is primer extension (PEX) which uses a primer with a complementary template with a 5' overhang. The polymerase extends the primer and incorporates the dN^{x}TP to its complementary nucleotide, thus allowing several fluorophores to be incorporated within the sequence[60,61]. Another methodology uses a modified PCR, allowing for long sequences to be rapidly synthesized using a modified dN^{x}TP[62]. However, both of these methods lack site-specific labeling at internal positions of DNA. Achieving optimal spacing of fluorophores to prevent contact quenching can also be challenging due to the reliance of a template sequence for these methods. One solution is to dope in natural dNTPs. Still, this

method would result in a heterogeneous population of labeled DNA sequences. Any differences in population would introduce error and reduce accuracy[60]. Finally, terminal labeling can be achieved using deoxynucleotidyl transferase for non-template 3' elongation. Typically 10-20 modified nucleobases can be incorporated using this method[63]. However, optimal spacing of the dyes to prevent quenching still remains a problem. A further analysis of internally labeled nucleic acids using a peptide nucleic acid oligomer will be discussed in Chapter 4.

2.2. RNA Labeling

RNA has many roles within the cellular environment. Through the sequencing of the human genome, scientists were surprised to find relatively few coding regions, which began to raise questions about the function of the non-coding RNA[41,64]. Subsequently, the exploration of mRNA, long non-coding RNA and small interfering RNA suggests a complicated image of RNAs function within the cell[65-67]. This has motivated scientists to develop new techniques to try to visualize RNA structure and function within the cell. Like DNA, RNA is not inherently fluorescent and requires biochemical modifications for analysis of its structure-function relationships. Similar to DNA, hybridization probes can the greatest probe specificity and versatility[68,69]. However, hybridization probes often perturb the native secondary structure making regulatory functions of RNA hard to study.

Alternative methods for RNA labeling require two steps: the genetically inserting an aptamer sequence to recognize a fluorophore into the 3'untranslated region of the RNA target and the addition or expression of the fluorophore (see Figure 6)[70,71]. RNA can fold into elaborate secondary structures creating recognition sites for proteins and small molecules to bind with high affinity (also known as aptamers). This property has been extended to create artificial binding sites for biorthogonal fluorogenic small molecules[72-74]. Therefore by genetically incorporating the aptamer tag into the 3'untranslated region of the RNA and adding a cell permeable fluorogenic dye, gene expression can be monitored (Figure 6A)[73]. Other

adaptations of this strategy use a protein binding region coupled with a GFP tagged RNA binding protein [75] (Figure 6B). This method does not require any additional chemical which could cause toxicity or perturbation to the cell.

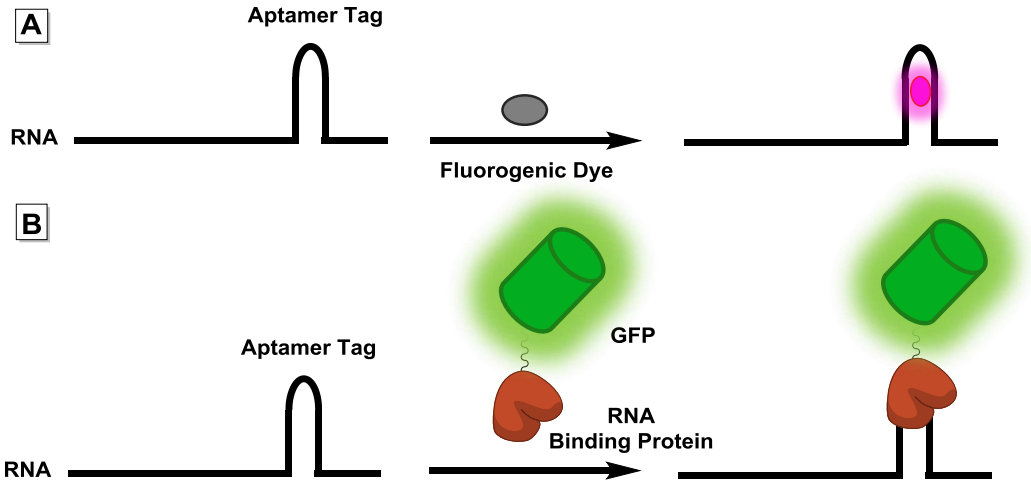


Figure 6: RNA labeling with aptamers (A) using a non-covalent fluorogenic dye with an aptamer; (B) using a selective RNA binding protein fused with GFP with an inserted hairpin aptamer. Figure adapted from Ref 70

2.3. Protein Labeling

Labeling proteins is extremely useful to detect and quantify the protein of interest and its interactions within the cell. A collection of chemical techniques have been used to covalently attach fluorescent tags to proteins. Some strategies include targeting fluorophores with electrophilic substituents to react with nucleophilic lysine or cysteine residues[76,77]. However, these strategies often lack site specificity which could perturb the function of the protein. In addition, the fluorescently tagged protein must be generated in vitro then added to the system of interest. In this section, we will discuss a few general strategies used to label a protein of interest.

2.3.1. Immunofluorescence

Nature has found a solution to target specific proteins by producing antibodies against a target antigen. These antibodies have two functional domains, a variable antigen binding

fragment (Fab) that recognizes the specific antigen, and a conserved fragment (Fc) that the immune system recognizes allowing a targeted response. Scientists have taken advantage of the antibody's specificity by selecting antibodies that are specific to the target[78]. In order to visualize the antigen/ target, the antibody is labeled with a fluorophore.

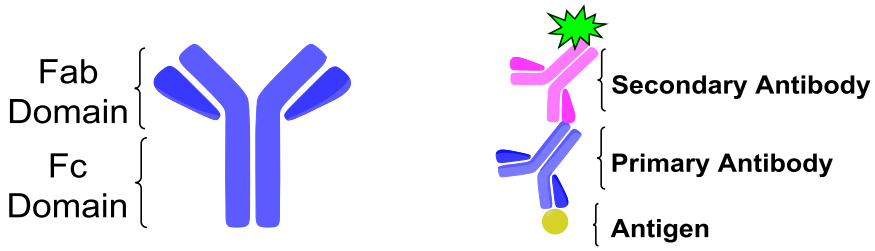


Figure 7: Generic structure of antibody (left); General scheme of labeling with secondary antibody (right).

There are two general strategies for antibody-based detection: direct or indirect immunofluorescence, each with their own merits. Direct immunofluorescence uses a fluorescently-labeled antibody to detect the analyte. This assay preparation is a fast two-step process: incubation with sample and removal of excess antibody. This simple design decreases non-specific targeting. However, this strategy is limiting and costly, requiring a new fluorophore labeled antibody for each analyte[79]. The second method is addition of a secondary antibody that recognizes the conserved fragment (Fc). It is a cost-effective solution allowing for the use of one fluorescently labeled antibody to target multiple primary antibodies. An added advantage includes an amplified signal for more sensitive detection where multiple secondary antibodies can bind to a primary. Nevertheless, both of methods have a few limitations, the antibodies cannot permeate across the cell membrane, restricting this method to surface proteins or fixed cells. Moreover due to the covalent nature of the fluorophore once the fluorophore is photo bleached it cannot be replaced[79].

2.3.2. Fluorescent Proteins

One of the most popular fluorescent proteins is Green Fluorescent Protein (GFP) initially found in nature in a small bioluminescent jellyfish. The beta-barrel protein protects the cyclized tri-amino acid chromophore from quenching from the exterior environment[37]. By genetically mutating a few amino acids various enhancements were achieved e.g. a variety of colors, increased photo stability and faster folding kinetics[80-82].

Fluorescent proteins can be designed to tag a protein of interest by fusing the fluorescent protein in an appropriate vector for the cell to express the desired protein of interest[83]. One potential disadvantage is the size of the fluorescent protein, which may affect the activity and function of the protein of interest. Another disadvantage is irreversible photo-bleaching of the chromophore which occurs during illumination and is due to photo oxidation.

2.3.3. Modular covalent protein/small molecule constructs

An alternative approach to label proteins combines the biocompatibility of a genetically encodable protein with the flexibility of a small molecule. A considerable number of protein tagging systems are now commercially available with SNAP-tag, Halo-Tag and FlAsH/ReAsH being the most widely used. SNAP-tag involves a human-alkylguanine-DNA-alkyltransferase (hAGT) that can be genetically fused to the protein of interest (Figure 8A)[84]. A fluorescent probe with an alkylated guanine can form a covalent bond with the protein allowing control of time and concentration of fluorophore. Halo-tag is a modified halo-alkane dehalogenase that was designed to covalently bind to a synthetic ligand, which can be conjugated to a fluorophore or other affinity tag[85] (Figure 8B). Both of these methods achieve chemical flexibility for a widespread of applications[86].

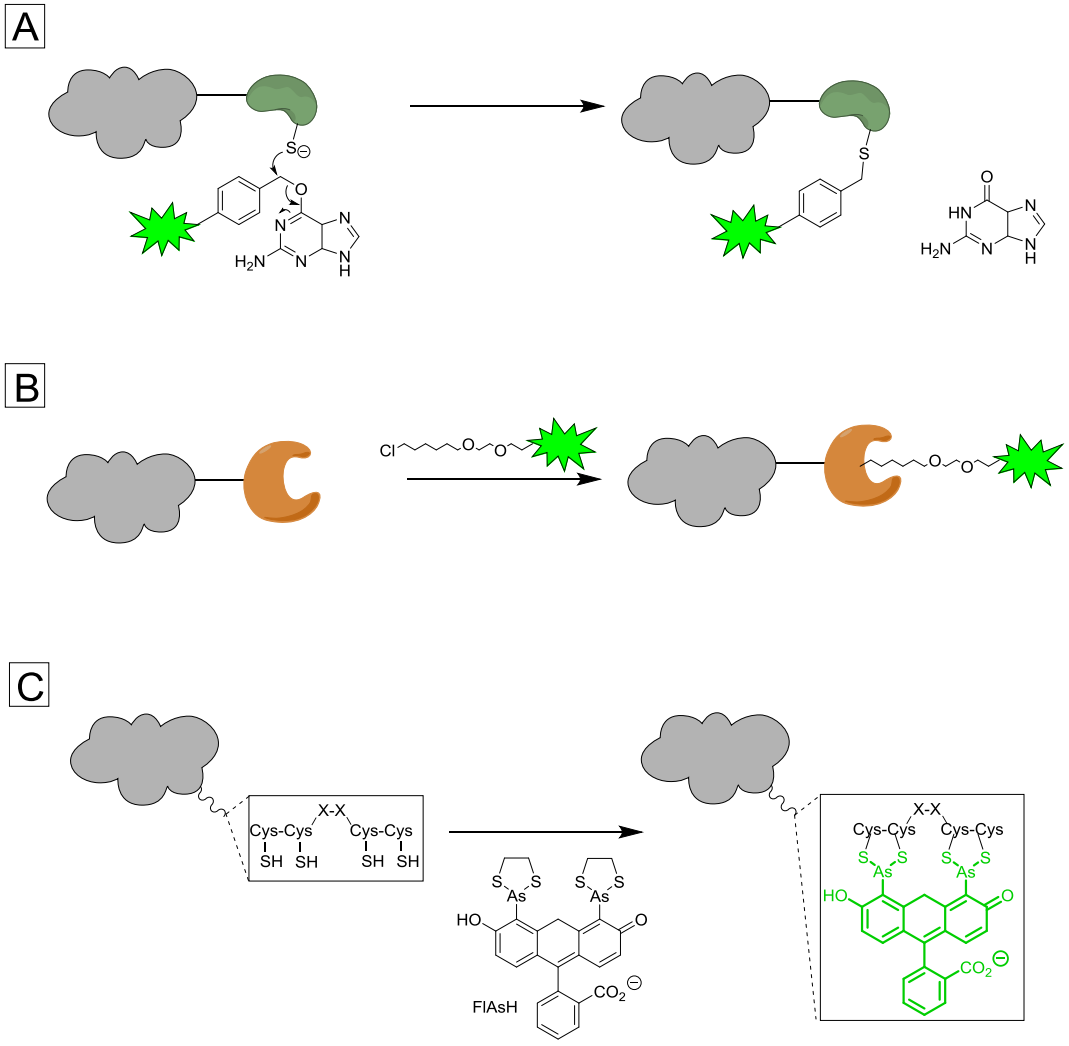


Figure 8: Modular covalent protein/small molecule constructs. (A) SNAP Tag (B) Halo Tag (C) FlAsH. Figure adapted from 86.

In place of a covalent bond, a fluorophore with metal-chelators introduces potential benefits. A pair of bi-arsenical labeling reagents based on the scaffold of fluorescein (FlAsH) and resorufin (ReAsH) are modified to contain two trivalent arsenic coordination sites[87,88] (Figure 8C). In the non-fluorescent state the arsenic is coordinated with ethane dithiol allowing a non-emissive relaxation. However, the compound becomes fluorescent when bound to a small tetra cysteine motif that can be genetically fused to the terminus of a protein. This fluorogenic strategy gives low background and more sensitive detection, with the two colors allowing for pulse chase experiments used in real time imaging. The combination of a small cell permeable

molecule and a small genetically encodeable six amino acid tag allows for live cell imaging. FlAsH/ ReAsH are 30 times smaller and have faster kinetics for fluorescent signal generation than GFP. However, the arsenic has potential cellular toxicity which can be alleviated with additional ethane dithiol added to the cell[87]. In addition, the cysteine content in the protein of interest as well as other proteins in the cell can lead to off target effects producing false signals.

2.3.4. Fluoromodules

Lastly, a non-covalent approach to protein-dye reporters was demonstrated by Molecular Biosensor and Imaging Center (MBIC) at Carnegie Mellon University. This model exploited the diversity of antibodies, coupled with carefully selected dyes that exhibited low fluorescence free in solution but were known to demonstrate large fluorescence enhancements once conformationally constrained[89-95]. These dyes were reported to bind selectively to protein partners with nanomolar affinity and up to thousand fold fluorescence enhancements. Due to the fluorogenic nature of the dye these assays were able to be performed in excess dye without the need to remove unbound dye. An additional advantage to a non-covalent reporter is the ability to restore fluorescence lost due to photo bleaching by the addition of new dye[94,96]. Chapter 2 will focus on developing new dyes for these fluoromodules.

3. Thesis Overview

Chapter 2 focuses on developing new fluorogenic dyes based on a thiazole orange scaffold. By modifying the substituents, a push-pull system yielded a fine tuning effect on the photophysics. There we focus on the fluorescence properties with DNA and protein partners. Chapter 3 and 4 focus on developing new designs and strategies for PNA hybridization probes. Chapter 3 uses fluorescently labeled γPNA and the fluorescence phenomenon FRET to generate a positive signal in the presence of the target with low background. Chapter 4 describes a new PNA monomer that allows internal, potentially high density labeling of PNA for brighter hybridization probes.

4. References

(1) Valeur, B.; Berberan-Santos, M. N.: *Molecular Fluorescence: Principals and Applications*; 2nd ed.; Wiley, 2013.

(2) Lakowicz, J. R.: *Principles of Fluorescence Spectroscopy*; 3rd ed.; Springer, 2006.

(3) Lavis, L. D.; Raines, R. T. Bright Ideas for Chemical Biology. *ACS Chem. Biol.* **2008**, *3*, 142-155.

(4) Teale, F. W. J.; Weber, G. Ultraviolet Fluorescence of the Aromatic Amino Acids. *Biochem. J.* **1957**.

(5) Weber, G. Intramolecular transfer of electronic energy in dihydro diphosphopyridine nucleotide. *Nature* **1957**, *4599*, 1409.

(6) Lavis, L. D.; Raines, R. T. Bright building blocks for chemical biology. *ACS Chem. Biol.* **2014**, *9*, 855-866.

(7) Brouwer, A. M. Standards for photoluminescence quantum yield measurements in solution (IUPAC Technical Report). *Pure App. Chem.* **2011**, *83*, 2213–2228.

(8) Grimm, J. B.; English, B. P.; Chen, J.; Slaughter, J. P.; Zhang, Z.; Revyakin, A.; Patel, R.; Macklin, J. J.; Normanno, D.; Singer, R. H.; Lionnet, T.; Lavis, L. D. A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat. Methods* **2015**, *12*, 244-250.

(9) Magde, D.; Wong, R.; Seybold, P. G. Fluorescence Quantum Yields and Their Relation to Lifetimes of Rhodamine 6G and Fluorescein in Nine Solvents: Improved Absolute Standards for Quantum Yields. *Photochem. Photobiol.* **2002**, *75*, 327-334.

(10) Beija, M.; Afonso, C. A.; Martinho, J. M. Synthesis and applications of Rhodamine derivatives as fluorescent probes. *Chem. Soc. Rev.* **2009**, *38*, 2410-2433.

(11) Karpovich, D. S.; Blanchard, G. J. Relating the polarity-dependent fluorescence response of pyrene to vibronic coupling. Achieving a fundamental understanding of the py polarity scale. *The Journal of Physical Chemistry* **1995**, *99*, 3951-3958.

(12) Haugland, R. P.: The Molecular Probes Handbook—A Guide to Fluorescent Probes and Labeling Technologies. 11 ed. ed.; Molecular Probes: Eugene, OR, 2010.

(13) Sun, W. C.; Gee, K. R.; Haugland, R. P. Synthesis of novel fluorinated coumarins: excellent UV-light excitable fluorescent dyes. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 3107-3110.

(14) Johnson, I. Review: Fluorescent probes for living cells. *Histochem. J.* **1998**, *30*, 123-140.

(15) Mujumdar, R. B.; Ernst, L. A.; Mujumdar, S. R.; Lewis, C. J.; Waggoner, A. S. Cyanine Dye Labeling Reagents: Sulfoindocyanine Succinimidyl Esters. *Bioconjugate Chem.* **1993**, *4*.

(16) Lavis, L. D.; Rutkoski, T. J.; Raines, R. T. Tuning the p a of Fluorescein to Optimize Binding Assays. *Anal. Chem.* **2007**, *79*, 6775-6782.

(17) Sun, W. C.; Gee, K. R.; Klaubert, D. H.; Haugland, R. P. Synthesis of Fluorinated Fluoresceins. *J. Org. Chem.* **1997**, *62*.

(18) Neckers, D. C. Rose Bengal. *J. Photochem. Photobiol. A. Chem.* **1989**, *47*, 1-29.

(19) Wysocki, L. M.; Grimm, J. B.; Tkachuk, A. N.; Brown, T. A.; Betzig, E.; Lavis, L. D. Facile and general synthesis of photoactivatable xanthene dyes. *Angew. Chem., Int. Ed.* **2011**, *50*, 11206-11209.

(20) Lavis, L. D.; Chao, T. T. Y.; Raines, R. T. Fluorogenic label for biomolecular imaging. *Chem. Biol.* **2006**, *1*, 252-260.

(21) Shieh, P.; Hangauer, M. J.; Bertozzi, C. R. Fluorogenic azidofluoresceins for biological imaging. *J. Am. Chem. Soc.* **2012**, *134*, 17428-17431.

(22) Grimm, J. B.; Sung, A. J.; Legant, W. R.; Hulamm, P.; Matlosz, S. M.; Betzig, E.; Lavis, L. D. Carbofluoresceins and carborhodamines as scaffolds for high-contrast fluorogenic probes. *ACS Chem. Biol.* **2013**, *8*, 1303-1310.

(23) Belov, V. N.; Mitronova, G. Y.; Bossi, M. L.; Boyarskiy, V. P.; Hebisch, E.; Geisler, C.; Kolmakov, K.; Wurm, C. A.; Willig, K. I.; Hell, S. W. Masked rhodamine dyes of five principal colors revealed by photolysis of a 2-diazo-1-indanone caging group: synthesis, photophysics, and light microscopy applications. *Chem.—Eur. J.* **2014**, *20*, 13162-13173.

(24) Smalley, M. K.; Silverman, S. K. Fluorescence of covalently attached pyrene as a general RNA folding probe. *Nucleic Acids Res.* **2005**, *34*, 152-166.

(25) Zhao, C.; Zhang, Y.; Wang, X.; Cao, J. Development of BODIPY-based fluorescent DNA intercalating probes. *J. Photochem. Photobio., A* **2013**, *264*, 41-47.

(26) Loudet, A.; Burgess, K. BODIPY dyes and their derivatives: syntheses and spectroscopic properties. *Chem. Rev.* **2007**, *107*, 4891-4932.

(27) Wells, G.; Suggitt, M.; Coffils, M.; Baig, M. A.; Howard, P. W.; Loadman, P. M.; Hartley, J. A.; Jenkins, T. C.; Thurston, D. E. Fluorescent 7-diethylaminocoumarin pyrrolobenzodiazepine conjugates: synthesis, DNA interaction, cytotoxicity and differential cellular localization. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2147-2151.

(28) Benson, R. C.; Kues, H. A. Absorption and Fluorescence Properties of Cyanine Dyes. *J. Chem. Eng. Data* **1977**, *22*.

(29) Mishra, A.; Behera, R. K.; Behera, P. K.; Mishra, B. K.; Behera, G. B. Cyanines during the 1990s: A Review. *Chem. Rev.* **2000**, *100*, 1973−2011.

(30) Jaiswal, J. K.; Mattoussi, H.; Mauro, J. M.; Simon, S. M. Long-term multiple color imaging of live cells using quantum dot bioconjugates. *Nat. Biotechnol.* **2002**, *21*, 47-51.

(31) Giepmans, B. N.; Deerinck, T. J.; Smarr, B. L.; Jones, Y. Z.; Ellisman, M. H. Correlated light and electron microscopic imaging of multiple endogenous proteins using Quantum dots. *Nat. Methods* **2005**, *2*, 743-749.

(32) Resch-Genger, U.; Grabolle, M.; Cavaliere-Jaricot, S.; Nitschke, R.; Nann, T. Quantum dots versus organic dyes as fluorescent labels. *Nat. Methods* **2008**, *5*, 763-775.

(33) Parak, W. J.; Pellegrino, T.; Plank, C. Labelling of cells with quantum dots. *Nanotechnology* **2005**.

(34) Liang, H.; Zhang, X. B.; Lv, Y.; Gong, L.; Wang, R.; Zhu, X.; Yang, R.; Tan, W. Functional DNA-containing nanomaterials: cellular applications in biosensing, imaging, and targeted therapy. *Acc. Chem. Res.* **2014**, *47*, 1891-1901.

(35) Selvan, S. T.; Tan, T. T. Y.; Yi, D. K.; Jana, N. R. Functional and multifunctional nanoparticles for bioimaging and biosensing. *Langmuir : the ACS journal of surfaces and colloids* **2010**, *26*, 11631-11641.

(36) Gupta, A. K.; Gupta, M. Synthesis and surface engineering of iron oxide nanoparticles for biomedical applications. *Biomaterials* **2005**, *26*, 3995-4021.

(37) Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **1997**, *67*, 509-544.

(38) Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. The fluorescent toolbox for assessing protein location and function. *Science* **2006**.

(39) Zheng, Q.; Juette, M. F.; Jockusch, S.; Wasserman, M. R.; Zhou, Z.; Altman, R. B.; Blanchard, S. C. Ultra-stable organic fluorophores for single-molecule research. *Chem. Soc. Rev.* **2014**, *43*, 1044-1056.

(40) Wu, X.; Zhu, W. Stability enhancement of fluorophores for lighting up practical application in bioimaging. *Chem. Soc. Rev.* **2014**.

(41) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J. The sequence of the human genome. *Science* **2001**.

(42) Watson, J. D.; Crick, F. H. C. Molecular structure of nucleic acids. *Nature* **1953**.

(43) Martin, R. M.; Leonhardt, H.; Cardoso, M. C. DNA labeling in living cells. *Cytometry* **2005**, *67*, 45-52.

(44) Armitage, B. A. Cyanine Dye–DNA Interactions: Intercalation, Groove Binding,and Aggregation. *Top. Curr. Chem.* **2005**, *253*, 55-76.

(45) Lee, L. G.; Chen, C.; Chiu, L. Thiazole Orange: A New Dye for Reticulocyte Analysis *Cytometry* **1986**, *7*, 508-517.

(46) Goldman, J. M.; Zhang, L. A.; Manna, A.; Armitage, B. A.; Ly, D. H.; Schneider, J. W. High affinity gammaPNA sandwich hybridization assay for rapid detection of short nucleic acid targets with single mismatch discrimination. *Biomacromolecules* **2013**, *14*, 2253-2261.

(47) Benson, S. C.; Rye, H. S.; Glazer, A. N.; Mathies, R. A. High-sensitivity capillary electrophoresis of double-stranded DNA fragments using monomeric and dimeric fluorescent intercalating dyes. *Anal. Chem.* **1994**.

(48) Cosa, G.; Focsaneanu, K. S.; McLean, J. R.; McNamee, J. P.; Scaiano, J. C. Photophysical properties of fluorescent DNA-dyes bound to single- and double-stranded DNA in aqueous buffered solution. *Photochem. Photobiol.* **2001**, *73*, 585-599.

(49) Silva, G. L.; Ediz, V.; Yaron, D.; Armitage, B. A. Experimental and Computational Investigation of Unsymmetrical Cyanine Dyes: Understanding Torsionally Responsive Fluorogenic Dyes. *J. Am. Chem. Soc.* **2007**, *129*, 5710-5718.

(50) Nygren, J.; Svanvik, N.; Kubista, M. The Interactions Between the Fluorescent Dye Thiazole Orange and DNA. *Biopolymers* **1998**, *46*, 39–51.

(51) Netzel, T. L.; Nafisi, K.; Zhao, M.; Lenhard, J. R. Base-content dependence of emission enhancements, quantum yields, and lifetimes for cyanine dyes bound to double-strand DNA: photophysical properties of …. *J. Phys. Chem.* **1995**.

(52) Krugh, T. R.; Wittlin, F. N.; Cramer, S. P. Ethidium bromide–dinucleotide complexes. Evidence for intercalation and sequence preferences in binding to double stranded nucleic acids. *Biopolymers* **1975**, *14*, 197-210.

(53) Stender, H.; Fiandaca, M.; Hyldig-Nielsen, J. J. PNA for rapid microbiology. *J. Microbiol. Methods* **2002**.

(54) Martí, A. A.; Jockusch, S.; Stevens, N.; Ju, J.; Turro, N. J. Fluorescent hybridization probes for sensitive and selective DNA and RNA detection. *Acc. Chem. Res.* **2007**, *40*, 402-409.

(55) Ranasinghe, R. T.; Brown, T. Fluorescence based strategies for genetic analysis. *Chem. Commun.* **2005**.

(56) Sapsford, K. E.; Berti, L.; Medintz, I. L. Materials for fluorescence resonance energy transfer analysis: beyond traditional donor–acceptor combinations. *Angew. Chem., Int. Ed.* **2006**.

(57) Tyagi, S.; Kramer, F. R. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* **1996**, *14*, 303-309.

(58) Randolph, J. B.; Waggoner, A. S. Stability, specificity and fluorescence brightness of multiply-labeled fluorescent DNA probes. *Nucleic acids research* **1997**, *25*.

(59) Teo, Y. N.; Kool, E. T. DNA-multichromophore systems. *Chem. Rev.* **2012**, *112*, 4221-4245.

(60) Hocek, M. Synthesis of base-modified 2'-deoxyribonucleoside triphosphates and their use in enzymatic synthesis of modified DNA for applications in bioanalysis and chemical biology. *J. Org. Chem.* **2014**, *79*, 9914-9921.

(61) Schlapak, R.; Kinns, H.; Wechselberger, C. Sizing Trinucleotide Repeat Sequences by Single-Molecule Analysis of Fluorescence Brightness. *Chem. Phys. Chem.* **2007**.

(62) Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Erlich, H. A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **1988**, *239*, 487-491.

(63) Horáková, P.; Macíčková-Cahová, H.; Pivoňková, H.; Spaček, J.; Havran, L.; Hocek, M.; Fojta, M. Tail-labelling of DNA probes using modified deoxynucleotide triphosphates and terminal

deoxynucleotidyl transferase. Application in electrochemical DNA hybridization and protein-DNA binding assays. *Org. Biomol. Chem.* **2011**, *9*, 1366-1371.

(64) Hannon, G. J.; Rossi, J. J. Unlocking the potential of the human genome with RNA interference. *Nature* **2004**.

(65) Ørom, U. A.; Derrien, T.; Beringer, M.; Gumireddy, K. Long noncoding RNAs with enhancer-like function in human cells. *Cell* **2010**.

(66) Mali, P.; Yang, L.; Esvelt, K. M.; Aach, J.; Guell, M.; DiCarlo, J. E. RNA-guided human genome engineering via Cas9. *Science* **2013**.

(67) Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **2004**.

(68) Femino, A. M.; Fay, F. S.; Fogarty, K.; Singer, R. H. Visualization of Single RNA Transcripts in Situ. *Science* **1998**, *280*, 585-590.

(69) Raj, A.; van den Bogaard, P.; Rifkin, S. A.; van Oudenaarden, A.; Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **2008**, *5*, 877-879.

(70) Armitage, B. A. Imaging of RNA in live cells. *Curr. Opin. Chem. Biol.* **2011**, *15*, 806-812.

(71) Novikova, I. V.; Afonin, K. A.; Leontis, N. B. New Ideas for in vivo Detection of RNA. *Biosensors* **2010**.

(72) Constantin, T. P.; Silva, G. L.; Robertson, K. L.; Hamilton, T. P.; Fague, K.; Waggoner, A. S.; Armitage, B. A. Synthesis of New Fluorogenic Cyanine Dyes and Incorporation into RNA Fluoromodules. *Org. Lett.* **2008**, *10*, 1561-1564.

(73) Pothoulakis, G.; Ceroni, F.; Reeve, B.; Ellis, T. The Spinach RNA aptamer as a characterization tool for synthetic biology. *ACS Synth. Biol.* **2013**.

(74) Famulok, M. Chemical biology: Green fluorescent RNA. *Nature* **2004**.

(75) Bertrand, E.; Chartrand, P.; Schaefer, M.; Shenoy, S. M.; Singer, R. S.; Long, R. M. Localization of ASH1 mRNA Particles in Living Yeast. *Mol. Cell* **1998**, *2*, 437–445.

(76) Ramil, C. P.; Lin, Q. Bioorthogonal chemistry: strategies and recent developments. *Chem. Commun.* **2013**, *49*, 11007-11022.

(77) Kim, E.; Lee, Y.; Lee, S.; Park, S. B. Discovery, understanding, and bioapplication of organic fluorophore: a case study with an indolizine-based novel fluorophore, seoul-fluor. *Acc. Chem. Res.* **2015**, *48*, 538-547.

(78) Schnell, U.; Dijk, F.; Sjollema, K. A.; Giepmans, B. N. G. Immunolabeling artifacts and the need for live-cell imaging. *Nat. Methods* **2012**.

(79) Stadler, C.; Rexhepaj, E.; Singan, V. R.; Murphy, R. F.; Pepperkok, R.; Uhlen, M.; Simpson, J. C.; Lundberg, E. Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nat. Methods* **2013**, *10*, 315-323.

(80) Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11984-11989.

(81) Hazelwood, K. L.; Davidson, M. W.; Tsien, R. Y. Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nature* **2008**.

(82) Matz, M. V.; Fradkov, A. F.; Labas, Y. A.; Savitsky, A. P.; Zaraisky, A. G.; Markelov, M. L.; Lukyanov, S. A. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* **1999**, *17*, 969-973.

(83) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasherf, D. C. Green Fluorescent protein as a marker for gene expression. *Science* **1994**.

(84) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature* **2003**.

(85) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.;

Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *Chem. Biol.* **2008**, *3*, 373-382.

(86) Jing, C.; Cornish, V. W. Chemical Tags for Labeling Proteins Inside Living Cells. *Acc. Chem. Res.* **2011**.

(87) Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. New Biarsenical Ligands and Tetracysteine Motifs for Protein Labeling in Vitro and in Vivo: Synthesis and Biological Applications. *J. Am. Chem. Soc.* **2002**, *124*, 6063-6076.

(88) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* **1998**, *281*, 269-272.

(89) Szent-Gyorgyi, C.; Schmidt, B. F.; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. A.; Woolford, C. A.; Yan, Q.; Vasilev, K. V.; Berget, P. B.; Bruchez, M. P.; Jarvik, J. W.; Waggoner, A. Fluorogen-activating single-chain antibodies for imaging cell surface proteins. *Nat. Biotechnol.* **2008**, *26*, 235-240.

(90) Saunders, M. J.; Block, E.; Sorkin, A.; Waggoner, A. S.; Bruchez, M. P. A bifunctional converter: fluorescein quenching scFv/fluorogen activating protein for photostability and improved signal to noise in fluorescence experiments. *Bioconjugate Chem.* **2014**, *25*, 1556-1564.

(91) Saunders, M. J.; Liu, W.; Szent-Gyorgyi, C.; Wen, Y.; Drennen, Z.; Waggoner, A. S.; Meng, W. S. Engineering fluorogen activating proteins into self-assembling materials. *Bioconjugate Chem.* **2013**, *24*, 803-810.

(92) Zanotti, K. J.; Silva, G. L.; Creeger, Y.; Robertson, K. L.; Waggoner, A. S.; Berget, P. B.; Armitage, B. A. Blue fluorescent dye-protein complexes based on fluorogenic cyanine dyes and single chain antibody fragments. *Org. Biomol. Chem.* **2011**, *9*, 1012-1020.

(93) Senutovitch, N.; Stanfield, R. L.; Bhattacharyya, S.; Rule, G. S.; Wilson, I. A.; Armitage, B. A.; Waggoner, A. S.; Berget, P. B. A variable light domain fluorogen activating protein homodimerizes to activate dimethylindole red. *Biochemistry* **2012**, *51*, 2471-2485.

(94) Shank, N. I.; Pham, H. H.; Waggoner, A. S.; Armitage, B. A. Twisted cyanines: a non-planar fluorogenic dye with superior photostability and its use in a protein-based fluoromodule. *J. Am. Chem. Soc.* **2013**, *135*, 242-251.

(95) Özhalici-Ünal, H.; Lee Pow, C.; Marks, S. A.; Jesper, L. D.; Silva, G. L.; Shank, N. I.; Jones, E. W.; Burnette, J. M., III; Berget, P. B.; Armitage, B. A. A Rainbow of Fluoromodules: A Promiscuous scFv Protein Binds to and Activates a Diverse Set of Fluorogenic Cyanine Dyes. *J. Am. Chem. Soc.* **2008**, *130*, 12620-12621.

(96) Shank, N. I.; Zanotti, K. J.; Lanni, F.; Berget, P. B.; Armitage, B. A. Enhanced Photostability of Genetically Encodable Fluoromodules Based on Fluorogenic Cyanine Dyes and a Promiscuous Protein Partner. *J. Am. Chem. Soc.* **2009**, *131*, 12961-12969.

Chapter 2: Spectral Fine Tuning of Cyanine Dyes: Electron Donor-Acceptor Substituted Analogues of Thiazole Orange

2.1. Chapter Summary

The introduction of electron donor and acceptor groups at strategic locations on a fluorogenic cyanine dye allows fine-tuning of the absorption and emission spectra while preserving the ability of the dye to bind to biomolecular hosts such as double-stranded DNA and a single-chain antibody fragment originally selected for binding to the parent unsubstituted dye, thiazole orange (TO). The observed spectral shifts are consistent with calculated HOMO-LUMO energy gaps and reflect electron density localization on the quinoline half of TO in the LUMO. A dye bearing donating methoxy and withdrawing trifluoromethyl groups on the benzothiazole and quinoline rings, respectively, shifts the absorption spectrum to sufficiently longer wavelengths to allow excitation at green wavelengths as opposed to the parent dye, which is optimally excited in the blue.

2.2. Introduction

Fluorogenic dyes are a special class of fluorophores that show distinctly different emission intensities depending on their local environment[1-7]. For example, unsymmetrical cyanine dyes exhibit low fluorescence quantum yields in fluid solution but strongly enhanced emission in viscous solution or other environments that conformationally constrain the dyes[8,9]. This phenomenon arises from a nonradiative twisting pathway about the central methine bridge that is inhibited when the dye is constrained[8,10]. The unsymmetrical cyanines (Chart 1) were originally used as fluorogenic DNA stains where intercalation into the DNA base pair stack leads to >100-fold fluorescence enhancements[11-13]. Subsequently, conjugation of unsymmetrical cyanines to various classes of molecules (e.g. DNA, peptides, PNA) yielded “light-up probes” that exhibit enhanced fluorescence upon binding to another molecule[14-24]. More recently, combination of fluorogenic cyanines and other dyes with single chain antibody fragment

partners has allowed creation of a modular catalogue of dye-protein complexes with absorption and emission spectra spanning most of the visible spectrum[25-34]. The dye and protein are each nonfluorescent separately, but become strongly fluorescent upon formation of a noncovalent complex, with quantum yields up to 100%[29].

Cyanine dye colors can be designed to extend over the entire visible and near-IR range through variation of the length of the central polymethine bridge (*n* = 1, 3, 5, 7, 9) and the identity of the heterocycles (e.g. dimethylindole, benzothiazole, benzoxazole, quinoline)[35,36]. For example, each increase in bridge length (e.g. *n* = 1 to *n* = 3) results in an approximate 100 nm red shift of the absorption and emission spectra. Addition of substituents to the polymethine bridge can also result in significant spectral shifts, although conformational and electronic factors can sometimes offset one another, leading to minimal change in color[27,30,37,38].

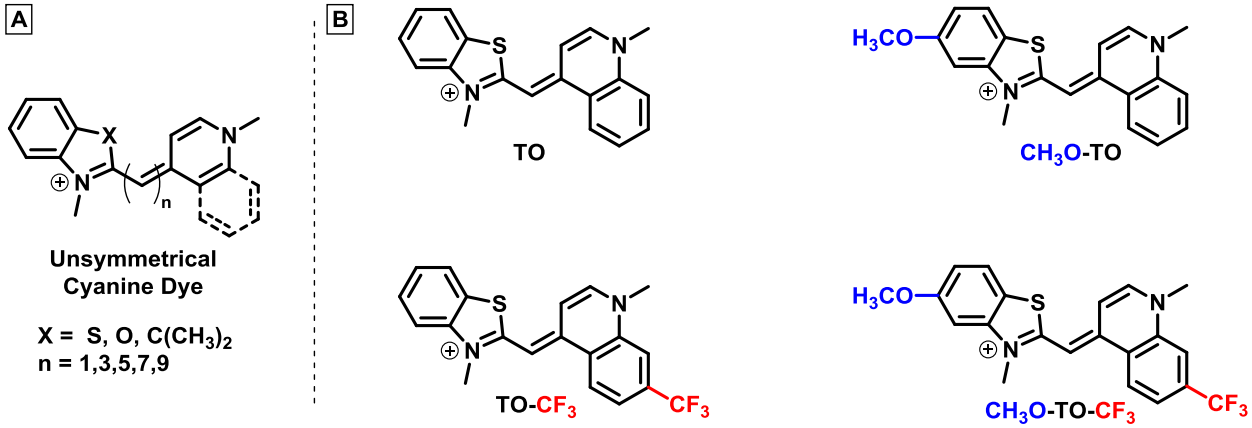


Chart 1: Unsymmetrical cyanine dyes (A) General structure of unsymmetrical cyanine dyes; (B) TO Dyes with EWG (Red) and/or EDG (Blue)

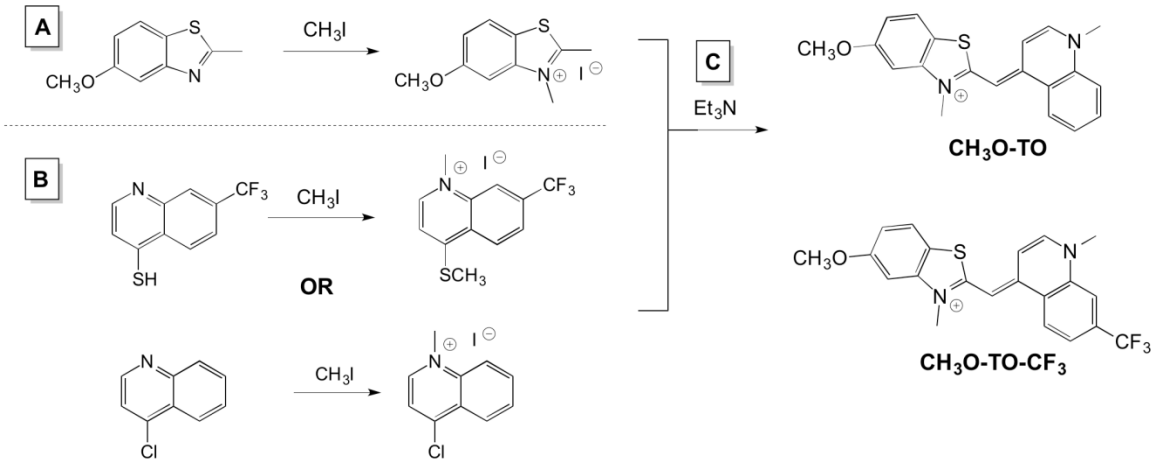
Alternatively, fine-tuning of the cyanine dye spectra can be accomplished through introduction of substituents on the aromatic heterocycles[10]. For example, substitution of electron-withdrawing fluorines for hydrogens on the benzothiazole ring of TO led to blue-shifted spectra, with the magnitude of the shift correlating with the number of fluorine atoms. On the other hand, substitution of a trifluoromethyl group on the quinoline ring of TO led to a red-shift. These observations were rationalized in terms of the frontier orbitals: the HOMO has more electron density on the benzothiazole ring, so EWGs on the benzothiazole will stabilize the HOMO more than the LUMO, leading to a larger HOMO-LUMO gap and therefore blue-shifted spectra. Conversely, EWGs on the quinoline will preferentially stabilize the LUMO, leading to a smaller HOMO-LUMO gap and red-shifted spectra.

The prior results for TO analogues lead to the prediction that EDGs on the benzothiazole ring will result in red-shifted spectra, due to preferential destabilization of the HOMO, thereby reinforcing the effects of EWGs on the quinoline ring. To test this prediction, we synthesized two new TO analogues bearing an electron-donating methoxy group on the benzothiazole ring and characterized their spectral properties computationally, in solution, and when bound to either DNA or a TO-binding protein.

2.3. Results

2.3.1. Synthesis of Thizaole Orange Derivatives

Two new methoxy-substituted TO derivatives, CH_3O-TO and CH_3O-TO-CF_3 (Chart 1) were synthesized as their iodide salts by a standard route[39] (Scheme 1): (A) *N*-methylation of 5-methoxy-2-methylbenzothiazole, (B) methylation of either 7-(trifluoromethyl)quinoline-4-thiol or 4-chloroquinoline, and (C) condensation in the presence of triethylamine to yield the final dyes.



Scheme 1: Synthesis of two methoxy substituted TO derivatives. (A) Alkylation of Benzothiazole, neat in Iodomethane at 100°C; (B) Alkylation of Quinoline with Iodomethane in Acetonitrile; (C) Condensation in the presence of Triethylamine in methanol.

2.3.2. Electronic Effects on Homo Lumo

Figure 1 shows the calculated frontier orbitals for TO, along with the effects of the substituents on the HOMO and LUMO energies and on the excitation energies predicted by TDDFT. The trends in the HOMO-LUMO gaps correlate well with the excitation energies, such that the spectral shifts can be understood in terms of the effects of the substituents on the frontier orbitals. These effects are consistent with EWGs (F and CF_3) preferentially stabilizing, and EDGs (methoxy) preferentially destabilizing, the frontier orbital whose density is largest on the heterocycle to which the group is attached. Fluorine substitution on the benzothiazole thus

preferentially stabilizes the HOMO and leads to blue-shifted absorption, while CF_3 substitution on the quinoline preferentially stabilizes the LUMO and leads to red-shifted absorption. Methoxy substitution on the benzothiazole preferentially destabilizes the HOMO, leading to red-shifted absorption, as observed previously by Ivanov and co-workers.[40] The substituent effects reinforce one another in CH_3O-TO-CF_3, leading to the largest red shift (Table 1).

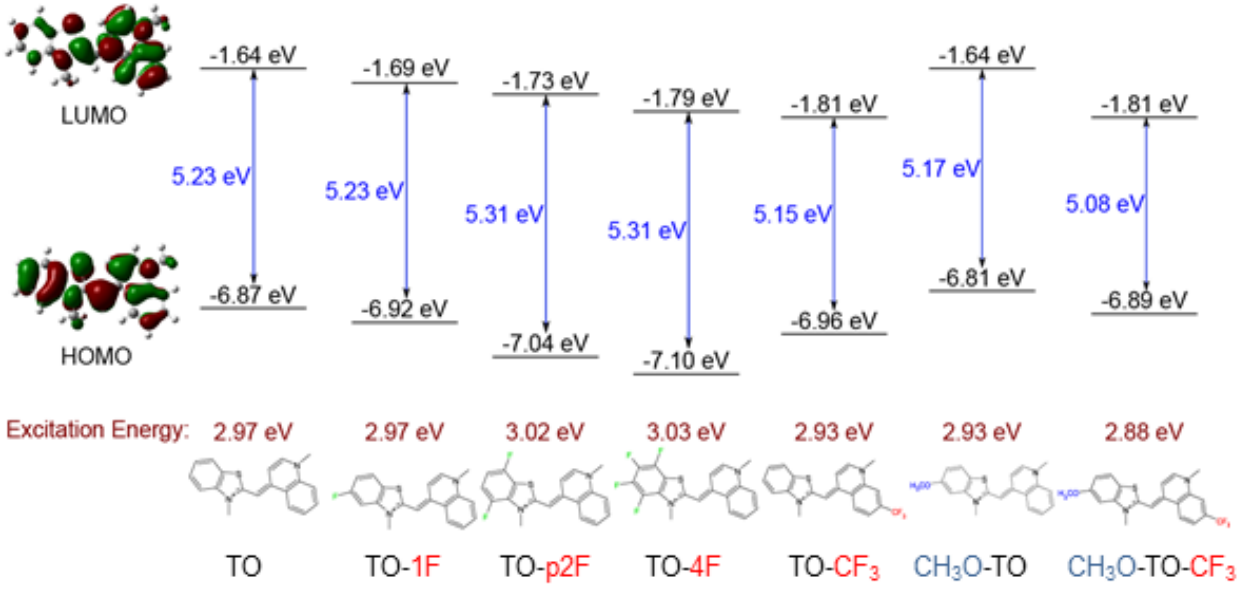


Figure 1: Frontier orbital analysis for TO and its donor/acceptor substituted analogues, including HOMO and LUMO orbitals of TO (far left), HOMO and LUMO energies (black), HOMO-LUMO gaps (blue arrows) and TDDFT excitation energies (red text). Figure and Calculations by Matteus Tanha and Dr. David Yaron.

2.3.3. Characterization in Various Solutions

Normalized UV-vis spectra for the new TO derivatives in methanol solution are shown in Figure 2A, along with the spectra for TO and the previously reported TO-CF_3. Trifluoromethylation of the quinoline group causes a 14 nm red shift in the absorption spectrum relative to TO. Introduction of a methoxy group at position 5 on the benzothiazole heterocycle causes a similar red shift of 10 nm, and the combination of both substituents, methoxy on the benzothiazole and trifluoromethyl on the quinoline, shows an additive effect with a red shift of 24 nm (Table 1).

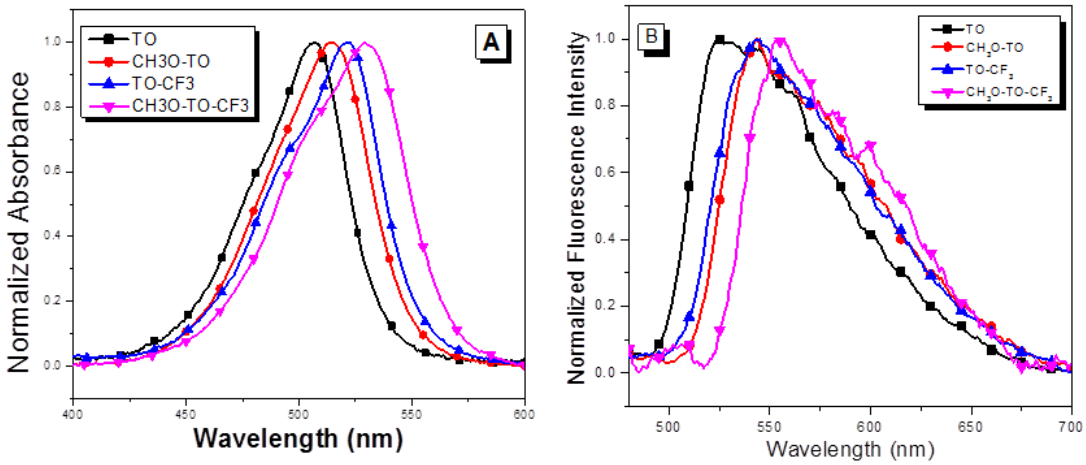


Figure 2. (A) Normalized absorbance spectra of TO dyes (1μM) in methanol. (B) Normalized fluorescence spectra of TO dyes (1μM) in 90% glycerol in water λ_{ex} = 470nm.

Similar results were observed in fluorescence emission spectra recorded for the dyes (Figure 2B and Table 1). Since TO and its analogues exhibit very low fluorescence in fluid solution, we measured emission spectra in a viscous solvent consisting of 90% glycerol in water. The spectra are relatively broad, but the emission clearly red-shifts in the order TO < TO-CF_3 ≈ CH_3O-TO < CH_3O-TO-CF_3.

Table 1. Extinction coefficients, wavelength maxima (nm) and fluorescence quantum yields for TO dyes in homogeneous solution. [a] in methanol, [b] in 90% glycerol in water

Dye	ε ($M^{-1}cm^{-1}$)[a]	Absλ_{max} [a]	Emλ_{max} [b]	ϕ_f [b]
TO	68,000	502	529	0.013
CH_3O-TO	68,600	512	545	0.009
TO-CF_3	65,000	516	547	0.009
CH_3O-TO-CF_3	58,000	526	560	0.007

The fluorescence quantum yields were also determined for TO analogs in 90% glycerol in water (Table 1; the values for TO and TO-CF_3 are lower than we previously reported; this could be due to slight variation in the glycerol content.) As previously reported the addition of the trifluoromethyl group to the quinoline side decreases the quantum yield due to reduction in the activation barrier for twisting about the central methine in the excited state. Addition of a

methoxy substituent to the benzothiazole group also decreases the quantum yield, and when both substituents are present the quantum yield decreases further.

2.3.4 Dye Aggregation in Aqueous Solvents

Studies have shown that cyanine dyes have a tendency to aggregate in aqueous solutions. These aggregate species (e.g. Dimers, H or J aggregates) can be observed in spectral shoulders or formation of new peaks as the concentration of dye increases[19,41]. The absorbance spectrum of a cyanine dye in methanol or ethanol normally will depict the monomeric dye. However in water as the concentration of dye increases typically the intensity of the λmax in methanol will decrease and a shoulder will start to appear. We first observed the absorbance spectra of each of the TO dyes at 10μM in methanol and 10% methanol in water to note any spectral changes (See Figure 3).

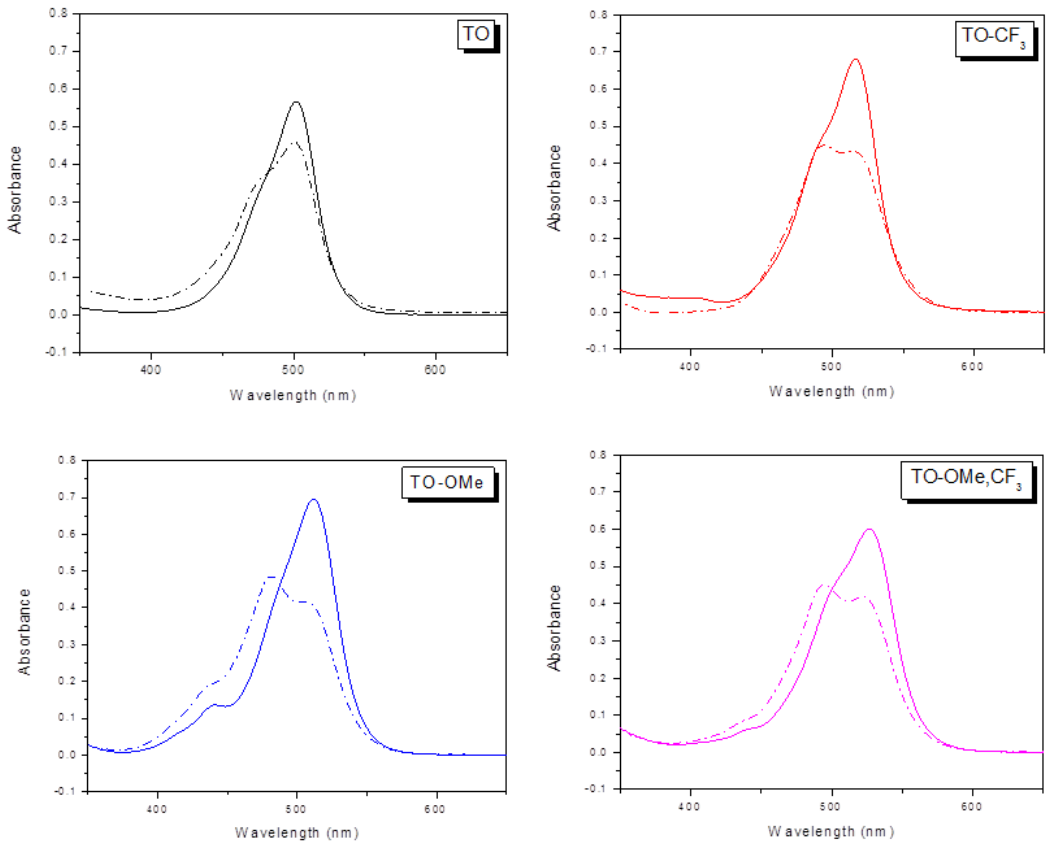


Figure 3: Absorbance Spectra of TO analogs (10μM) in Methanol (Solid) and 10% Methanol in Water (Dashed).

As seen in Figure 3, there is a change in absorbance spectra for each dye. TO shows the least change, however there is a slight decrease at the λmax and an appearance of a shoulder at 478nm. TO-CF_3, CH_3O-TO and CH_3O-TO-CF_3 all exhibit a decrease at λmax and the formation of a second peak at a shorter wavelength. It is suggested that this new maximum due to formation of an H-dimer or aggregate.

Interpretation of the spectra suggests that the methoxy substituent increases the aggregation due to the increased absorbance of the dimeric peak at 481nm and the decrease of the bathochromic monomeric peak at 506 nm. In order to compare quantitatively, a titration of dye (0-10μM) in a solution of 10% methanol in water was performed and the absorbances at the monomer and dimer wavelength were recorded. The $Abs_{dimer}/Abs_{monomer}$ indicates the change in species; this ratio was plotted as a function of dye concentration (Figure 4) where a positive slope would indicate a preference for dimer species at high concentrations

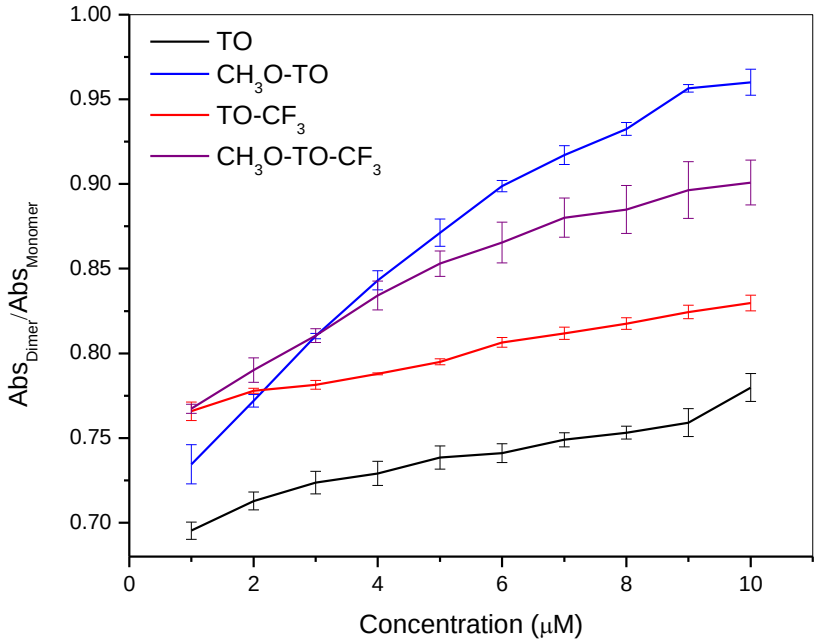


Figure 4: TO dyes in 10% methanol in water. Plot the relationship between dimer:monomer absorbance as the dye concentration increases. (0-10μM). For specific wavelengths for each dye see experimental.

As illustrated in Figure 4, CH_3O-TO showed a substantial increase which indicates a preference for the dimeric species in aqueous solvents. TO showed the least change, which is in agreement with the absorbance spectra in Figure 3. TO-CF_3 and CH_3O-TO-CF_3 showed intermediate aggregation which suggests that the methoxy substituent may encourage dimerization in aqueous solvents. This is consistent with a trend observed in symmetric dyes where the di-substituted methoxy $DiSC_2$ (see Figure 5) showed a strong tendency to aggregate in aqueous solvent when compared to the difluorinated DiSC2 in water[42].

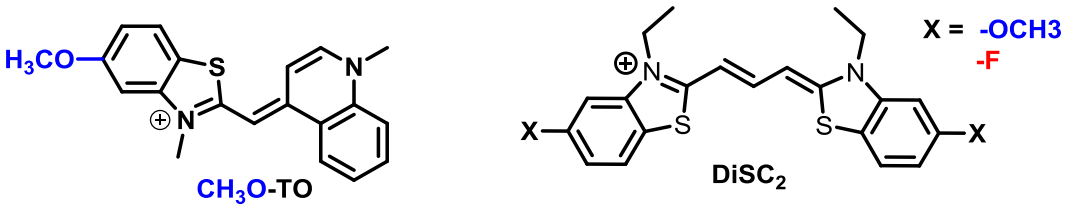


Figure 5: Comparison of structure of methoxy-thiazole orange and di-methoxy $DiSC_2$

This is consistent with a trend observed in symmetric dyes where the di-substituted methoxy $DiSC_2$ showed a strong tendency to aggregate in aqueous solvent when compared to the difluorinated DiSC2 in water[42].

2.3.5 Thiazole Orange Derivatives and Biomolecules

The noncovalent binding interactions between the TO analogues and two biomolecular hosts, namely double-stranded DNA and a single-chain antibody fragment (scFv) protein called K7, which was previously shown to bind to a range of unsymmetrical cyanine dyes, were also investigated. K_D values for the four dyes to K7 are all in the low nanomolar range (Table 2). Figure 6 shows unnormalized absorbance spectra for the four dyes in glycerol/water or excess DNA or K7. The spectra exhibit the same progressive red-shift as the donor-acceptor nature of the substituents increases that was evident in methanol. Also, all four dyes exhibit greater

vibronic structure upon binding to DNA and K7, as evidenced by the more prominent short wavelength shoulders observed in the corresponding absorbance spectra compared with the glycerol spectra, indicating that the dyes experience a narrower range of conformations when bound to the biopolymers.

Table 2. Wavelength maxima (nm), fluorescence quantum yields and binding affinities for TO dyes in the presence of DNA and a dye-binding protein. [a] KD values in nM

Dye	$\lambda_{Abs}/\lambda_{Em}$	ϕ_f (DNA)	$\lambda_{Abs}/\lambda_{Em}$ (Protein)	ϕ_f (Protein)	K_D (Protein)[a]
TO	511/528	0.14	502/518	0.17	2.1 ± 0.5
CH_3O-TO	520/543	0.10	511/534	0.11	3.8 ± 0.6
TO-CF_3	527/548	0.05	516/535	0.21	1.0 ± 0.2
CH_3O-TO-CF_3	536/561	0.04	527/548	0.17	5.4 ± 1.2

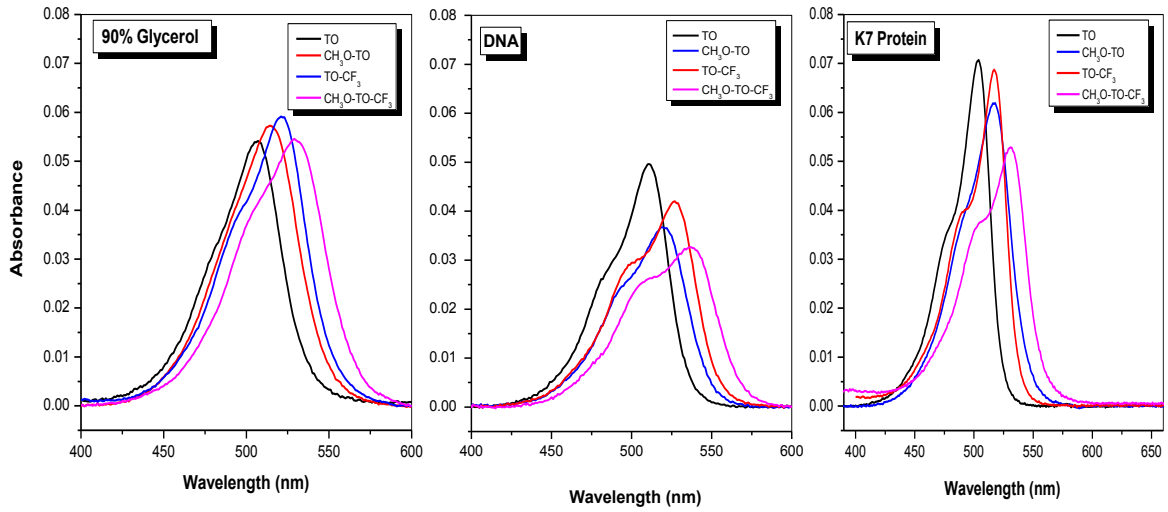


Figure 6. Visible absorbance spectra for TO dyes recorded in 90% glycerol (left), calf thymus DNA (middle) or K7 scFv protein (right). [Dye] = 1 μM, [DNA] = 100 μM base pairs, [K7] = 1 μM.

Fluorescence spectra recorded for the four dyes in methanol, glycerol, DNA and K7 protein are overlaid in Figure 7. The spectra were corrected for absorbance differences at the excitation wavelength and therefore reflect differences in quantum yields. In all four cases, the brightest fluorescence is observed when the dye is bound to the scFv protein, with intensities

10-20-fold brighter than in glycerol. Relative to TO, substitution with CH_3O causes approximately 2-fold decrease in fluorescence for both DNA and K7. In contrast, substitution with CF_3 results in nearly 2-fold *increase* in K7 fluorescence but 5-fold *decrease* in DNA fluorescence relative to TO. The disubstituted dye exhibits K7 fluorescence that is comparable to TO, but the lowest DNA fluorescence of any of the dyes. Overall, these results indicate that the protein tolerates both substituents reasonably well, but methoxy substitution on the benzothiazole ring is detrimental to fluorescence activation by DNA. (Under the conditions of the experiment, all dyes were completely bound by the DNA and protein, so the differences in intensities reflect quantum yield differences rather than affinity differences.)

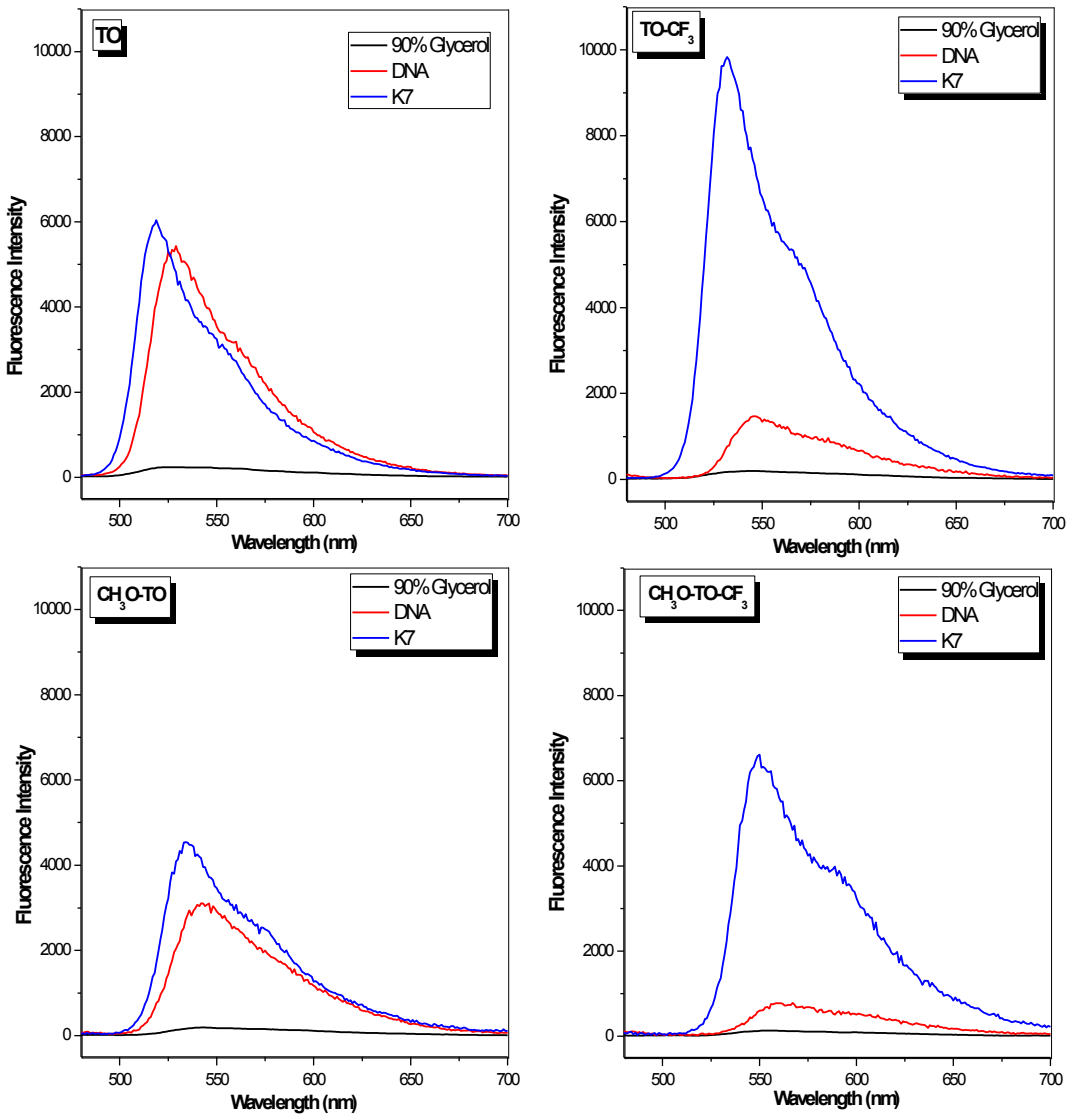


Figure 7. Fluorescence spectra for the four dyes recorded in 90% glycerol (blue), calf thymus DNA (red) or K7 scFv protein (violet). [Dye] = 1 μM, [DNA] = 100 μM base pairs, [K7] = 1 μM. λ_{ex} = 470 nm; spectra were corrected for differences in absorbance.

2.3.6 Thiazole Orange Derivatives and Imaging

The red-shifted fluorescence exhibited by CH_3O-TO-CF_3 is of practical significance. In particular, whereas TO is efficiently excited by a 488 nm laser, CH_3O-TO-CF_3 can be excited by a 561 nm laser with greater than 60-fold higher extinction coefficient than TO (Table 3), more

than compensating for the lower quantum yields for CH_3O-TO-CF_3 (Table 2). Thus, fine-tuning of the TO absorption spectrum through the rational introduction of donor and acceptor groups allows different excitation sources to be used without significantly changing the dye structure, permitting promiscuous receptors such as the K7 protein to be used as a multi-color fluorescent tag differentiated by the fluorogenic dye used for imaging.

Table 3. Extinction coefficients (M^{-1} cm^{-1}) at common laser wavelengths of TO dyes bound to CT DNA and scFv protein K7

	CT DNA			scFv protein K7		
Dye	ε_{488}	ε_{532}	ε_{561}	ε_{488}	ε_{532}	ε_{561}
TO	30,000	11,000	160	44,000	3,800	120
CH_3O-TO	22,000	27,000	1,100	40,000	30,000	1,700
TO-CF_3	22,000	40,000	2,700	43,000	26,000	476
CH_3O-TO-CF_3	15,000	32,000	10,000	25,000	48,000	6,800

In order to visually demonstrate the spectral fine tuning potential of these dyes, we tested the two extrema, TO and CH_3O-TO-CF_3, as stains for mammalian cells. We imaged HeLa cells stained with 500 nM dye without washing at two different excitation wavelengths, 488 and 561 nm. As seen in Figure 8, the dyes stain both nuclei and cytoplasm, similar to other fluorogenic cyanines.[43,44] As expected, TO gives brighter staining upon excitation at 488 nm while CH_3O-TO-CF_3 is significantly brighter for 561 nm excitation

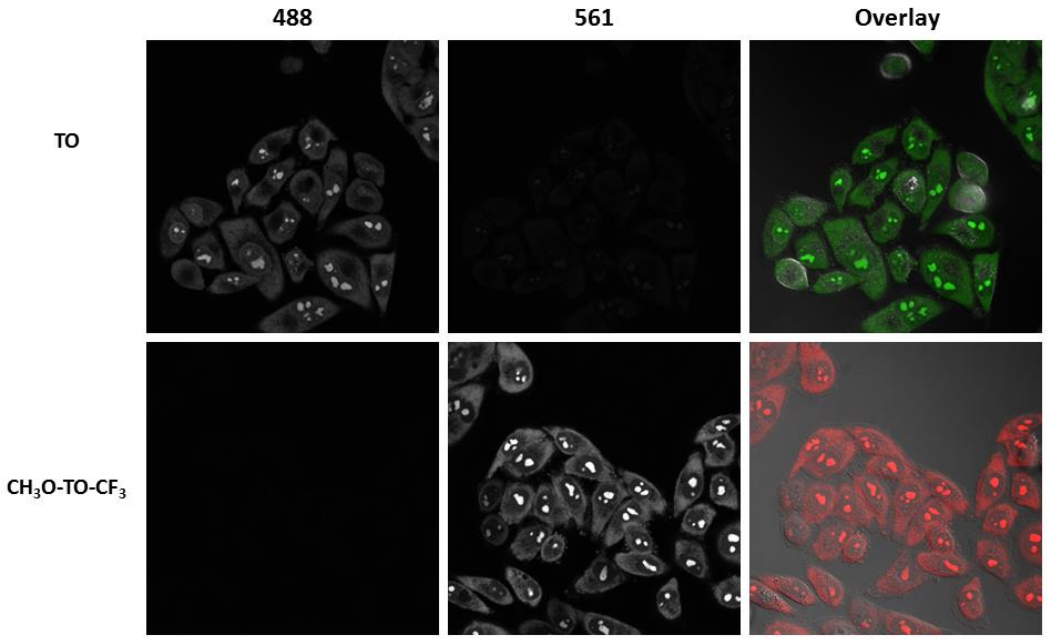


Figure 8. Fluorescence images of CH_3O-TO-CF_3 and TO with HeLa cells excited at 488nm (left) and 561nm (middle). The overlay image (right) shows the DIC with fluorescence in false color from excitation at both 488nm (Green), and 561nm (Red). Samples were not washed before imaging. [Dye] = 500 nM.

2.4 Discussion

Spectral tuning of the cyanine dye spectra can be accomplished through introduction of substituents on the aromatic heterocycles. A series of studies have investigated this effect for several families of cyanine dyes. Neutral unsymmetrical cyanine dyes like merocyanines rely heavily on electron donating and withdrawing components to give rise to dramatic spectral shifts upon changes in solvent polarity. In contrast, similar modifications of cationic unsymmetrical and symmetrical cyanine dyes gave rise to modest spectral shifts, up to ~18nm . To our knowledge the 24 nm red shift exhibited by CH_3O-TO-CF_3 surpasses reported shifts from other di-substituted cationic cyanine dyes, both symmetrical and asymmetrical.

Early studies focused on fluorogenic unsymmetrical cyanine dyes by investigating the electronic effects of substituents on the benzothiazole component of thiazole orange. The addition of an electron donating group, for example an ether, yielded a modest red shift of 15nm when compared to the parent dye which is consistent with our findings. Further studies showed the

family of TO dyes could exhibit a broad red shift up to 50nm with ssDNA and RNA, attributing the fluorescence due to electrostatic interactions. However, few studies have looked at the effect of a push pull system with an unsymmetrical cyanine dye and their interactions with various biomolecules; this study broadens the use asymmetric cyanine dyes.

Our study provides insight for further studies into finely tuning cyanine dyes through EWG/EDG substituents. This strategy can be applied to thiazole orange derivatives (e.g TO-3) as well as other asymmetrical dyes. For example oxazole thiazole blue or OTB is an interesting parent dye[29]. A protein-dye fluoromodule based on OTB was found to have an exceptional high quantum yield (~1.00). By introducing EWG and EDG at strategic locations we may be able to finely tune the spectral properties to expand the catalog of protein-dye pairs.

2.5 Conclusion

The results reported here show the synthesis and spectral characterization of a family of dyes with placement of electron donating and withdrawing groups in order to finely tune the absorbance and fluorescence spectra. It was also discovered that the dyes retained their advantageous properties of a substantial fluorescence enhancement upon noncovalent binding interactions with DNA and protein in addition to a red shift in absorbance and fluorescence spectra. The ability to precisely adjust the excitation and emission wavelengths without eliminating specific biomolecule binding properties can be used to optimize fluorescent labeling reagents for microscopy and cytometry applications.

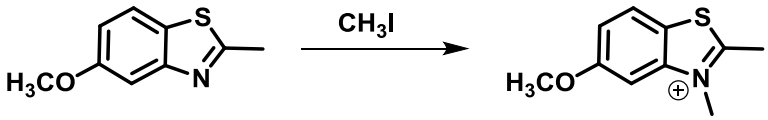
2.6 Experimental

2.6.1 General Experimental

All reagents were purchased from Sigma Aldrich or Alfa Aesar and purity was checked by NMR. Solvents were ACS grade. 4-chloro-1-methylquinolinium iodide (Q) was obtained from Dr. N. Shank. UV-Vis spectra were recorded on a CARY-300 Bio UV-visible spectrophotometer, Fluorescence spectra were recorded on a CARY-300 Bio fluorimeter, 1H and 13C NMR spectra were run on a Brucker Avance spectrometer at 500 and 75.47 MHz, respectively. Chemical shifts are reported as δ values (ppm) with reference to the residual solvent peaks. ESI-MS spectra were taken in a Finnigan ESI/APCI Ion Trap Mass Spectrometer on positive ion mode.

2.6.2 Synthesis of Thiazole Orange Derivatives

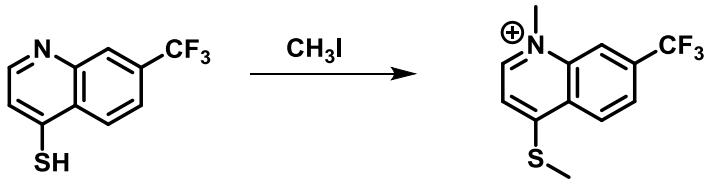
Synthesis of 5-methoxy-2,3-dimethylbenzothiazol-3-ium iodide (CH_3O-BT).



5-methoxy-2-methylbenzothiazole (2.00 g, 11.2 mmol) was dissolved in Iodomethane (7.043 mL) and microwave irradiated for 35 minutes at 100 °C. The precipitated product was collected through filtration and washed with cold ether and dried under vacuum. A chalky white solid was obtained (2.1347 g, 60%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.29 (d, J = 9.00 Hz,1H); 7.78 (d, J = 2.3 Hz, 1H); 7.43 (dd, J = 9.0, 2.3 Hz, 1H); 4.17 (s, 3H); 3.97 (s, 3H); 3.138 (s, 3H). ^{13}C NMR (75 MHz, DMSO) δ 177.58, 161.12, 143.64, 125.60, 120.88, 118.26, 100.47, 59.90, 36.63, 17.55. *ESI-MS* (positive*) m/z* calcd for $C_{10}H_{12}NOS^{+}$: 194.2; Found : 194.1.

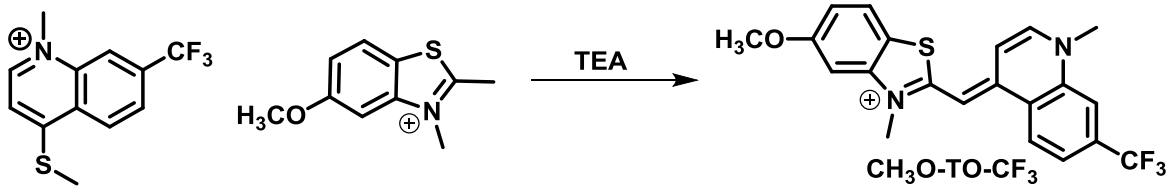
Synthesis of 1-methyl-4-(methylthio)-7-trifluoromethylquinolinium iodide ***(Q-CF$_3$).***



To a solution of 7-trifluoromethylquinolinethiol (1.00g, 4.328 mmol) dissolved in acetonitrile (5 mL), Iodomethane (326 µL, 0.0366 mg, 5.194 mmol) was added under argon. After 1 hr of stirring, the product precipitated as a dark red solid. The mixture was then stirred for an additional 3 days. The resulting reaction mixture was filtered and washed with acetonitrile. The mother liquid was recrystallized using acetonitrile and the crystals were filtered and washed with cold ether. A red solid was obtained (120 mg, 15 %)

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.29 (d, J = 6.6 Hz, 1H); 8.77(s, 1H); 8.67 (d, J = 8.8 Hz, 1H); 8.297 (dd, J = 8.8, 1.4 Hz, 1H); 8.03 (d, J = 6.6 Hz, 1H); 4.573 (s, 3H); 2.944 (s, 3H). ^{13}C NMR (75 MHz, DMSO) δ 164.3, 148.9, 136.6, 127.2, 125.5, 118.8, 117.7, 44.04, 15.12. *ESI-MS* (positive) *m/z* calcd for $C_{12}H_{11}F_3NS^+$: 258.1; Found: 258.2.

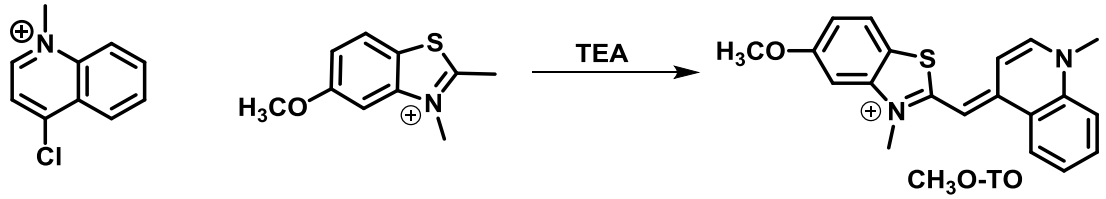
Synthesis of CH$_3$O-TO-CF$_3$:



CH$_3$O-BT (33.4 mg, 0.104 mmol) and **Q-CF$_3$** (40.0 mg, 0.104 mmol) were dissolved in methanol (2 mL). Triethylamine (73.3 µL, 52.61 mg, 0.52 mmol) was added drop wise over a minute and the product precipitated as a red solid. The reaction mixture was then stirred for 17 hours. To the reaction mixture ether (15 mL) was added drop wise and a red precipitate was filtered (20 mg, 20% yield).

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.03 (d, *J* = 8.9 Hz, 1H), 8.60 (d, *J* = 7.3 Hz, 1H), 8.31 (s, 1H), 8.01 (d, *J* = 8.7 Hz, 2H), 7.45 (d, *J* = 2.2 Hz, 1H), 7.34 (d, *J* = 7.2 Hz, 1H), 7.11 (dd, *J* = 8.8, 2.3 Hz, 1H), 7.03 (s, 1H), 4.21 (s, 3H), 4.11 (s, 3H), 3.95 (s, 3H). ^{13}C NMR (75 MHz, DMSO) δ 162.52, 160.79, 147.43, 146.14, 142.35, 138.45, 132.40, 127.79, 126.71, 124.11, 122.38, 116.28, 116.06, 113.64, 108.59, 98.98, 89.81, 56.57, 55.35, 42.73, 34.78. *ESI-MS* (positive*) m/z* calcd for $C_{21}H_{18}F_3N_2OS^+$: 403.1 *m/z* Found: 403.2.

Synthesis of CH_3O-TO



: **CH_3O-BT** (100.0mg, 0.328 mmol) and 4-chloro-1-methylquinolinium iodide, (105.0mg, 0.328 mmol) were dissolved in 3mL of methanol. Triethylamine (231.2µL, 1.64 mmol) was added dropwise over 2 minutes to the yellow reaction mixture. Upon the addition of TEA the reaction mixture turned dark red. The reaction was stirred for 23 hours at room temperature and 20 mL of ether was added to the reaction mixture and a red precipitate was filtered and washed with cold ether (104.6mg, 67%)

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.81 (1H,d, J=8.15); 8.59 (1H,d,J=7.2); 8.05 (2H,m); 7.93 (1H,d,J=8.7); 7.78 (1H,t,J=7.8); 7.38 (1H,d, J= 2.2); 7.33 (1H,d,J=7.15); 7.05 (1H,dd, J1=8.7,J2=2.2); 6.95 (1H,s); 4.17 (3H,s); 4.03 (3H,s); 3.92 (3H,s). ^{13}C NMR (75 MHz, DMSO) δ 161.4, 160.7, 148.6, 145.4, 142.4, 138.5, 133.6, 127.3, 125.9, 124.5, 123.9, 118.7, 115.6, 112.8, 108.2, 98.8, 88.6, 56.5, 42.8, 34.4 *ESI-MS* (positive) *m/z* calcd for $C_{20}H_{19}N_2OS^+$: 335.2, Found: 335.1.

2.6.3 Characterization in Various Solutions

Spectroscopic Studies

Dye solutions. Dye stock solutions (~1 mM) were prepared in DMSO and stored in the dark at 0 °C; concentrations were determined using extinction coefficients reported in the literature (TO:

ε_{502} = 63,000 $M^{-1}cm^{-1}$; TO-CF_3: ε_{516} = 65,000 $M^{-1}cm^{-1}$). Extinction coefficients for new compounds were determined in methanol CH_3O-TO: ε_{512} = 68,600 ± 1,100 $M^{-1}cm^{-1}$; CH_3O-TO-CF_3 : ε_{526} = 58,000 ± 2,600 $M^{-1}cm^{-1}$. Fluorescence emission spectra were collected on a Cary Eclipse Fluorescence spectrophotometer at 25° C. UV-Vis spectra were recorded on a Varian Cary 300 Bio UV-Vis spectrophotometer.

Dye Aggregation in Aqueous Solvents (Appendix)

Dye solutions were prepared in a 10% methanol in water solution. Dye titrations were performed from 0μM to 10μM for each dye. UV-Vis spectra were collected at room temperature.

Dye	$\lambda_{Monomer}$ (nm)	λ_{Dimer} (nm)
TO	501	478
CH_3O-TO	506	482
TO-CF_3	514	493
CH_3O-TO-CF_3	522	494

2.6.4 Fluorescence Quantum Yield

Fluorescence quantum yields for all dyes were determined relative to standards fluorescein and lucifer yellow, which were chosen due to their spectral overlap with TO analogs ($\Phi_{fluorescein}$ = 0.79 in 0.1 N NaOH[45]; $\Phi_{lucifer\ yellow}$ = 0.21 in nanopure H_2O at 22 °C[46]). Serial dilutions of the standards were made and the absorbance and fluorescence (ex 470) were measured. Fluorescence and absorbance measurements were performed with the dye in a solution of 90% glycerol in water as well as dye in the presence of excess soluble protein or calf thymus DNA. All standard and sample solutions had a maximum absorbance that did not exceed 0.1. All measurements were performed in triplicate to determine mean quantum yields and standard deviations.

Each fluorescence spectrum was integrated using Origin software. Then the integrated fluorescence intensity was plotted vs absorbance at 470 nm and a trendline calculated with y intercept = 0. The slope of this line is the gradient (*grad*). Using the gradient of the unknown (*grad x*) as well as the standard ($grad_{std}$) in addition to the refractive index of their respective solvents (*n*), the unknown quantum yield (Φ_x) was determined using the equation below.

$$\Phi_x = \Phi_{std} * \left(\frac{grad_x}{grad_{std}}\right) * \left(\frac{n_x}{n_{std}}\right)^2$$

2.6.5 Biomolecular recognition: DNA and ScFv

DNA solutions. DNA was purchased from GE-Biosciences, calf thymus DNA (ct-DNA) was used as received. Its concentration was determined in PBS buffer at 260 nm using the manufacturer extinction coefficient (ε_{260}= 13,200 $M^{-1}cm^{-1}$) per base pair.

Protein Solutions. Soluble scFv K7 was prepared as described previously.[26] Protein concentration was determined using a Nanodrop spectrometer (ε_{280} = 36,040 $M^{-1}cm^{-1}$).

2.6.6 Determination of Kd Values

Range Finding Assay. A serial dilution of 10:1 (dye:scFv) as well as a control serial dilution of just dye were made, starting at 10 μM dye to 2 nM. The resulting samples were allowed to equilibrate in eppendorf tubes at room temperature overnight. Then the background signal was corrected by subtracting the fluorescence intensity of just dye from the fluorescence intensity of the complex. The corrected fluorescence was multiplied by the dilution factor and then normalized. The expected results should show a definite plateau at 1, this indicates that the system is above the K_D and the assumption that a 1:1 binding relationship is true. However the point in which the normalized fluorescence decreases indicates that the system is nearing the

K_D and the complex starts to dissociate. We use this curve to estimate the concentration in which half of the dye or protein is bound.

K_D Determination for Dye and Soluble FAP. For each dye, a parallel dilution series was performed using 16 different concentrations of dye that was centered around the estimated K_D (determined by range finding assay). From that dilution series two different samples were generated: a control dye in wash buffer, as well as dye:K7 in wash buffer. The concentration of K7 was kept constant for each dye sample and the experiment was performed in triplicate. The resulting data were fit using the ligand depletion model (Saturation binding equation) in Graph Pad.

2.6.7 Cell Imaging

HeLa cells were plated on MatTec dishes were incubated with a 1:1000 dilution of either TO or CH_3O-TO-CF_3 for 30 minutes at 37°C in a humid tissue culture incubator at 37°C in the presence of 9% CO_2. The dyes were washed off using 1 mL of complete DMEM. Following incubation with the dyes, the cells were imaged live using a Nikon A1 confocal microscope. Dyes were excited with a 488 nm and 561 nm laser and the emissions were collected with a band pass filter between the wavelengths of 505-550 nm and 603-678 nm, respectively. Imaging parameters were kept identical for each dish of cells. Images were processed using ImageJ software where a false color which represented the emission window was applied.

2.7 Acknowledgment

We are grateful to Prof. Simon Watkins and Morgan Jessup of the Center for Biologic Imaging at the University of Pittsburgh for cell imaging.

2.8 Appendix

2.8.1 Fluorescence Quantum Yield Data

Quantum Yields were determined as described in the methods section below is the raw data. All samples were excited at 470nm.

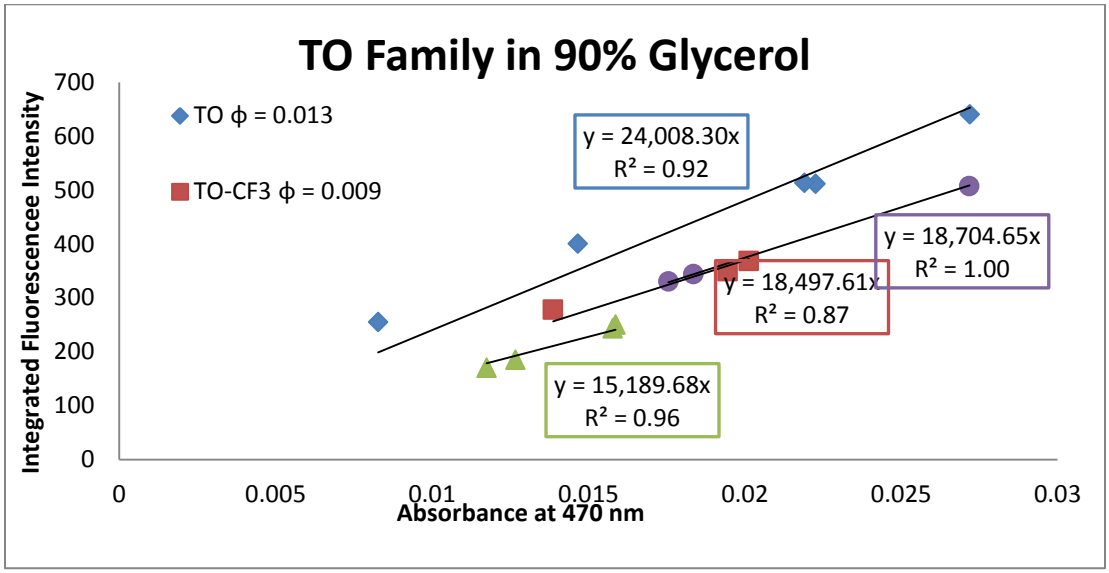


Figure S1. Linear plots for TO Dye family in 90% glycerol. The gradient for each sample was used to calculate the quantum yield using equation [eq.1]

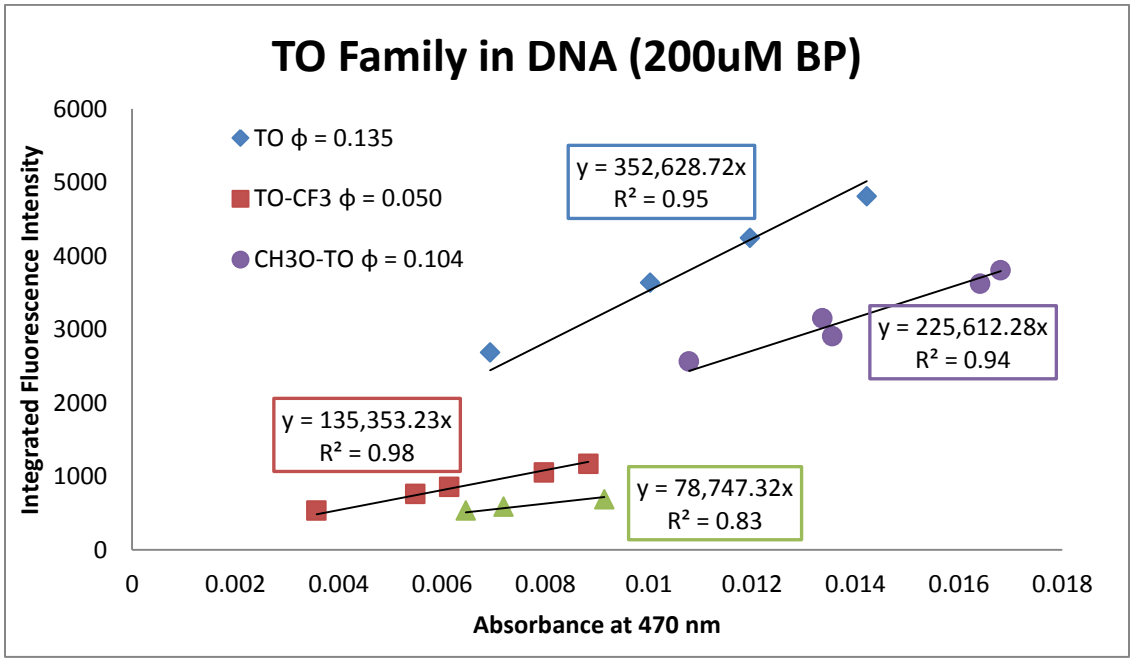


Figure S2. Linear plots for TO Dye family in CT DNA (200μM Base Pairs).

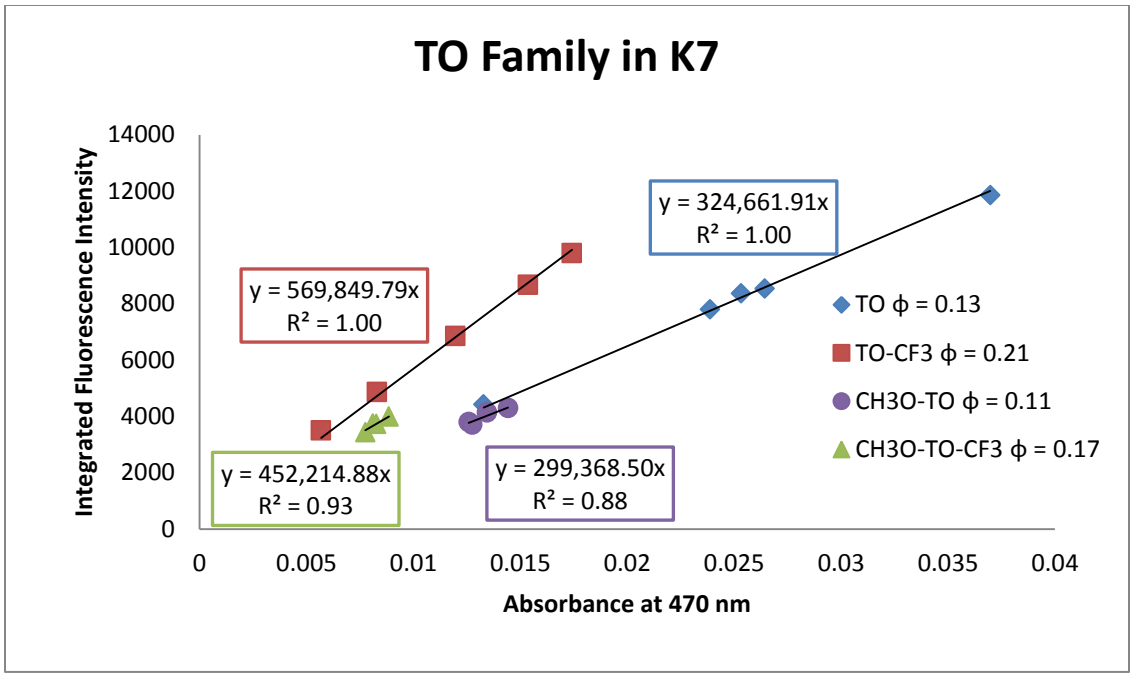


Figure S3. Linear plots for TO Dye family in excess soluble protein K7. Concentrations of soluble protein K7 (TO, TO-CF_3)=2μM; CH_3O- TO = 3μM; CH_3O- TO- CF_3 = 3.5μM

2.3.4. Fluorogen binding affinity to soluble FAPs Assay

In order to determine the fluorogen binding affinity to the soluble protein, an initial range finding was performed to determine an estimate concentration in which half of the dye or protein is bound. This concentration was used in the subsequent K_D assay. Method is described in Section 2.6.6.

Range Finding assay

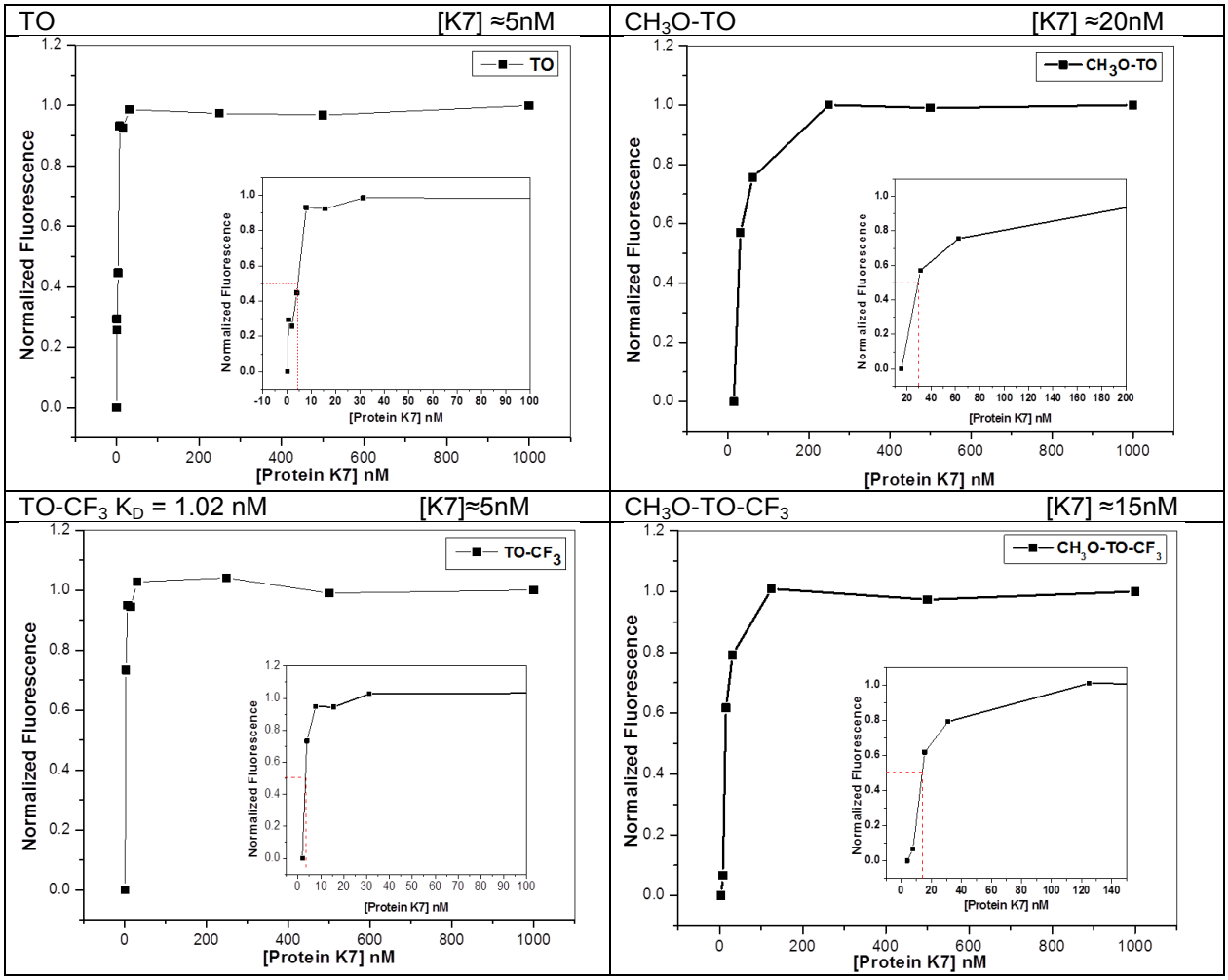


Figure S4. Range Finding Data for TO Dye family and soluble protein K7. Inset shows estimate protein concentrations when approximately 50% is bound. Samples were excited at 490nm and emission was monitored at the lambda max for each dye. (TO= 520nm; CH_3O-TO = 536 nm; TO-CF_3 = 536; CH_3O-TO- CF_3 = 552nm)

Kd Determination for Dye and Soluble FAP

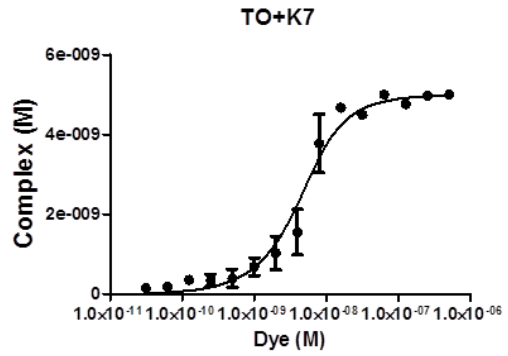
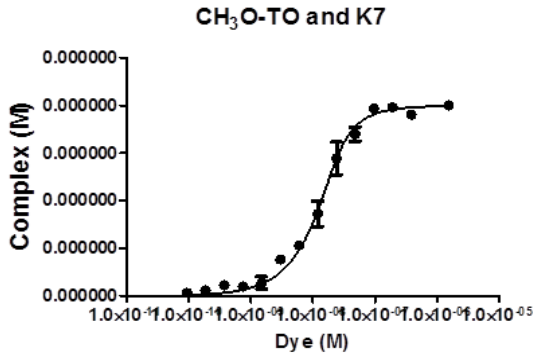
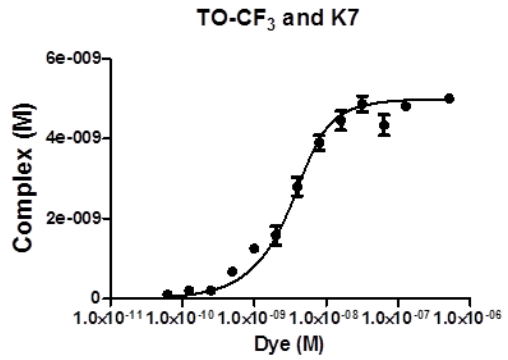
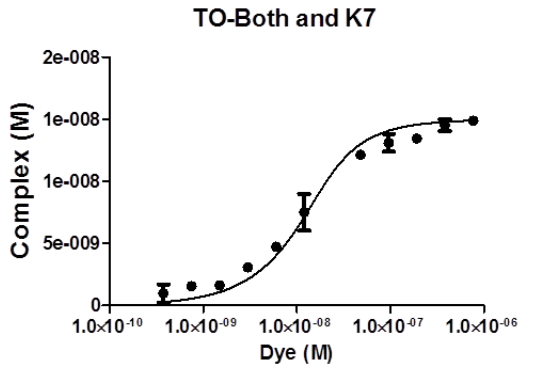
TO K_D = 2.14 nM [K7]=5nM	CH_3O-TO K_D = 3.78 nM [K7]=20nM
	
TO-CF_3 K_D = 1.02 nM [K7]=5nM	CH_3O-TO-CF_3 K_D = 5.4 nM [K7]=15nM
	

Figure S5. Fluorescence titrations of TO dye into soluble K7 (protein concentrations were determined by range finding assay see **Figure S5**). Samples were excited at 490nm and emission was monitored at lambda max for each dye. (TO= 520nm; CH_3O-TO = 536 nm; TO-CF_3 = 536; CH_3O-TO- CF_3 = 552nm)

4.NMR Spectra:

4.1 CH_3O-TO-CF_3

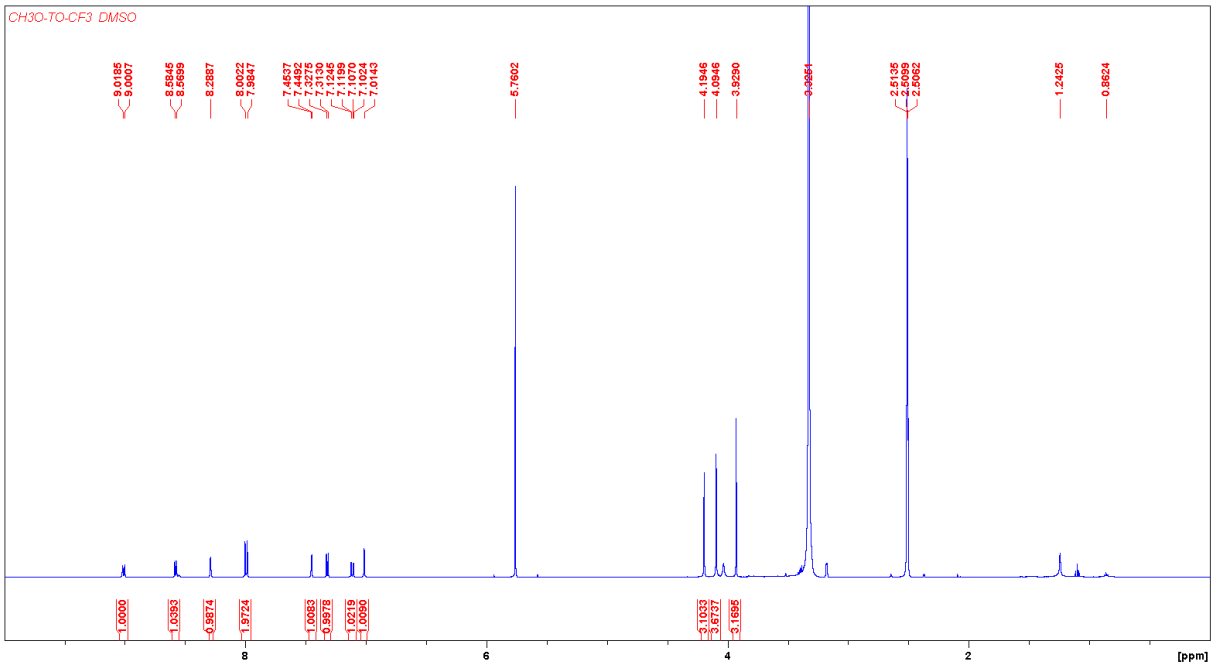


Figure S6. 1H NMR spectrum of dye **CH_3O-TO-CF_3** (500MHz, $(CD_3)_2SO$)

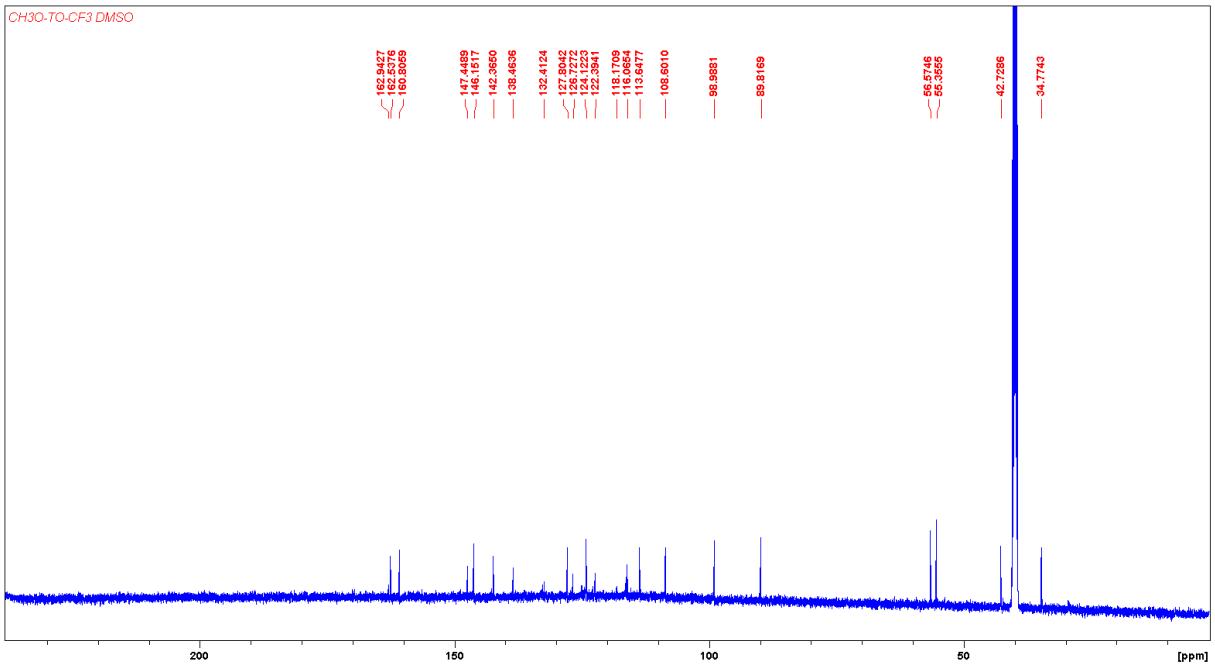


Figure S7. 13C NMR spectrum of dye **CH_3O-TO-CF_3** (500MHz, $(CD_3)_2SO$)

4.2 CH_3O-TO

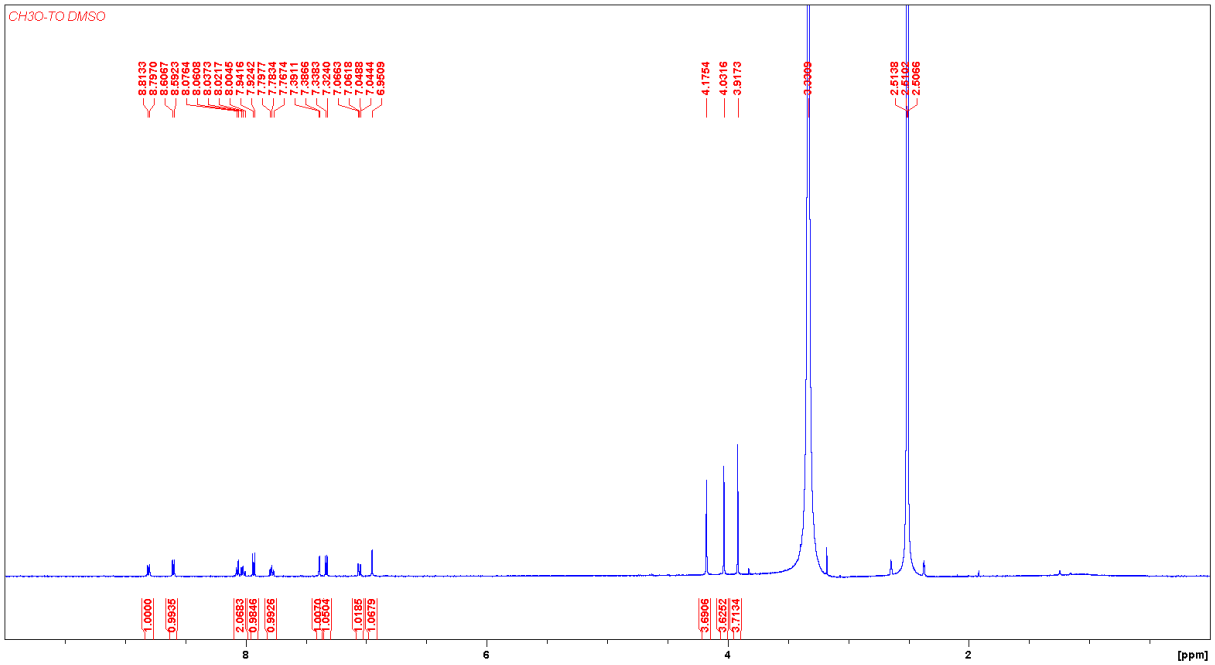


Figure S8. 1H NMR spectrum of dye **CH_3O-TO** (500MHz, $(CD_3)_2SO$)

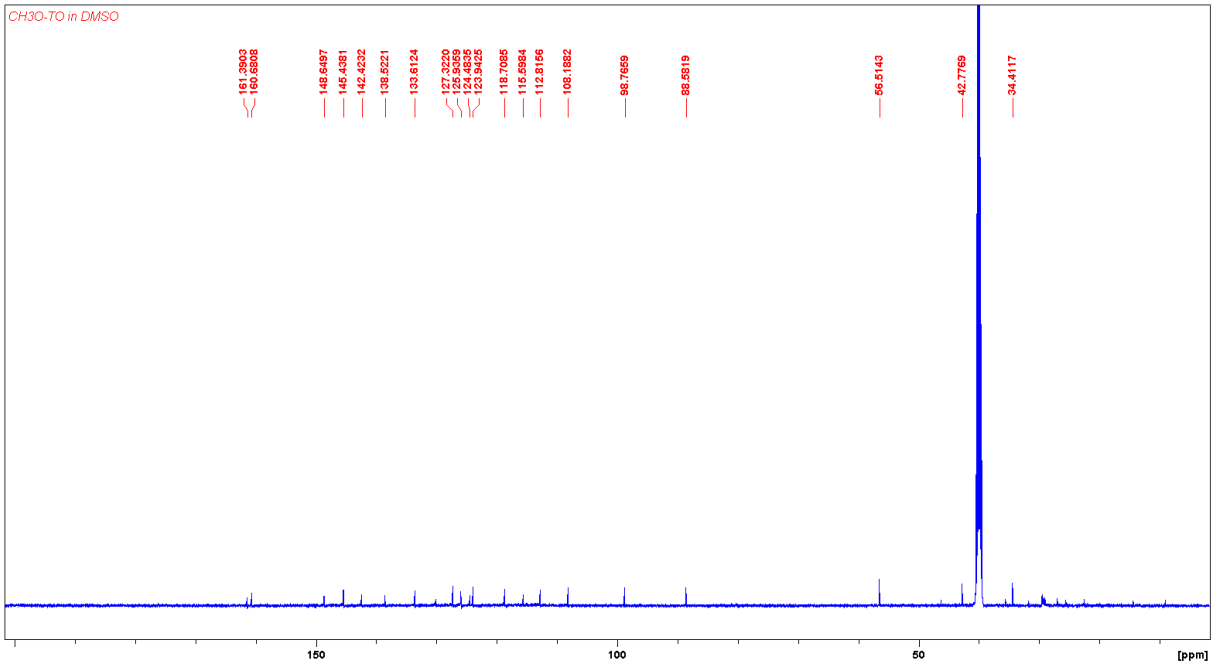


Figure S9. 13C NMR spectrum of dye **CH_3O-TO** (75MHz, $(CD_3)_2SO$)

4.3 Q-CF_3

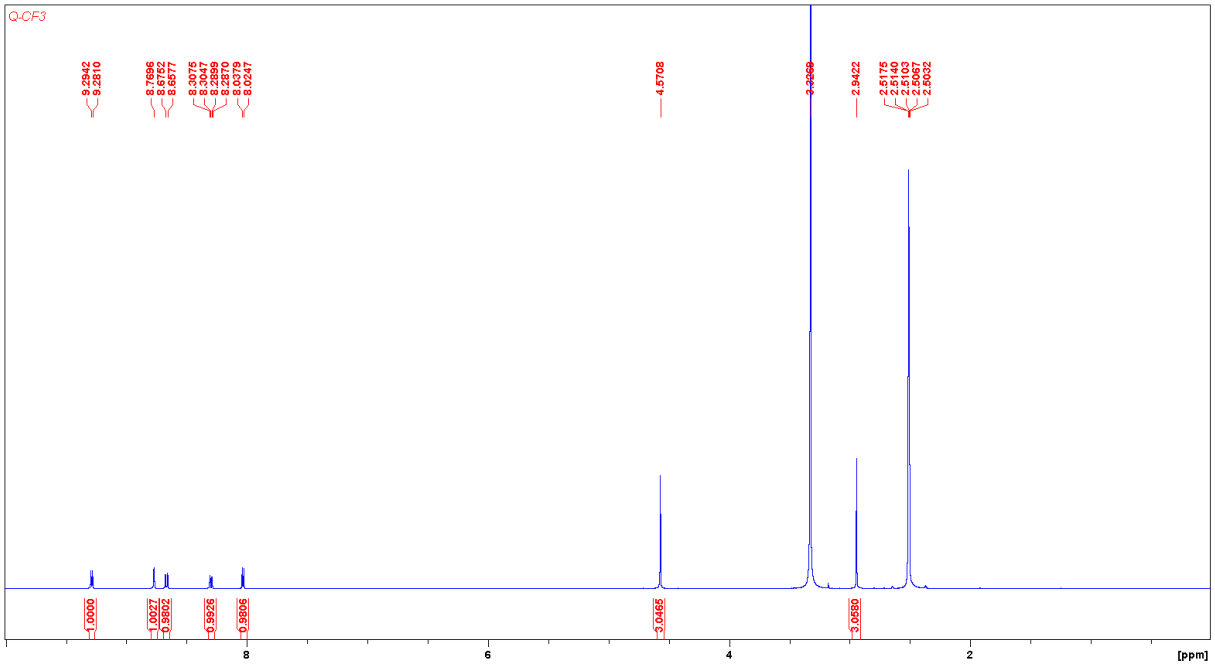


Figure S10. 1H NMR spectrum of hemi-dye ***Q-CF_3*** (500MHz, $(CD_3)_2SO$)

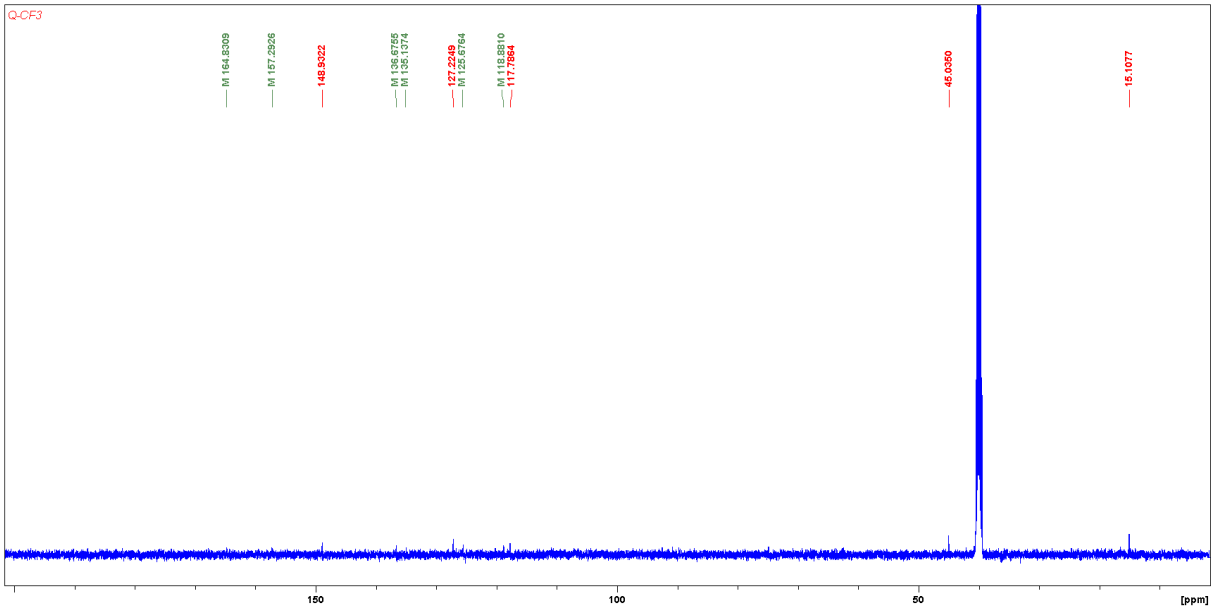


Figure S11. 13C NMR spectrum of hemi-dye ***Q-CF_3*** (75MHz, $(CD_3)_2SO$)

4.4 CH_3O-BT

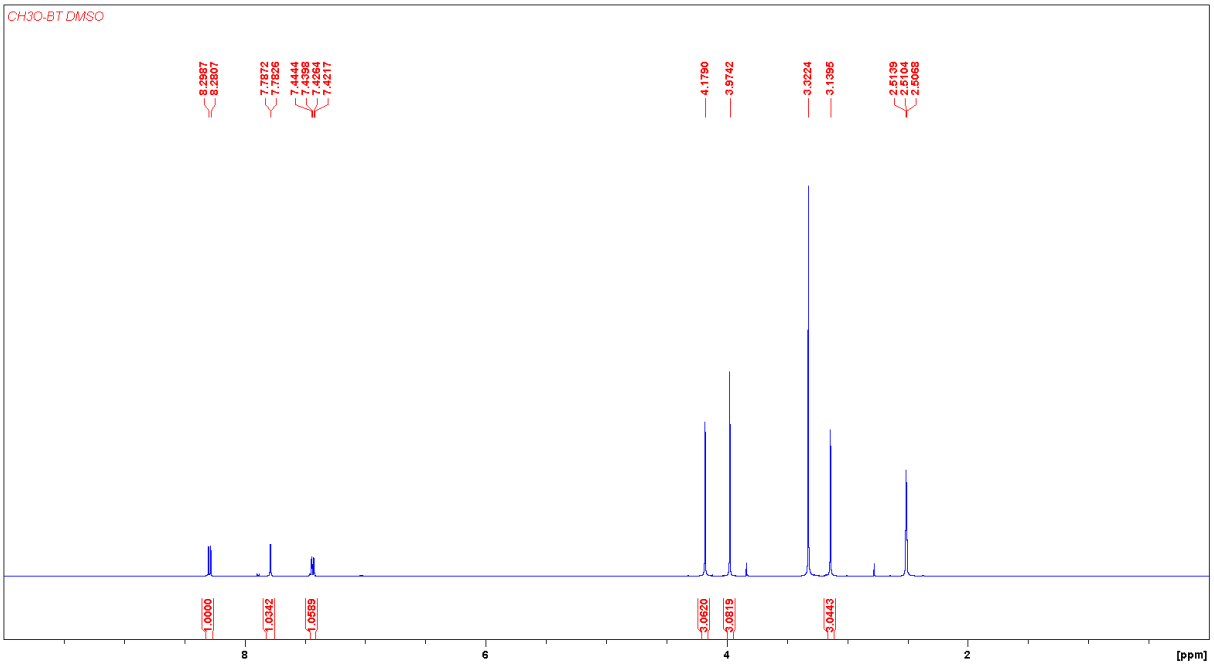


Figure S12. 1H NMR spectrum of hemi-dye ***CH_3O-BT*** ($(CD_3)_2SO$)

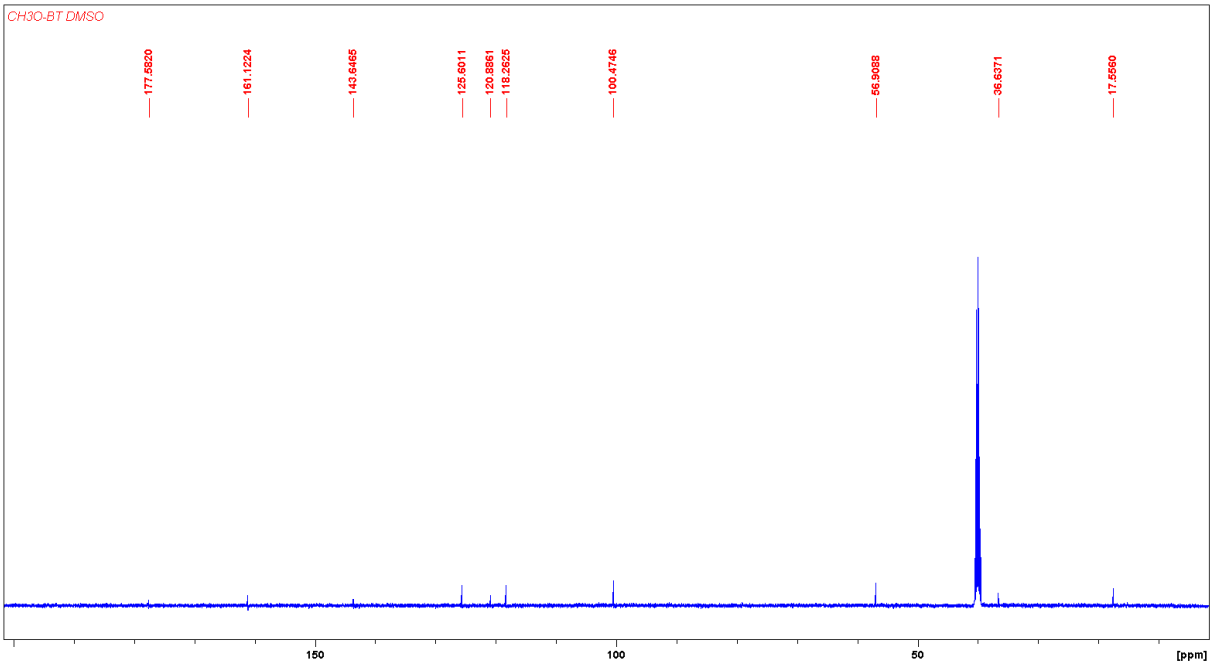


Figure S13. 13C NMR spectrum of hemi-dye ***CH_3O-BT*** ($(CD_3)_2SO$)

5 LCQ- ESI –MS Spectra

5.1 CH_3O-BT

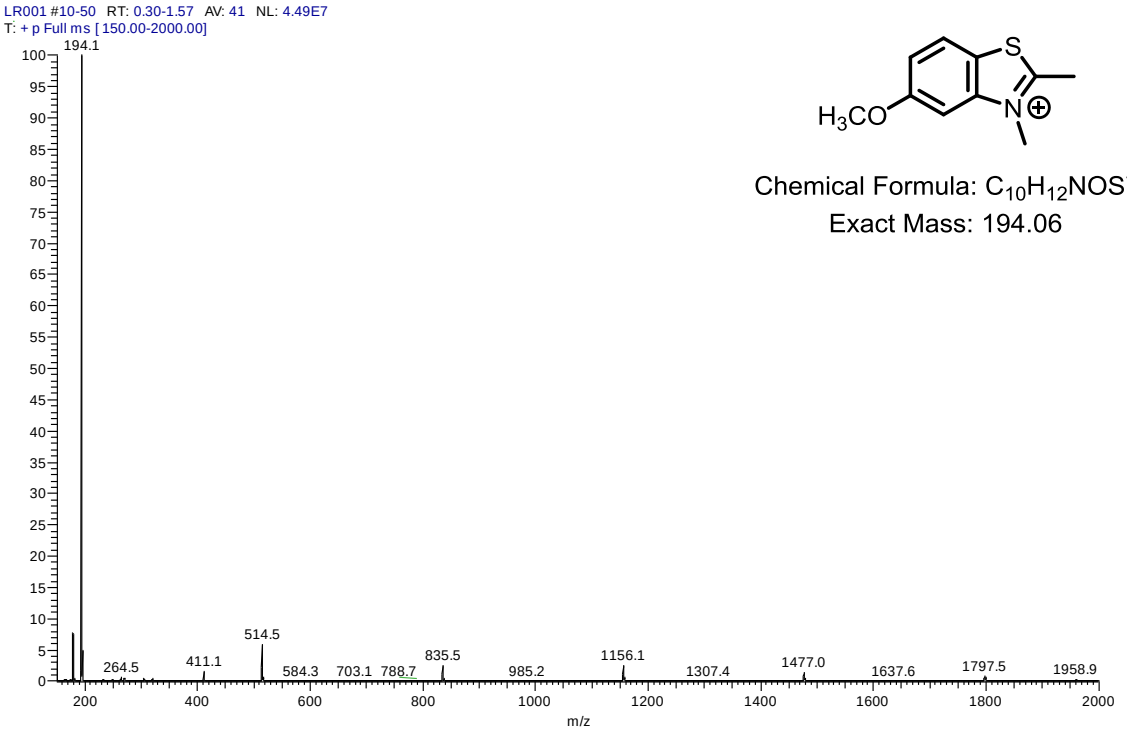


Figure S14. ESI- MS of hemi-dye **CH_3O-BT**

5.2 *Q-CF_3*

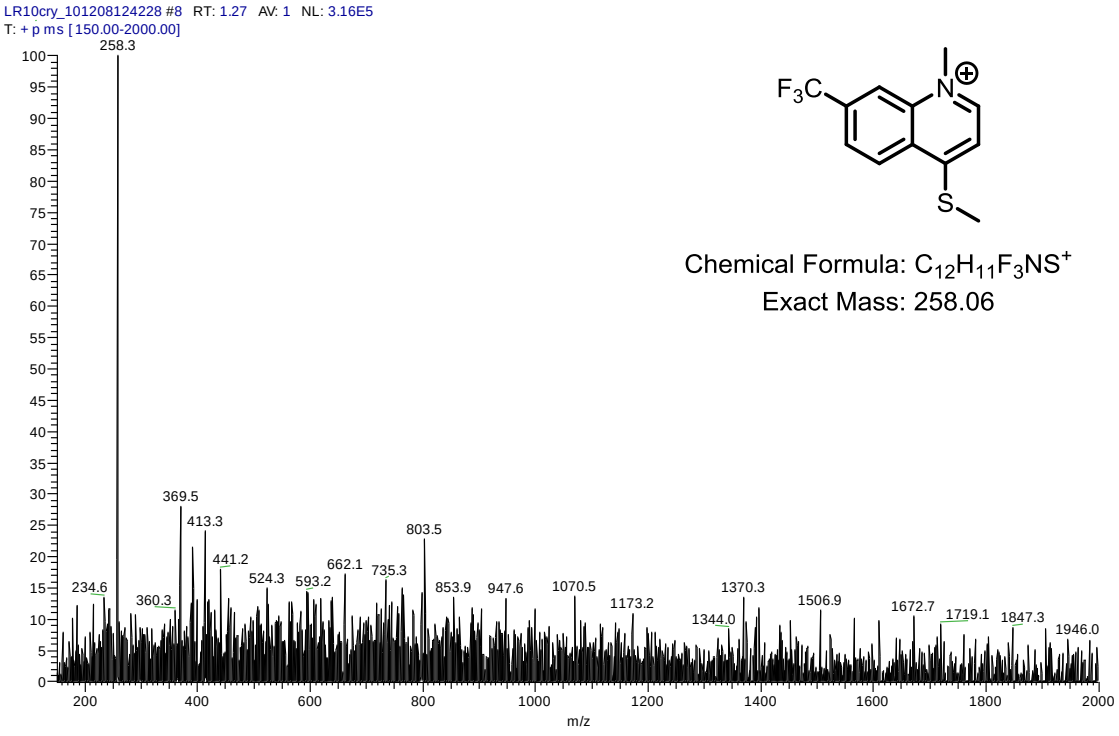


Figure S15. ESI- MS of hemi-dye ***Q-CF_3***

5.3 CH_3O-TO-CF_3

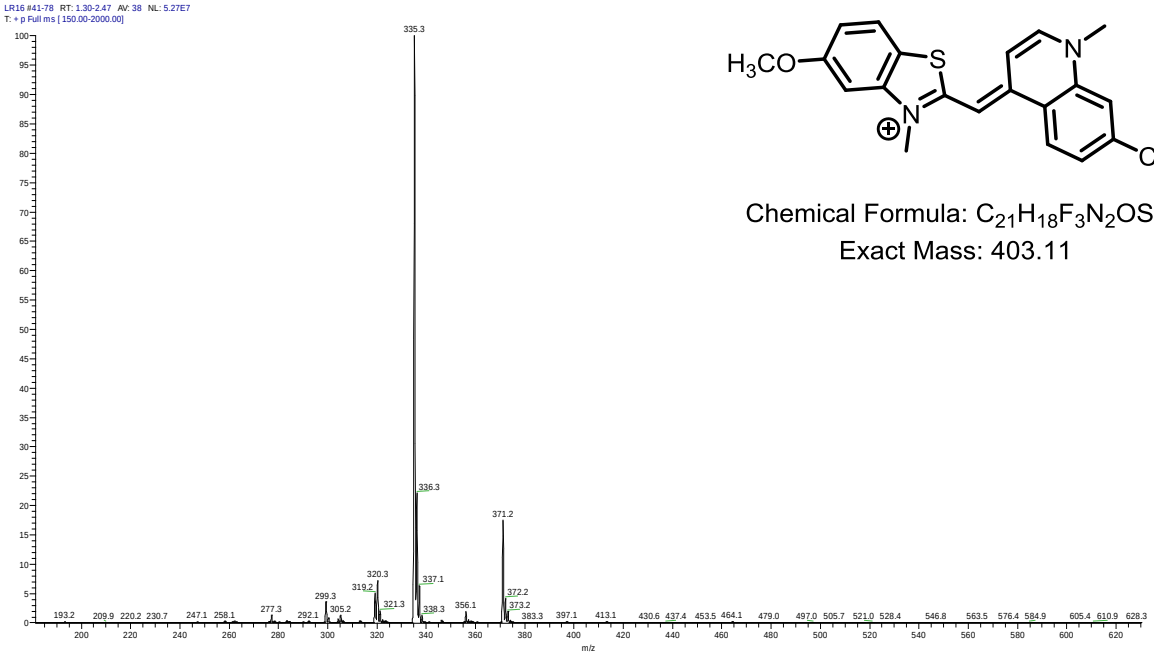


Figure S16. ESI- MS of dye **CH_3O-TO-CF_3**

5.4 CH_3O-TO

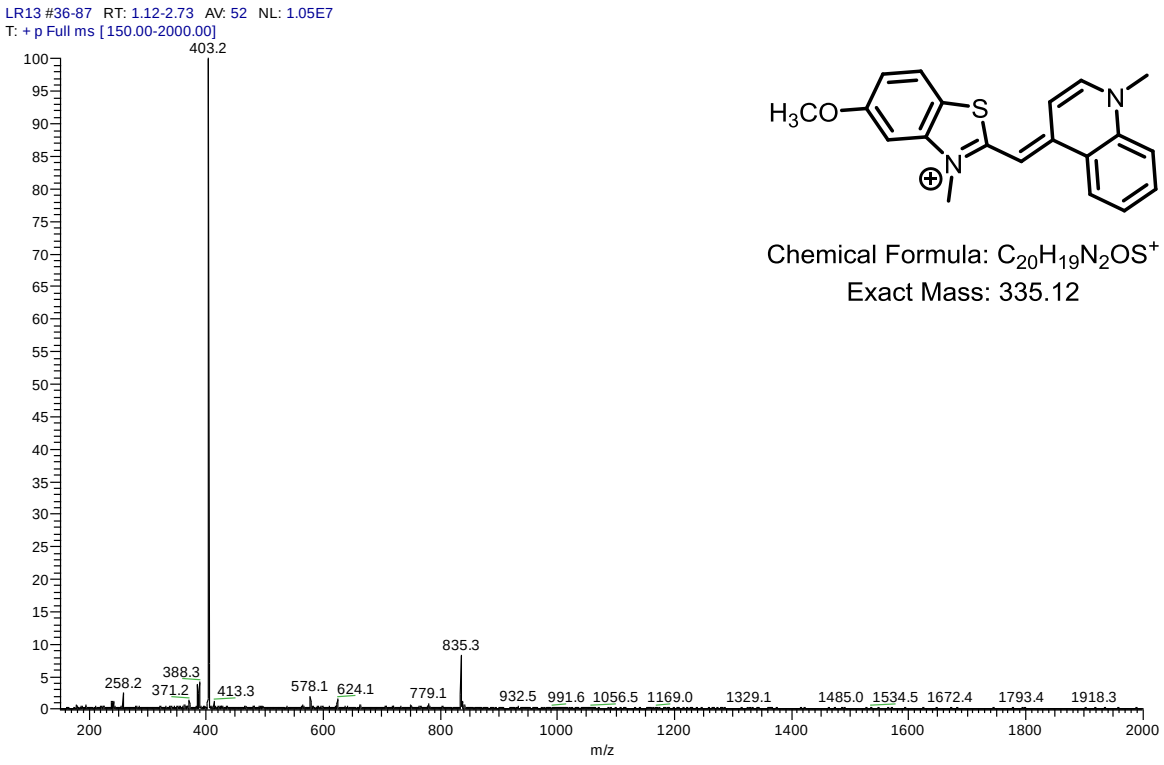


Figure S17. ESI- MS of dye **CH_3O-TO**

References

(1) Lavis, L. D.; Raines, R. T. Bright Ideas for Chemical Biology. *ACS Chem. Biol.* **2008**, *3*, 142-155.

(2) Nadler, A.; Schultz, C. The power of fluorogenic probes. *Angew. Chem., Int. Ed.* **2013**, *52*, 2408-2410.

(3) Pawar, M. G.; Srivatsan, S. G. An Environment-Responsive Fluorescent Nucleoside Analog Probe for Studying Oligonucleotide Dynamics in a Model Cell-Like Compartment. *J. Phys. Chem. B* **2013**, *117*, 14273-14282.

(4) Suh, E. H.; Liu, Y.; Connelly, S.; Genereux, J. C.; Wilson, I. A.; Kelly, J. W. Stilbene Vinyl Sulfonamides as Fluorogenic Sensors of and Traceless Covalent Kinetic Stabilizers of Transthyretin that Prevent Amyloidogenesis. *J. Am. Chem. Soc.* **2013**, *135*, 17869-17880.

(5) Paige, J. S.; Wu, K. Y.; Jaffrey, S. R. RNA Mimics of Green Fluorescent Protein. *Science* **2011**, *333*, 642-646.

(6) Debler, E. W.; Kaufmann, G. F.; Meijler, M. M.; Heine, A.; Mee, J. M.; Pljevaljcic, G.; Di Bilio, A. J.; Schultz, P. G.; P., M. D.; Janda, K. D.; Wilson, I. A.; Gray, H. B.; Lerner, R. A. Deeply Inverted Electron-Hole Recombination in a Luminescent Antibody-Stilbene Complex. *Science* **2008**, *319*, 1232-1235.

(7) Li, X.; Gao, X.; Shi, W.; Ma, H. Design strategies for water-soluble small molecular chromogenic and fluorogenic probes. *Chem. Rev.* **2014**, *114*, 590-659.

(8) Cao, J.; Wu, T.; Hu, C.; Liu, T.; Sun, W.; Fan, J.; Peng, X. The nature of the different environmental sensitivity of symmetrical and unsymmetrical cyanine dyes: an experimental and theoretical study. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13702-13708.

(9) Biancardi, A.; Biver, T.; Marini, A.; Mennucci, B.; Secco, F. Thiazole orange (TO) as a light-switch probe: a combined quantum-mechanical and spectroscopic study. *Phys. Chem. Chem. Phys.* **2011**, *13*, 12595-12602.

(10) Silva, G. L.; Ediz, V.; Yaron, D.; Armitage, B. A. Experimental and Computational Investigation of Unsymmetrical Cyanine Dyes: Understanding Torsionally Responsive Fluorogenic Dyes. *J. Am. Chem. Soc.* **2007**, *129*, 5710-5718.

(11) Lee, L. G.; Chen, C.; Chiu, L. Thiazole Orange: A New Dye for Reticulocyte Analysis *Cytometry* **1986**, *7*, 508-517.

(12) Benson, S. C.; Mathies, R. A.; Glazer, A. N. Heterodimeric DNA-binding dyes designed for energy transfer: stability and applications of the DNA complexes. *Nucleic Acids Res.* **1993**, *21*, 5720-5726.

(13) Martin, R. M.; Leonhardt, H.; Cardoso, M. C. DNA labeling in living cells. *Cytometry* **2005**, *67*, 45-52.

(14) Nygren, J.; Svanvik, N.; Kubista, M. The Interactions Between the Fluorescent Dye Thiazole Orange and DNA. *Biopolymers* **1998**, *46*, 39–51.

(15) Svanvik, N.; Westman, G.; Wang, D.; Kubista, M. Light-up probes: thiazole orange-conjugated peptide nucleic acid for detection of target nucleic acid in homogeneous solution. *Anal. Biochem.* **2000**, *281*, 26-35.

(16) Svanvik, N.; Nygren, J.; Westman, G.; Kubista, M. Free-Probe Fluorescence of Light-up Probes. *J. Am. Chem. Soc.* **2001**, *123*, 803-809.

(17) Carreon, J. R.; Mahon, K. P., Jr.; Kelley, S. O. Thiazole orange-peptide conjugates: sensitivity of DNA binding to chemical structure. *Org. Lett.* **2004**, *6*, 517-519.

(18) Kohler, O.; Jarikote, D. V.; Seitz, O. Forced intercalation probes (FIT Probes): thiazole orange as a fluorescent base in peptide nucleic acids for homogeneous single-nucleotide-polymorphism detection. *Chembiochem* **2005**, *6*, 69-77.

(19) Armitage, B. A. Cyanine Dye–DNA Interactions: Intercalation, Groove Binding,and Aggregation. *Top. Curr. Chem.* **2005**, *253*, 55-76.

(20) Constantin, T. P.; Silva, G. L.; Robertson, K. L.; Hamilton, T. P.; Fague, K.; Waggoner, A. S.; Armitage, B. A. Synthesis of New Fluorogenic Cyanine Dyes and Incorporation into RNA Fluoromodules. *Org. Lett.* **2008**, *10*, 1561-1564.

(21) Ditmangklo, B.; Boonlua, C.; Suparpprom, C.; Vilaivan, T. Reductive alkylation and sequential reductive alkylation-click chemistry for on-solid-support modification of pyrrolidinyl peptide nucleic acid. *Bioconjugate Chem.* **2013**, *24*, 614-625.

(22) Hovelmann, F.; Gaspar, I.; Ephrussi, A.; Seitz, O. Brightness enhanced DNA FIT-probes for wash-free RNA imaging in tissue. *J. Am. Chem. Soc.* **2013**, *135*, 19025-19032.

(23) Xu, B.; Wu, X.; Yeow, E. K.; Shao, F. A single thiazole orange molecule forms an exciplex in a DNA i-motif. *Chem. Commun.* **2014**, *50*, 6402-6405.

(24) Hovelmann, F.; Gaspar, I.; Loibl, S.; Ermilov, E. A.; Roder, B.; Wengel, J.; Ephrussi, A.; Seitz, O. Brightness through local constraint--LNA-enhanced FIT hybridization probes for in vivo ribonucleotide particle tracking. *Angew. Chem., Int. Ed.* **2014**, *53*, 11370-11375.

(25) Szent-Gyorgyi, C.; Schmidt, B. F.; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. A.; Woolford, C. A.; Yan, Q.; Vasilev, K. V.; Berget, P. B.; Bruchez, M. P.; Jarvik, J. W.; Waggoner, A. Fluorogen-activating single-chain antibodies for imaging cell surface proteins. *Nat. Biotechnol.* **2008**, *26*, 235-240.

(26) Özhalici-Ünal, H.; Lee Pow, C.; Marks, S. A.; Jesper, L. D.; Silva, G. L.; Shank, N. I.; Jones, E. W.; Burnette, J. M., III; Berget, P. B.; Armitage, B. A. A Rainbow of Fluoromodules: A Promiscuous scFv Protein Binds to and Activates a Diverse Set of Fluorogenic Cyanine Dyes. *J. Am. Chem. Soc.* **2008**, *130*, 12620-12621.

(27) Shank, N. I.; Zanotti, K. J.; Lanni, F.; Berget, P. B.; Armitage, B. A. Enhanced Photostability of Genetically Encodable Fluoromodules Based on Fluorogenic Cyanine Dyes and a Promiscuous Protein Partner. *J. Am. Chem. Soc.* **2009**, *131*, 12961-12969.

(28) Szent-Gyorgyi, C.; Schmidt, B. F.; Fitzpatrick, A. J.; Bruchez, M. P. Fluorogenic Dendrons with Multiple Donor Chromophores as Bright Genetically Targeted and Activated Probes. *J. Am. Chem. Soc.* **2010**, *132*, 11103–11109.

(29) Zanotti, K. J.; Silva, G. L.; Creeger, Y.; Robertson, K. L.; Waggoner, A. S.; Berget, P. B.; Armitage, B. A. Blue fluorescent dye-protein complexes based on fluorogenic cyanine dyes and single chain antibody fragments. *Org. Biomol. Chem.* **2011**, *9*, 1012-1020.

(30) Shank, N. I.; Pham, H. H.; Waggoner, A. S.; Armitage, B. A. Twisted cyanines: a non-planar fluorogenic dye with superior photostability and its use in a protein-based fluoromodule. *J. Am. Chem. Soc.* **2013**, *135*, 242-251.

(31) Yan, Q.; Schmidt, B. F.; Perkins, L. A.; Naganbabu, M.; Saurabh, S.; Andrekob, S. K.; Bruchez, M. P. Near-instant surface-selective fluorogenic protein quantification using sulfonated triarylmethane dyes and fluorogen activating proteins. *Org. Biomol. Chem.* **2015**, DOI: 10.1039/C1034OB02309A.

(32) Wang, Y.; Telmer, C. A.; Schmidt, B. F.; Franke, J. D.; Ort, S.; Arndt-Jovin, D. J.; Bruchez, M. P. Fluorogen Activating Protein−Affibody Probes: Modular, No-Wash Measurement of Epidermal Growth Factor Receptors. *Bioconjugate Chem.* **2014**, DOI: 10.1021/bc500525b.

(33) Schwartz, S. L.; Yan, Q.; Telmer, C. A.; Lidke, K. A.; Bruchez, M. P.; Lidke, D. S. Fluorogen-Activating Proteins Provide Tunable Labeling Densities for Tracking FcepsilonRI Independent of IgE. *ACS Chem. Biol.* **2014**, DOI: 10.1021/cb5005146.

(34) Saunders, M. J.; Block, E.; Sorkin, A.; Waggoner, A. S.; Bruchez, M. P. A bifunctional converter: fluorescein quenching scFv/fluorogen activating protein for photostability and improved signal to noise in fluorescence experiments. *Bioconjugate Chem.* **2014**, *25*, 1556-1564.

(35) Mishra, A.; Behera, R. K.; Behera, P. K.; Mishra, B. K.; Behera, G. B. Cyanines during the 1990s: A Review. *Chem. Rev.* **2000**, *100*, 1973−2011.

(36) Haugland, R. P.: The Molecular Probes Handbook—A Guide to Fluorescent Probes and Labeling Technologies. 11 ed. ed.; Molecular Probes: Eugene, OR, 2010.

(37) Cao, J.; Hu, C.; Sun, W.; Xu, Q.; Fan, J.; Song, F.; Sun, S.; Peng, X. The mechanism of different sensitivity of meso-substituted and unsubstituted cyanine dyes in rotation-restricted environments for biomedical imaging applications. *RSC Adv.* **2014**, *4*, 13385–13394.

(38) Toutchkine, A.; Nguyen, D. V.; Hahn, K. M. Merocyanine Dyes with Improved Photostability. *Org. Lett.* **2007**, *9*, 2775-2777.

(39) Winstead, A. J.; Fleming, N.; Hart, K.; Toney, D. Microwave Synthesis of Quaternary Ammonium Salts. *Molecules* **2008**, *13*, 2107-2113.

(40) Timcheva, I. I.; Maximova, V. A.; Deligeorgiev, T. G.; Gadjev, N. I.; Sabnis, R. W.; Ivanov, I. G. Fluorescence Spectral Characteristics of Novel Asymmetric Monomethine Cyanine Dyes in Nucleic Acid Solutions. *FEBS Lett.* **1997**, *405*, 141-144.

(41) West, W.; Pearce, S. The dimeric state of cyanine dyes. *The Journal of Physical Chemistry* **1965**.

(42) Stadler, A. L.; Renikuntla, B. R.; Yaron, D.; Fang, A. S.; Armitage, B. A. Substituent effects on the assembly of helical cyanine dye aggregates in the minor groove of a DNA template. *Langmuir : the ACS journal of surfaces and colloids* **2011**, *27*, 1472-1479.

(43) Zhang, S.; Fan, J.; Li, Z.; Hao, N.; Cao, J.; Wu, T.; Wang, J.; Peng, X. A Bright Red Fluorescent Cyanine Dye for Live-Cell Nucleic Acid Imaging, with High Photostability and a Large Stokes Shift. *J. Mat. Chem. B* **2014**, *2*, 2688-2693.

(44) Horobin, R. Where Do Dyes Go Inside Living Cells? Predicting Uptake, Intracellular Localisation,

and Accumulation Using QSAR Models. *Color.Technol.*, *130*, 155-173.

(45) Brouwer, A. M. Standards for photoluminescence quantum yield measurements in solution (IUPAC Technical Report). *Pure App. Chem.* **2011**, *83*, 2213–2228.

(46) Stewart, W. W. Synthesis of 3,6-Disulfonated 4-Aminonaphthalimides. *J. Am. Chem. Soc.* **1981**, *103*, 7615-7620.

Chapter 3: Labeling Nucleic Acids with Dye-Labeled PNA Probes

3. Chapter Summary

In this chapter we study a modified binary probe used to label DNA telomeric repeats in fixed cells and tissue. By utilizing γPNAs we were able to shorten and simplify the binary probe design with two terminally-labeled γPNA 9-mer probes designed to hybridize in tandem allowing for a FRET signal to indicate hybridization to the target. First, the γPNA probes were analyzed for affinity, structure, and fluorescence properties. Next, the γPNA probes were applied for the analysis of telomeric DNA in genomic DNA, as well as cellular imaging using fluorescence in-situ hybridization (FISH).

3.1. Introduction

Fluorescently-labeled probes that target biomolecules are a crucial tool for biomedical research, allowing for sensitive and diverse information-rich analysis of complex biological samples. To label biomolecules, fluorescent dyes can be covalently attached to a targeting moiety such as a small molecule, peptide, or nucleic acid. In order to target a nucleic acid, a complementary nucleic acid or synthetic analog is often end-labeled with a fluorescent dye to allow detection upon hybridization (Figure 1A)[1-3]. These fluorescently-labeled oligonucleotides have been used in many biological applications[4] including: the detection of DNA mutations, ligation and cleavage[5,6], gene expression[7], monitoring RNA splicing[8], and fluorescence in-situ hybridization (FISH)[1,3,9].

A widely-used fluorescent phenomenon exploited to increase sensitivity in assays for biomolecular research and probe technologies is Förster resonance energy transfer or FRET. FRET is a phenomenon where a donor molecule in an excited state, instead of releasing energy via fluorescence, transfers its energy to an acceptor molecule through a non-radiative process. The transfer of energy excites the acceptor molecule, which can subsequently release a photon (fluorescence) as it returns back to the ground state. FRET requires the presence of two dyes

that have reasonable spectral overlap, where the emission of the donor dye overlaps with the absorption spectrum of an acceptor dye. This overlap allows several vibrational transitions to have the same energy and therefore be in resonance. In order for this resonance energy transfer to occur the two dyes must be in close physical proximity (a few nanometers). Energy transfer efficiency can be defined in eqn 1.

$$E = \frac{1}{1+\left(\frac{r}{R_0}\right)^6} \qquad \text{Eqn 1}$$

The variable r is the distance between donor and acceptor and R_0 is the Förster radius. The Förster radius is the critical distance with 50% transfer efficiency. It is important to note the relationship to distance and radius is inversed to the sixth-power; therefore there are dramatic differences in FRET efficiency near the Förster radius. Due to the heavy reliance on distance, this fluorescence phenomenon has been used as a spectral ruler to determine distances between fluorophores[10]. This phenomenon has also been used to study conformational changes in biomolecules, protein interactions, and enzyme kinetics[11].

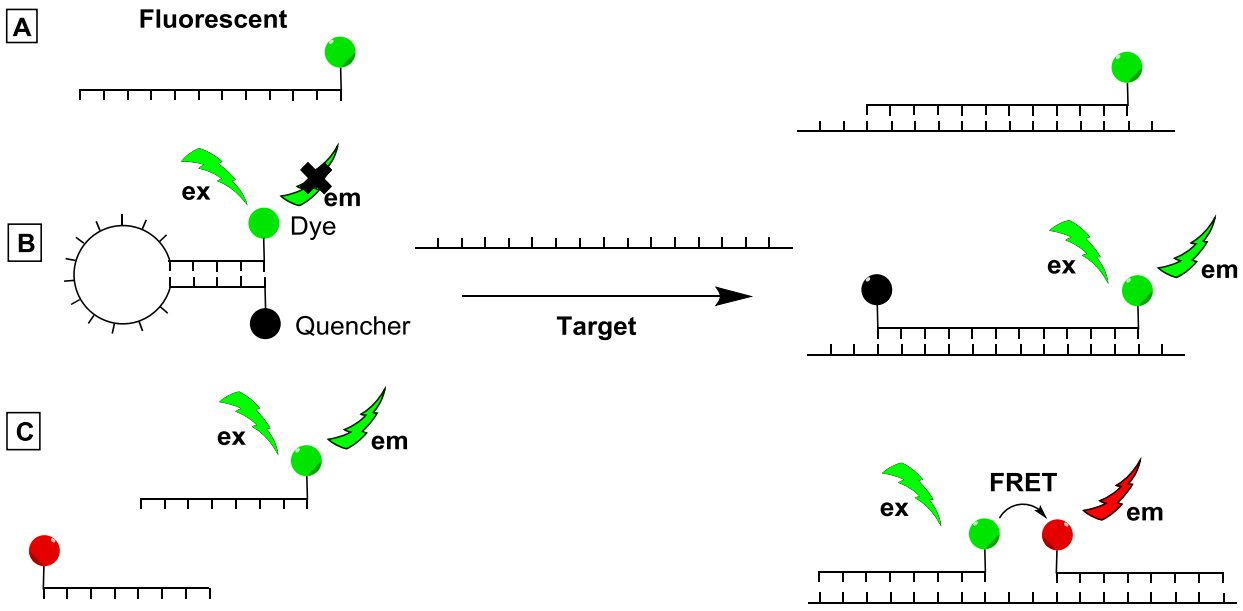
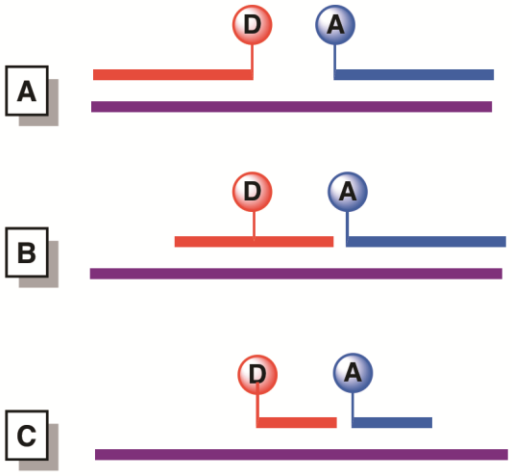


Figure 1: Hybridization probe strategies: (A) Traditional fluorescently-labeled probe (B) Molecular beacon with dye/quencher pair (C) Binary probes with FRET pair

Background fluorescence from unbound probes in FISH experiments is typically minimized through washing to remove unbound probes[12]. In order to reduce the high background of terminally-labeled probes, new generations of fluorescently-labeled probes were designed[13,14]. Molecular beacons and binary probes continue to be popular strategies, each with their own merit. A molecular beacon is typically a single nucleic acid designed with a hairpin/loop with either a fluorophore/quencher or FRET system at the 5'/3' end of the sequence (Figure 1B)[15-17]. In the presence of the target the hairpin is opened, allowing for fluorescence signal to be monitored. This system theoretically allows for low background and fast detection but nonspecific opening of the hairpin has been shown to yield false positive signals[18,19]. Alternatively binary probes are comprised of two dye-labeled strands that are designed to hybridize adjacently on the target (Figure 1C). Both probes are modified with a fluorescent dye (either a donor or acceptor). In the absence of the target, donor excitation should yield no FRET due to the short lifetime of the excited donor and the large distance between the probes at low nanomolar concentrations. However, in the presence of the target, the donor and acceptor are hybridized and are in close proximity, allowing for efficient FRET resulting in quenching of the donor emission and sensitization of the acceptor fluorescence[17,20,21]. The red-shifted acceptor emission can be easily observed with minimal background. This two probe system allows for selective assay, with low background because two hybridization events must occur in close proximity to yield a positive signal.

Previously reported binary probes comprise of long 15-18 oligonucleotide sequences that attempt to maximize FRET signal by attaching the donor at the 3' end of probe A and the acceptor at the 5' end of probe B. In an effort to reduce non-FRET quenching, which would result in weaker acceptor signal, the design allows a few nucleotide separation (1-5 bases) between the two probes (Scheme 1A)[8,17]. However, the unrecognized spacer between the two probes allows mutations, insertions, and deletions to go undetected. To close this gap, other studies have introduced internally dye-labeled nucleic acid probes, allowing recognition of a

contiguous target (Scheme 1B)[22,23]. While this strategy reduces the donor/acceptor distance, the introduction of dyes at internal positions is often costly.



Scheme 1: Designs for binary probes: (A) traditional 3'/5' terminally labeled probes (B) internally/ terminal labeled probe (C) shorter 3' terminally labeled probes

A third option that combines the benefits of the previous two schemes is shown in Scheme 1C: this allows the recognition of a contiguous target without requiring internal labeling. In order to maximize acceptor signal, relatively short probes must be used, which in turn requires high hybridization affinity. While there are other high-affinity DNA analogues, (e.g. LNA[24,25]) we chose γ-modified PNA for the probes. Introduction of a chiral center at the γ position of the traditional PNA backbone induces a helical structure that pre-organizes a γPNA probe for hybridization to complementary DNA or RNA, leading to exceptionally high affinities[26-28]. Recently, our group demonstrated that relatively short (12mer and even 6mer) γPNA *miniprobes* can be used to fluorescently label telomeric repeats in chromosomal DNA[29].

Telomeres are nucleoprotein structures located at the end of linear chromosomes. They are comprised of a repeating single stranded DNA sequence (TTAGGG) and several DNA binding proteins that help protect the DNA from damage[30,31]. Together, they act as a protective cap to the genetic material of the chromosome. [32] In human somatic cells, the telomeres

shorten with each round of replication due to DNA polymerase's inability to fully extend to the end (commonly referred to as the 'end replication' problem). Once the telomeres reach a critically short length and the proteins can no longer efficiently protect the DNA, the cell undergoes senescence[33]. Rapidly dividing cells such as stem cells, germ cells and cancer cells have other mechanisms to solve the end replication problem[34]. The 'end replication' problem is one source for short telomeres; however, several studies point to oxidative stress having a larger contribution to shorten telomeres. Oxidative stress has been shown to cause DNA mutations including double strand breaks and nucleobase modifications, the most common being oxidation of guanine forming an 8-oxoG[35]. Studies suggest a disproportional amount of oxidative damage to occur at the telomeres due to the G rich nature of the sequence, proposing that telomeres can act as stress sensors for the cell[35-37]. In addition to several diseases that show a strong relationship to telomere shortening. [38,39]

Mechanistic questions and disease implications have motivated scientists to develop techniques to determine the length of the telomeres. [34] Existing techniques to measure telomere length range from qualitative techniques like Southern blot to quantitative techniques (e.g. Quantiative PCR, Flow-FISH and Q-FISH). The most common method is fluorescence in-situ hybridization (FISH) which uses a fluorescently-labeled PNA to hybridize to the telomeric repeats in fixed permeablilized cells[9].

This study focuses on minimizing the background signal produced from free or unhybridized probes for labeling telomeric DNA. To address this issue we adjust the design of a traditional binary probe. In order to simplify the synthesis we use two terminally dye-labeled probes designed to hybridize in tandem. Since FRET is distance dependent, in order to achieve efficient energy transfer, we shortened the probes to a 9-mer. We used γPNA, a high affinity DNA analog, to rescue any loss of affinity due to probe length. This design is especially attractive for telomeric repeats where the donor/acceptor probes will alternate. This alternating pattern allows one donor to transfer energy to two acceptors facilitating more efficient energy

transfer at longer wavelengths where lower background fluorescence can be detected. As seen in previous designs of binary probes, using a FRET signal as a reporter requires a two-probe hybridization process to occur, which reduces false positive signals from off-target effects[18,19]. We test this design in telomeric DNA detection assays and staining in cells and tissues.

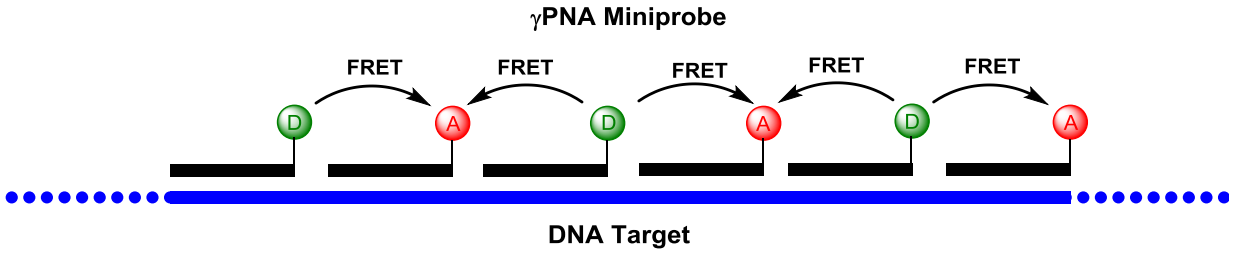
3.2. Design and Rationale

Traditional PNA probes used for telomeric labeling utilize an 18-mer oligomer, allowing hybridization of one dye-labeled probe for every three telomeric repeats[9]. Recently we have reported on a 12-mer γPNA oligonucleotide that allowed one dye labeled probe for every two repeats[29]. Further, by designing shorter 9-mer γPNA probes, we allow one dye-labeled probe per 1.5 repeats. Due to the nature of the hexameric repeat target, the 9-mer PNA probes were designed to have relatively similar GC content (see Chart 1).

Chart 1: Sequences of γPNA and DNA used in study

Name	Sequence
Telo-A	H_2N- TCC CAA TCC-H
Telo-B	H_2N- CAA TCC CAA-H
Telo-A-Cy3	H_2N- TCC CAA TCC-Cy3
Telo-B-Cy5	H_2N- CAA TCC CAA-Cy5
Telo-3 (DNA)	5'-(AGG GTT)$_3$-3'
Telo-6 (DNA)	5'-(AGG GTT)$_6$-3'

Due to the two different sequences we were able to attach either a donor (Cy3) or acceptor (Cy5) to the terminus of each sequence allowing for a binary probe system in which the donor (Telo-A-Cy3) and acceptor (Telo-B-Cy5) strands hybridize in alternating fashion along the telomere target (see Scheme 2). This alternating fashion allows a donor to be in close proximity to two acceptors resulting in enhanced FRET. By designing a binary probe we aimed to reduce background signal and remove a washing step in FISH assays.



Scheme 2: Design of FRET miniprobes on a repeating DNA target

3.3. Results

The results are divided into two parts: first, the physical characterization of the γPNA FRET probes with DNA and secondly, the application of the γPNA FRET probes in a homogeneous solution assay for the detection of telomeric DNA in genomic DNA; in addition to the detection telomeric DNA in biological samples (e.g. cell and tissue samples).

3.3.1. Physical Characterization

In the following sections the γPNA FRET miniprobes' physical properties are studied. We analyzed the γPNA FRET miniprobes affinity to several DNA telomeric repeats in varying length using UV melting analysis. Then we confirmed the secondary structure of the γPNA FRET miniprobes as well as the target sequences using CD. Then we studied the FRET efficiency of the γPNA FRET miniprobes for three DNA targets that varied in length to simulate the effect of a repeating target.

3.3.2. UV Melting Curves

γPNA hybridization properties were assessed using UV melting curve analysis monitoring at 260nm. γPNA oligomers characteristically bind with high affinity to complementary targets thus yielding high melting temperature (T_M) values[27].

First, we analyzed the thermal stability of each individual dye-labeled γPNA probe that was hybridized to an 18-mer complementary DNA target designed to hybridize a full FRET pair (Telo-3, Chart 1). Telo-A-Cy3 and Telo-B-Cy5 gave T_ms of 71.5 and 72.5 °C respectively (Figure 2 and Table 1). The similar T_m values for each individual probe reflects their similar G-C sequence composition. Upon introduction of both dye-labeled probes, thus forming a full FRET pair, the thermal stability was comparable to the individual probes at 69.6°C.

Table 1: UV melting temperatures (°C) recorded for PNA/DNA duplexes

Name	Telo-$_{18}$ (DNA)	Telo-$_{36}$ (DNA)
Telo-A-Cy3 / Telo-B-Cy5	69.6 ± 0.8	71.5 ± 0.9
Telo-A / Telo-B	64.1 ± 0.3	65.9 ± 0.3
Telo-A-Cy3	71.5 ± 0.8	-
Telo-B-Cy5	72.5 ± 0.3	-

Next, we analyzed the effect of target length on the thermal stability by increasing the target from 18 nucleotides to 36 nucleotides (Telo-6), allowing for four probes to occupy one target. The longer target showed a 4°C increase when compared to the 18-mer target. This increase in thermal stability is consistent with previous studies that showed evidence of a cooperative effect upon binding multiple PNAs to a single DNA target[29].

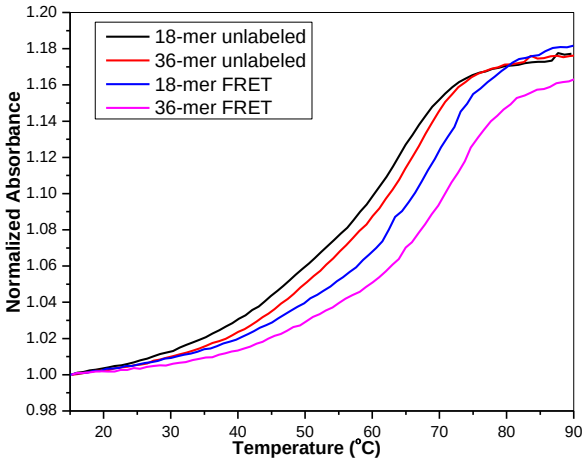


Figure 2 : UV melting curves monitored at 260nm. Samples contained dye-labeled and unlabeled γPNA Telo-A and B hybridized with Telo-3 and Telo-6 DNA targets (as indicated in the legend). Samples were buffered in 10mM HCl; 0.1 mM EDTA; 100mM KCl.

Finally, we examined the effect of the dye on the thermal stability. It was observed that the unlabeled probes, when hybridized to either the Telo-3 or Telo-6, showed a T_m of 64.1°C and 65.9°C respectively. The thermal stability of unlabeled probe is approximately 5°C less than the dye-labeled probe irrespective of the target, suggesting that the addition of a dye provides additional thermal stability.

Overall the UV melting experiments confirm that the dye-labeled γPNA probes can hybridize to a complementary DNA target to form hetero-duplexes, and the addition of the dye and a cooperative effect due to end stacking may provide extra thermal stability. Both of these properties are important for targeting repeating sequences.

3.3.3. Circular Dichroism Spectropolarimetry

Circular Dichroism spectropolarimetry (CD) experiments are particularly useful to determine the secondary structure of oligonucleotides in both single stranded and double stranded forms. The CD experiments were performed at room temperature with unlabeled γPNA Telo-A and Telo-B (2uM each) with and without DNA Telo-6.

The telomeric DNA is a well-known quadruplex forming sequence, meaning that the secondary structure will be different than a single stranded DNA target. The CD spectrum shows distinct positive peaks at 260nm and 295nm, and negative band around 240nm (see Figure 3); suggesting a mixed parallel/antiparallel hybrid-quadruplex, as seen in previous studies[40]. Typically, B-form DNA shows a distinct CD spectrum with positive peaks around 275nm, 220nm and a negative peak around 240nm indicating a helical conformation[41]. As illustrated in Figure 3, the unlabeled γPNA Telo-A and Telo-B contains positive peaks around 275nm, 220nm and a negative peak around 210nm. As seen in previous studies, the presence of these peaks suggests pre-organization of B-form helical conformation of the γPNA; which is distinctly different than the DNA target.

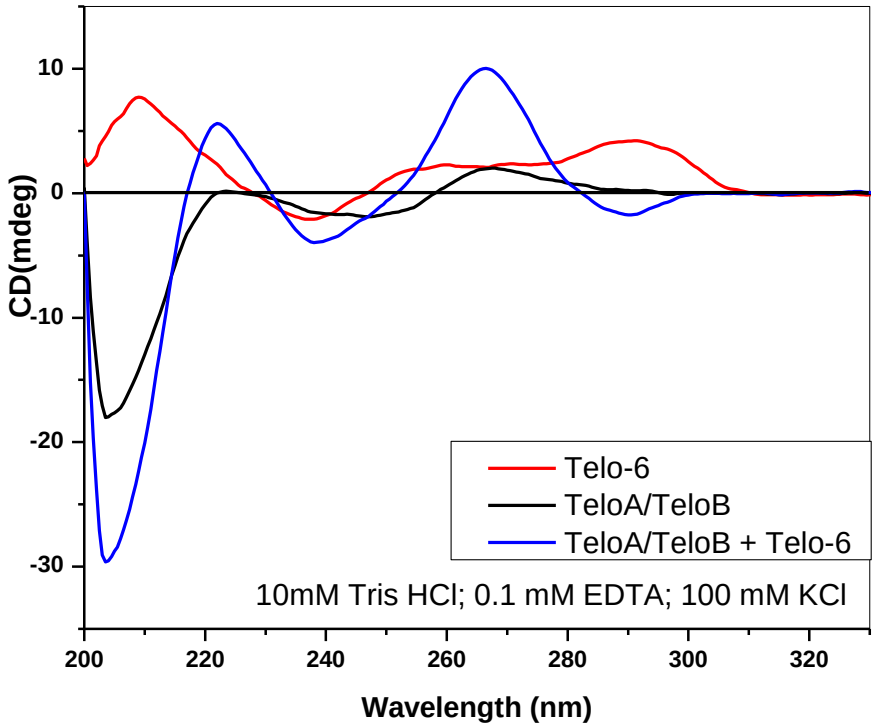


Figure 3: CD Spectra of various samples containing Telo-6 (Red), complementary PNA (2uM), (Black) and Telo-6 (1uM) with complementary PNA (2uM) (Blue). All samples were pre-annealed. Buffer : 10mM HCl; 0.1 mM EDTA; 100mM KCl

Conversely, upon hybridization of the DNA target, the positive peaks at 220 and 260 and negative peaks at 240 and 290 remain, consistent with right-handed helical behavior. Enhancement of the dip at 240 indicates a tighter helical pitch which is needed to bind to a DNA target as described in previous studies. The CD spectra indicate pre-organized secondary structures of γPNA which is consistent with the literature[27].

3.3.4. Fluorescence Spectroscopy

In these fluorescence experiments, the aforementioned γPNA probes (Chart 1) were designed to hybridize to adjacent sequences on telomeric DNA. These probes were expected to undergo efficient FRET due to the relatively short distance between the donor (Cy3) and the acceptor (Cy5) dyes when co-hybridized to the same DNA target. When the FRET probes are excited at 500nm in the absence of the target DNA strand a strong fluorescence signal from Cy3

is observed, with minimal signal from the acceptor, Cy5 (Figure 4A). Upon introduction of a DNA target (Telo-3), FRET is activated by the close proximity of the probes leading to quenching of the Cy3 donor signal and sensitization of the Cy 5 acceptor signal. The FRET efficiency, determined by quenching of the donor, was calculated using the formula: $E = 1 - \frac{I_{Cy5}}{I_{Cy3}}$.

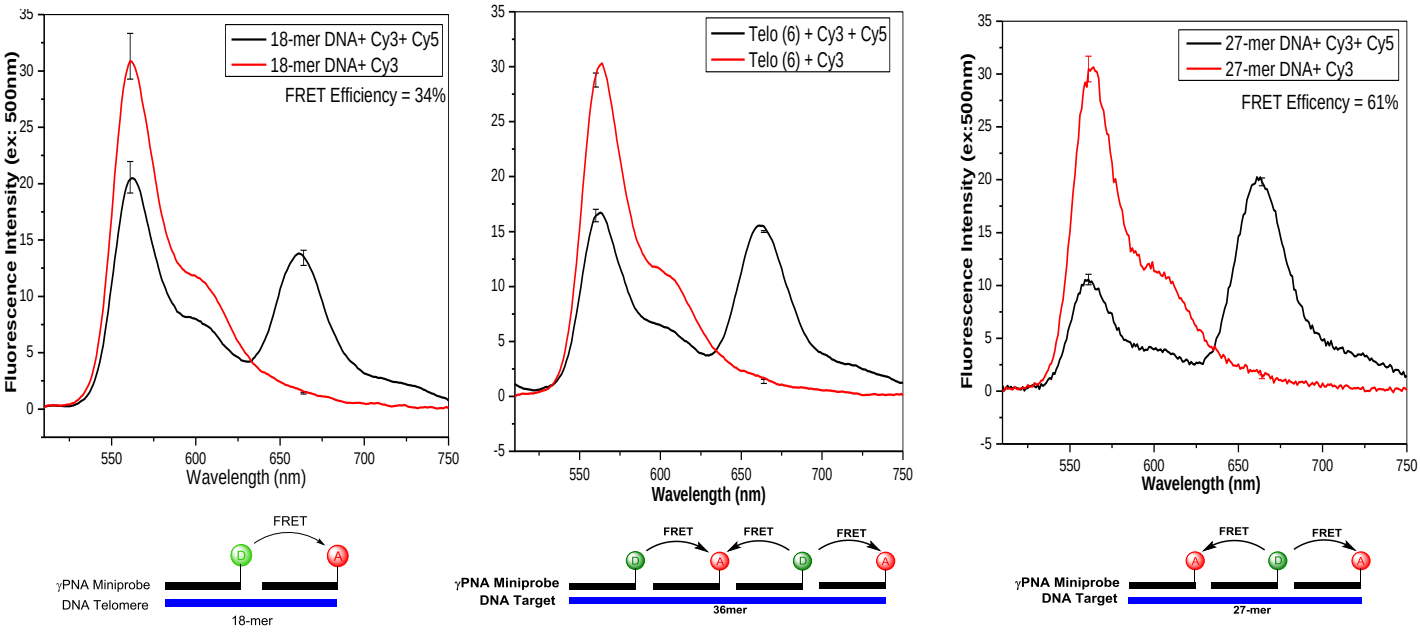


Figure 4: Fluorescence spectra of FRET Probes with varying DNA targets: (A) Telo-3 with Telo-A-Cy3 + Telo-B-Cy5; (B) Telo-6 with Telo-A-Cy3 + Telo-B-Cy3; (C) Telo-4.5 with Telo-A-Cy3 + Telo-B-Cy3. The schematic below each spectrum illustrates the donor/acceptor placement according to the sequence of the DNA. All samples were averaged (D-A) and (D). Samples were excited at 500 nm.

The DNA target Telo-3 allows hybridization of one complete FRET system (i.e. one donor and one acceptor γPNA). This target is equal to three telomeric repeats or 18 nucleotides and yields a 34% FRET efficiency based on the extent of Cy3 quenching. As illustrated in Figure 4B, extending the target length from three repeats to six repeats (Telo-6) allows for two copies of each γPNA to hybridize and the FRET efficiency increases to 43%. As shown in the schematic under the spectra, it is reasonable to attribute the enhanced FRET to the additional acceptor dye in close proximity to the second Cy3.

Although FRET is observed, we expected higher FRET efficiency than 34% for Telo-3. In Telo-3, the donor and acceptor are 9 base pairs apart, which excluding linkers places the dyes approximately 29 Å apart. Using the equation for FRET efficiency (E) ($E = \frac{1}{1+\left(\frac{r}{R_0}\right)^6}$) (Cy3-Cy5 R_o=60 Å;) we predicted a FRET efficiency of 98%. A few factors may account for the difference in experimental vs calculated FRET efficiencies. The donor quenching was monitored using Cy3 labeled Telo-A γPNA and an unlabeled Telo B γPNA to account for any differences in quantum yield when fully hybridized. The transition dipole orientation factor(κ^2) or the relative orientation of the transition moments between the donor emission and acceptor absorption is typically assumed to be 2/3[42]. However, this assumption may fail in certain designs especially with restricted dyes. Allowing the dye more rotational freedom by using a longer linker may increase κ^2, which will increase the FRET efficiency.

In order to study the 'end effect' we removed the end donor to compare the energy transfer contribution of a donor at the end vs in between two acceptors. To study this effect we used an intermediate DNA sequence: a 27 nucleotide target that could hybridize only one Cy3-labeled γPNA bounded by two Cy5-labeled γPNAs. This system showed an increase in FRET efficiency to 61% (Figure 4C). In an extended telomere target, where tens or even hundreds of γPNAs could hybridize, the "end effects" due to inefficient quenching of a terminal Cy3-labeled probe should be insignificant. Moreover, the efficient Cy5 sensitization observed upon hybridization should allow for an excess of probe to be present during assays, in addition to eliminating the need for a washing step during Q-FISH analysis of telomeric DNA.

3.4. Results- Applications for FRET Miniprobes

In order to demonstrate the utility of the FRET miniprobes, we tested the system in both a homogeneous solution assay as well as in fixed cells and tissues.

3.4.1. Fluorescence Detection Telomeric DNA

An application of these FRET miniprobes is the fluorescence-based detection of DNA in vitro. This idea of using a binary probe with a FRET signal has been well-established using DNA or PNA probes for sequence detection in PCR. An advantage to this system is the dramatic difference in acceptor signal with and without the target allowing clear detection of analyte. One could envision an assay where samples of genomic DNA could be extracted from blood samples and the telomere repeats could be quantified by adding the FRET miniprobe in a high-throughput fashion using a plate reader. This assay would be an alternative to q-PCR and Southern blot approaches currently utilized. As a proof of concept, we demonstrated the utility of this assay by detecting a telomeric repeat (Telo-6) of DNA by first determining the limit of detection, and optimizing the assay for high-throughput analysis.

3.4.2. LOD for Telomeric DNA Analysis

In this study, we tested the limit of detection for our assay using DNA Telo-6 with 6 telomeric repeats designed to hybridize two donors and acceptors (see Figure 4B for illustration). This sequence was used to represent a DNA sequence with multiple repeats in a situation where multiple miniprobes could bind to a continuous repeating target. Equal molar concentration of each FRET miniprobe (100nM) was used. The excitation wavelength was optimized to 470nm to prevent directly exciting the acceptor which would result in false-positive signal and higher background fluorescence. The fluorescence intensity for the acceptor was measured at 660nm. As seen in Figure 5, the fluorescence spectra clearly show a Cy5 sensitization at 660nm with a linear increase with target concentration, suggesting value for DNA/RNA quantification assays. We titrated known amounts of Telo-6, which were used to

generate a calibration curve (see Figure 5). The limit of detection (as determined by the fluorescence intensity at 660nm of a sample without DNA; LOD= AVE + (STD*3)) using the calibration curve was calculated to be 1.72nM.

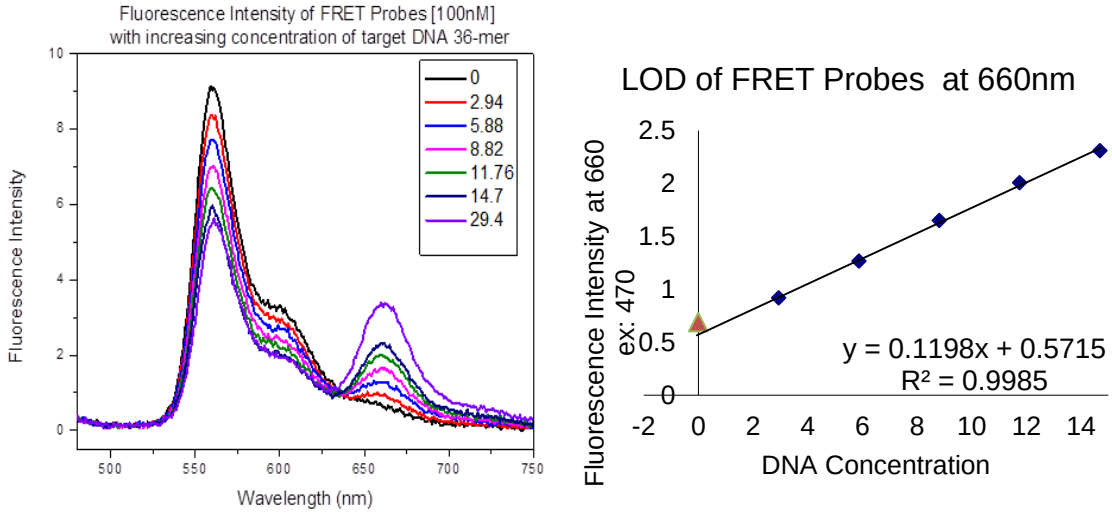


Figure 5: (Left) Fluorescence Spectra of FRET miniprobes (100nM) with varying concentrations of Telo-6 DNA. Samples were excited at 470nm. (Right) Calibration curve generated from the aforementioned titrations; graph of fluorescence intensity at 660nm vs Telo-6 DNA concentration (nM). Samples were buffered in 10mM HCl; 0.1 mM EDTA; 100mM KCl.

The aforementioned experiments were performed in a cuvette, which does not translate well for high throughput analysis. Translating the assay to a high throughput assay using a 96-well plate required some optimization. There was a striking difference observed in the fluorescence spectra between the quartz cuvette and the plate reader (see Figure 6A). Looking at several variables, it was found that addition of a surfactant (0.1% Triton-X) to the buffer rescued fluorescence response of the FRET miniprobes with varying concentrations of DNA. The following experiment shows the fluorescence spectra of Telo-6 samples at known concentrations using a 96-well plate. These preliminary results provide insights for using this type of assay for detecting telomeric repeats in genomic DNA and suggest feasibility for translating to a high-throughput assay.

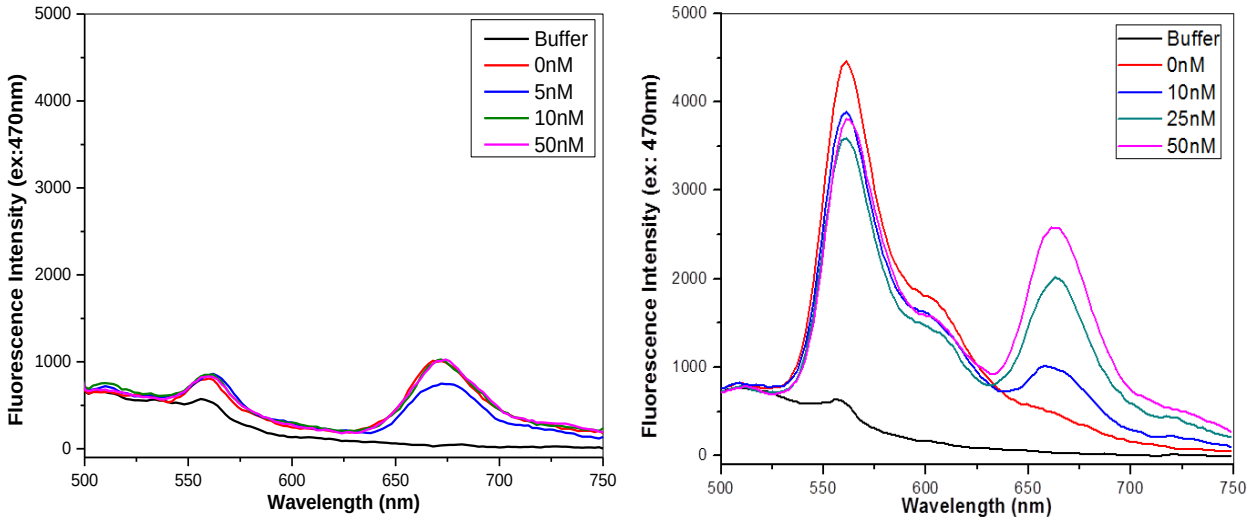


Figure 6: Fluorescence spectra of FRET miniprobes (100nM) with varying concentrations of Telo-6 DNA (nM) (A) without surfactant and (B) with surfactant (0.1% Trixon-X). Samples were prepared in buffer (10mM HCl; 0.1 mM EDTA; 100mM KCl), in a 96-well plate and excited at 470nm.

3.4.3. Intracellular DNA Detection of Telomeric Repeats (Dr. Patty Opresko's Lab)

In order to demonstrate the utility of the FRET miniprobes in biological samples, Dr. Patty Opresko (University of Pittsburgh), was able to image telomeres in BJ skin fibroblasts (Figure 7), U2OS cells (Figure 8), and mammalian tissue samples prepared from mouse colon (Figure 9) using standard FISH protocols.

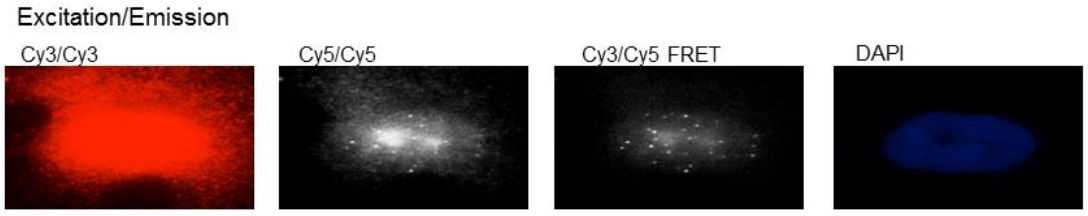


Figure 7: The images show an interphase cell nucleus from telomerase expressing BJ skin fibroblasts stained with a mixture of the Cy3 and Cy5 labeled γPNA 9mers that are complementary to telomeric DNA and were left unwashed. The image was captured under three excitation/emission conditions as indicated using an inverted fluorescence microscope, and a 60x lens. Images and analysis by Patricia L. Opresko, Connor T. Murphy and Melinda Sager.

As illustrated in the interphase cell nucleus images in Figure 7, the red indicates the Cy3 emission channel and white shows the Cy5 emission suggesting that both probes are present.

Under FRET conditions the Cy3 dye is excited and donates energy to a neighboring Cy5 dye which is visualized in the Cy5 emission channel. Unbound probe was not washed away as observed upon direct excitation and visualization of Cy3/Cy3 (ref) or Cy5/Cy5 (white). The FRET condition (Cy3/Cy5) reveals distinct telomeric foci and loss of the unbound probe signal.

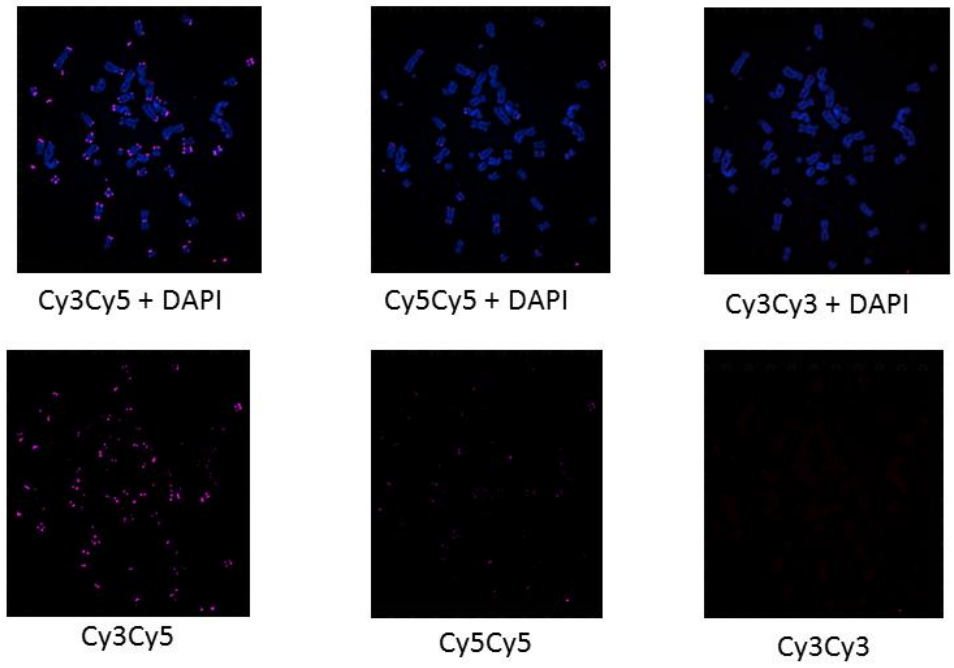


Figure 8: Metaphase chromosome spreads prepared from U2OS cells and stained with a mixture of the Cy3 and Cy5 labeled γPNA 9mers that are complementary to telomeric DNA. The image was captured under three excitation/emission conditions as indicated. Samples were counterstained with DAPI. Images and analysis by Patricia L. Opresko, Connor T. Murphy and Melinda Sager.

As illustrated in Figure 8, directly exciting both Cy3 and Cy5 and capturing their respective images, marginal signal was observed. However, upon exciting Cy3 and capturing the image through the FRET channel, telomeric DNA fragments were able to be observed.

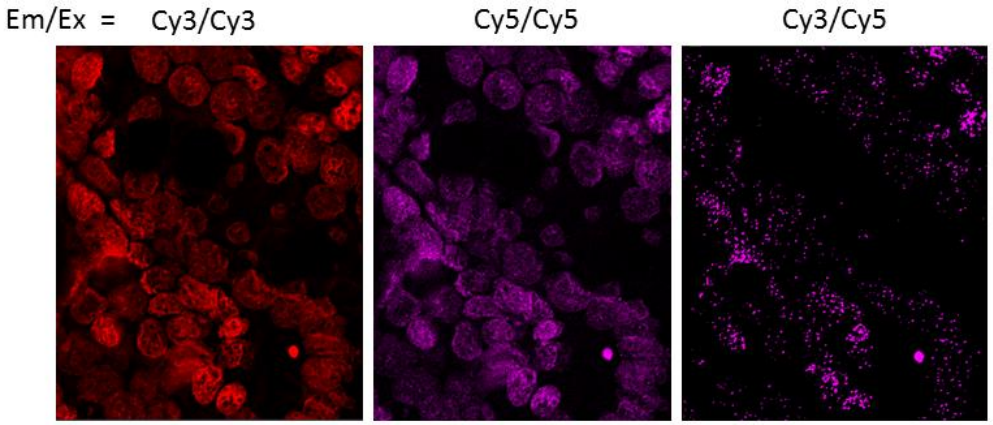


Figure 9: Mouse colon tissue stained with FRET 9mer γPNA miniprobes. Images were deconvoluted to remove out-of-focus light. Images and analysis by Patricia L. Opresko, Connor T. Murphy and Melinda Sager.

The FRET miniprobes successfully stained telomeres in interphase nuclei, metaphase chromosome spreads and tissue. These biological assays are promising and will continue to be an area for future development.

3.5. Discussion

Telomere length has been suggested to be a reliable indicator of the health of a cell. Studies have found correlations of shortened telomeres with increased oxidative stress[35], disease[38], and even life stress[43]. These correlations encourage scientists to develop reliable techniques to measure the length of telomeres.

Current techniques to measure the telomere length include: southern blots, Q-PCR, Q-FISH and Flow-FISH, each with their own merits and limitations. Southern blots of genomic DNA digested with restrictive enzymes require large quantities of cells and provide an average of the number of repeats. Quantitative PCR (Q-PCR) methods are generally fast and accurate but do not translate well to telomeric repeats due to the high guanine content. Q-FISH and Flow-FISH rely on dye labeled PNA to provide an accurate signal to produce reliable data. Initial studies using dye labeled PNA reported intense staining of telomeres but required optimizations in the FISH protocol to reduce the high background, when compared to dye labeled DNA or RNA[9]. Extensive studies have tried to optimize the Q-FISH and Flow-FISH

protocol to obtain the greatest signal to noise ratio[12]. With the use of the FRET miniprobes, minimal washing was required to obtain signal from the telomeres. The low background signal and minimal washing that FRET miniprobes provide could be an interesting advancement in Flow-FISH for telomere length measurements particularly for high-throughput analysis of telomere length of a cellular population.

By utilizing γPNAs we were able to shorten and simplify the binary probe design with two terminally-labeled γPNA probes, designed to produce a FRET signal when hybridized in tandem. A promising feature of this design is the FRET signal enhancement achieved due to the alternating motif. Coupling this design with a repeating target increases the likelihood of a donor with two adjacent acceptors ensuring improved FRET signal enhancement. A binary system requires two hybridization events, which increase the accuracy of detection and allows the assay to have excess probes to ensure complete binding to the target. This design allows for applications for nucleic acid detection in vitro as well as in vivo. Binary FRET probes have been used extensively in vitro. To the best of our knowledge FRET probes have not been used to measure telomeric repeats.

The in vitro results emphasize the idea that even in the presence of excess probes, signal from the target results in a distinguishable fluorescence signal at low nanomolar concentrations. We performed supplementary studies to analyze telomeric RNA. Physical characterization showed similar CD structure and higher affinity (T_M) for the telomeric RNA target, in addition to a similar fluorescence response. These preliminary results suggest further applications for the Telo-FRET miniprobes. Due to having high-affinity, γPNA allowed for shortening of the binary probes while retaining a FRET signal and recognizing a continuous sequence. We explored the FRET miniprobes strategy with other repeating sequences such as the ALS hexanucleotide repeat. Two 9-mer FRET γPNA miniprobes were designed to hybridize to the hexanucleotide repeat of C_2G_4. Preliminary experiments of the physical and fluorescence

properties are comparable to that of the telomere target (Appendix). These findings suggest that the concept of FRET miniprobes can be transferred to different targets as long as the probes have similar G/C content.

3.6. Future Implications

Our FRET miniprobe strategy can also be applied to non-repeating sequences, thus expanding the application to mRNA labeling. However, in order for this strategy to be utilized for mRNA detection, we must first minimize non-specific binding to prevent misleading signals. In addition, the development of a computational method to design γPNA probe sequences to have similar affinity, so there would be less reliance on G/C content would be beneficial. While expanding to more complicated sequences may allow for shorter probes to be utilized, previous studies suggests that a fully-modified 6-mer is may be too short to hybridize with high affinity to a telomeric target without additional modifications to the nucleobase (e.g. G-Clamp)[29]. However, different sequences should be explored to find a thermodynamic limit for RNA detection in FISH assays.

As illustrated in the tissue imaging experiments, some non-specific binding was observed which required removal of unbound probe. Future generations of γPNA FRET miniprobes should optimize off-target effects by incorporating a structured probe, either a toehold or molecular beacon design. This additional structure limits the non-specific binding that could occur in the complicated environment of a cell.

Although the low nanomolar limit of detection found for the homogeneous assay is promising, future work should focus on optimizing the experimental conditions to detect telomeric repeats at lower concentrations. In addition, further studies should aim to validate our construct with samples generated from genomic DNA, and compare assay accuracy to other well-established methods such as Q-PCR and Southern blot. Overall, the in vitro results provide a framework for future studies towards optimizing a high throughput assay for telomeric

DNA detection and quantification. Although this study improves the signal-to-noise ratio by incorporating short probes with a FRET signal, the brightness, and therefore the sensitivity, is comparable to the commercially-available 18-mer PNA probe. Future studies should explore multi-dye labeled PNA probes, to increase the brightness of the probe. In addition, strategically placed dyes on an internally-labeled probe could contribute to more efficient energy transfer. A promising approach to internally-labeled PNAs for multi-dye PNA probes is presented in Chapter 4.

3.7. Conclusions

Here we offer an alternative method to detect nucleic acid targets, using simple terminally labeled γPNA probes as shortened binary probes. The system provides a modified method for low background detection of telomeric DNA for cell and tissue imaging. We present preliminary studies that suggest potential high throughput screens for telomeric DNA in genomic DNA samples, cells and tissues. Our design of FRET miniprobes could eventually lead to high throughput gene expression screens providing information rich analysis of the physical state of the patient leading towards more personalized medicine.

3.8. Experimental

3.8.1. *Biophysical Experimental*

The DNA sequences were designed to have either 3 or 6 repeats and therefore designed to hybridize one FRET pair probe, or two FRET pair probes (see table 1). For UV and Fluorescence experiments the samples were prepared in a buffer containing 10 mM Tris-HCl, 0.1 mM EDTA, and 100 mM KCl. The same samples were used in UV melting experiments and fluorescence experiments. All samples contained concentrations of 3 uM repeat sequences, meaning that for the full target (Telo-6): 0.5uM = 36 nucleobases = 6 repeats = hybridized to 2 probe donors and 2 probe acceptors each 1uM (see scheme 1). All experiments were run in

triplicate. γPNA probes were purchased from PNA Innovations, DNA sequences were purchased from IDT.

3.8.2. UV Melting Curves

Thermal UV melting experiments were recorded on a Varian Cary 300 Bio UV-Vis spectrophotometer using a thermoelectrically controlled Peltier cell holder. UV melting experiments were performed by heating the samples up to 90°C and slowly cooling (1°C per min) to 15°C, then the samples were allowed to equilibrate for 5 mins at 15°C then slowly heated to 90°C (rate= 1°C / min) and the absorbance at 260nm was collected. The resulting data was normalized and the first derivative was plotted. The maximum of the first derivative was recorded as the T_M.

3.8.3. Circular Dichroism Spectropolarimetry

CD measurements were performed on a Jasco J-715 CD spectropolarimeter equipped with peltier water circulating temperature controller. Samples were prepared by mixing 2uM Telo A and Telo B and 1uM 36-mer DNA target in buffer containing 10mM Tris HCl; 0.1 mM EDTA; 100 mM KCl. All samples were annealed by heating up to 95°C for 2 minutes and then slowly cooled to room temperature. Each spectrum represents an average of 6 scans, baseline corrected and collected at 25°C.

3.8.4. Fluorescence Spectroscopy

Fluorescence emission spectra were collected on a Cary Eclipse Fluorescence spectrophotometer at 25° C by exciting the samples at 520nm and 600nm and reading from 530-800nm and 610-800nm respectively. UV absorbance spectra from 750-240nm were also collected for each sample.

3.8.5. LOD for Telomeric DNA Analysis

Samples containing equal molar concentration of 100nM of each Probe (TeloA (Cy3), TeloB (Cy5)) were prepared. The absorbance and fluorescence spectra were recorded on a Tecan Infinite M1000 plate spectrometer and the samples were excited at 470nm. The intensity at 660nm for the sample without DNA was measured and the AVE and STD was found (LOD = AVE + (STD*3)). A calibration curve was then generated using known concentrations of DNA Telo-6 and the fluorescence the intensity at 660nm was plotted against DNA. A calibration curve was generated. The LOD was found using calibration curve.

3.8.6. DNA Detection Assay

DNA Samples were prepared in 4 different concentrations (10, 25, 50, 75 nM) in Buffer (10 mM Tris HCl; 0.1mM EDTA; 100mM KCl; 0.1% Triton-X). Samples were prepared in the 96-well plate with 48uL of DNA sample and mixed with 2uL of 2uM solution of FRET miniprobes, for a final probe concentration of 100nM. Fluorescence was measured using Tecan Infinite M1000 plate spectrometer at room temperature. Samples were excited at 470nm and intensities were measured at 550nm (Cy3) and 664nm (Cy5) as well as full spectrum scan from 495-750 nm. Each experiment was performed in triplicate.

3.9. Appendix

3.9.1. Telomeric RNA

Telomeric repeat-containing RNA (TERRA) are long noncoding RNAs transcribed from telomeric regions and studies have suggested alternative roles in regulation pathways[44]. Because of emerging evidence of signaling and regulation we wanted to explore the how our FRET miniprobes respond to TERRA. Here we present preliminary biophysical studies of two telomeric RNA targets (3 and 6 repeats) with our FRET miniprobes.

UV melting experiments

The thermal stability of the γ PNA probes hybridized with an RNA Telo-6 target was determined to be 65.4°C as seen in Figure S1 and Table S1. A simple comparison of the T_M between the DNA and RNA target shows the thermal stability to be essentially the same. Overall the UV melting experiment confirms that the γPNA probes can hybridize to a complementary DNA target to form hetero-duplexes.

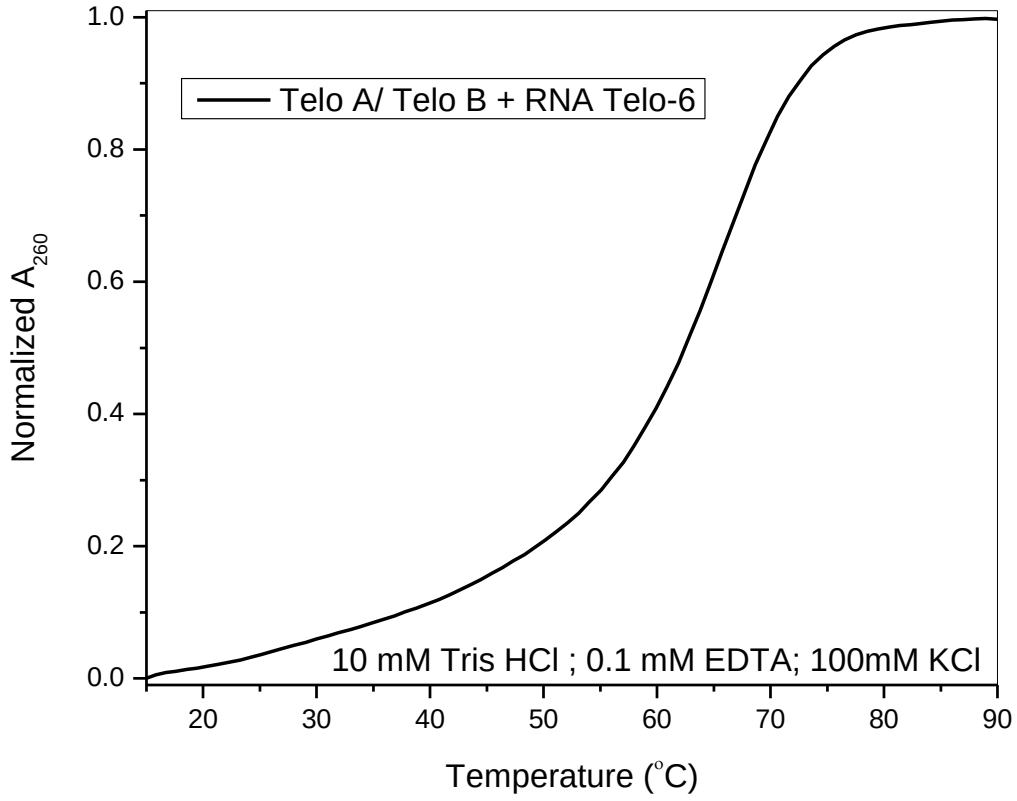


Figure S1: UV melting curves monitored at 260nm. Samples contained 1uM Telo-6 RNA hybridized with 2uM Telo PNA A/B and were buffered in 10mM Tris HCl; 0.1mM EDTA; 100mM KCl.

Table S1: UV melting temperatures (°C) recorded for PNA/DNA duplexes

Name	Telo-$_{36}$ (RNA)	Telo-$_{36}$ (DNA)
Telo-A / Telo-B	65.4 ± 0.3	65.9 ± 0.3

Circular Dichroism Spectropolarimetry

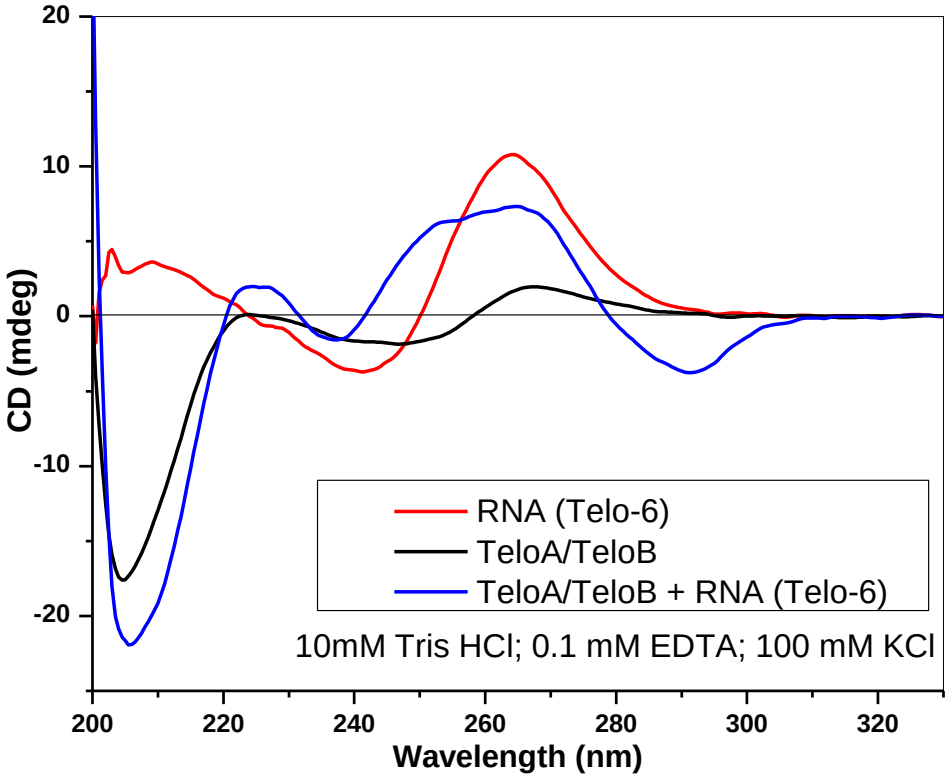


Figure S2: CD Spectra of various samples containing Telo-6 RNA (Red), complementary PNA (2uM), (Black) and Telo-6 (1uM) with complementary PNA (2uM) (Blue). All samples were pre-annealed. Buffer : 10mM HCl; 0.1 mM EDTA; 100mM KCl

The telomeric RNA is a well-known quadruplex forming sequence, typically RNA forms parallel quadruplexes. The CD spectrum shows distinct positive peak at 260nm and negative band around 240nm (see Figure S2); suggesting a parallel quadruplex, as seen in previous studies[40]. As shown before the γPNA shows distinct peaks suggests pre-organization of B-form helical conformation of the γPNA; which is distinctly different than the RNA target. Conversely, upon hybridization of the RNA target, the positive peaks at 220 and 260 and negative peaks at 240 and 290 remain, consistent with right-handed helical behavior. Overall the CD spectra behave as expected and are consistent with the literature.

Fluorescence Experiments

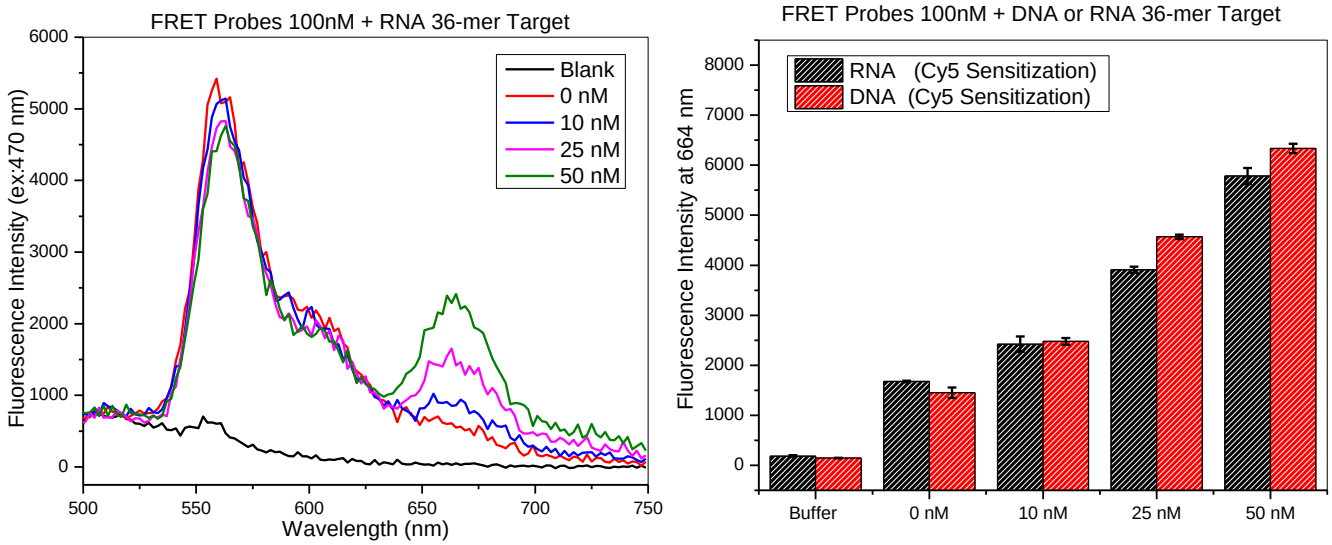


Figure S3: (Left) Fluorescence spectra of FRET miniprobes (100nM) with varying concentrations of Telo-6 RNA (nM). Samples were prepared in buffer (10mM HCl; 0.1 mM EDTA; 100mM KCl; 0.1% Triton-X), in a 96-well plate and excited at 470nm. (Right) Comparison of FRET signal (at 664nm)

The following experiment shows the fluorescence spectra of Telo-6 RNA samples at known concentrations using a 96-well plate. As seen with the DNA target, there is an increase in FRET signal upon addition of the complementary target (Figure S3). These samples were not annealed, indicating that the FRET γPNA were able to invade any secondary structure and bind to the target. As observed in Figure S3, the increase in FRET signal is comparable in DNA and RNA targets. Overall, preliminary experiments exploring the physical properties of the FRET miniprobes with an RNA telomeric target indicate minimal differences between DNA and RNA targets. These preliminary results provide insights for further development of a homogeneous solution based assay.

3.9.2. ALS (DNA and RNA Hexameric Target)

Recent studies have suggested that an expansion of a noncoding hexanucleotide repeat (GGGGCC) may be strongly associated with the neurodegenerative disease ALS. Healthy individuals reported 2-23 repeating units, while ALS patients reported 700-1600 repeating units[45]. Although the mechanism is not well understood, it is suggested that the repeating units generation a cytotoxic RNA foci as observed in FISH assays[46]. Here we present preliminary data looking at the biophysical characterization of ALS targeting FRET miniprobes with DNA and RNA targets (See Chart 1).

Chart 1: Sequences of γPNA and DNA used in study

Name	Sequence
ALS-A-Cy3	H_2N- CCC CGG CCC-Cy3
ALS-B-Cy5	H_2N- CGG CCC CGG-Cy5
ALS-$_{18}$ (DNA)	5'-GGG GCC GGG GCC GGG GCC-3'
ALS-$_{36}$ (DNA)	5'-(GGG GCC GGG GCC GGG GCC)$_2$-3'

UV Melting Curves

The thermal stability of the γ PNA probes hybridized with either a DNA or an RNA target with either 3 or 6 repeats could not be determined at the current conditions (See Figure S4). This high T_M indicating an exceptionally stable hetero-duplex is not surprising due to the prevalence of G/C base pairs.

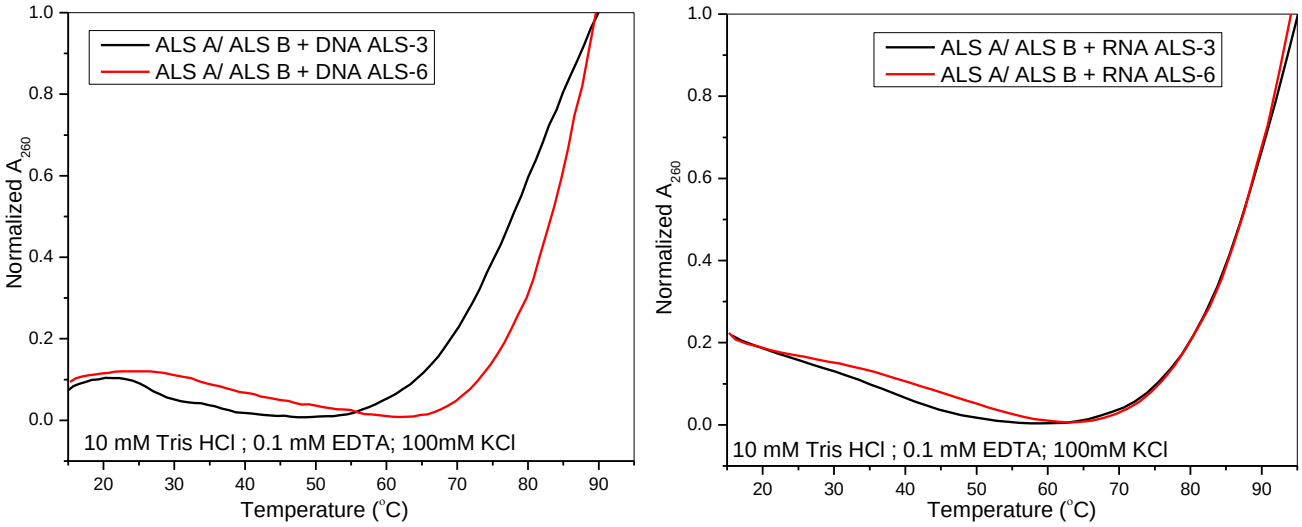


Figure S4: UV melting curves monitored at 260nm. Samples contained unlabeled γPNA ALS-A and B hybridized with ALS-3 and ALS-6 DNA (Left) and RNA (Right) targets (as indicated in the legend). [PNA]=1 uM; [ALS-3]=1uM; [ALS-3]=0.5uM Samples were buffered in 10mM HCl; 0.1 mM EDTA; 100mM KCl.

Fluorescence Spectroscopy

Fluorescence experiments exploring the FRET miniprobes with varied DNA targets of 3 or 6 repeats. As seen in section 3.3.4, introduction of a DNA target (ALS-3 or ALS-6), FRET is activated by the close proximity of the probes leading to quenching of the Cy3 donor signal and sensitization of the Cy 5 acceptor signal. There was not a significant difference between the two DNA targets (ALS-3, ALS-6) with 49 % and 44% FRET efficiency respectively.

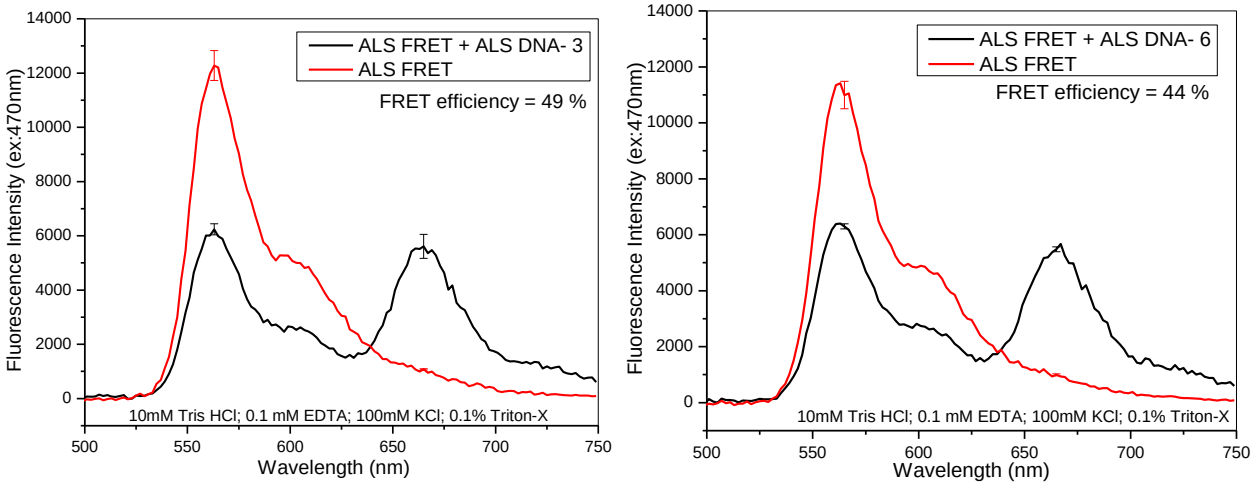


Figure S5: Fluorescence spectra of FRET Probes with varying DNA targets: (A) ALS-3 with ALS-A-Cy3 + ALS-B-Cy5; (B) ALS-6 with ALS-A-Cy3 + ALS-B-Cy3; All samples were averaged (D-A) and (D). Samples were excited at 470 nm.

ALS Circular Dichroism Spectropolarimetry

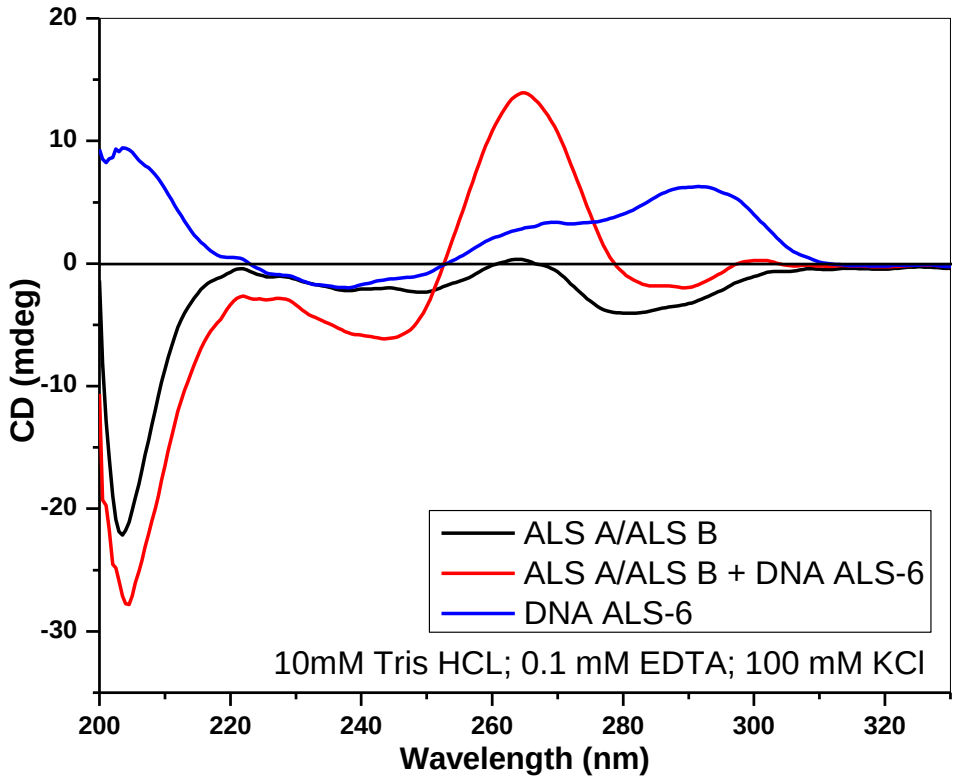


Figure S6: CD Spectra of various samples containing ALS DNA-36 (Red), complementary PNA (2uM), (Black) and ALS DNA-36 (1uM) with complementary PNA (2uM) (Blue). All samples were pre-annealed.

The ALS hexanucleotide repeat has a complicated secondary structure, there is evidence of both hairpin-like structures as well as quadruplex structures[46]. As seen in Figure S6, the DNA suggests the formation of a hybrid parallel/antiparallel quadruplex as described in section 3.3.2 (positive bands at 295, 260 and a negative band at 240nm), although the peaks are not as strong as the telomeric DNA target, which suggest a more complicate picture

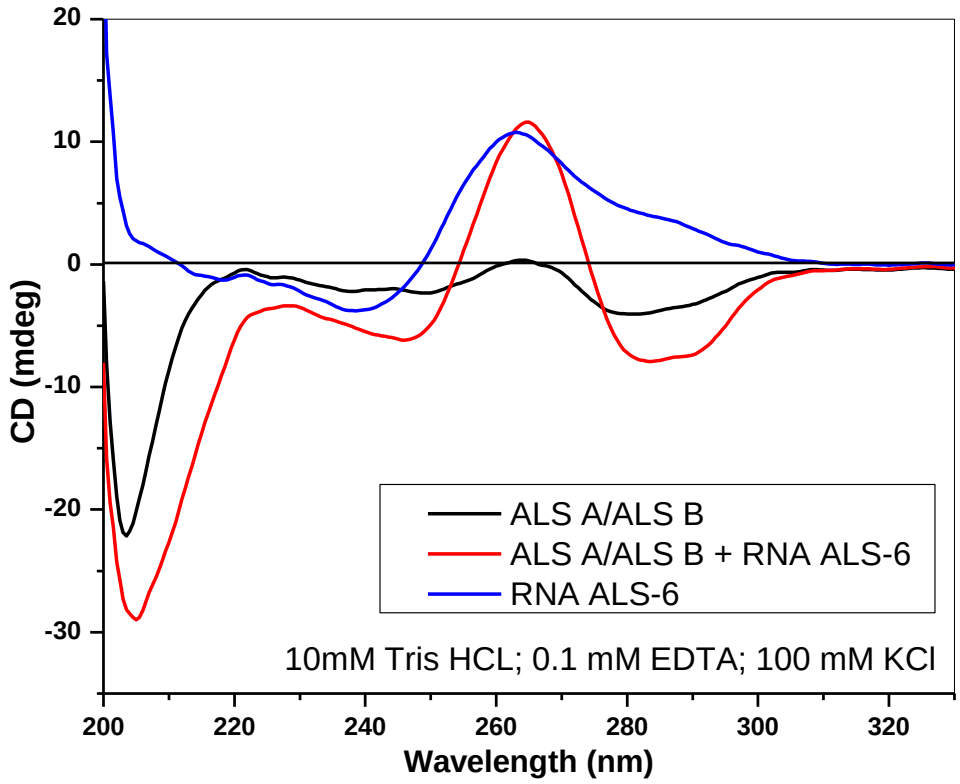


Figure S7: CD Spectra of various samples containing ALS RNA-36 (Red), complementary PNA (2uM), (Black) and ALS RNA-36 (1uM) with complementary PNA (2uM) (Blue). All samples were pre-annealed.

The RNA hexanucleotide repeat is also known to form a hybrid of parallel quadruplex and hairpin structures. RNA forms parallel quadruplexes and the CD spectrum shows distinct positive peak at 260nm and negative band around 240nm (see Figure S7); suggesting a parallel quadruplex, as seen in previous studies[40].

As shown before the γPNA shows distinct peaks suggests pre-organization of B-form helical conformation of the γPNA; which is distinctly different than the both the DNA and RNA target (see Figure S6 and S7). Conversely, formation of the hetero-duplex both DNA/PNA and RNA/PNA, the positive peaks at 260 and negative peaks at 240 and 290 remain, consistent with right-handed helical behavior. Overall the CD spectra behave as expected and are consistent with the literature.

3.10. References

(1) Raj, A.; van den Bogaard, P.; Rifkin, S. A.; van Oudenaarden, A.; Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **2008**, *5*, 877-879.

(2) Itzkovitz, S.; van Oudenaarden, A. Validating transcripts with probes and imaging technology. *Nat. Methods* **2011**, *8*, S12-19.

(3) Femino, A. M.; Fay, F. S.; Fogarty, K.; Singer, R. H. Visualization of Single RNA Transcripts in Situ. *Science* **1998**, *280*, 585-590.

(4) Marras, S. A.; Tyagi, S.; Kramer, F. R. Real-time assays with molecular beacons and other fluorescent nucleic acid hybridization probes. *Clin. Chim. Acta.* **2005**, *363*, 48-60.

(5) Ranasinghe, R. T.; Brown, T. Fluorescence based strategies for genetic analysis. *Chem. Commun.* **2005**, 5487-5502.

(6) Holland, P. M.; Abramson, R. D. Detection of specific polymerase chain reaction product by utilizing the 5'----3'exonuclease activity of Thermus aquaticus DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **1991**.

(7) Tsuji, A.; Koshimoto, H.; Sato, Y.; Hirano, M.; Sei-Iida, Y. Direct observation of specific messenger RNA in a single living cell under a fluorescence microscope. *Biophys. J.* **2000**.

(8) Robertson, K. L.; Yu, L.; Armitage, B. A.; Lopez, J.; Peteanu, L. A. Fluorescent PNA Probes as Hybridization Labels for Biological RNA. *Biochemistry* **2006**.

(9) Lansdorp, P. M.; Verwoerd, N. P.; Rijke, F. M.; Dragowska, V.; Little, M. T.; Dirks, R. W.; Raap, A. K.; H.J., T. Heterogeneity in telomere length of human chromosomes. *Hum. Mol. Genet.* **1996**, *5*, 685-691.

(10) Kalinin, S.; Peulen, T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **2012**, *9*, 1218-1225.

(11) Sapsford, K. E.; Berti, L.; Medintz, I. L. Materials for fluorescence resonance energy transfer analysis: beyond traditional donor–acceptor combinations. *Angew. Chem., Int. Ed.* **2006**, *45*, 4562 – 4588.

(12) Baerlocher, G. M.; Mak, J.; Tien, T.; Lansdorp, P. M. Telomere length measurement by fluorescence in situ hybridization and flow cytometry: Tips and pitfalls. *Cytometry* **2002**, *47*, 89-99.

(13) Martí, A. A.; Jockusch, S.; Stevens, N.; Ju, J.; Turro, N. J. Fluorescent hybridization probes for sensitive and selective DNA and RNA detection. *Acc. Chem. Res.* **2007**, *40*, 402-409.

(14) Tsourkas, A.; Behlke, M. A.; Xu, Y.; Bao, G. Spectroscopic Features of Dual Fluorescence/Luminescence Resonance Energy-Transfer Molecular Beacons. *Anal. Chem.* **2003**, *75*, 3697-3703.

(15) Tyagi, S.; Kramer, F. R. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* **1996**, *14*, 303-309.

(16) Santangelo, P.; Nitin, N.; Bao, G. Nanostructured probes for RNA detection in living cells. *Ann. Biomed. Eng.* **2006**, *34*, 39-50.

(17) Guo, J.; Ju, J.; Turro, N. J. Fluorescent hybridization probes for nucleic acid detection. *Anal. Bioanal. Chem.* **2012**, *402*, 3115-3125.

(18) Molenaar, C.; Marras, S. A.; Slats, J. C. M. Linear 2′ O-Methyl RNA probes for the visualization of RNA in living cells. *Nucleic acids Res.* **2001**.

(19) Santangelo, P. J.; Nix, B.; Tsourkas, A.; Bao, G. Dual FRET molecular beacons for mRNA detection in living cells. *Nucleic Acids Res.* **2003**, *32*.

(20) Stoop, M.; Leumann, C. J. Homo-DNA templated chemistry and its application to nucleic acid sensing. *Chem. Commun.* **2011**, *47*, 7494-7496.

(21) Marti, A. A.; Li, X.; Jockusch, S.; Stevens, N.; Li, Z.; Raveendra, B.; Kalachikov, S.; Morozova, I.; Russo, J. J.; Akins, D. L.; Ju, J.; Turro, N. J. Design and characterization of two-dye and three-dye binary fluorescent probes for mRNA detection. *Tetrahedron* **2007**, *63*, 3591-3600.

(22) Sei-Iida, Y. Real-time monitoring of in vitro transcriptional RNA synthesis using fluorescence resonance energy transfer. *Nucleic Acids Res.* **2000**, *28*, 59-59.

(23) Tsuji, A.; Sato, Y.; Hirano, M.; Suga, T.; Koshimoto, H. Development of a time-resolved fluorometric method for observing hybridization in living cells using fluorescence resonance energy transfer. *Biophys. J.* **2001**.

(24) Robertson, K. L.; Thach, D. C. LNA flow–FISH: A flow cytometry–fluorescence in situ hybridization method to detect messenger RNA using locked nucleic acid probes. *Anal. Biochem.* **2009**, *390*, 109-114.

(25) Koshkin, A. A.; Singh, S. K.; Nielsen, P. E.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C.; Wengel, J. LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron* **1998**, *54*, 3607-3630.

(26) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. A Simple γ-Backbone Modification Preorganizes Peptide Nucleic Acid into a Helical Structure. *J. Am. Chem. Soc.* **2006**, *128*, 10258-10267.

(27) Sahu, B.; Sacui, I.; Rapireddy, S.; Zanotti, K. J.; Bahal, R.; Armitage, B. A.; Ly, D. H. Synthesis and characterization of conformationally preorganized, (R)-diethylene glycol-containing gamma-peptide nucleic acids with superior hybridization properties and water solubility. *J. Org. Chem.* **2011**, *76*, 5614-5627.

(28) He, W.; Crawford, M. J.; Rapireddy, S.; Madrid, M.; Gil, R. R.; Ly, D. H.; Achim, C. The structure of a gamma-modified peptide nucleic acid duplex. *Mol. Biosyst.* **2010**, *6*, 1619-1629.

(29) Pham, H. H.; Murphy, C. T.; Sureshkumar, G.; Ly, D. H.; Opresko, P. L.; Armitage, B. A. Cooperative hybridization of gammaPNA miniprobes to a repeating sequence motif and application to telomere analysis. *Org. Biomol. Chem.* **2014**, *12*, 7345-7354.

(30) Griffith, J. D.; Comeau, L.; Rosenfield, S.; Stansel, R. M.; Bianchi, A.; Moss, H.; de Lange, T. Mammalian telomeres end in a large duplex loop. *Cell* **1999**, *97*, 503-514.

(31) Blackburn, E. H.; Szostak, J. W. The Molecular Structure of Centromeres and Telomeres. *Annu. Rev. Biochem.* **1984**, *53*, 163-194.

(32) de Lange, T. How telomeres solve the end-protection problem. *Science* **2009**, *326*, 948-952.

(33) Yu, G. L.; Bradley, J. D.; Attardi, L. D.; Blackburn, E. H. In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs. *Nature* **1990**, *344*, 126-132.

(34) von Zglinicki, T.; Martin-Ruiz, C. M. Telomeres, cell senescence and human ageing. *Sig. Transd.* **2005**, *3*, 103-114.

(35) von Zglinicki, T. Oxidative stress shortens telomeres. *Trends Biochem. Sci.* **2002**, *27*, 339-344.

(36) Chen, Q.; Fischer, A.; Reagan, J. D.; Yan, L. J.; Ames, B. N. Oxidative DNA damage and senescence of human diploid fibroblast cells. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 4337-4341.

(37) Shosuke, K.; Shinji, O. Mechanism of Telomere Shortening by Oxidative Stress. *Ann. N.Y. Acad. Sci.* **2004**, *1019*, 278-284.

(38) Armanios, M. Syndromes of telomere shortening. *Annu. Rev. Genomics. Hum. Genet.* **2009**, *10*, 45-61.

(39) Sahin, E.; Colla, S.; Liesa, M.; Moslehi, J.; Muller, F. L.; Guo, M.; Cooper, M.; Kotton, D.; Fabian, A. J.; Walkey, C.; Maser, R. S.; Tonon, G.; Foerster, F.; Xiong, R.; Wang, Y. A.; Shukla, S. A.; Jaskelioff, M.; Martin, E. S.; Heffernan, T. P.; Protopopov, A.; Ivanova, E.; Mahoney, J. E.; Kost-Alimova, M.; Perry, S. R.; Bronson, R.; Liao, R.; Mulligan, R.; Shirihai, O. S.; Chin, L.; DePinho, R. A. Telomere dysfunction induces metabolic and mitochondrial compromise. *Nature* **2011**, *470*, 359-365.

(40) Maurizot, J. C.: Circular Dichroism of Nucleic Acids: Nonclassical Conformations and Modified Oligonucleotides. In *Circular Dichroism*; Berova, N., Nakanishi, K., Woody, R. W., Eds.; Wiley-VCH, 2000; pp 719-740.

(41) Johnson, C. W.: CD of Nucleic Acids. In *Circular Dichroism*; Berova, N., Nakanishi, K., Woody, R. W., Eds.; Wiley-VCH: New York, 2000; pp 703-718.

(42) Valeur, B.; Berberan-Santos, M. N.: *Molecular Fluorescence: Principals and Applications*; 2nd ed.; Wiley, 2013.

(43) Epel, E. S.; Blackburn, E. H.; Lin, J.; Dhabhar, F. S.; Adler, N. E.; Morrow, J. D.; Cawthon, R. M. Accelerated telomere shortening in response to life stress. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17312-17315.

(44) Luke, B.; Lingner, J. TERRA: telomeric repeat-containing RNA. *EMBO J.* **2009**, *28*, 2503-2510.

(45) DeJesus-Hernandez, M.; Mackenzie, I. R.; Boeve, B. F.; Boxer, A. L.; Baker, M.; Rutherford, N. J.; Nicholson, A. M.; Finch, N. A.; Flynn, H.; Adamson, J.; Kouri, N.; Wojtas, A.; Sengdy, P.; Hsiung, G. Y.; Karydas, A.; Seeley, W. W.; Josephs, K. A.; Coppola, G.; Geschwind, D. H.; Wszolek, Z. K.; Feldman, H.; Knopman, D. S.; Petersen, R. C.; Miller, B. L.; Dickson, D. W.; Boylan, K. B.; Graff-Radford, N. R.; Rademakers, R. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **2011**, *72*, 245-256.

(46) Su, Z.; Zhang, Y.; Gendron, T. F.; Bauer, P. O.; Chew, J.; Yang, W. Y.; Fostvedt, E.; Jansen-West, K.; Belzil, V. V.; Desaro, P.; Johnston, A.; Overstreet, K.; Oh, S. Y.; Todd, P. K.; Berry, J. D.; Cudkowicz, M. E.; Boeve, B. F.; Dickson, D.; Floeter, M. K.; Traynor, B. J.; Morelli, C.; Ratti, A.; Silani, V.; Rademakers, R.; Brown, R. H.; Rothstein, J. D.; Boylan, K. B.; Petrucelli, L.; Disney, M. D. Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* **2014**, *83*, 1043-1050.

Chapter 4: Synthesis of Alkyne Modified Pyrimidines for High Density Labeling of PNA Hybridization Probes

4. Chapter Summary

Fluorophore-modified nucleic acids have been frequently utilized in various biological applications, including in the detection and quantification of DNA and RNA. Often the fluorophore is covalently attached on the end of the oligomer, however, this limits the number of fluorophores and ultimately the brightness of the probe. By attaching the fluorophore to a modified nucleobase multiple fluorescent dyes can be incorporated into the interior of a hybridization probe.

In this study we have synthesized PNA probes with multiple internal fluorescent dyes using the Huisgen-Meldal Sharpless "Click" reaction between an azide-modified fluorescent dye and alkynyl-uracil-containing PNA. As a proof of concept we have synthesized a dye-modified PNA monomer and have incorporated it into an oligomer. It was determined that the PNA:DNA duplex stability was moderately affected by the presence of the dye and an increase of fluorescence upon hybridization was observed. We have demonstrated the feasibility of this approach which suggests that several fluorescent dyes can be introduced into a PNA probe at internal positions in order to enhance the brightness of DNA/RNA-targeting PNA probes, provided self-quenching can be avoided.

4.1. Introduction

Nucleic acid hybridization probes allow scientists to visualize a nucleic acid of interest in order to detect and quantify DNA or RNA[1]. Typically, nucleic acid hybridization probes are designed as a complementary oligonucleotide with a detectable marker[2,3]. Such probes allow for monitoring in-situ gene expression[4], localization[5], and RNA splicing events[6] . An ideal hybridization probe has a high affinity for the target and is stable within an assay. Although there are several high-affinity oligonucleotides (e.g LNA[7,8]) we have chosen to focus on peptide

nucleic acids (PNA). PNAs are synthetic nucleic acids comprised of a nucleobase attached to an unnatural, neutral, peptide-like backbone that allows PNA to be nuclease-resistant and reduces charge repulsion; allowing the oligonucleotide to bind with high affinity to its nucleic acid target[9,10]. These properties make PNA a highly-attractive probe for targeting nucleic acids[11,12].

A common strategy for labeling nucleic acids often focuses on attachment of a fluorescent dye to the terminal end of a complementary sequence. However, this terminal-labeling strategy restricts the number of fluorophores per probe, which ultimately limits the brightness of the probe. Brightness is the result of the probe's ability to absorb light and release that energy back into light at a different wavelength. This property helps determine the limit of detection of the probe. For instance, a signal may not be sufficiently bright to detect if a nucleic acid is expressed in small amounts or is not strongly localized to a specific region. Although there are many commercially-available dyes in a variety of colors and physical properties, there is always a concern for the overall brightness of the label[13,14]. In order to increase the brightness of hybridization probes, scientists have incorporated modified nucleobases internally to accomplish high density multi-fluorophore probes[15].

Several tactics to increase probe brightness have led to attaching fluorophores at internal positions of oligomers. Although there have been many studies using DNA or RNA as a scaffold[15], here we will focus on using PNA as a scaffold in order to exploit its high-affinity targeting. One strategy replaces the nucleobase with a fluorophore, exploiting the systematic spacing of the backbone. Using a fluorogenic dye (e.g. thiazole orange) this strategy has been used to detect single nucleotide polymorphisms with improved single to noise; however this strategy removes base-pair recognition which limits the applications[16-18]. A second strategy attaches a fluorophore to the backbone. An example of the second strategy utilizes a Lysine γ substituted PNA as an attachment point for ligands or fluorophores[19]. By attaching a fluorene at the γ position a stemless molecular beacon was developed showing a 4-fold increase upon

hybridization with the target[20]. Further generations of this strategy covalently attached thiazole orange, a fluorogenic dye, using a long linker to allow for intercalation upon hybridization to the target for an enhanced fluorescence (42 times brighter). However, only a 30% increase of fluorescence was observed with two thiazole orange dyes. The suboptimal fluorescence enhancement was attributed to contact quenching from the long linker[20]. These studies suggest that the γ position of the PNA backbone may be too flexible to prevent contact quenching for high-density labeling.

A third approach strategically attaches the fluorophore to the nucleobase without interfering with Watson-Crick base pairs. In general, attachment to the 5-position of pyrimidines and 7-position of deazapurines (see Figure 1), allows the fluorophore to project into the major groove of the duplex, minimizing steric interferences. This strategy was pioneered in DNA and several studies have shown its versatility with both solid-phase synthesis and enzymatic labeling capabilities[21-23].

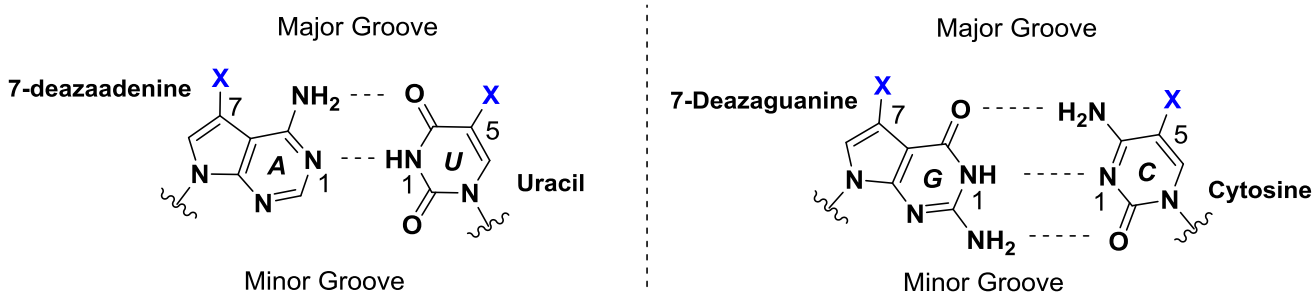
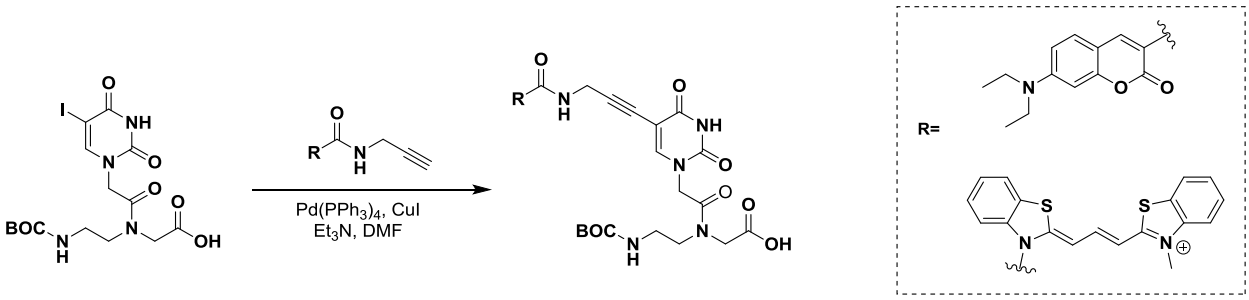


Figure 1: Watson-Crick base pairs with pyrimidines and deazapurines, Dashed lines indicate hydrogen bonds, the blue highlighted positions indicate that X substituent would project into the major grove.

One example in the PNA world uses the Sonogashira reaction to attach a terminal-alkyne small molecule to a halogenated PNA monomer[24]. This reaction scheme was modified to attach multiple fluorophores to a PNA. Previous work reported the synthesis of a dye-modified PNA monomer by using a terminal-alkyne dye and 5-iodo-PNA. The dye-modified PNA

monomer was incorporated into a small family of multi-labeled PNAs that showed minimal quenching[25].



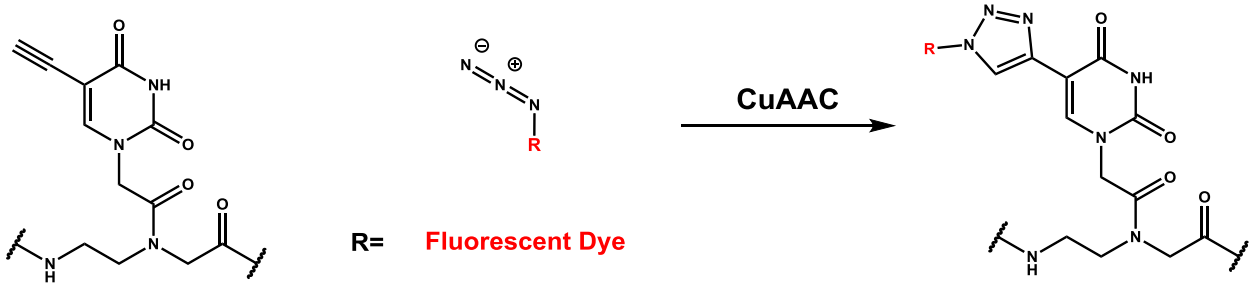
Scheme 1: Synthetic scheme using Sonogashira reaction to attach dye figure adapted from Ref 25.

However, this strategy required the synthesis of alkyne-modified dyes which limited the utility of this method. Using a different route to covalently attach the dye to the PNA monomer could permit the use of commercially-available dyes, which expand the utility of the oligomer.

The term “click chemistry” was coined by Barry Sharpless to define reactions that are high yielding and selective. Although there are many examples of ‘click chemistry,’ the Cu (I) catalyzed azide-alkyne Huisigen 1,3 dipolar cycloaddition reaction has attracted the most attention[26-28]. Research using this reaction has focused on its versatility and bio-orthogonality, establishing new labeling strategies for proteins[29] and oligonucleotides[30]. Several studies have functionalized nucleobases with alkynes and incorporated the nucleobases into DNA oligomers [31-34]. They observed minimal effects on stability and selectivity. Others explored the synthetic diversity of click chemistry by reacting the azide with the alkyne-modified oligonucleotide before, during, and/or after oligomer synthesis[23,31,35]. In addition, the rising popularity of ‘click chemistry’ as a fluorescent labeling strategy has supported an interest in the development of azide-functionalized dyes[32,35-39].

4.2. Design and Rationale

A modular design for a dye labeled PNA monomer for internal labeling could expand the potential of PNA hybridization probes. We aimed to label our hybridization probes at internal positions to optimize the number of dyes that could be attached to a single probe. To selectively recognize a complementary target, we intended to retain Watson-Crick base pairs. To reduce quenching, we chose to modify the nucleobase, over the more flexible γ position in the PNA backbone. For synthetic simplicity, we chose uracil because it did not require orthogonal protecting groups for primary amines.



Scheme 2: Modular design of alkyne PNA with azide fluorescent dye using Cu (I) catalyzed azide- alkyne Husigen 1,3 dipolar cycloaddition

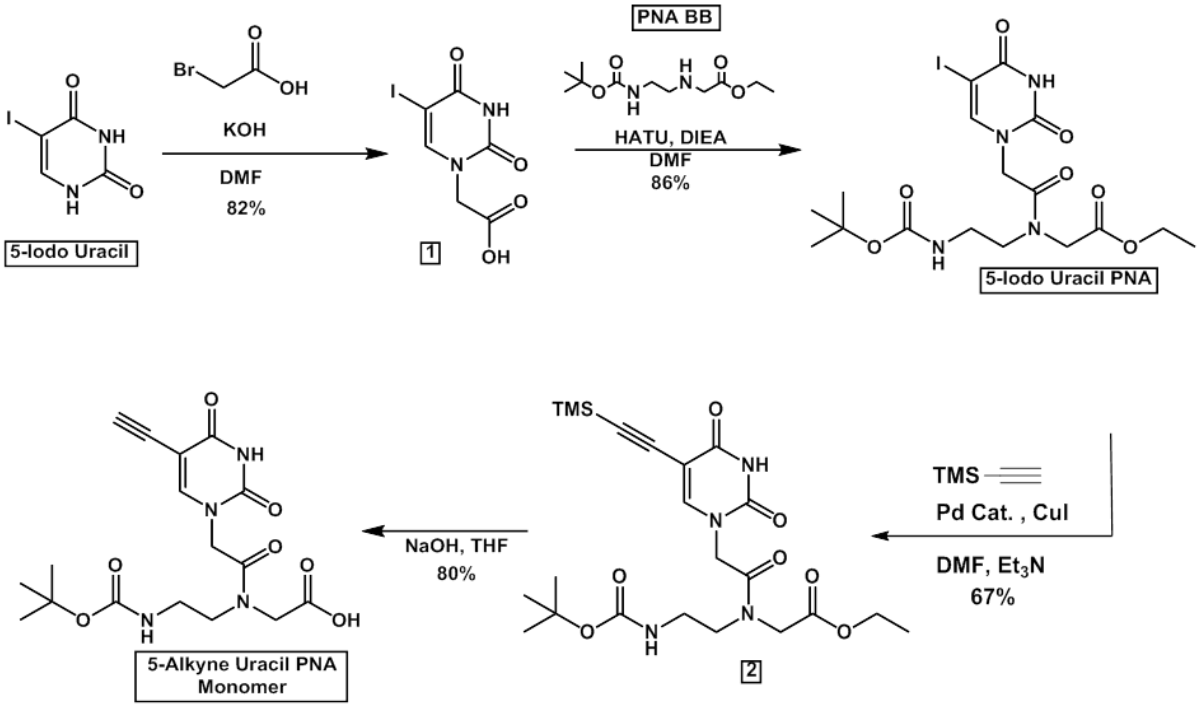
We adopted a promising approach in the DNA world where at the C-5 position of purine, a robust triazole linkage was formed using Cu(I)-mediated Huisgen cyclization. We chose the shorter alkyne linker over the longer, more flexible octadiynyl linker because we hypothesized that a short, rigid linker between the dye and nucleobase would minimize quenching between fluorophores. In addition, we wanted to take advantage of the growing popularity of click chemistry as a common fluorescent labeling method which enhances the dye possibilities (in the literature as well as commercially-available). Therefore, our first step was to synthesize the key alkyne monomer. Then, using click chemistry, various dye modified monomers could be synthesized and incorporated into various oligomers and their physical properties could be studied.

4.3. Results- Synthesis

First, we synthesized the key intermediate, alkynyl PNA monomer; then the dye was covalently attached to the monomer using Cu(I)-mediated Huisgen cyclization. To study the effect of the modification, we synthesized two mono-labeled PNA oligomers (internal and terminal).

4.3.1. Synthesis of Alkyne Uracil PNA Monomer

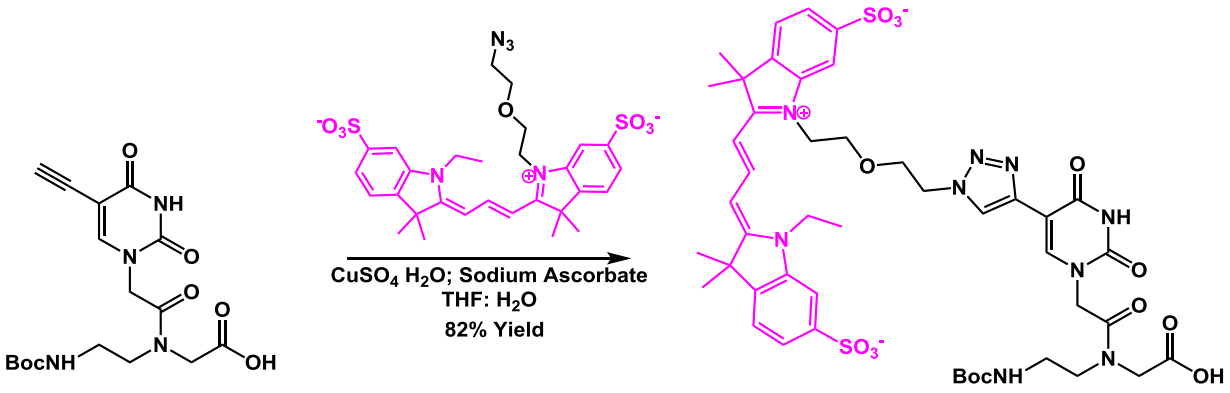
The synthesis of the Boc-protected 5-alkyne uracil PNA monomer was accomplished in four steps from the commercially available 5-iodo uracil. The 5-iodo uracil PNA was obtained according to standard procedures for Boc-protected PNA monomers[10,40]. The TMS-protected alkyne was then coupled to the monomer using the Sonogashira cross-coupling reaction. The TMS and ester protecting groups could be deprotected using basic hydrolysis conditions.



Scheme 3: Synthesis of 5-Alkyne Uracil PNA

4.3.2. Synthesis of Dye modified PNA monomers

The attachment of the dye relies on azide alkyne Huisgen cycloaddition, the PNA monomer has a terminal alkyne; therefore, the only requirement for a dye is the presence of an azide. Initially, a coumarin with an aromatic azide was used however the chemical properties were not optimal for use as a hybridization probe (see Appendix).



Scheme 4: Synthesis of Cy3- Uracil PNA Monomer

Cy3 azide (Scheme 4) was graciously gifted to the lab from Dr. B. Schmidt. This dye has two interesting modifications: two aromatic sulfonate groups, which increase the water solubility of the dye and may reduce interaction with the DNA backbone, and a short 5-atom flexible linker with a terminal azide. Using traditional azide alkyne Huisgen cycloaddition conditions, the dye modified monomer was synthesized with an 82% yield.

4.3.3. Synthesis and Characterization of Mono Dye Labeled PNAs

Two 12-mer mono-labeled PNAs using the Cy3-Uracil PNA monomer were synthesized using traditional solid-phase synthesis[41]. Central placement of the modified monomer was designed to test the location where it we expected the most perturbation in the stability of a duplex. The terminal position was synthesized as a control, to look at the effect with only one neighbor. All the PNAs were purified using HPLC and characterized using MALDI (See Chart 1).

Chart 1: Sequences of γPNA and DNA used in study and extinction coefficients at 260nm

Name	Sequence	ε at 260 ($M^{-1}cm^{-1}$)	MALDI (m/z)	Calc (m/z)
Cy3- Internal	Lys GCATCUCTCACT	103,593	3950.4	3948
Cy3-Terminal	Lys GCATCTCTCACU	103,593	3851.7	3948
Control	Lys GCATCTCTCACT	106,500	3309.9	3310

4.4. Results - Physical Characterization

The mono-labeled PNAs were characterized to study the effect of the internal modification. Factors that were of particular importance for determining the usefulness of the probe were stability, brightness, and specificity. The stability was assessed using UV thermal melting experiments with complementary DNA. The brightness was analyzed using fluorescence spectroscopy. The quantum yield of the Cy3 was determined for the single-stranded PNA in addition to the PNA/DNA duplex. Finally, specificity was analyzed by studying the effect of a mismatch in the DNA target that was assessed through UV thermal melting and fluorescence spectroscopy.

4.4.1. UV Melting Curves

PNA hybridization properties were assessed using UV-melting-curve analysis, monitoring at 260nm. PNA oliogomers characteristically bind with high affinity to complementary targets, thus yielding high melting temperature (T_M) values[42].

First, we analyzed the thermal stability of the unlabeled PNA probe that was hybridized with a complementary DNA target which yielded a T_M of 60.9°C (See Figure 2, Table 1). Then, the terminally-labeled PNA probe was hybridized to the same complementary target. The thermal stability was comparable at 61.8°C. Finally, the internal-labeled PNA probe hybridized to its complementary target showed a slight decrease in thermal stability at 56.7 °C.

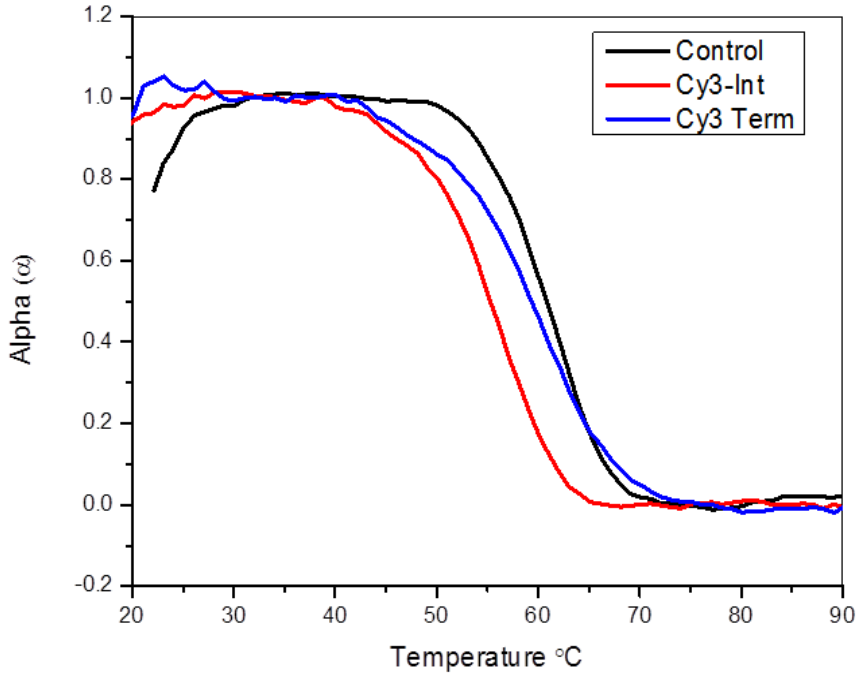


Figure 2: UV melting curves monitored at 260nm. Samples contained unlabeled (Control), internal and terminal dye-labeled PNA (1uM) hybridized with complementary DNA (1uM) target (as indicated in the legend). Samples were buffered in 10 mM sodium phosphate buffer.

Table 1: UV melting temperatures (°C) recorded for PNA/DNA duplexes

Name	Sequence	T_M
PNA Control	**Lys GCATCTCTCACT**	**60.9 ±0.2**
PNA Cy3-Int	**Lys GCA TCU CTC ACT**	**56.7 ±0.6**
PNA Cy3-Term	**Lys GCA TCT CTC ACU**	**61.8 ± 0.6**

The 4 degree difference between the unlabeled and internally labeled PNA suggests that the modification may slightly perturb the hybridization with the complementary DNA. Conversely, the terminal-labeled PNA does not exhibit the same behavior, which suggests that crowding at the internal position may be the source of the destabilization. Overall, both dye-labeled and unlabeled PNA have comparable melting temperatures and are well above physiological temperature indicating acceptable thermal stability for further studies.

4.4.2. Fluorescence Spectroscopy

In these fluorescence experiments we studied the effect of incorporating the dye into a PNA. As seen in Figure 2 and Table 2, the Cy3 monomer has the lowest fluorescence and quantum yield: 0.08. The quantum yield of Cy3 increases to 0.18 when the monomer was incorporated at an internal position in the PNA oligomer. Upon hybridization to a complementary target, the quantum yield of Cy3 doubles to 0.39.

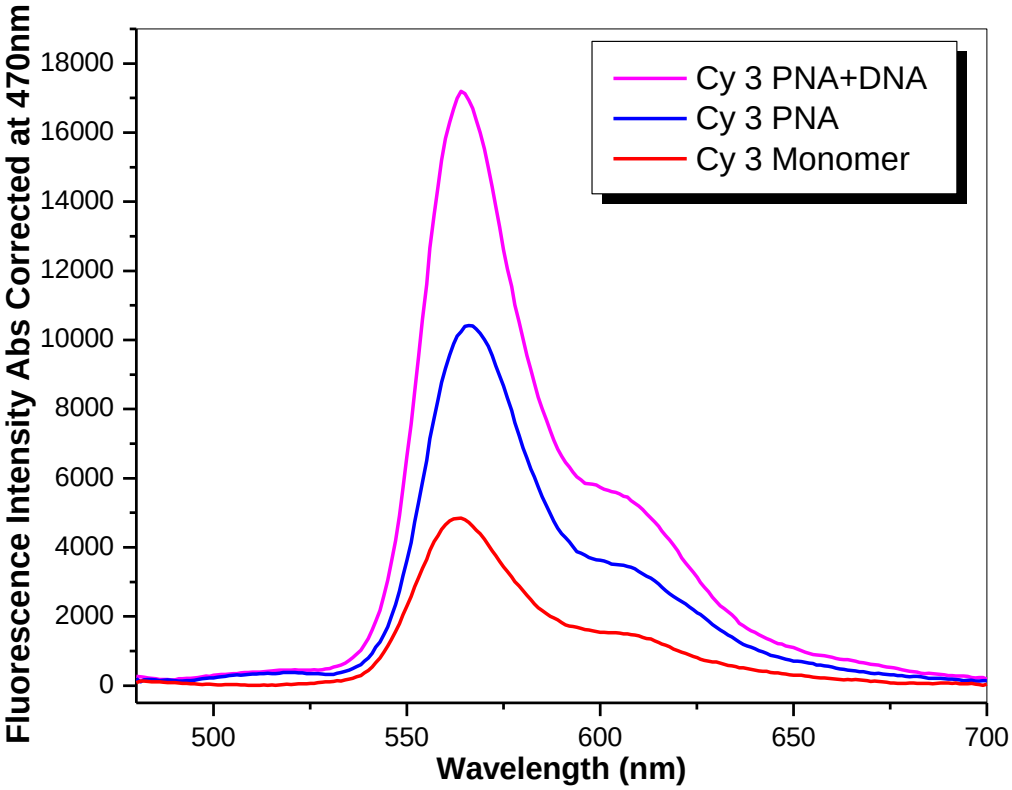


Figure 3: Fluorescence spectra of Cy3 as a progression of incorporation into a PNA probe. Samples were excited at 470nm; spectra were absorbance corrected. (A) Cy 3 monomer; (B) Internally labeled Cy3 PNA; (C) Internally labeled Cy3 PNA with Complementary DNA.

Table 2: Quantum yield of Cy3-PNA monomer, internal PNA and Internal PNA/DNA duplex

PNA	Quantum Yield
Cy 3 Monomer	**0.08**
Cy3 Int PNA	**0.19**
Cy3 Int PNA-DNA	**0.39**

An increase in fluorescence signal upon hybridization to the target is a favorable response for a hybridization probe. The increase of quantum yield of the Cy3 may reflect the physical environment surrounding the Cy3. It is known that in aqueous solution PNA forms a globular structure that is sequence-dependent[43,44]. It has also been shown that the quantum yield of Cy3 in glycerol increases to 0.51 (from 0.04 in PBS)[45]. The globular PNA environment may give the similar effect, resulting in an increase of fluorescence. Furthermore, when the Cy3 Int PNA hybridizes to its complementary DNA, the quantum yield doubles from its single-stranded PNA. The increase of quantum yield for a hybridized probe further indicates a change in the surrounding environment. Because the dye was designed to project into the major groove of the duplex, we suggest the increase of fluorescence is due to a more restrictive environment. Studies using Cy3 and a longer linker (12 atoms) do not show this increase[46], which may support our hypothesis that a shorter linker may improve the overall design of PNA hybridization probes. An increase in quantum yield improves the overall brightness of the probe which is an advantageous property; however, this fluorogenic property may be position and sequence specific.

We also looked at the internal Cy3 ability to transfer energy when hybridized to a Cy5 labeled DNA target. As seen in Figure 4, although the dye is internally positioned, once excited it can transfer energy to an acceptor yielding a 64% FRET efficiency. This energy transfer upon hybridization was anticipated but may indicate further opportunities for use in a FRET probe.

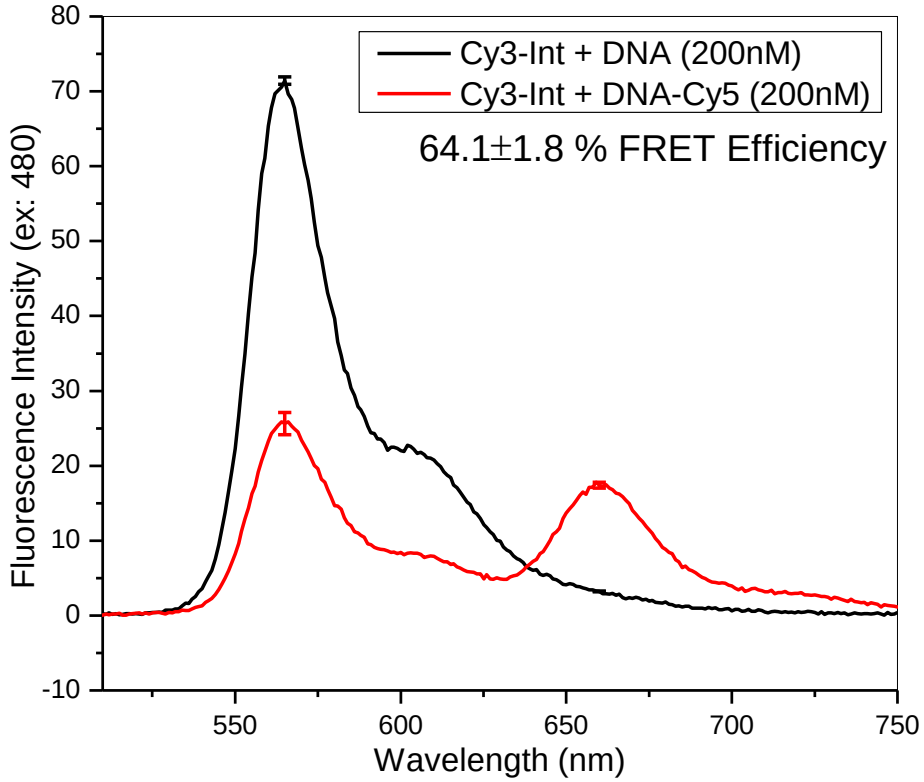


Figure 4: Fluorescence spectra of Cy 3 labeled PNA with DNA labeled with and without 5'Cy5 (Red and black respectively). Samples were buffered in 10mM sodium phosphate buffer and annealed down to room temperature; then measurement was taken at 20°C, ex: 480nm. FRET efficiency was calculated by quenching of the donor.

4.4.3. Duplex Stability and Specificity

UV melting experiments were performed to test the effect of the Cy3 modification on sequence specificity. DNA targets were designed to have a mismatch directly across from the internal position (see experimental section for DNA sequences). The ΔTm was calculated to be the difference between the perfect match and mismatch for each sequence. As detailed in Table 2, overall, each PNA was able to distinguish between a single base mismatch and a perfect match. The terminal Cy3 does not show significant differences in ΔT_M when compared to the control PNA. However, the Cy3 internal has a greater effect of stabilizing the mismatch when compared to the control and the terminally-labeled Cy3; as demonstrated with smaller ΔT_M. A higher $\Delta T_{M,}$ indicating a destabilizing effect in the presence of a mismatch, would be ideal.

Nevertheless, there is still on average a 12°C difference between perfect match and subsequent mismatches.

Table 2: UV melting temperatures (°C) recorded for PNA/DNA duplexes and ΔT_M for mismatch DNA targets.

Name	Sequence	Tm	Δ Tm		
		PM-A	MM-T	MM-G	MM-C
PNA Control	Lys GCA TCT CTC ACT	60.9 ±0.2	12.4 ±2.4	13.4 ±0.5	22.5 ±0.4
PNA Cy3-Int	Lys GCA TCU CTC ACT	56.7 ±0.6	10.7 ±1.4	11.9 ±0.4	15.9 ±0.3
PNA Cy3-Term	Lys GCA TCT CTC ACU	61.8 ± 0.6	14.0 ± 0.5	15.6 ± 2.2	17.6 ± 0.5

The fluorescence spectra of the aforementioned PNA probes with their mismatch DNA targets were collected (see Figure 5 and Table 3). Looking at the Cy3 Internal PNA, we can observe a change in fluorescence intensity that can discriminate between a perfect match and a single base mismatch when located across from the modification. There is minimal difference between the purine or pyrimidine mismatches. A single base mismatch would create a distortion in the duplex which could alter the environment of the major groove. If the Cy3 is located in the major groove then a difference in environment would result in a change in the quantum yield.

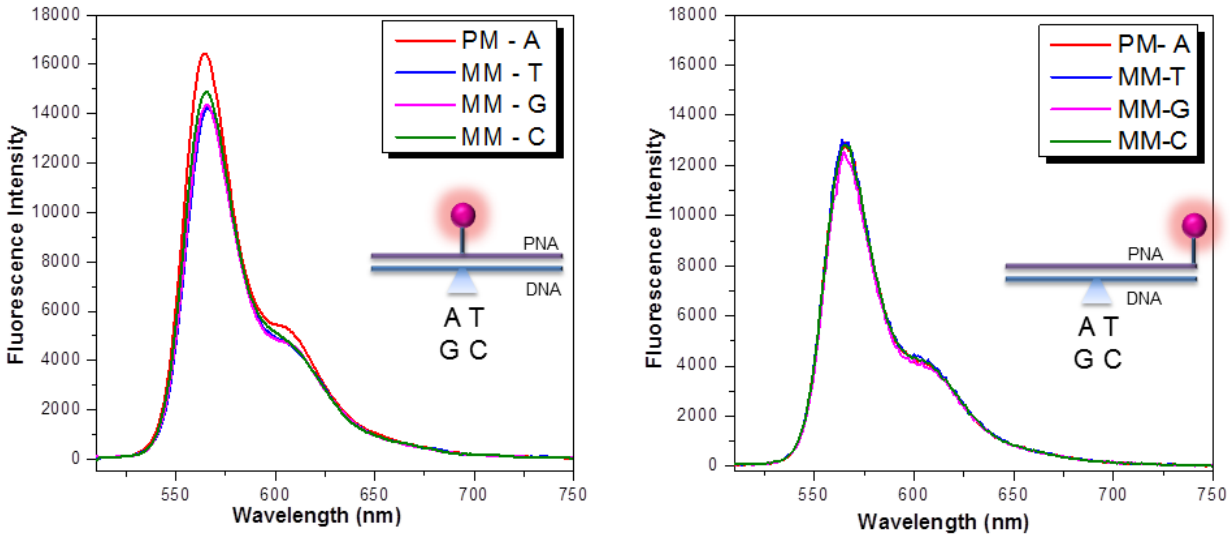


Figure 5: Fluorescence Spectra of Cy-3 Internal (Left) and Cy-3 Terminal (Right) PNA hybridized with perfect match target and subsequent mismatches (as indicated in legend). Samples were excited at 500nm and the spectrum was absorbance corrected. The schematic illustrates the location of the mismatch relative to the modification. [PNA]= 200nM; [DNA] =200nM; Buffered in 10mM sodium phosphate, pH =7.

Table 3: Quantum Yield for Internal and Terminal Cy3 PNA with DNA targets

Name	Sequence	Quantum Yield				
		Probe	PM-A	MM-T	MM-G	MM-C
PNA Cy3-Int	Lys GCA TCT CTC ACT	0.19	0.39	0.33	0.34	0.35
PNA Cy3-Term	Lys GCA TCT CTC ACT	0.15	0.26	0.27	0.26	0.27

Looking at the terminal Cy3 PNA, a single mismatch did not significantly change the fluorescence intensity. This could be attributed to the longer distance between the dye and the mismatch. In addition, the fluorescence of the terminal Cy3 was lower, relative to duplexes with Cy3- Int. This could be attributed to the less restrictive environment compared to the internal position.

Overall, subtle differences were observed upon hybridization of the Cy3-labeled probes with single-base mismatches. As seen in the UV melting experiments, the probe specificity was modestly reduced upon internal labeling with Cy3. The fluorescence experiments showed that mismatches opposite the Cy3 modification resulted in a modest decrease of fluorescence intensity.

4.5. Dual Modified PNA

A preliminary set of PNAs were synthesized to look at multi-fluorophore probes. A PNA probe with two modifications, two nucleobases apart and the unlabeled control were synthesized (for sequence information see Table 4). The PNAs were purified using HPLC and characterized using MALDI (see Table 4).

PNA hybridization properties were assessed using UV melting curve analysis monitoring at 260nm. As indicated in Table 4, the unlabeled PNA hybridized to a complementary DNA gave T_M of 62.9°C. The dual-modified PNA probe showed a modest decrease in thermal stability with a T_M of 56.6°C. A 6 degree difference in thermal stability is encouraging as a preliminary result.

This suggests that further modifications may result in only a moderate decrease of thermal stability.

Table 4: PNA sequences, MALDI, and UV melting temperatures (°C) recorded for PNA/DNA duplexes

Name	Sequence	MALDI (m/z)	Calc (m/z)	T_M
Control	**Lys GCA CTT TTC ACG**	**3309.9**	**3310**	**62.9 ± 0.4**
Cy3-2T	**Lys GCA C<u>T</u>T T<u>T</u>C ACG**	**4622.5**	**4624**	**56.6 ± 2.0**

The fluorescence spectra were collected of the dual-modified PNA with and without a complementary DNA. As illustrated in Figure 6, we observed a decrease in fluorescence when hybridized to a DNA target. A decrease in fluorescence intensity upon hybridization is contrary to the previous mono-labelled PNAs. This decrease of fluorescence may be due to contact quenching since the dyes are spaced only 2 nucleobases apart. In addition, the 5-atom linker between the dye and thiazole may allow quenching to occur.

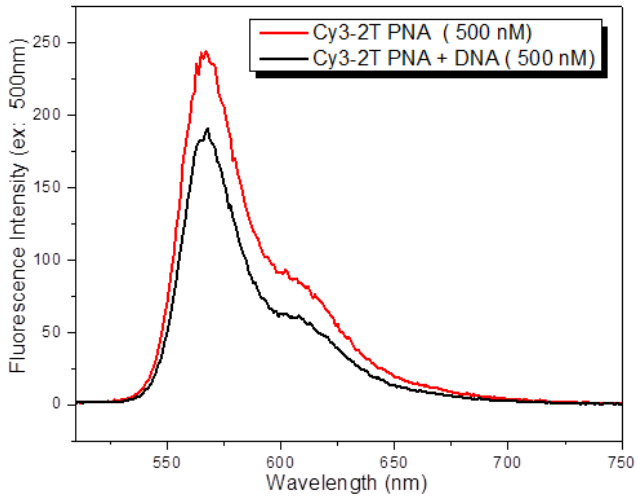


Figure 6: Fluorescence Spectra of dual labeled PNA with and without complementary DNA. Samples were excited at 500nm and buffered in 10mM sodium phosphate buffer. [PNA]=500nM; [DNA]=500nM.

Looking at the absorbance and fluorescence spectra of the mono vs dual modified PNAs, there is a dramatic difference observed (as seen in Figure 6). The absorbance spectra

shows a clear increase in absorbance which would be expected due to the increase of dyes on a single probe. However, upon excitation the two PNAs hybridized to their target show comparable fluorescence. This lack of fluorescence response observed for the dual-modified PNA suggests that significant quenching occurs. It is noted that this may not be the ideal comparison, because the two PNAs have different sequences which may have an influence on fluorescence behavior. However, we can conclude that two nucleobases apart is too close with the current dye-linker. This suggests further experiments can be done to optimize the distance and reduce quenching.

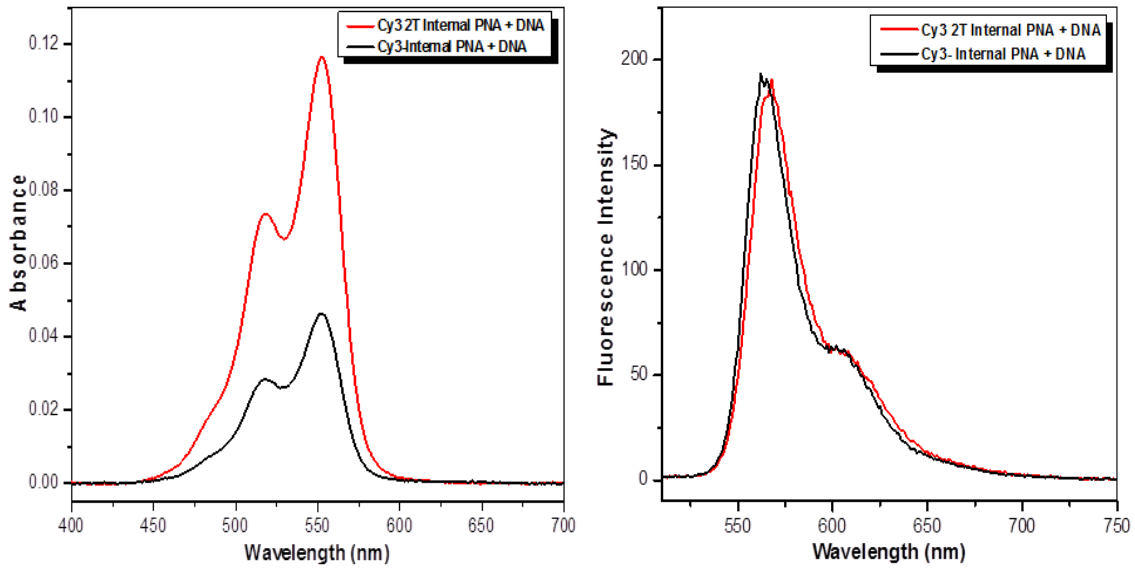


Figure 7 : Absorbance (left) and fluorescence (right) spectra of Cy3-Internal PNA (Black) and Dual Cy3 PNA (Red) hybridized with complementary DNA target. [PNA]=500nM; [DNA]=500nM. Samples were buffered in 10mM sodium phosphate and excited at 500nm.

4.6. Discussion

A common challenge among fluorescent probes is to optimize the brightness in order to increase sensitivity of the assay. Increased brightness can be achieved by precisely controlling the position of dyes within the hybridization probe. Brightness depends on optimizing two

variables: the extinction coefficient and the quantum yield of the probe. By increasing the local concentration of dyes the extinction coefficient can be improved. However, studies have shown that increasing the dye density may have deleterious effects on quantum yield or the stability of the probe to hybridize to the target[32,46]. Therefore by precisely controlling the position of the dyes we can start to study the effect of position, spacing, dye, and linker in order to optimize brightness in a hybridization probe.

Here we have demonstrated preliminary studies of the use of click chemistry to develop a modular dye-labeling system for PNA hybridization probes. We have synthesized an alkyne-modified monomer and functionalized the monomer with two different dyes (Cy3 and Coumarin). Then, we incorporated the dye-labeled monomer into several PNA oligomers and analyzed the thermal stability, target specificity, and fluorescence properties. Overall, the dye-modified PNAs showed moderate decrease in the duplex stability, suggesting the dye/linker having minimal effect. Previously, a study of Cy3 labeling DNA used an amino-modified uridine analogue with flexible linkers that ranged from 12-18 atoms long. Looking at the probe stability, it was suggested that the long linker had minimal effect on the T_M, however it was suggested that the negatively charged Cy3 had a slight destabilizing effect (approximately 2°C per Cy3)[46]. This may explain the decrease of the duplex stability observed with the internally modified mono and dual PNA probes in the current study. One approach to increase the duplex stability of PNA probes incorporates a chiral center at the gamma position of the PNA backbone. It has been shown that incorporating a γ -modification can increase the duplex stability[47]. .

The hybridization to the target significantly affects the fluorescence behavior of the probe. Previous work with DNA hybridization probes have shown that the fluorescence intensity can vary depending on the label density, fluorophore charge, and linker length[46,48,49]. Preliminary studies suggest this may also be the case for PNA hybridization probes. The fluorescence enhancement for the mono internal-labeled probe is promising; however, the fluorescence

quenching for the dual internal-labeled probes leaves room for improvement. Different linker lengths as well as dye density should be explored.

4.7. Future Implications

Future work should look into incorporating a hybrid of γ-modified PNAs with dye-labeled PNAs. As illustrated in figure 8, where (A) represents the current dye-labeled PNA oligomers, (B) represents a hybrid with γ PNA for G, C, A, T but incorporates the dye-labeled uracil without any γ modification, (C) represents a fully γ-modified PNA oligomer. Incorporating γ-modified monomers into the sequence increases the thermal stability, thus allowing for shorter probes.

(A) Lys- GCAUCTCTCUCACT
(B) Lys- G*C*A*U C*T*C*T*C*U C*A*C*T*
(C) Lys- G*C*A*U*C*T*C*T*C*U*C*A*C*T*

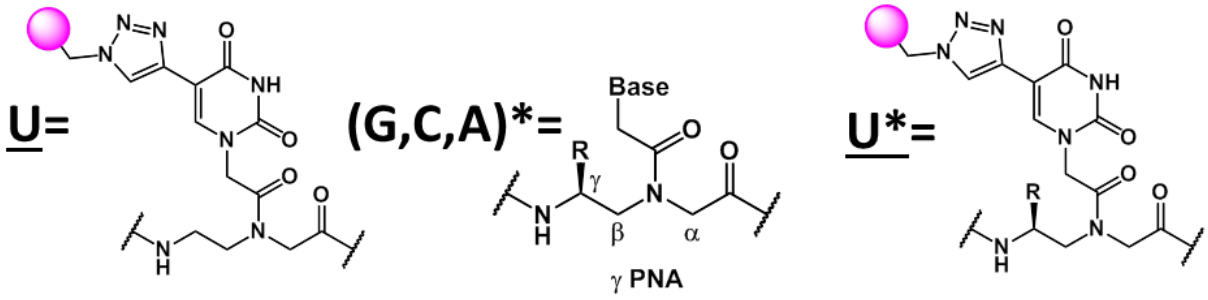


Figure 8: Future generations of dye labeled PNAs with incorporating γ-modified PNAs (top); structures of PNA monomers used in future generations. U= current dye labeled PNA monomer; (G,C,A)*= γ -modified PNA monomers; U*= Future dye labeled γ -modified PNA monomer

The overall challenge in the field of hybridization probes is not only to develop bright dyes but also develop spectrally orthogonal probes for multi-color imaging. Although we have looked at using coumarin (appendix), future studies should look at different fluorescent dyes, which allows multi-channel imaging, in addition to exploiting other fluorescence phenomena (i.e. FRET). Some dyes are not compatible with BOC/Cbz protecting group chemistry used in traditional solid phase synthesis. Therefore in order to use milder cleavage conditions an Fmoc/Boc PNA monomer should be synthesized. Another option would be to look into attaching

the dye after cleavage. Future generations would focus on expanding to other nucleobases, by synthesizing an alkyne-Cytosine or a 7-deazapurine alkyne derivative.

In general, this strategy could be used to increase the sensitivity of fluorescence based nucleic acid detection assays. It has been shown that multi-labeled DNA probes have increased brightness in FISH assays[50]. PNA hybridization probes have been used to detect and quantify telomeric repeats[12]. The commercially available hybridization probe is an 18-mer PNA with a terminal fluorescent dye (Lys-$(C_3TA_2)_3$). We attempted to make an 18-mer tri-modified PNA (Lys-$(C_3\underline{\mathbf{U}}A_2)_3$) but the synthesis was unsuccessful.

In addition, future generations of this dye-labeled monomer could also be used to improve the limit of detection of the FRET miniprobes (Chapter3). Where short terminal-labeled 9-mer γ PNA probes were designed to hybridize in tandem, allowing a FRET signal to indicate hybridization to the target (See Figure 9A). FRET efficiencies up to 60% were achieved. FRET is distance-dependent, by reducing the distance between Donor and Acceptor by incorporating dyes at both internal and terminal positions the FRET efficacy can be improved. In addition more dyes per probe will result in a brighter signal.

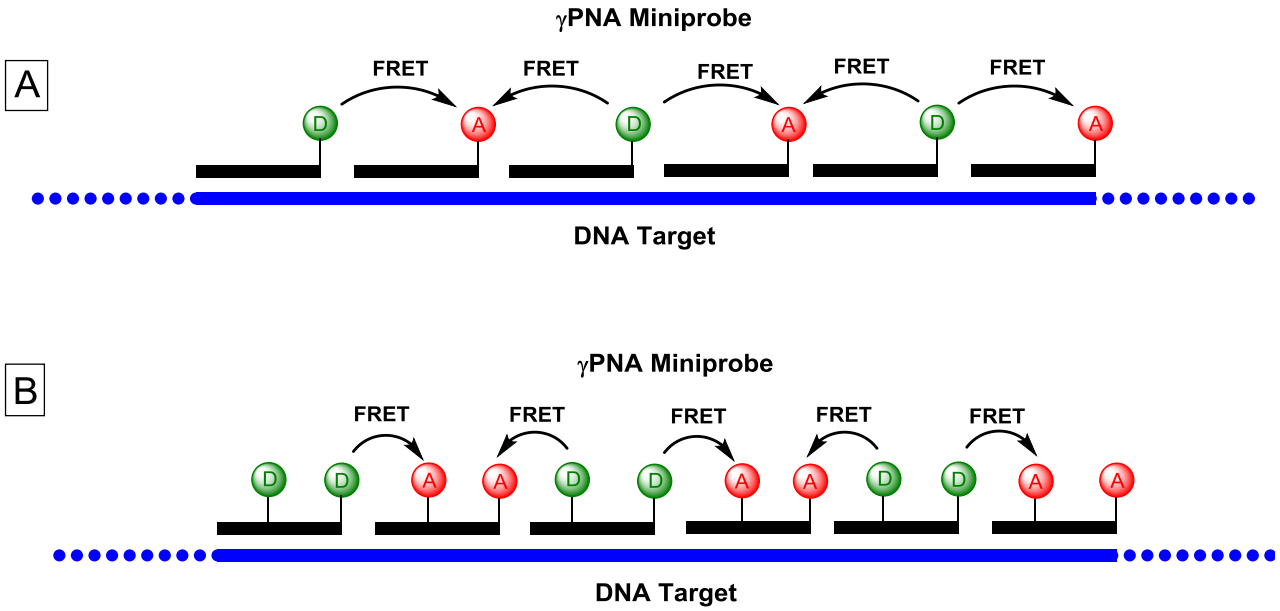


Figure 9: Design of FRET miniprobes on a repeating DNA target (A) Current design as described in Chapter 2 (B) Future design incorporating internal and terminal dye labeling

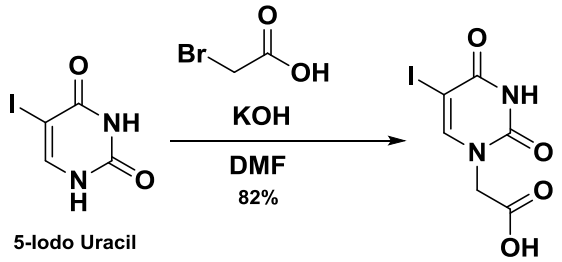
4.8. Conclusion

Nucleic acid hybridization probes are often labeled by covalently attaching a fluorophore at the terminal end of oligomer, thereby limiting the number of fluorophores per probe. However, if a fluorophore was attached to the nucleobase without disturbing Watson Crick base pairing, multiple dyes could be incorporated into one probe. Attachment of fluorophores on DNA using an alkyne-modified nucleobase has been reported by several groups. This strategy has not yet been applied to PNA, which has a higher affinity to DNA. Here we have designed and synthesized a novel alkyne modified PNA monomer, as well as new dye-labeled Cy3 PNA monomers, for internal labeling of PNA. We have demonstrated the feasibility of this approach which suggests that several fluorescent dyes can be introduced into a PNA probe at internal positions in order to enhance the brightness of DNA/RNA-targeting PNA probes.

4.9. Experimental

4.9.1. Synthesis of Alkyne PNA monomer

Synthesis of 5-iodo uracil acetic acid

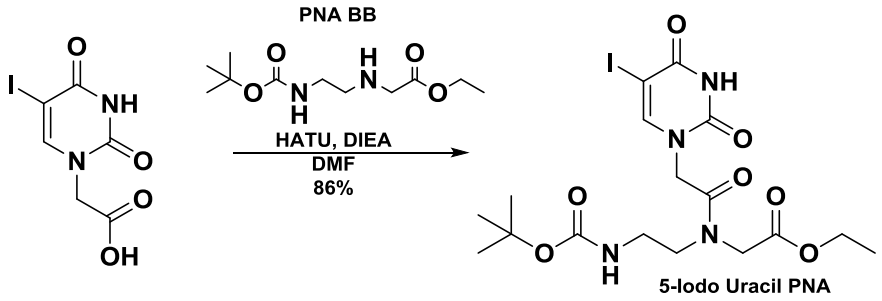


5-iodo-uracil (2.37g, 0.01 mmol, 1 eqiv) was dissolved in KOH/H_2O solution (2.13g,0.038 mmol, 3.8 equiv in 10mL of H_2O) and heated to 40^{o}C. Then bromo-acetic acid was dissolved in 10 mL of H_2O and added dropwise. The reaction mixture was stirred for 30 min, then cooled overnight (placed in refrigerator to cool). The pH was adjusted to 5 and a precipate formed and was

filtered. The filtrate was adjusted to a pH of 2 and more precipitate was formed and filtered. Fractions were dried on vacuum. 2.250g (76% Yield)

^{1}H (DMSO) δ : 11.70 (s,1H), 8.17 (s, 1H), 4.35 (s,2H)

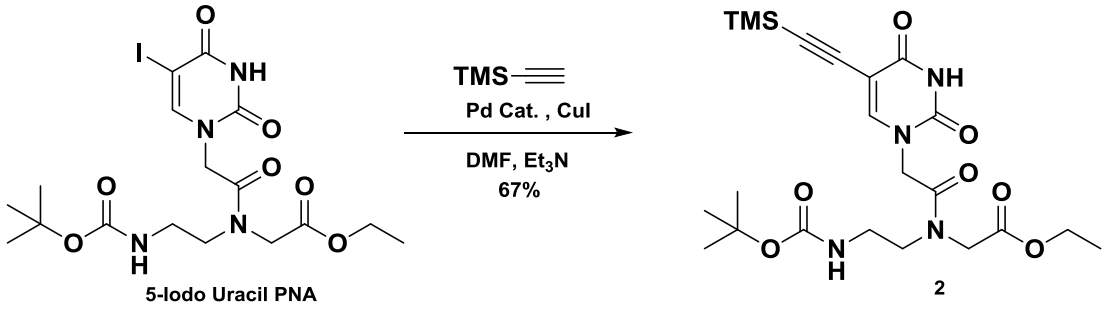
Synthesis of 5-Iodo Uracil PNA



5-iodo uracil acetic acid (1.5 eq, 2.78 g, 9 mmol) was prepared in DIPEA (1.8 eq, 1.908 mL, 11 mmol) and 2,6-lutidine (1.5 eq, 1.058 mL, 9 mmol). HATU (1.4 eq, 3.293, 8.5 mmol) in DMF (10 mL) was added to the initial solution. After stirring for 15 minutes, a solution of PNA backbone (1 eq, 1.5 g, 6 mmol) and DIPEA (1.3 eq, 1.378 mL, 7.9 mmol) was added. After stirring at room temperature for a period of two days, the mixture was quenched with water (~400 mL). After stirring for 20 minutes, the reaction mixture was extracted with ethyl acetate and washed with brine using a series of manual separatory-funnel extractions. The organic layers were combined, dried using Na_2SO_4, and the solvent was removed.

^{1}H NMR ($CDCl_3$, 300 MHz) δ: A mixture of 2 rotamers are present in the NMR spectra. 1.29 (t, 3H, J = 7.18 Hz), 1.45/1.47 (s, 9H), 3.34 (q, 2H, J = 5.7 Hz), 3.54 (t, 2H, J = 5.7 Hz), 4.07/4.19 (s, 2H), 4.23 (q, 2H, J = 7.16), 4.49/4.64 (s, 2H), 4.98/5.56 (t, 1H, J = 6.18), 7.60/7.66 (s, 1H), 8.99/9.02 (s, 1H)

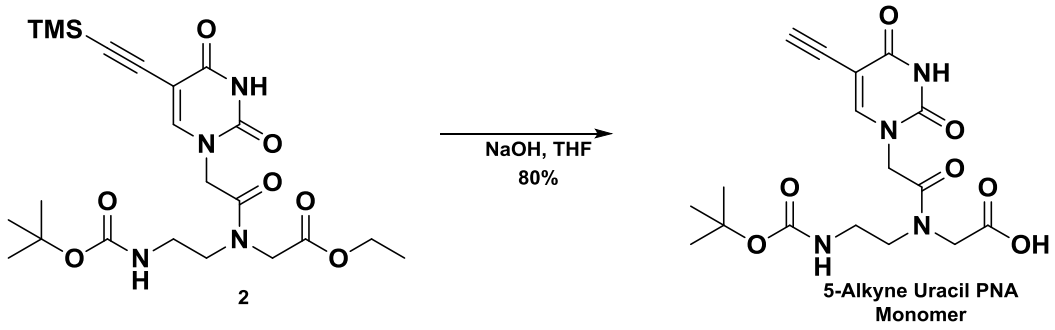
Synthesis of 5-TMS-Alkyne-Uracil- PNA (2)



5-iodo uracil PNA was suspended in THF degassed with Ar, then $Pd(PPh_3)_4$ and CuI were added and the reaction mixture was degassed again with Ar. TMS acetylene and triethyamine were added and the reaction mixture was heated at 40^{o}C overnight. The resulting reaction mixture was filtered through celite and rinsed with THF, the mother liquid was condensed and purified by column chromatography (EtOAc:Hex).

1H NMR (CDCl3, 300 MHz) δ: 0.24 (s, 9H), 1.31 (t, 3H), 1.46/1.48 (s, 9H), 3.34 (q, 2H), 3.54 (t, 2H), 4.07/4.19 (s, 2H), 4.22 (q, 2H), 4.47/4.62 (s, 2H), 4.95/5.58 (t, 1H), 7.43/7.50 (s, 1H), 8.63 (s, 1H).

Synthesis of 5-Alkyne-Uracil- PNA



A sample of the 5-alkyne uracil PNA (0.2196 g, 0.52 mmol, 1.0 eq.) was dissolved in approximately 30 mL THF. To this solution, 2M NaOH (16 mL) was slowly added over 15 mins. The reaction mixture was compared to the reactant through TLC (10% MeOH/DCM). Approximately 75 mL of water was added, and the resulting mixture was extracted with ethyl

acetate. The aqueous layer, with an initial pH of 8, was collected and titrated until it reached a pH level of 4. The mixture was extracted with ethyl acetate (3x75 mL), and the combined organic layers was dried using sodium sulfate. The solvent was removed and the crude product was dried on high vacuum.

^{1}H NMR (CD3OD 300 MHz) d: 1.44/1.45 (s, 9H), 3.07/3.20 (t, 2H), 3.49 (q, 2H), 3.55 (s, 1H), 4.12/4.27 (s, 2H), 4.63/4.81 (s, 2H), 6.48/6.57 (t, 1H), 7.80/7.83 (s, 1H), 9.33 (s, 1H).

Synthesis of Cy3-Uracil-PNA-OH

Alkyne-Uracil-PNA-OH and Cy3 azide were dissolved in 500μL of methanol. Then, in a separate vial, copper sulfate and sodium ascorbate were both dissolved in 500μL of water and vortexed until the solution turned orange. The resulting solution was then added to the dye-monomer mixture. The vial was stirred at room temperature and the reaction was monitored via TLC. After 2 hours the reaction was determined to be finished and the reaction mixture was condensed. The crude was filtered through a small plug of silica using 70% MeOH/DCM as solvent. The resulting fractions were tested using ESI and NMR (250mg, 83% yield). The extinction coefficient for the monomer at 260nm was determined to be 8,054 $M^{-1}cm^{-1}$.

ESI negative mode: Calc 1022 M/z; Found 1022.5

4.9.2. PNA Synthesis

PNA oligomers were synthesized via solid phase synthesis using Boc/Cbz protected PNA monomers. Oligomers were purified using reverse-phase high performance liquid chromatography (HPLC), and characterized by matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry (Applied biosystems Voyager DE sSTR) using α-cyano-4-hydroxycinnamic acid as the matrix dissolved in 0.1% TFA in 50:50 Acetonitrile/H_2O.

4.9.3. UV Melting Curves

Samples were prepared in 10mM sodium phosphate buffer. Thermal UV melting experiments were performed by heating the samples up to 90°C and slowly cooling (1°C per min) to 15°C, then the samples were allowed to equilibrate for 5 minutes at 15°C then slowly heated to 90°C (rate= 1°C / min) and the absorbance at 260nm was collected. The resulting data was normalized and the first derivative was plotted. The maximum of the first derivative was recorded as the T_M. [PNA]= 1μM; [DNA]=1μM

4.9.4. Fluorescence Spectroscopy

Fluorescence emission spectra were collected on a Cary Eclipse Fluorescence spectrophotometer at 25° C by exciting the samples at 470nm. UV absorbance spectra from 750-240nm were also collected for each sample. [PNA]=200nM; [DNA] =200nM

Chart 2: Sequences of PNA and DNA used in mismatch experiments. Bold U indicates the position of the internal dye-labeled monomer, underline indicates mismatch

Name	Sequence
PNA	Lys GCA TC**U** CTC ACT
Perfect Match DNA	5' – CGT AGA GAG TGA – 3'
Mismatch T DNA	5' – CGT AG<u>T</u> GAG TGA – 3'
Mismatch C DNA	5' – CGT AG<u>C</u> GAG TGA – 3'
Mismatch G DNA	5' – CGT AG<u>G</u> GAG TGA – 3'

4.9.5. Quantum Yield

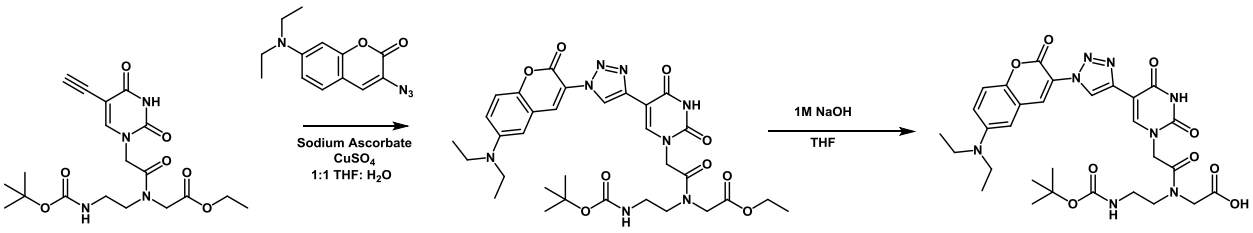
Fluorescence quantum yields for Cy3 derivatives were determined using the relative method. The relative method relies on the idea that at the same excitation conditions, and using a known standard, the unknown quantum yield can be determined quantitatively. The standards, Fluorescein and Lucifer Yellow, were chosen due to their spectral overlap with Cy3 analogs (Φfluorescein: 0.79 in 0.1 N NaOH; Φlucifer yellow : 0.21 in Nanopure H_2O at 22°C). A serial dilution of both standards were made and the absorbance and fluorescence (ex: 470) were measured. Fluorescence and absorbance measurements were performed in 10mM

sodium phosphate buffer and all standard and sample solutions had a maximum absorbance that did not exceed 0.1. All measurements were performed in triplicate and all error was below 10%.

In Origin, each fluorescence spectrum was integrated. Then, the integrated fluorescence intensity vs absorbance at 470nm was plotted and a trendline calculated with y intercept=0. The slope of this line is the gradient (grad). Using the gradient of the unknown ($grad_x$) as well as the standard ($grad_{std}$) in addition to the refractive index of their respective solvents (n), the unknown quantum yield (Φ_x) was determined using the equation below.

$$\Phi_x = \Phi_{std} * (grad_x / (grad_{std}) * (n_x / n_{std})^2$$

4.10. Chapter 4 Appendix



Scheme S1: Synthesis of Coumarin Uracil PNA monomer

Synthesis of Coumarin Uracil PNA monomer

Coumarin was synthesis according to Silvakumar et.al[36].

Alkyne-PNA-Ester (100mg) and Coumarin- Azide (62.2mg) were dissolved in 500µL THF, a solution of $CuSO_4$ and sodium ascorbate in 500µL H_2O was added and the reaction mixture was allowed to stir overnight and was monitored via TLC. The solution was condensed, ethyl acetate was added and the organic layers were washed with H_2O and brine dried and condensed. The crude was further purified by column chromatography 5% methanol in DCM. 125mg 77 % yield

Coumarin- PNA Ester (76mg) was dissolved in THF (2mL) and 1M NaOH (2mL) was added at 0°C. The reaction mixture was monitored via TLC until complete. Ethyl acetate was added and the aqueous was collected and neutralized to pH of 4. Then extracted with ethyl acetate, the organic layer was collected and dried. 68mg; 94% yield.

^{1}H NMR (MeOD, 300 MHz) δ: A mixture of 2 rotamers are present in the NMR spectra. 1.24 (t, 6H), 1.46 (s, 9H), 3.54 (q, 4H), 3.54 (t, 2H, J = 5.7 Hz), 4.14/4.30 (s, 2H), 4.79/4.96 (s, 2H), 6.64(s,1H), 6.84 (d, 1H), 7.54(d, 1H), 8.32 (t, 1H), 8.37 (s, 1H) 8.88(s, 1H)

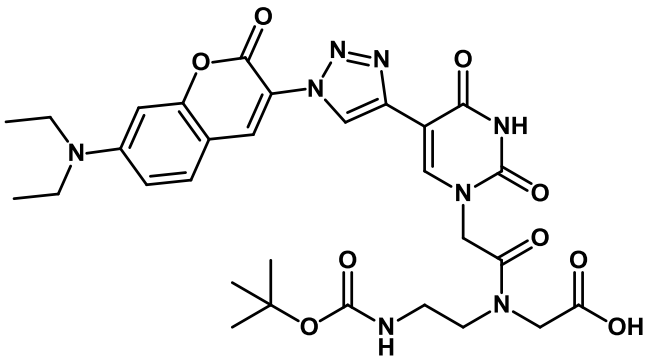


Figure S1: Structure of Coumarin T PNA

Table S1: PNA sequences, the three sequences should show the effect of dye spacing.

Table 1: PNA Sequences ** = MALDI; [a] = ESI, M^{+2} Species; [b] = ESI, M^{+3}Species*				
PNA	Sequence	Spacing	Mass (m/z)	Calc (m/z)
PNA-00	Lys-CGATTTTCGA	NA	2846.9*	2847.9
PNA-47	Lys-CGATTTTCGA	2	1128.4[b] /1692.5[a]	1127.8 /1691.7
PNA-57	Lys-CGATTTTCGA	1	1128.6[b] /1692.9[a]	1127.8 /1691.7
PNA-67	Lys-CGATTTTCGA	0	1128.5[b] /1692.9[a]	1127.8 /1691.7

UV Melting Experiments

Samples were prepared in a 1:1 PNA:DNA ratio (2μM PNA: 2μM DNA) and prepared in 10 mM tris, 5 mM $MgCl_2$ buffer. The fluorescence and absorbance of each sample was measured before and after the addition of DNA to observe any spectral change. Then each sample was heated at 93°C and cooled to 15 °C at a rate of 1 degree per minute to anneal the DNA/ PNA strands. Then the samples were slowly heated at a rate of 1 degree per minute to 95 °C. Each sample was normalized from 0 to 1, and temperature in which 50 % of the complex is single stranded was determined.

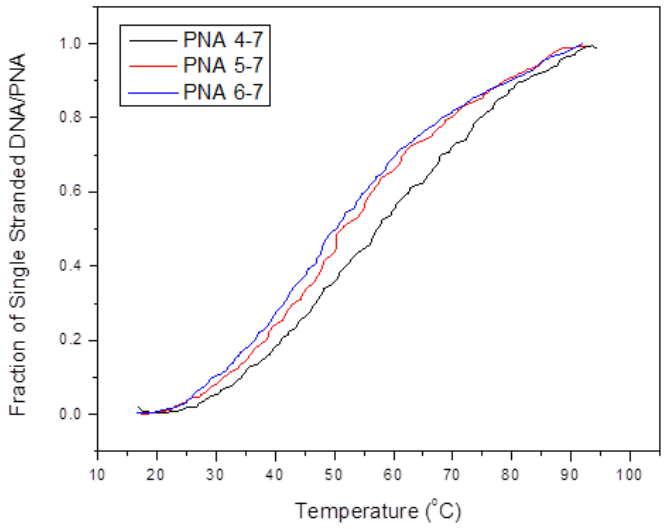


Figure S2: UV melting curves monitored at 260nm. Samples contained dye-labeled PNA (2μM) hybridized with complementary DNA (2μM) target (as indicated in the legend). Buffered in 10 mM tris, 5 mM $MgCl_2$ buffer

Table S2: PNA sequences, Dye spacing and UV melting Temperature (T_M)

PNA	Sequence	Spacing	T_m 10-mer DNA
PNA-47	Lys-CGATTTTCGA	2	56
PNA-57	Lys-CGATTTTCGA	1	51
PNA-67	Lys-CGATTTTCGA	0	50

It was observed that the spacing may affect the stability of the complex. PNA 47, containing the largest gap, shows the highest T_M of 56 °C and is most stable when hybridized to the DNA target. PNA-57 and PNA-67 show similar thermal stability at 51°C and 50°C respectively. However all modified PNAs show a stable duplex above 37°C.

Fluorescence Properties

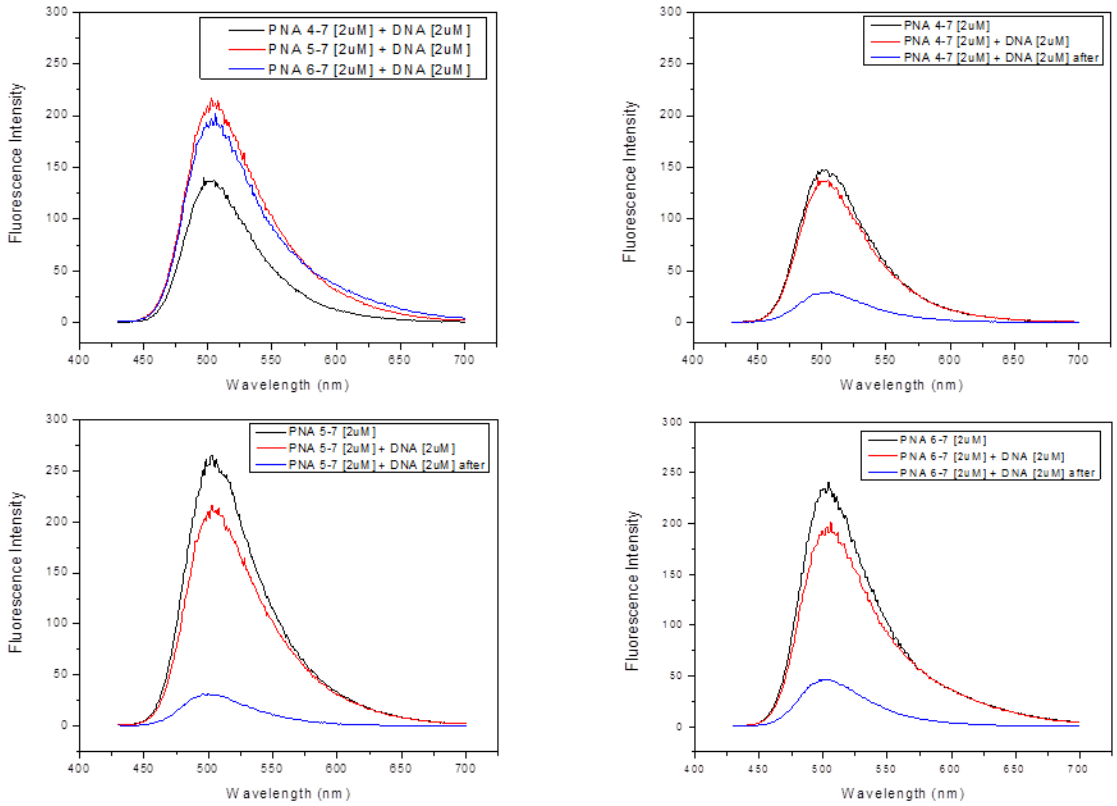


Figure S3: Upper Left: Comparison of Dye labeled PNA hybridized to complementary DNA; Upper Right: PNA 4-7; Lower Left: PNA 57; Lower Right : PNA All samples were excited at 420nm.

The fluorescence spectra as seen in Figure S3 above shows an interesting behavior. The upper right spectra shows the comparison of all PNA/DNA duplexes. It is observed that PNA 57 and PNA 67 are the brightest with PNA 47 having the lowest fluorescence intensity. The remaining spectra are a comparison of each PNA in different aspects of the melting experiment. The highest intensity is the black line which is the PNA free in solution. Upon addition of the complementary DNA sequence the fluorescence is quenched. This may be due to the change in environment. After the melting experiment was performed a large decrease in fluorescence intensity was observed for all PNAs. This led me to look at the absorbance spectra for answers.

Absorbance Spectra of Coumarin PNA

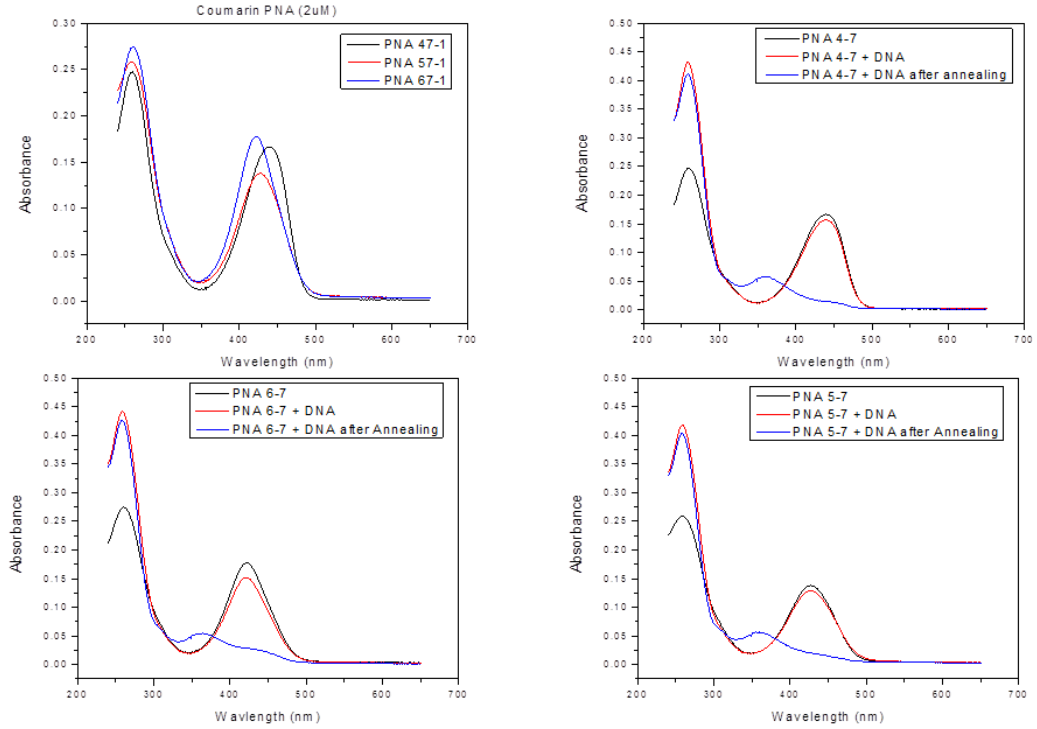


Figure S4: Upper Left: ALL PNA+ DNA Upper Right: PNA 4-7; Lower Left: PNA 57; Lower Right : PNA

The absorbance spectra above also show interesting behavior. The upper left spectra overlays all the PNA/DNA hetero-duplexes. It is interesting to note that not all the PNA's have the same maximum. The remaining spectra show a comparison of each PNA during the thermal melting experiment. It is observed upon addition of DNA a slight decrease in absorbance, which would explain the decrease in the fluorescence spectra after DNA. Interestingly, after annealing the absorbance spectra were different regardless of dye position. There was a clear hypsochromic shift, which could be interpreted as degradation of the coumarin. This suggests that the coumarin is not thermally stable.

Thermal Stability Test

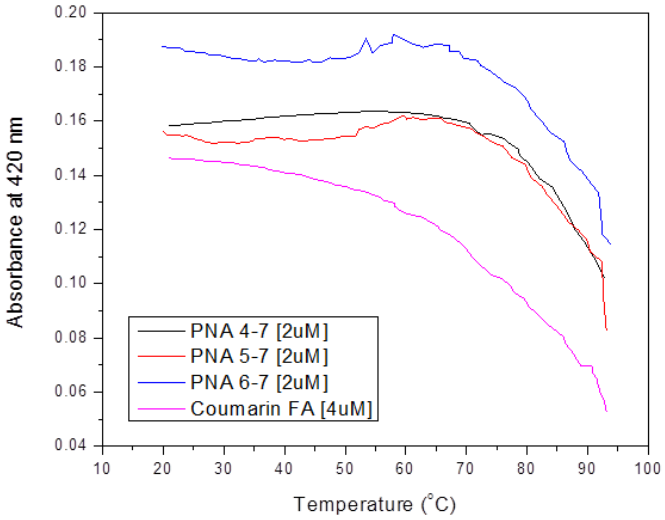


Figure S5: UV melting curves monitored at 420 nm.

In order to test the thermal stability of the coumarin PNA, an additional thermal melting experiment was performed monitoring at 420nm the λ max of coumarin. As seen in Figure S5 samples are heated from 15°C to 93°C, as the temperature increases the absorbance at 420nm starts to decrease. This suggests that the there is a change in the electronic structure of coumarin. This also indicates that this coumarin is not an ideal dye for these experiments, since it is hydrolytically unstable.

Although the coumarin proved to be defective, these experiments showed the utility of the monomer with different dyes and linkers.

References

(1) Itzkovitz, S.; van Oudenaarden, A. Validating transcripts with probes and imaging technology. *Nat. Methods* **2011**, *8*, S12-19.

(2) Guo, J.; Ju, J.; Turro, N. J. Fluorescent hybridization probes for nucleic acid detection. *Anal. Bioanal. Chem.* **2012**, *402*, 3115-3125.

(3) Martí, A. A.; Jockusch, S.; Stevens, N.; Ju, J.; Turro, N. J. Fluorescent hybridization probes for sensitive and selective DNA and RNA detection. *Acc. Chem. Res.* **2007**, *40*, 402-409.

(4) Tsuji, A.; Koshimoto, H.; Sato, Y.; Hirano, M.; Sei-Iida, Y. Direct observation of specific messenger RNA in a single living cell under a fluorescence microscope. *Biophys. J.* **2000**.

(5) Raj, A.; van den Bogaard, P.; Rifkin, S. A.; van Oudenaarden, A.; Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **2008**, *5*, 877-879.

(6) Robertson, K. L.; Yu, L.; Armitage, B. A.; Lopez, J.; Peteanu, L. A. Fluorescent PNA Probes as Hybridization Labels for Biological RNA. *Biochemistry* **2006**.

(7) Robertson, K. L.; Thach, D. C. LNA flow–FISH: A flow cytometry–fluorescence in situ hybridization method to detect messenger RNA using locked nucleic acid probes. *Anal. Biochem.* **2009**, *390*, 109-114.

(8) Koshkin, A. A.; Singh, S. K.; Nielsen, P. E.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C.; Wengel, J. LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron* **1998**, *54*, 3607-3630.

(9) Nielsen, P. E.: *Peptide Nucleic Acids: Protocols and Applications*; Horizon Bioscience, 2004.

(10) Hyrup, B.; Nielsen, P. E. Peptide nucleic acids (PNA): synthesis, properties and potential applications. *Bioorg. Med. Chem.* **1996**.

(11) Sietz, O. Solid-Phase Synthesis of Doubly Labeled Nucleic Acids as Probes for the real time detection of hybridization *Chem. Commun.* **2000**, *39*.

(12) Lansdorp, P. M.; Verwoerd, N. P.; Rijke, F. M.; Dragowska, V.; Little, M. T.; Dirks, R. W.; Raap, A. K.; H.J., T. Heterogeneity in telomere length of human chromosomes. *Hum. Mol. Genet.* **1996**, *5*, 685-691.

(13) Lavis, L. D.; Raines, R. T. Bright Ideas for Chemical Biology. *ACS Chem. Biol.* **2008**, *3*, 142-155.

(14) Lavis, L. D.; Raines, R. T. Bright building blocks for chemical biology. *ACS Chem. Biol.* **2014**, *9*, 855-866.

(15) Teo, Y. N.; Kool, E. T. DNA-multichromophore systems. *Chem. Rev.* **2012**, *112*, 4221-4245.

(16) Köhler, O.; Jarikote, D. V.; Seitz, O. Forced intercalation probes (FIT Probes): thiazole orange as a fluorescent base in peptide nucleic acids for homogeneous single-nucleotide-polymorphism detection. *Chembiochem* **2004**, *6*, 69-77.

(17) Kummer, S.; Knoll, A.; Socher, E.; Bethge, L.; Herrmann, A.; Seitz, O. Fluorescence Imaging of Influenza H1N1 mRNA in Living Infected Cells Using Single-Chromophore FIT-PNA. *Angew. Chem. Int. Ed.* **2011**, *50*, 1931-1934.

(18) Knoll, A.; Kummer, S.; Hövelmann, F.; Herrmann, A.; Seitz, O. Life Cell Imaging of mRNA Using PNA FIT Probes. *Concepts and Case …* **2014**, 351-363.

(19) Englund, E. A.; Appella, D. H. Synthesis of γ-substituted peptide nucleic acids: a new place to attach fluorophores without affecting DNA binding. *Organic Lett.* **2005**.

(20) Englund, E. A.; Appella, D. H. Gamma-substituted peptide nucleic acids constructed from L-lysine are a versatile scaffold for multifunctional display. *Angewandte Chemie* **2007**, *46*, 1414-1418.

(21) Pujari, F. S. a. S. S. Azide-Alkyne “Click” Conjugation of 8-Aza-7-deazaadenine-DNA: Synthesis, Duplex Stability, and Fluorogenic Dye Labeling. *Bioconjug. Chem.* **2010**, *21*, 1629–1641.

(22) Seela, F.; Jiang, D.; Xu, K. 8-Aza-2'-deoxyguanosine: base pairing, mismatch discrimination and nucleobase anion fluorescence sensing in single-stranded and duplex DNA. *Organic & biomolecular chemistry* **2009**, *7*, 3463-3473.

(23) Johannes Gierlich, G. A. B., Philipp M. E. Gramlich, David M. Hammond, and Thomas Carell. Click Chemistry as a Reliable Method for the High-Density Postsynthetic Functionalization of Alkyne-Modified DNA. *Org. Lett.* **2006**, *8*, 3639-3642.

(24) Hudson, R. H. E.; Li, G.; Tse, J. The use of Sonogashira coupling for the synthesis of modified uracil peptide nucleic acid. *Tetrahedron Lett.* **2002**, *43*.

(25) Shank, N. I. Design, Synthesis and Characterization of Fluorescent Probes for Biomolecule Detection. Ph.D., Carnegie Mellon University, 2012.

(26) Lahann, J. Click Chemistry: A Universal Ligation Strategy for Biotechnology and Materials Science. *Click Chemistry for Biotechnology and Materials …* **2009**.

(27) Kolb, H. C.; Finn, M. G.; Sharpless, B. K. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie International Edition* **2001**, *40*, 2004-2021.

(28) Moses, J. E.; Moorhouse, A. D. The growing applications of click chemistry. *Chem. Soc. Rev.* **2007**, *36*, 1249-1262.

(29) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, B. K.; Finn, M. G. Bioconjugation by copper (I)-catalyzed azide-alkyne [3+ 2] cycloaddition. *Journal of the American Chemical Society* **2003**, *125*, 3192-3193.

(30) Binder, W. H.; Sachsenhofer, R.: *Click Chemistry on Supramolecular Materials*, 2009.

(31) Gramlich, P. M.; Wirges, C. T.; Manetto, A.; Carell, T. Postsynthetic DNA modification through the copper-catalyzed azide-alkyne cycloaddition reaction. *Angew. Chem. Int. Ed.* **2008**, *47*, 8350-8358.

(32) Stadler, A. L.; Delos Santos, J. O.; Stensrud, E. S.; Dembska, A.; Silva, G. L.; Liu, S.; Shank, N. I.; Kunttas-Tatli, E.; Sobers, C. J.; Gramlich, P. M.; Carell, T.; Peteanu, L. A.; McCartney, B. M.; Armitage, B. A. Fluorescent DNA nanotags featuring covalently attached intercalating dyes: synthesis, antibody conjugation, and intracellular imaging. *Bioconj. Chem.* **2011**, *22*, 1491-1502.

(33) Frank Seela, V. R. S., and Padmaja Chittepu. Modification of DNA with Octadiynyl Side Chains: Synthesis, Base Pairing, and Formation of Fluorescent Coumarin Dye Conjugates of Four Nucleobases by the Alkyne-Azide “Click” Reaction. *Bioconjug. Chem.* **2008**, *19*.

(34) He, J.; Seela, F. Propynyl groups in duplex DNA: stability of base pairs incorporating 7‐substituted 8‐aza‐7‐deazapurines or 5‐substituted pyrimidines. *Nucleic Acids Res.* **2002**, *30*, 5485-5496.

(35) Astakhova, I. K.; Wengel, J. Interfacing click chemistry with automated oligonucleotide synthesis for the preparation of fluorescent DNA probes containing internal xanthene and cyanine dyes. *Chemistry* **2013**, *19*, 1112-1122.

(36) Krishnamoorthy Sivakumar, F. X., Brandon M. Cash, Su Long, Hannah N. Barnhill, and Qian Wang. A Fluorogenic 1,3-Dipolar Cycloaddition Reaction of 3-Azidocoumarins and Acetylenes. *Org. Lett.* **2004**, *6*, 4603-4606.

(37) Xie, F.; Sivakumar, K.; Zeng, Q.; Bruckman, M. A.; Hodges, B.; Wang, Q. A fluorogenic ‘click’ reaction of azidoanthracene derivatives. *Tetrahedron* **2008**, *64*, 2906-2914.

(38) Shieh, P.; Hangauer, M. J.; Bertozzi, C. R. Fluorogenic azidofluoresceins for biological imaging. *J. Am. Chem. Soc.* **2012**, *134*, 17428-17431.

(39) Shieh, P.; Bertozzi, C. R. Design strategies for bioorthogonal smart probes. *Organic & biomolecular chemistry* **2014**, *12*, 9307-9320.

(40) Feagin, T. A.; Shah, N. I.; Heemstra, J. M. Convenient and scalable synthesis of fmoc-protected Peptide nucleic Acid backbone. *Journal of nucleic acids* **2012**, *2012*, 354549.

(41) Braasch, D. A.; Nulf, C. J.; Corey, D. R. Synthesis and purification of peptide nucleic acids. *Curr. Protocols in nucleic acid chemistry* **2002**, *Chapter 4*.

(42) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. PNA hybridizes to complementary oligonucleotides obeying the Watson Crick hydrogen-bonding rules. *Nature* **1993**, *365*, 566-568.

(43) Tackett, A. J. Non-Watson-Crick interactions between PNA and DNA inhibit the ATPase activity of bacteriophage T4 Dda helicase. *Nucleic Acids Res.* **2002**, *30*, 950-957.

(44) Svanvik, N.; Nygren, J.; Westman, G.; Kubista, M. Free-Probe Fluorescence of Light-up Probes. *J. Am. Chem. Soc.* **2001**, *123*, 803-809.

(45) Mujumdar, R. B.; Ernst, L. A.; Mujumdar, S. R.; Lewis, C. J.; Waggoner, A. S. Cyanine Dye Labeling Reagents: Sulfoindocyanine Succinimidyl Esters. *Bioconjug. Chem.* **1993**, *4*, 105-111.

(46) Randolph, J. B.; Waggoner, A. S. Stability, specificity and fluorescence brightness of multiply-labeled fluorescent DNA probes. *Nucleic acids research* **1997**, *25*, 2923–2929.

(47) Sahu, B.; Sacui, I.; Rapireddy, S.; Zanotti, K. J.; Bahal, R.; Armitage, B. A.; Ly, D. H. Synthesis and characterization of conformationally preorganized, (R)-diethylene glycol-containing gamma-peptide nucleic acids with superior hybridization properties and water solubility. *J. Org. Chem.* **2011**, *76*, 5614-5627.

(48) Gerowska, M.; Hall, L.; Richardson, J.; Shelbourne, M.; Brown, T. Efficient reverse click labeling of azide oligonucleotides with multiple alkynyl Cy-Dyes applied to the synthesis of HyBeacon probes for genetic analysis. *Tetrahedron* **2012**.

(49) Richardson, J. A.; Gerowska, M.; Shelbourne, M.; French, D.; Brown, T. Six Colour HyBeacon Probes for Multiplex Genetic Analysis. *ChemBioChem* **2010**, *11*, 2530-2533.

(50) Femino, A. M.; Fay, F. S.; Fogarty, K.; Singer, R. H. Visualization of Single RNA Transcripts in Situ. *Science* **1998**, *280*, 585-590.