Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA

A Dissertation

Presented in Partial Fulfillment of the Requirements for the

Degree of Doctorate of Philosophy

with a

Major in Chemistry

in the

College of Graduate Studies

University of Idaho

by

Brooke Alyssa Anderson

Major Professor: Patrick J. Hrdlicka, Ph.D.

Committee Members: Jakob Magolan, Ph.D.; Eric Brauns, Ph.D.; Deborah Stenkamp, Ph.D.

Department Administrator: Ray von Wandruszka, Ph.D.

March 2015

Authorization to Submit Dissertation

This dissertation of Brooke A. Anderson, submitted for the degree of Doctor of Philosophy with a Major in Chemistry and titled "Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor:	_____ Patrick J. Hrdlicka, Ph.D.	Date: _____
Committee Members:	_____ Jakob Magolan, Ph.D.	Date: _____
	_____ Eric Brauns, Ph.D.	Date: _____
	_____ Deborah Stenkamp, Ph.D.	Date: _____
Department Administrator:	_____ Ray von Wandruszka, Ph.D.	Date: _____

Abstract

The main purpose of the work described in this dissertation is to develop oligonucleotide-based probes that can target genomic DNA. The development of probes capable of interrupting the flow of genetic information in living organisms have become an interesting field of research due to their potential as diagnostic and fundamental research tools, and – the grand challenge – therapeutics that can combat diseases of genetic origin. There is an extensive need to expand the current toolbox of double-stranded DNA (dsDNA) targeting probes to enable high specificity targeting at physiologically relevant conditions without sequence limitations. The Hrdlicka lab focuses on the development of a novel DNA targeting methodology utilizing energetically activated DNA duplexes, which potentially overcome the limitations of current DNA recognition strategies (e.g., triplex-forming oligonucleotides, polyamides, and peptide nucleic acids). This approach originally utilized N2′-pyrene-functionalized 2′-amino-α-L-LNA nucleotides as the key activating modifications. However, these building blocks are synthetically difficult to make impeding the full characterization of this novel DNA recognition strategy. Identification of simpler and more readily accessible scaffolds therefore presented itself as a highly desirable goal in order to conduct structure-property relationship studies with the aim of optimizing the dsDNA binding affinity of Invader probes. The work presented in this dissertation describes the synthesis and characterization of oligonucleotides and Invader probes based on (i) N2'-pyrene-functionalized 2'-amino-α-L-LNA adenosine, (ii) N2′-pyrene-/perylene-/coronene-functionalized 2′-*N*-methyl-2′-aminouridine monomers, to study the influence of intercalator size on dsDNA recognition efficiency, (iii) phosphorothioate DNA backbones, to improve pharmacokinetic properties, (iv) S2'-pyrene-functionalized 2'-thiouridine, to study the effect of electronegativity of the 2'-sugar atom on DNA recognition

efficiency, (v) pseudo-complementary Invader building blocks, to further increase the binding affinity of Invader probes. The long-term goal of this research project is to develop simple nucleic acid probes that allow for sequence-unrestricted targeting of double-stranded DNA and to apply these probes as tools in molecular biology, nucleic acid diagnostics, and novel gene therapeutics.

Acknowledgements

I would like to express my deep appreciation and gratitude to my advisor, Dr. Patrick J. Hrdlicka, for his support all the way from the beginning, as an undergraduate researcher in his laboratory, to encouraging me to consider applying to the Ph.D program at the Univesity of Idaho, and through completion of my doctoral degree. Without his guidance and mentorship, I would not be the nucleic acid chemist I have become today.

I would like to thank my committee members, Dr. Jakob Magolan, Dr. Eric Brauns, and Dr. Deborah Stenkamp for their friendly guidance, and constructive suggestions and comments to revise my research proposal and dissertation. In similar vein, I would like to thank my fellow colleagues from over the years, Dr. Saswata Karmakar, Dale C. Guenther, Dr. Michael E. Østergaard, Dr. Mamta Kaura, and Dr. Sujay P. Sau for their support, helpful disscussions and contributations to my intellectural growth throughout the course of my doctoral degree. Additionally, I would like to acknowledge and thank Dr. Lee Deobald and Dr. Alexander Blumenfeld for their generous help with mass spectrometry and NMR spectroscopy.

A very special thank you to Dr. Richard Williams and Dr. Dan Edwards for being great sources of knowledge and always being available for discussing any problems I have in (or out) of the laboratory. Additional thanks to Dr. Williams, Dr. Hrdlicka, and Dr. Edwards for helping my postdoctoral job search by providing what I hear are spectacular letters of recommendation. I am sincerely grateful to all of you.

My gratitude goes out to my many friends who always found a way to help me relax, to put a smile on my face, who made sure I had something to eat on those late nights in the laboratory, who listened to my woes during times of scientific difficulties during my years of research, and

who helped me keep my head up while assuring me that there is a light at the end of every tunnel. I couldn't ask for better, more understanding friends and I feel very blessed for the support I have received from each and every one of them.

Finally, I would like to thank my parents and siblings for their never ending support and for encouraging me to work hard and to keep pushing forward to reach my goals. They are my source of pride and strength and are the reason I have accomplished so much. To them I dedicate this dissertation.

Dedication

For my family and many friends, for their never ending support and for encouraging me to never give up.

Vita

2004-2008: B.Sc. Chemistry – Professional, University of Idaho, Moscow, Idaho, USA

2008-2015: Ph.D. Candidate – Chemistry, University of Idaho, Moscow, Idaho, USA

2007-2014: Teaching Assistant (Chem 111 and Chem 112) and Research Assistant (Nucleic Acid Chemistry, Dr. Patrick J. Hrdlicka's laboratory), Department of Chemistry, University of Idaho, Moscow, Idaho, USA

Honors

2014: Cooley-Juve Award (Outstanding Teaching Assistant in Chemistry)

2013: Sigma Xi Research Society Award

2012: Student Grant Program ($3000), University of Idaho, Moscow, Idaho, USA

2010: INBRE Graduate Student Fellowship, University of Idaho, Moscow, Idaho, USA

2009: Student Grant Program ($3000), University of Idaho, Moscow, Idaho, USA

2008: ASUI & Senator Larry Craig Leadership Award

2008: Merck Index Award

2008: Brian and Gayle Hill Undergraduate Research Fellowship

2004: Freshmen Chemist of the Year, University of Idaho, Moscow, Idaho, USA

Publications

Brooke A. Anderson, Jared J. Onley, Patrick J. Hrdlicka. "Mixed-sequence Recognition of Double-stranded DNA using Energetically Activated Duplex Probes Modified with N2′-Pyrene-, Perylene-, or Coronene-functionalized 2′-*N*-methyl-2′-amino-DNA Monomers," *In preparation.*

Brooke A. Anderson and Patrick J. Hrdlicka "2′-Atom Substitutions and Their Affect on Invader-Mediated DNA Recognition Efficiency," *In preparation.*

Brooke A. Anderson and Patrick J. Hrdlicka. "The Merger of Two Strategies for Recognition of Double-stranded DNA: Pseudocomplementary Invaders," *In preparation.*

Brooke A. Anderson, Saswata Karmakar, Patrick J. Hrdlicka. "Double-stranded DNA Recognition by Enzymatically Stable Phosphorothioate Invader Probes," *In preparation.*

Dale C. Guenther, Grace H. Anderson, Saswata Karmakar, **Brooke A. Anderson**, Bradley A. Didion, Wei Guo, John P. Verstegen, Patrick J. Hrdlicka. "Optimized Invader Probes for Mixed-sequence Recognition of Chromosomal DNA," *In preparation.*

Dale C. Guenther, Pawan Kumar, **Brooke A. Anderson**, Patrick J. Hrdlicka "Amino acid functionalized LNA: Positively poised for antisense applications**,**" *Chem. Commun.*, **2014**, 50, 9007-9009.

Pawan Kumar, Bharat Baral, **Brooke A. Anderson**, Dale C. Guenther, Michael E. Østergaard, Pawan K. Sharma, and Patrick J. Hrdlicka "C5-Alkynyl-Functionalized α-L-LNA: Synthesis, Thermal Denaturation Experiments and Enzymatic Stability," *J. Org. Chem.,* **2014**, 5062-5073.

Pawan Kumar, Michael E. Østergaard, Bharat Baral, **Brooke A. Anderson**, Dale C. Guenther, Mamta Kaura, Daniel J. Raible, Pawan K. Sharma, and Patrick J. Hrdlicka. "Synthesis and Biophysical Properties of C5-Functionalized LNA (Locked Nucleic Acid)," *J. Org. Chem.*, **2014**, 5047-5061.

Nicolai K. Andersen‡, **Brooke A. Anderson**‡, Jesper Wengel, and Patrick J. Hrdlicka "Synthesis and Characterization of Oligodeoxyribonucleotides Modified with 2'-Amino-alpha-L-LNA Adenine Monomers: High-affinity Targeting of Single-Stranded DNA," *J. Org. Chem.*, **2013**, 78, 12690-12702. ‡Contributed equally to study.

Sujay P. Sau, Andreas S. Madsen, Peter Podbevsek, Nicolai K. Andersen, T. Santhosh Kumar, Sanne Andersen, Rie L. Rathje, **Brooke A. Anderson**, Dale C. Guenther, Saswata Karmakar, Pawan Kumar, Janez Plavec, Jesper Wengel, and Patrick J. Hrdlicka "Identification and Characterization of Second-Generation Invader Locked Nucleic Acids (LNAs) for Mixed-Sequence Recognition of Double-Stranded DNA," *J. Org. Chem.*, **2013**, 78, 9560-9570.

Saswata Karmakar, **Brooke A. Anderson**, Rie L. Rathje, Sanne Andersen, Troels B. Jensen, Poul Nielsen, and Patrick J. Hrdlicka. "High-Affinity DNA Targeting Using Readily Accessible Mimics of N2'-Functionalized 2'-Amino-α-L-LNA," *J. Org. Chem.*, **2011**, 76, 7119-7131.

Michael E. Østergaard, Pawan Kumar, Bharat Baral, Dale C. Guenther, **Brooke A. Anderson**, F. Marty Ytreberg, Lee Deobald, Andrzej J. Paszczynski, Pawan K. Sharma, and Patrick J. Hrdlicka. "C5-Functionalized DNA, LNA, and a-L-LNA: Positional Control of Polarity Sensitive Fluorophores Leads to Improved SNP-Typing," *Chem. A. Eur. J.*, **2011**, 17, 3157-3165.

Michael E. Østergaard, Pawan Kumar, Bharat Baral , Daniel J. Raible , T. Santhosh Kumar, **Brooke A. Anderson**, Dale C. Guenther, Lee Deobald, Andrzej J. Paszczynski, Pawan K. Sharma, Patrick J. Hrdlicka. "C5-Functionalized LNA: Unparalleled Hybridization Properties and Enzymatic Stability," *ChemBioChem*, **2009**, 10, 2740-2741.

Sujay P. Sau, Pawan Kumar, **Brooke A. Anderson**, Michael E. Østergaard, Lee Deobald, Andrzej Paszczynski, Pawan K. Sharma and Patrick J. Hrdlicka "Optimized DNA-targeting using triplex forming C5-alkynyl functionalized LNA," *Chem Commun.*, **2009**, 6756-6758.

Bharat Baral, Pawan Kumar, **Brooke A. Anderson**, Michael E. Østergaard, Pawan K. Sharma and Patrick J. Hrdlicka. **"**Optimized synthesis of [3-^{15}N]-labeled uridine phosphoramidites," *Tetrahedron Letters*, **2009**, 50, 5850-5852.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

A	Adenin-9-yl
Ac	Acetyl
aq.	Aqueous
Bn	Benzyl
Bz	Benzoyl
C	Cytosin-1-yl
Cas	CRISPR Associated
cDNA	Complementary DNA
CRISPR	Clustered Regularly Interspaced Short Palendromic Repeat
cRNA	Complementary RNA
Cor	Coronene
D	2,6-Diamino Adenosine
DA	Deviation from Additivity
DCE	1,2-Dichloroethane
DCM	Dichloromethane
DIPEA	N,N-diisopropylethylamine
DMAP	Dimethylaminopyridine
DMSO	Dimethylsulfoxide
DMTr	4,4′-Dimethoxytrityl
DNA	Deoxyribonucleic Acid
DNase	Deoxyribonuclease
dsDNA	Double-stranded DNA
EDTA	N,N,N′,N′- ethylenediamine tetraacetic acid
ESI	Electrospray Ionization
EtBr	Ethidium Bromide
Fmoc	9-Fluorenylmethoxycarbonyl
G	Guanin-9-yl
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HP	Hairpin
HPLC	High Pressure Liquid Chromatography

INA	Intercalating Nucleic Acid
LNA	Locked Nucleic Acid
MALDI	Matrix Assisted Laser Desorption Ionization
mC	5-methylcytosine
mRNA	Messenger RNA
MS	Mass Spectrometry
Ms	Methanesulfonyl
NMR	Nuclear Magnetic Resonance
nt	Nucleotide
ON	Oligodeoxyribonucleotide
PAGE	Polyacrylamide Gel Electrophoresis
pc	Pseudo-complementary
Pery	Perylene
PNA	Peptide Nucleic Acid
PO	Phosphodiester
PS	Phosphorothioate
Py	Pyrene
RNA	Ribonucleic Acid
rt	Room Temperature
S	C2-Thio Thymine
SNP	Single Nucleotide Polymorphism
ss	Single-stranded
SVPDE	Snake Venom Phosphodiesterase Enzyme
T	Thymin-1-yl
TA	Thermal Advantage
TBE	Tris/Borate/EDTA
TBM	Tris/Borate/Magnesium
T_m	Thermal Denaturation Temperature
TALEN	Transcription Activator-Like Effector Nuclease
TFO	Triplex Forming Oligonucleotide
TLC	Thin Layer Chromatography

Tris	Tris(hydroxymethyl)aminomethane
U	Uracil-1-yl
UV	Ultra Violet
ZFN	Zinc-Finger Nuclease

CHAPTER 1: Chemical Strategies for Recognition of Double-stranded DNA

1.1 Nucleic Acid Structure and Function

Nucleic acids are arguably the most important class of biomolecules that exist within the cells of living organisms. Governing the expression of proteins in cells, it is nucleic acids that hold the blueprint for an individual organism's biochemistry. DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) are the carriers of genetic information that encode for proteins in the cells of all Kingdom's of life. These proteins are workhorse molecules that carry out many catalytic and regulatory functions in cells. DNA and RNA are composed of polymers of nucleotides. Nucleotides consist of a) a heterocyclic nucleobase, b) a sugar moiety, and c) a phosphodiester backbone. Although the sugar moiety and the phosphate backbone play a vital role in the overall structure of the DNA or RNA, the genetic information is encoded in the sequence of the nucleobases. Adenine (A), cytosine (C), and guanine (G) are found in DNA/RNA while thymine (T) and uracil (U) are found only in DNA or RNA, respectively (Figure 1-1). The difference at the C2′-position of the furanose sugar and the C5-position of the thymine/uracil nucleobase are the distinguishing factors between DNA and RNA. RNA is more hydrophilic than DNA due to the additional C2′-hydroxyl group and the lack of a C5-methyl group. Moreover, the presence or absence of a C2′-hydroxyl group results in stereoelectronic differences that manifest themselves as differences in DNA and RNA duplex geometry.

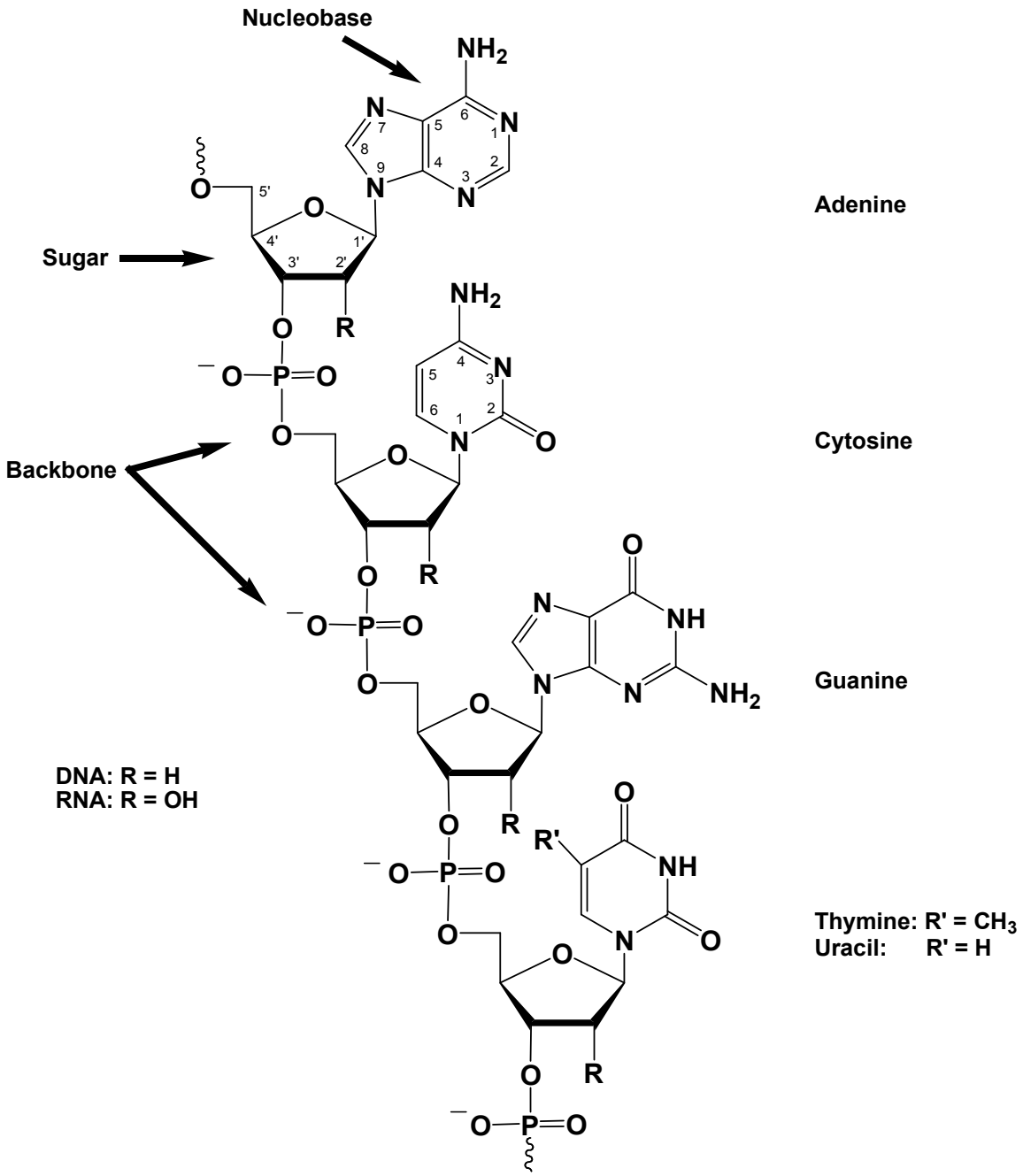


Figure 1-1: Single-stranded segment of DNA/RNA of sequence 5′-(…ACG T/U…)-3′.

The genetic information of a living organism is stored in the nucleus of mammalian cells as double-stranded DNA (dsDNA). Sequence information transfer from DNA to RNA to protein is highly organized. Inside the nucleus of the cell, transcription of DNA catalyzed by a DNA-

dependent RNA polymerase generates complementary RNA. The primary RNA transcript is then trimmed and processed inside the nucleus to generate messenger RNAs (mRNAs), which are transported into the cytoplasm of the cell and translated by ribosomes and transfer RNAs (tRNAs) to generate polypeptides which fold to produce active proteins (Figure 1-2).

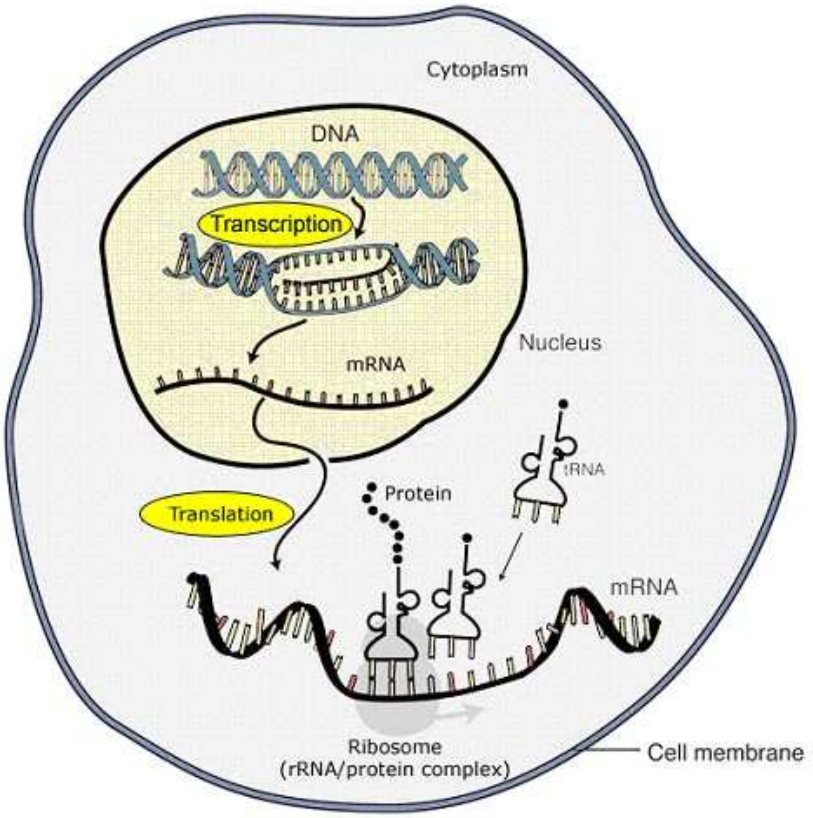


Figure 1-2: The central dogma of molecular biology.[1]

Misregulation of gene expression can result in diseases, such as the wrong protein being formed in a particular cell or too much or too little of the right protein, metabolite, or chemical messenger being formed. Accordingly, there has been great interest in developing ligands that can modulate the level of gene expression. Medicinal chemistry has traditionally focused on

the development of small molecules that directly target disease-related proteins. A conceptually more elegant strategy is to decrease or altogether prevent the expression of the gene coding for a particular disease-related protein. In this so-called antigene (ag) approach, transcription is prevented through the binding of chemically modified ligands to dsDNA. Molecules that bind dsDNA with high affinity and specificity have a remarkable potential as therapeutic agents as well as molecular diagnostic probes.

Modulation of gene expression, and more specifically transcription, has been realized via a variety of different approaches (Figure 1-3).[2] Binding of a high affinity ligand or oligonucleotide (ON) to either a promotor region or a transcription factor binding site of dsDNA can result in inhibition of transcription initiation by preventing the RNA polymerase complex or other transcription factors from binding. Likewise, binding to a transcribed region of the targeted gene can arrest transcription elongation. DNA binding ligands modified with a mutagen can induce site-specific DNA damage (i.e., site-specific mutagenesis). Upon inducing mutagenesis, introduction of a donor DNA sequence can result in homologous recombination, resulting in DNA repair. Fluorophore-labeled ligands can be used as molecular diagnostic probes, for example, for detection of genomic mutations such as single nucleotide polymorphisms (SNPs). Such ligands would be beneficial for clinical diagnostics and development of molecular assays for analyzing genomic variations.

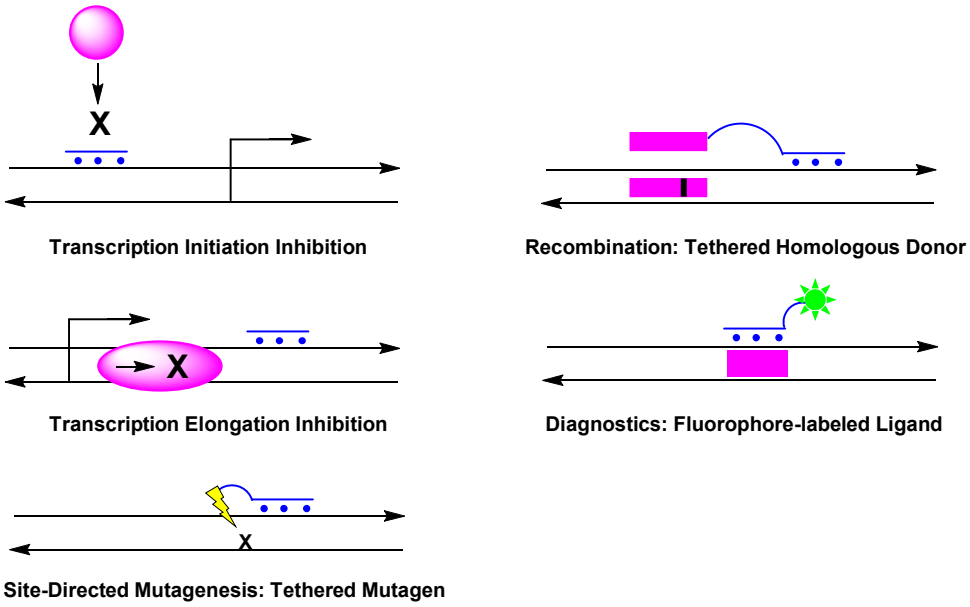


Figure 1-3: Pathways for modulation of gene expression and gene diagnostics.

Development of probes for site-specific targeting of DNA represents one of the most daunting challenges of contemporary biological chemistry. Although there is a limited amount of targets in every cell compared to the number of RNA targets or protein targets, the compact supercoiled packing of DNA in the nucleus of the cell renders recognition of dsDNA very challenging. However, progress has been achieved with compounds that recognize accessible features from (1) the DNA grooves, such as triplex forming oligonucleotides (TFOs) and minor groove binding polyamides, (2) by a combination of triplex/duplex invasion such as peptide nucleic acids (PNAs), or (3) by invasion of pre-existing Watson-Crick base pairing such as pseudo-complementary DNA/PNA (pcDNA/pcPNA), Invader nucleic acids, intercalating nucleic acids (INAs), or (4) by pathways that are not fully understood, such as Zorro locked nucleic acids (LNAs), agLNAs, and agPNAs. All of these are discussed in more detail in the following sections.

1.2 Triplex Forming Oligonucleotides (TFOs)

1.2.1 Structure and Limitations

Triplex forming oligonucleotides (TFOs) have been one of the most studied approaches for recognition of dsDNA.[3] TFOs bind to the major groove of a DNA duplex target region and form non-covalent Hoogsteen base-pairs to polypurine sequences (Figure 1-4). The resulting triplex region is stable enough to interfere with initiation and elongation of gene transcription.[4] Purines are the only nucleotides with the capacity of forming Hoogsteen bonds as they, unlike pyrimidines, have additional hydrogen bond acceptor and donor regions available for triplex binding situated in the major groove of the duplex.

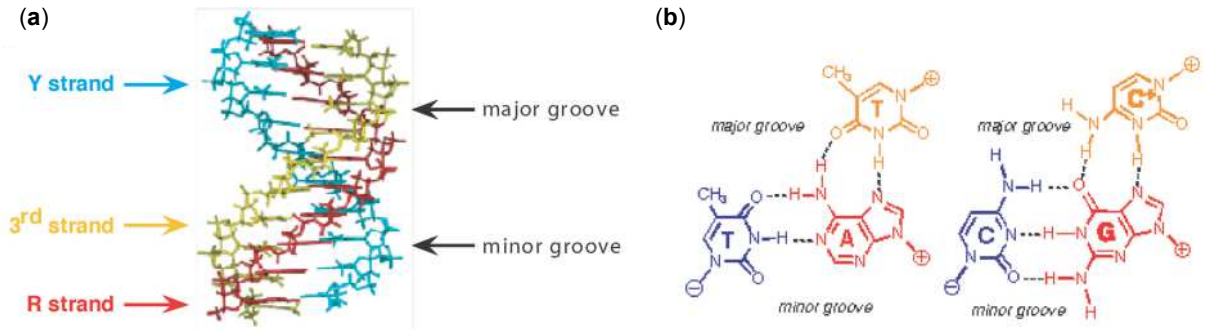


Figure 1-4: Illustration of triplex structure. (a) The third strand is shown in yellow, whereas the oligopurine and the oligopyrimidine strand are shown in red and blue, respectively. (b) The triplet motifs of a pyrimidine triplex are shown, TA·T and CG·C^+, respectively from left to right. "·" denotes Hoogsteen bonds. (Reproduced with permission from ref. 3. Copyright 2008 Oxford University Press).

Three main classes of triplex forming motifs are known that differ with respect to the base composition and orientation of the TFO (Figure 1-5).[5] The nucleobases of the TFO (in yellow)

bind to the homopurine strand (red strand) of the target DNA via Hoogsteen or reverse Hoogsteen hydrogen bonds and follow a parallel or an antiparallel orientation relative to the target. TC-motif TFOs bind in a parallel orientation to form T·AT and C^{+}·GC Hoogsteen triplets. GT-motif TFOs can bind in both parallel and antiparallel orientations, resulting in Hoogsteen or reverse Hoogsteen hydrogen bonds of the form G·GC and T·AT triplets. GA-motif TFOs are in antiparallel orientation forming reverse Hoogsteen hydrogen bonds resulting in G·GC and A·AT triplets. Vekhoff et. al. have developed a set of simple rules for predicting the best triple-helical binding motifs based off the sequence composition of the target DNA.[6] The most stable triplexes are antiparallel GU-motif TFOs and TC-motif TFOs. To determine between GU- and TC- motif TFOs, the target length and the percentage G of the target and the TFO are taken into effect.

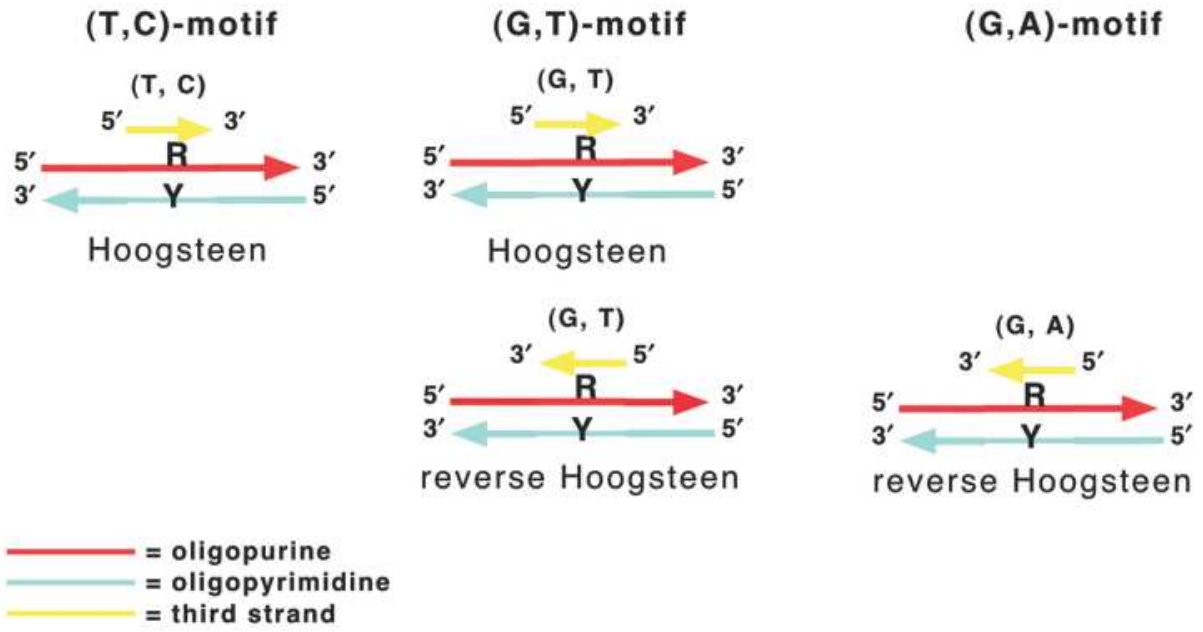


Figure 1-5: Orientation of the three triplex motifs. (Reproduced with permission from ref. 3. Copyright 2008 Oxford University Press).

Stable triplex formation requires a contiguous homopurine stretch, as pyrimidine interruptions within the purine stretch can inhibit triplex formation, which limits the number of suitable target sites.[7-9] A serious limitation for the formation of stable triplexes using TC-motif TFOs is the requirement that the N3 atom of cytosine must be protonated to enable formation of Hoogsteen hydrogen bonds (C^+·GC base pair, Fig. 1-4b). At physiological pH, this protonation only occurs to a minor extent[3] due to the low pK_a of cytosine (~4.35) and TC-motif TFOs generally only form stable triplexes at pH < 6.0. GT- and GA-motif TFOs exhibit other limitations, such as formation of secondary structures at physiological concentrations of monovalent and divalent cations, which impedes triplex structure formation.[6] Even when formed, triplexes are relatively labile under physiological conditions due to unfavorable charge repulsion between the three negatively charged strands. The nuclear environment for in vivo applications presents additional challenges, including the TFO needing to be nuclease resistant in order for it to interfere with biological processes that act on DNA. Despite these limitations, triplex forming oligonucleotide technology has been extensively studied over the last few decades with many of the limitations being alleviated by the use of chemically modified building blocks.

1.2.2 Modifications to Increase Triplex Formation and Stability

Significant progress has been made in developing chemically modified TFOs to improve triplex stability (Figure 1-6). For example, 5-methylcytosine (mC) is one of the most commonly used base modifications for alleviating the pH restriction of TC-motif TFOs.[10] Cytosine has a pK_a of ~4.35, although this is likely higher in TFOs, explaining why triplexes form at a pH up to 6.0.[11] 5-Methylcytosine has a slightly higher pK_a than cytosine (~4.50) and TFOs modified with this

unit are found to more stable at physiological pH. It has been suggested that the increase in stability is a result of the extra spine of methyl groups within the DNA major groove resulting in a favorable entropy change upon duplex formation[12] rather than the increased pK_a. Thymidine has also been replaced by 2′-deoxyuridine (dU)[13] or affinity-enhancing building blocks such as 5-propynyl-2′-deoxyuridine (pU)[14] or positively charged 5-propargylamine-2′-deoxyuridine (UP)[15-17]. The stabilizing effects exerted by these modifications is ascribed to favorable π-π stacking of the modified base with neighboring bases in the triplex.[17-19]

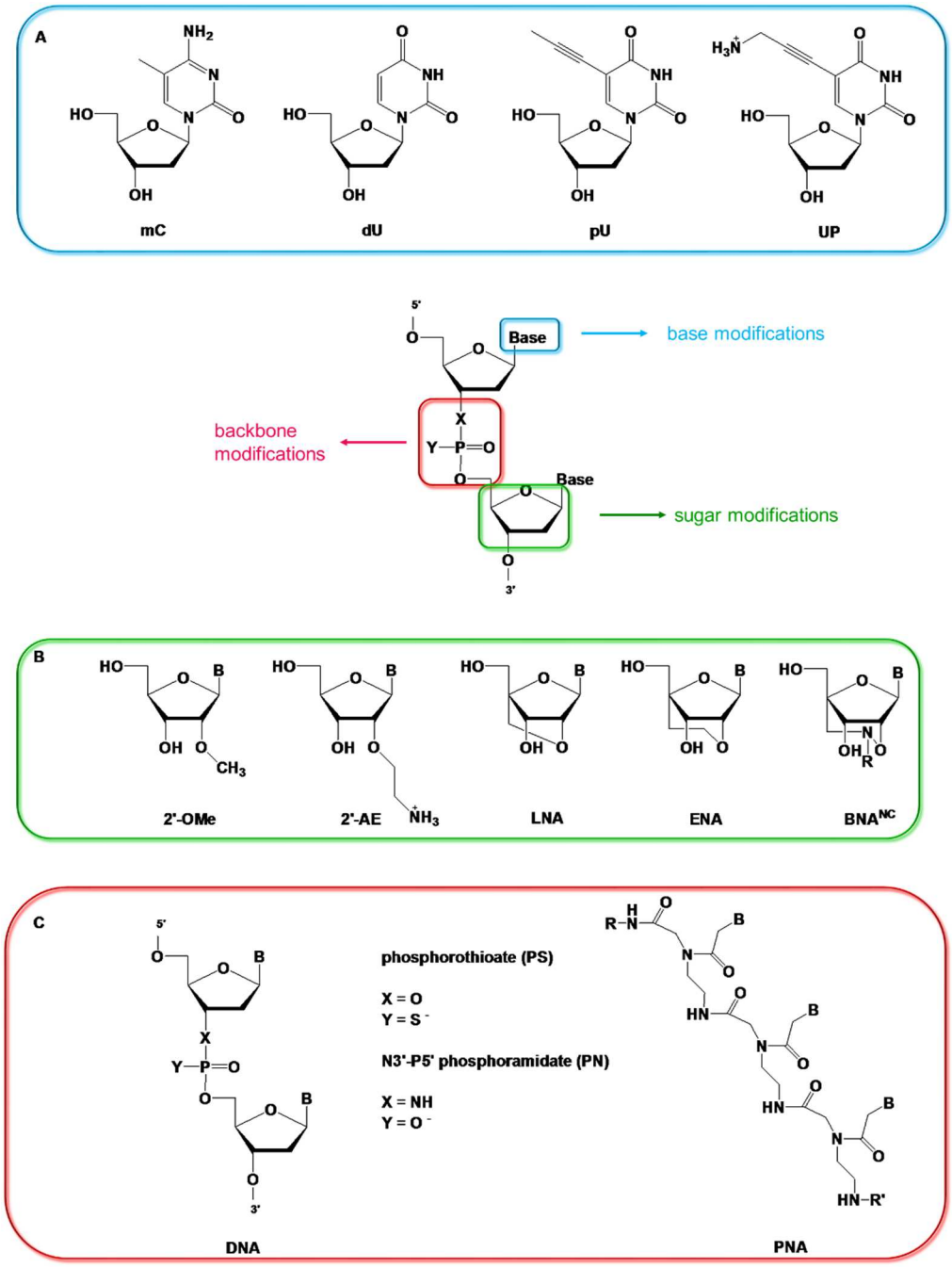


Figure 1-6: Chemical modifications of TFOs. (A) Nucleobase modifications; (B) sugar modifications; and (C) backbone modifications.

An alternative strategy to increase the dsDNA affinity of TFOs is by incorporation of sugar-modified nucleotides. Numerous useful monomers have been reported, including 2′-*O*-methyl RNAs (2′-OMe)[20,21], 2′-aminoethyl RNAs (2′-AE)[16], and conformationally restricted monomers such as O2′,O4′-methylene linked locked nucleic acid (LNA)[22-25], O2′,O4′-ethylene linked nucleic acid (ENA)[26], and a variety of alternative six-membered bridged nucleic acids (BNAs),[27,28] such as 2′,4′-BNA^{NC} monomers containing an N-O linkage, with a functionalizable nitrogen atom on the bridge[29] (Figure 1-6B). These modifications share the characteristic of being conformationally restricted in an RNA-like C3′-*endo* conformation, resulting in enhanced triplex stabilization. RNA-like modifications in the TFO have been observed to form more stable triplexes than their 2′-deoxy counterparts[30] as the overall helical structure of the underlying duplex is between an A and a B form.[31] The 2′-aminoethyl ribose analogue combines the C3′-*endo* character with a positive charge, as the amine is protonated at physiological pH, which reduces charge repulsion and increases resistance to enzyme mediated degradation in living cells.[32] LNA containing TC-motif TFOs induce significant triplex stabilization at physiological conditions in vials, but must be carefully designed to contain alternate LNA and DNA bases. Extensive LNA modification results in very rigid structures that prevent triplex binding[33] or which reduce the biological activity of TFOs in vivo due to nonspecific target binding.[34] ENA-modified TFOs result in slightly less stable triplexes than LNA-modified TFOs, although fully ENA-modified TFOs can form stable triplexes at physiological pH due to their slightly less rigid structure unlike fully modified LNA-TFOs. BNA^{NC} modified TC-motif TFOs have greater affinity for DNA and higher nuclease stability than either LNAs or ENAs. Their rigidity is similar to that of ENA enabling fully BNA^{NC}[NH] (in which the cyclic amine is not functionalized) modified TC-motif TFOs to form stable triplexes at physiological pH.

TFO building blocks that are modified in the nucleobase as well as the sugar region have shown promising dsDNA affinity.[19,35] For example, monomers that combine C5-alkynyl modifications with 2′-*O*-alkyl modifications, such as 2′-*O*-methylribose and 2′-*O*-aminoethylribose, have been shown to display significantly increased affinity toward DNA targets relative to the corresponding monomers without nucleobase modifications.[36-38] Extending this strategy, Hrdlicka and coworkers developed C5-alkynyl functionalized LNAs and α-L-LNAs that improve the stability of LNA-modified triplexes (Figure 1-7).[39,40]

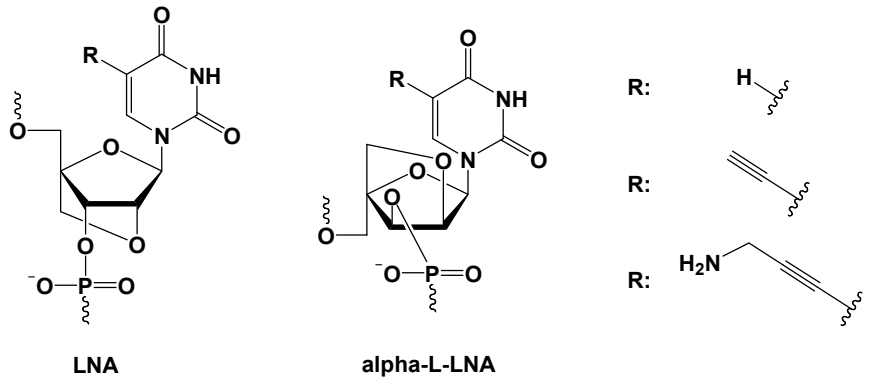


Figure 1-7: Structures of C5-alkynyl-functionalized LNA and α-L-LNA.

Incorporation of C5-ethynyl and C5-propargylamine LNA U monomers into TC-motif TFOs results in improved thermal stability of triplexes and increased binding fidelity relative to correspondingly TC-motif TFOs modified with canonical LNA-T monomers.[39] TFOs that were densely functionalized with C5-propargylamine LNA U resulted in higher stabilization than C5-ethynyl LNA U, presumably due to protonation of the C5-amino residue. Triplexes modified with C5-ethynyl and C5-propargylamine α-L-LNA-U result in increased thermal stability and increased discrimination toward DNA duplexes with homopurine mismatch sites. The stabilization of LNA and α-L-LNA modified triplexes is directly correlated to the dissociation rates of the triplexes, i.e., the more stable the TFO, the slower the dissociation rate.

Relative to C5-alkynyl-functionalized LNA TFOs, the correspondingly modified α-L-LNA TFOs resulted in slightly faster triplex dissociation rates indicating that C5-alkynyl-functionalized monomers with a DNA-like conformation (α-L-LNA) are less stable than modifications with an RNA-like conformation (LNA).[40]

Backbone modifications constitute the third major modification strategy. Phosphorothioate (PS) backbone modifications have been prominently used in place of the natural phosphodiester (PO) backbone as they provide increased nuclease resistance.[41-43] The phosphoramidate (PN) cationic backbone has also been studied and results in favorable electrostatic interactions giving rise to enhanced triplex stability, enzymatic stability and cellular permeability.[44,45] Triplex forming peptide nucleic acids (PNAs) in which the nucleobases are attached to an *N*-(2-aminoethyl)glycine unit form highly stable triplexes[46] and display excellent nuclease and protease resistance.[47] The neutral backbone of PNA is believed to reduce the electrostatic repulsion of the triplexes.[5]

1.2.3 Overcoming Poly-Guanine Secondary Structure Formation

GA- and GT-motif TFOs mainly bind in an antiparallel orientation via reverse Hoogsteen bonds. The stability of triplexes formed with these G-rich TFOs has been shown to be high; however, some G-rich TFOs engage in auto-association processes at physiological potassium and magnesium cation concentrations, which compete with triplex formation. The use of 2′-deoxy-6-thioguanosine (sG), 7-deaza-2′-deoxyguanosine (dzaG), or 7-deaza-2′-deoxyxanthosine (dzaX) inhibits secondary structure formation (e.g. G-tetrad structures) in G-rich TFOs, which facilites triplex formation (Figure 1-8).[48-50]

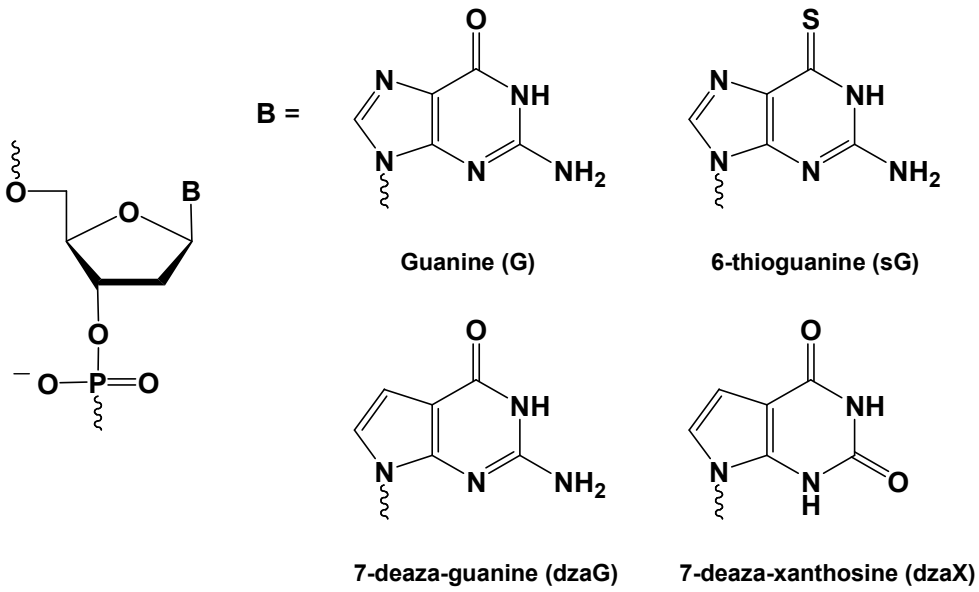


Figure 1-8: Structures of 2′-deoxy-sugars with modified bases for overcoming G-tetrad structure formation in G-rich TFOs.

It is hypothesized that 6-thioguanine inhibits G-tetrad formation at physiological potassium concentrations due to the increased radius and decreased electronegativity of the sulfur atom, rendering it less likely to interact via ion-dipole interactions with potassium ions , thus facilitating triple helix formation in physiological buffers.[49,50] Likewise, replacement of the N7 of guanine with a carbon eliminates the ability of the TFO to form a G-quartet, while retaining the ability to form a dzaG·GC reverse Hoogsteen bond.[48] Neighboring thymines are known to enhance the stability of G-quartets, whereas other bases do not.[51] Replacement of 7-deazaxanthine (dzaX) for thymine in GT-motif TFOs results in a more purine-like nucleobase and decreases the propensity of the ODNs to form secondary structures while increasing the

triplex binding affinity due to decreased repulsive forces between the phosphodiester backbone of the third strand and the phosphodiester backbone of the duplex.[48]

1.2.4 Targeting Pyrimidine Interruption Sites

One of the most severe limitations of TFOs is that they only can bind to long polypurine stretches. Even a single pyrimidine interruption site can decrease the stability of the triplex to a point where it does not form. To overcome this, several modified oligonucleotides have been introduced. Introduction of W-shaped nucleoside analogs (WNAs) increases the affinity of the TFO to the pyrimidine interruption site.[52,53] The addition of a benzene ring increases base-stacking interactions whereas the nucleobase, which is separated from the base-stacking benzene, can form Hoogsteen hydrogen bonds to the purine located in the interruption site. Therefore, introduction of WNAs into TFOs increases the dsDNA affinity of purine-rich TFOs; WNA-βT and WNA-βC, in particular, recognize TA and CG interruptions with high selectivity in the polypurine target site (Figure 1-9a).[52,53]

Imanishi, Obika and coworkers have developed LNA monomers in which the natural nucleobase has been replaced with artificial nucleobases to enable recognition of interruption sites under physiological conditions (Figure 1-9b).[54-57] TFOs modified with LNAs having artificial nucleobases such as 2-pyridone[54], 5-methyl-2-pyridone[55], 1-isoquinolone[56], and 2-pyridine[57] selectively recognize CG interruption sites with higher affinity than their DNA counterparts. Thymine can recognize CG interruption sites with low affinity, but lacks selectivity as it binds more strongly to AT base pairs. 2-pyridone and 5-methyl-2-pyridone LNA monomers produce TFOs that are more stable and more selective for recognition of CG interruption sites than DNA-T (singly-modified TFOs result in ΔT_m = +11.0 to +12.0 °C relative

to a DNA-T modification in the same sequence under the same conditions).[55] The stability arises from the 2-carbonyl group of the 2-pyridone and 5-methyl-2-pyridone units engaging in Hoogsteen-bonding to the C4-amine hydrogen of cytosine in the CG interruption site. The ability to discriminate AT sites from CG sites stems from the lack of the N3-atom and C4-carbonyl groups in 2-pyridone, which are crucial for forming bonds to AT base pairs. Introduction of 1-isoquinolone LNA monomers into TFOs results in less effective recognition of dsDNA with CG interruption sites, but much more CG selective recognition relative to 2-pyridone LNA modified TFOs.[56] The lower affinity against targets with CG interruption sites is likely a result of steric hindrance of the 1-isoquinolone. At the same time, this steric hindrance accounts for the increased selectivity of the modification to CG sites over AT sites as 1-isoquinolone interferes with the 5-methyl group of T at an AT site. Interestingly, increasing the number of 2-pyridone or 5-methyl-2-pyridone LNA modifications and of regularly stabilizing LNA-T modifications allows for high affinity targeting of dsDNA targets with multiple CG interruption sites under physiological conditions.[55] LNA-2-pyridine behaves similarly to LNA-2-pyridone, resulting in identical thermal stabilities under identical conditions.[57]

The same group also introduced alternative LNAs monomers for targeting TA interruption sites, namely LNA-bP and LNA-Tz (bP = 4-(3-benzamidophenyl)-2-pyridone whereas Tz = 2-(N-methylbenzamido)thiazole).[58] These modifications do not recognize interruption sites using hydrogen bonding interactions like the modifications recognizing CG interruption sites. Instead, their selective affinity arises from partial intercalation of the nucleobase surrogate bP and van der Waals interactions of Tz. Tz was shown to be more selective for TA sites than bP, but was less stabilizing, with a Tz-modified TFO resulting in triplexes with T_m's (thermal denaturation temperatures) that are ~7 °C lower than corresponding triplexes involving a bP-

modified TFO.[58] Although the series of artificial nucleobase modified LNA TFOs have been shown to target sequences containing one to three pyrimidine interruption site, sequence limitations persist and recognition of a mixed-sequence DNA targets remains challenging.

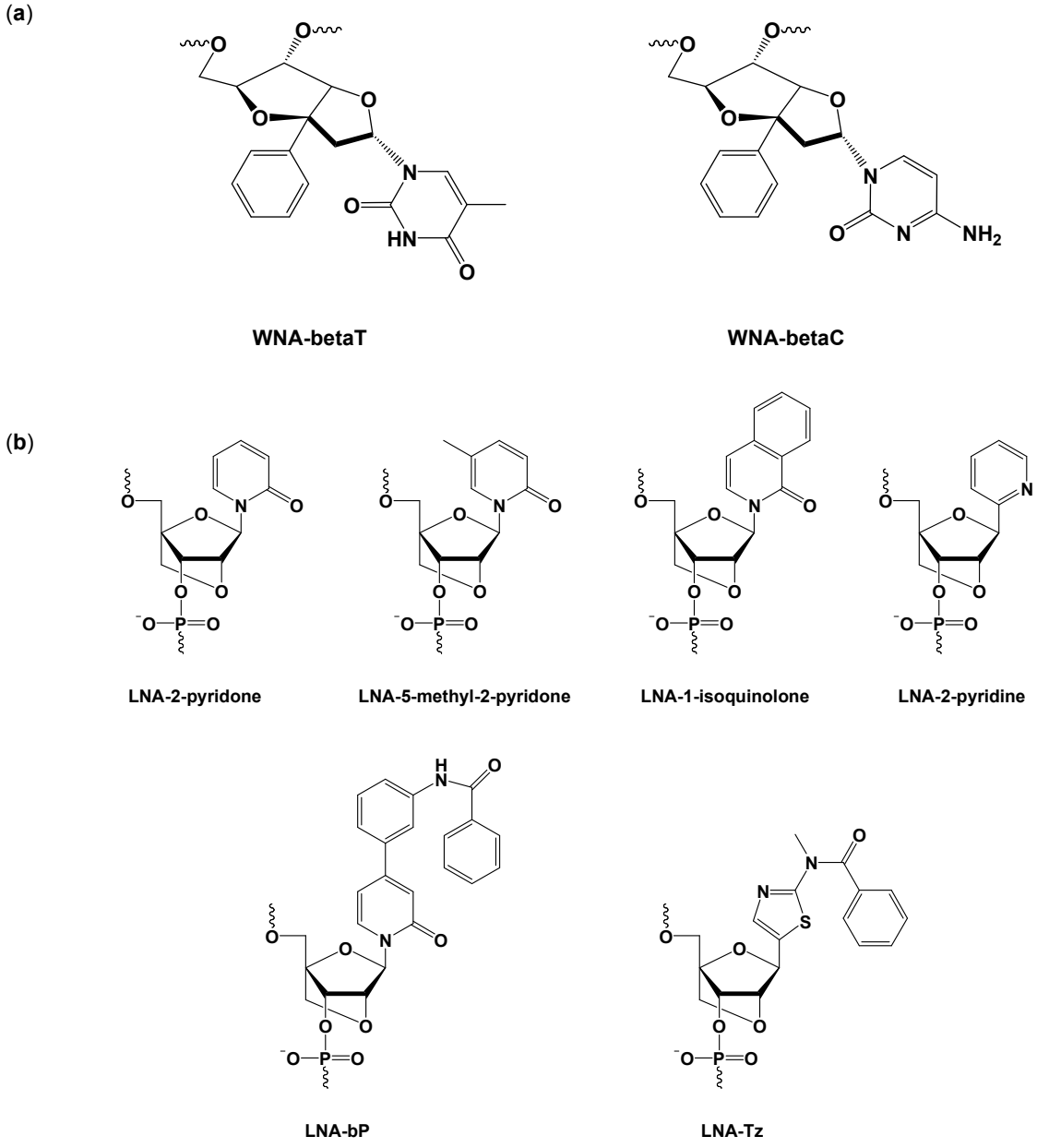


Figure 1-9: Monomers for targeting TFO pyrimidine interruption sites. (a) WNA nucleosides; (b) Nucleobase-modified LNA nucleosides.

Rusling et. al. have reported four base recognition of dsDNA at physiological pH using TFOs modified with four different synthetic nucleotides (BAU, ^{Me}P, ^{A}PP, and S), each specifically targeting AT, GC, CG, or TA base pairs (Figure 1-10).[59] These nucleotides are able to form stable triplexes at mixed-sequence target sites at physiological pH, although their affinity is ~100-fold lower than at pH 5.0 due to insufficient ^{Me}P protonation at pH 7.0. Additionally, the selectivity of S for a TA inversion site is minimal, as it also recognizes CG base pairs with similar affinity. There is therefore still a need for the development of TFO monomers that enable mixed-sequence recognition of DNA targets at physiological conditions.

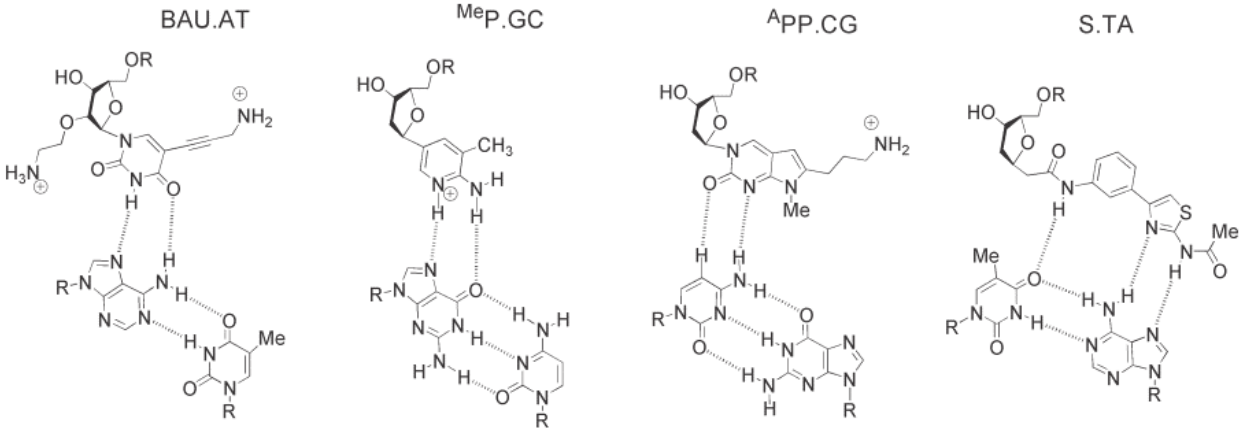


Figure 1-10: Chemical structure of the four-base triplets for recognition of mixed-sequence TFO targets using Rusling's approach. (Reproduced with permission from ref. 59. Copyright 2005 Oxford University Press).

1.2.5 Alternate Strand Hoogsteen Triplexes - Crossover Strategies

When two short oligopurine tracts are present on alternate strands of dsDNA, it is possible to use two covalently linked TFOs for recognition of the dsDNA target (Figure 1-11). For

example, Horne and Dervan have used an alternate strand approach in which a 9-mer TC-motif TFO was linked via a 3'-3' phosphodiester and 1,2-dideoxy-*D*-ribose to a second 9-mer TC-motif TFO.[60] The bidirectional TFO oligonucleotide binds parallel to both purine regions of the target (Figure 1-11a). When this TFO was linked to an EDTA·Fe group, efficient double-stranded cleavage of the target site in plasmid DNA was observed.[60]

Filichev et. al. focused on a 5'-5' linkage utilizing a series of intercalators at the linker position of alternate strand T-rich TC-motif TFOs to further stabilize the triplex (Figure 1-11c).[61] TFOs with linkers **1** – **3** (shown in Figure 1-11c) formed stable triplexes at pH 6.0, but at pH 7.2 only the TFO with linker **2** formed a stable triplex. Molecular modelling studies indicates that the two naphthalene rings of linker **2** engage in more efficient π-π stacking with the alternate strand nucleobases of the TFO relative to the benzene rings in linkers **1** and **3**. The larger aromatic surface area results in much higher thermal stability of linker **2** relative to linkers **1** and **3**. The alternate strand TFOs with intercalator-functionalized linkers allows for triplex formation at physiological pH. However, if the T-rich pyrimidine TFO strand contains cytosine or 5-methyl-cytosine, physiological pH targeting remains limited.

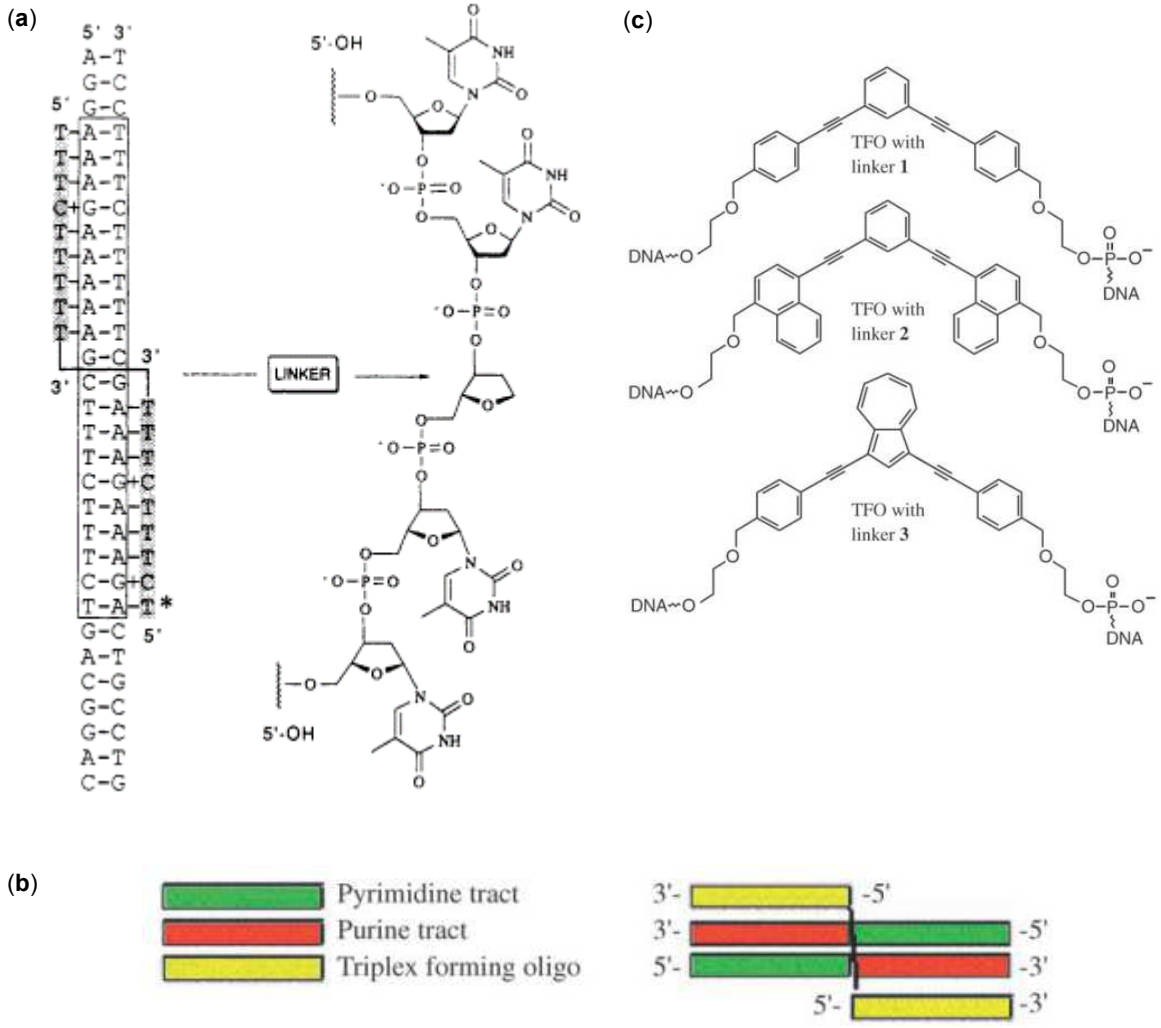


Figure 1-11: Alternate strand TFOs. (a) Illustration of the 5'-(pyrimidine)$_9$-linker-(pyrimidine)$_9$-5' TFO containing a 3'-3' phosphodiester and 1,2-dideoxy-*D*-ribose linker binding alternate strand purine stretches of target DNA. (Reprinted with permission from ref. 60. Copyright 1990 American Chemical Society). (b) Representation of the 5'-5' alternate strand triplex context. (c) Intercalator based linkers for 5'-5' alternate strand TFOs. (Reprinted with permission from ref. 61. Copyright 2006 John Wiley and Sons).

1.3 Minor Groove Binding Polyamides

1.3.1 Structure and Limitations

While TFOs have been developed to target DNA via recognition of the major groove, recognition of the DNA minor groove is also possible using minor groove binding polyamides. There are many natural proteins that bind to dsDNA, with the majority of DNA binding proteins relying on major groove contacts. However, nature has also evolved a series of proteins that can recognize base pair signatures from the minor groove, such as TATA-box binding proteins and integration host factors responsible for DNA organization and transcriptional regulation.[62] Certain crescent shaped natural products, such as netropsin and distamycin A, also bind to the minor groove of dsDNA via hydrogen bonding interactions to exposed electron lone pairs in the minor groove. These natural products preferentially bind to A/T tracts with modest affinity and specificity.[63] Analysis of the DNA interactions of these molecules, which are based mostly on repeating *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) units (Figure 1-12), and the convenience of these polyamides to synthetic manipulation, has allowed for extensive investigations into the principles of minor groove binding interactions.[63]

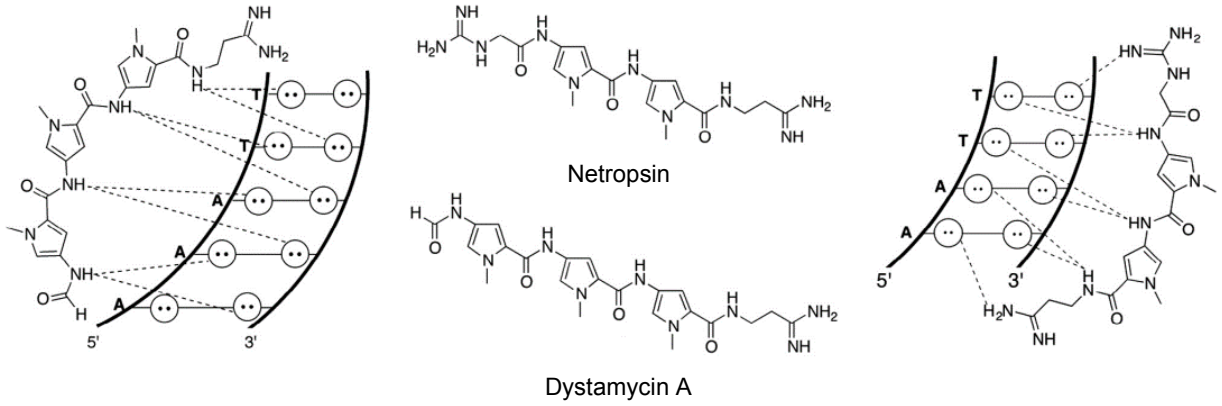


Figure 1-12: DNA minor-groove binding small molecules. Netropsin and distamycin A are shown with diagrammatic representations of molecular interactions involved in DNA binding. Circles with dots represent lone pairs of N(3) of purine and O(2) of pyrimidine. Potential hydrogen bonds between small molecules and the minor groove of DNA are represented with dotted lines. (Reprinted with permission from ref. 52. Copyright 2013 Elsevier Ltd).

More than two decades of research in the Dervan group has resulted in a new paradigm for sequence-specific recognition of virtually any predetermined DNA sequence via the minor groove.[65-71] The Watson-Crick base pairs have specific signatures of hydrogen acceptors and donors facing the minor groove (Figure 1-13a). There have been many studies devoted to establishing the pairing rules for polyamides with DNA. DNA binding polyamides composed of N-methylpyrrole (Py) and N-methylimidazole (Im) are crescent shaped molecules that can bind to the minor groove of DNA as antiparallel dimers (Figure 1-13b). DNA association is driven by a combination of van der Waals and hydrogen bonding interactions with unsymmetrical pairings of Py and Im (i.e., Py/Im or Im/Py in which Py and Im are paired on opposite of one another) with DNA resulting in the sequence-specificity of GC and CG base pairs (Figure 1-13).[70] The use of symmetrical pairing of Py/Py appears to be degenerate,

resulting in targeting of both AT and TA base pairs due to their similarity as from the minor groove (Figure 1-13).

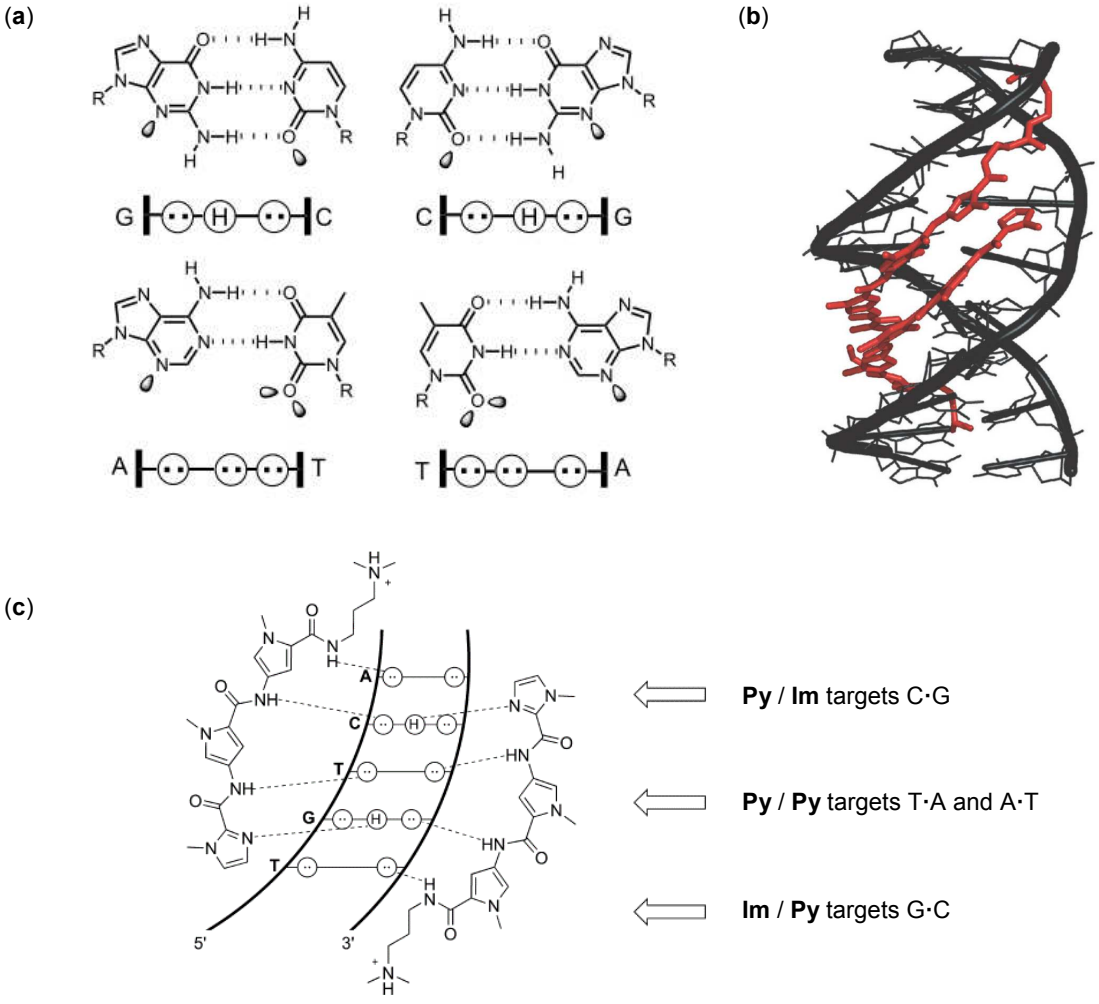


Figure 1-13: (a) Minor groove hydrogen-bonding patterns of Watson-Crick base pairs. Circles with dots represent lone pairs on N3 of purines and O2 of pyrimidines and circles containing H represent the N2 hydrogen of guanine. (Reprinted with permission from ref. 68. Copyright 2003 Elsevier Ltd). (b) Crystal structure of a polyamide dimer on a DNA decamer in red and black, respectively. (Reprinted with permission from ref. 64. Copyright 2013 Elsevier Ltd). (c) Pairing rules derived from 2:1 polyamide:DNA complexes. Potential hydrogen bonds between the

polyamides and the minor groove of DNA are represented with dotted lines. (Reproduced with permission from ref. 3. Copyright 2008 Oxford University Press).

If the two antiparallel polyamide strands are linked via a γ-aminobutyric acid (γ-turn) unit, molecules are obtained that bind with 100-fold greater affinity toward dsDNA than unlinked homodimers, with the γ-turn demonstrating selectivity for AT over GC base pairs.[71]

Nonetheless, minor groove binding polyamides display a variety of limitations. For example, Py/Py targets both AT and TA base pairs, which potentially decreases binding specificity and could lead to non-target binding interactions. The size of the target region that can be targeted is another limitation. Beyond five contiguous rings, the binding affinity of polyamides decreases dramatically[72] since the pitch of DNA is less than that of the polyamide dimers. As the polyamide lengthens, it becomes "overwound" with respect to the duplex, such that the shape of the polyamide is no longer complementary to DNA resulting in decreased affinity against longer target regions.[73] Although Py-Im polyamides that bind six base pairs of DNA were shown to inhibit transcription of a specific gene in cell culture[65], polyamides recognizing longer DNA sequences, e.g., 15-16 base pairs, would be expected to provide more specific biological activity. Similarly, certain DNA sequences, especially those including G/C tracts, have been challenging sites for hairpin polyamide recognition. High resolution crystal structures of different polyamide dimer-DNA complexes suggest that G/C containing complexes exhibit large negative propeller twist, resulting in shifted hydrogen bond locations relative to those shown in Figure 1-13, and decreased binding affinity.[73-75]

1.3.2 Overcoming Limitations

Structure-property studies on minor-groove binding polyamides have yielded three key advances. The first was identification of *N*-methyl-3-hydroxypyrrole (Hp) as a thymine-selective recognition element when paired across from Py (as shown in Figure 1-14). The Hp/Py modification allows for differentiation of the TA base pair relative to the other base pairs, whereas a Py/Hp modification allows differentiation of AT from other base pairs, albeit Hp polyamides bind with lower affinity than their Py counterparts.[76] The Hp modification thus provided a solution to the A/T degeneracy problem encountered when using Py/Py paired polyamides. Secondly, introduction of β-alanine (β) as a recognition element reduces the curvature mismatch between long polyamides and dsDNA targets. β-alanine has the same recognition profile as a Py subunit; however, its greater flexibility allows the polyamide to match the pitch and curvature of long stretches of DNA duplexes. In one case, a Py-, Im-, and β-containing polyamide was designed such that it recognized a 16 base pair mixed-sequence HIV-1 target site with affinity that resembled that of natural DNA-binding proteins, however the specificity was relatively poor.[77] The flexible β-modification has also allowed for high affinity targeting of DNA containing G/C tracts by allowing Im rings to orient better.[75] The third major advance was the use of diaminobutyric acid (DABA) in place of the γ-aminobutyric acid hairpin linker.[78] The linker, containing a cationic amino group, allowed for higher affinity recognition of DNA. The addition of an amino group also allows for functionalization of the polyamide conjugates. These conjugates, depending on their chirality, can impart improved biological function to hairpin polyamides such as improved cellular and nuclear uptake.[79]

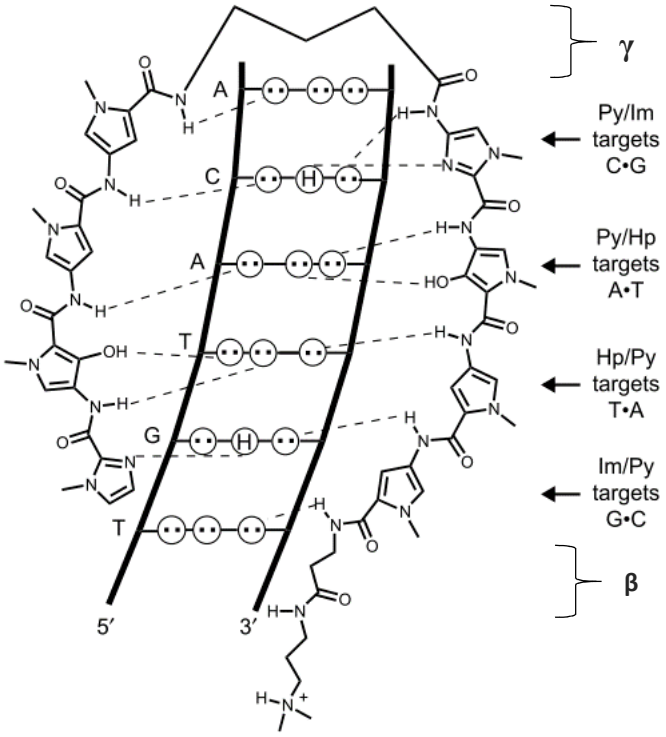


Figure 1-14: Binding model for the complex formed between ImHpPyPy-γ-ImHpPyPy-β-Dp and a 5′-TGTACA-3′ sequence. Putative hydrogen bonds are shown as dashed lines. (Adapted with permission from ref. 68. Copyright 2003 Elsevier Ltd).

1.4 Peptide Nucleic Acids (PNAs)

1.4.1 PNA Structure, Properties, and Limitations

Peptide nucleic acids (PNAs) are an interesting class of dsDNA-targeting agents developed by Nielsen and coworkers.[80] In PNAs, the entire sugar-phosphate backbone is replaced by an achiral *N*-(2-aminoethyl)glycine unit rendering these oligomers electrostatically neutral (Figure 1-15a). Fully modified PNAs, which are assembled using solid-phase peptide chemistry, bind to complementary DNA with high affinity and improved mismatch discrimination relative to

unmodified DNA reference strands.[81] In addition to targeting dsDNA via triplex formation, they can also target DNA via a range of duplex invasion mechanisms (Figure 1-15b).[82] The neutral backbone of PNA gives rise to desirable properties, such as excellent biostability[47], low toxicity, and the aforementioned high affinity to target DNA. However, the neutral backbone also introduces some drawbacks as well, such as poor aqueous solubility and cellular uptake. Other limitations include sequence limitations and the tendency for self-hybridization. Various PNA-based dsDNA targeting strategies and limitations will be discussed in further detail herein.

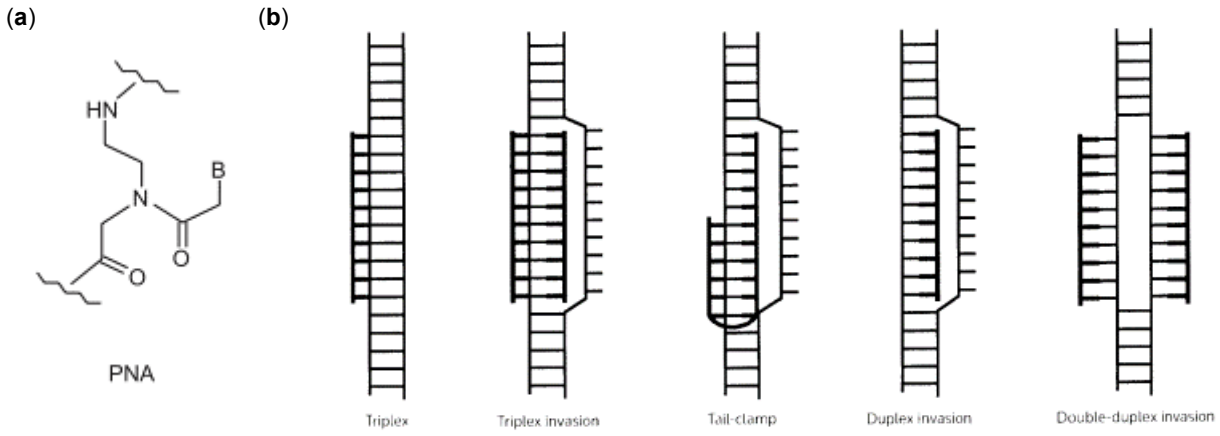


Figure 1-15: (a) Chemical structure of PNA and (b) modes of PNA directed antigene action. (Reprinted with permission from ref. 82. Copyright 1999 Elsevier Ltd).

1.4.2 Triplex Based Strategies Using PNAs

Homopyrimidine PNAs have been shown to bind to dsDNA via standard triplex formation, forming a DNA:DNA·PNA triplex (Fig. 1-15, triplex).[83] However, homopyrimidine PNA can also bind via triplex invasion resulting in a PNA:DNA·PNA triplex (Fig 1-15, triplex invasion) depending on PNA concentration, oligomer length, composition, and reaction time.[83] In this

binding mode, one strand of PNA forms a triplex via Hoogsteen binding to target dsDNA and then a second PNA strand invades the same DNA target region via Watson-Crick base pairing (Figure 1-15, triplex invasion).[84] Strand invasion is enhanced by joining the two PNA strands via a flexible linker (Figure 1-16a). The joining of the two strands reduces the entropic penalty upon binding the second PNA strand resulting in enhanced hybridization properties to DNA.[84,85r] These so-called bis-PNAs have been shown to recognize dsDNA fragments with different distinct complexes (Figure 1-16b) being possible depending on the PNA concentration.[86] A drawback to this approach is that the requirement for polypurine targets and protonation of PNA-C persist. This latter requirement can be alleviated by the use of pseudoisocytosine (pseudo-iC), which eliminates the pH sensitivity of binding and increases the binding affinity toward cDNA targets (Figure 1-16c).[85,86]

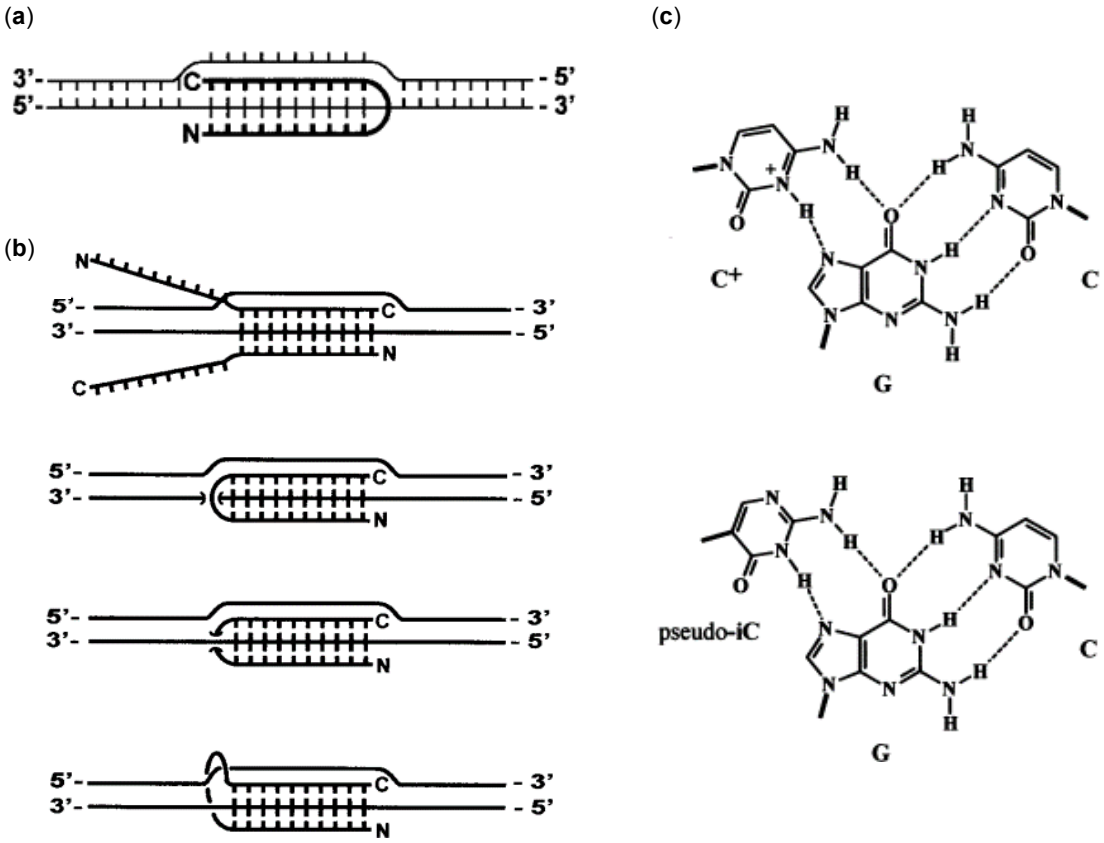


Figure 1-16: (a) Schematic of a triplex invasion complex displaying the internal PNA·DNA-PNA triplex and unbound DNA strand displaced. [Reprinted with permission from ref. 86. Copyright 2001 Elsevier Ltd). (b) Proposed structural isomers of triplex invasion complexes. (Reprinted with permission from ref. 86. Copyright 2001 Elsevier Ltd). (c) Watson-Crick and Hoogsteen bonding of C^+·GC and pseudo-iC·GC base pairs. (Reproduced with permission from ref. 85. Copyright 1995 Oxford University Press).

To alleviate some of the sequence-requirements and increase the versatility of strand invasion by PNAs, tail-clamp PNAs were developed. Tail-clamp PNAs combine a short TC-motif TFO (5-6 base pairs) with a mixed-sequence strand, which allows for invasion of DNA by a combined duplex/triplex strand invasion mode (Figure 1-17).[87] Thus, shorter polypurine regions are required, which increases the number of potential targets. These tail-clamp PNAs

have been used to inhibit transcription.[88] Although tail-clamp PNAs expand the range of sequences that are accessible to PNAs, they still have some sequence limitations, invasion is limited to promotor sequences or open transcription complexes, and nuclear uptake remains an issue.[88]

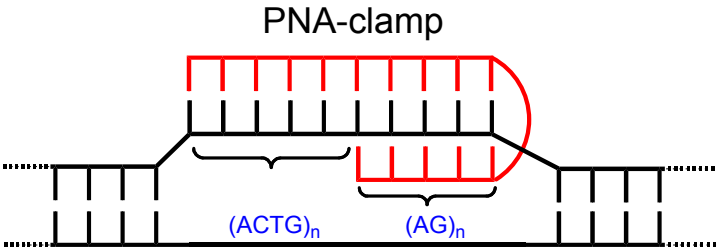


Figure 1-17: (a) Schematic representation of PNA tail-clamp binding to dsDNA.

1.4.3 Mixed-Sequence Targeting of DNA via Strand Invasion (agPNAs and γ-PNAs)

Probes that recognize mixed-sequence *B*-DNA via duplex invasion are particularly attractive due to the predictability of the Watson-Crick base-pairing rules. Early studies by Radding and coworkers on DNA recognition showed that ssDNA can bind to supercoiled DNA by hybridizing to one of the strands of a DNA duplex while displacing the other strand (creating a D-loop), in a process known as strand invasion (Figure 1-18).[90] This is a difficult process as strand invasion must have enough energy to overcome the preexisting Watson-Crick base pairs of the duplex to initiate recognition. This DNA recognition process is therefore slow, inefficient, and restricted to partially single-stranded segments of DNA (e.g., highly AT-rich sequences or cruciforms) [91,92] unless that probe strand is modified with affinity-enhancing building blocks. Rate constants for recognition of DNA can be enhanced up to 48000-fold by

attachment of cationic peptides as favorable electrostatic interactions are introduced[91], yet still only supercoiled DNA could be efficiently recognized.[92] Due to these difficulties, mixed-sequence single-stranded PNA (ssPNA) were targeted toward DNA complements and have been shown to have enough free energy for duplex invasion.[93-95] Strand-invasion by PNA is more efficient than strand invasion by DNA as the uncharged PNA backbone is not electrostatically repelled by the polyanionic backbone of the duplex target, resulting in faster association kinetics and slower dissociation kinetics. In one study, Xhang et. al. used mixed-sequence PNAs to target supercoiled DNA; however, recognition via strand invasion was only efficient at inverted repeats or AT rich sites.[95] In the same study, PNA-peptide conjugates were shown to invade DNA at sequences that lack inverted repeats, including G/C rich sites indicating that invasion efficiency can be enhanced by addition of cationic peptides.[95] However, the invasion complex is unstable because a single-stranded portion remains unpaired and binding efficiency is lowered at increasing monovalent and divalent cation concentrations.

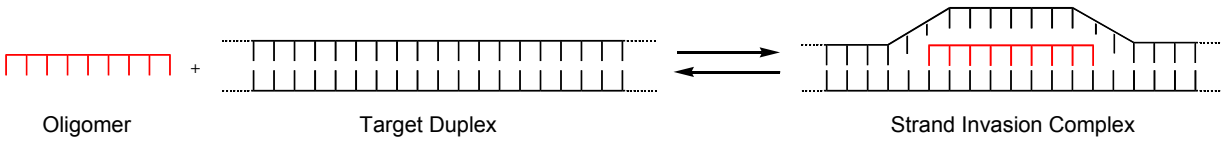


Figure 1-18: Schematic of oligomer-mediated single-duplex invasion into dsDNA.

Until recently, DNA recognition via the single duplex invasion mechanism shown in Figure 1-18 received little attention. However, encouraging results with the so-called γ-PNAs have reignited interest back into this dsDNA recognition approach. Regular single-stranded PNAs do not have a well-defined conformation due to their achiral nature.[96-99] By analogy to oligonucleotides[100] it was reasoned that stronger DNA affinity could be attained if

conformationally biased or restricted PNA monomers were used. Suitably preorganized PNA would require minimal structural reorganization prior to complexation with DNA targets, translating into more favorable binding energetics and kinetics.[100] One of the most interesting PNA backbone modifications features a methyl group at the γ-position of the *N*-aminoethyl glycine backbone (γ-PNA, Figure 1-19). The chiral single-stranded γ-PNA folds into a right-handed helical structure.[101] (*S*)-γ-methyl-PNAs display exceptional affinity for DNA and have been demonstrated to recognize mixed-sequence *B*-DNA in a site-specific manner through direct Watson-Crick base pairing via single duplex invasion.[102] Conjugation of γ-PNA to DNA intercalating moieties, such as acridine, further increases their affinity toward DNA while maintaining site-specific strand invasion into *B*-DNA.[103] This intercalation likely minimizes PNA-DNA end fraying, providing the necessary binding energy for the PNA to fully invade the DNA sequence. However, like other PNA binding modes, DNA strand invasion by γ-PNAs is most efficient at low ionic strengths. When physiological salt conditions are used, little to no recognition of *B*-DNA is observed.[104] It was not clear from these studies whether the lack of binding was due to target inaccessibility (lack of base pair opening) or lack of binding free energy.

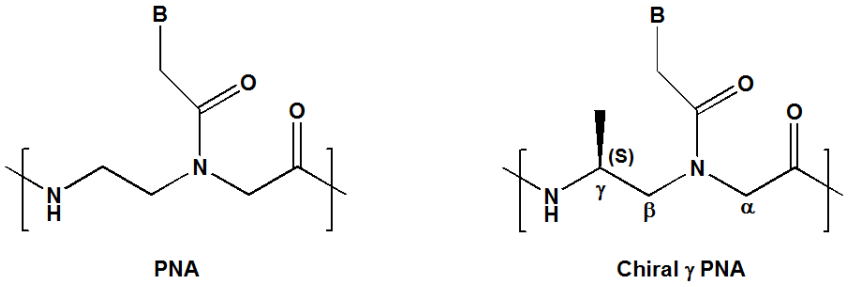


Figure 1-19: Structures of PNA and Chiral γ-PNA

Ly and coworkers also studied G-clamp γ-PNAs.[104,105] A G-clamp is produced by replacing the cytosine nucleobase with 9-(2-guanidinoethoxy) phenoxazine, a cytosine analog that can form five hydrogen bonds with guanine in addition to providing extra base stacking as a result of the expanded phenoxazine ring[106] (Figure 1-20). A single cytosine replacement with a G-clamp in a PNA-DNA duplex results in dramatically increased thermal stability by as much at 23 °C.[107,108]

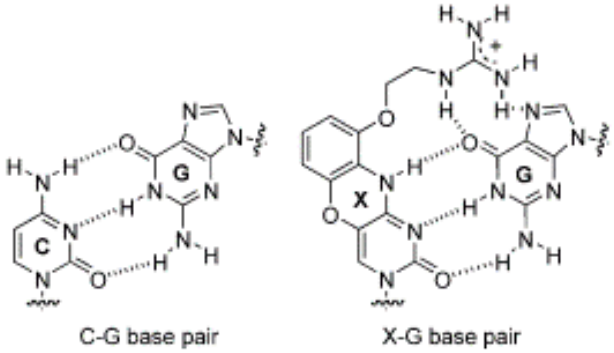


Figure 1-20: Chemical structure of C-G and X-G base pairs. X = 9-(2-guanidinoethoxy) phenoxazine (G-clamp). (Reprinted with permission from ref. 104. Copyright 2008 John Wiley and Sons).

Under low salt conditions, the conditions in which the majority of PNAs have shown efficient DNA recognition, single-stranded G-clamp modified γ-PNAs invade mixed-sequence 171-mer dsDNA.[104] γ-PNAs with higher G-clamp content result in increased invasion efficiency. However, under simulated physiological conditions (10 mM sodium phosphate, 2 mM $MgCl_2$, 150 mM KCl, pH = 7.4) very little strand invasion was observed when two G-clamps were used in a mixed-sequence 10-mer γ-PNA.[104] Addition of a third G-clamp into the 10-mer γ-PNA, resulted in high invasion efficiency under simulated physiological conditions. However, at higher PNA concentrations, aggregation and non-specific binding of the G-clamp modified γ-

PNA strand was observed.[105] These limitations can be alleviated by attachment of hydrophilic groups such as diethylene glycol to the γ-position of the PNA backbone (miniPEG-γPNA)[109] or $_{L}$-serine derived γ-PNA.[101] These results suggest that the DNA duplex is sufficiently dynamic to permit strand invasion provided the binding free energy can be met to displace the native cDNA strand.[105]

1.4.4 pcPNAs

Probes that recognize mixed-sequence *B*-DNA via double-duplex invasion are particularly appealing because the D-loop also would be engaged in base pairing, potentially leading to stabilization of recognition complexes. Developed by Gamper and coworkers in 1996, pseudo-complementary (pc) DNA are double-stranded probes that consist of 2′-deoxy-2-aminoadenosine (D) and 2′-deoxy-2-thiouracil (S) in place of 2′-deoxy-adenosine (A) and 2′-deoxy-thymine (T), respectively.[110] Unlike A:T base pairs, which form two hydrogen bonds, D:S pairs are destabilized due to a steric clash between the 2-amino group of D and the 2-thio group of S. However, D:T and S:A form stable base pairs, with the D:T base pair being particularly stable as three hydrogen bonds are formed (Figure 1-21). Therefore, complementary pcDNA strands have reduced affinity for one another but higher affinity for unmodified complementary DNA. These pcDNAs have been shown to invade the termini of homologous duplexes to form stable three-arm junctions under conditions where unmodified DNAs failed.[110]

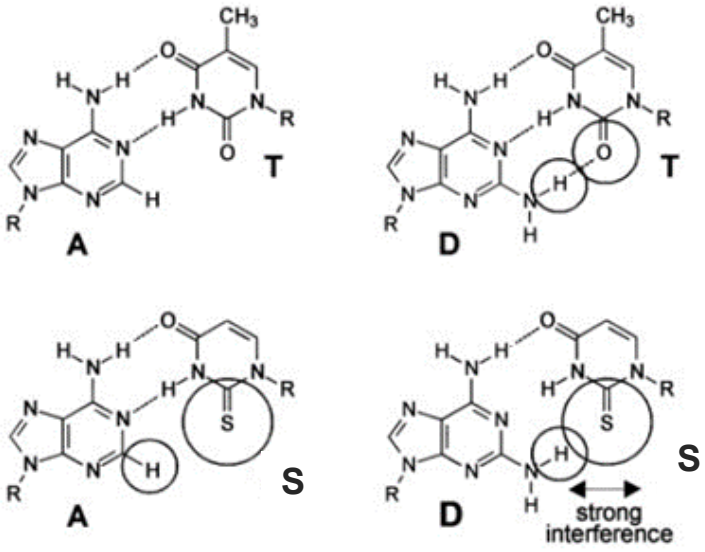


Figure 1-21: Watson-Crick base pairs between adenine (A) or 2-aminoadenine (D) and thymine (T) or 2-thiouracil (S). (Adapted with permission from ref. 111. Copyright 2009 Oxford University Press).

This concept has been extended into PNAs, and pcPNAs[112] have been shown to enable recognition of internal regions of mixed-sequence dsDNA via double-duplex invasion (Figure 1-15b, double-duplex invasion)[113], including breathed plasmid DNA.[84] pcPNAs have been shown to block binding of proteins to DNA.[112,114] Glazer and coworkers stimulated homologous recombination in mammalian cells by targeted correction of a thalassemia-associated β-globin mutation by pcPNAs.[111] Homologous recombination with pcPNAs results in 3- to 5-fold higher recombination frequencies than either TFOs or bis-PNAs. In another study, Komiyama and coworkers synthesized artificial nucleases using pcPNAs conjugated to terminal monophosphate or iminodiacetate groups attached to the terminus. Upon addition of Ce(IV)/EDTA, site-selective hydrolysis of the dsDNA target was observed, including site-selective scission of genomic DNA in *E. Coli* (4.6 Mbp).[115] The use of pcPNAs alleviates the

self-hybridization potential between PNA strands and relieves the sequence-restrictions for other triplex-based PNA invasion techniques. However, self-inhibitory effects observed at high pcPNA concentrations, slow invasion kinetics, and requirement for low ionic strengths and non-physiological temperatures impose limitations on pcPNA-mediated strand invasion in biological media.[116] Moreover, the target sites must have an A/T content of at least 40% for efficient invasion to take place, due to the absence of good G/C pseudocomplementary base pairs.[112] Sumaoka and Komiyama have recently shown pcPNA mediated double-duplex invasion at physiologically relevant ionic strengths and temperatures is efficient if recognition experiments are conducted at "molecular crowding effects" conditions (i.e. containing poly(ethylene glycol) as a molecular crowder to simulate intracellular environments).[117] These results present a chemical basis for the intracellular activity of pcPNAs in living cells previously reported by Glazer and coworkers.[111]

1.5 Zorro LNA and bis-LNA

Smith and coworkers introduced a novel DNA-targeting approach utilizing a LNA-modified double-stranded oligonucleotide construct termed Zorro LNA.[118,119] This approach involves two partially complementary LNA/DNA mixmer oligonucleotides, which are connected to each other through a seven base pair duplex region having long 3′-overhangs. Zorro LNAs sequence-specifically invade dsDNA via the 3′-overhangs, which bind to either strand of the DNA. One 3′-overhang binds mixed-sequence DNA targets via Watson-Crick binding while the other binds to a purine-rich strand as a TC-motif TFO resulting in the alleged "Z" shaped invasion complex (Figure 1-22). The Zorro LNA construct was shown to invade plasmid DNA and efficiently block RNA-polymerase dependent gene transcription in mammalian cells in a

sequence-specific manner.[18,119] Limitations to this strategy are the requirement for TFO binding sites and the potential for self-hybridization with highly LNA-modified strands. Moreover, successful gene regulation via Zorro LNAs requires two sequential Zorro LNA binding sites as a target, which limits the use of this targeting strategy. The reasoning behind the latter requirement is not fully understood.

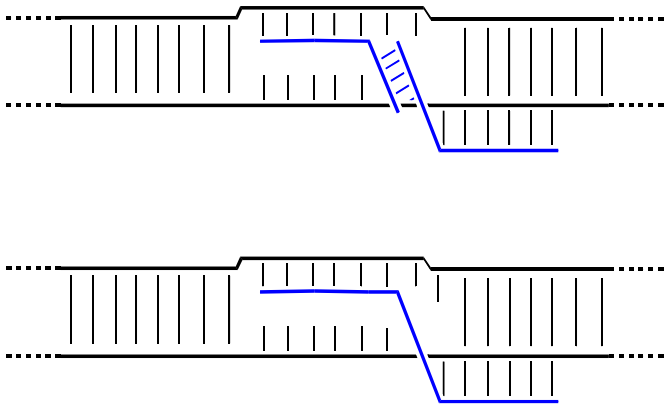


Figure 1-22: Strand-Invasion complex by Zorro-LNA (top) and ssZorro LNA (bottom).

In 2011, Smith and coworkers further optimized the double-stranded Zorro LNA construct to single-stranded Zorro LNAs (ssZorro), which displayed enhanced strand invasion into duplex DNA. The first generation Zorro LNAs were cumbersome due to intra-molecular binding between bases in the arms and bases in the linker region. They also required a pre-annealing step for hybridization of the two ONs in order to create the Zorro construct. Therefore, to make the process more straightforward for a simple and predictable ON-based therapeutic, the ssZorro was introduced. Replacement of the double-stranded linker with various covalent linkers resulted in improved dsDNA invasion efficiency and kinetics.[120] It was found that linker length and linker type, with the exception for detrimental hydrophobic long alkyl linkers or

aromatic linkers, have no significant effect on invasion efficiency and kinetics of ssZorro-LNAs.

In a similar binding mechanism, bisLNAs have been introduced and shown to invade dsDNA via CT-motif triplex binding and duplex invasion.[121] Bis-LNAs resemble bis-PNAs (bis-LNA-A, Figure 1-23) and tail-clamp PNAs (bis-LNA-B, Figure 1-23), but contain DNA/LNA mixmers such as those in Zorro LNAs. Tail-clamp bis-LNA-B was shown to invade supercoiled duplex DNA under physiological pH and ion concentrations whereas bis-LNA-A was much less efficient, indicating the Watson-crick tail is essential for strong binding.[121] Shortening the tail results in less efficient invasion, whereas shortening the triplex binding area (by three nucleotides) can increase invasion/binding efficiency, while removing the triplex binding region entirely results in very mild invasion of supercoiled DNA and no invasion of linearized DNA. This DNA recognition approach falls to the limitations of requiring a short polypurine tract for triplex binding, has shown non-specific binding interactions at high probe concentrations, and targeting linearized DNA remains a challenge. Therefore, further improvement on bis-LNA design is needed to expand the use of this targeting strategy.

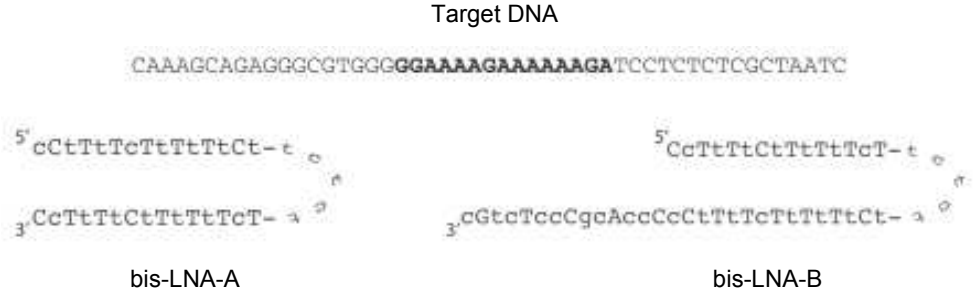


Figure 1-23: Sequence of promoter region of target DNA sequence. In bold is highlighted the 15-mer target-site (top). Also shown are the bis-LNA construct bis-LNA-A and the Watson-

Crick arm extended bis-LNA-B (bottom). Uppercase letters denote DNA whereas lowercase letters denote LNA modifications. (Adapted with permission from ref. 121. Copyright 2013 Oxford University Press).

1.6 ZFNs, TALENs and CRISPR-Cas Genome Engineering Tools

There are many site-specific dsDNA-binding proteins that are essential for the biological functions of DNA. These DNA-binding proteins include transcription factors, repair proteins, and DNA maintenance enzymes (nucleases, topoisomerases, helicases), all of which are encoded for in the human genome.[122] The high affinity and sequence specificity of natural DNA-binding proteins to their dsDNA targets has spurred the development of a highly interesting class of genome engineering tools, which include artificial Zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regulatory interspaced short palindromic repeat (CRISPR)/Cas based RNA-guided DNA endonucleases. These chimeric nucleases are composed of programmable, sequence-specific DNA binding modules linked to a non-specific DNA cleavable domain.

Zinc-finger domains are the largest family of DNA binding domains in the human genome (Figure 1-24).[122] Each domain is a small peptide consisting of 30 amino acid residues folded into a structure that is stabilized by chelation of a zinc ion to two cysteine and two histidine residues.[123] The specificity of the zinc finger is based on the amino acid sequence in the α-helix region of the residue, which scans along the major groove of DNA duplexes and typically binds via contact to 3 bases. Zinc-finger domains have been developed that target nearly all of the 64 possible triplet regions, with all 5′-GNN-3′, most 5′-ANN-3′ and 5′-CNN-3′, and some 5′-TNN-

3′ type sequences fully identified.[124] These domains can be combined into arrays to form multi-zinc-finger structures that can site-specifically recognize longer DNA fragments. Engineered zinc-fingers are often conjugated to a DNA cleavage domain to generate zinc-finger nucleases, which can interfere with transcription factor binding or introduce site-specific double-stranded breaks (DSBs) to DNA, and thereby manipulate the genome of mammalian cells.[125] ZFN-induced DSBs may induce site-specific mutagenesis by recruiting natural DNA repair enzymes resulting in homologous recombination in the presence of donor DNA, or non-homologous end joining. This methodology has been used for replacement of a defective gene with a normal allele at its natural chromosomal location via homologous recombination.[126] However, a limitation of this methodology is it can induce significant toxicity due to excessive cleavage, presumably, at unintended sites, although modification of the cleavage domain can reduce toxicity.[125,127,128] Other limitations include the highly variable efficiency of gene targeting in different cell types, which may be due to different target accessibility and differences in the efficiency of ZFN and donor DNA cellular delivery. In the majority of cases, the ZFN and donor DNA are introduced as viral vectors, with the choice of vector being dependent on the cell or tissue being studied. In other cases, the exogenous DNA coding for the nuclease is taken up quite efficiently by the cell, with ZFN-induced homologous recombination taking place at a wide range of frequencies depending on cell type.[129]

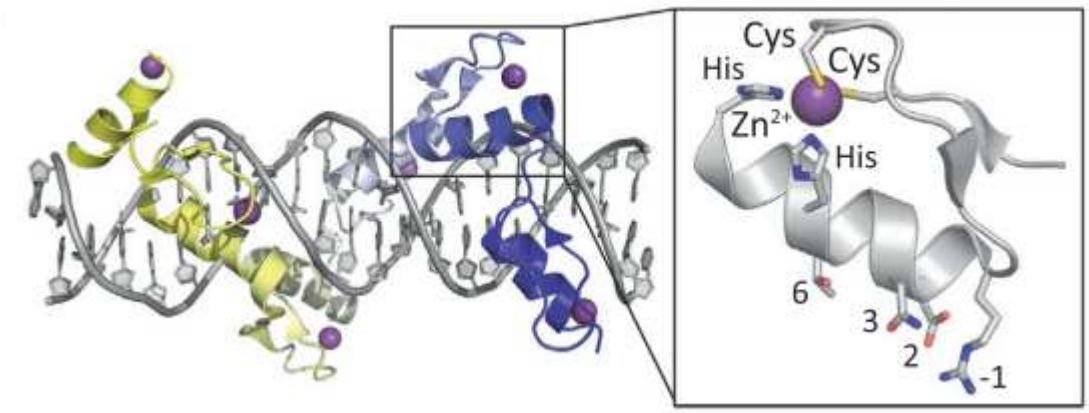


Figure 1-24: Designed zinc-finger protein in complex with target DNA (grey). Each zinc-finger consists of approximately 30 amino acids in a ββα arrangement (inset). Surface residues (-1, 2, 3, and 6) contacting DNA are shown as sticks. Each zinc-finger domain contacts 3 or 4 bp in the major groove of DNA. The side chains of the conserved Cys and His residues are depicted as sticks in complex with a Zn^{2+} ion (purple). (Reprinted with permission from ref. 129. Copyright 2013 Elsevier Ltd).

Transcription activator-like effectors (TALEs) are naturally occurring proteins in plant pathogenic bacteria of the genus *Xanthomonas*, which contain DNA binding domains composed of a series of 33-35 amino acid repeat domains that recognize a single base pair via repeat-variable diresidues (RVDs) located at the 12 and 13 positions (Figure 1-26).[130,131] Specific RVDs have been identified for contacting A, C, and T nucleobases via hydrogen bonding or van der Waals interactions, but lack of a residue that specifically recognizes G limits the broader applications of TALEs.[132,133] Like zinc-finger proteins, TALE repeats can be linked together to recognize contiguous DNA sequences and when linked to a nuclease (TALEN) can cleave target DNA.[134,135] In recent years, TALENs have be engineered to bind practically any desired DNA sequence.[129] When TALENs are introduced into cells, they can be used for

genome editing via introduction of double-stranded breaks. Similar to ZFNs, after the TALEN induces a DSB in the target DNA, cells respond with repair mechanisms resulting in gene corrections via non-homology end joining or, if donor DNA is introduced, homology-directed repair. A limitation to this approach is insufficient specificity of the TALENs, as they can cleave non-target sites resulting in multiple DSBs that can overwhelm the cells natural repair machinery resulting in chromosomal damage and cell death.[135] TALENs also have a minor sequence limitation, inasmuch as the TALE binding sites must start with a T nucleotide.[129]

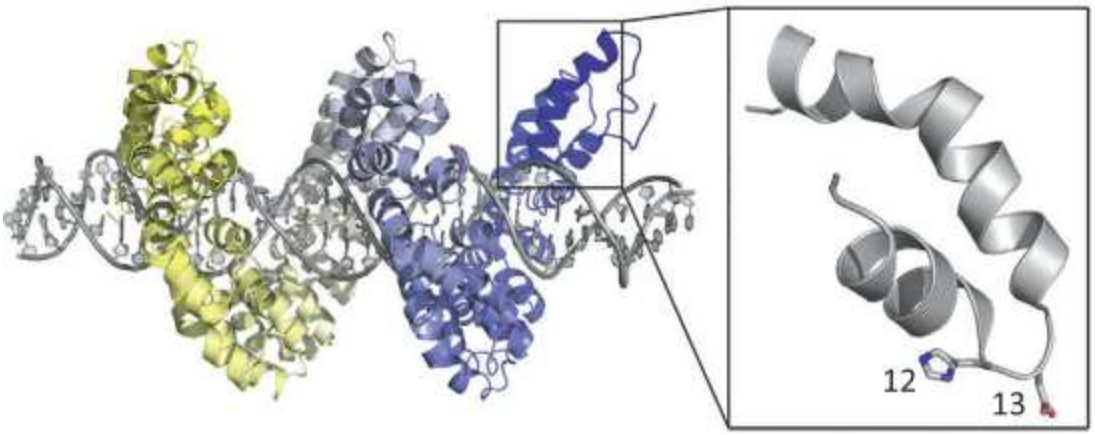


Figure 1-25: TALE protein in contact with target dsDNA (gray). The two repeat-variable diresidues are shown as sticks in the inset. (Reprinted with permission from ref. 129. Copyright 2013 Elsevier Ltd).

The newest addition to the class of engineered DNA-binding proteins is the clustered regulatory interspaced short palindromic repeat (CRISPR)/Cas system. This system was recently discovered in bacteria and most archaea and is involved in an interference pathway that protects the host cells from invasive species and provides a form of acquired immunity.[136,137] CRISPRs are short repeats of DNA sequences separated by short spacer fragments of DNA (Figure 1-26).

The spacer fragments are presumably taken up from invasive elements and incorporated into the CRISPR system resulting in DNA-encoded immunity.[r136,138] CRISPR loci range in size from 21-48 base pair repeats separated by spacer sequences of similar length. The spacer sequences encode for CRISPR-RNAs (crRNA) and repeat sequences encode for guide RNAs which are complexed with CRISPR-associated (Cas) proteins, which silence foreign DNA via RNA-guided DNA cleavage with Cas nucleases. The repeat sequences guide the RNA, whereas the spacer sequences, derived from foreign DNA species, are the sequences responsible for the defensive capacity of CRISPRs in bacteria and archaea (i.e. by binding to matching invasive DNA from which they were derived). Designer CRISPR/Cas systems have been developed to harness the capabilities of prokaryotic CRISPR/Cas systems for introducing DSBs into target mammalian DNA, and have been introduced into zebrafish[139], bacterial cells[140], and human cells[141-43] via plasmids encoding the Cas9 endonuclease and containing spacer sequences that are complementary to target DNA sites. The designer CRISPR/Cas systems have been shown to successfully interrupt gene function for gene editing purposes (Figure 1-26). However, there are concerns about specificity as spacers containing mutations retained their activity on their mismatched targets.[142] This approach requires more exhaustive studies to determine the specificity of crRNAs to their target DNA and to discover an effective delivery method for human trials.

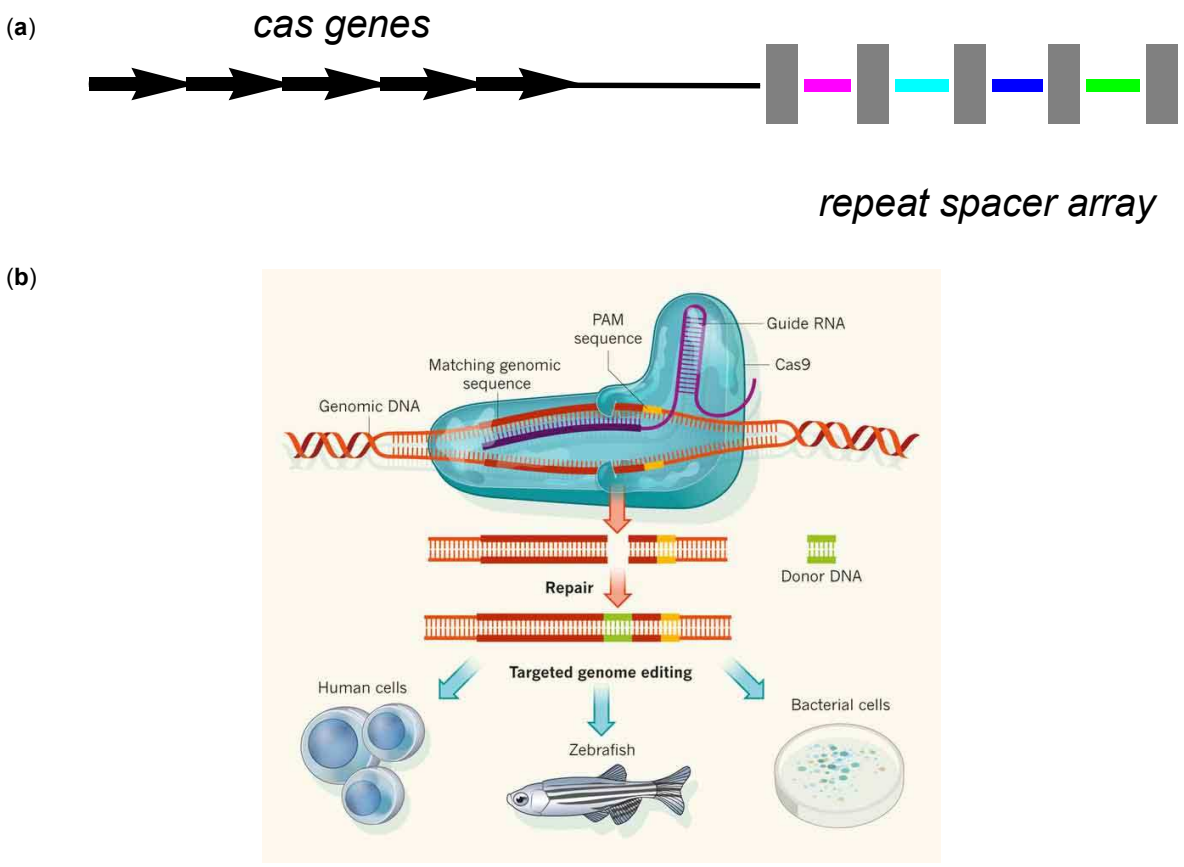


Figure 1-26: (a) Simplified diagram of a CRISPR locus. Repeats are shown as grey boxes and spacers as colored bars. (b) Mechanism of genomic editing using the CRISPR/Cas machinery. (Reprinted with permission from ref. 144. Copyright 2013 Nature Publishing Group).

1.7 Intercalator Modified Duplex Invasion Probes

1.7.1 Invader Probes

Hrdlicka and coworkers introduced Invader probes as a fundamentally different approach toward mixed-sequence recognition of dsDNA.[145] These probes recognize dsDNA via energetically activated duplexes modified with intercalator-functionalized nucleotides (for an illustration of the concept, see Figure 1-27). These double-stranded probes are inherently

activated for recognition of dsDNA due to the +1 zipper arrangement of the intercalator-functionalized nucleotides. This structural motif induces a duplex destabilization as the intercalators are forced into the same region causing violation of the 'nearest neighbor exclusion principle'[146], which states that intercalators, at most, bind to every second base pair of a DNA duplex due to limits in helix expandability, i.e., local duplex unwinding resulting in an "energetic hotspot". On the other hand, the two strands that constitute the Invader probe exhibit exceptional affinity toward complementary DNA, as they are able to site-specifically position intercalators into the duplex core resulting in efficient π-π-stacking with neighboring nucleobases. The advantages of this approach are a) the strand invasion potential can be tuned by varying the number of hotspots in the invader strand, b) there are no inherent sequence limitations, and c) the recognition process can take place under physiological conditions.

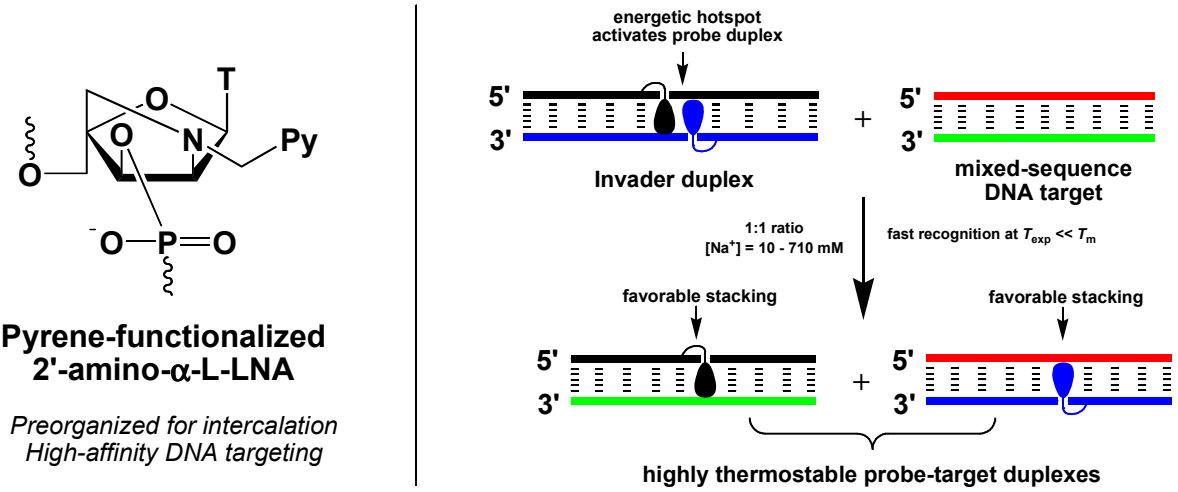


Figure 1-27: Illustration of Invader concept for recognition of mixed-sequence dsDNA. Droplets denote pyrene intercalating moiety.

The stability difference between Invader probes and dsDNA target regions on one side, and probe-target duplexes on the other side, provides the energy to drive recognition of

isosequential dsDNA target hairpins[147-150] and chromosomal DNA targets[151] which will be discussed in more detail in the following sections.

1.7.1.1 First Generation Invader Building Blocks

Initially, N2′-pyrene-functionalized LNAs were used as the key activating components of Invader probes (Figure 1-27). The synthetic route to the corresponding phosphoramidites is challenging, involving ~20 synthetic steps and overall reaction yields of <3%.[152,153] Numerous N2′-functionalized-2′-amino-α-L-LNA monomers with different linker chemistries were developed via this route (Figure 1-28).

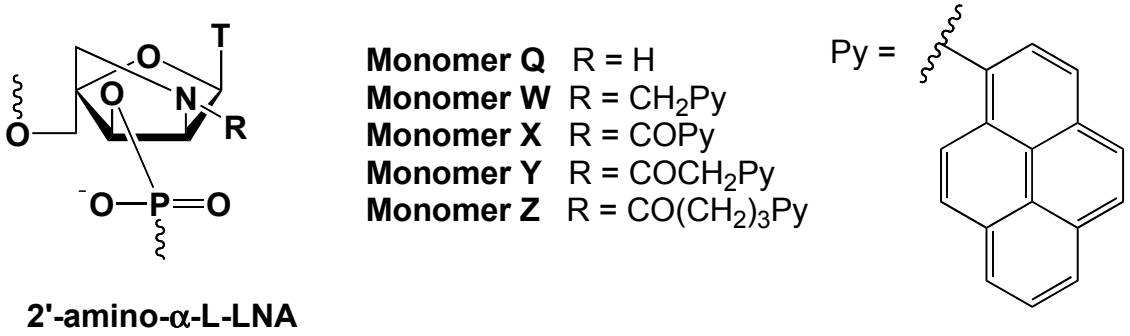


Figure 1-28: Structures of 2′-*N*-functionalized-2′-amino-α-L-LNA-T monomers.

Upon incorporation of monomers **Q**-**Z** into mixed-sequence 9-mer ONs, the effect on thermal stability toward DNA and RNA targets was evaluated by UV thermal denaturation melting experiments and compared to unmodified reference duplexes. Monomers containing pyrene intercalating moieties resulted in exceptional affinity toward cDNA targets, whereas monomer **Q** which does not contain an intercalator, resulted in similar affinity against DNA complements as unmodified DNA strands (ΔT_m from -2.0 to +2.5 °C) indicating that addition of a 2′-amino-α-L-LNA skeleton is not stabilizing by itself. However, incorporation of PyMe/PyCO/PyAc-

functionalized 2′-amino-α-L-LNA monomers **W**/**X**/**Y** results in dramatic increases in stability of the modified duplexes with cDNA relative to unmodified reference duplexes (ΔT_m from +6.5 to +19.5 °C). More moderate increases in stability were observed for PyBu monomer **Z** (ΔT_m up to +6.5 °C). The observed affinity trends toward DNA complements (**X**>**Y**>**W**>>**Z**) suggest that 1) carbonyl linkers are preferred over alkyl linkers of the same length (**X**>**W**) and 2) shorter linkers are preferred over longer linkers (**X**>**Y**>>**Z**).

Next, the effect on RNA affinity of strands modified with monomers **Q**-**Z** was evaluated. Incorporation of monomer **Q** showed slightly increased thermal affinity to cRNA (ΔT_m from +1.0 to +4.5 °C) whereas a single incorporation of N2′-pyrene-functionalized-2′-amino-α-L-LNA monomers **W**-**Z** resulted in much lower affinity increases toward cRNA targets than toward cDNA targets (ΔT_m vs. cRNA, from -5.0 to +12.0 °C).[153] Accordingly, N2′-pyrene-functionalized 2′-amino-α-L-LNAs exhibit a marked DNA selectivity, i.e., a positive $\Delta\Delta T_m$/mod (DNA-RNA) = ΔT_m/mod (DNA) - ΔT_m/mod (RNA), especially with PyMe and PyCO monomers **W** and **X** ($\Delta\Delta T_m$/mod (DNA-RNA) between 6 and 9 °C). This indicates that monomers with shorter linkers are preferred over monomers with longer linker for DNA targeting applications.

The pronounced DNA selectivity of N2′-pyrene-functionalized 2′-amino-α-L-LNAs **W**, **X**, and **Y** suggests that pyrene intercalation is a likely binding mode. To corroborate this hypothesis, additional UV-absorption experiments were performed. UV-absorption in the 320-360 nm range of singly-modified strands with monomers **W**, **X**, and **Y** show moderate to large bathochromic shifts upon binding to cDNA strands ($\Delta\lambda_{max}$ = 1 - 5 nm) whereas smaller increases are generally observed vs cRNA strands ($\Delta\lambda_{max}$ = 0 - 6 nm).[153] Similarly, the shifts in absorbance showed pronounced hypochromic shifts upon binding to cDNA and cRNA targets

indicating strong electronic interactions between the pyrene and the nucleobases of the duplex and accordingly the pyrene moieties are assumed to be intercalated into the duplex.

Force-field calculations were carried out on duplexes modified with various N2′-pyrene-functionalized 2′-amino-α-L-LNAs. In agreement with biophysical data, the lowest energy structure of the duplex between 2′-N-(pyren-1-yl)methyl-2′-amino-α-L-LNA-T and cDNA suggests that the pyrene residue is intercalating into the duplex core (Figure 1-29a). The pyrene engages in extensive π-π stacking with the neighboring nucleobases resulting in the highly stabilized duplexes discussed previously. Closer scrutiny of the molecular arrangement of these moieties in DNA duplexes suggest that the nucleobase and the pyrene are efficiently locked relative to each other (Figure 1-29b) as a consequence of the conformationally restricted 2-oxo-5-azabicyclo[2.2.1]heptane skeleton and the short linker between the pyrene and the sugar skeleton.[153] This precise positioning of the intercalator into the duplex core has many potential interesting applications within nucleic acid based diagnostics and therapeutics.

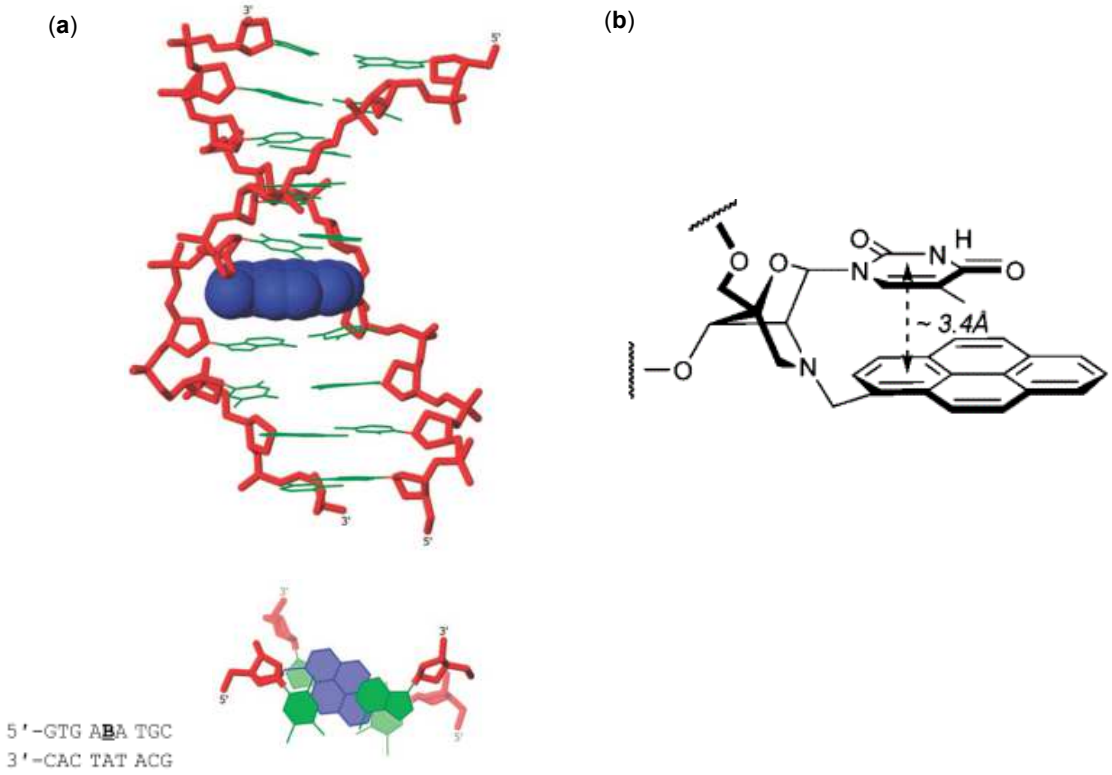


Figure 1-29: (a) Lowest energy structure of a 9-mer mixed-sequence DNA duplex containing a single modification of 2′-*N*-(pyren-1-yl)methyl-2′-amino-α-L-LNA-T. Upper: side view of duplex; Lower: top view of central duplex region. (b) Illustration of direct positioning of the pyrene moiety in the duplex core by 2′-N-(pyren-1-yl)methyl-2′-amino-α-L-LNA-T. Color code: sugar phosphate backbone (red); pyren-1-ylmethyl moiety (blue); nucleobases (green). H atoms, Na +ions, and bond orders have been omitted for clarity. The sequence of the duplex is given in the lower left hand corner; B = site of modification. (Reprinted with permission from ref. 153. Copyright 2009 American Chemical Society).

DNA duplexes with a +1 interstrand zipper arrangement of N2′-pyrene-functionalized-2′-amino-α-L-LNA modifications (lower left hand corner of Figure 1-30) display lower T_m's due to formation of an energetic hotspot.[148] This is presumably due to a high local density of

intercalators violating the 'nearest neighbor exclusion principle' resulting in a locally perturbed region in the DNA duplex.[146] This hypothesis was supported by fluorescence spectroscopy experiments which resulted in increased pyrene fluorescence and excimer formation at $\lambda \sim 490$ nm relative to probe-target complexes indicating the two pyrenes are in close proximity to one another.[148] Force-field calculations (Figure 1-30) also indicate destabilization is due to close proximity of intercalated pyrene moieties resulting in molecular crowding, duplex extension, and unwinding.[148] This model accounts for the observed excimer emission and labile nature of the Invader probes.

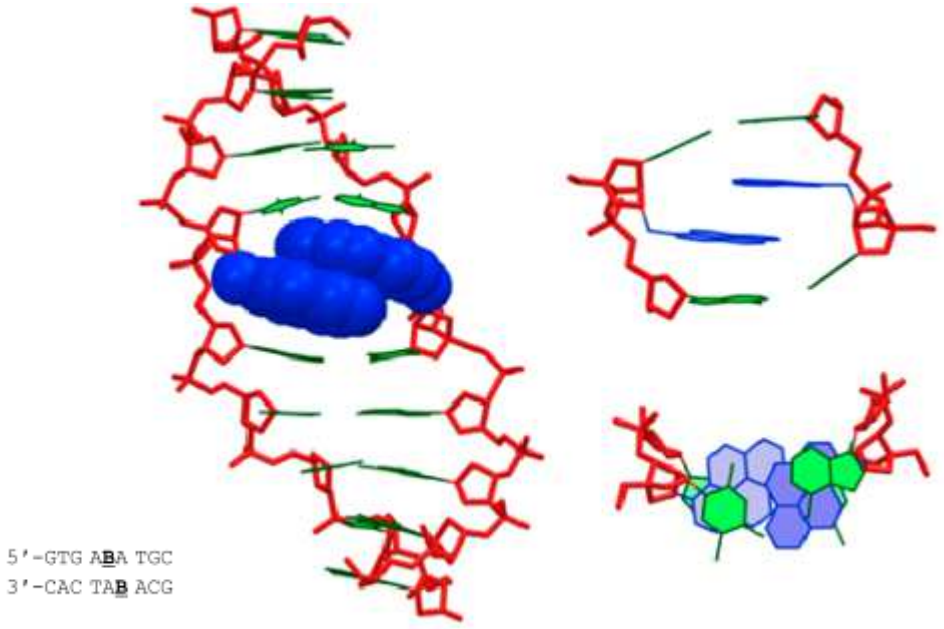


Figure 1-30: Lowest energy structure of 9-mer +1 interstrand zipper duplex modified with 2′-*N*-(pyren-1-yl)methyl-2′-amino-α-L-LNA-T. Left: side view of the duplex. Upper right: alternative representation of the central duplex region. Bottom right: top view of the central duplex region. Color code: sugar phosphate backbone (red); pyren-1-ylmethyl moieties (blue); nucleobases (green). H atoms, Na +ions, and bond orders have been omitted for clarity. The

sequence of the duplex is given in the lower left hand corner; B = site of modification. (Reprinted with permission from ref. 148. Copyright 2013 American Chemical Society).

To further characterize these probes, the dsDNA recognition characteristics of double-stranded probes with interstrand arrangements of pyrene-functionalized monomers was evaluated via analysis of thermodyanamic parameters, a fluorescence-based recognition assay, and a hairpin invasion assay. Evaluation of binding free energies for duplex formation, ΔG, indicates that +1 zipper probes are energetically activated relative to probe-target duplexes.[147] From these values, ΔG_{rec} is calculated and used for approximating the dsDNA recognition potential of Invader probes (ΔG_{rec} = ΔG(probe strand 1:ssDNA) + ΔG(probe strand 2:ssDNA) - ΔG(dsDNA target) - ΔG(Invader probe duplex)). ΔG_{rec} takes into account the energy of the probe, the energy of the target region, and the energy of the probe-target duplexes. The more negative the value, the higher the potential for binding to target dsDNA region. Highly negative values are observed for all +1 interstrand zippers modified with pyrene-functionalized Invader LNAs indicating that the probe is activated for recognition of dsDNA.[147] Probes of longer sequences with more energetic hotspots display the highest potential for dsDNA targeting.

A fluorescence-based assay was originally used to experimentally test the dsDNA recognition characteristics of Invader LNAs. This assay is based on monitoring the disappearance of the pyrene excimer emission signal of the +1 interstrand zipper probe as a function of time upon incubation with its target DNA.[r] Analysis of the time at which 50% and 75% excimer signal was lost was used to assess the recognition kinetics of DNA invasion with Invader LNA probes. Invader probes containing two consecutive positioned hotspots resulted in faster kinetics than Invaders with one hotspot or two separated hotspots (4 bp separation).[147]

To evaluate the dsDNA recognition potential of Invader LNAs in a more challenging assay, an electrophoretic mobility shift assay was developed. A digoxigenin (DIG)-labeled DNA hairpin comprised of a 9-mer double-stranded stem linked via a T_{10} loop served as a model dsDNA target (Figure 1-31). Incubation of N2′-pyrene-functionalized 2′-amino-α-L-LNA Invader probes with the DNA hairpin under physiological conditions resulted in dose-dependent formation of slower moving recognition complexes.[148] Between 38-48% recognition was observed with 100-fold molar excess of N2′-pyrene-functionalized 2′-amino-α-L-LNA Invader probes relative to the hairpin target.[148] These results indicate that the use of Invader LNA activated probes in a +1 interstrand arrangement is a promising strategy for mixed-sequence dsDNA recognition via a dual duplex invasion mechanism that relies on differences in thermostability between probes and probe-target duplexes.

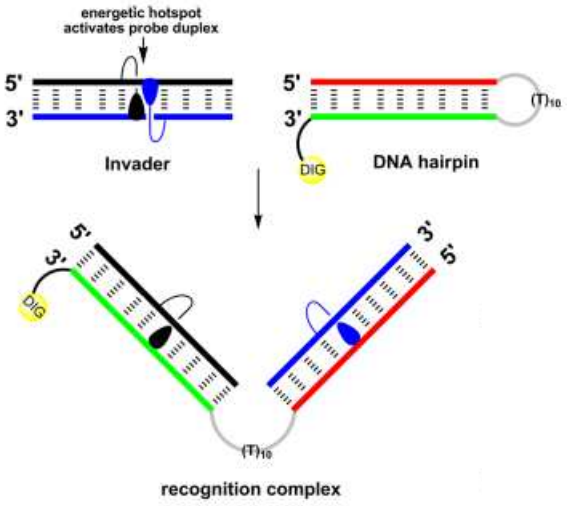


Figure 1-31: Illustration of dsDNA-recognition process by invaders using an electrophoretic mobility shift assay. DIG = digoxigenin.

1.7.1.2 Second Generation Invader Building Blocks

Due to the challenging and lengthy protocols required for synthesizing N2′-pyrene-functionalized 2′-amino-α-L-LNA building blocks, alternative nucleosides were pursued that are more readily accessible, while retaining the dsDNA-recognition characteristics of the 1st generation Invader building blocks. Based on literature reports, O2′-pyrene-functionalized RNA and N2′-pyrene-functionalized-2′-*N*-methyl-2′-amino DNA (Figure 1-32) were identified as likely candidates.

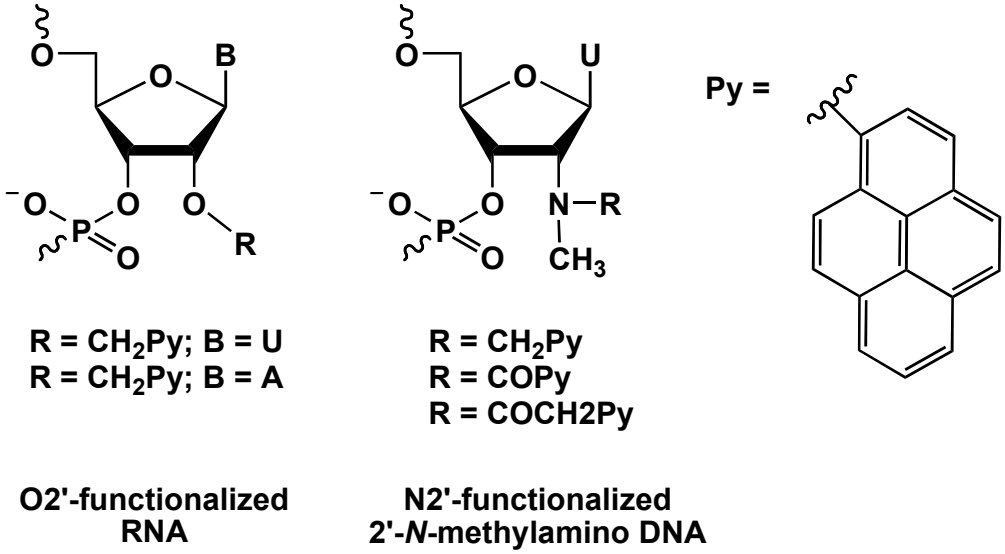


Figure 1-32: Structures of 2nd generation Invader building blocks.

Ribonucleotides that are functionalized at the O2′-position can be synthesized via a short and robust synthetic route (four steps from uridine), allowing the introduction of different intercalating moieties.[154] ONs modified with O2′-(pyren-1-yl)methyl-RNA-U monomers display efficient and specific binding to mixed-sequence cDNA with similar affinity as ONs modified with N2′-pyrene-functionalized-2′-amino-α-L-LNA-T (ΔT_m of up to +14.0 °C and +19.5 °C, respectively).[154] The high affinity toward DNA targets is presumably also due to

pyrene intercalation into the duplex core and was verified through force-field calculations (Figure 1-33a), thermal denaturation temperatures and absorbance, and fluorescence spectroscopy experiments. Similar to first generation Invader probes, +1 interstrand zippers of O2′-(pyren-1-yl)methyl-RNA-U results in thermodynamically unstable duplexes as verified through thermal denaturation temperatures and molecular modeling (Figure 1-33b).[148,150] The pyrene-pyrene interaction perturbs the stacking between the pyrene and the neighboring nucleobases and as a result, the flanking base pairs are strongly buckled, resulting in thermodynamically destabilized duplexes. Incubation of 2nd generation Invaders with DIG-labeled hairpins in a similar manner as previously discussed results in dose-dependent recognition of hairpin DNA as determined via polyacrylamide gel electrophoresis experiments. Similar recognition efficiencies are observed as with Invader probes based on the original 2′-amino-α-L-LNA-T building blocks.[148] Not only was it demonstrated that these next generation probe architectures enable efficient and highly specific dsDNA recognition of DNA hairpins in cell free assays, but also in chromosomal DNA in fixed interphase nuclei and metaphase spreads.[151] Fluorophore-labeled Invader probes (Cy3-Invaders) containing 3 hotspots designed to target the Y-chromosome of bovine DNA were used in a non-denaturing FISH (fluorescence in situ hybridization) experiments. Incubation of fixed interphase and metaphase nuclei spreads with Cy3-labeled Invaders under non-denaturing conditions results in highly specific localized nuclear Cy3 fluorescence as a result of Invader-mediated dsDNA recognition in a system where LNAs, PNAs, and polyamides have failed, establishing proof-of-concept for Invader-mediated recognition of mixed-sequence dsDNA in biological media (Figure 1-33c).[151]

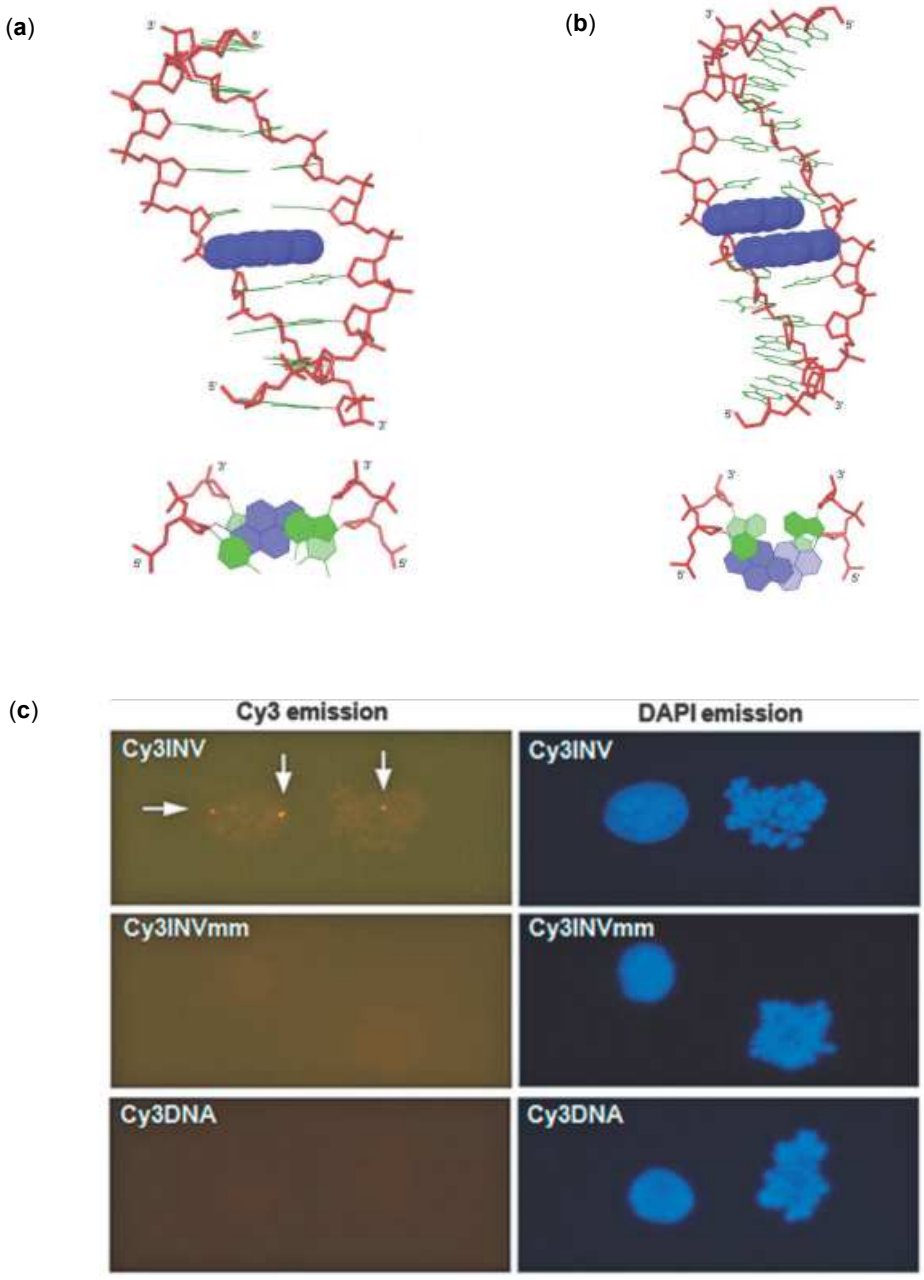


Figure 1-33: (a) Lowest energy structure of 9-mer mixed-sequence DNA (single incorporation of O2′-pyren-1-ylmethyl RNA U) with cDNA. (Reprodued from ref. 150. Copyright 2014 published by the Royal Society of Chemistry). (b) Lowest energy structure of 9-mer O2′-pyren-1-ylmethyl RNA U zipper probe. Upper: side view of duplex; Lower: top view of central duplex region (Ref. 150. Copyright 2014 $_{\text{published}}$ by the Royal Society of Chemistry). (c) Representative

images from non-denaturaing FISH experiments. Cy3INV: sequence-matched Invader, Cy3INVmm: fully base-paired but triply mismatched Invader, and Cy3DNA: unmodified analogue of sequence-matched Invader. Images viewed by using Cy3 (left) or DAPI (right) filter settings. (Reprinted with permission from ref. 151. Copyright 2013 John Wiley and Sons).

The influence of nucleobases on dsDNA recognition efficiency was systematically studied using double-stranded probes modified with all four nucleobase moieties of 2′-O-(pyren1-yl)methyl RNA monomers to identify an inherent limitations in the Invader design.[149] Thermal denaturation, binding energy, and recognition experiments using Invaders with +1 interstrand zippers of 2'-O-(pyren-1-yl)methyl-RNA A/C/G/U monomers resulted in Invader design guidelines and showed there are only a few restrictions in the design of Invader probes for dsDNA recognition.[149] Invaders with interstrand zippers comprised of C and/or U monomers result in the most efficient recognition of dsDNA, whereas Invaders comprised of G or A monomers are the least activated constructs for dsDNA recognition. The insight gained from this study will drive the design of effective Invaders for future applications in molecular biology, nucleic acid diagnostics and biotechnology

N2′-pyrene-functionalized 2′-*N*-methyl-2′-amino DNAs (Figure 1-32) were discovered hand-in-hand with O2′-pyrene-functionalized RNAs as alternative Invader building blocks. They are synthesized via a short and robust method (7 synthetic steps from uridine) to give access to N2′-pyrene-functionalized 2′-*N*-methyl-2′-amino-DNA-U monomers with multiple pyrene-linker chemistries.[154] ONs modified with these monomers display high affinity and specificity for cDNA targets as determined through thermal denaturation, absorbance, and fluorescence spectroscopy experiments. Monomers in which the pyrene moiety is attached via an alkyl linker induce greater duplex stability vs cDNA than monomers with a carbonyl linker.[154] The

increased affinity for cDNA is due to pyrene intercalation into the duplex core in a similar manner as discussed with N2′-pyrene-functionalized 2′-amino-α-L-LNA and 2′-O-pyrene-functionalized RNA. DNA duplexes with +1 interstrand zipper arrangements of 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-U are highly energetically activated with similar dsDNA recognition potential as original N2′-pyrene-functionalized-2′-amino-α-L-LNA-T building blocks.[148] DNA recognition experiments with probes containing +1 interstrand zipper moieties of 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-U result in dose-dependent formation of dsDNA-recognition complexes in the cell-free hairpin assay with similar recognition efficiencies as Invader probes based on the 1st generation monomers indicating this modification is in fact a structural and functional mimic of N2′-pyrene-functionalized-2′-amino-α-L-LNA-T.[148]

Identification of simpler scaffolds enables systematic structure-property relationship studies to be performed with the goal of optimizing the recognition efficiency of Invader probes. Ongoing studies include evaluating these next-generation probes on longer sequences with more hotspots for high affinity and specificity binding, modifying the hotpots with different intercalator sizes and substituents, and evaluating them as transcriptional inhibitors in in vitro transcription assays to identify the optimal Invader design for use as diagnostic tools and as gene modulatory agents.

1.7.2 Intercalating Nucleic Acids (INAs)

A related dsDNA targeting approach based on intercalating nucleic acids (INAs)[155] appeared in scientific literature shortly after the Invader concept was introduced.[145] Proposed by Christensen and Pedersen, INAs are ONs modified with non-nucleosidic 1-*O*-(pyren-1-

ylmethyl)-glycerol units (Figure 1-34a) that are placed as bulge insertions into DNA.[156] These ONs show increased affinity toward ssDNA targets from +2.0 to +6.7 °C per modification (depending on the proximity of the INA modifications to one another) with specificity toward DNA vs RNA targets.[156] This increased affinity toward DNA is a result of the pyrene moiety intercalating into the DNA duplex and interacting with the neighboring nucleic acid base pairs, as determined through molecular modeling (Figure 1-34b), thermal denaturation experiments, and fluorescence spectroscopy.[156] Similar to Invader probes, INA-modified ONs can be utilized to recognize dsDNA targets through incorporation into energetically activated duplexes.[155] The activation of the probe arises from the bulged intercalators being positioned opposite of one another in the probe strand, which results in duplex destabilization due to the intercalators being forced into the same region. However, each INA has increased binding affinity toward complementary ssDNA, which drives the recognition event (see Figure 1-34c).[155]

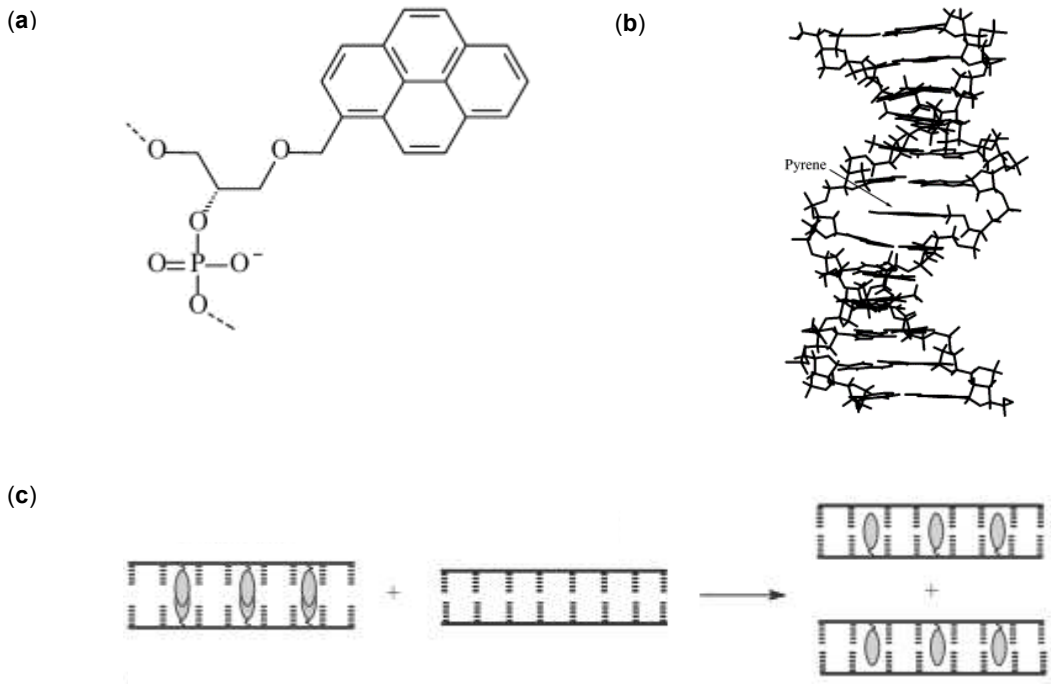


Figure 1-34: (a) Structure of INA monomer. (b) Macromodel calculated structure of 12-mer DNA with 13-mer INA/DNA duplex with the sequence 5′-AGCTTGCTTGAG-3′ + 5′-CTCAAGXCAAGCT-3′, X = INA. (c) illustration of INA mediated strand invasion. The ovals denote the INA monomer inserted as a bulge into ONs. (Reproduced with permission from ref. 155. Copyright 2005 and ref. 156. Copyright 2002 Oxford University Press).

INAs were tested in a transcriptional inhibition assay to ascertain their use as gene modulatory tools. Inhibition of transcription was observed for an 8-mer dsDNA containing multiple INA modifications when targeting the transcription start site of 128-mer dsDNA.[155] However, the observed inhibition was believed to be a result of nonspecific binding of the modified DNA to the RNA polymerase as they were targeting an open transcription complex resulting in ssDNA and dissociated INAs.[155] In the same study, a single-stranded INA with 3 bulge insertions was targeted downstream from the transcription start site which resulted in a transcriptional halt and

a truncated RNA transcript. This targeting was believed to be a result of local melting of the dsDNA by RNA polymerase making the target site more accessible and allowing the INA to bind.[155] These findings encourage the further use of INAs in RNA transcription assays for the development of new therapeutic tools.

1.8 Conclusion

All the different types of gene-regulatory methodologies mentioned remain targets of intense chemical and biological research, with the advantages and limitations of each type of targeting-strategy being identified. As dsDNA targeting methodologies become more refined and more is understood regarding these recognition complexes, the development in biotechnology and drug discovery will generate increased interest. The design and characterization of high affinity and specificity targeting agents will undoubtedly result in gene regulatory approaches that may hold promise for novel drug discovery and the future of genomic engineering.

1.9 References

(1) Illustration adapted from the National Human Genome Research Institute (NHGRI). Genetic Illustrations entry for mRNA. Illustrations are copyright free.

(2) Rogers, F. A.; Lloyd J. A.; Glazer, P. M. *Curr. Med. Chem.*: *Anti-Cancer Agents* **2005**, 5, 319 – 326.

(3) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L; Arimondo, P. B. *Nucleic Acids Res.* **2008**, 36, 5123 – 5138.

(4) Besch, R.; Giovannangeli, C.; Degitz, K. *Curr. Drug Targets* **2004**, 5, 691 – 703.

(5) Buchini, S; Leumann,C. J. *Curr. Opin. Chem. Biol.* **2003**, 7, 717 – 726.

(6) Vekhoff, P.; Ceccaldi, A.; Polverary, D.; Pylouster, J.; Pisano, C.; Arimondo, P. B. *Biochemistry* **2008**, 47, 12277 – 12289.

(7) Cheng, A. J.; Van Dyke, M. W. *Nucleic Acids Res.* **1994**, 22, 4742 – 4747.

(8) Gowers, D. M.; Fox, K. R. *Nucleic Acids Res.* **1997**, 25, 3787 – 3794.

(9) Orson, F. M.; Klysik, J.; Bergstrom, D. E.; Ward, B.; Glass, G. A.; Hua, P.; Kinsey, B. M.; *Nucleic Acids Res.* **1999**, 27, 810 – 816.

(10) Lee, J. S.; Woodsworth, M. L.; Latimer, L. J.; Morgan, A. R. *Nucleic Acids Res*. **1984**, 12, 6603 – 6614.

(11) Cassidy, S. A.; Slickers, P.; Trent, J. O.; Capaldi, D. C.; Roselt, P. D.; Reese, C. B.; Neidle, S.; Fox, K. R. *Nucleic Acids. Res.* **1997**, 25, 4891 – 4898.

(12) Xodo, L.E.; Manzini, G.; Quadifoglio, F.; van der Marel, G.; van Boom, J. *Nucleic Acids Res.* 1991, 19, 5625 – 5631.

(13) Mills, M.; Arimondo, P. B.; Lacroix, L.; Garestier, T.; Klump, H.;Mergny, J. L. *Biochemistry* **2002**, 41, 357 – 366.

(14) Froehler, B. C.; Wadwani, S.; Terhorst, T. J.; Gerrard, S. R. *Tetrahedron Lett.* **1992**, 33, 5307 – 5310.

(15) Bijapur, J.; Keppler, M. D.; Bergqvist, S.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **1999**, 27, 1802 – 1809.

(16) Sollogoub, M.; Darby, R. A.; Cuenoud, B.; Brown, T.; Fox, K. R. *Biochemistry* **2002**, 41, 7224 – 7231.

(17) Brazier, J. A.; Shibata, T.; Townsley, J.; Taylor, B. F.; Frary, E.; Williams, N. H.; Williams, D. M. *Nucleic Acids Res*. **2005**, 33, 1362 – 1371.

(18) Alam, M. R.; Majumdar, A.; Thazhathveetil, A. K.; Liu, S. T.; Liu, J. L.; Puri, N.; Cuenoud, B.; Sasaki, S.; Miller, P. S.; Seidman, M. M. *Biochemistry* **2007**, 46, 10222 – 10233.

(19) Rusling, D. A.; Le Strat, L.; Powers, V. E.; Broughton-Head, V. J.; Booth, J.; Lack, O.; Brown, T; Fox, K. R. *FEBS Lett.* **2005**, 579, 6616 – 6620.

(20) Shimizu, M.; Konishi, A.; Shimada, Y.; Inoue, H.; Ohtsuka, E. *FEBS Lett.* 1992, 302, 155 – 158.

(21) Escude,C.; Sun, J. S.; Rouge, M.; Garestier, T; Helene, C. *C. R. Acad. Sci. Paris*, *Serie III* **1992**, 315, 521 – 525.

(22) Brunet, E.;Corgnali, M.; Perrouault, L.; Roig, V.; Asseline, U.; Sorensen, M. D.; Babu, B. R.; Wengel, J.; Giovannangeli, C. *Nucleic Acids Res.* **2005**, 33, 4223 – 4234.

(23) Petersen, M.; Wengel, J. *Trends Biotechnol.* 2003, 21, 74 – 81.

(24) Torigoe, H.; Hari, Y.; Sekiguchi, M.; Obika, S.; Imanishi, T. *J. Biol. Chem*., **2001**, 276, 2354 – 2360.

(25) Brunet, E.; Alberti, P.; Perrouault, L.; Babu, R.; Wengel, J.; Giovannangeli, C. *J. Biol. Chem*., **2005**, 280, 20076 – 20085.

(26) Koizumi, M.; Morita, K.; Daigo, M.; Tsutsumi, S.; Abe, K.; Obika, S.; Imanishi, T. *Nucleic Acids Res*. **2003**, 31, 3267 – 3273.

(27) Obika, S.; Rahman, A. M. A.; Fujisaka, A.; Kawada, Y.; Baba, T.; Imanishi, T. *Heterocycles* **2010**, 81, 1347 – 1392

(28) Albaek, A.; Petersen, M.; Nielsen, P. *J. Org. Chem.* **2006**, 71, 7731.

(29) Rahman, S. M.; Seki, S.; Obika, S.; Yoshikawa, H.; Miyashita, K.; Imanishi, T. *J. Am. Chem. Soc.* **2008**, 130, 4886 – 4896.

(30) Roberts, R. W.; Crothers, D. M. *Science* **1992**, 258, 1463 – 1466.

(31) Tarkoy, M.; Phipps, A. K.; Schultze, P.; Feigon, J. *Biochemistry* **1998**, 37, 5810 – 5819.

(32) Seidman, M. M.; Puri, N.; Majumdar, A.; Cuenoud, B.; Miller, P. S.; Alam, R. *Ann. N Y Acad. Sci*., **2005**, 1058, 119 – 127.

(33) Obika, S.; Uneda, T.; Sugimoto, T.; Nanbu, D.; Minami, T.; Doi, T.; Imanishi, T. *Bioorg. Med. Chem.* **2001**, 9, 1001 – 1011.

(34) Alam, M. R.; Majumdar, A.; Thazhathveetil, A. K.; Liu, S. T.; Liu, J. L.; Puri, N.; Cuenoud, B.; Sasaki, S.; Miller, P. S.; Seidman, M. M. *Biochemistry* **2007**, 46, 10222 – 10233.

(35) Li, H.; Miller, P. S.; Seidman, M. M. *Org. Biomol. Chem.* **2008**, 6, 4212 – 4217.

(36) Lacroix, L.; Lacoste, J.; Reddoch, J. F.; Mergny, J.-L.; Levy, D. D.; Seidman, M. M.; Matteucci, M. D.; Glazer, P. M. *Biochemistry* **1999**, 38, 1893 – 1901.

(37) Osborne, S. D.; Powers, V. E. C; Rusling, D. A.; Lack, O.; Fox, K. R.; Brown, T. *Nucleic Acids Res.* **2004**, 32, 4439 – 4447.

(38) Rusling, D. A.; Broughton-Head, V. J.; Tuck, A.; Khairallah, H.; Osborne, S. D.; Brown, T.; Fox, K. R. *Org. Biomol. Chem.* **2008**, 6, 122 – 129.

(39) Sau, S. P.; Kumar, P.; Anderson, B. A.;Ostergaard, M. E.; Deobald, L.; Paszczynski, A.; Sharma, P. K.; Hrdlicka, P. J. *Chem. Commun*. **2009**, 6756 – 6758.

(40) Sau, S. P.; Kumar, P.; Sharma, P. K.; Hrdlicka, P. J.; *Nucleic Acids Res.* **2012**, 40, e162.

(41) Xodo, L.; Alunni-Fabbroni, M.; Manzini, G.; Quadrifoglio, F. *Nucleic Acids Res.* **1994**, 22, 3322 – 3330.

(42) Cogoi, S.; Rapozzi, V.; Quadrifoglio, F.; Xodo, L. *Biochemistry* **2001**, 40 1135 – 1143.

(43) McGuffie, E. M.; Pacheco, D.; Carbone, G. M.; Catapano, C. V.; *Cancer Res.* **2000**, 60, 3790 – 3799.

(44) Michel, T.; Debart, F.; Heitz, F.; Vasseur, J .J. *Chembiochem* **2005**, 6, 1254 – 1262.

(45) Michel, T.; Martinand-Mari, C.; Debart, F.; Lebleu, B.; Robbins, I.; Vasseur, J. J. *Nucleic Acids Res.* **2003**, 31, 5282 – 5290.

(46) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, 254, 1497 – 1500.

(47) Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchard, O.; Sonnichsen, S. H.; Nielsen P. E. *Biochem. Pharmacol*. **1994**, 48, 1310 – 1313.

(48) Milligan, J. F.; Krawczyk, S. H.; Wadwani, S.; Matteucci, M. D. *Nucleic Acids Res.* **1993**, 21, 327 – 333.

(49) Olivas, W. M. Maher, L. J. III. *Nucleic Acids Res.* **1995**, 23, 1936 – 1941.

(50) Rao, T. S.; Durland,R. H.; Seth, D. M.; Myrick, M.A.; Bodepudi, V.; Revankar, G. R.; *Biochemistry* **1995**, 34 765 – 772.

(51) Cheong, C.; Moore, P. B. *Biochemistry* **1992**, 31, 8406 – 8414.

(52) Sasaki, S.; Taniguchi, Y.; Takahashi, R.; Senko, Y.; Kodama, K.; Nagatsugi, F.; Maeda, M. *J. Am. Chem. Soc*. **2004**, 126, 516 – 528.

(53) Taniguchi, Y.; Uchida, Y.; Takaki, T.; Aoki, E.; Sasaki S. *Bioorg. Med. Chem.* **2009**, 17, 6803 – 6810.

(54) Obika, S.; Hari, Y.; Sekiguchi, M.; Imanishi, T. *Angew. Chem. Int. Ed.* **2001**, 40, 2079 – 2081.

(55) Obika, S.; Hari, Y.; Sekiguchi, M.; Imanishi, T. *Chem. Eur. J.* **2002**, 8, 4796 – 4802.

(56) Hari, Y.; Obika, S.; Sekiguchi, M.; Imanishi, T. *Tetrahedron* **2003**, 59, 5123 – 5128.

(57) Hari, Y.; Matsugu, S.; Inohara, H.; Hatanaka, Y.; Akabane, M.; Imanishi, T.; Obika, S. *Org. Biomol. Chem.* **2010**, 8, 4176 – 4180.

(58) Obika, S.; Inohara, H.; Hari, Y.; Imanishi, T. *Bioorg. Med. Chem.* **2008**, 16, 2945 – 2954.

(59) Rusling, D. A.; Powers, V. E. C.; Ranasinghe, R. T.; Wang, Y.; Osborne, S. D.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **2005**, 33, 3025 – 3032.

(60) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, 112, 2435 – 2437.

(61) Filichev, V. V.; Nielsen, C. M.; Bomholt, N.; Jessen, C. H.; Pedersen, E. B. *Angew. Chem., Int. Ed*., **2006**, 45, 5311 – 5315.

(62) Bewley, C. A.; Gronenborn, A. M.; Clore, G. M. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, 27, 105 – 131.

(63) Bailly, C.; Chaires, J. B. *Bioconj. Chem*. **1998**, 9, 513 – 538 and references therein.

(64) Blackledge, M. S.; Melander, C. *Bio. Med. Chem.* **2013**, 21, 6101-6114.

(65) Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **199**7, 387, 202 – 205.

(66) Dervan, P. B. *Bioorg. Med. Chem*. **2001**, 9, 2215 – 2235.

(67) Dervan, P. B.; Bürli, R. W. *Curr. Opin. Chem. Biol.* **1999**, 3, 688 – 693.

(68) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, 13, 284 – 299.

(69) Meier, J. L.; Yu, A. S.; Korf, I.; Segal, D. J.; Dervan, P. B. *J. Am. Chem. Soc.* **2012**, 134, 17814 – 17822.

(70) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, 114, 8783 – 8794.

(71) Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, 116, 7983 – 7988.

(72) Kelly, J. J.; Baird, E. E.; Dervan, P. B.; *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93, 6981 – 6985.

(73) Kielkopf, C.L.; Baird, E. E.; Dervan, P. D.; Rees, D. C. *Nat. Struct. Biol.* **1998**, 5, 104 – 109.

(74) Kielkopf C. L.; White, S.; Szewczyk, J. W,; Turner, J. M,; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science* **1998**, 282, 111 – 115.

(75) Kielkopf, C. L,; Bremer, R. E.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *J. Mol. Biol.* **2000**, 295, 557 – 567.

(76) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, 391, 468 – 471.

(77) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, 120, 3534 – 3535.

(78) Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, 120, 1382 – 1391.

(79) Meier, J. L.; Montgomery, D. C.; Dervan, P. B. *Nucleic Acids Res.* **2012**, 40, 2345 – 2356.

(80) Nielsen, P.; Egholm, M.; Berg, R.; Buchardt, O. *Science* **1991**, 254, 1497 – 1500.

(81) Uhlmann, E.; Peymann, A.; Breipohl, G.; Will, D. W.; *Angew. Chem. Int. Ed.* **1998**, 37, 2796 – 2823.

(82) Nielsen, P. E. *Curr. Opin. Struct. Biol.* **1999**, 9, 353 – 357.

(83) Hansen, M. E.; Bentin, T.; Nielsen, P. E. *Nucleic Acids Res*. **2009**, 37, 4498 – 4507.

(84) Kaihatsu, K.; Janowski, B. A.; Corey, D R. *Chem. Biol.* **2004**, 11, 749 – 758.

(85) Egholm, M.; Christensen, L.; Dueholm, K. L.; Buchardt, O.; Coull, J.; Nielsen, P. E. *Nucleic Acids Res.* **1995**, 23, 217 – 222.

(86) Hansen, G. I.; Bentin, T.; Larsen, H. J.; Nielsen, P. E. *J. Mol. Biol.* **2001**. 307, 67 – 74.

(87) Bentin, T.; Larsen, H. J.; Nielsen, P. E. *Biochemistry* **2003**, 42, 13987 – 13995.

(88) Kaihatsu, K.; Shah, R. H.; Zhao, X.; Corey, D. R. *Biochemistry* **2003**, 42, 13996 – 14003.

(89) Simon, P.; Cannata, F.; Concordet, J.; Giovannangeli, C. *Biochimie* **2008**, 90, 1109 – 1116.

(90) Holloman, W.K.; Wiegand, R.; Hoesli, C.; Radding, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, 72, 2394 – 2398.

(91) Corey, D.R. *J. Am. Chem. Soc*. **1995**, 117, 9373 – 9374.

(92) Ishihara, T., Corey, D. R. *J. Am. Chem. Soc.* **1999**, 121, 2012 – 2020.

(93) Smulevitch, S. V.; Simmons, C. G.; Norton, J. C.; Wise, T. W.; Corey, D. R. *Nat. Biotechnol.* **1996**, 14, 1700 – 1705.

(94) Bentin, T.; Nielsen, P. E. *Biochemistry* **1996**, 35, 8863 – 8869.

(95) Zhang, X.; Ishihara, T.; Corey, D. R. *Nucleic Acids Res.* **2000**, 28, 3332 – 3338.

(96) Gildea, B. D.; Coull, J. M.; Hyldig-Nielsen, J. J.; Fiandaca, M. J. WO A-9921881, **1999**.

(97) Tackett, A. J.; Corey, D. R.; Raney, K. D. *Nucleic Acids Res.* **2002**, 30, 950 – 957.

(98) Seitz, O. *Angew. Chem., Int. Ed.* **2000**, 39, 3249 – 3252.

(99) Ranasinghe, R. T.; Brown, L. J.; Brown, T. *Chem. Commun.* **2001**, 1480 – 1481.

(100) Kool, E. T. *Chem. Rev.* **1997**, 97, 1473 – 1487.

(101) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, 128, 10258 – 10267.

(102) He, G.; Rapireddy, S.; Bahal, R.; Sahu, B.' Ly, D. H. *J. Am. Chem. Soc.* **2009**, 131, 12088 – 12090.

(103) Rapireddy, S.; He,G.; Roy, S.; Armitage, B. A.; Ly, D. H. *J. Am. Chem. Soc.* **2007**, 129, 15596 – 15600.

(104) Chenna, V.; Rapireddy, S.; Sahu, B.; Ausin, C.; Pedroso, E.; Ly, D. H. *ChemBioChem* **2008**, 9, 2388 – 2391.

(105) Rapireddy, S.; Bahal, R.; Ly, D. H. *Biochemistry* **2011**, 50, 3913 – 3918.

(106) Wilds, C. J.; Maier, M. A.; Tereshko, V.; Manoharan, M.; Egli, M. *Angew. Chem.* **2002**, 114, 123 – 125; *Angew. Chem. Int. Ed.* **2002**, 41, 115 – 117.

(107) Rajeev, K. G.; Maier, M. A.; Lesnik, E. A.; Manoharan, M.; *Org. Lett*. **2002**, 4, 4395 – 4398.

(108) Ortega, J-A.; Blas, J. R.; Orozco, M.; Grandas, A.; Pedroso, E.; Robles, J. *Org. Lett.* **2007**, 9, 4503 – 4506.

(109) Bahal, R.; Sahu, B.; Rapireddy, S.; Lee, C-M.;. Ly, D. H. *ChemBioChem* **2012**, 13, 56 – 60.

(110) Kutyavin, I. V.; Rhinehart, R. L.; Lukhtanov, E. A.; Gorn, V. V.; Meyer, R. B. Jr.; Gamper, H. B. *Biochemistry* **1996**, 35, 11170 – 11176.

(111) Lonkar, P.; Kim, K.; Kuan, J. Y.; Chin, J. Y.; Rogers, F. A.; Knauert, M. P.; Kole, R.; Nielsen, P. E.; Glazer, P. M. *Nucleic Acids Res.* **2009**, 37, 635 – 3644.

(112) Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 11804 – 11808.

(113) Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corradini, R.; Sforza, S.; Komiyama, M. *Nucleic Acids Res.* **2008**, 36, 1464 – 1471.

(114) Protozanova, E.; Demidovl, V. V.; Nielsen, P. E.; Frank- Kamenetskii, M. D. *Nucleic Acids Res.* **2003**, 31, 3929 – 3935.

(115) Yamamoto, Y.; Mori, M.; Aiba, Y.; Tomita, T.; Chen, W.; Zhou, J.; Uehara, A.; Ren, Y.; Kitamura, Y.; Komiyama, M. *Nucleic Acids Res.* **2007**, 35, e53.

(116) Demidov, V. V.; Protozanova, E.; Izvolsky, K. I.; Price, C.; Nielsen, P. E.; Frank-Kamenetskii, M. D. *Proc*. *Natl*. *Acad*. *Sci*. *U.S.A.* **2002**, 99, 5953 – 5958.

(117) Sumaoka, J.; Komiyama, M. *Chem. Lett.* **2014**, 43, 1581 – 1583.

(118) Ge, R.; Heinonen, J. E.; Svahn, M. G.; Mohamed, A. J.; Lundin, K. E.; Smith, C. I. E. *FASEBJ*. **2007,** 21, 1902 – 1914.

(119) Ge, R.; Svahn, M. G.; Simonson, O. E.; Mohamed, A. J.; Lundin, K. E.; Smith, C. I. E. *J. Gene Med*. **2008**, 10, 101 – 109.

(120) Zaghloul, E. M.; Madsen, A. S.; Moreno, P. M. D.; Oprea, I. I.; El-Andaloussi, S.; Bestas, B.; Gupta, P.; Pedersen, E. B.; Lundin, K. E.; Wengel, J.; Smith, C. I. E. *Nucleic Acids Res.* **2011**, 39, 1142 – 1154.

(121) Moreno, P. M. D.; Geny, S.; Pabon, Y. V.; Bergquist, H.; Zaghloul, E. M.; Rocha, C. S. J.; Oprea, I. I.; Bestas, B.; Andaloussi, S. E. L.; Jørgensen, P. T.; Pedersen, E. B.; Lundin, K. E.; Zain, R.; Wengel, J.; Smith, C. I. E. *Nucleic Acids Res*. **2013**, 41, 3257 – 3273.

(122) Venter, J. C. *Science* **2001**, 291, 1304 – 1351.

(123) Wolfe, S. A.; Nekludova, L.; Pabo, C. O. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, 29, 183 – 212.

(124) Ghosh, I.; Stains, C. I.; Ooi A. T.; Segal, D. J. *Mol. Biosyst.* **2006**, 2, 551 – 560.

(125) Carroll, D. *Gene Ther.* **2008**, 15, 1463 – 1468.

(126) Uil, T. G.; Haisma, H. J.; Rots, M. G. *Nucleic Acids Res.* **2003**, 31, 6064 – 6078.

(127) Beumer, K.; Bhattacharyya, G.; Bibikova, M.; Trautman, J. K.; Carroll, D. *Genetics* **2006**, 172, 2391 – 2403.

(128) Szczepek, M.; Brondani, V.; Buchel, J.; Serrano, L,; Segal, D. J.; Cathomen, T. *Nat. Biotechnol.* **2007**, 25, 786 – 793.

(129) Gaj, T.; Gersbach, C. A.; Barbas III, C. F. *Trends Biotechnol*. **2013**, 31, 397 – 405.

(130) Mak, A. N.; Bradley, P.; Cernadas, R. A.;Bogdanove, A. J.; Stoddard, B. L. *Science*, **2012**, 335, 716–719.

(131) Deng, D. A.; Yan, C.; Pan, X.; Mahfouz, M.; Wang, J.; Zhu, J.; Shi, Y.; Yan, N. *Science,* **2012**, 335, 720-723.

(132) Moscou, M. J.; Bogdanove, A. J.; *Science* **2009**, 326, 1501 – 1508.

(133) Boch, J.; Scholze, H.; Schornack, S.; Landgraf, A.; Hahn, S.; Kay, S.; Lahaye, T.; Nickstadt, A.; Bonas, U. *Science,* **2009**, 326, 1509 – 1512.

(134) Christian, M.; Cermak, T.; Doyle, E. L.; Schmidt, C.; Zhang, F.; Hummel, A.; Bogdanove, A. J.; Voytas, D. F. *Genetics* **2010**, 186, 757 – 761.

(135) Mussolino, C.; Morbitzer, R.; Lutge, F.; Dannemann, N.; Lahaye, T.; Cathomen, T. *Nucleic Acids Research*, **2011**, 21, 9283 – 9293.

(136) Marraffini, L. A.; Sontheimer, F. J. *Nat. Rev. Genet.* **2010**, 11, 181 – 190.

(137) Horvath, P.; Barrangou, R. *Science* **2010**, 327, 167 – 170.

(138) Wiedenheft, B.; Sternberg, S. H.; Doudna, J. A. *Nature* **2012**, 482, 331 – 338.

(139) Hwang, W.Y; Fu, Y,; Reyon D.; Maeder, M. L.; Tsai, S. Q.; Sander, J. D.; Petersen, R. T.; Yeh, J. J.; Young, J. K. *Nat. Biotechnol.* **2013**, 31, 227 – 229.

(140) Jiang, W.; Bikard, D.; Cox, D.; Zhang, F.; Marraffini, L. A. *Nat. Biotechnol*. **2013**, 31, 233 – 239.

(141) Cho, S.W.; Kim, S.; Kim, J. M.; Kim, J. S. *Nat. Biotechnol*. **2013**, 31, 230 – 232.

(142) Cong, L.; Ran, F. A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P. D.; Wu, X.; Jiang, W.; Marraffini, L. A.; Zhang, F. *Science* **2013**, 339, 819 – 823.

(143) Mali, P.; Yang, L.; Esvelt, K. M.; Aach, J.; Guell, M.; DiCarlo, J. E.; Norville, J. E.; Church, G. M. *Science* **2013**, 339, 823 – 826.

(144) Charpentier, E.; Doudna, J. A. *Nature* **2013**, 495, 50 – 51.

(145) Hrdlicka, P. J.; Kumar, T. S.; Wengel, J. *Chem. Commun*. **2005**, 4279 − 4281.

(146) Crothers, D. M. *Biopolymers* **1968**, 6, 575 – 584.

(147) Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org*. *Biomol*. *Chem*. **2010**, 8, 2028 – 2036.

(148) Sau, S. P.; Madsen, A. S.; Podbevsek, P.; Andersen, N. K.; Kumar, T. S.; Andersen, S.; Rathje, R. L.; Anderson, B. A.; Guenther, D. C.; Karmakar, S.; Kumar, P.; Plavec, J.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem*. **2013**, 78, 9560 – 9570.

(149) Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *J. Org*. *Chem*. **2013**, 78, 12040 – 12048.

(150) Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J. *Org*. *Biomol*. *Chem*. **2014**, 12, 7758 – 7773.

(151) Didion, B. A.; Karmakar, S.; Guenther, D. C.; Sau, S.; Verstegen, J. P.; Hrdlicka, P. J. *ChemBioChem* **2013**, 14, 1534 – 1538.

(152) Kumar, T. S.; Madsen, A. S.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem.* **2006**, 71, 4188 – 4201.

(153) Kumar, T. S.; Madsen, A. S.; Østergaard, M. E.; Sau, S. P.; Wengel, J.; Hrdlicka, P. J. *J. Org*. *Chem*. **2009**, 74, 1070 – 1081.

(154) Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem*. **2011**, 76, 7119 – 7131.

(155) Filichev, V. V.; Vester, B.; Hansen, L. H.; Pederson, E. B. *Nucleic Acids Res.* **2005**, 33, 7129 – 7137.

(156) Christenesen, U. B; Pedersen, E. *Nucleic Acids Res.* **2002**, 30, 4918 – 4925.

CHAPTER 2: Synthesis and Characterization of Oligodeoxyribonucleotides Modified with 2'-Amino-α-L-LNA Adenine Monomers: High-affinity Targeting of Single-Stranded DNA

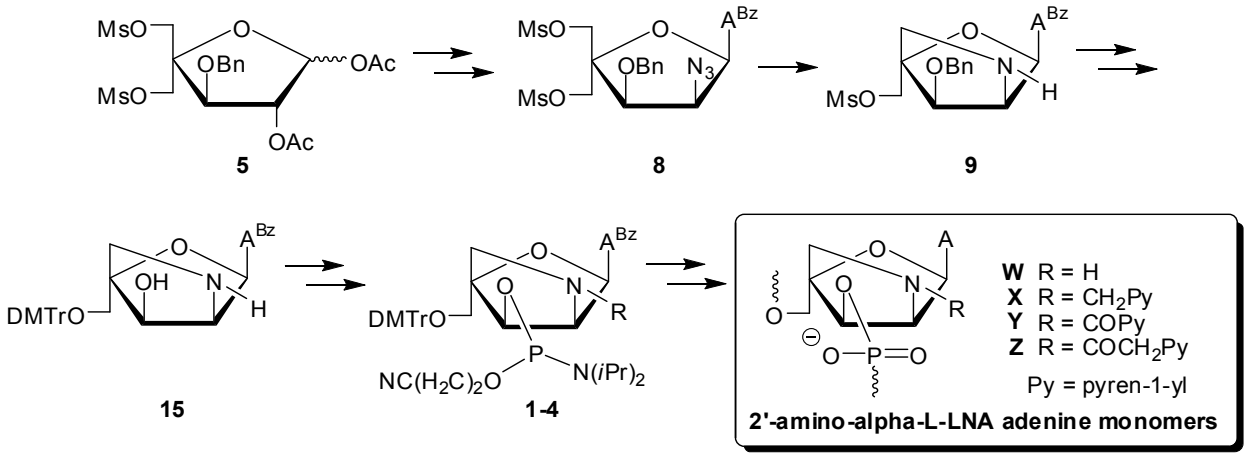
Nicolai K. Andersen,[†,§] Brooke A. Anderson,[‡,§] Sujay P. Sau,[‡] Jesper Wengel,[†] Patrick J. Hrdlicka[‡,*]

[†] Nucleic Acid Center, Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, 5230 Odense, Denmark

[‡] Department of Chemistry, University of Idaho, Moscow, ID 83844-2343, USA

[§] contributed equally to study

*Adapted with permission from *J. Org. Chem.* **2013**, 78, 12690 – 12702. Copyright 2013. American Chemical Society.



Abstract

Development of conformationally restricted nucleotide building blocks continues to attract considerable interest due to their successful use within antisense, antigene and other gene-

targeting strategies. Locked nucleic acid (LNA) and its diastereomer α-L-LNA are two interesting examples hereof. Oligonucleotides modified with these units display greatly increased affinity toward nucleic acid targets, improved binding specificity and enhanced enzymatic stability relative to unmodified strands. Here, we present the synthesis and biophysical characterization of oligodeoxyribonucleotides (ONs) modified with 2'-amino-α-L-LNA adenine monomers **W**-**Z**. The synthesis of target phosphoramidites **1**-**4** initiates from pentafuranose **5**, which upon Vorbrüggen glycosylation, O2'-deacylation, O2'-activation and C2'-azide introduction yields nucleoside **8**. A one-pot tandem Staudinger/intramolecular nucleophilic substitution converts **8** into 2'-amino-α-L-LNA adenine intermediate **9**, which after a series of non-trivial protecting group manipulations affords key intermediate **15**. Subsequent chemoselective N2'-functionalization and O3'-phosphitylation gives targets **1**-**4** in ~1-3% overall yield over eleven steps from **5**. ONs modified with pyrene-functionalized 2'-amino-α-L-LNA adenine monomers **X**-**Z** display greatly increased affinity toward DNA targets (ΔT_m/modification up to +14 °C). Results from absorption and fluorescence spectroscopy suggest that the duplex stabilization is a result of pyrene intercalation. These characteristics render N2'-pyrene-functionalized 2'-amino-α-L-LNA of considerable interest for DNA-targeting applications.

2.1 Introduction

Major efforts have been devoted over the past twenty years to develop conformationally restricted nucleotides. Oligonucleotides that are modified with these building blocks often display markedly increased affinity toward nucleic acid targets, improved discrimination of non-targets and greater resistance against enzymatic degradation relative to reference strands.[1] Such oligonucleotides are accordingly widely used for nucleic acid targeting applications in

molecular biology, biotechnology and medicinal chemistry.[2] Locked Nucleic Acid (LNA) is amongst the most prominent members of this compound class due to its extraordinary affinity toward DNA and RNA complements (Figure 2-1); thus, increases in thermal denaturation temperatures (T_m's) of up to +10 °C have been observed per modification.[3-5] The diastereomeric α-L-LNA shares many characteristics with LNA, including very high affinity toward DNA/RNA targets, but is less well characterized due to more limited commercial availability (Fig. 2-1).[6] The interesting properties of LNA, α-L-LNA and other conformationally restricted nucleotides, has spurred development of many analogues.[1,7]

As part of our ongoing interest in LNA chemistry and diagnostic applications of pyrene-functionalized oligonucleotides,[2c,8] we recently pursued the development of N2'-pyrene-functionalized 2'-amino-α-L-LNA thymine monomers (Fig. 2-1).[9] Oligodeoxyribonucleotides (ONs) modified with these monomers display remarkable affinity toward complementary DNA as the pyrene moieties are preorganized to intercalate and engage in stacking with neighboring nucleobases upon duplex formation.[9b] Thus, increases in T_m's of up to 20 °C per modification have been observed for short ONs modified with these building blocks. We have taken advantage of these characteristics and have developed probes for a variety of diagnostic applications. For example, ONs modified with two next-nearest neighbor incorporations of 2'-*N*-(pyren-1-yl)acetyl-2'-amino-α-L-LNA-T monomers are promising tools for detection of single nucleotide polymorphisms, which are the most prevalent type of genetic mutation in the human genome. These probes efficiently discriminate between complementary and single-base mismatched DNA/RNA targets through differences in pyrene excimer emission levels. In another example, the fluorescence of ONs modified with 2'-*N*-(pyren-1-yl)acetyl-2'-amino-α-L-LNA-T monomers was found to increase 16-fold upon hybridization with DNA targets

featuring abasic sites, which are lesions resulting in genomic mutations and emergence of cancers if unrepaired.[11] We have also utilized N2'-pyrene-functionalized 2'-amino-α-L-LNA-T monomers as the key activating components of the so-called Invader probes, which recognize mixed-sequence double-stranded DNA.[12] Unfortunately, the synthesis of the N2'-pyrene-functionalized 2'-amino-α-L-LNA thymine phosphoramidites is challenging (~20 steps, <4% overall yield from diacetone-α-D-glucose), mainly due to unsuccessful attempts of introducing the necessary C2'-azido group at the nucleoside level without concomitant O2,O2'-anhydronucleoside formation.[9a] This forced us to introduce the azido group at the carbohydrate level, which resulted in a loss of anchimeric assistance from the O2-position during Vorbrüggen glycosylation and the formation of anomeric nucleoside mixtures. We hypothesized that these synthetic difficulties might be overcome with the corresponding adenine derivates, as formation of anhydronucleosides is unlikely. Easier access to N2'-pyrene-functionalized 2'-amino-α-L-LNA monomers is desirable in order to evaluate the diagnostic potential of these building blocks in greater detail.

In the present article, we report the synthesis and characterization of ONs modified with four different 2'-amino-α-L-LNA adenine monomers **W**-**Z** (Fig. 2-1). The ONs are characterized via thermal denaturation, UV-Vis and fluorescence experiments and shown to display extraordinary thermal affinity toward complementary DNA (ΔT_m/modification up to 14.0 °C) and photophysical characteristics consistent with intercalative binding modes for the pyrene moieties.[13]

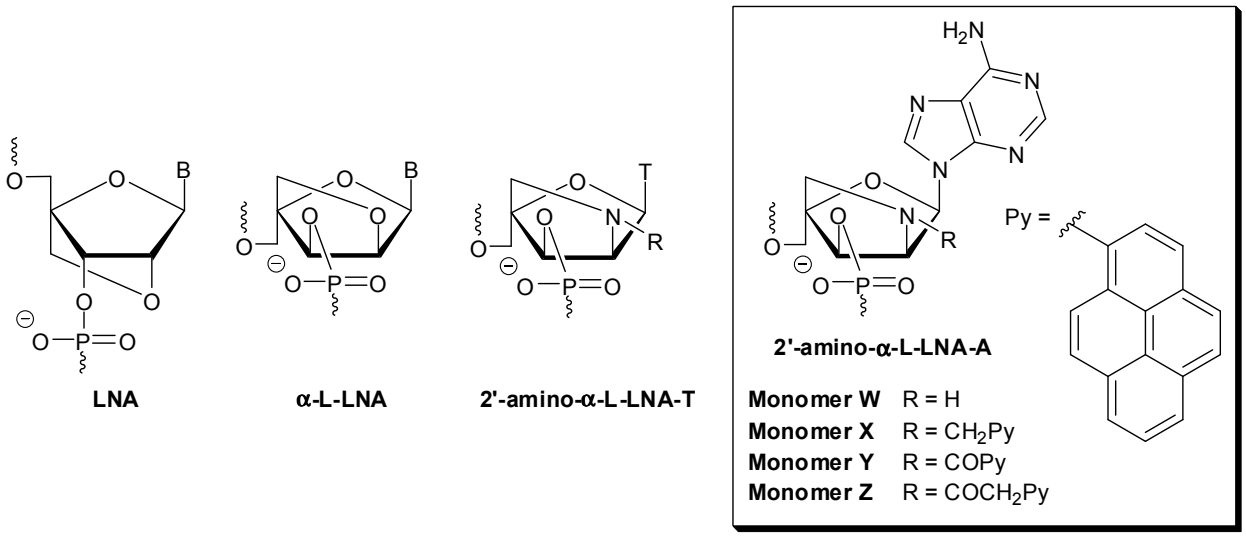


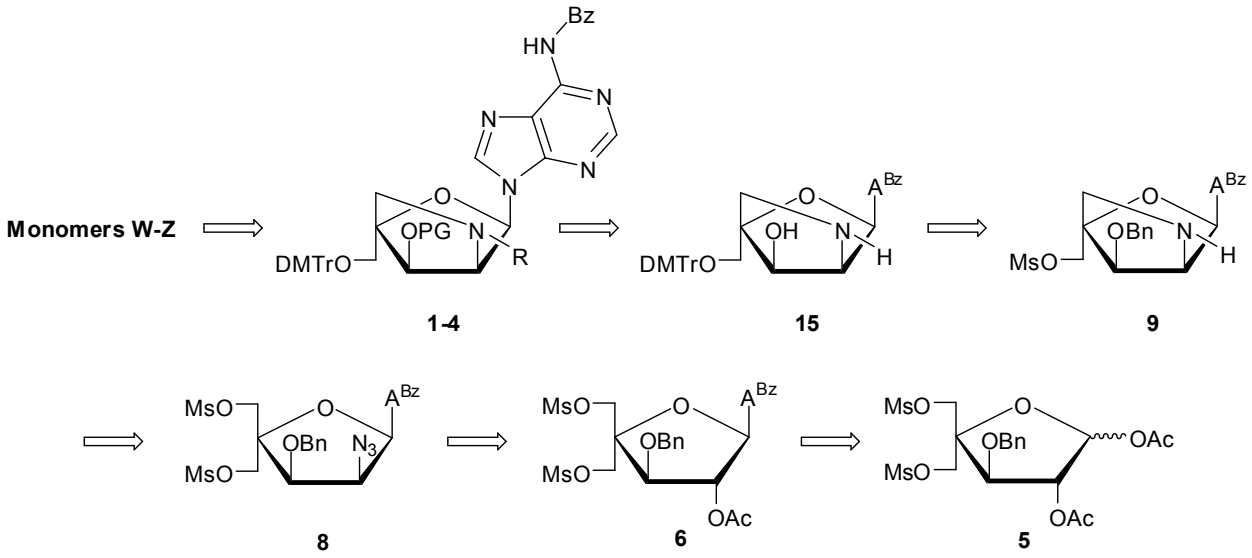
Figure 2-1: Structures of LNA, α-L-LNA and 2'-amino-α-L-LNA monomers studied herein. T = thymin-1-yl.

2.2 Results and Discussion

2.2.1 Retrosynthetic Analysis of 2'-amino-α-L-LNA-A Monomers

Inspired by our previously reported synthesis of the corresponding thymine monomers,[9] we identified adenine derivative **15** as a suitable substrate for chemoselective N2'-functionalization and subsequent O3'-phosphitylation, which was expected to yield target 6-*N*-benzoyl-adenin-9-yl nucleosides **1**-**4** (Scheme 2-1). We surmised that key intermediate **15** could be obtained from nucleoside **9** via a series of protecting group manipulations. The O3'-benzyl group was selected due to i) its stability in acidic and basic reaction conditions, allowing it to be introduced at the beginning of the synthetic route, and ii) its ability to be removed under conditions that only have a minimal effect on the 6-*N*-benzoyl group of adenine.[14] We expected to construct the 2-oxo-5-azabicyclo[2.2.1]heptane skeleton of **9** in a similar manner as originally reported for 2'-amino-β-D-LNA,[15] i.e., via an one-pot tandem Staudinger/intramolecular nucleophilic

substitution reaction, which revealed nucleoside **8** as an early target. In contrast to our synthesis of the corresponding thymine monomer where the C2'-azido group had to be introduced at the carbohydrate stage,[9a] we decided to introduce the 2'-azido group of **8** at the nucleoside stage, anticipating that selective O2'-deacylation of known nucleoside **6**[6b] and subsequent O2'-activation and introduction of the azido group would provide **8**. Importantly, the anchimeric assistance provided by the O2-substituent during the Vorbrüggen glycosylation of **5** to give **6** prevents formation of anomeric mixtures, which is one of the drawbacks in the synthesis of 2'-amino-α-L-LNA-T monomers.[9a]

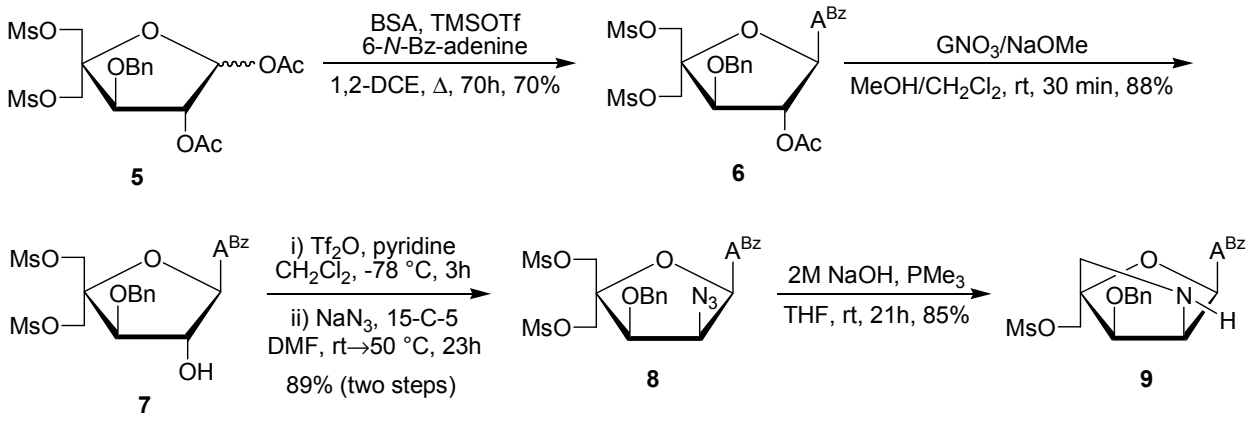


Scheme 2-1: Retrosynthetic analysis for 2'-amino-α-L-LNA adenine monomers **W**-**Z**. PG = $PN(iPr)_2OCH_2CH_2CN$; R = $H/CH_2Py/COPy/COCH_2Py$; py = pyren-1-yl; A^{Bz} = 6-*N*-benzoyl-adenin-9-yl .

2.2.2 Synthesis of Key Intermediate 15

Fully protected pentafuranose **5**, which is obtained from diacetone-α-D-glucose in six steps and ~30% overall yield,[6b,16] serves as the starting material for the synthesis of key intermediate **15**

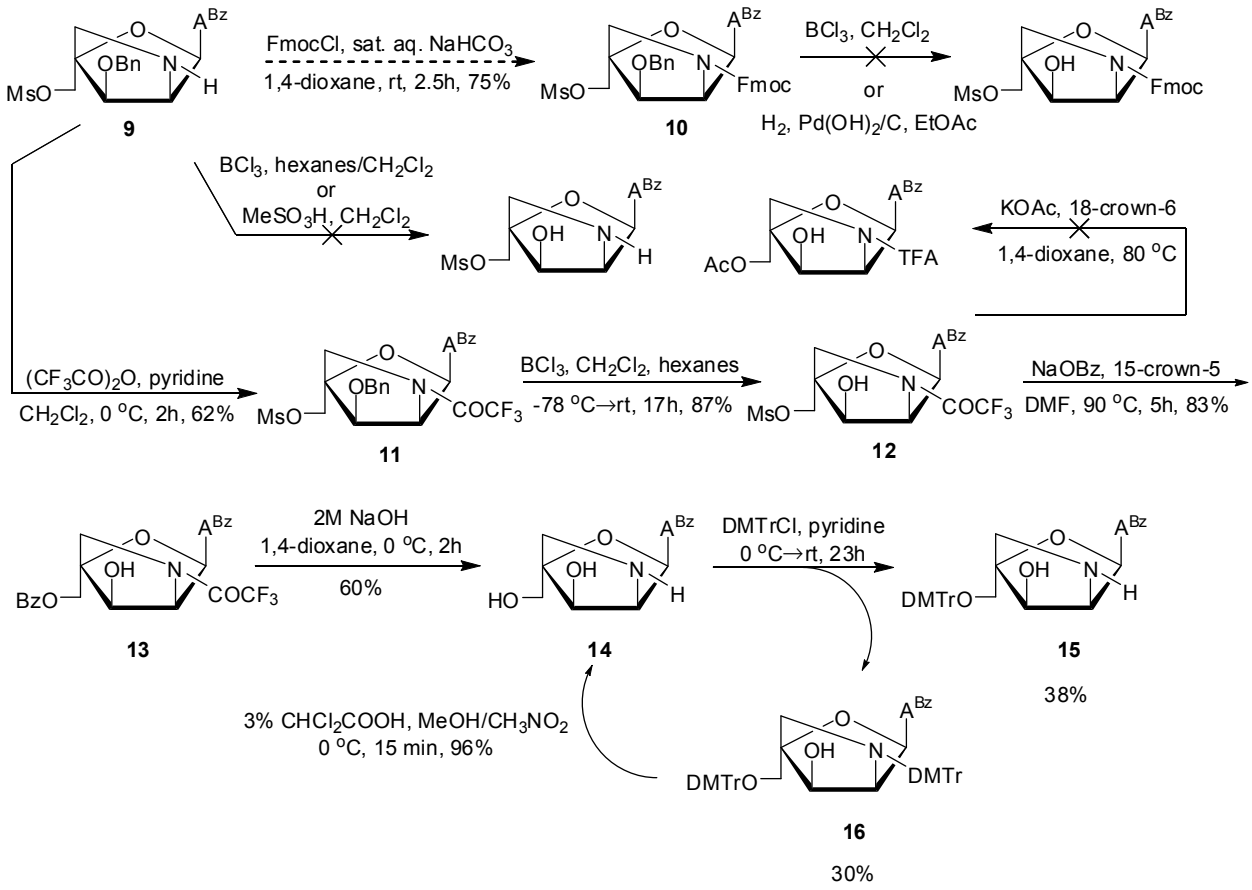
(Scheme 2-2). Glycosyl donor **5** was converted into alcohol **7** following protocols that deviated from the known route[6b] in the following manner: i) the use of trimethylsilyl triflate as Lewis acid and 1,2-dicholoroethane as reaction solvent during Vorbrüggen glycosylation of **5** affords **6** in higher yield and with easier workup than the original protocol involving tin(IV) chloride and acetonitrile (70% vs 57%, respectively), and ii) the use of the guanidinium nitrate/sodium methoxide reagent mixture[17] results in slightly more efficient O2'-deacylation of **6** than the original protocol entailing dilute methanolic ammonia (88% vs 79% yield); however, the dilute reaction conditions of the former approach rendered it less practical for large scale reactions. Alcohol **7** was O2'-triflated and treated with sodium azide and 15-crown-5 in DMF at elevated temperatures to furnish azide **8** in excellent yield (89% over two steps). IR spectra of **8** verified the presence of the azide group (sharp band at 2115 cm^{-1}). Nucleoside **8** was converted into the desired 2'-amino-α-L-LNA intermediate **9** via a one-pot tandem Staudinger/intramolecular nucleophilic substitution reaction using alkaline trimethylphosphine[15] in a satisfying 85% yield.[18] The 2-oxo-5-azabicyclo[2.2.1]heptane skeleton and stereochemical configuration of **9** was verified by NOE difference experiments on downstream products (*vide infra*).



Scheme 2-2: Synthesis of intermediate **9**. BSA = *N*,*O*-bis(trimethylsilyl)acetamide; 1,2-DCE = 1,2-dichloroethane; A^{Bz} = 6-*N*-benzoyl-adenin-9-yl; GNO_3 = guanidinium nitrate; 15-C-5: 15-crown-5.

The protecting group manipulations needed to convert nucleoside **9** into key intermediate **15** proved surprisingly challenging (Scheme 2-3). For example, O3'-debenzylation of bicyclic nucleoside **9** was unsuccessful using boron tricholoride or methanesulfonic acid in dichloromethane. Similarly, the corresponding N2'-Fmoc-protected nucleoside **10** - obtained by reacting **9** with 9'-fluorenylmethyl chloroformate under Schotten-Baumann conditions (results not shown) - failed to undergo O3'-debenzylation with BCl_3 in dichloromethane or by standard hydrogenolysis. Additional attempted strategies are outlined in the Supporting information (Scheme S1).[33] Instead, it proved necessary to protect the N2'-position of bicyclic nucleoside **9** as a trifluoroacetamide to give nucleoside **11** in 62% yield, which was then O3'-debenzylated with BCl_3 in hexanes to give nucleoside **12** in 87% yield (Scheme 2-3). Protection of the O3'-position as a napthyl ether presents itself as an attractive alternative option as it can be cleaved using DDQ,[7a] but this option was not considered at the time of synthesis. Subsequent

nucleophilic substitution of the C5'-mesylate group of **12** was attempted using potassium acetate and 18-crown-6, but this resulted in the formation of numerous side products according to analytical TLC and the approach was abandoned (Scheme 2-3). Instead, treatment of **12** with sodium benzoate and 15-crown-5 in hot DMF afforded O5'-benzoylated nucleoside **13** in 83% yield. Subsequent treatment of nucleoside **13** with aqueous sodium hydroxide in 1,4-dioxane effected the cleavage of both the O5'-benzoyl and N2'-trifluroacetamide protecting groups to afford polar amino alcohol **14** in 60% yield after column chromatography. The O5'-hydroxyl group of nucleoside **14** was subsequently protected as the 4,4'-dimethoxytrityl (DMTr) ether using standard conditions to afford key intermediate **15**. Curiously, the reaction proceeded in no more than 38% yield as significant amounts of N2',O5'-di-DMTr-protected nucleoside **16** were produced as well (30% yield). Efforts to optimize the yield of **15** were unsuccessful (e.g., slow addition of DMTr-Cl at low temperatures). However, it was possible to recycle nucleoside **16** into amino diol **14** using 3% dichloroacetic acid in a mixture of methanol and nitromethane[19] in up to 96% yield. An alternative strategy toward **15**, involving N2'-Fmoc protection of **14** and subsequent O5'-DMTr-protection and N2'-Fmoc deprotection, was dismissed due to inefficient O5'-DMTr protection (~40% yield, results not shown). Thus, the preferred route (**5** $\rightarrow\!\!\rightarrow$ **9** $\rightarrow$ **11** $\rightarrow\!\!\rightarrow$ **15**) affords intermediate **15** in ~5% overall yield from **5**, not taking the recycling step **16** $\rightarrow$ **14** into account.



Scheme 2-3: Synthesis of key intermediate **15.** A^{Bz} = 6-*N*-benzoyl-adenin-9-yl; DMTr = 4,4′-dimethoxytrityl; FmocCl = 9'-fluorenylmethyl chloroformate.

2.2.3 Structural Verification of 2'-amino-α-L-LNA Configuration.

In agreement with previously reported ^{1}H NMR signals of other α-L-LNA nucleosides,[6b,9] the ^{1}H NMR signals of H1', H2' and H3' of 2'-amino-α-L-LNA nucleosides appear as singlets or narrow doublets (J < 2 Hz),[20] since the torsion angles described by H1'-C1'-C2'-H2' and H2'-C2'-C3'-H3' are fixed in +*gauche* and -*gauche* conformations, respectively. The structure of bicyclic nucleoside **14** was ascertained by NOE difference spectroscopy. NOE contacts between H1'/H2' (7%), H1'/H3' (5%) and H2'/H3' (2%) suggest a *cis* relationship between these protons

(Fig. 2-2). Since the stereochemical configuration at C3' is defined by the choice of starting material and remains unchanged throughout synthesis, H1' and H2' must be pointing “down”, which confirms the nucleobase as pointing “up” and hence establishes the 2'-amino-α-L-LNA configuration. This is substantiated by signal enhancements between H5"$_A$/H8' (6%) indicating a *cis* relationship between the nucleobase and H5" (H5"$_A$ is tentatively assigned as the H5" closest to the nucleobase).

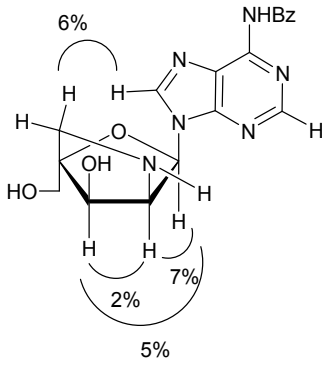
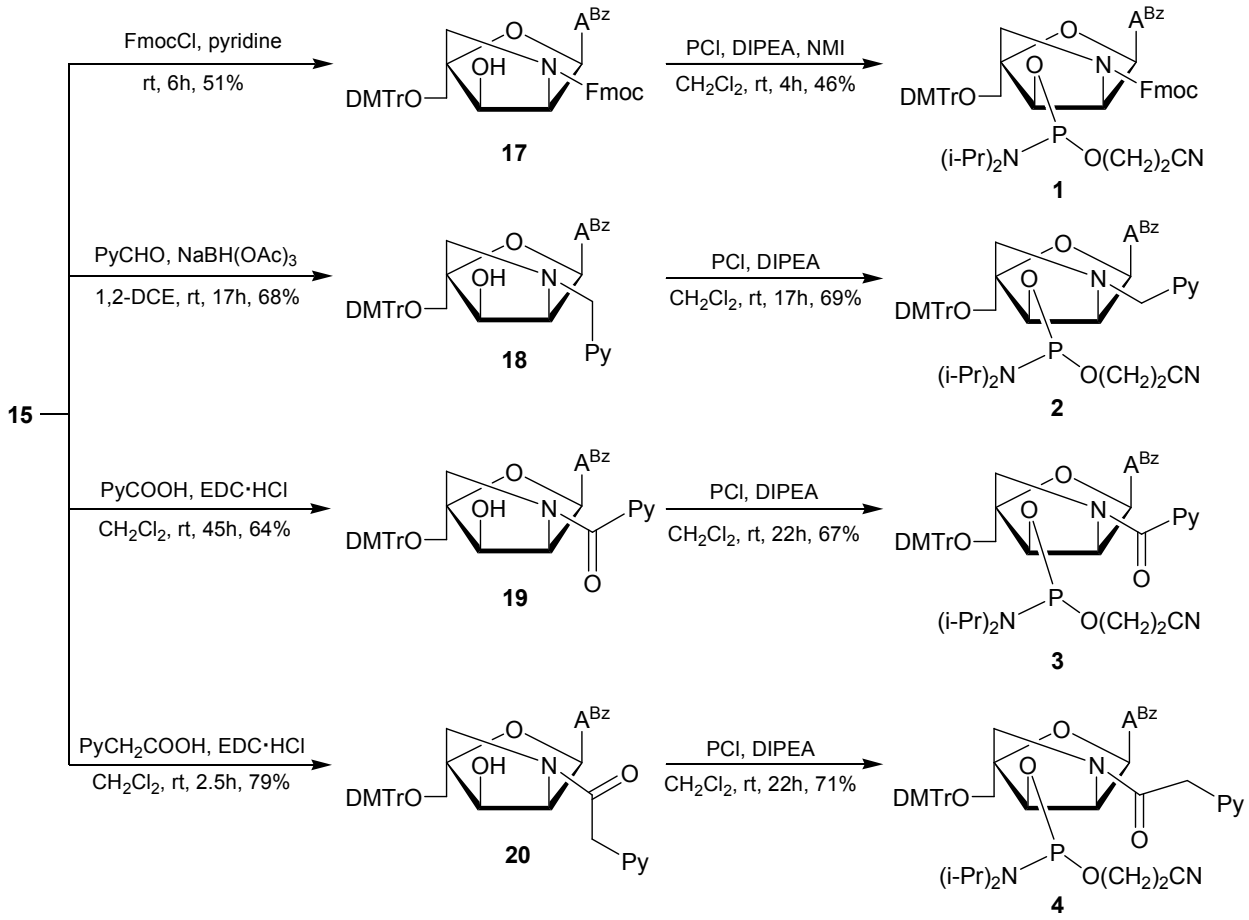


Figure 2-2: Key NOE contacts in nucleoside 14.

2.2.4. Synthesis of Phosphoramidite Building Blocks 1-4

Chemoselective N2'-functionalization of key intermediate **15** to give nucleosides **17**-**20** was realized: i) using 9'-fluorenylmethyl chloroformate (**17**: 51% yield; Schotten-Baumanm conditions could not be used due to the low solubility of **15** in dioxane/water), ii) via reductive amination using 1-pyrenecarboxaldehyde and sodium triacetoxyborohydride[21] as the reducing agent (**18**: 68% yield), or iii) via EDC-mediated coupling of 1-pyrenecarboxylic acid or 1-pyreneacetic acid (**19**: 64% yield; **20**: 79% yield) (Scheme 2-4). Subsequent phosphitylation using 2-cyanoethyl-*N*,*N*'-diisopropylchlorophosphoramidite afforded target compounds **1**-**4** in 46-71% yield after column chromatography and precipitation.



Scheme 2-4: Synthesis of phosphoramidites **1-4**. A^{Bz} = 6-*N*-benzoyl-adenin-9-yl; DIPEA = *N*,*N*'-diisopropylethylamine; DMTr = 4,4'-dimethoxytrityl; EDC·HCl = 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride; FmocCl = 9'-fluorenylmethyl chloroformate; NMI = *N*-methylimadazole; PCl = 2-cyanoethyl-*N*,*N*'-diisopropylchlorophosphoramidite; Py = pyren-1-yl.

2.2.5. Synthesis of Modified ONs and Experimental Design

Phosphoramidites **1-4** were used in machine-assisted solid-phase DNA synthesis (0.2 μmol scale) to incorporate monomers **W**-**Z** into ONs using the following hand-coupling conditions

(activator; coupling time; stepwise coupling yields): monomer **W** (pyridinium hydrochloride; 30 min; ~82%) and monomers **X**-**Z** (pyridinium hydrochloride; 15 min; ~95%). Following workup and HPLC purification, the composition and purity of all modified ONs was ascertained by MALDI MS analysis (Tables S1 and S2 in the Supporting Information[33]) and ion-pair reversed-phase HPLC, respectively. ONs containing a single incorporation in the 5'-GTG **B**TA TGC context are denoted **W1**, **X1**, **Y1** and **Z1**, respectively. Similar conventions apply for ONs in the **B2**-**B6** series (Table 2-1). Reference DNA and RNA strands are denoted **D1**/**D2** and **R1**/**R2**, respectively. The following descriptive nomenclature is also used: N2'-PyMe (**X**-series), N2'-PyCO (**Y**-series) and N2'-PyAc (**Z**-series). We have previously used this 9-mer mixed-sequence context to study the hybridization properties of ONs modified with N2'-pyrene-functionalized 2'-amino-α-L-LNA thymine monomers, which facilitates direct comparison.[9]

2.2.6. Thermal Denaturation Experiments – Thermal Affinity Toward Complementary DNA/RNA.

The thermostability of duplexes between **W**/**X**/**Y**/**Z**-modified ONs and complementary DNA/RNA was evaluated by determining their thermal denaturation temperature (T_m) in medium salt buffer ([Na^+] = 110 mM, pH 7.0). Changes in T_m's of modified duplexes are discussed relative to T_m's of unmodified reference duplexes.

ONs with one or two incorporations of 2'-amino-α-L-LNA adenine monomer **W** display very similar thermal affinity toward complementary DNA as unmodified reference ONs (ΔT_m between -1.0 to +0.5 °C, Table 1). Thus, unlike the corresponding 2'-oxy-α-L-LNA adenine monomer,[6b,22] the 2-oxo-5-azabicyclo[2.2.1]heptane skeleton of monomer **W** is not inherently

beneficial for DNA duplex formation. We have made similar observations with ONs modified with the thymine counterpart of monomer **W**.[9a] This indicates that the unfunctionalized 2'-nitrogen of 2'-amino-α-L-LNA monomers either stabilizes the single-stranded ON or destabilizes the duplex, e.g., by perturbing the hydration spine in the major groove. In contrast, ONs modified with N2'-pyrene-functionalized 2'-amino-α-L-LNA adenine monomers **X**, **Y**, or **Z**, display greatly increased thermal affinity toward DNA targets (ΔT_m/modification between +2.5 and +14 °C, Table 2-1). The observed T_m trends for singly modified ONs (**Y** > **Z** ≥ **X**) indicate that: i) monomers, in which the pyrene moiety is attached to the 2-oxo-5-azabicyclo[2.2.1]heptane skeleton via an alkanoyl linker, induce greater thermostabilization than corresponding monomers employing alkyl linkers (N2'-PyCO **Y** > N2'-PyMe **X**) and ii) monomers with short alkanoyl linkers result in more thermostable DNA duplexes than the corresponding monomers with longer alkanoyl linkers (N2'-PyCO **Y** > N2'-PyAc **Z**). We have observed similar structure-property relationships with the analogous thymine monomers, which – based on results from UV-Vis spectroscopy and molecular modeling – were explained by differential placement of the pyrene moiety for affinity-enhancing intercalation.[9b] Similar structural underpinnings are likely in effect for monomers **X**-**Z** (*vide infra*). However, the electron density of the pyrene moieties, which differs between the three monomers, may also be a contributing factor to the differential duplex stabilization.

An interesting trend is observed for the doubly modified ONs. Thus, incorporation of a second N2'-PyMe **X** or N2'-PyCO **Y** monomer does not result in additionally increased thermal affinity against DNA targets, whereas a second incorporation of the more flexible N2'-PyAc **Z** monomer does (compare ΔT_m for the **B1**-**B3** and **B4**-**B6** series, Table 2-1). We hypothesize that the former observation is due to violation of the 'nearest neighbor exclusion principle', which

states that free intercalators - at most - bind to every second base pair of a DNA duplex due to limits in local expandability of duplexes.[23] Extrapolating the principle to tethered intercalators and assuming 3'-intercalative binding modes of the pyrene moieties of monomers **X**-**Z** (*vide infra*), the resulting duplexes involving **B3** and **B6** would feature a localized region with two intercalators in an area defined by four base pairs, which is at the saturation threshold. We speculate that the greater flexibility of the N2'-PyAc **Z** monomer, allows the modified DNA duplexes to structurally compensate for any stress induced by the high intercalator density.

Table 2-1: T_m values of duplexes between **B1**-**B6** and complementary DNA targets.[a]

			ΔT_m/°C			
ON	**Duplex**	**B =**	**W**	**X**	**Y**	**Z**
B1 **D2**	5'-GTG **B**TA TGC 3'-CAC TAT ACG		-0.5	+5.0	+11.0	+6.0
B2 **D2**	5'-GTG AT**B** TGC 3'-CAC TAT ACG		+0.5	+7.0	+14.0	+7.5
B3 **D2**	5'-GTG **B**T**B** TGC 3'-CAC TAT ACG		-1.0	+5.0	+13.5	+16.0
D1 **B4**	5'-GTG ATA TGC 3'-CAC T**B**T ACG		±0.0	+6.5	+11.5	+7.5
D1 **B5**	5'-GTG ATA TGC 3'-CAC TAT **B**CG		±0.0	+5.5	+12.0	+6.0
D1 **B6**	5'-GTG ATA TGC 3'-CAC T**B**T **B**CG		-2.0	+5.0	+13.5	+14.5

[a] ΔT_m = change in T_m relative to unmodified reference duplex **D1**:**D2** ($T_m \equiv 27.5$ °C); T_m's determined as the maximum of the first derivative of melting curves (A_{260} vs T) recorded in medium salt buffer ($[Na^+] = 110$ mM, $[Cl^-] = 100$ mM, pH 7.0 (NaH_2PO_4/Na_2HPO_4)), using 1.0 µM of each strand. T_m's are averages of at least two measurements within 1.0 °C; A/C/G/T = adenin-9-yl/cytosin-1-yl/guanin-9-yl/thymin-1-yl DNA monomers. For structures of monomers **W**-**Z**, see Figure 2-1. Data for the **B1**/**B2**/**B4**/**B5**-series (for monomers **X**/**Y**/**Z**) have previously been reported in reference 12b.

Interestingly, the modified ONs display rather different thermal denaturation characteristics with RNA targets. Thus, ONs modified with 2'-amino-α-L-LNA adenine monomer **W** display

significantly higher affinity against RNA than DNA targets (ΔT_m/modification between +1.5 and +7.0 °C, Table 2-2). Further, **X**- and **Y**-modified ONs display markedly lower thermal affinity toward RNA than DNA targets, with N2'-PyMe **X**-modified DNA:RNA duplexes displaying similar thermostability as unmodified reference duplexes (ΔT_m/modification between -0.5 and +2.5 °C, Table 2-2) and N2'-PyCO **Y**-modified duplexes being moderately stabilized (ΔT_m/modification between +1.5 and +6.5 °C, Table 2-2). N2'-PyAc **Z**-modified ONs also display lower affinity toward RNA than DNA targets, but result in more thermostable heteroduplexes (ΔT_m/modification between +4.0 and +8.0 °C, Table 2-2). These differences may again reflect the greater flexibility of monomer **Z**, which allows the pyrene moieties to adopt more favorable positions for affinity-enhancing intercalation (*vide infra*). DNA-selective hybridization (Table 2-3) - as seen for **X**/**Y**/**Z**-modified ONs and ONs modified with the thymine counterparts[9b] - is often observed for ONs modified with intercalating moieties,[24] as intercalators favor the *B*-type helix geometry of DNA:DNA duplexes over the more compressed *A*/*B*-type helix geometry of DNA:RNA duplexes.[25]

Table 2-2: T_m values of duplexes between **B1-B6** and complementary RNA targets.[a]

			ΔT_m/°C			
ON	**Duplex**	**B =**	**W**	**X**	**Y**	**Z**
B1 **R2**	5'-GTG **B**TA TGC 3'-CAC UAU ACG		+2.5	-1.0	+1.5	+4.0
B2 **R2**	5'-GTG AT**B** TGC 3'-CAC UAU ACG		+2.5	+1.0	+5.0	+4.5
B3 **R2**	5'-GTG **B**T**B** TGC 3'-CAC UAU ACG		+3.0	-1.5	+4.5	+11.0
R1 **B4**	5'-GUG AUA UGC 3'-CAC T**B**T ACG		+4.5	-0.5	+2.5	+6.5
R1 **B5**	5'-GUG AUA UGC 3'-CAC TAT **B**CG		+7.0	+2.5	+6.5	+8.0
R1 **B6**	5'-GUG AUA UGC 3'-CAC T**B**T **B**CG		+7.0	±0.0	+7.0	+14.5

[a] ΔT_m = change in T_m relative to unmodified reference duplexes **D1**:**R2** ($T_m \equiv 26.0$ °C) and **R1**:**D2** ($T_m \equiv 24.5$ °C); for conditions of thermal denaturation experiments, see Table 1.

Table 2-3: DNA-selectivity of **B1-B6**.[a]

			$\Delta\Delta T_m$ (DNA-RNA)/°C			
ON	**Sequence**	**B =**	**W**	**X**	**Y**	**Z**
B1	5'-GTG **B**TA TGC		-3.0	+6.0	+9.5	+2.5
B2	5'-GTG AT**B** TGC		-2.0	+6.0	+9.0	+3.0
B3	5'-GTG **B**T**B** TGC		-4.0	+6.5	+9.0	+5.0
B4	3'-CAC T**B**T ACG		-4.5	+7.0	+9.0	+1.0
B5	3'-CAC TAT **B**CG		-7.0	+2.0	+5.5	-2.0
B6	3'-CAC T**B**T **B**CG		-9.0	+5.0	+6.5	±0.0

[a] DNA selectivity defined as $\Delta\Delta T_m$ (DNA-RNA) = ΔT_m (vs DNA) - ΔT_m (vs RNA).

2.2.7. Thermal Denaturation Studies – Mismatch Discrimination.

The binding specificity of centrally modified ONs was determined against DNA/RNA strands with mismatched nucleotides opposite of the **W**-**Z** monomers. **W4** displays improved

discrimination of mismatched DNA targets relative to reference strand **D2**, with the AG-mismatch being discriminated particularly well (Table 2-4). **X4**/**Y4**/**Z4** display surprisingly efficient discrimination of mismatched DNA targets - and AG mismatches in particular (Table 2-4) - considering the likely intercalative binding mode of the pyrene moieties, which is known to often decrease base pairing fidelity.[9b,26] However, ONs with **X**/**Y**/**Z** monomers positioned as next-nearest neighbors display poor discrimination of centrally mismatched DNA target (**B3** series, Table S4 in the Supporting Information[33]), indicating that modification patterns have a major influence on binding specificity. Discrimination of RNA mismatches is generally less efficient with **W4**/**X4**/**Y4** than with reference ONs, but improved with **Z4** (Table 2-5).

Table 2-4: Discrimination of mismatched DNA targets by **B4**-series and reference ONs.[a]

			DNA: 5'-GTG A**B**A TGC			
			T_m/°C	ΔT_m/°C		
ON	**Sequence**	**B** =	T	A	C	G
D2	3'-CAC TAT ACG		**27.5**	-17.0	-15.5	-9.0
W4	3'-CAC T**W**T ACG		**27.5**	-20.0	-17.0	-16.0
X4	3'-CAC T**X**T ACG		**34.0**	-16.5	-7.5	-17.0
Y4	3'-CAC T**Y**T ACG		**39.0**	-21.0	-12.0	-17.0
Z4	3'-CAC T**Z**T ACG		**35.0**	-21.5	-14.0	-14.5

[a]For conditions of thermal denaturation experiments, see Table 1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched DNA:DNA duplex.

Table 2-5: Discrimination of mismatched RNA targets by **B4**-series and reference ONs.[a]

			RNA: 5'-GUG A**B**A UGC			
			T_m/°C	ΔT_m/°C		
ON	**Sequence**	**B** =	U	A	C	G
D2	3'-CAC TAT ACG		**24.5**	-15.0	-15.0	-11.0
W4	3'-CAC T**W**T ACG		**29.0**	-13.5	-13.0	-11.5
X4	3'-CAC T**X**T ACG		**24.0**	-10.5	-8.0	-10.0
Y4	3'-CAC T**Y**T ACG		**27.0**	-10.5	-11.0	-8.5
Z4	3'-CAC T**Z**T ACG		**31.0**	<-19.0	-15.5	-14.5

[a] For conditions of thermal denaturation experiments, see Table 1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched RNA:DNA duplex.

2.2.8. Optical Spectroscopy

UV-Vis absorption and steady-state fluorescence emission spectra of **X**/**Y**/**Z**-modified ONs and the corresponding duplexes with DNA/RNA targets were recorded to gain additional insight into the binding modes of the pyrene moieties of monomers **X**/**Y**/**Z**. Bathochromic shifts of pyrene absorption maxima, which are indicative of strong interactions between pyrenes and nucleobases,[27] were generally observed upon hybridization with complementary DNA/RNA ($\Delta\lambda_{max}$ = 0-6 nm, Table 2-6, Figure 2-3, and Figures S2 and S3 in the Supporting Information[33]), with the most pronounced increases being observed for duplex formation of **Y**-modified ONs with DNA. These results are consistent with intercalative binding modes for the pyrene moieties of monomers **X**/**Y**/**Z** and in agreement with our previous observations with the corresponding thymine analogues.[9b]

Table 2-6: Absorption maxima in the 300-400 nm region for ONs modified with N2'-pyrene functionalized 2'-amino-α-L-LNA adenine monomers **X**/**Y**/**Z** in the presence or absence of complementary DNA/RNA targets.[a]

			$\lambda_{max}[\Delta\lambda_{max}]$ (nm)								
		B=	**X**			**Y**			**Z**		
ON	Sequence		SSP	+DNA	+RNA	SSP	+DNA	+RNA	SSP	+DNA	+RNA
B1	5'-GTG **B**TA TGC		348	350 [+2]	350 [+2]	349	352 [+3]	351 [+2]	348	350 [+2]	351 [+3]
B2	5'-GTG AT**B** TGC		348	350 [+2]	349 [+1]	349	353 [+4]	351 [+2]	350	351 [+1]	351 [+1]
B3	5'-GTG **B**T**B** TGC		347	349 [+2]	348 [+1]	349	351 [+2]	351 [+2]	347	350 [+3]	351 [+4]
B4	3'-CAC T**B**T ACG		348	350 [+2]	348 [±0]	347	353 [+6]	352 [+5]	348	351 [+3]	351 [+3]
B5	3'-CAC TAT **B**CG		348	350 [+2]	350 [+2]	350	351 [+1]	350 [±0]	348	351 [+3]	351[+3]
B6	3'-CAC T**B**T **B**CG		348	350 [+2]	348 [±0]	348	352 [+4]	353 [+5]	348	351 [+3]	351 [+3]

[a] $\Delta\lambda_{max}$ = change in absorption maximum relative to single stranded probe (SSP). Measurements were performed at 5 °C except for single-stranded **X**/**Y**-modified probes, which were recorded at room temperature. Buffer conditions are as for thermal denaturation experiments.

Steady-state fluorescence emission spectra were recorded at 5 °C using an excitation wavelength of λ_{ex} = 350 nm. Hybridization of N2'-PyMe **X**-modified or N2'-PyAc **Z**-modified ONs with complementary DNA/RNA results in decreased emission (Figure 2-3 and Figures S4 and S7 in the Supporting Information[33]), which is consistent with intercalation-induced quenching by flanking nucleobases.[27,28] The resulting duplexes are only weakly fluorescent and exhibit defined vibronic bands at ~380 nm and ~400 nm along with a shoulder at 420-425 nm. On the other hand, a range of responses is observed upon hybridization of N2'-PyCO **Y**-modified ONs with complementary DNA/RNA, varying from ~50% reduction to ~3-fold enhancement of emission (Figure S5 in the Supporting Information[33]). The resulting duplexes are strongly fluorescent (nearly two orders of magnitude more fluorescent than **X**-modified duplexes) with little or no vibronic fine structure ($\lambda_{em,max}$ ~400 nm, broad band). The lack of a consistent response upon hybridization to DNA/RNA presumably is not due to different pyrene

binding modes in the resulting duplexes but rather a reflection of different fluorescence intensities of the single-stranded probes, possibly due to different sequence-dependent rotational barriers around the Py-CO bond. Thus, singly modified DNA duplexes display similar emission intensities (Figure S6 in the Supporting Information[33]), suggesting that the pyrene moieties are in similar microenvironments in the duplex, which is consistent with an intercalative binding mode.

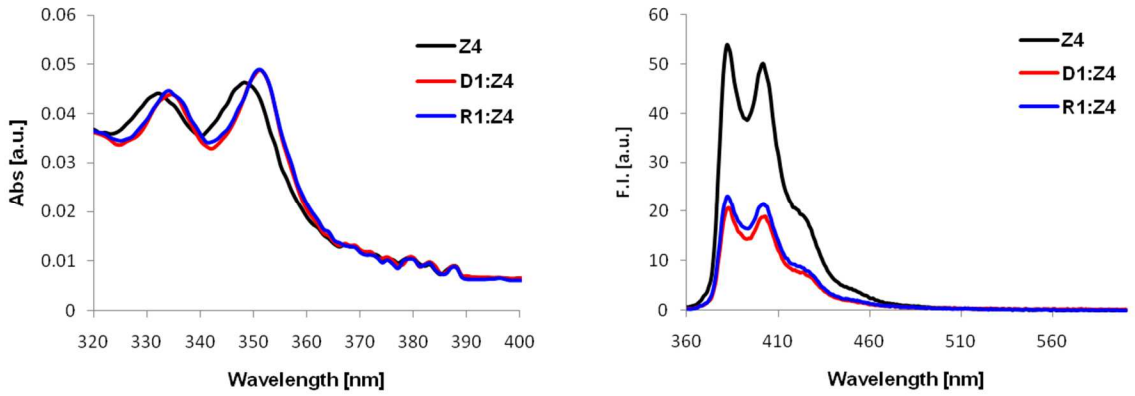


Figure 2-3: UV-vis absorption (left) and steady-state fluorescence emission (right) spectra of **Z4** in the presence or absence of complementary DNA/RNA. T_{exp} = 5 °C; each ON used at 1 μM concentration in T_m buffer; λ_{ex} = 350 nm (fluorescence).

To further substantiate the proposed intercalative binding mode of N2'-PyCO monomer **Y**, we synthesized three additional **Y**-modified ONs in which the 3'-flanking nucleotide was varied systematically (**Y7**-**Y9**, Table 2-7). ONs with 3'-flanking purines display greater relative increases in thermal affinity against DNA targets than the corresponding ONs with 3'-flanking pyrimidines (compare ΔT_m for **Y7**:**D3** and **Y9**:**D5** vs **Y8**:**D4** and **Y4**:**D1**, Table 2-7). This is

consistent with the proposed binding mode as 3'-intercalating pyrenes are expected to interact more strongly with large purines.

Table 2-7: T_m values for duplexes between **Y4**/**Y7**/**Y8**/**Y9** and complementary DNA targets.[a]

ON	Duplex	ΔT_m/°C	reference T_m/°C
D3 **Y7**	5'-GTG TT ATGC 3'-CAC A**Y** TACG	+17.0	27.5
D4 **Y8**	5'-GTG GT ATGC 3'-CAC C**Y** TACG	+10.0	33.0
D5 **Y9**	5'-GTG CT ATGC 3'-CAC G**Y** TACG	+19.5	24.0[b]
D1 **Y4**	5'-GTG AT ATGC 3'-CAC T**Y** TACG	+11.5	27.5

[a]ΔT_m = change in T_m relative to unmodified reference duplex. For experimental conditions, see Table 2-1**.**
[b]The low T_m of this reference duplex was confirmed by two independent operators using strands from three different commercial batches.

Steady-state fluorescence emission experiments provide further evidence for an intercalative binding mode of the pyrenes as the emission levels of these four DNA duplexes decrease in the anticipated order of nucleobase quenching efficiency (from least to most quenched: 3'-flanking A > T > C > G, Fig. 2-4).[28b,29]

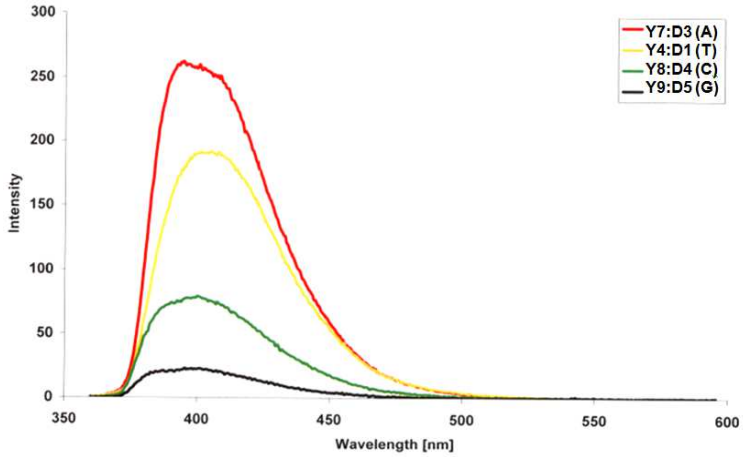


Figure 2-4: Fluorescence emission spectra of duplexes between **Y4**/**Y7**/**Y8**/**Y9** and complementary DNA (nucleotide flanking the Y monomer on its 3'-side is listed in the parenthesis). T_{exp} = 5 °C; each ON used at 0.15 μM concentration in T_m buffer; λ_{ex} = 350 nm.

2.2.9 Invader-mediated Recognition of DNA Hairpins

2.2.9.1 Introduction

Based on the high thermal stability of the 2'-N-pyrene-functionalized-2'-amino-α-L-LNA adenine monomers toward cDNA, we chose the study these nucleotides as components of probes for recognition of dsDNA (Figure 2-5). Briefly described, Invader probes are double-stranded probes that are activated for dsDNA recognition through modification with one or more "+1 interstrand zippers" of intercalator-functionalized nucelotides (For a definition on zipper nomenclature, see ref. 34). This particular monomer arrangement results in duplex destabilization, presumably because the pyrene moieties are forced to intercalate into the same region, leading to excessive local duplex unwinding and the formation of "energetic hotspots" (Figure 2-5).[12] On the other hand, the two strands that constitute an Invader probe display very

strong affinity toward complementary single-stranded DNA (ssDNA) as a result of efficient pyrene intercalation and π–π stacking with neighboring base pairs (Figure 2-5).[9b,12]

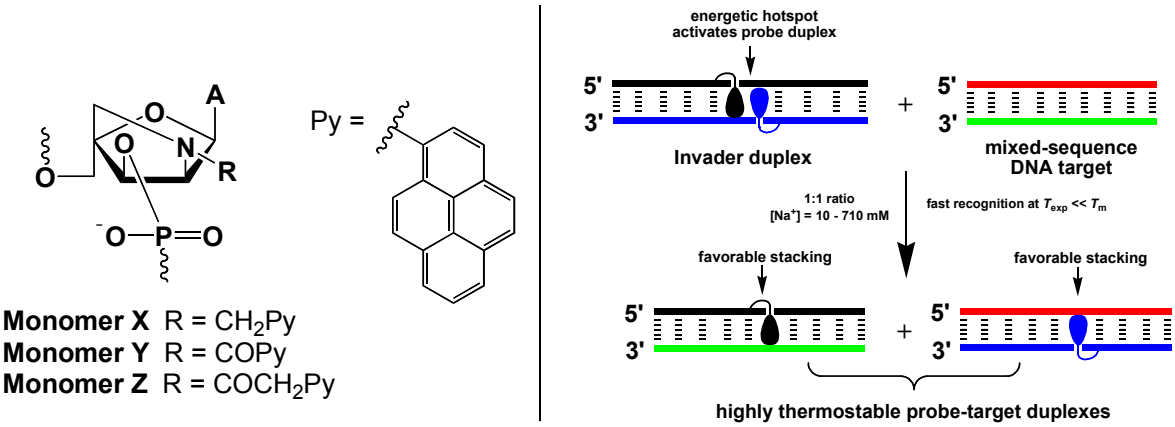


Figure 2-5: Illustration of the Invader approach for mixed-sequence recognition of dsDNA and structures of monomers used herein. Droplets denote intercalating pyrene moieties.

2.2.9.2 Thermal Stability and *TA* Values of Interstrand Zipper Arrangements

The thermostability of DNA duplexes with different interstrand zipper arrangements of pyrene-functionalized monomers was measured to identify monomers and probe architectures that are activated for dsDNA recognition via the Invader strategy (Table 2-8). To estimate a probes propensity for dsDNA recognition, we use the term *thermal advantage* (*TA*) given as $TA = \Delta T_m$ (ON_A:cDNA) + ΔT_m(cDNA:ON_B) - ΔT_m (ON_A:ON_B), where ON_A:ON_B is a duplex with an interstrand zipper arrangement of monomers. A positive *TA* value suggest that a probe is highly activated for recognition of isosequential dsDNA (via the process depicted in Figure 2-5) since the products of the recognition reaction (i.e., probe-target duplexes) are more stable than the reactants (i.e., double-stranded probes and target duplexes).

Double-stranded probes with +1 interstrand zippers of N2'-pyrene-functionalized 2'-amino-α-L-LNA adenine monomers are more thermolabile and energetically activated for dsDNA recognition than probes with other zipper configurations (compare T_m's and *TA* values for **B1**:**B4** and **B2**:**B5** relative to other probe duplexes, Table 2-8). N2'-pyrenecarbonyl 2'-amino-α-L-LNA adenine modified duplexes **Y2**:**Y5** and **Y1**:**Y4** are the most strongly activated probes for dsDNA-recognition in this series (*TA* trend: **Y**>**Z**≥**X**, Table 2-8).

Table 2-8: Thermal Stability and *TA* values of Interstrand Zipper Arrangements.[a]

				ΔT_m/°C[*TA* /°C]		
ON	**ZP**	**Sequence**	**B =**	**X**	**Y**	**Z**
B1	**+3**	5′-GTG **B**TA TGC		+12.0	+13.0	+16.0
B5		3′-CAC TAT **B**CG		[-1.5]	[+10.0]	[-4.0]
B1	**+1**	5′-GTG **B**TA TGC		**-7.0**	**-8.0**	**-7.5**
B4		3′-CAC T**B**T ACG		**[+18.5]**	**[+30.5]**	**[+21.0]**
B2	**+1**	5′-GTG AT**B** TGC		**-5.0**	**-7.0**	**-8.0**
B5		3′-CAC TAT **B**CG		**[+17.5]**	**[+33.0]**	**[+21.5]**
B2	**-1**	5′-GTG AT**B** TGC		+14.0	+23.5	+22.0
B4		3′-CAC T**B**T ACG		[-0.5]	[+2.0]	[-7.0]

[a] ΔT_m = change in T_m value relative to the unmodified reference duplex **D1**:**D2** (T_m = 29.5 ° C); see Table 2-1 for the experimental conditions.

The results indicate that high dsDNA-targeting potential is an inherent characteristic of probes with +1 interstrand zippers of intercalator-functionalized nucleotides.

2.2.9.3 Recognition of DNA Hairpins by Invader Probes

Next, we examined the dsDNA-targeting characteristics of double-stranded probes with interstrand arrangements of 2'-N-pyrene-functionalized-2'-amino-α-L-LNA adenine monomers. Assuming that efficient recognition of dsDNA targets requires probes that are

thermally activated and display low thermostability (i.e., $TA \gg 0$ °C and $\Delta T_m \leq 0$ °C), we decided to focus our efforts on probes with +1 interstrand monomer arrangements. We therefore set out to experimentally test the recognition efficiency of these probes using a 3′-digoxigenin (DIG) labeled DNA hairpin (**DH1**) – comprised of a 9-mer double-stranded mixed sequence stem that is linked by a T_{10} loop – as a model dsDNA target (Figure 2-6a). Incubation of **DH1** with **X1**:**X4** or **Z1**:**Z4** in HEPES buffer at ambient temperature for 3 hours resulted in relatively poor recognition efficiency (<20% recognition at 100-fold excess), whereas Invader probes based on the corresponding thymine monomers resulted in 38-48% recognition at 100-fold excess depending on the pyrene linker chemistry.[12b]

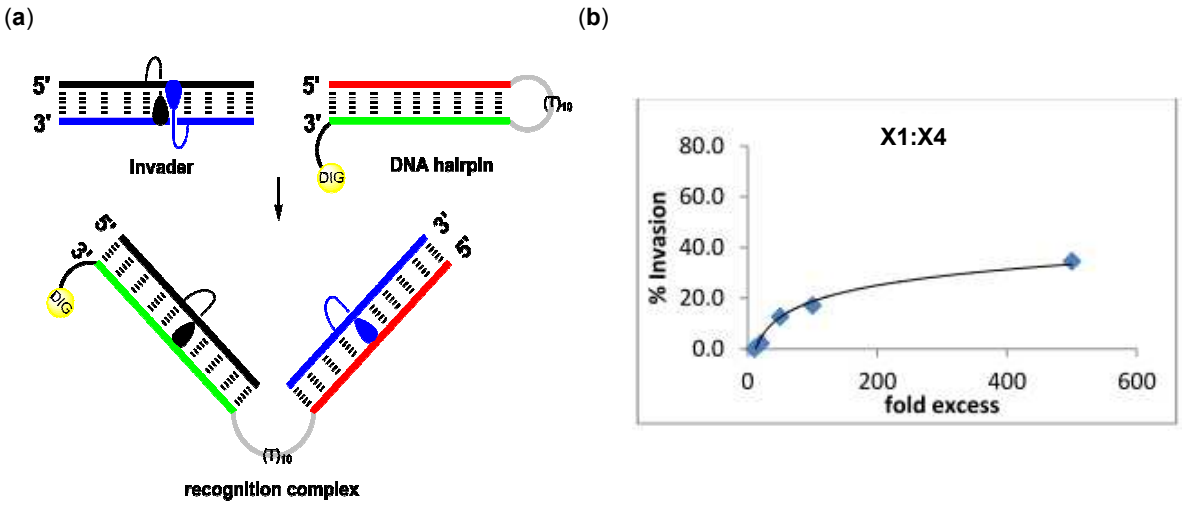


Figure 2-6: Recognition of DNA hairpins using activated double-stranded probes.* (a) Illustration of recognition process; (b) dose-response curve for recognition of DNA hairpin **DH1** by Invader **X1**:**X4**. Probe target incubation: 3 h at 20 °C, 15% non-denaturing PAGE. DIG = digoxigenin. [*This mobility shift assay was conducted by Sujay P. Sau].[12b]

2.3. Conclusions

ONs modified with N2'-pyrene-functionalized 2'-amino-α-L-LNA adenine monomers **X**-**Z** display very high affinity toward DNA targets. The DNA-selective hybridization, together with hybridization-induced bathochromic shifts of pyrene absorption maxima and quenching of pyrene fluorescence, is indicative of intercalative binding modes for the pyrene moieties. ONs with such characteristics are likely to be of significant interest for applications in nucleic acid diagnostics and biotechnology.[30] However, the synthesis of the corresponding phosphoramidites is very challenging due to non-trivial group manipulations (<1% overall yield from diacetone-α-D-glucose), which emphasizes the need for a more efficient synthetic route toward these building blocks or access to functional analogues of N2'-functionalized 2'-amino-α-L-LNA monomers, which are easier to synthesize. Invaders probes modified with the 2'-*N*-pyrene-functionalized-2'-amino-α-L-LNA adenine result in less efficient recognition of dsDNA in our DNA hairpin assay relative to the correspondingly modified thymine Invaders. These results indicate there is a need for identification of simpler and more readily accessible scaffolds in order to conduct structure-property relationship studies with the aim of optimizing the dsDNA binding affinity of Invader probes.

2.4. Experimental Section

2.4.1. Synthesis of N2'-pyrene-functionalized 2'-amino-α-L-LNA adenine Phosphoramidites

9-[2-*O*-Acetyl-3-*O*-benzyl-5-*O*-methanesulfonyl-4-*C*-methanesulfonyloxymethyl-α-L-threo-pentofuranosyl]-6-*N*-benzoyladenine (**6**). 6-*N*-benzoyladenine (28.6 g, 0.12 mol) and glycosyl donor **5**[6b] (40.6 g, 79.8 mmol) were co-evaporated with 1,2-dichloroethane (2×150

mL) and resuspended in anhydrous 1,2-dichloroethane (270 mL). To this was added *N*,*O*-bis(trimethylsilyl)acetamide (BSA, 49.2 mL, 0.20 mol) and the suspension was heated at reflux until turning homogenous. The solution was then cooled to rt, TMSOTf (43.1 mL, 0.24 mol) was slowly added, and the reaction mixture heated at reflux for 70h. The mixture was then cooled to rt and slowly poured into a solution of sat. aq. $NaHCO_3$ and crushed ice (500 mL, 1:1, v/v). Additional crushed ice (~400 mL) was added and the mixture was stirred for 30 min. The resulting precipitate was removed via filtration and the filtrate was extracted with CH_2Cl_2 (1.5 L). The organic layer was washed with sat. aq. $NaHCO_3$ (2×500 mL) and the combined aqueous layer back-extracted with CH_2Cl_2 (2×500 mL). The combined organic layer was evaporated to afford a crude residue, which was purified by silica gel column chromatography (0-10% *i*-PrOH in CH_2Cl_2, v/v) to provide nucleoside **6** (38.3 g, 70%) as a white foam. R_f = 0.4 (5% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m*/*z* 712.1380 ($[M+Na]^+$, $C_{29}H_{31}N_5O_{11}S_2 \cdot Na^+$, Calc. 712.1354). The observed ^{13}C NMR data (75.5 MHz, $CDCl_3$) are in good agreement with previously reported data for this compound.[6b]

9-[3-*O*-Benzyl-5-*O*-methanesulfonyl-4-*C*-methanesulfonyloxymethyl-α-L-threo-pentofuranosyl]-6-*N*-benzoyladenine (7). Fully protected nucleoside **6** (4.66 g, 6.76 mmol) was dissolved in solution[17] of guanidinium nitrate (4.91 g, 40.2 mmol) and NaOMe (0.24 g, 4.44 mmol) in MeOH:CH_2Cl_2 (450 mL, 9:1, v/v). The reaction mixture was stirred at rt for 30 min at which point sat. aq. NH_4Cl (200 mL) was added. The resulting white precipitate was filtered off, washed with CH_2Cl_2, and the filtrate concentrated. The aqueous layer was extracted with CH_2Cl_2 (5×200 mL) and the combined organic layers were evaporated to near dryness. The resulting crude residue was purified by silica gel column chromatography (0-4% MeOH in

CH_2Cl_2, v/v) to afford alcohol **7** (3.84 g, 88%) as a white foam. R_f = 0.5 (10% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m*/*z* 670.1234 ($[M+Na]^+$, $C_{27}H_{29}N_5O_{10}S_2 \cdot Na^+$, Calc. 670.1248). The observed ^{13}C NMR data (75.5 MHz, $CDCl_3$) are in good agreement with previously reported data for this compound.[6b]

9-[2-*C*-Azido-3-*O*-benzyl-2-deoxy-4-*C*-methanesulfonyloxymethyl-5-*O*-methanesulfonyl-α-L-*erythro*-pentofuranosyl]-6-*N*-benzoyladenine (**8**). Alcohol **7** (18.6 g, 28.7 mmol) was co-evaporated with anhydrous pyridine (50 mL) and dissolved in anhydrous CH_2Cl_2 (200 mL). The solution was cooled to -78 °C and anhydrous pyridine (7.3 mL, 90.6 mmol) was added, followed by dropwise addition of trifluoromethanesulfonyl anhydride (Tf_2O, 9.90 mL, 58.9 mmol). The reaction mixture was allowed to warm to rt and stirred for 3h. At this point, crushed ice (50 mL) was added, and the layers were separated. The organic layer was washed with sat. aq. $NaHCO_3$ (2×50 mL), evaporated to near dryness, and co-evaporated with abs. EtOH (2×50 mL) to afford the crude O2'-triflate as a lightly brown residue, which was used in the next step without further purification.

NaN_3 (19.6 g, 0.30 mol) and 15-crown-5 (6.0 mL, 0.30 mol) were added to a solution of the crude O2'-triflate in anhydrous DMF (300 mL). The reaction mixture was first stirred at rt for 15h, then for an additional 8h at 50 °C. After cooling the mixture to rt, solids were filtered off and washed with EtOAc, and the combined organic layer was concentrated until near dryness. The concentrate was taken up in EtOAc (200 mL) and brine (200 mL), the layers were separated, and the aqueous layer extracted with EtOAc (4×100 mL). The combined organic layer was evaporated to near dryness and the resulting crude residue was purified by silica gel column chromatography (0-90% EtOAc in petroleum ether, v/v) to afford azide **8** (17.9 g, 89%

over two steps) as a white solid material. R_f = 0.4 (EtOAc); IR (KBr): 2116 cm^{-1} (N_3); MALDI-HRMS *m*/*z* 695.1295 ($[M+Na]^+$, $C_{27}H_{28}N_8O_9S_2 \cdot Na^+$, Calc. 695.1313); ^1H NMR (300 MHz, DMSO-d_6) δ 11.30 (s, 1H, ex), 8.79 (s, 1H), 8.54 (s, 1H), 8.05 (d, 2H, *J* = 7.0 Hz), 7.30-7.67 (m, 8H), 6.74 (d, 1H, *J* = 4.4 Hz), 5.08-5.16 (m, 1H), 4.89 (d, 1H, *J* = 5.1 Hz), 4.74-4.83 (m, 2H), 4.69-4.72 (d, 1H, *J* = 11.4 Hz), 4.47-4.51 (d, 1H, *J* = 11.4 Hz), 4.42 (s, 2H), 3.28 (s, 3H), 3.24 (s, 3H); ^{13}C NMR (75.5 MHz, DMSO-d_6) δ151.9, 150.4, 142.7, 136.9, 133.2, 132.5, 128.5, 128.0, 81.9, 81.7, 80.1, 73.5, 68.3, 61.9, 36.98, 36.94. The carbonyl group of the 6-*N*-benzoyl group was not visible.

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-7-benzyloxy-1-methanesulfonyloxymethyl-2-oxa-5-azabicyclo[2.2.1]heptane (**9**). Aqueous NaOH (2M, 38.9 mL, 77.8 mmol) and trimethylphosphine (1M in THF, 77.8 mL, 77.8 mmol) were added to an ice-cold solution of azido nucleoside **8** (35.2 g, 51.8 mmol) in THF (500 mL). The reaction mixture was allowed to warm up to rt and was stirred at this temperature for 21h. The mixture was then evaporated to near dryness and the resulting crude residue taken up in EtOAc (200 mL) and brine (200 mL). After separating the layers, the aqueous layer was extracted with MeOH:CH_2Cl_2 (3×200 mL, 2:8, v/v). The combined organic layer was evaporated to dryness, and the resulting crude purified by silica gel column chromatography (0-5% MeOH in CH_2Cl_2, v/v) to provide bicyclic nucleoside **9** (24.2 g, 85%) as a solid brown material. R_f = 0.3 (EtOAc); MALDI-HRMS *m*/*z* 573.1517 ($[M+Na]^+$, $C_{26}H_{26}N_6O_6S \cdot Na^+$, Calc. 573.1527); ^1H NMR[31] (500 MHz, DMSO-d_6) δ 11.17 (s, 1H, ex, NH), 8.77 (s, 1H, H8), 8.73 (s, 1H, H2), 8.06 (d, 2H, *J* = 7.0 Hz, Bz), 7.54-7.67 (m, 3H, Bz), 7.29-7.47 (m, 5H, Ph), 6.52 (d, 1H, *J* = 1.8 Hz, H1'), 4.72-4.76 (d, 1H, *J* = 11.7 Hz, CH_2Ph), 4.62-4.67 (d, 1H, *J* = 11.7 Hz, CH_2Ph), 4.57-4.60 (d, 1H, *J* = 11.7 Hz, H5'$_a$),

4.49-4.53 (d, 1H, J = 11.7 Hz, H5'$_b$), 4.45 (s, 1H, H3'), 3.93 (br s, 1H, H2'), 3.28-3.31 (m, 1H, H5"$_a$, partial overlap with H_2O), 3.22 (s, 3H, CH_3SO_2), 3.10-3.13 (d, 1H, J = 9.9 Hz, H5"$_b$); ^{13}C NMR (125 MHz, DMSO-d_6) δ165.5, 152.1, 151.4 (C2), 150.0, 143.1 (C8), 137.8, 133.3, 132.3 (Bz), 128.4 (Ar), 128.2 (Ar), 127.62 (Ar), 127.60 (Ar), 125.1, 87.2, 84.3 (C1'), 80.4 (C3'), 71.0 (CH_2Ph), 66.8 (C5'), 59.8 (C2'), 51.1 (C5"), 36.8 (CH_3SO_2). The ^{1}H and ^{13}C NMR data are in reasonable agreement with previously reported data from the patent literature.[18] The 2-oxo-5-azabicyclo[2.2.1]heptane skeleton and stereochemistry of 9 was verified via NOE experiments on downstream product **14**.

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-7-benzyloxy-1-methanesulfonyloxymethyl-5-trifluoroacetyl-2-oxa-5-azabicyclo[2.2.1]heptane (**11**). Bicyclic nucleoside 9 (8.68 g, 15.8 mmol) was co-evaporated with pyridine (2×20 mL) and dissolved in anhydrous CH_2Cl_2 (200 mL) and anhydrous pyridine (5.09 mL, 63 mmol). The solution was cooled to 0 °C and trifluoroacetic acid anhydride (4.45 mL, 31.5 mmol) was added. The reaction mixture was stirred for 2 h at 0 °C, at which point crushed ice (50 mL) was added. The layers were separated and the organic layer was washed with sat. aq. $NaHCO_3$ (2×50 mL). The combined aqueous layer was back-extracted with CH_2Cl_2 (2×100 mL) and MeOH:CH_2Cl_2 (100 mL, 2:8, v/v), and the combined organic layer evaporated to near dryness. The resulting crude residue was sequentially co-evaporated with toluene (50 mL) and abs. EtOH:toluene (50 mL, 1:1, v/v) and purified by silica gel column chromatography (0-100% EtOAc in petroleum ether, v/v) to afford fully protected nucleoside **11** (6.27 g, 62%) as a white foam. R_f = 0.5 (10% MeOH:EtOAc, v/v). Physical data for mixture of rotamers (~4:6 by ^{1}H NMR): MALDI-HRMS m/z 647.1541 ([M+H]$^+$, $C_{28}H_{25}F_3N_6O_7S{\cdot}H^+$, Calc. 647.1530); ^{1}H NMR[31,32] (500 MHz, DMSO-d_6) δ11.21 (s,

0.6H, ex, NH_B), 11.19 (s, 0.4H, ex, NH_A), 8.78 (s, 0.6H, $H2_B$), 8.76 (s, 0.4H, $H2_A$), 8.63 (s, 0.4H, $H8_A$), 8.60 (s, 0.6H, $H8_B$), 8.05 (d, 2H, J = 7.0 Hz, Bz-$_{A+B}$), 7.52-7.67 (m, 3H, Bz-$_{A+B}$), 7.31-7.41 (m, 5H, Ph-$_{A/B}$), 6.83 (d, 0.6H, J = 1.1 Hz, $H1'_B$), 6.80 (d, 0.4H, J = 1.1 Hz, $H1'_A$), 5.26 (s, 0.4H, $H2'_A$), 5.17 (s, 0.6H, $H2'_B$), 4.84 (s, 0.6H, $H3'_B$), 4.82 (s, 0.4H, $H3'_A$), 4.62-4.79 (m, 4H, CH_2Ph-$_{A+B}$, $H5'_{A+B}$), 4.53 (d, 0.4H, J = 10.6 Hz, $H5''_A$), 4.35 (d, 0.6H, J = 12.1 Hz, $H5''_B$), 4.07 (d, 0.4H, J = 10.6 Hz, $H5''_A$), 3.91 (d, 0.6H, J = 12.1 Hz, $H5''_B$), 3.28 (s, 3H, CH_3SO_2); ^{13}C NMR (125 MHz, DMSO-d_6) δ 165.49, 165.48, 155.3 (q, J = 37 Hz, $COCF_3$), 155.0 (q, J = 37 Hz, $COCF_3$), 151.8 ($C2_B$), 151.65 ($C2_A$), 151.59, 150.4, 150.3, 141.2 ($C8_B$), 141.0 ($C8_A$), 137.2, 137.0, 133.3, 132.4 (Bz), 128.43 (Bz), 128.40 (Ar), 128.39 (Ar), 128.38 (Ar), 128.34 (Ph), 127.91 (Ph), 127.88 (Ph), 127.53 (Ph), 127.51 (Ph), 125.5, 125.3, 115.5 (q, J = 288 Hz, CF_3), 115.2 (q, J = 288 Hz, CF_3), 86.16, 86.15, 84.9 ($C1'_A$), 84.0 ($C1'_B$), 79.2 ($C3'_B$), 77.4 ($C3'_A$), 71.6 (CH_2Ph), 65.3 (C5'), 65.0 (C5'), 63.0 ($C2'_B$), 61.4 ($C2'_A$), 53.2 ($C5''_B$), 53.1 ($C5''_A$), 37.0 (CH_3SO_2); ^{19}F NMR (376 MHz, DMSO-d_6) δ -71.3 (CF_3-$_B$), -72.1 (CF_3-$_A$).

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-7-hydroxy-1-methanesulfonyloxymethyl-5-trifluoroacetyl-2-oxa-5-azabicyclo[2.2.1]heptane (**12**). Fully protected nucleoside **11** (19.6 g, 30.4 mmol) was co-evaporated with 1,2-dichloroethane (3×100 mL) and dissolved in anhydrous CH_2Cl_2 (600 mL). The solution was cooled to -78 °C and BCl_3 (1M solution in hexanes, 370 mL, 0.37 mol) was added. The reaction mixture was allowed to warm to rt and was stirred for 17h. The mixture was then cooled to 0 °C and crushed ice (800 mL) was slowly added. The layers were separated and the organic phase was washed with sat. aq. $NaHCO_3$ (2×300 mL). The combined aqueous layer was back-extracted with EtOAc (4×500 mL) and the combined organic phase was evaporated to near dryness. The resulting residue was purified using silica

gel column chromatography (0-20% MeOH in CH_2Cl_2, v/v) to afford alcohol **12** (14.7 g, 87%) as a white solid material. R_f = 0.3 (50% acetone in CH_2Cl_2, v/v). Physical data for mixture of rotamers (~4:6 by ^{1}H NMR): MALDI-HRMS *m*/*z* 579.0871 ([M+Na]$^+$, $C_{21}H_{19}F_3N_6O_7S{\cdot}Na^+$, Calc. 579.0880); ^{1}H NMR[31,32] (500 MHz, DMSO-d_6) δ11.21 (s, 1H, ex, NH_{A+B}), 8.76 (s, 0.6H, $H2_B$), 8.75 (s, 0.4H, $H2_A$), 8.60 (s, 0.4H, $H8_A$), 8.56 (s, 0.6H, $H8_B$), 8.04 (d, 2H, *J* = 7.7 Hz, Bz_{A+B}), 7.48-7.66 (m, 3H, Bz_{A+B}), 6.83 (d, 0.6H, *J* = 1.4 Hz, $H1'_B$), 6.80 (d, 0.4H, *J* = 1.4 Hz, $H1'_A$), 6.72 (d, 0.6H, ex, *J* = 4.1 Hz, 3'-OH_B), 6.68 (d, 0.4H, ex, *J* = 4.1 Hz, 3'-OH_A), 4.91 (br s, 0.4H, $H2'_A$), 4.82 (d, 0.6H, *J* = 4.1 Hz, $H3'_B$), 4.76 (d, 0.4H, *J* = 4.1 Hz, $H3'_A$), 4.70 (br s, 0.6H, $H2'_B$), 4.60-4.67 (m, 2H, $H5'_{A+B}$, 4.46 (d, 0.4H, *J* = 10.3 Hz, $H5''_A$), 4.26 (d, 0.6H, *J* = 11.7 Hz, $H5''_B$), 4.03 (d, 0.6H, *J* = 10.3 Hz, $H5''_A$), 3.86 (d, 0.5H, *J* = 11.7 Hz, $H5''_B$), 3.29 (s, 3H, CH_3SO_2); ^{13}C NMR (125 MHz, DMSO-d_6) δ165.5, 155.3 (q, *J* = 36.6 Hz, $COCF_3$), 155.1 (q, *J* = 36.6 Hz, $COCF_3$), 151.7 ($C2_{A/B}$), 151.6 ($C2_{A/B}$), 150.3, 150.2, 141.2 ($C8_B$), 141.0 ($C8_A$), 132.4 (Bz), 128.44 (Bz), 128.42 (Bz), 128.39 (Bz), 128.38 (Bz), 125.5, 125.3, 115.5 (q, *J* = 288 Hz, CF_3), 115.2 (q, *J* = 288 Hz, CF_3), 87.0, 85.8, 84.8 ($C1'_A$), 83.9 ($C1'_B$), 72.5 ($C3'_B$), 70.7 ($C3'_A$), 65.7 (C5'), 65.4 (C5'), 65.3 ($C2'_B$), 63.7 ($C2'_A$), 52.8 ($C5''_B$), 52.6 ($C5''_A$), 37.0 (CH_3SO_2); ^{19}F NMR (376 MHz, DMSO-d_6) δ-71.1 (CF_3-$_B$), -72.1 (CF_3-$_A$).

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-1-benzoyloxymethyl-7-hydroxy-5-trifluoroacetyl-2-oxa-5-azabicyclo[2.2.1]heptane (**13**). NaOBz (2.99 g, 20.8 mmol) and 15-crown-5 (2.07 mL, 10.4 mmol) were added to a solution of alcohol **12** (5.79 g, 10.4 mmol) in anhydrous DMF (100 mL). The reaction mixture was first stirred at 90 °C for 5h and then at rt for additional 18h. The mixture was concentrated to nearly dryness and taken up in EtOAc and brine. The layers were separated and the aqueous layer was extracted with EtOAc (4×200 mL).

The combined organic layer was evaporated to near dryness, and the resulting residue purified by silica gel column chromatography (0-3.5% *i*-PrOH in $CHCl_3$, v/v) to afford O5'-benzoylated nucleoside **13** (5.05 g, 83%) as a white foam. R_f = 0.4 (10% *i*-PrOH in $CHCl_3$, v/v). Physical data for the mixture of rotamers (~4.5:5.5 by ^{1}H NMR): MALDI-HRMS *m*/*z* 605.1337 ($[M+Na]^+$, $C_{27}H_{21}F_3N_6O_6{\cdot}Na^+$ Calc. 605.1367); ^{1}H NMR[31,32] (500 MHz, DMSO-d_6) δ11.21 (br s, 1H, ex, NH_{A+B}), 8.69 (s, 1H $H2_{A+B}$), 8.59 (s, 0.45H, $H8_A$), 8.54 (s, 0.55H, $H8_B$), 8.05-8.13 (m, 4H, Bz_{A+B}), 7.49-7.72 (m, 6H, Bz_{A+B}), 6.86 (d, 0.55H, J = 1.7 Hz, $H1'_B$), 6.83 (d, 0.45H, J = 1.7 Hz, $H1'_A$), 6.60-6.80 (br s, 1H, ex, 3'-OH_{A+B}), 4.93 (s, 0.55 H, $H3'_B$), 4.89-4.91 (m, 0.9H, $H2'_A$, $H3'_A$), 4.74-4.79 (m, 1.0H, $H5'_{A+B}$), 4.68 (br s, 0.55H, $H2'_B$), 4.58-4.64 (m, 1.0H, $H5'_{A+B}$), 4.54 (d, 0.45H, J = 10.7 Hz, $H5''_A$), 4.35 (d, 0.55H, J = 11.5 Hz, $H5''_B$), 4.10 (d, 0.45H, J = 10.7 Hz, $H5''_A$), 3.93 (d, 0.55H, J = 11.5 Hz, $H5''_B$); ^{13}C NMR (125 MHz, DMSO-d_6) δ166.5, 165.3, 155.1 (2q, J = 36 Hz, $COCF_3$), 151.8 ($C2_{A/B}$), 151.6 ($C2_{A/B}$), 151.4, 140.6 ($C8_{A/B}$), 140.4 ($C8_{A/B}$), 134.5, 134.4, 133.6 (Bz), 131.9 (Bz), 129.55 (Bz), 129.54 (Bz), 129.18, 129.17, 128.7 (Bz), 128.4 (Bz), 128.2 (Bz), 125.5, 125.3, 115.5 (2q, CF_3, J = 286 MHz, CF_3), 87.4, 86.0, 84.5 ($C1'_A$), 83.7 ($C1'_B$), 72.9 ($C3'_B$), 71.0 ($C3'_A$), 65.3 ($C2'_B$), 63.7 ($C2'_A$), 60.7 (C5'), 60.2 (C5'), 53.0 ($C5''_B$), 52.9 ($C5''_A$) – minor impurities were observed at 121.3, 119.9 and 79.1 ppm; ^{19}F NMR (376 MHz, DMSO-d_6) δ-71.1 (CF_3-$_B$), -72.0 (CF_3-$_A$).

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-7-hydroxy-1-hydroxymethyl-2-oxa-5-azabicyclo [2.2.1]heptane (**14**). From **13**: Aqueous NaOH (2M, 17.5 mL, 0.35 mol) was added to an ice-cold solution of alcohol **13** (3.41 g, 5.85 mmol) in 1,4-dioxane and water (225 mL, 2:1, v/v). The reaction mixture was stirred at 0 °C for 2h, at which point sat. aq. NH_4Cl (25 mL) was added. The solution was evaporated to dryness, and the resulting crude was adsorbed on

silica gel and purified by silica gel column chromatography (0-20% MeOH in CH_2Cl_2, v/v) to afford amino alcohol **14** (1.33 g, 60%) as a white solid material. R_f = 0.4 (20% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m/z* 405.1294 ($[M+Na]^+$, $C_{18}H_{18}N_6O_4 \cdot Na^+$, Calc. 405.1282); ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.17 (br s, 1H, ex, NH), 8.73 (s, 1H, A^{Bz}), 8.71 (s, 1H, A^{Bz}), 8.05 (d, 2H, *J* = 8.1 Hz, Bz), 7.49-7.67 (m, 3H, Bz), 6.44 (d, 1H, *J* = 1.9 Hz, H1'), 5.70 (d, 1H, ex, *J* = 4.0 Hz, 3'-OH), 4.82 (t, 1H, ex, *J* = 5.5 Hz, 5'-OH), 4.30 (d, 1H, *J* = 4.0 Hz, H3'), 3.69 (d, 2H, *J* = 5.5 Hz, H5'), 3.51 (s, 1H, H2'), 3.15-3.20 (d, 1H, *J* = 10.6 Hz, H5''), 2.94-3.01 (d, 1H, *J* = 10.6 Hz, H5''); ^{13}C NMR (75.5 MHz, DMSO-d_6) δ 165.5, 152.0, 151.2, 149.8, 143.2, 133.3, 132.3, 128.4, 125.2, 91.3, 83.9, 73.4, 62.2 (C5''), 58.3 (C2'), 50.6 (C5').

From **16**: Nucleoside **16** (1.00 g, 1.01 mmol) was dissolved in a solution of $CHCl_2COOH$/MeOH/CH_3NO_2 (50 mL, 3:5:92, v/v/v) and stirred at 0 °C for 15 min. Sat. aq. $NaHCO_3$ was carefully added to neutralize the solution and the mixture was evaporated to dryness and adsorbed on silica gel. The resulting residue was purified by silica gel column chromatography (0-20% MeOH in CH_2Cl_2, v/v - initially built including 1% Et_3N) to afford amino diol **14** as a white solid material (0.37 g, 96%).

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-1-(4,4'-dimethoxytrityloxymethyl)-7-hydroxy-2-oxa-5-azabicyclo[2.2.1]heptane (**15**). Amino alcohol **14** (1.25 g, 3.27 mmol) was co-evaporated with anhydrous pyridine (30 mL) and redissolved in anhydrous pyridine (65 mL). The solution was cooled using a ice-salt mixture and DMTrCl (1.55 g, 4.58 mmol) was added in one portion (addition over several portions did not influence reaction outcome). The reaction mixture was warmed to rt and stirred for 23 h, at which point MeOH (10 mL) was added. The mixture was diluted with EtOAc (100 mL) and washed with sat. aq. $NaHCO_3$ (2×15 mL). The

combined aqueous layer was back-extracted with EtOAc (3×30 mL) and the combined organic phase was evaporated to dryness and subsequently co-evaporated with abs. EtOH:toluene (3×50 mL, 2:1, v/v). The resulting residue was purified by silica gel column chromatography (0-8% MeOH in CH_2Cl_2, v/v - initially built with 0.5% Et_3N, v/v) and subsequently co-evaporated with abs. EtOH:toluene (2×50 mL, 1:1, v/v) to afford nucleoside **15** (0.85 g, 38%) as a white solid material, along with nucleoside 16 (0.97 g, 30%) as a yellow foam. Physical data for 15: R_f = 0.4 (10% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m/z* 707.2609 ($[M+Na]^+$, $C_{39}H_{36}N_6O_6 \cdot Na^+$, Calc. 707.2589); ^{1}H NMR[31] (500 MHz, DMSO-d_6) δ 11.16 (br s, 1H, ex, NH), 8.78 (s, 1H, H8), 8.73 (s, 1H, H2), 8.06 (d, 2H, *J* = 7 Hz, Bz), 7.63-7.67 (t, 1H, *J* = 7.7 Hz, Bz), 7.54-7.58 (d, 2H, *J* = 7.0 Hz, Bz), 7.19-7.44 (m, 9H, DMTr), 6.86-6.92 (m, 4H, DMTr), 6.54 (d, 1H, *J* = 1.9 Hz, H1'), 5.66 (d, 1H, ex, *J* = 4.8 Hz, 3'-OH), 4.40 (d, 1H, *J* = 4.8 Hz, H3'), 3.74 (s, 6H, CH_3O), 3.50 (br s, 1H, H2'), 3.30-3.32 (m, 1H, H5' – overlap with H_2O), 3.25-3.27 (d, 1H, *J* = 10.7 Hz, H5'), 3.19 (br s, 2H, H5''), 2.90 (br s, 1H, ex, 2'-NH); ^{13}C NMR (125 MHz, DMSO-d_6) δ 165.5, 158.0, 152.2, 151.3 (C2), 149.9, 144.8, 143.3 (C8), 135.5, 135.4, 133.4, 132.3 (Ar), 129.71 (Ar), 129.68 (Ar), 128.45, 128.39 (Ar), 127.8 (Ar), 127.7 (Ar), 126.6 (Ar), 125.2, 113.2 (DMTr), 89.6, 85.1, 84.1 (C1'), 74.0 (C3'), 62.2 (C2'), 61.2 (C5'), 55.0 (CH_3O), 51.2 (C5'').

Physical data for (1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-5-(4,4'-dimethoxytrityl)-1-(4,4'-dimethoxytrityloxymethyl)-7-hydroxy-2-oxa-5-azabicyclo[2.2.1]heptane (**16**): R_f = 0.5 (3% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m/z* 1009.3900 ($[M+Na]^+$, $C_{60}H_{54}N_6O_8 \cdot Na^+$, Calc. 1009.3895); ^{1}H NMR[31] (500 MHz, DMSO-d_6) δ 11.33 (br s, 1H, ex, NH), 9.13 (s, 1H, H8), 8.77 (s, 1H, H2), 8.10 (d, 2H, *J* = 7.5 Hz, Bz), 7.64-7.67 (t, 1H, *J* = 7.5 Hz, Bz), 7.52-7.58

(t, 2H, J = 7.5 Hz, Bz), 6.78-7.41 (m, 22H, DMTr), 6.58 (d, 1H, J = 1.1 Hz, H1'), 6.53 (d, 2H, J = 9.1 Hz, DMTr), 6.43 (d, 2H, J = 9.1 Hz, DMTr), 4.26 (d, 1H, J = 5.0 Hz, H3'), 3.82 (d, 1H, J = 9.9 Hz, H5''), 3.75 (br s, 1H, H2'), 3.73 (s, 6H, CH_3O), 3.65 (s, 3H, CH_3O), 3.62 (s, 3H, CH_3O), 3.55 (d, 1H, ex, J = 5.0 Hz, 3'-OH), 3.17-3.19 (d, 1H, J = 11.0 Hz, H5'), 3.09-3.11 (d, 1H, J = 11.0 Hz, H5'), 2.98 (d, 1H, J = 9.9 Hz, H5''); ^{13}C NMR (125 MHz, DMSO-d_6) δ165.7, 158.0, 157.05, 156.96, 152.0, 151.6 (C2), 150.3, 145.2, 144.8, 142.9 (C8), 136.7, 136.2, 136.0, 135.4, 135.3, 133.3, 132.4 (Bz), 130.5 (DMTr), 130.2 (DMTr), 129.8 (DMTr), 129.7 (DMTr), 128.8 (DMTr), 128.44 (Bz), 128.39 (Bz), 127.73 (DMTr), 127.70 (DMTr), 127.1 (DMTr), 126.5 (DMTr), 126.3, 125.5 (DMTr), 123.8 (DMTr), 113.1 (DMTr), 112.5 (DMTr), 112.4 (DMTr), 88.4, 86.6 (C1'), 85.2, 74.3, 74.0 (C3'), 64.3 (C2'), 60.7 (C5'), 55.0 (C5''), 54.9 (CH_3O), 54.8 (CH_3O), 54.75 (CH_3O), 54.73 (CH_3O). A trace impurity of pyridine was identified in the ^{13}C NMR spectrum at 149.5 ppm.

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-1-(4,4'-dimethoxytrityloxymethyl)-5-(9'-fluorenylmethoxycarbonyl)-7-hydroxy-2-oxa-5-azabicyclo[2.2.1]heptane (**17**). Amino alcohol **15** (200 mg, 0.29 mmol) was co-evaporated in anhydrous pyridine (2×2 mL) and re-dissolved in anhydrous pyridine (1.5 mL). The solution was cooled to 0 °C and 9'-fluorenylmethyl chloroformate (100 mg, 0.38 mmol) added hereto. The reaction mixture was warmed to rt and stirred for 6h, at which point it was diluted with EtOAc (30 mL) and washed with sat. aq. $NaHCO_3$ (30 mL). The aqueous layer was back-extracted with EtOAc (25 mL) and the combined organic layer evaporated to near dryness. The resulting crude was co-evaporated with abs. EtOH:toluene (2×6 mL, 2:1, v/v) and purified by silica gel column chromatography (30-100% EtOAc in petroleum ether, v/v) to afford target nucleoside **17** (136 mg, 51%) as a

white foam. R_f = 0.4 (EtOAc). Physical data for mixture of rotamers (~1:1.2 by ^{1}H NMR): MALDI-HRMS *m/z* 929.3241 ([M + Na]$^+$, $C_{54}H_{46}N_6O_8 \cdot Na^+$ Calc. 929.3269); ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.29 (br s, ex), 8.74 (s, 1H), 8.72 (s, 1.2H), 8.55 (s, 1.2H), 8.49 (s, 1H), 6.90-8.08 (m, 57.2H), 6.80 (d, 1.2H, *J* = 1.7 Hz), 6.73 (d, 1H, *J* = 1.7 Hz), 6.25 (d, 1.2H, ex, *J* = 4.4 Hz), 6.22 (d, 1H, ex, *J* = 4.4 Hz), 4.50 (br s, 1.2H), 4.45 (br s, 1H), 4.22 (d, 1H, *J* = 6.9 Hz), 4.15 (d, 1.2H, *J* = 6.9 Hz), 3.87 (d, 1H, *J* = 6.9 Hz), 3.83 (d, 1.2H, *J* = 6.9 Hz), 3.62-3.76 (m, 17.6H), 3.34-3.43 (m, 6.6H); ^{13}C NMR (75.5 MHz, DMSO-d_6) δ 165.66, 165.63, 158.1, 154.7, 154.6, 151.8, 151.5, 150.3, 150.2, 144.7, 143.7, 143.6, 143.5, 142.8, 141.3, 141.2, 140.6, 140.4, 140.3, 135.3, 135.2, 133.3, 132.4, 129.7, 128.8, 128.4, 127.9, 127.7, 127.6, 127.5, 127.2, 127.1, 126.9, 126.7, 125.4, 125.2, 125.1, 124.9, 124.6, 121.3, 119.98, 113.2, 88.9, 88.4, 85.4, 84.9, 84.7, 84.6, 72.5, 72.1, 66.8, 66.7, 63.9, 63.4, 60.6, 60.5, 55.0, 52.7, 52.6, 46.4, 45.9.

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-1-(4,4'-dimethoxytrityloxymethyl)-7-hydroxy-5-(pyren-1-yl)methyl-2-oxa-5-azabicyclo[2.2.1]heptane (**18**). Nucleoside **15** (225 mg, 0.33 mmol) was co-evaporated with anhydrous 1,2-dichloroethane (2×5 mL) and re-dissolved in anhydrous 1,2-dichloroethane (4 mL). 1-Pyrenecarboxaldehyde (115 mg, 0.49 mmol) and $NaBH(OAc)_3$ (104 mg, 0.49 mmol) were added, and the resulting suspension was stirred at rt for 17 h, at which point sat. aq. $NaHCO_3$ (30 mL) was added. The mixture was extracted with CH_2Cl_2 (3×15 mL) and the combined organic layers dried (Na_2SO_4) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-5% MeOH in CH_2Cl_2, v/v) to afford N2'-alkylated nucleoside **18** as a white foam (201 mg, 68%). R_f = 0.4 (EtOAc); MALDI-HRMS *m/z* 921.3326 ([M+Na]$^+$, $C_{56}H_{46}N_6O_6 \cdot Na^+$, Calc. 921.3371); ^{1}H NMR[31] (500 MHz, DMSO-d_6) δ 11.18 (s, 1H, ex, NH), 8.53 (s, 1H, H2), 8.46 (s, 1H, H8), 7.97-

8.24 (m, 8H, Ar), 7.60-7.85 (m, 6H, Ar), 7.41-7.44 (m, 2H, DMTr), 7.19-7.32 (m, 7H, DMTr), 6.85-6.92 (m, 4H, DMTr), 6.50 (d, 1H, J = 1.9 Hz, H1'), 6.14 (d, 1H, ex, J = 3.6 Hz, 3'-OH), 4.72-4.75 (d, 1H, J = 12.5 Hz, CH_2Py), 4.60 (d, 1H, J = 3.6 Hz, H3'), 4.52-4.56 (d, 1H, J = 12.5 Hz, CH_2Py), 3.73 (s, 6H, CH_3O), 3.66 (s, 1H, H2'), 3.34-3.45 (m, 3H, 2x H5', H5''), 3.17-3.20 (d, 1H, J = 10.3 Hz, H5''); ^{13}C NMR (125 MHz, DMSO-d_6) δ165.6, 158.0, 151.7, 151.0 (C2), 149.9, 144.8, 142.8 (C8), 135.5, 135.4, 133.6, 133.2, 132.4 (Ar), 130.6, 130.2, 130.0, 129.71 (DMTr), 129.68 (DMTr), 128.7, 128.5 (Ar), 128.4 (Ar), 127.8 (DMTr), 127.7 (DMTr), 127.6 (Ar), 127.2 (Ar), 126.9 (Ar), 126.8 (Ar), 126.6 (DMTr), 125.9 (Ar), 125.5, 124.9 (Ar), 124.8 (Ar), 124.2 (Ar), 123.9, 123.7, 123.1 (Ar), 113.2 (DMTr), 90.3, 85.2, 84.8 (C1'), 75.4 (C3'), 66.3 (C2'), 61.3 (C5'), 59.2 (C5''), 58.2 (CH_2Py), 55.0 (CH_3O).

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-1-(4,4'-dimethoxytrityloxymethyl)-7-hydroxy-5-(pyren-1-yl)carbonyl-2-oxa-5-azabicyclo[2.2.1]heptane (**19**). Amino alcohol **15** (0.41 g, 0.59 mmol) was co-evaporated with anhydrous 1,2-dichloroethane (2 × 10 mL), dissolved in anhydrous CH_2Cl_2 (11.9 mL), and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochlorid (EDC·HCl, 226 mg, 1.19 mmol) and 1-pyrenecarboxylic acid (0.29 g, 1.19 mmol) added hereto. The reaction mixture was stirred for 45h at rt, at which point it was diluted with CH_2Cl_2 (50 mL) and washed with water (20 mL). The two phases were separated and the aqueous phase was back-extracted with CH_2Cl_2 (3 × 50 mL). The combined organic phase was evaporated to dryness and the resulting residue was purified by silica gel column chromatography (0-99% EtOAc and 1% pyridine in petroleum ether, v/v). The resulting product was co-evaporated with abs. EtOH:toluene (2 × 50 ml, 1:1, v/v) to afford nucleoside **19** (343 mg, 64%) as a yellow solid material. Physical data for the mixture of rotamers (~0.4:1 by ^1H

NMR): R_f: 0.5 (5% MeOH: CH_2Cl_2, v/v); MALDI-HRMS *m/z* 935.3138 ($[M+Na]^+$, $C_{56}H_{44}N_6O_7 \cdot Na^+$ Calc. 935.3164). ^{1}H NMR[31,32] (300 MHz, DMSO-d_6) δ 11.21 (br s, 1.4H, ex, NH), 8.55 (s, 1H, A^{Bz}-$_B$), 8.53 (s, 0.4H, A^{Bz}-$_A$), 8.45 (s, 1H, A^{Bz}-$_B$), 8.40 (s, 0.4H, A^{Bz}-$_A$), 6.82-8.23 (m, ~37.8H, Ar_{A+B}), 6.50 (d, 1H, J = 2.2 Hz, H1'$_B$), 6.40 (d, 0.4H, J = 1.8 Hz, H1'$_A$), 6.15 (d, 1H, ex, J = 4.0 Hz, 3'-OH_B), 6.07 (d, 0.4H, ex, J = 3.7 Hz, 3'-OH_A), 4.42-4.77 (m, 4.2H, H3'$_{A+B}$, H5'$_{A+B}$), 3.55-3.75 (m, ~9.8H, CH_3O_{A+B}, H2'$_{A+B}$), 3.15-3.44 (m, ~2.8H, H5"$_{A+B}$); ^{13}C NMR (75.5 MHz, DMSO-d_6) δ 165.5, 157.9, 157.7, 151.7, 150.98, 149.8, 148.2, 144.7, 142.8, 140.1, 135.4, 135.3, 133.5, 133.1, 132.3, 130.5, 130.2, 129.9, 129.6, 128.8, 128.6, 128.4, 127.8, 127.6, 127.5, 127.3, 127.2, 126.9, 126.5, 126.3, 125.9, 125.5, 124.9, 124.8, 124.78, 124.2, 123.8, 123.7, 123.1, 113.1, 112.7, 92.0, 90.3, 85.1, 84.7, 79.8, 75.3, 74.7, 66.4, 61.2, 59.1, 58.6, 58.3, 58.2, 54.6.

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-1-(4,4'-dimethoxytrityloxymethyl)-7-hydroxy-5-(pyren-1-yl)acetyl-2-oxa-5-azabicyclo[2.2.1]heptane (**20**). Amino alcohol **15** (0.25 g, 0.37 mmol) was co-evaporated with anhydrous 1,2-dichloroethane (2×5 mL) and dissolved in anhydrous CH_2Cl_2 (10 mL). To this was added 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl, 108 mg, 0.55 mmol) and 1-pyreneacetic acid (145 mg, 0.55 mmol). The reaction mixture was stirred for 2.5h at rt, at which point it was diluted with CH_2Cl_2 (20 mL) and washed with water (2×10 mL). The aqueous layer was back-extracted with CH_2Cl_2 (10 mL) and the combined organic layer evaporated to dryness. The resulting crude residue was purified by silica gel column chromatography (0-4% MeOH in CH_2Cl_2, v/v) to afford nucleoside **20** as a white solid (0.27 g, 79%). Physical data for the mixture of rotamers (~0.5:1.0 by ^{1}H NMR): R_f = 0.5 (10% MeOH:CH_2Cl_2, v/v); MALDI-HRMS *m/z*

949.3321 ([M+Na]$^+$, $C_{57}H_{46}N_6O_7 \cdot Na^+$, Calc. 949.3320); ^{1}H NMR[31,32] (500 MHz, DMSO-d_6) δ 11.29 (br s, 1.5H, ex, NH_{A+B}), 8.97 (s, 1H, $H8_B$), 8.73 (s, 0.5H, $H2_A$), 8.71 (s, 1H, $H2_B$), 8.55 (s, 0.5H, $H8_A$), 6.91-8.32 (m, 40.5H, Ar_{A+B}), 6.83 (d, 0.5H, J = 1.7 Hz, $H1'_A$), 6.80 (d, 1H, J = 1.4 Hz, $H1'_B$), 6.32 (d, 0.5H, ex, J = 4.4 Hz, 3'-OH_A), 6.19 (d, 1H, ex, J = 4.4 Hz, 3'-OH_B), 5.07 (s, 0.5H, $H2'_A$), 4.71-4.73 (m, 1.5H, $H2'_B$, $H3'_A$), 4.67 (d, 1H, J = 10.4 Hz, $H5''_B$), 4.63 (d, 1H, J = 4.4 Hz, $H3'_B$), 4.55-4.60 (d, 1H, J = 17.0 Hz, $CH_2Py\text{-}_B$), 4.36-4.40 (d, 1H, J = 17.0 Hz, $CH_2Py\text{-}_B$), 4.11 (d, 0.5H, J = 16.2 Hz, $CH_2Py\text{-}_A$), 3.96-4.03 (m, 1.5H, $H5''_A$, $H5''_B$), 3.66-3.77 (m, 9.5H, CH_3O, $H5''_A$), 3.34-3.48 (m, 3.5H, $H5'_{A+B}$, $CH_2Py\text{-}_A$); ^{13}C NMR (125 MHz, DMSO-d_6) δ169.9, 169.6, 165.6, 165.5, 158.1, 158.0, 152.0, 151.7 ($A^{Bz}\text{-}_A$), 151.5, 151.4 ($A^{Bz}\text{-}_B$), 150.6, 150.2, 144.71, 144.68, 141.7 ($A^{Bz}\text{-}_B$), 141.5 ($A^{Bz}\text{-}_A$), 135.4, 135.2, 133.54, 133.47, 132.5 (Ar), 130.70, 130.69, 130.28, 130.1, 129.91, 129.88, 129.82 (Ar), 129.79 (Ar), 129.77, 129.74 (Ar), 129.68, 129.62, 129.4, 129.2, 128.8, 128.7, 128.58 (Ar), 128.56 (Ar), 128.54 (Ar), 128.46 (Ar), 128.2 (Ar), 127.85 (Ar), 127.77 (Ar), 127.72 (Ar), 127.4 (Ar), 127.27 (Ar), 127.2 (Ar), 127.1 (Ar), 126.9, 126.8 (Ar), 126.68 (Ar), 126.2, 126.0 (Ar), 125.9 (Ar), 125.4, 125.3, 125.00 (Ar), 124.98 (Ar), 124.82 (Ar), 124.76 (Ar), 124.7 (Ar), 124.6 (Ar), 124.5 (Ar), 124.4 (Ar), 124.3 (Ar), 124.1, 123.89, 123.79, 123.76, 123.72, 123.2 (Ar), 113.2 (Ar), 89.0, 88.4, 85.5, 85.4, 84.73 ($C1'_A$), 84.66 ($C1'_B$), 72.9 ($C3'_A$), 72.0 ($C3'_B$), 64.5 ($C2'_A$), 61.8 ($C2'_B$), 60.9 ($C5'_B$), 60.8 ($C5'_A$), 55.0 (CH_3O), 52.8 ($C5''_B$), 52.3 ($C5''_A$), 37.8 ($CH_2Py\text{-}_B$), 37.7 ($CH_2Py\text{-}_A$). Trace impurities of dichloromethane and 1-pyreneacetic acid were identified in the ^{13}C NMR spectrum. The compound was used in the next step without further purification.

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-7-[2-cyanoethoxy(diisopropylamino)phosphinoxy]-1-(4,4'-dimethoxytrityloxymethyl)-5-(9'-fluorenylmethoxycarbonyl)-2-oxa-5-azabicyclo[2.2.1]heptane (**1**). Nucleoside **17** (225 mg, 0.25 mmol) was co-evaporated with anhydrous 1,2-dichloroethane (2×4 mL) and re-dissolved in anhydrous CH_2Cl_2 (3.5 mL). Anhydrous *N*,*N*'-diisopropylethylamine (DIPEA, 195 μL, 1.12 mmol) and *N*-methylimidazole (NMI, 16 μL, 0.20 mmol) were added followed by dropwise addition of 2-cyanoethyl-*N*,*N*'-diisopropylchlorophosphoramidite (111 μL, 0.50 mmol). The reaction mixture was stirred for 4h at rt, evaporated to dryness, and the resulting residue purified by silica gel column chromatography (0-2% MeOH in CH_2Cl_2, v/v, initially built in 0.5% Et_3N) and subsequent precipitation from CH_2Cl_2/petroleum ether to afford target amidite **1** (126 mg, 46%) as a white foam. Physical data for mixture of rotamers: R_f = 0.5 (5% MeOH in CH_2Cl_2, v/v); ESI-HRMS *m*/*z* 1129.4356 ($[M+Na]^+$, $C_{63}H_{63}N_8O_9P{\cdot}Na^+$, Calc. 1129.4348); ^{31}P NMR (121 MHz, $CDCl_3$) δ150.32, 150.26, 150.1, 149.5.

(1*S*,3*R*,4*S*,7R)-3-(6-*N*-Benzoyladenin-9-yl)-7-[2-cyanoethoxy(diisopropylamino)phosphinoxy]-1-(4,4'-dimethoxytrityloxymethyl)-5-(pyren-1-yl)methyl-2-oxa-5-azabicyclo[2.2.1]heptane (**2**). Nucleoside **18** (155 mg, 0.17 mmol) was co-evaporated with 1,2-dichloroethane (2×5 mL) and dissolved in a 20% *N*,*N*'-diisopropylethylamine solution in anhydrous CH_2Cl_2 (5.0 mL, v/v). To this was added 2-cyanoethyl-*N*,*N*'-diisopropylchlorophosphoramidite (0.10 mL, 0.44 mmol) and the reaction mixture was stirred at rt for 17h whereupon the reaction mixture was quenched with abs EtOH (1 mL) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-90% EtOAc in petroleum ether, v/v) and precipitated from

CH_2Cl_2/petroleum ether to afford amidite **2** as a white solid material (131 mg, 69%). R_f = 0.7 (5% MeOH in CH_2Cl_2, v/v); ESI-HRMS *m*/*z* 1099.4642 ($[M+H]^+$, $C_{65}H_{63}N_8O_7P{\cdot}H^+$, Calc.1099.4630); ^{31}P NMR (121 MHz, $CDCl_3$) δ 150.6, 148.4.

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-7-[2-cyanoethoxy(diisopropylamino)phosphinoxy]-1-(4,4'-dimethoxytrityloxymethyl)-5-(pyren-1-yl)carbonyl-2-oxa-5-azabicyclo[2.2.1]heptane (**3**). Nucleoside **19** (0.30 g, 0.33 mmol) was co-evaporated with 1,2-dichloroethane (2×5 mL) and dissolved in 20% *N*,*N*'-diisopropylethylamine in anhydrous CH_2Cl_2 (3.4 mL, v/v). To this was added 2-cyanoethyl-*N*,*N*'-diisopropylchlorophosphoramidite (0.11 mL, 0.49 mmol) and the reaction mixture was stirred for 22 h at rt, at which point abs. EtOH (2 mL) was added. The mixture was taken up in CH_2Cl_2 (50 mL) and washed with sat. aq. $NaHCO_3$ (50 mL) and brine (50 mL). The combined aqueous layer was back-extracted with CH_2Cl_2 (2×20 ml) and the combined organic layer evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-99% EtOAc in petroleum ether containing 1% pyridine, v/v/v), co-evaporated with abs. EtOH:toluene (2×50 mL, 1:1, v/v) and precipitated from EtOAc/hexane to afford phosphoramidite **3** as a white foam (247 mg, 67%). Physical data for the mixture of rotamers: R_f = 0.5 (EtOAc); ESI-HRMS *m*/*z* 1113.4462 ($[M+H]^+$, $C_{65}H_{61}N_8O_8P{\cdot}H^+$, Calc.1113.4422); ^{31}P NMR (121 MHz, $CDCl_3$) δ 152.3, 151.8, 150.2.

(1*S*,3*R*,4*S*,7*R*)-3-(6-*N*-Benzoyladenin-9-yl)-7-[2-cyanoethoxy(diisopropylamino)phosphinoxy]-1-(4,4'-dimethoxytrityloxymethyl)-5-(pyren-1-yl)acetyl-2-oxa-5-azabicyclo[2.2.1]heptane (**4**). Nucleoside **20** (235 mg, 0.25 mmol) was co-evaporated with anhydrous 1,2-dichloroethane (2×5 mL) and re-dissolved in anhydrous CH_2Cl_2 (5 mL). Anhydrous *N*,*N*'-diisopropylethylamine (220 μL, 1.27 mmol) was added, followed by dropwise addition of 2-cyanoethyl-*N*,*N*'-diisopropylchlorophosphoramidite (115 μL, 0.51 mmol). The reaction mixture was stirred at rt for 22 h, at which point abs. EtOH (1 mL) was added. The solution was evaporated to dryness and the resulting crude residue purified by silica gel column chromatography (0-2% MeOH in CH_2Cl_2, v/v) and precipitation from CH_2Cl_2/petroleum ether to afford target amidite **4** (203 mg, 71%) as a white foam. Physical data for the mixture of rotamers: R_f = 0.6 (5% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m*/*z* 1149.4358 ($[M+Na]^+$, $C_{66}H_{63}N_8O_8P{\cdot}Na^+$, Calc. 1149.4399); ^{31}P NMR (121 MHz, $CDCl_3$) δ 150.4, 150.31, 150.26, 148.1.

2.4.2. Protocol - Synthesis and Purification of ONs

W/X/Y/Z-modified oligodeoxyribonucleotides were made on an automated DNA synthesizer using 0.2 μmol scale succinyl linked LCAA-CPG (long chain alkyl amine controlled pore glass) columns with a pore size of 500Å. Phosphoramidites **1-4** were incorporated into ONs using the following hand-coupling conditions (activator; coupling time; stepwise coupling yield): monomer **W** (pyridinium hydrochloride; 30 min; ~82%) and monomers **X**-**Z** (pyridinium hydrochloride; 15 min; ~95%). Standard protocols for incorporation of DNA phosphoramidites were used. Modified ONs were deprotected using 32% aq. NH_3 (55 °C, 2-12 h) and purified (DMTr-ON) by ion-pair reversed-phase HPLC using either an ammonium formate –

acetonitrile gradient or a triethylammonium acetate - acetonitrile gradient, which was followed by detritylation (80% aq. AcOH, 20 min) and precipitation (abs. EtOH or acetone, −18 °C, 12 h). The composition of the modified ONs was verified by MALDI-MS analysis (Tables S1 and S2 in the Supporting Information[33]) recorded either in positive or negative ion mode on a Quadrupole Time-Of-Flight Tandem Mass Spectrometer equipped with a MALDI source using 3-hydroxypicolinic acid as a matrix. Purity (generally >90%) was verified either by the ion-pair reverse phase HPLC system running in analytical mode or ion-exchange HPLC using a Tris-Cl/EDTA - NaCl gradient.

2.4.3. Protocol - Thermal Denaturation Studies

Concentrations of ONs were calculated using the following extinction coefficients (OD_{260}/μmol): G, 12.0; A, 15.2; T, 8.4; U, 10.0; C, 7.1; pyrene, 22.4. ONs (each strand at 1.0 μM) were thoroughly mixed, denatured by heating and subsequent cooling to the starting temperature of the experiment. Quartz optical cells with a path length of 10.0 mm were used. Thermal denaturation temperatures (T_m's/°C) were measured on a temperature-controlled UV-Vis spectrophotometer and determined as the maximum of the first derivative of the thermal denaturation curve (A_{260} vs. T) recorded in medium salt buffer (T_m buffer: 100 mM NaCl, 0.1 mM EDTA and pH 7.0 adjusted with 10 mM NaH_2PO_4/5 mM Na_2HPO_4). The temperature of the denaturation experiments ranged from at least 15 °C below T_m to 20 °C above T_m (although not below 3 °C). A temperature ramp of 0.5 °C/min or 1.0 °C/min was used in the experiments. Reported thermal denaturation temperatures are an average of two measurements within ±1.0 °C.

2.4.4. Protocol – UV-Vis Absorption Spectroscopy

UV-Vis spectra were recorded on a spectrophotometer at 5 °C (except for single-stranded **X**/**Y**-modified probes, which were recorded at room temperature) using each strand at 1 μM concentration in T_m buffer and quartz cells with 1 cm path lengths.

2.4.5. Protocol – Steady-state Fluorescence Emission Spectroscopy

Steady-state fluorescence spectra were recorded at 5 °C using an excitation wavelength of λ_{ex} = 350 nm, each strand at 1 μM concentration in T_m buffer (except with **Y**-modified ONs where strands were used at 0.15 μM concentration), and quartz cells with 1 cm path lengths.

2.4.6. Protocol – Electrophoretic Mobility Shift Assay

These assays, which were conducted in lieu of footprinting experiments to avoid the use of ^{32}P-labeled targets, were performed in a similar manner as previously described.[8b] Thus, ∼100 pmol samples of unmodified **DH1** were 3′-DIG-labeled using the second-generation DIG gel shift kit (Roche Applied Bioscience) as recommended. Equal volumes of 100 nM solutions of DIG-labeled dsDNA targets and probe solutions (concentrations: 0.5, 1, 5, 10, and 50 μM) in 1 × HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, pH 7.2, 10% sucrose, 1 mg/mL spermine tetrahydrochloride) were mixed and incubated for 3 h at room temperature before being loaded on a 15% nondenaturing polyacrylamide gel. After 2 − 3 h of electrophoresis at 100 V in a cold room (∼4 °C) using TBM (89 mM Tris, 89 mM boric acid, 10 mM $MgCl_2$) as a running buffer, the nucleic acid complexes were transferred to a positively charged nylon membrane by electroblotting and processed as recommend by the manufacturer of the DIG gel

shift kit. The chemiluminescence was captured on an X-ray film, and the bands were quantified using Quantity One software.

Supporting Information

General experimental section; additional synthetic strategies; NMR spectra for all new compounds; MS data for all modified ONs; representative thermal denaturation curves; additional T_m data; representative absorption and fluorescence emission spectra; evaluation of SNP-discriminatory potential. This material is available free of charge via Reference 33.

Acknowledgments

PJH appreciates financial support from Award Number R01 GM088697 from the National Institute of General Medical Sciences, National Institutes of Health, and from the Institute of Translational Health Sciences (ITHS) (supported by grants UL1 RR025014, KL2 RR025015, and TL1 RR025016 from the NIH National Center for Research Resources). JW appreciates financial support from The Danish National Research Foundation and the Danish Agency for Science Technology and Innovation. We thank Dr. T. Santhosh Kumar (Univ. Southern Denmark) for providing us with starting material and Dr. Lee Deobald (EBI Murdock Mass Spectrometry Center, Univ. Idaho) for mass spectrometric analyses.

2.5 References

(1) For reviews on conformationally restricted nucleotides, see e.g., (a) Leumann, C. J. *Bioorg*. *Med*. *Chem*. **2002**, *10*, 841-854. (b) Obika, S.; Abdur Rahman, S. M.; Fujisaka, A.; Kawada, Y.; Baba T.; Imanishi, T. *Heterocycles* **2010**, *81*, 1347-1392. (c) Prakash, T. P. *Chem*. *Biodiv*.

2011, *8*, 1616-1641. (d) Deleavey, G. F.; Damha, M. J. *Chem*. *Biol*. **2012**, *19*, 937-954. (e) Zhou, C.; Chattopadhyaya, J. *Chem*. *Rev*. **2012**, *112*, 3808-3832.

(2) (a) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res*. **2008**, *36*, 5123-5138. (b) Bennett, C. F.; Swayze, E. E. *Annu*. *Rev*. *Pharmacol*. *Toxicol*. **2010**, *50*, 259-293. (c) Østergaard, M. E.; Hrdlicka, P. J. *Chem*. *Soc*. *Rev*., **2011**, *40*, 5771-5788. (d) Watts, J. K.; Corey, D. R. *J*. *Pathol*. **2012**, *226*, 365-379. (e) Matsui, M.; Corey, D. R. *Drug Disc*. *Today* **2012**, *17*, 443-450. (f) Dong, H.; Lei, J.; Ding, L.; Wen, Y.; Ju, H.; Zhang, X. *Chem*. *Rev*. **2013**, *113*, 6207-6233.

(3) Singh, S. K.; Nielsen, P.; Koshkin, A. A.; Wengel, J. *Chem*. *Commun*. **1998**, 455-456.

(4) Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J.-I.; Morio, K.-I.; Doi, T.; Imanishi, T. *Tetrahedron Lett*. **1998**, *39*, 5401-5404.

(5) Kaur, H.; Babu, B. R.; Maiti, S. *Chem*. *Rev*. **2007**, *107*, 4672-4697.

(6) (a) Rajwanshi, V. K.; Håkansson, A. E.; Dahl, B. M.; Wengel, J. *Chem*. *Commun*. **1999**, 1395-1396. (b) Sørensen, M. D.; Kværnø, L.; Bryld, T.; Håkansson, A. E.; Verbeure, B.; Gaubert, G.; Herdewijn, P.; Wengel, J. *J*. *Am*. *Chem*. *Soc*. **2002**, *124*, 2164-2176.

(7) Recent examples include: (a) Seth, P. P.; Vasquez, G.; Allerson, C. A.; Berdeja, A.; Gaus, H.; Kinberger, G. A.; Prakash, T. P.; Migawa, M. T.; Bhat, B.; Swayze, E. E. *J*. *Org*. *Chem*. **2010**, *75*, 1569-1581. (b) Li, Q.; Yuan, F.; Zhou, C.; Plashkevych, O.; Chattopadhyaya, J. *J*. *Org*. *Chem*. **2010**, *75*, 6122-6140. (c) Liu, Y.; Xu, J.; Karimiahmadabadi, M.; Zhou, C.; Chattopadhyaya, J. *J*. *Org*. *Chem*. **2010**, *75*, 7112-7128. (d) Upadhayaya, R.; Deshpande, S. A. Li, Q.; Kardile, R. A.; Sayyed, A. Y.; Kshirsagar, E. K.; Salunke, R. V.; Dixit, S. S.; Zhou, C.; Foldesi, A.; Chattopadhyaya, J. *J*. *Org*. *Chem*. **2011**, *76*, 4408-4431. (e) Shrestha, A. R.; Hari, Y.; Yahara, A.; Osawa, T.; Obika, S. *J*. *Org*. *Chem*. **2011**, *76*, 9891-9899. (f) Hanessian, S.;

Schroeder, B. R.; Giacometti, R. D.; Merner, B. L.; Østergaard, M. E.; Swayze, E. E.; Seth, P. P. *Angew*. *Chem*., *Int*. *Ed*. **2012**, *51*, 11242-11245. (g) Haziri, A. I.; Leumann, C. J. *J. Org*. *Chem*. **2012**, *77*, 5861-5869. (h) Gerber, A-B.; Leumann, C. J. *Chem*. *Eur*. *J*. **2013**, *19*, 6990-7006. (i) Morihiro, K.; Kodama, T.; Kentefu; Moai, Y.; Veedu, R. N.; Obika, S. *Angew*. *Chem*. *Int*. *Ed*. **2013**, *52*, 5074-5078. (j) Hari, Y.; Osawa, T.; Kotobuki, Y.; Yahara, A.; Shrestha, A. R.; Obika, S. *Bioorg*. *Med*. *Chem*. **2013**, *21*, 4405-4412. (k) Hari, Y.; Morikawa, T.; Osawa, T.; Obika, S. *Org*. *Lett*, **2013**, *15*, 3702-3705. (l) Migawa, M. T.; Prakash, T. P.; Vasquez, G.; Seth, P. P.; Swayze, E. E. *Org*. *Lett*. **2013**, *15*, 4316-4319. (m) Hanessian, S.; Schroeder, B. R.; Merner, B. L.; Chen. B.; Swayze, E. E.; Seth, P. P. *J. Org*. *Chem*. **2013**, *78*, 9051-9063. (n) Hanessian, S.; Wagger, J.; Merner, B. L.; Giacometti, R. D.; Østergaard, M. E.; Swayze, E. E.; Seth, P. P. *J. Org*. *Chem*. **2013**, *78*, 9064-9075.

(8) For recent examples, see: (a) Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org*. *Chem*. **2011**, *76*, 7119-7131. (b) Didion, B. A.; Karmakar, S.; Guenther, D. C.; Sau, S.; Verstegen, J. P.; Hrdlicka, P. J. *ChemBioChem*, **2013**, *14*, 1534-1538. (c) Denn, B.; Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *Chem*. *Commun*. **2013**, *49*, 9851-9853.

(9) (a) Kumar, T. S.; Madsen, A. S.; Wengel, J.; Hrdlicka, P. J. *J. Org*. *Chem*. **2006**, *71*, 4188-4201. (b) Kumar, T. S.; Madsen, A. S.; Østergaard, M. E.; Sau, S. P.; Wengel, J.; Hrdlicka, P. J. *J. Org*. *Chem*. **2009**, *74*, 1070-1081.

(10) Kumar, T. S.; Wengel, J.; Hrdlicka, P. J. *ChemBioChem* **2007**, *8*, 1122-1125.

(11) Kumar, T. S.; Madsen, A. S.; Østergaard, M. E.; Wengel, J.; Hrdlicka, P. J. *J. Org*. *Chem*. **2008**, *73*, 7060-7066.

(12) (a) Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2010**, *8*, 2028-2036. (b) Sau, S. P.; Madsen, A. S.; Podbevsek, P.; Andersen, N. K.; Kumar, T. S.; Andersen, S.; Rathje, R. L.; Anderson, B. A.; Guenther, D. C.; Karmakar, S.; Kumar, P.; Plavec, J.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, *78*, 9560-9570.

(13) A conference proceeding outlining a portion of of this study has been published, see: Andersen, N. K.; Wengel, J.; Hrdlicka, P. J. *Nucleosides Nucleotides Nucleic Acids* **2007**, *26*, 1415-1417.

(14) Wenska, M.; Honcharenko, D.; Pathmasiri, W.; Chattopadhyaya, J. *Heterocycles* **2007**, *73*, 303-324.

(15) Rosenbohm, C.; Christensen, S. M.; Sørensen, M. D.; Pedersen, D. S.; Larsen, L-E.; Wengel, J.; Koch, T. *Org. Biomol. Chem.* **2003**, *1*, 655-663.

(16) Hrdlicka, P. J.; Andersen, N. K.; Jepsen, J. S.; Hansen, F. G.; Haselmann, K. F.; Nielsen, C.; Wengel, J. *Bioorg. Med. Chem.* **2005**, *77*, 2597-2621.

(17) (a) Ellervik, U.; Magnusson, G. *Tetrahedron Lett.* **1997**, *38*, 1627-1628. (b) Rigoli, J. W.; Østergaard, M. E.; Canady, K. M.; Guenther, D. C.; Hrdlicka, P. J. *Tetrahedron Lett.* **2009**, *50*, 1751-1753.

(18) The conversion of **7** into **9** has been outlined in the patent literature, see: Sørensen, M. D.; Wengel, J.; Koch, T.; Christensen, S. M.; Rosenbohm, C.; Pedersen, D. S. WO 2003095467 A1.

(19) Takaku, H.; Morita, K.; Sumiuchi, T. *Chem Lett.* **1983**, 11, 1661-1664.

(20) The signal for H3' appears as a doublet in nucleosides with unmodified 3'-OH groups.

(21) Abdel-Magid, A. F.; Carson, K. G.; Harris, B. D.; Maryanoff, C. A.; Shah, R. D. *J. Org. Chem.* **1996**, *61*, 3849-3862.

(22) Hakansson, A. E.; Wengel, J. *Bioorg*. *Med*. *Chem*. *Lett*. **2001**, *11*, 935-938.

(23) Crothers, D. M. *Biopolymers* **1968**, *6*, 575-584.

(24) (a) Yamana, K.; Iwase, R.; Furutani, S.; Tsuchida, H.; Zako, H.; Yamaoka, T.; Murakami., A. *Nucleic Acids Res*. **1999**, *27*, 2387-2392. (b) Christensen, U. B.; Pedersen, E. B. *Nucleic Acids Res*. **2002**, *30*, 4918–4925. (c) Bryld, T.; Højland, T.; Wengel, J. *Chem*. *Commun*. **2004**, 1064-1065.

(25) Marin, V.; Hansen, H. F.; Koch, T. R.; Armitage, B. A. *J. Biomol. Struc. Dyn.*, **2004**, *21*, 841-850.

(26) (a) Korshun, V. A.; Stetsenko, D. A.; Gait, M. J. *J. Chem*. *Soc*., *Perkin Trans*. *1* **2002**, 1092-1104. (b) Dohno, C.; Saito, I. *ChemBioChem* **2005**, *6*, 1075-1081.

(27) (a) Dougherty, G.; Pilbrow, J. R. *Int*. *J. Biochem*. **1984**, *16*, 1179-1192. (b) Nakamura, M.; Fukunaga, Y.; Sasa, K.; Ohtoshi, Y.; Kanaori, K.; Hayashi, H.; Nakano, H.; Yamana, K. *Nucleic Acids Res*. **2005**, *33*, 5887-5895. (c) Asanuma, H.; Fujii, T.; Kato, T.; Kashida, H. *J. Photochem. Photobiol. C*. **2012**, *13*, 124-135.

(28) (a) Okamoto, A.; Kanatani, K.; Saito, I. *J*. *Am*. *Chem*. *Soc*. **2004**, *126*, 4820-4827. (b) Østergaard, M. E.; Kumar, P.; Baral, B.; Guenther, D. C.; Anderson, B. A.; Ytreberg, F. M.; Deobald, L.; Paszczynski, A. J. Sharma, P. K.; Hrdlicka, P. J. *Chem*. *Eur*. *J*. **2011**, *17*, 3157-3165

(29) Manoharan, M.; Tivel, K. L.; Zhao, M.; Nafisi, K.; Netzel, T. L. *J*. *Phys*. *Chem*. **1995**, *99*, 17461-17472. (b) Seo, Y. J.; Ryu, J. H. Kim, B. H. *Org*. *Lett*., **2005**, **7**, 4931-4933.

(30) Persil, O.; Hud, N. V. *Trends Biotechnol*. **2007**, *25*, 433-436.

(31) Assignments of ^{1}H NMR signals (and of the corresponding ^{13}C signals) of H5', H5" and CH_2Ph (if present) are interchangeable.

(32) The least predominant rotamer is denoted as 'A'.

(33) Andersen, N. K.; Anderson, B. A.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem.*, **2013**, 78, 12690-12702.

(34) The following nomenclature describes the relative arrangement between two monomers positioned on opposing strands in a duplex. The number n describes the distance measured in number of base pairs and has a positive value if a monomer is shifted toward the 5′-side of its own strand relative to a second reference monomer on the other strand or a negative value if a monomer is shifted toward the 3′-side of its own strand relative to a second reference monomer on the other strand.

CHAPTER 3: Mixed-sequence Recognition of Double-stranded DNA using Energetically Activated Duplex Probes Modified with N2′-Pyrene-, Perylene-, or Coronene-functionalized 2′-*N*-methyl-2′-amino-DNA Monomers

Brooke A. Anderson, Jared J. Onley and Patrick J. Hrdlicka

Department of Chemistry, University of Idaho, Moscow, Idaho 83844, United States

*Manuscript in preparation

Abstract

Invader probes have been introduced as alternatives to polyamides, triplex-forming oligonucleotides and peptide nucleic acids for recognition of chromosomal DNA targets. These double-stranded probes are activated for DNA recognition through incorporation of +1 interstrand zippers of pyrene-functionalized nucleotides. This motif forces the intercalating pyrene moieties into the same region, resulting in perturbation and destabilization of the probe duplex. In contrast, the two probe strands display very high affinity toward complementary DNA. The energy difference between probe duplexes and recognition complexes provides the driving force for DNA recognition. In the present study, we set out to explore the properties of Invaders based on larger intercalators, i.e., perylene and coronene, expecting that the larger π-surface area would result in additional destabilization of the probe duplex and additional stabilization of probe-target duplexes, and in effect increase the thermodynamic driving force for DNA recognition. Toward this end, we synthesized 2′-*N*-methyl-2′-amino-2′-deoxyuridine

phosphoramidites that are functionalized at the N2′-position with pyrene, perylene or coronene moieties and incorporated these monomers into oligodeoxyribonucleotides (ONs). The resulting ONs and Invader probes were characterized by thermal denaturation experiments, analysis of thermodynamic parameters, absorption and fluorescence spectroscopy, and DNA recognition experiments. Invaders functionalized with N2′-perylene intercalators resulted in the largest duplex stabilization, the strongest activation for dsDNA recognition, and the greatest DNA recognition efficiency in this series.

3.1 Introduction

Development of strategies for recognition of specific double-stranded DNA (dsDNA) regions continues to be an area that attracts considerable attention due to the prospect of tools for enabling applications in biological sciences and medicine, including regulation of gene expression via transcriptional interference, detection of chromosomal DNA targets, and correction of genetic mutations.[1-7] Classic approaches toward these ends entail the use of triplex forming oligonucleotides (TFOs)[8] or peptide nucleic acids (PNAs),[9] minor-groove binding polyamides,[10,11] or engineered proteins such as zinc finger nucleases or transcription activator-like effector nucleases (TALENs).[3,12] While major progress has been made with these probe technologies, they are not without limitations. For example, triplex-based approaches require the dsDNA targets to contain an extended purine-rich region, polyamides typically only recognize short target regions, and the construction of engineered proteins entails non-trivial molecular cloning techniques. A plethora of alternative approaches that address some of these shortcomings have been developed,[12-23] including pseudocomplementary PNA,[24-27] γ-PNA[28,29]

and the CRISPR/Cas (clustered, regularly interspaced, short palindromic repeat/CRISPR-associated protein) systems.[30] However, an unmet need for oligonucleotide-based probes that enable rapid, efficient and site-specific mixed-sequence recognition of dsDNA target regions at physiological conditions, still exists.

Our laboratory is pursuing a fundamentally different approach toward this goal, which entails the use of energetically activated DNA duplexes.[31-36] These so-called Invader probes are modified with +1 interstrand zippers arrangements of suitable intercalator-functionalized nucleotides (Figure 3-1 – see Supporting information for a definition of the zipper nomenclature). This particular structural motif forces the intercalators into the same region of the synthetic DNA duplex, leading to a violation of the 'nearest neighbor exclusion principle',[37] according to which the highest intercalator density that can be accommodated in a DNA duplex is one intercalator for every two base pairs. As a result, duplexes with +1 interstrand zippers arrangements of intercalator-functionalized nucleotides are significantly perturbed and destabilized.[31-36] On the other hand, the two strands of the energetically activated duplex display exceptionally high affinity toward complementary DNA (cDNA), as duplex formation results in strongly stabilizing interactions between intercalators and neighboring base pairs (Figure 3-1). The energy difference between the Invader probe and the probe-target duplexes provides the driving force for recognition of dsDNA via dual duplex invasion.[31-36] Invader probes have already been used for recognition of 28-mer mixed-sequence dsDNA fragments specific to food pathogens,[34] and for detection of gender-specific chromosomal DNA at non-denaturing conditions.[35]

First-generation Invader probes were based on 2′-*N*-(pyren-1-yl)-2′-amino-α-L-LNA (locked nucleic acid) monomers.[31] However, the challenging synthesis of these building blocks,[38,39]

prompted us to identify more readily available Invader monomers. Two viable candidates emerged from initial screens, i.e., 2′-*O*-(pyren-1-yl)methyl-RNA and 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-DNA monomers (Figure 3-1).[32] Straightforward access to these building blocks[33,40] has enabled us to conduct structure-property relationship studies, in which the influence of the nucleobase[33] or the orientation of the pyrene relative to the sugar skeleton[36] on dsDNA recognition, was delineated.

In the present study, we set out to study the influence of intercalator size on the dsDNA recognition efficiency of Invader probes. Thus far, we have used pyrene-functionalized nucleotides as the key activating components of Invader probes. However, it is known that the surface area of pyrene (~220 $Å^2$) is smaller than the area occupied by natural base pairs (~270 $Å^2$).[41] The use of building blocks with larger intercalators therefore presents itself as a promising strategy for i) additional destabilization of Invader probes (more pronounced violation of the 'nearest neighbor exclusion principle'), ii) increasing the cDNA affinity of individual Invader strands (more efficient intercalator-nucleobase stacking), and, consequentially, iii) increasing the thermodynamic driving force of Invader-mediated dsDNA-recognition. Toward this end, we synthesized 2′-*N*-methyl-2′-amino-2′-deoxyuridine nucleotides that are N2′-functionalized with pyrene, perylene or coronene moieties (Figure 3-1) and incorporated these building blocks into oligodeoxyribonucleotides (ONs). The resulting ONs and Invader probes are characterized by means of thermal denaturation experiments, analysis of thermodynamic parameters, absorption and fluorescence spectroscopy, and model dsDNA recognition experiments.

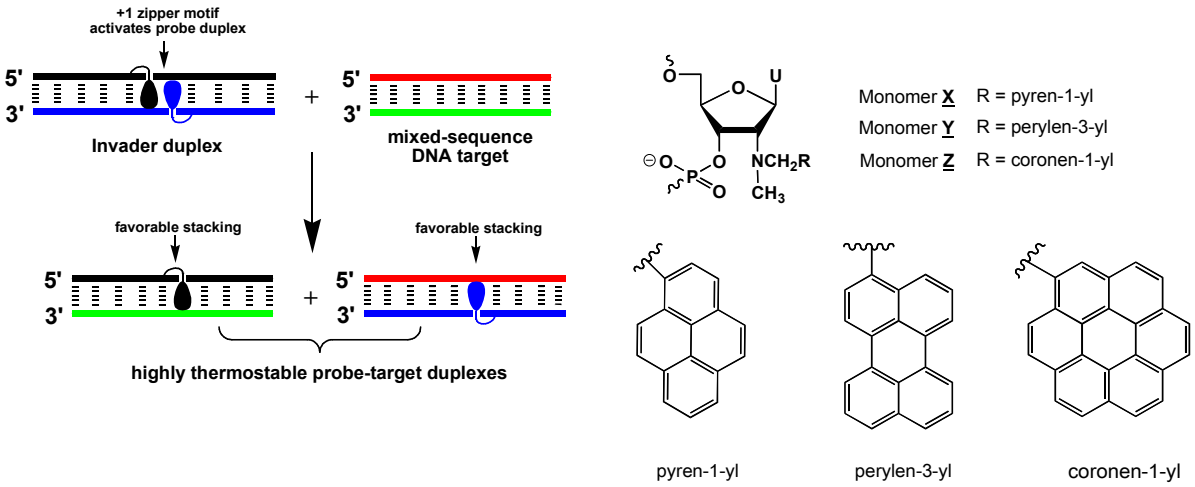


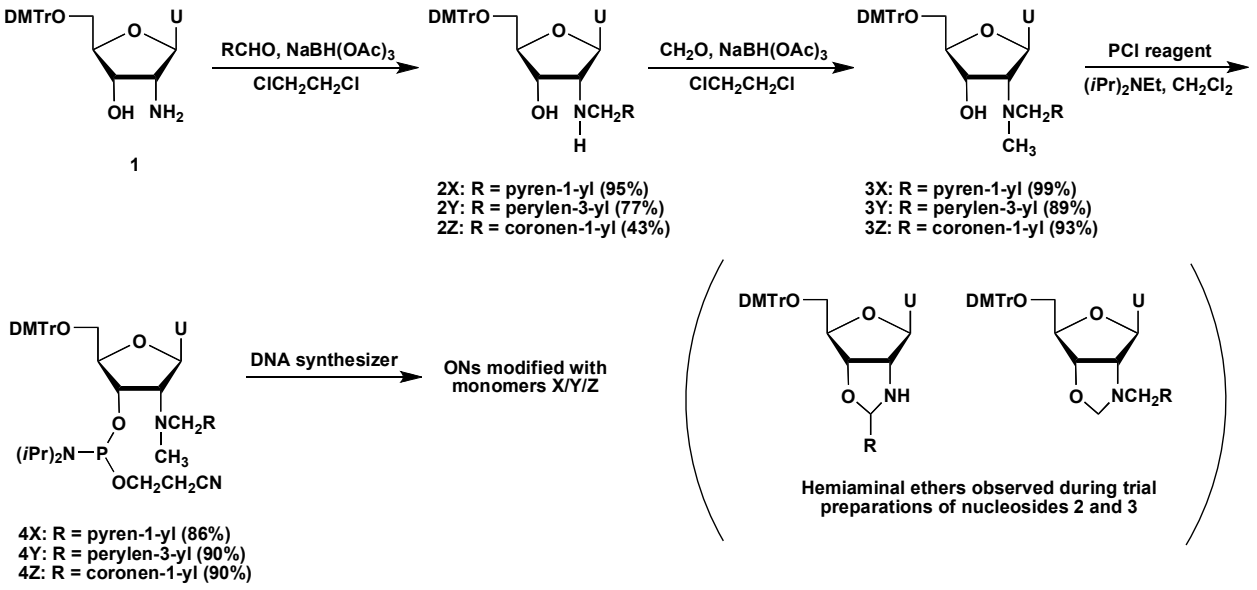
Figure 3-1: Illustration of the Invader approach for recognition of mixed-sequence DNA and structures of monomers used herein. Droplets denote intercalating moieties.

3.2 Results and Discussion

3.2.1 Synthesis of N2′-functionalized 2′-N-methyl-2′-aminodeoxyuridine Phosphoramidites

Our original synthesis of phosphoramidite **4X** proceeded in only ~10% overall yield over seven steps from uridine, largely due to moderate yields during direct N2′-alkylation of 2′-amino-2′-deoxy-2′-*N*-methyl-5′-*O*-(4,4′-dimethoxytrityl)uridine (46% yield, $PyCH_2Cl/Et_3N$/THF/80 °C).[40] Reductive alkylation using 1-pyrenecarbaldehyde and sodium triacetoxyborohydride or sodium cyanoborohydride offered no improvement due to concomitant formation of a cyclic N2′,O3′-hemiaminal ether, which probably was formed due to steric crowding at the 2′-position.[40] Motivated by previous reports on reductive alkylations on less hindered 2′-amino-2′-deoxyuridines,[42] we set out to devise a route to **4X** in which *N*-arylation is carried out prior to *N*-methylation and which can be adapted for the synthesis of **4Y** and **4Z**. O5′-DMTr protected

2′-amino-2′-deoxyuridine **1** – obtained in 53% yield from uridine over four steps[43] – was used as a starting material (Scheme 3-1). Reductive alkylation of **1** using sodium triacetoxyborohydride[44] and the appropriate aromatic aldehyde afforded nucleosides **2X-2Z** (43-95%, Scheme 3-1). It is noteworthy that the reaction yield decreases with increasing bulk of the aromatic moiety. Subsequent reductive methylation using sodium triacetoxyborohydride and formaldehyde furnished nucleosides **3X-3Z** in excellent yields. We found it necessary to use an excess of sodium triacetoxyborohydride to minimize formation of cyclic N2′,O3′-hemiaminal ethers during N2′-alkylations (Scheme 3-1). Treatment of nucleosides **3X-3Z** with 2-cyanoethyl-*N*,*N*-diisopropylchorophosphoramidite (PCl reagent) and *N*,*N*-diisopropylethylamine gave target phosphoramidites **4X-4Z** in high yields. This new route to **4X** is a major improvement over existing routes[42] (~52% yield from uridine over six steps versus 5-10 % yield from uridine over seven or eight steps).



Scheme 3-1: Synthesis of N2′-functionalized 2′-amino-2′-deoxy-2′-*N*-methyluridine phosphoramidites. U = uracil-1-yl; DMTr = 4,4′-dimethoxytrityl; PCl reagent = 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite.

3.2.2 Synthesis of Modified ONs and Experimental Design

Phosphoramidites **4X**, **4Y** and **4Z** were used in machine-assisted solid phase DNA synthesis to incorporate monomers **X**-**Z** into oligodeoxyribonucleotides (ONs) using extended hand-coupling (15 min) and the following activators: **4X** (5-[3,5-bis(trifluoromethyl)phenyl]-1*H*-tetrazole, ~99% coupling yield), **4Y** (pyridinium hydrochloride, ~90% coupling yield) and **4Z** (5-[3,5-bis(trifluoromethyl)phenyl]-1*H*-tetrazole, ~80% coupling yield). Suitable activators were identified through screening of common activators (results not shown). The identity and purity of the modified ONs was established through MALDI-TOF (Table 3-S1) and ion-pair

reverse phase HPLC (>85% purity), respectively. **Y**-modified ONs were found to be light-sensitive and were therefore stored in the dark until use.

Novel monomers **Y** and **Z** were studied in the same 9-mer mixed-sequence contexts that we have used for evaluation of other potential Invader building blocks.[32] Previously reported data for **X**-modified ONs are included to facilitate direct comparison. ONs containing a single incorporation in the 5′-G**B**G ATA TGC context are denoted **X1**, **Y1** and **Z1**. Similar conventions apply for the **B2**-**B6** series (Table 3-1). Reference DNA and RNA strands are denoted **D1**/**D4** and **R1**/**R4**, respectively (see footnote, Table 3-1).

3.2.3 Thermostability of Duplexes Between Modified ONs and Complementary DNA/RNA

Thermal denaturation temperatures (T_m's) of duplexes between **B1**-**B6** and complementary DNA or RNA (cDNA/cRNA) were determined in medium salt phosphate buffer ([Na^+] = 110 mM, pH 7.0[*]). ONs with one incorporation of monomer **Y** or **Z** form exceptionally stable duplexes with cDNA (ΔT_m from +7.5 to +21.0 °C, Table 3-1). In fact, duplexes modified with perylene monomer **Y** are 5-10 °C more stable than the corresponding pyrene-modified duplexes (T_m trend: **Y** ≥ **Z** > **X**) and are even slightly more stable than duplexes modified with 2′-*N*-(pyren-1-yl)carbonyl-2′-amino-α-L-LNA-T monomers,[32] which are among the most strongly stabilizing modified nucleotides reported till date. Incorporation of a second monomer as a next-nearest neighbor results in near-additive increases in T_m's (compare ΔT_m's for **B4**-, **B5**- and **B6**-series, Table 3-1). The degree of stabilization is strongly dependent on the sequence context,

[*] For representative thermal denaturation curves, see Supporting Information Figure 3-S1.

which is consistent with observations with other intercalator-modified ONs.[33,45,46] For example, ONs in which the modification is flanked by 3′-purines form more stable duplexes than if flanked by 3′-pyrimidines (e.g., compare ΔT_m's for **B2-** and **B4-**series, Table 3-1). This suggests that the aromatic moieties intercalate in the 3′-direction as particularly strong π-π-stacking interactions are possible with purines.[46]

Duplexes with cRNA are much less stable, and in some cases destabilized relative to reference duplexes (ΔT_m = -6.5 to +14.0 °C, Table 3-1; trend: **Z** > **Y** > **X**), which is another indicator of intercalation, as intercalators generally favor the less compressed *B*-type helix geometry of DNA:DNA duplexes. [38-40,45,47-49] As a consequence, these ONs display substantial selectivity for DNA targets, defined as $\Delta\Delta T_m$ (DNA-RNA) = ΔT_m (vs cDNA) - ΔT_m (vs cRNA) > 0 °C, with **Y**-modified ONs displaying particularly remarkable DNA selectivity ($\Delta\Delta T_m$ (DNA-RNA) between 11.0 and 23.5 °C, Table 3-S2).

Table 3-1: Thermal denaturation temperatures of duplexes between **B1**-**B6** and cDNA or cRNA relative to reference duplexes.[a]

			ΔT_m [°C]					
			+ cDNA			+ cRNA		
ON	Sequence	B =	**X**	**Y**	**Z**	**X**	**Y**	**Z**
B1	5′-G**B**G ATA TGC		+5.0[b]	+11.5	+7.5	-2.0[b]	-6.5	+0.5
B2	5′-GTG A**B**A TGC		+15.0	+20.0	+21.0	+3.0	+7.0	+14.0
B3	5′-GTG ATA **B**GC		+9.0	+16.0	+14.0	-0.5	+2.5	+1.0
B4	3′-CAC **B**AT ACG		+1.5[b]	+11.5	+7.5	-6.5[b]	-4.0	+1.0
B5	3′-CAC TA**B** ACG		+15.0	+20.0	+20.0	+3.0[b]	+9.0	+11.0
B6	3′-CAC **B**A**B** ACG		+14.0[b]	+31.0	+24.5	-3.0[b]	+7.5	+7.0

[a]ΔT_m = change in T_m relative to reference duplexes **D1**:**D4** ($T_m \equiv 29.5$ °C), **D1**:**R4** ($T_m \equiv 27.5$ °C) or **R1**:**D4** ($T_m \equiv 27.5$ °C), where **D1**: 5′-GTG ATA TGC, **D4**: 3′-CAC TAT ACG, **R1**: 5′-GUG AUA UGC and **R4**: 3′-CAC UAU ACG; T_m's are determined as the maximum of the first derivative of melting curves (A_{260} vs T) recorded in medium salt phosphate buffer ($[Na^+] = 110$ mM, $[Cl^-] = 100$ mM, pH 7.0 (NaH_2PO_4/Na_2HPO_4)), using 1.0 µM of each strand. Reported T_m's are averages of at least two measurements within 1.0 °C; A = adenin-9-yl DNA monomer, C = cytosin-1-yl DNA monomer, G = guanin-9-yl DNA monomer, T = thymin-1-yl DNA monomer. For structures of monomers **X**-**Z**, see Figure 3-1.
[b]Data previously reported in reference 40.

3.2.4 Binding Specificity

The binding specificity of centrally modified ONs (**B2**-series) was studied using DNA strands with mismatched nucleotides opposite to the modification (Table 3-2). **X2**/**Y2**/**Z2** discriminate C- and T-mismatched DNA targets with similar efficiency as unmodified **D1** whereas G-mismatched DNA targets are poorly discriminated, indicating that the wobble base pair is greatly stabilized by the intercalating pyrene moiety.

Table 3-2: Discrimination of mismatched DNA targets by **X2**/**Y2**/**Z2** and reference strands.[a]

		DNA: 3′-CAC T**B**T ACG			
		T_m [°C]	ΔT_m [°C]		
ON	**Sequence** **B** =	**A**	C	G	T
D1	5′-GTG ATA TGC	**29.5**	-16.5	-9.5	-17.0
X2[b]	5′-GTG A**X**A TGC	**44.5**	-23.0	-3.5	-13.0
Y2	5′-GTG A**Y**A TGC	**49.5**	-19.5	-4.0	-17.0
Z2	5′-GTG A**Z**A TGC	**50.5**	-15.5	-3.0	-15.5

[a]For experimental conditions, see Table 3-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched duplex.
[b]From reference 40.

ONs with two modifications positioned as next-nearest neighbors (**B6**-series) display improved discrimination of DNA targets with a mismatched nucleotide opposite to the central 2′-deoxyadenosine residue, with binding fidelity decreasing in the order: **Y6** > **Z6** ≥ **X6** (Table 3-3). DNA strands with mismatched A- and G-nucleotides are particularly efficiently discriminated. For data with mismatched RNA, see Figures 3-S3 and 3-S4 in the Supporting Information.

These results indicate that **X**/**Y**/**Z**-modified ONs should be designed in a manner that places likely single nucleotide polymorphism (SNP) sites opposite to canonical 2′-deoxyribonucleotides rather than the modified monomers, to ensure optimal thermal discrimination of mismatched targets.

Table 3-3: Discrimination of mismatched DNA targets by **X6**/**Y6**/**Z6** and reference strands.[a]

			DNA: 5′-GTG A**B**A TGC			
			T_m [°C]	ΔT_m [°C]		
ON	**Sequence**	**B** =	**T**	A	C	G
D4	3′-CAC TAT ACG		**29.5**	-17.0	-15.5	-9.0
X6[b]	3′-CAC **X**A**X** ACG		**43.5**	-21.5	-10.5	-13.5
Y6	3′-CAC **Y**A**Y** ACG		**60.5**	-25.5	-22.5	-18.0
Z6	3′-CAC **Z**A**Z** ACG		**54.0**	-22.5	-16.5	-12.5

[a]For experimental conditions, see Table 3-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched duplex.
[b]From reference 40.

3.2.5 Photophysical Characterization of Modified ONs and Duplexes with Complementary DNA/RNA.

UV-Vis absorption and steady-state fluorescence emission spectra of **Y**- or **Z**-modified ONs were recorded in the absence or presence of complementary DNA/RNA to gain further insight into the binding modes of the attached aromatic hydrocarbons. Hybridization of **Y**- or **Z**-modified ONs with cDNA/cRNA results in minor bathochromic shifts of the hydrocarbon absorption maxima ($\Delta\lambda_{max}$ = 0-3 nm, Table 3-4, Figures 3-S2 and 3-S3), which is indicative of ground-state electronic interactions between hydrocarbons and nucleobases and, accordingly, intercalation.[50,51] However, smaller bathochromic shifts are observed than for the pyrene-modified **X1**-**X6**. We speculate that this is because the perylene and coronene moieties are not fully contained within the duplex core.[52] Structural studies beyond the scope of the present study are needed to verify this hypothesis.

Table 3-4: Absorption maxima in the 300-500 nm region for **X**/**Y**/**Z**-modified ONs and the corresponding duplexes with complementary DNA or RNA.[a]

		λ_{max} [$\Delta\lambda_{max}$]/nm								
	B =	**X**[b]			**Y**			**Z**		
ON	Sequence	SSP	+cDNA	+cRNA	SSP	+cDNA	+cRNA	SSP	+cDNA	+cRNA
B1	5′-G**B**G ATA TGC	349	353 [+4]	351 [+2]	448	450 [+2]	450 [+2]	312	314 [+2]	313 [+1]
B2	5′-GTG A**B**A TGC	348	353 [+5]	351 [+3]	451	453 [+2]	453 [+2]	313	314 [+1]	313 [±0]
B3	5′-GTG ATA **B**GC	349	353 [+4]	354 [+5]	451	452 [+1]	452 [+1]	313	314 [+1]	313 [±0]
B4	3′-CAC **B**AT ACG	349	354 [+5]	349 [±0]	450	452 [+2]	452 [+2]	312	314 [+2]	313 [+1]
B5	3′-CAC TA**B** ACG	348	354 [+6]	352 [+4]	450	450 [±0]	452 [+2]	313	314 [+1]	313 [±0]
B6	3′-CAC **B**A**B** ACG	ND	ND	ND	449	451 [+2]	452 [+3]	310	313 [+3]	313 [+3]

[a] Measurements were performed at 5 °C (**X**, **Y**) or 10 °C (**Z**) using a spectrophotometer and quartz optical cells with 1.0 cm path lengths. For buffer conditions, see Table 3-1. ND = not determined.
[b]Data for the **X**-series, with the exception of **X3**, have been previously reported in reference 40.

Steady-state fluorescence emission spectra (λ_{ex} = 420 nm, T = 5 °C) of duplexes between perylene-modified **Y1**-**Y6** and cDNA/cRNA feature two peaks at ~460 nm and ~490 nm (Figure 3-2 and 3-S4). Hybridization with DNA/RNA targets generally results in moderately increased fluorescence intensity (0.8- to 4.4-fold), with more pronounced increases being observed upon DNA binding. Similar trends have been observed with other perylene-functionalized ONs in which hybridization-induced intercalation is a likely binding mode.[36] Unlike pyrene,[50,53-56] the fluorescence intensity of perylene is strongest in hydrophobic environments and much less sensitive to quenching by flanking nucleobases.[56-58.]

Fluorescence emission spectra of the coronene-modified **Z1**-**Z6** display three main emission peaks at ~435 nm, ~455 nm and ~483 nm along with several shoulders when excited at λ_{max} =

310 nm (T = 10 °C), which corresponds to a remarkable Stokes shift of > 125 nm (Figure 3-2 and 3-S5). Hybridization with cDNA/cRNA only has a minor impact on fluorescence intensity, ranging from slight decreases (**Z5** vs cDNA) to moderate increases (**Z6** vs cDNA/cRNA). Although only a few studies have been conducted with coronene-modified ONs,[59,60] it is interesting to note that isosequential ONs modified with the closely related 2′-*O*-(coronen-1-yl)methyluridines display virtually identical photophysical characteristics,[60] which underscores intercalation as a likely binding mode.

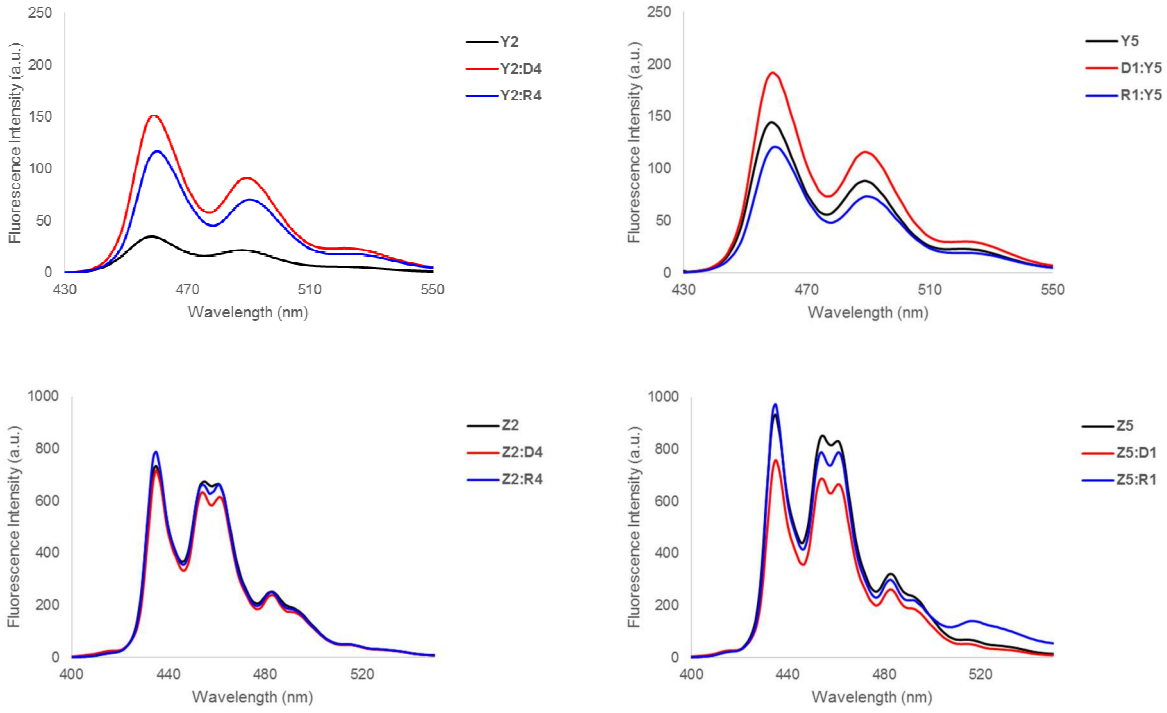


Figure 3-2: Steady-state fluorescence emission spectra of representative **Y**/**Z**-modified ONs and the corresponding duplexes with DNA/RNA targets. Spectra were recorded at 5 °C (**Y**-modified) or 10 °C (**Z**-modified) using λ_{ex} = 420 and 310 nm for **Y**- and **Z**-modified ONs, respectively. Each strand was used at 1.0 µM concentration in T_m buffer. Note that different axis scales are used.

3.2.6 Biophysical Properties of Duplexes with Interstrand Zippers of X/Y/Z-monomers

After establishing that the perylene and coronene moieties of **Y**- and **Z**-modified ONs most likely intercalate upon duplex formation – a pre-requisite for their potential use as Invader modifications – we went on to study double-stranded probes with different interstrand zipper arrangements of these monomers as potential dsDNA-targeting probes (Table 3-5). The term thermal advantage (TA = ΔT_m (ON_A:cDNA) + ΔT_m(cDNA:ON_B) - ΔT_m (ON_A:ON_B), where ON_A:ON_B is a duplex with an interstrand zipper arrangement of monomers), serves as a first approximation to describe the energy difference between the 'products' and 'reactants' of the recognition process, with more positive values signifying greater dsDNA recognition potential (i.e. via the process depicted in Figure 3-1)

Double-stranded probes with +1 monomer zippers are more thermolabile and energetically activated for dsDNA-recognition than probes with other zipper configurations (compare T_m's and TA values for **B2**:**B5** relative to other probe duplexes, Table 3-5), which mirrors the trends with other Invader probes.[31,32,35,36] Perylene-modified duplex **Y2**:**Y5** is the most strongly activated probe in this series. The coronene-modified **Z2**:**Z5** displays lower dsDNA-targeting potential as the probe duplex is unusually stable (TA trend: **Y2**:**Y5**>**X2**:**X5**≥**Z2**:**Z5**, Table 3-5). The above T_m based conclusions are corroborated by thermodynamic parameters for duplex formation, which were derived from denaturation curves via line fitting.[61] Thus, formation of duplexes between **Y**- or **Z**-modified ONs and cDNA is considerably more favorable than formation of unmodified reference duplexes ($\Delta\Delta G^{293}$ between -24 and -6 kJ/mol, first and second ΔG^{293} columns, Table 3-5), and more favorable than the corresponding **X**-modified duplexes. The stabilization is enthalpic in origin ($\Delta\Delta H$ < 0 kJ/mol in most cases, Table 3-S5). Formation of **B2**:**B5** duplexes – as well as +2 zipper duplexes **B1**:**B4** – is 25-35 kJ/mol less

favorable than duplexes with other interstrand zipper arrangements of **Y**- or **Z**-monomers (compare $\Delta\Delta G^{293}$ values in third ΔG^{293} column, Table 3-5). The energetic activation of the **B2**:**B5** probes is weakly enthalpic in nature ($\Delta\Delta H \geq 0$ kJ/mol for **B2**:**B5**, Table 3-S5). Consequentially, **B2**:**B5** probes – and to a far lesser degree **B1**:**B4** probes – display favorable binding energy for recognition of isosequential dsDNA targets as estimated by ΔG_{rec}^{293} (**ON$_A$**:**ON$_B$**) = ΔG^{293} (**ON$_A$**:cDNA) + ΔG^{293} (cDNA:**ON$_B$**) - ΔG^{293} (**ON$_A$**:**ON$_B$**) - ΔG^{293} (dsDNA) (i.e., $\Delta G_{rec}^{293} << 0$ kJ/mol, Table 3-5). The trend in the ΔG_{rec}^{293} values (**Y2**:**Y5**>**X2**:**X5**>**Z2**:**Z5**, Table 3-5) identifies the perylene-modified **Y2**:**Y5** as the most strongly activated probe for dsDNA-recognition among the studied duplexes.

Table 3-5: Biophysical properties of **X**/**Y**/**Z**-modified probe duplexes.[a]

ON	ZP	Probe duplex	T_m (°C)	*TA* (°C)	ΔG^{293}[$\Delta\Delta G^{293}$] (kJ/mol)			ΔG^{293}_{rec} (kJ/mol)	λ_{max} (nm)
					upper ON vs cDNA	lower ON vs cDNA	probe duplex		
X1 **X5**	**+4**	5′-G**X**G ATA TGC 3′-CAC TA**X** ACG	49.0	-1.5	-51±1 [-6]	-64±1 [-19]	N/A	-	353
X1 **X4**	**+2**	5′-G**X**G ATA TGC 3′-CAC **X**AT ACG	28.0	+8.0	-51±1 [-6]	-48±1 [-3]	-44±1 [+1]	-10	350
X2 **X5**	**+1**	5′-GTG A**X**A TGC 3′-CAC TA**X** ACG	**28.5**	**+31.0**	**-65±1 [-20]**	**-64±1 [-19]**	**-44±0 [+1]**	**-40**	**345**
X2 **X4**	**-1**	5′-GTG A**X**A TGC 3′-CAC **X**AT ACG	42.5	+2.5	-65±1 [-20]	-48±1 [-3]	-54±1 [-9]	-14	352
Y1 **Y5**	**+4**	5′-G**Y**G ATA TGC 3′-CAC TA**Y** ACG	54.0	+8.5	-51±0 [-6]	-69±1 [-24]	-84±3 [-39]	+9	452
Y1 **Y4**	**+2**	5′-G**Y**G ATA TGC 3′-CAC **Y**AT ACG	34.5	+18.0	-51±0 [-6]	-58±1 [-13]	-49±0 [-4]	-15	450
Y2 **Y5**	**+1**	5′-GTG A**Y**A TGC 3′-CAC TA**Y** ACG	**32.5**	**+37.5**	**-69±3 [-24]**	**-69±1 [-24]**	**-49±1 [-4]**	**-44**	**448**
Y2 **Y4**	**-1**	5′-GTG A**Y**A TGC 3′-CAC **Y**AT ACG	57.5	+1.0	-69±3 [-24]	-58±1 [-13]	-74±4 [-29]	-8	452
Z1 **Z5**	**+4**	5′-G**Z**G ATA TGC 3′-CAC TA**Z** ACG	61.5	-4.5	-57±0 [-12]	-66±1 [-21]	-79±1 [-34]	+1	314
Z1 **Z4**	**+2**	5′-G**Z**G ATA TGC 3′-CAC **Z**AT ACG	43.5	+1.5	-57±0 [-12]	-57±1 [-12]	-55±1 [-10]	-14	313
Z2 **Z5**	**+1**	5′-GTG A**Z**A TGC 3′-CAC TA**Z** ACG	**41.0**	**+29.5**	**-68±1 [-23]**	**-66±1 [-22]**	**-54±1 [-9]**	**-35**	**309**
Z2 **Z4**	**-1**	5′-GTG A**Z**A TGC 3′-CAC **Z**AT ACG	62.5	-4.5	-68±1 [-23]	-57±1 [-12]	-80±1 [-35]	±0	314

[a] ZP = zipper. For conditions of thermal denaturation and absorption experiments, see Table 3-1 and Table 3-4, respectively. $DA_{ONA:ONB} = \Delta T_m$ (ON_A:ON_B) – [ΔT_m (ON_A:cDNA) + ΔT_m(cDNA:ON_B)]. $\Delta\Delta G^{293}$ is measured relative to ΔG^{293} for **D1**:**D4** = -45 kJ/mol. ΔG^{293}_{rec} (**ON_A**:**ON_B**) = ΔG^{293} (**ON_A**:cDNA) + ΔG^{293} (cDNA:**ON_B**) - ΔG^{293} (**ON_A**:**ON_B**) - ΔG^{293} (dsDNA). "±" denotes standard deviation. N/A = the absence of a clear lower base line precluded determination of this value. T_m's and *DA*'s for all **X**-modified duplexes except those involving **X3**, have been previously published in reference 32, but are included to facilitate direct comparison.

The results from the present and earlier studies[31-36] clearly demonstrate that high dsDNA-targeting potential is an inherent characteristic of probes with +1 interstrand zippers of monomers with intercalating moieties. Presumably, this monomer configuration forces two intercalators into the same region within the duplex core, which leads to a violation of the

'nearest-neighbor exclusion principle'[37] and structural perturbation of the duplex.[32,36] These effects also manifest themselves in the absorption maxima of the intercalators (Figure 3-S6 in Supporting Information). Thus, significantly blue-shifted maxima are observed for **B2**:**B5** probes relative to probes with other zipper configurations (compare λ_{max} for **B2**:**B5** and other probe duplexes (Table 3-5)), which indicates decreased interactions with neighboring nucleobases due to duplex perturbation. Moreover, **B2**:**B5** probes also exhibit distinctive steady-state fluorescence emission spectra as compared to probes with other zipper configurations (Figure 3-3). Thus, **X2**:**X5** displays the highest fluorescence intensity as perturbation of intercalation decreases pyrene-nucleobase interactions resulting in reduced fluorescence quenching. Conversely, **Y2**:**Y5** and **Z2**:**Z5** display low fluorescence intensity as intercalation-mediated duplex perturbation exposes the fluorophores to the polar, and in this case, quenching grooves. It is also interesting to point out that the emission spectrum of **Z2**:**Z5** contains less vibrational fine structure, which further indicates structural perturbation.

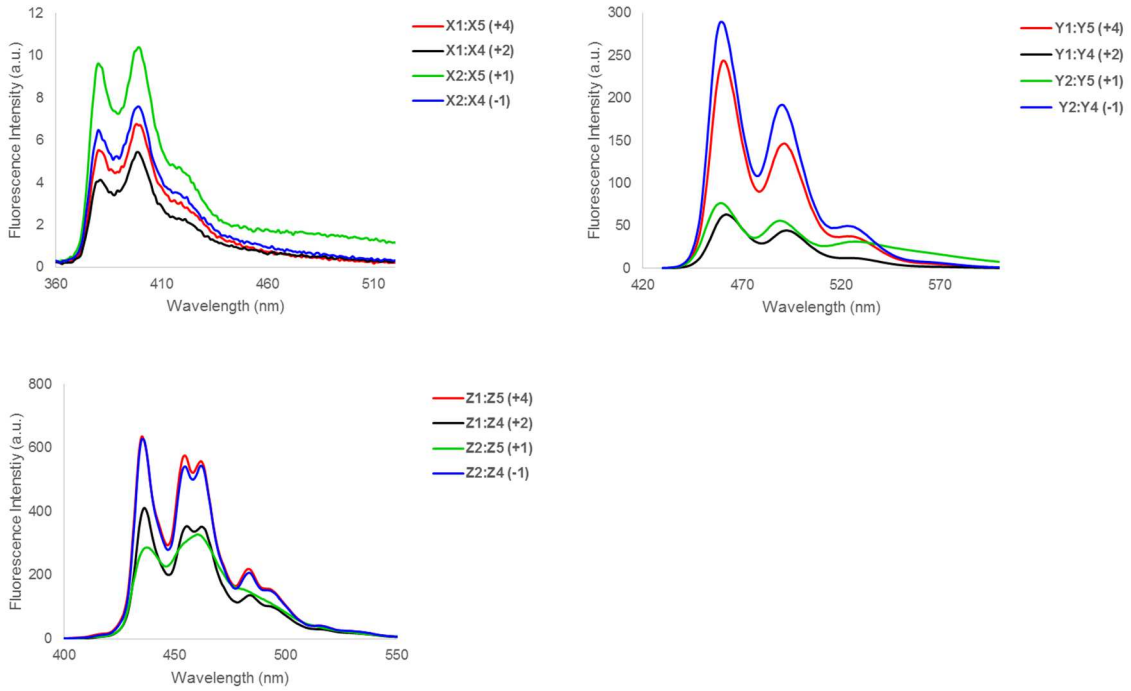


Figure 3-3: Steady-state fluorescence emission spectra of duplexes with different interstrand zippers of **X**, **Y**, or **Z** monomers (zipper type indicated in parentheses). For experimental conditions, see Figure 3-2. Spectra for **X**-modified duplexes were previously reported in reference 32 but are included for comparison. Please note that different axis scales are used.

3.2.7 Recognition of DNA Hairpins using Energetically Activated Probe Duplexes

The *TA* and ΔG^{293}_{rec} values, in concert, clearly point toward double-stranded probes with +1 interstrand zipper configurations of monomers **X**/**Y**/**Z** as the most thermodynamically activated constructs for dsDNA recognition. We therefore set out to experimentally test the recognition efficiency of these probes using a 3′-digoxigenin (DIG) labeled DNA hairpin (DH) – comprised of a 9-mer double-stranded mixed sequence stem that is linked by a T_{10} loop – as a model dsDNA target (Figure 3-4a). Incubation of **DH1** with **Y2**:**Y5**, **Z2**:**Z5**, or benchmark Invader

X2:**X5** in HEPES buffer at ambient temperature for 12-16 hours, resulted in dose-dependent formation of a more slowly migrating band in non-denaturing PAGE gels, which is indicative of ternary recognition complex formation (Figure 3-4c). Analysis of the corresponding dose-response curves revealed that **X2**:**X5**, **Y2**:**Y5**, and **Z2**:**Z5** display C_{50} values of ~0.80 μM, ~0.45 μM, and ~0.63 μM, respectively (Figure 3-4d). It is particularly noteoworthy that even as little as 0.5 molar equivalents of **Y2**:**Y5** or **Z2**:**Z5** resulted in ~20% recognition of **DH1**. Complete recognition is accomplished when Invaders are used at 100-fold molar excess relative to **DH1** (Figure 3-4d). Less recognition is observed when using shorter incubation times (3h), especially at higher concentrations, presumably due to slow reaction kinetics (Figure 3-S7). However, we have shown that recognition kinetics can be dramatically improved through incorporation of additional energetic hotspots.[31]

As a control, single-stranded ONs **X2**/**X5**/**Y2**/**Y5**/**Z2**/**Z5** were incubated with **DH1** for 12-16 h under otherwise identical conditions. Significantly less efficient dsDNA recognition was observed (C_{50} between 4.0 μM and >17.2 μM, Figures 3-S8 and 3-S9), underlining that both strands comprising an Invader probe are necessary to drive dsDNA recognition to completion. Lastly, the binding specificities of **Y2**:**Y5**, **Z2**:**Z5**, and benchmark Invader **X2**:**X5** were examined by incubating the probes with DNA hairpins **DH2** and **DH3**, which are fully base-paired but which deviate in the nucleotide sequence at one or two positions relative to the Invader probes (underlined residues indicate sequence deviations, Figure 3-4b). Remarkably, even when using **X2**:**X5**, **Y2**:**Y5** or **Z2**:**Z5** at a 500-fold molar excess, mismatched DNA hairpins are not recognized, while efficient recognition of matched **DH1** is observed (Figure 3-4e). This demonstrates that Invader-mediated dsDNA-recognition proceeds with excellent fidelity.

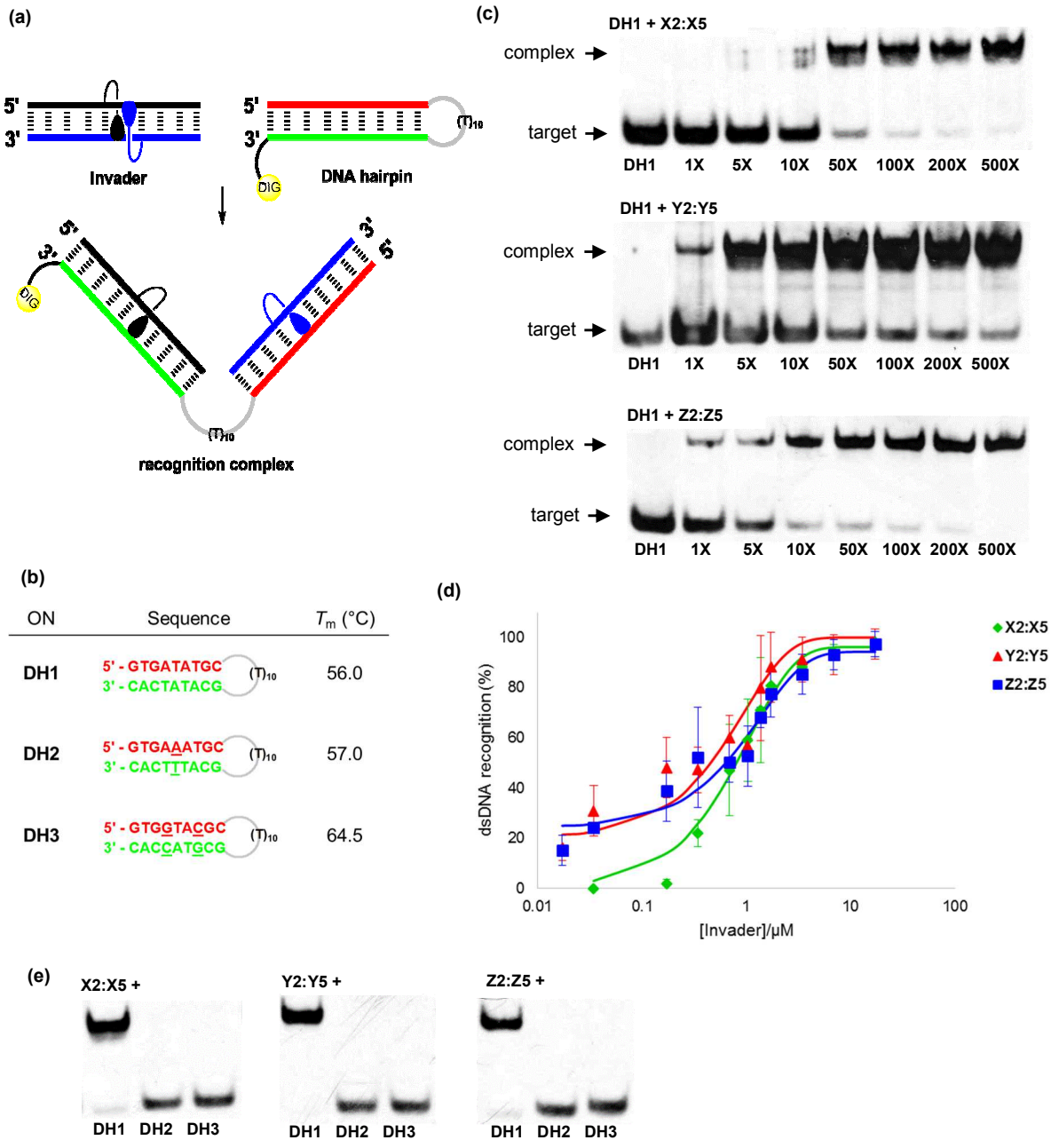


Figure 3-4: Recognition of DNA hairpins using activated double-stranded probes. (a) Illustration of recognition process; (b) sequences of DNA hairpins with iso-sequential (**DH1**) or mismatched stems (**DH2** and **DH3**) – underlined nucleotides indicate positions of mismatches relative to probes. (c) representative electrophoretograms illustrating recognition of **DH1** using 1- to 500-fold excess of **X2**:**X5**, **Y2**:**Y5**, or **Z2**:**Z5**; (d) dose-response curves

(average of at least three independent experiments, error bars represent standard deviation); (e) electrophoretograms illustrating incubation of **DH1**-**DH3** with 500-fold molar excess of **X2**:**X5**, **Y2**:**Y5**, or **Z2**:**Z5**. Experimental conditions for electrophoretic mobility shift assay: separately pre-annealed targets (34.4 nM) and probes (variable concentrations) were incubated 12-16h at ambient temperature in 1X HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 10% sucrose, 1.4 mM spermine tetrahydrochloride, pH 7.2) and then run on 16% non-denaturing PAGE (performed at 70V, 2.5h, ~4 °C) using 0.5x TBE as a running buffer (45 mM Tris, 45 mM boric acid, 1mM EDTA); DIG: digoxigenin.

3.3 Conclusions

Efficient synthetic routes to N2′-pyrene/perylene/coronene-functionalized 2′-*N*-methyl-2′-aminodeoxyuridine phosphoramidites have been developed. ONs that are modified with these building blocks form very stable duplexes with cDNA (ΔT_m/modification between +1.5 °C and +21.0 °C), with greater stabilization being observed with the larger aromatic hydrocarbons. The observed trends in absorption and fluorescence emission upon hybridization with cDNA strongly suggests that the extraordinary duplex stabilization is due to intercalation of the labels. DNA duplexes with +1 interstrand zipper arrangements of these monomers are much less stable, but their stability increases with intercalator size. The results from the present and earlier studies[31-36] clearly demonstrate that the activated nature is an inherent property of double-stranded probes with +1 interstrand zippers of intercalator-functionalized monomers. As a consequence of these stability trends, Invaders based on N2′-perylene-functionalized 2′-*N*-methyl-2′-amino-DNA monomers were predicted to be most strongly activated for dsDNA

recognition. Experiments using DNA hairpins as model dsDNA targets confirmed this and, moreover, showed that mixed-sequence recognition of dsDNA proceeds with excellent fidelity. Invaders based on N2′-perylene-functionalized 2′-*N*-methyl-2′-amino-DNA monomers therefore present themselves as particularly interesting probes for dsDNA-targeting applications in molecular biology, nucleic acid diagnostics, and biotechnology.

3.4 Experimental Section

3.4.1 Synthesis of N2'-intercalator-functionalized 2'-amino-2'-deoxyuridine Phosphoramidites

2′-Amino-2′-deoxy-2′-*N*-(pyren-1-ylmethyl)-5′-*O*-(4,4′-dimethoxytrityl)uridine (2X). Nucleoside **1** (200 mg, 0.37 mmol) was coevaporated with anhydrous 1,2-dichloroethane (2 × 3 mL) and redissolved in anhydrous 1,2-dichloroethane (2 mL). To this was added $NaBH(OAc)_3$ (120 mg, 0.55 mmol) and 1-pyrenecarboxaldehyde (105 mg, 0.44 mmol) and the reaction mixture was stirred under an argon atmosphere at room temperature for 5 h. Saturated aqueous $NaHCO_3$ (25 mL) was added and the aqueous layer was extracted with CH_2Cl_2 (3 × 15 mL). The organic layers were dried (Na_2SO_4) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-4% MeOH in CH_2Cl_2, v/v) to afford **2X** (0.27 g, 95%) as a white foam. R_f = 0.5 (5% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m/*z 782.2849 ($[M+Na]^+$, $C_{47}H_{41}N_3O_7{\cdot}Na^+$, calcd 782.2837); ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.37 (br s, ex, 1H, NH(U)), 8.45 (d, 1H, J = 9.1 Hz, Py), 8.26-8.29 (m, 2H, Py), 8.16-8.20 (m, 2H, Py), 8.14 (ap s, 2H, Py), 8.04-8.10 (m, 2H, Py), 7.60 (d, 1H, J = 8.2 Hz, H6), 7.28-7.32 (m, 2H, DMTr), 7.21-7.26 (m, 2H, DMTr), 7.15-7.20 (m, 5H, DMTr), 6.77-6.83 (m, 4H, DMTr),

5.92 (d, 1H, J = 6.3 Hz, H1′), 5.63 (d, ex, 1H, J = 5.8 Hz, 3′-OH), 5.19 (d, 1H, J = 8.2 Hz, H5), 4.49-4.59 (m, 2H, CH_2Py), 4.24-4.28 (m, 1H, H3′), 4.01-4.05 (m, 1H, H4′), 3.68 (s, 3H, CH_3O), 3.66 (s, 3H, CH_3O), 3.39-3.44 (m, 1H, H2′), 3.24-3.28 (dd, 1H, J = 10.6 Hz, 4.0 Hz, H5′), 3.16-3.20 (dd, 1H, J = 10.6 Hz, 3.3 Hz, H5′), 2.58-2.65 (m, ex, 1H, $NHCH_2$); ^{13}C NMR (125 MHz, DMSO-d_6) δ 162.8, 158.05, 158.02, 150.7, 144.4, 140.2 (C6), 135.3, 135.0, 134.1, 130.8, 130.3, 130.0, 129.7 (Ar), 129.6 (Ar), 128.5, 127.8 (Ar), 127.6 (Ar), 127.4 (Ar), 127.1 (Ar), 126.8 (Ar), 126.7 (Ar), 126.6 (Ar), 126.1 (Ar), 125.03 (Ar), 124.98 (Ar), 124.6 (Ar), 124.1, 124.0, 123.3 (Ar), 113.2 (Ar), 113.1 (Ar), 101.5 (C5), 87.3 (C1′), 85.9, 84.1 (C4′), 68.4 (C3′), 63.5 (C2′), 63.4 (C5′), 54.9 (OCH_3), 48.8 (CH_2Py).

2′-Amino-2′-deoxy-2′-*N*-(perylen-3-ylmethyl)-5′-*O*-(4,4′-dimethoxytrityl)uridine (2Y). Nucleoside **1** (0.28 g, 0.50 mmol) was coevaporated with anhydrous 1,2-dichloroethane (2 × 5 mL) and redissolved in anhydrous 1,2-dichloroethane (5 mL). To this was added $NaBH(OAc)_3$ (0.75 g, 3.53 mmol) followed by slow addition of 3-perylenecarboxaldehyde[45] (185 mg, 0.66 mmol) over 1.5 h. The reaction mixture was stirred under an argon atmosphere at room temperature for 22 h at which point it was diluted with EtOAc (30 mL) and washed with saturated aqueous $NaHCO_3$ (2 × 20 mL). The combined aqueous layers were back-extracted with EtOAc (3 × 15 mL) and the combined organic layers were dried (Na_2SO_4) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-100% EtOAc in petroleum ether, v/v) to afford **2Y** (0.32 g, 77%) as a yellow foam. R_f = 0.3 (5% MeOH in CH_2Cl_2, v/v); MALDI-HRMS m/z 832.3020 ($[M+Na]^+$, $C_{51}H_{43}N_3O_7{\cdot}Na^+$, calcd 832.2993); ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.36 (br s, ex, 1H, NH(U)), 8.36 (t, 2H, J = 7.5 Hz, Pery), 8.29 (d, 1H, J = 7.5 Hz, Pery), 8.22 (d, 1H, J = 7.5 Hz, Pery), 8.03 (d, 1H, J = 8.5

Hz, Pery), 7.76-7.81 (m, 2H, Pery), 7.60-7.63 (d, 1H, J = 8.0 Hz, H6), 7.51-7.56 (m, 4H, Pery), 7.15-7.35 (m, 9H, DMTr), 6.81-6.86 (m, 4H, DMTr), 5.87 (d, 1H, J = 6.0 Hz, H1′), 5.63 (d, ex, 1H, J = 5.0 Hz, 3′-OH), 5.25 (d, 1H, J = 8.0 Hz, H5), 4.24-4.28 (m, 1H, H3′), 4.19-4.23 (m, 2H, CH_2Pery), 4.00-4.06 (m, 1H, H4′), 3.68 (s, 3H, CH_3O), 3.66 (s, 3H, CH_3O), 3.33-3.39 (m, 1H, H2′), 3.25-3.30 (m, 1H, H5′), 3.17-3.22 (m, 1H, H5′); ^{13}C NMR (125 MHz, DMSO-d_6) δ 162.8, 158.06, 158.03, 150.7, 144.4, 140.2 (C6), 135.9, 135.4, 135.1, 134.2, 132.6, 130.8, 130.6, 130.5, 129.7 (DMTr), 129.6 (DMTr), 128.2, 127.8 (Ar), 127.6 (Ar), 126.8 (Pery), 126.7 (Pery), 126.6 (Pery), 126.5 (Pery), 123.9 (Pery), 120.65 (Pery), 120.60 (Pery), 120.4 (Pery), 120.1 (Pery), 113.2 (DMTr), 113.1 (DMTr), 101.5 (C5), 87.2 (C1′), 85.9, 84.1 (C4′), 68.3 (C3′), 63.5 (C2′), 63.4 (C5′), 54.9 (CH_3O), 48.7 (CH_2Pery). A minor impurity of EtOAc was identified.[46]

2′-Amino-2′-deoxy-2′-*N*-(coronen-1-ylmethyl)-5′-*O*-(4,4′-dimethoxytrityl)uridine (2Z). Nucleoside **1** (0.30 g, 0.55 mmol) was coevaporated with anhydrous 1,2-dichloroethane (2 × 5 mL) and redissolved in anhydrous 1,2-dichloroethane (2.5 mL). This was slowly added over 1 h to a stirred solution of $NaBH(OAc)_3$ (240 mg, 1.10 mmol) and 1-coronenecarboxaldehyde[47] (0.27 g, 0.82 mmol) in anhydrous 1,2-dichloroethane (3 mL). The reaction mixture was stirred under an argon atmosphere at room temperature for 14 h at which point it was diluted with CH_2Cl_2 (40 mL) and washed with saturated aqueous $NaHCO_3$ (2 × 20 mL) and H_2O (20 mL). The organic layer was dried (Na_2SO_4) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-1.5% MeOH in CH_2Cl_2, v/v – initially built with 0.5% Et_3N) to afford **2Z** (205 mg, 43%) as a pale yellow foam. R_f = 0.8 (10% MeOH in CH_2Cl_2, v/v); MALDI-HRMS m/z 880.2998 ($[M+Na]^+$, $C_{55}H_{43}N_3O_7 \cdot Na^+$, calcd 880.2993); ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.45 (br s, ex, 1H, NH(U)), 9.16-9.18 (d, 1H, J = 9.0 Hz, Cor),

8.91-9.00 (m, 7H, Cor), 8.90 (d, 1H, J = 1.7 Hz, Cor), 8.88 (s, 1H, Cor), 8.74-8.76 (d, 1H, J = 8.5 Hz, Cor), 7.63 (d, 1H, J = 8.2 Hz, H6), 7.04-7.27 (m, 9H, DMTr), 6.69 (d, 2H, J = 9.0 Hz, DMTr), 6.65 (d, 2H, J = 9.0 Hz, DMTr), 6.07 (d, 1H, J = 6.5 Hz, H1′), 5.74 (d, ex, 1H, J = 4.4 Hz, 3′-OH – overlap with residual CH_2Cl_2), 5.14 (d, 1H, J = 8.2 Hz, H5), 4.98 (dd, 2H, J = 13.7 Hz, 4.1 Hz, CH_2Cor), 4.33-4.39 (m, 1H, H3′), 4.09-4.12 (m, 1H, H4′), 3.60-3.63 (m, 1H, H2′), 3.53 (s, 3H, CH_3O), 3.47 (s, 3H, CH_3O), 3.25-3.30 (dd, 1H, J = 10.5 Hz, 4.2 Hz, H5′), 3.18-3.22 (dd, 1H, J = 10.5 Hz, 3.5 Hz, H5′), 2.91-2.97 (m, ex, 1H, NH); ^{13}C NMR (125 MHz, DMSO-d_6) δ 162.9, 157.95, 157.89, 150.8, 144.3, 140.2 (C6), 135.3, 135.0, 134.7, 129.6 (DMTr), 129.5 (DMTr), 128.24, 128.15, 128.04, 127.97, 127.73, 127.70 (DMTr), 127.6 (DMTr), 126.9, 126.6 (DMTr), 126.3 (Cor), 126.24, 126.23 (Cor), 126.21 (Cor), 126.1 (Cor), 126.0 (Cor), 125.2 (Cor), 122.4 (Cor), 122.0, 121.7, 121.6, 121.4, 121.3, 120.9, 113.05 (DMTr), 113.01 (DMTr), 101.5 (C5), 87.3 (C1′), 85.9, 84.3 (C4′), 68.5 (C3′), 63.7 (C2′), 63.5 (C5′), 54.8 (CH_3O), 54.7 (CH_3O), 49.5 (CH_2Cor).

General procedure for preparation of nucleosides 3 (description for ~1 mmol scale). The appropriate nucleoside **2** was dissolved in anhydrous 1,2-dichloroethane. To this was added $NaBH(OAc)_3$, followed by dropwise addition of 37% aqueous solution of CH_2O (stabilized with ~12% MeOH) over 30 seconds. The reaction mixture was then stirred under an argon atmosphere at room temperature until analytical TLC indicated completion (quantities and reaction times are specified below). The reaction mixture was then worked up and purified as specified below to afford nucleosides **3** (yields specified below).

2′-Amino-2′-deoxy-2′-*N*-methyl-2′-*N*-(pyren-1-ylmethyl)-5′-*O*-(4,4′-dimethoxytrityl)uridine (3X). Nucleoside **2X** (1.30 g, 1.71 mmol), $NaBH(OAc)_3$ (3.63 g, 17.1 mmol), CH_2O (37% solution, 130 μL, 2.57 mmol) and anhydrous 1,2-dichloroethane (12 mL) were reacted as described above (4 h). Saturated aqueous $NaHCO_3$ (100 mL) was added very slowly and the aqueous layer was extracted with CH_2Cl_2 (2 × 50 mL). The combined organic layers were dried (Na_2SO_4) and evaporated to dryness, and the resulting residue was purified by silica gel column chromatography (0-5% MeOH in CH_2Cl_2, v/v) to afford **3X** (1.35 g, quant.) as a white foam. R_f = 0.4 (5% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m/*z 796.2969 ($[M+Na]^+$, $C_{48}H_{44}N_3O_7 \cdot Na^+$, calcd 796.2993); ^{13}C NMR is in agreement with previous data.[40]

2′-Amino-2′-deoxy-2′-*N*-methyl-2′-*N*-(perylen-3-ylmethyl)-5′-*O*-(4,4′-dimethoxytrityl)uridine (3Y). Nucleoside **2Y** (1.00 g, 1.23 mmol), $NaBH(OAc)_3$ (2.61 g, 12.3 mmol), CH_2O (37% solution, 100 μL, 1.86 mmol) and anhydrous 1,2-dichloroethane (20 mL) were reacted as described above (7 h). The reaction mixture was diluted with EtOAc (100 mL) and very slowly washed with saturated aqueous $NaHCO_3$ (2 × 75 mL). The combined aqueous layer was back-extracted with EtOAc (3 × 30 mL) and the combined organic layers were dried (Na_2SO_4) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-60% EtOAc in petroleum ether, v/v) to afford **3Y** (0.91 g, 89%) as a bright yellow foam. R_f = 0.4 (60% EtOAc in petroleum ether, v/v); MALDI-HRMS *m/*z 846.3174 ($[M+Na]^+$, $C_{52}H_{45}N_3O_7 \cdot Na^+$, calcd 846.3150); ^1H NMR (500 MHz, DMSO-d_6) δ 11.39 (br d, ex, J = 2.0 Hz, NH(U)), 8.33-8.37 (m, 2H, Pery), 8.29 (d, 1H, J = 8.0 Hz, Pery), 8.23 (d, 1H, J = 8.0 Hz, Pery), 8.08 (d, 1H, J = 8.0 Hz, Pery), 7.76-7.80 (m, 2H, Pery), 7.60 (d, 1H, J = 8.0 Hz, H6), 7.42-7.55 (m, 4H, Pery), 7.18-7.40 (m, 9H, DMTr), 6.83-6.91 (m, 4H, DMTr), 6.39

(d, 1H, J = 8.5 Hz, H1′), 5.48 (d, ex, 1H, J = 5.5 Hz, 3′-OH), 5.43 (dd, 1H, J = 8.0 Hz, 2.0 Hz, H5), 4.37-4.41 (m, 1H, H3′), 4.10-4.18 (2d, 2H, J = 13.3 Hz, CH_2-pery), 4.02-4.06 (m, 1H, H4′), 3.71 (s, 3H, CH_3O), 3.70 (s, 3H, CH_3O), 3.35-3.41 (m, 1H, H2′), 3.27-3.31 (m, 1H, H5′ - partial overlap with H_2O signal), 3.15-3.19 (m, 1H, H5′), 2.37 (s, 3H, CH_3); ^{13}CNMR (125 MHz, DMSO-d_6) δ 162.7, 158.09, 158.08, 150.5, 144.5, 140.1 (C6), 135.4, 135.1, 134.5, 134.2, 133.0, 130.64, 130.58, 130.4, 130.0, 129.73 (DMTr), 129.67 (DMTr), 129.6, 128.3, 128.1 (Pery), 127.83 (DMTr), 127.78 (DMTr), 127.72 (Pery), 127.65 (DMTr), 127.62, 126.84 (Pery), 126.80 (Pery), 126.7 (DMTr), 126.3 (Pery), 124.7 (Pery), 120.6 (Pery), 120.4 (Pery), 120.0 (Pery), 113.21 (DMTr), 113.19 (DMTr), 102.0 (C5), 85.9, 85.1 (C4′), 83.2 (C1′), 71.2 (C3′), 67.6 (C2′), 64.1 (C5′), 57.7 (CH_2Pery), 55.0 (CH_3O), 38.6 (NCH_3 – overlap with DMSO-d_6 signal).

2′-Amino-2′-deoxy-2′-*N*-(coronen-1-ylmethyl)-2′-*N*-methyl-5′-*O*-(4,4′-dimethoxytrityl)uridine (3Z). Nucleoside **2Z** (120 mg, 0.14 mmol), $NaBH(OAc)_3$ (0.39 g, 1.40 mmol), CH_2O (37% solution, 12 μL, 0.21 mmol) and anhydrous 1,2-dichloroethane (2 mL) were reacted as described above (5 h). The reaction mixture was diluted with CH_2Cl_2 (20 mL) and very slowly washed with saturated aqueous $NaHCO_3$ (2 × 20 mL). The aqueous was back-extracted with CH_2Cl_2 (2 × 10 mL) and the combined organic layers were dried (Na_2SO_4) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-2% MeOH in CH_2Cl_2, v/v – initially built with 0.5% Et_3N, v/v) to afford **3Z** (113 mg, 93%) as a pale yellow foam. R_f = 0.7 (5% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m/*z 894.3161 ($[M+Na]^+$, $C_{56}H_{45}N_3O_7 \cdot Na^+$, calcd 894.3155); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.10 (br s, 1H, ex, NH(U)), 8.91-8.95 (d, 1H, J = 8.8 Hz, Cor), 8.70 (d, 1H, J = 8.8 Hz, Cor),

8.50-8.67 (m, 7H, Cor), 8.46 (br s, 1H, Cor), 8.39-8.42 (d, 1H, *J* = 8.5 Hz, Cor), 7.91 (d, 1H, *J* = 8.0 Hz, H6), 7.30-7.35 (m, 2H, DMTr), 7.16-7.27 (m, 7H, DMTr – partial overlap with $CDCl_3$), 6.71-6.78 (m, 5H, DMTr + H1′), 5.34 (d, 1H, *J* = 8.0 Hz, H5), 4.93 (d, 1H, *J* = 12.5 Hz, CH_2Cor), 4.45 (d, 1H, *J* = 12.5 Hz, CH_2Cor), 4.17-4.22 (m, 1H, H3′), 4.12-4.15 (m, 1H, H4′), 3.98 (br s, 1H, ex, 3′-OH), 3.71 (s, 3H, CH_3O), 3.69 (s, 3H, CH_3O), 3.58-3.62 (m, 1H, H2′), 3.41-3.45 (dd, 1H, *J* = 10.5 Hz, 2.8 Hz, H5′), 3.33-3.37 (dd, 1H, *J* = 10.5 Hz, 2.8 Hz, H5′), 2.54 (s, 3H, NCH_3); ^{13}CNMR (125 MHz, $CDCl_3$) δ 163.2, 158.98, 158.95, 150.6, 144.5, 140.7 (C6), 135.4, 135.2, 131.1, 130.4 (DMTr), 130.3 (DMTr), 128.8, 128.6, 128.54, 128.50, 128.4 (Ar), 128.2 (Ar), 128.1 (Ar), 127.7, 127.4 (DMTr), 127.3, 126.5 (Cor), 126.4 (Cor), 126.3 (Cor), 126.2 (Cor), 126.1 (Cor), 126.0 (Cor), 125.8 (Cor), 123.0, 122.5, 122.3, 122.2, 122.0 (Cor), 113.5 (DMTr), 103.3 (C5), 87.5, 85.4 (C4′), 84.9 (C1′), 70.7 (C2′), 70.5 (C3′), 63.7 (C5′), 60.2 (CH_2Cor), 55.4 (CH_3O), 40.5 (NCH_3).

General procedure for preparation of nucleosides 4 (description for ~1 mmol scale). The appropriate nucleoside **3** was coevaporated with anhydrous 1,2-dichloroethane (10 mL) and redissolved in anhydrous CH_2Cl_2. To this was added anhydrous *N*,*N*-diisopropylethylamine (DIPEA) followed by dropwise addition of 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (PCl reagent) and the reaction mixture was allowed to stir under an argon atmosphere at room temperature until analytical TLC indicated complete conversion (quantities and reaction times are specified below). Unless otherwise mentioned, cold EtOH (1 mL) was added and all solvents were evaporated off. The resulting residue was purified by silica gel column chromatography and subsequent precipitation from CH_2Cl_2 and petroleum ether to afford the desired phosphoramidite **4**.

2′-Amino-2′-deoxy-2′-*N*-methyl-2′-*N*-(pyren-1-yl-methyl)-3′-O-(*N*,*N*-diisopropylamino-2-cyanoethoxyphosphinyl)-5′-O-(4,4′-dimethoxytrityl)uridine (4X). Nucleoside **3X** (1.34 g, 1.73 mmol), PCl reagent (0.77 mL, 3.46 mmol), anhydrous DIPEA (1.50 mL, 8.67 mmol) and anhydrous CH_2Cl_2 (20 mL) were reacted and worked up as described above (2.5 h). Purification by silica gel column chromatography (0-50% EtOAc in petroleum ether, v/v) and precipitation from CH_2Cl_2 and petroleum ether afforded nucleoside **4X** as a white foam (1.45 g, 86%). R_f = 0.5 (50% EtOAc in petroleum ether, v/v); MALDI-HRMS *m/*z 996.4083 ($[M+Na]^+$, $C_{57}H_{60}N_5O_8P{\cdot}Na^+$, calcd 993.4077); ^{31}P NMR (121 MHz, $CDCl_3$) δ 151.0, 149.8. ^{31}P NMR data are in agreement with literature data.[40,42]

2′-Amino-2′-deoxy-2′-*N*-methyl-2′-*N*-(perylen-3-ylmethyl)-3′-*O*-(*N*,*N*-diisopropylamino-2-cyanoethoxyphosphinyl)-5′-*O*-(4,4′-dimethoxytrityl)uridine (4Y). Nucleoside **3Y** (0.40 g, 0.49 mmol), PCl reagent (220 μL, 0.97 mmol), anhydrous DIPEA (0.34 mL, 1.94 mmol) and anhydrous CH_2Cl_2 (5 mL) were reacted as described above (2 h). Absolute EtOH (~1 mL) was added and the reaction mixture was diluted with CH_2Cl_2 (30 mL) and washed with saturated aqueous $NaHCO_3$ (20 mL). The aqueous layer was back-extracted with CH_2Cl_2 (2 × 10 mL) and the combined organic layer was dried (Na_2SO_4) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-50% EtOAc in petroleum ether, v/v) to afford nucleoside **4Y** (0.45 g, 90%) as a bright yellow foam. R_f = 0.4 (60% EtOAc in petroleum ether, v/v); MALDI-HRMS *m/*z 1046.4272 ($[M+Na]^+$, $C_{61}H_{62}N_5O_8P{\cdot}Na^+$, calcd 1046.4228); ^{31}P NMR (121 MHz, $CDCl_3$) δ 150.9, 149.7.

2′-Amino-2′-deoxy-2′-*N*-(coronen-1-ylmethyl)-2′-*N*-methyl-3′-*O*-(*N*,*N*-diisopropylamino-2-cyanoethoxyphosphinyl)-5′-*O*-(4,4′-dimethoxytrityl)uridine (4Z). Nucleoside **3Z** (0.27 g, 0.31 mmol), PCl reagent (210 μL, 0.93 mmol), anhydrous DIPEA (0.27 mL, 1.55 mmol) and anhydrous CH_3CN (1.5 mL) were reacted and worked up as described above (2.5 h). Purification by silica gel column chromatography (0-1% MeOH in CH_2Cl_2, v/v - initially built with 0.5% Et_3N, v/v) and precipitation from CH_2Cl_2 and petroleum ether afforded nucleoside **4Z** (0.30 g, 90%) as a pale yellow foam. R_f = 0.6 (3% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m/*z 1072.4399 ($[M+Na]^+$, $C_{65}H_{62}N_5O_8P{\cdot}Na^+$, calcd 1072.4409); ^{31}P NMR (121 MHz, $CDCl_3$) δ 151.0, 149.8.

3.4.2 Protocol - Synthesis and Purification of ONs

Modified ONs were synthesized on a 0.2 μmol scale using a DNA synthesizer and succinyl linked LCAA-CPG (long chain alkyl amine controlled pore glass) columns with a pore size of 500Å. Standard protocols for incorporation of DNA monomers were used. The following hand-coupling conditions were used for incorporation of monomers **X**-**Z** (coupling time; activator; coupling yield): **4X** (15 min; 5-[3,5-bis(trifluoromethyl)phenyl]-1*H*-tetrazole; ~99%), **4Y** (15 min; pyridinium hydrochloride; ~90%) and **4Z** (15 min; 5-[3,5-bis(trifluoromethyl)phenyl]-1*H*-tetrazole; CH_2Cl_2; ~80%). All modified phosphoramidites were used at 50-fold molar excess and 0.05 M concentration in CH_3CN (**4X**) or CH_2Cl_2 (**4Y**/**4Z**). Extended oxidation (45s) was used. Cleavage from solid support and removal of protecting groups was accomplished upon treatment with 32% aq. ammonia (55 °C, 12 h). ONs were purified in the DMT-on mode via ion-pair reverse phase HPLC (C18 column) using a 0.05 M triethylammonium acetate - water/acetonitrile gradient. This was followed by detritylation (80% aq. AcOH) and

precipitation (NaOAc/$NaClO_4$/acetone, -18 °C for 12-16 h). The identity of synthesized ONs was established through MALDI-MS analysis (Table 3-S1) recorded in positive ions mode on a quadrupole time-of-flight tandem mass spectrometer equipped with a MALDI source using anthranilic acid, 3-hydroxypicolinic acid (3-HPA) or 2′,4′,6′-trihydroxyacetophenone (THAP) as matricies. Purity was verified by ion-pair reverse phase HPLC running in analytical mode (>85%). ONs modified with monomer **Y** were stored in the dark (wrapped in aluminum foil) to prevent light-induced bleaching/degradation of the fluorophore. ONs stored in this manner were stable for at least 12 months (>85% purity).

3.4.3 Protocol - Thermal Denaturation Studies

ON concentrations were estimated using the following extinction coefficients for DNA (OD/μmol): G (12.01), A (15.20), T (8.40), C (7.05); RNA (OD/μmol): G (13.70), A (15.40), U (10.00), C (9.00); and hydrocarbons (OD/μmol): pyrene (22.4)[66], perylene (33.2)[67] and coronene (36.0).[59] Strands were thoroughly mixed and denatured by heating to 70-85 °C, followed by cooling to the starting temperature of the experiment. Quartz optical cells with a path length of 1.0 cm were used. Thermal denaturation temperatures (T_m's) of duplexes (1.0 μM final concentration of each strand) were measured using a UV/VIS spectrophotometer equipped with a 12-cell Peltier temperature controller and determined as the maximum of the first derivative of the thermal denaturation curve (A_{260} vs. T) recorded in medium salt phosphate buffer (T_m buffer: 100 mM NaCl, 0.1 mM EDTA and pH 7.0 adjusted with 10 mM Na_2HPO_4 and 5 mM Na_2HPO_4). The temperature of the denaturation experiments ranged from at least 15 °C below T_m to 20 °C above T_m (although not below 3 °C). A temperature ramp of 0.5 °C/min was used in all experiments. Reported T_m's are averages of two experiments within ± 1.0 °C.

3.4.4 Protocol - Determination of Thermodynamic Parameters

Thermodynamic parameters for duplex formation were determined through baseline fitting of denaturation curves (van't Hoff analysis) using software provided with the UV/Vis spectrometer. Bimolecular reactions, two-state melting behavior, and a heat capacity change of $\Delta C_p = 0$ upon hybridization were assumed.[61] A minimum of two experimental denaturation curves were each analyzed at least three times to minimize errors arising from baseline choice. Averages and standard deviations are listed.

3.4.5 Protocol - Absorption Spectra

UV-vis absorption spectra (range 200-600 nm) were recorded at 5 °C (**X**- and **Y**-modified ONs/duplexes) or 10 °C (**Z**-modified ONs/duplexes) using the same samples and instrumentation as in the thermal denaturation experiments.

3.4.6 Protocol - Steady-state Fluorescence Emission Spectra

Steady-state fluorescence emission spectra of ONs modified with monomers **X**-**Z** and the corresponding duplexes with complementary DNA/RNA targets, were recorded in non-deoxygenated thermal denaturation buffer (each strand at 1.0 μM concentration) and obtained as an average of five scans using an excitation wavelength of λ_{ex} = 350, 420 or 310 nm for **X**-, **Y**- or **Z**-modified ONs, respectively. Excitation and emission slits of 5.0 nm and 2.5 nm, respectively were used along with a scan speed of 600 nm/min. Experiments were determined at 5 °C (**X**/**Y**) or 10 °C (**Z**) under N_2 flow to ascertain maximal hybridization of probes to DNA/RNA targets.

3.4.7 Protocol - Electrophoretic Mobility Shift Assay

This assay was performed essentially as previously described.[35] Unmodified DNA hairpins **DH1**-**DH3** were obtained from commercial sources and used without further purification. The DNA hairpins were 3′-DIG-labeled using the 2nd generation DIG Gel Shift Kit (Roche Applied Bioscience) per the manufacturer's recommendation. DIG-labeled ONs obtained in this manner were diluted and used without further purification in the recognition experiments. Pre-annealed probes (85 °C for 10 min, cooled to room temperature over 15 min) and DIG-labeled DNA hairpins (34.4 nM) were mixed and incubated in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 10% sucrose, 1.44 mM spermine tetrahydrochloride, pH 7.2) for the specified time at ambient temperature (~21±3 °C). The reaction mixtures were then diluted with 6x DNA loading dye (Fermentas) and loaded onto a 16% non-denaturing polyacrylamide gel. Electrophoresis was performed using a constant voltage of 70 V for 2.5 h at ~4 °C using 0.5x TBE as a running buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA). Gels were blotted onto positively charged nylon membranes (Roche Applied Bioscience) using constant voltage with external cooling (100V, ~4 °C). The membranes were exposed to anti-digoxigenin-AP F_{ab} fragments as recommended by the manufacturer of the DIG Gel Shift Kit, transferred to a hybridization jacket, and incubated with the substrate (CSPD) in detection buffer for 10 min at 37 °C. The chemiluminescence of the formed product was captured on X-ray film, which was developed using an X-Omatic 1000A X-ray film developer (Kodak). The resulting bands were quantified using Image J software. Invasion efficiency was determined as the intensity ratio between the recognition complex band and the total lane. An average of three independent experiments is reported along with standard deviations. Non-linear regression was used to fit

data points from dose-response experiments, using a script written for the "Solver" module in Microsoft Office Excel.

3.4.8 Explanation of Zipper Nomenclature

The following nomenclature describes the relative arrangement between two monomers positioned on opposing strands in a duplex. The number *n* describes the distance measured in number of base pairs and has a positive value if a monomer is shifted toward the 5′-side of its own strand relative to a second reference monomer on the other strand. Conversely, *n* has a negative value if a monomer is shifted toward the 3′-side of its own strand relative to a second reference monomer on the other strand.

Acknowledgements

This study was supported by Award Number GM088697 from the National Institute of General Medical Sciences, National Institutes of Health. Scholarships from the National Science Foundation under award number 0648202 and the Department of Defense ASSURE (Awards to Stimulate and Support Undergraduate Research Experiences) Program (J.J.O) are appreciated. We thank Dr. Alex Blumenfeld (Dept. Chemistry, Univ. Idaho) and Dr. Lee Deobald (EBI Murdock Mass Spectrometry Center, Univ. Idaho) for assistance with NMR and mass spectrometric analysis, and Prof. Carolyn Bohach (Food Science, Univ. Idaho) for access to gel documentation stations.

3.5 Supporting Information

3.5.1 General Experimental Section

Reagents and solvents were commercially available, of analytical grade and used without further purification. Petroleum ether of the distillation range 60-80 °C was used. Solvents were dried over activated molecular sieves: CH_3CN (3Å); CH_2Cl_2, 1,2-dichloroethane, and *N,N*-diisopropylethylamine (4Å). Water content of anhydrous solvents was verified on Karl-Fisher apparatus. Reactions were monitored by TLC using silica gel coated plates with a fluorescence indicator (SiO_2-60, F-254) which were visualized a) under UV light and/or b) by dipping in 5% conc. H_2SO_4 in absolute ethanol (v/v) followed by heating. Silica gel column chromatography was performed with Silica gel 60 (particle size 0.040–0.063 mm) using moderate pressure (pressure ball). Evaporation of solvents was carried out under reduced pressure at temperatures below 45 °C. After column chromatography, appropriate fractions were pooled, evaporated and dried at high vacuum for at least 12h to give the obtained products in high purity (>95%) as ascertained by 1D NMR techniques. Exchangeable (ex) protons were detected by disappearance of signals upon D_2O addition. Assignments of NMR spectra are based on 2D spectra (COSY, HSQC) and DEPT-spectra. Quaternary carbons are not assigned in ^{13}C NMR but verified from HSQC and DEPT spectra (absence of signals). MALDI-HRMS spectra of compounds were recorded on a mass spectrometer using 2,5-dihydroxybenzoic acid as a matrix and polyethylene glycol (PEG 600) as an internal calibration standard.

3.5.2 Additional Tables, Figures, and Discussion

Table 3-S1. MALDI-MS of modified ONs.[a]

ONs	Sequence	Calc. *m*/*z* [M+H]	Found *m*/*z* [M+H]
X3	5′-GTG ATA **X**GC	2983	2983
Y1	5′-G**Y**G ATA TGC	3033	3032
Y2	5′-GTG A**Y**A TGC	3033	3033
Y3	5′-GTG ATA **Y**GC	3033	3032
Y4	3′-CAC **Y**AT ACG	2962	2962
Y5	3′-CAC TA**Y** ACG	2962	2962
Y6	3′-CAC **Y**A**Y** ACG	3241	3243
Z1	5′-G**Z**G ATA TGC	3081	3081
Z2	5′-GTG A**Z**A TGC	3081	3081
Z3	5′-GTG ATA **Z**GC	3081	3081
Z4	3′-CAC **Z**AT ACG	3010	3009
Z5	3′-CAC TA**Z** ACG	3010	3011
Z6	3′-CAC **Z**A**Z** ACG	3337	3337

[a] For structures of monomers **X**, **Y** and **Z**, see Figure 3-1 in the main manuscript. **X1**/**X2**/**X4**/**X5**/**X6** were made in a previous study.[40]

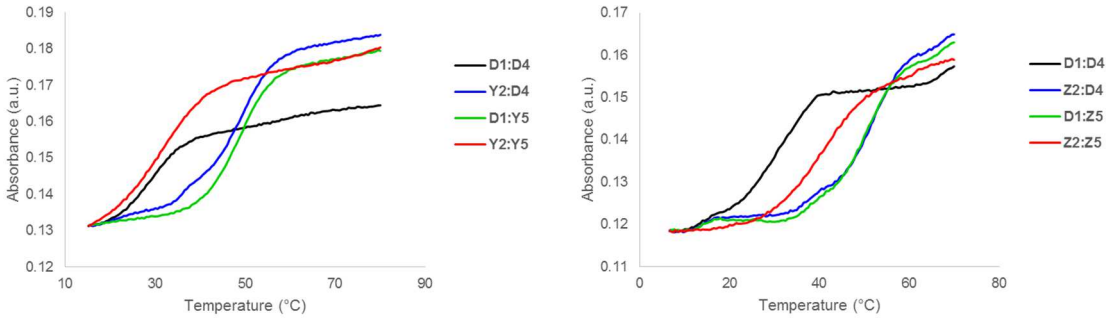


Figure 3-S1: Representative thermal denaturation curves of **Y**/**Z**-modified duplexes and reference duplex **D1**:**D4**. For experimental conditions, see Table 3-1.

Table 3-S2: DNA selectivity of **B1-B6**.[a]

			$\Delta\Delta T_m$ (DNA-RNA) [°C]		
ON	**Duplex**	**B =**	**X**	**Y**	**Z**
B1	5'-G**B**G ATA TGC		+7.0	+18.0	+7.0
B2	5'-GTG A**B**A TGC		+12.0	+13.0	+7.0
B3	5'-GTG ATA **B**GC		+9.5	+13.5	+13.0
B4	3'-CAC **B**AT ACG		+8.0	+15.5	+6.5
B5	3'-CAC TA**B** ACG		+12.0	+11.0	+9.0
B6	3'-CAC **B**A**B** ACG		+17.0	+23.5	+17.5

[a] DNA selectivity defined as $\Delta\Delta T_m$ (DNA-RNA) = ΔT_m (vs cDNA) - ΔT_m (vs cRNA).

Additional discussion of binding specificity.

Centrally modified **X2**/**Y2**/**Z2** display less efficient discrimination of mismatched RNA targets compared with reference strand **D1**, except when a mismatched G is opposite of the coronene-functionalized monomer **Z** (Table 3-S3).

Table 3-S3: Discrimination of mismatched RNA targets by **X2**/**Y2**/**Z2** and reference ONs.[a]

			RNA: 3'-CAC U**B**U ACG			
			T_m/°C	ΔT_m/°C		
ON	**Sequence**	**B** =	A	C	G	U
D1	5'-GTG ATA TGC		**27.5**	<-17.5	-4.5	<-17.5
X2[b]	5'-GTG A**X**A TGC		**30.5**	-16.5	-0.5	-13.0
Y2	5'-GTG A**Y**A TGC		**34.5**	-13.5	-1.0	-14.5
Z2	5'-GTG A**Z**A TGC		**41.5**	-13.0	-9.0	-14.5

[a] For conditions of thermal denaturation experiments, see Table 3-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched DNA:RNA duplex.
[b]Data previously reported in reference 40. Included to facilitate direct comparison.

Doubly modified ONs (**B6**-series) very efficiently discriminate RNA targets with a single mismatched nucleotide opposite of the central 2′-deoxyriboadenosine (Table 3-S4). Similar observations were made with mismatched DNA targets (Table 3-4).

Table 3-S4: Discrimination of mismatched RNA targets by **X6**/**Y6**/**Z6** and reference ONs.[a]

			RNA : 5'-GUG A**B**A ACG			
			T_m [°C]	ΔT_m [°C]		
ON	Sequence	**B** =	U	A	C	G
D4	3'-CAC TAT ACG		**27.5**	-7.0	-3.0	-4.5
X6	3'-CAC **X**A**X** ACG		**24.5**	<-12.5	<-12.5	-5.0
Y6	3'-CAC **Y**A**Y** ACG		**35.0**	-15.5	-12.5	-6.5
Z6	3'-CAC **Z**A**Z** ACG		**34.5**	-9.0	-8.0	-7.0

[a] For conditions of thermal denaturation experiments, see Table 3-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched DNA:RNA duplex.

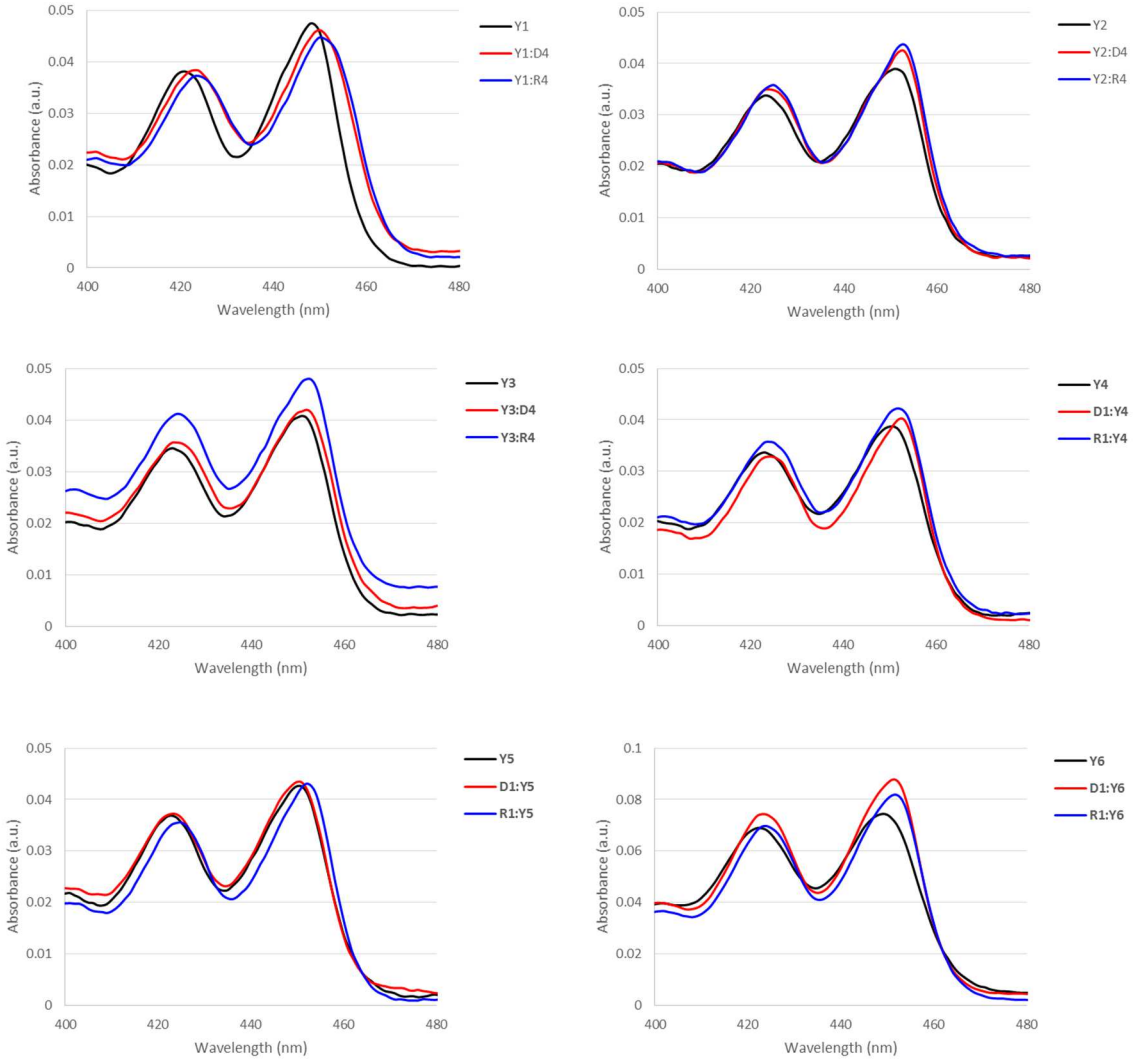


Figure 3-S2: Absorbance spectra of single-stranded **Y1-Y6** and their corresponding duplexes with DNA/RNA targets. Spectra were recorded at T = 5 °C using each strand at 1.0 μM concentration in T_m buffer. Note that a different Y-axis scale is used for **Y6**.

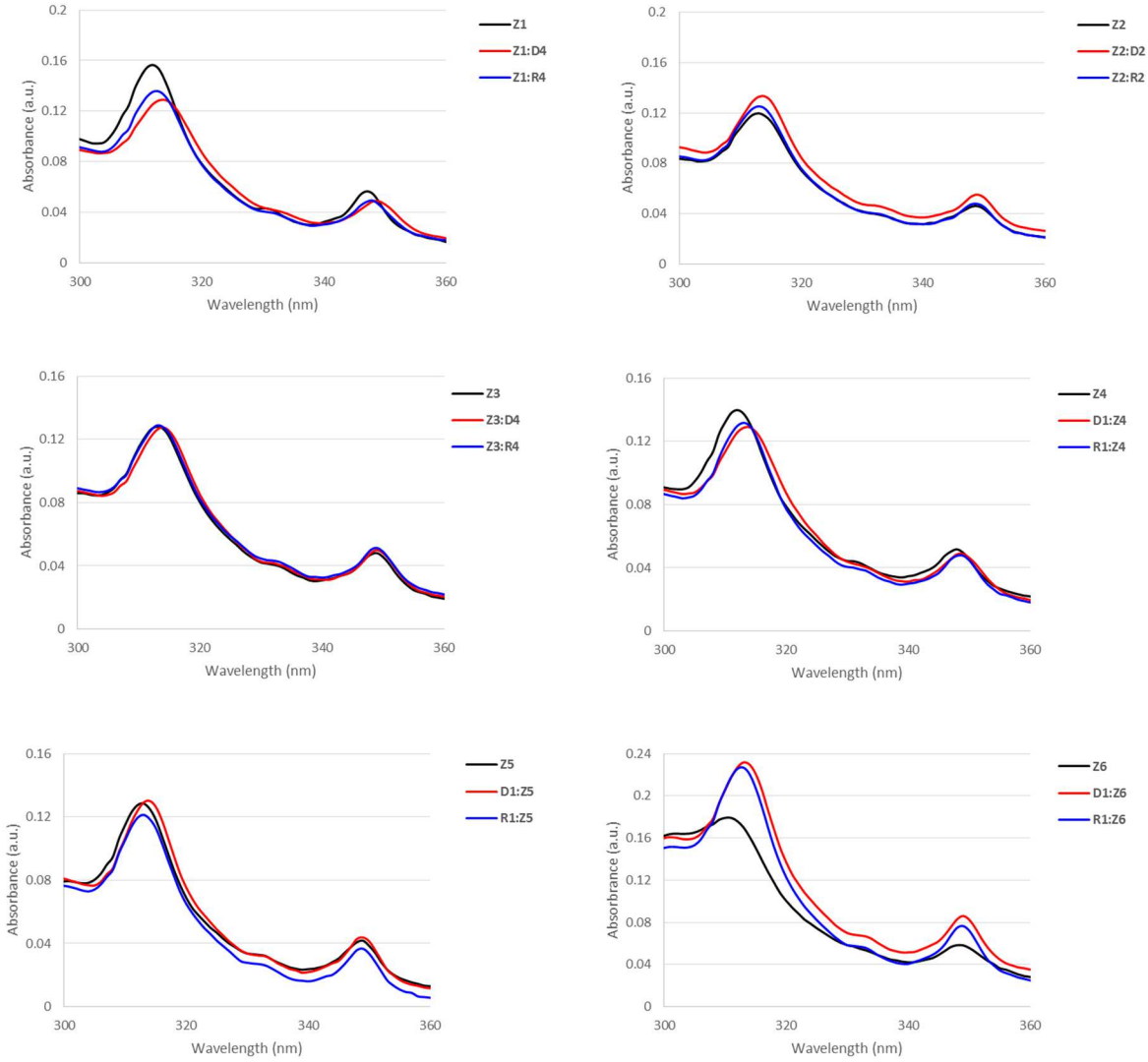


Figure 3-S3: Absorbance spectra of single-stranded **Z1**-**Z6** and their corresponding duplexes with DNA/RNA targets. Spectra were recorded at T = 10 °C using each strand at 1.0 µM concentration in T_m buffer. Note that different Y-axis scales are used.

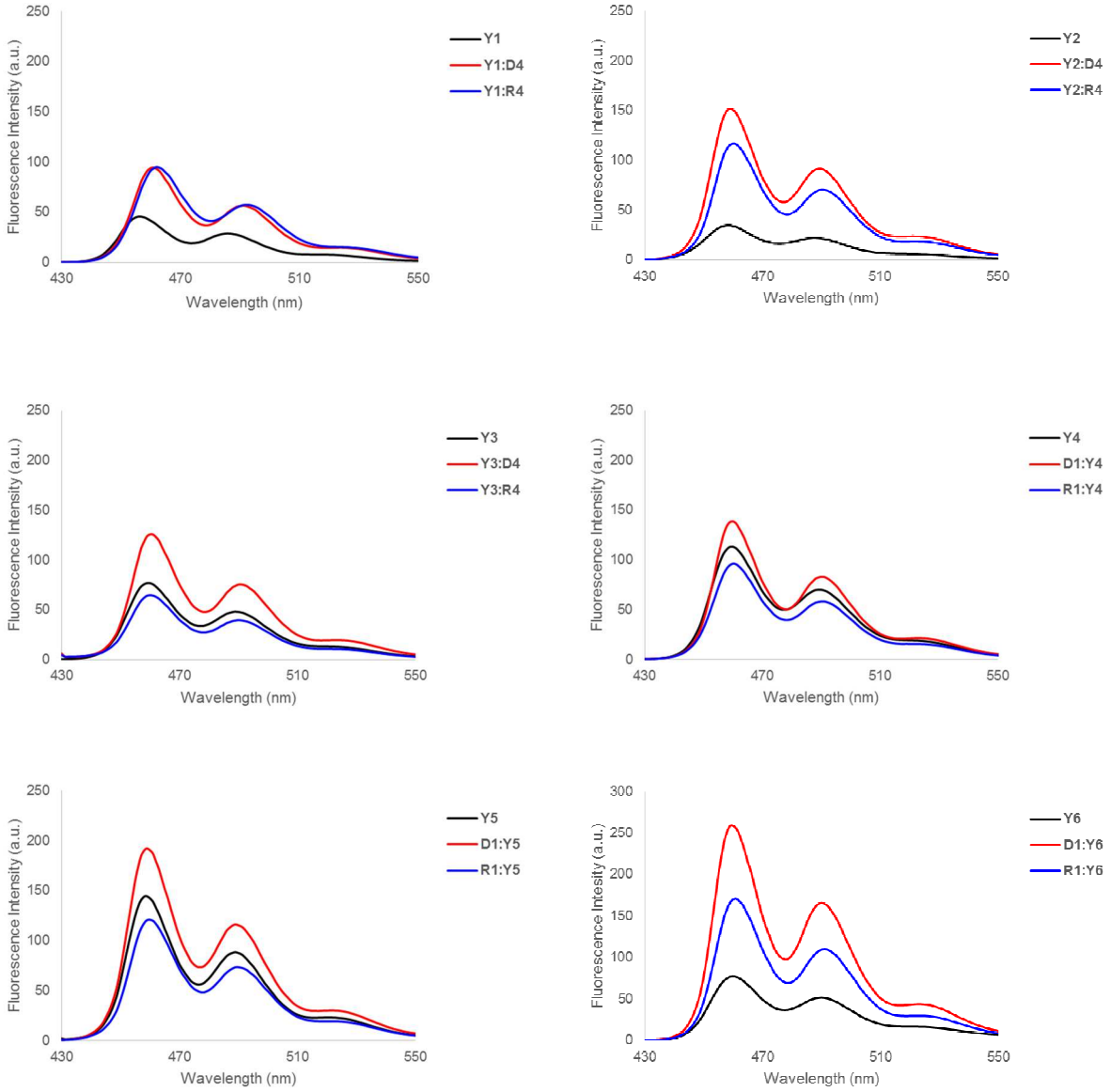


Figure 3-S4: Steady-state fluorescence emission spectra of ONs **Y1-Y6** and the corresponding duplexes with cDNA/cRNA targets. Spectra were recorded at $T = 5$ °C using $\lambda_{ex} = 420$ nm and each strand at 1.0 μM concentration in T_m buffer. Note that different Y-axis scales are used. No signal was observed above 550 nm.

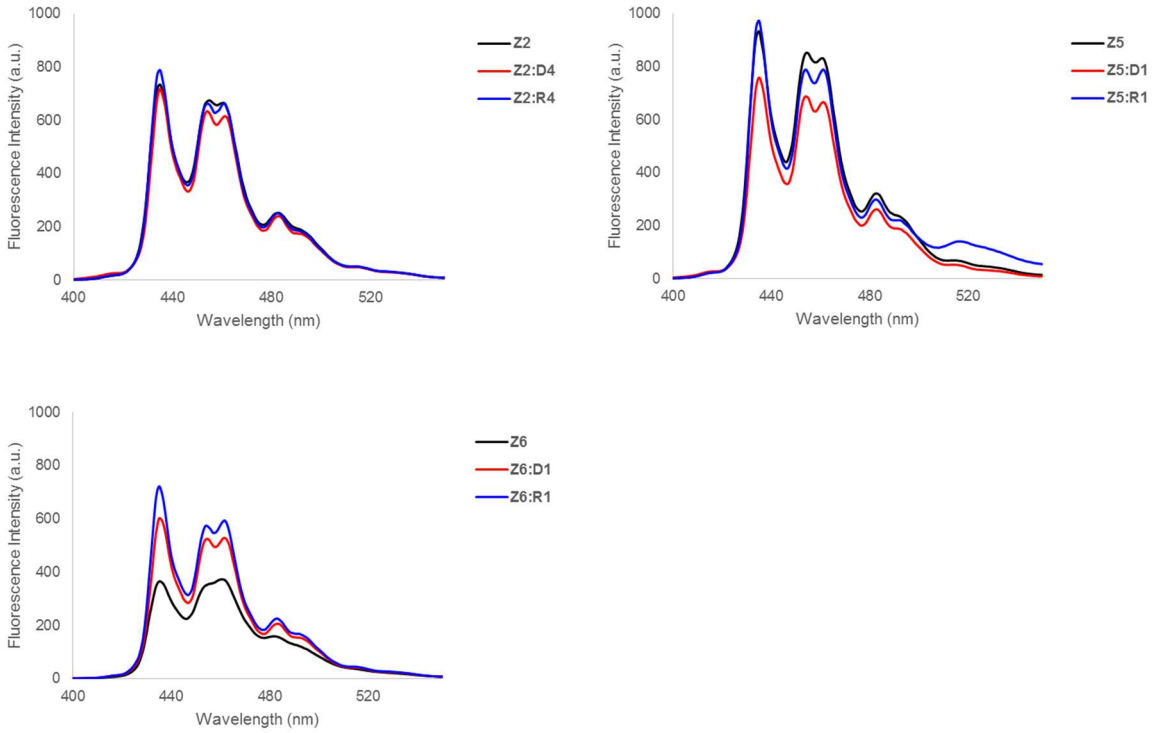


Figure 3-S5: Steady-state fluorescence emission spectra of select **Z**-modified ONs and the corresponding duplexes with cDNA/cRNA targets. Spectra were recorded at T = 10 °C using λ_{ex} = 310 nm and each strand at 1.0 μM concentration in T_m buffer.

Table 3-S5: Change in enthalpy upon duplex formation (ΔH) and change in enthalpy upon probe recognition of iso-sequential dsDNA target **D1**:**D4** (ΔH_{rec}).[a]

ON	ZP	Sequence	ΔH [$\Delta\Delta H$] (kJ/mol) upper ON vs cDNA	lower ON vs cDNA	probe duplex	ΔH_{rec} (kJ/mol)
X1 **X5**	**+4**	5′-G**X**G ATA TGC 3′-CAC TA**X** ACG	-290±14 [-4]	-377±11 [-91]	N/A	-
X1 **X4**	**+2**	5′-G**X**G ATA TGC 3′-CAC **X**AT ACG	-290±14 [-4]	-284±16 [+2]	-299±16 [-13]	+11
X2 **X5**	**+1**	5′-GTG A**X**A TGC 3′-CAC TA**X** ACG	**-389±13 [-103]**	**-377±11 [-91]**	**-262±4 [+24]**	**-218**
X2 **X4**	**-1**	5′-GTG A**X**A TGC 3′-CAC **X**AT ACG	-389±13 [-103]	-284±16 [+2]	-266±16 [+20]	-121
Y1 **Y5**	**+4**	5′-G**Y**G ATA TGC 3′-CAC TA**Y** ACG	-268±3 [+18]	-358±15 [-72]	-254±24 [+32]	-86
Y1 **Y4**	**+2**	5′-G**Y**G ATA TGC 3′-CAC **Y**AT ACG	-268±3 [+18]	-363±25 [-77]	-236±6 [+50]	-109
Y2 **Y5**	**+1**	5′-GTG A**Y**A TGC 3′-CAC TA**Y** ACG	**-356±26 [-70]**	**-358±15 [-72]**	**-286±7 [±0]**	**-142**
Y2 **Y4**	**-1**	5′-GTG A**Y**A TGC 3′-CAC **Y**AT ACG	-356±26 [-70]	-363±25 [-77]	-340±33 [-54]	-93
Z1 **Z5**	**+4**	5′-G**Z**G ATA TGC 3′-CAC TA**Z** ACG	-336±6 [-50]	-327±8 [-41]	-350±5 [-64]	-27
Z1 **Z4**	**+2**	5′-G**Z**G ATA TGC 3′-CAC **Z**AT ACG	-336±6 [-50]	-328±8 [-42]	-247±11 [+39]	-131
Z2 **Z5**	**+1**	5′-GTG A**Z**A TGC 3′-CAC TA**Z** ACG	**-335±11 [-49]**	**-327±8 [-41]**	**-270±7 [+16]**	**-106**
Z2 **Z4**	**-1**	5′-GTG A**Z**A TGC 3′-CAC **Z**AT ACG	-335±11 [-49]	-328±8 [-42]	-350±10 [-64]	-27

[a] $\Delta\Delta H$ is measured relative to ΔH for **D1**:**D4** = -286 kJ/mol. ΔH_{rec} = ΔH (upper strand vs cDNA) + ΔH (lower strand vs cDNA) - ΔH (probe duplex) - ΔH (dsDNA target). "±" denotes standard deviation. N/A = the lack of a clear lower base line prevented determination of this value.

Table 3-S6: Change in entropy at 293K upon duplex formation ($-T^{293}\Delta S$) and change in entropy upon probe recognition of iso-sequential dsDNA target **D1**:**D4** ($-T^{293}\Delta S_{rec}$).[a]

Duplex	Zipper	Sequence	$-T^{293}\Delta S$ [$\Delta(T^{293}\Delta S)$] (kJ/mol) upper strand vs cDNA	lower strand vs cDNA	probe duplex	$-T^{293}\Delta S_{rec}$ (kJ/mol)
X1 **X5**	**+4**	5′-G**X**G ATA TGC 3′-CAC TA**X** ACG	239±13 [-1]	313±10 [+73]	N/A	-
X1 **X4**	**+2**	5′-G**X**G ATA TGC 3′-CAC **X**AT ACG	239±13 [-1]	236±17 [-4]	255±15 [+15]	-20
X2 **X5**	**+1**	5′-GTG A**X**A TGC 3′-CAC TA**X** ACG	**324±11 [+84]**	**313±10 [+73]**	**218±3 [-22]**	**+179**
X2 **X4**	**-1**	5′-GTG A**X**A TGC 3′-CAC **X**AT ACG	324±11 [+84]	236±17 [-4]	211±15 [-29]	+109
Y1 **Y5**	**+4**	5′-G**Y**G ATA TGC 3′-CAC TA**Y** ACG	217±3 [-23]	289±13 [+49]	371±21 [+131]	-105
Y1 **Y4**	**+2**	5′-G**Y**G ATA TGC 3′-CAC **Y**AT ACG	217±3 [-23]	305±24 [+65]	187±6 [-53]	+95
Y2 **Y5**	**+1**	5′-GTG A**Y**A TGC 3′-CAC TA**Y** ACG	**287±23 [+47]**	**289±13 [+49]**	**237±11 [-3]**	**+99**
Y2 **Y4**	**-1**	5′-GTG A**Y**A TGC 3′-CAC **Y**AT ACG	287±23 [+47]	305±24 [+65]	261±17 [+21]	+91
Z1 **Z5**	**+4**	5′-G**Z**G ATA TGC 3′-CAC TA**Z** ACG	279±6 [+39]	261±7 [+21]	271±4 [+31]	+29
Z1 **Z4**	**+2**	5′-G**Z**G ATA TGC 3′-CAC **Z**AT ACG	279±6 [+39]	271±7 [+31]	192±10 [-48]	+118
Z2 **Z5**	**+1**	5′-GTG A**Z**A TGC 3′-CAC TA**Z** ACG	**267±10 [+27]**	**261±7 [+21]**	**216±7 [-24]**	**+72**
Z2 **Z4**	**-1**	5′-GTG A**Z**A TGC 3′-CAC **Z**AT ACG	267±10 [+27]	271±7 [+31]	270±8 [+30]	+28

[a] $\Delta(T^{293}\Delta S)$ is measured relative to $-T^{293}\Delta S$ for **D1**:**D4** = 240 kJ/mol. $-T^{293}\Delta S_{rec} = \Delta(T^{293}\Delta S)$ (upper ON vs cDNA) + $\Delta(T^{293}\Delta S)$ (lower ON vs cDNA) - $\Delta(T^{293}\Delta S)$ (probe duplex). "±" denotes standard deviation. N/A = the lack of a clear lower base line prevented determination of this value.

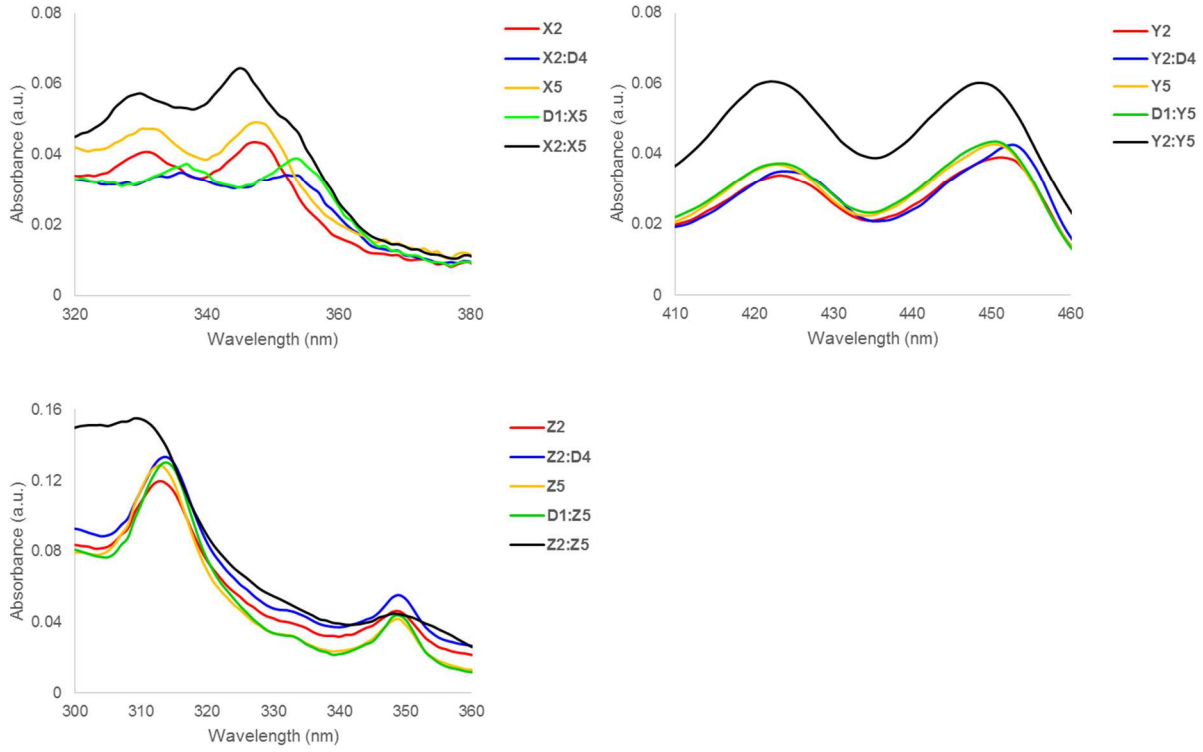


Figure 3-S6: Absorption spectra of representative Invaders, duplexes between probe strands and cDNA, and single-stranded probes (SSP). **X** and **Y** spectra were recorded at T = 5 °C whereas **Z** spectra were recorded at T = 10 °C using each strand at 1.0 μM concentration in T_m buffer.

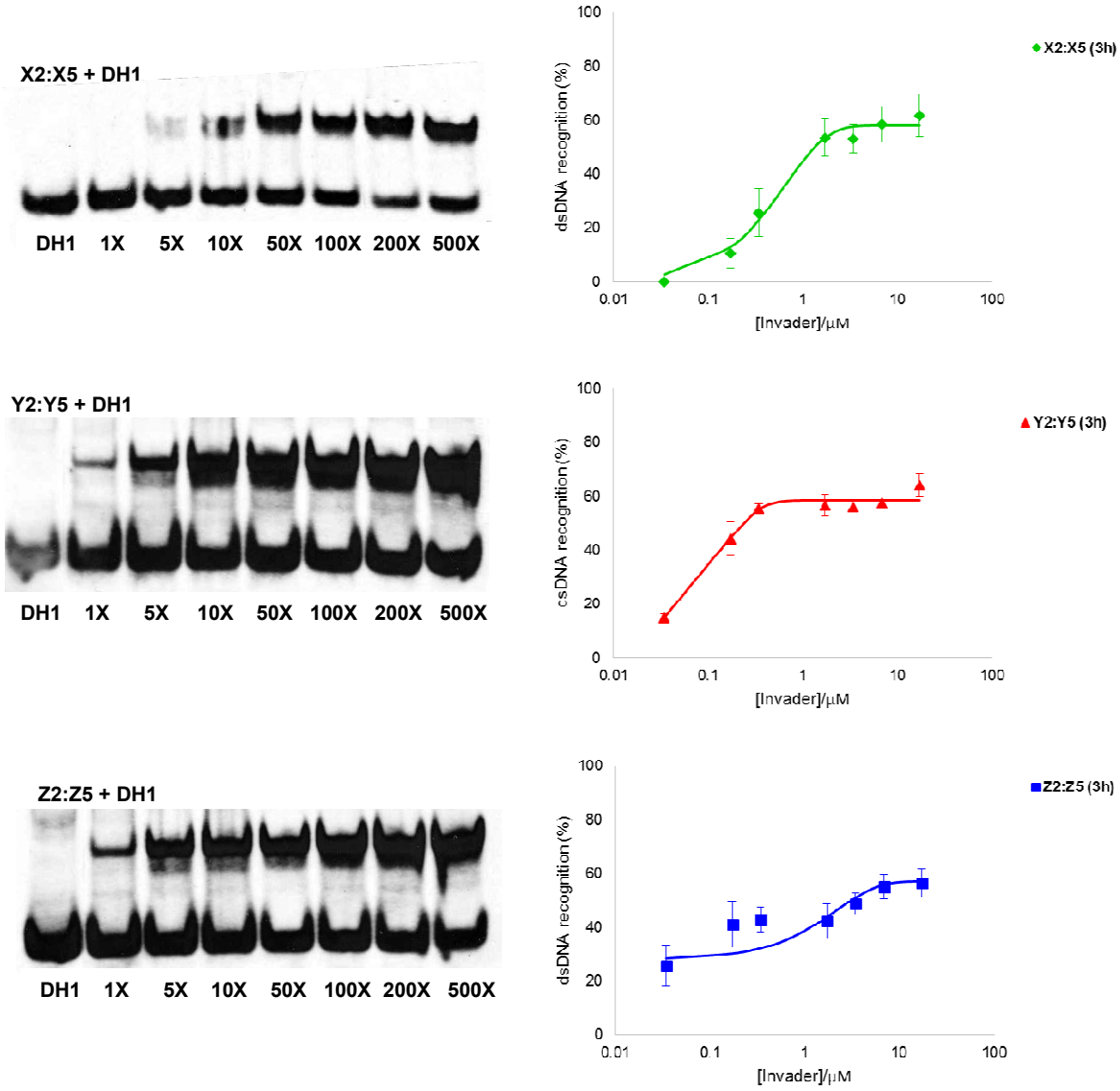


Figure 3-S7: Representative electrophoretograms and dose-response curves (average of at least three independent experiments, error bars represent standard deviation) for recognition of **DH1** using 1- to 500-fold excess of **X2**:**X5**, **Y2**:**Y5**, or **Z2**:**Z5**. Experimental conditions are as described in Figure 3-4 except for a shorter incubation time (3h).

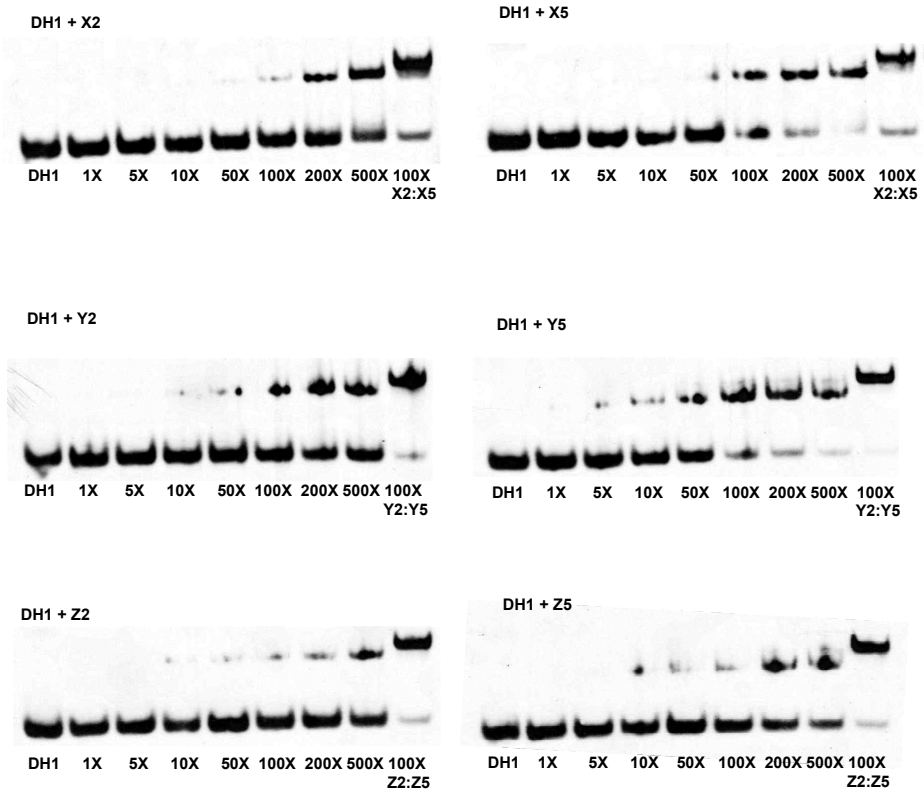


Figure 3-S8: Recognition of DNA hairpin **DH1** using single-stranded ONs. Representative gel electropheretograms illustrating recognition of **DH1** using 1- to 500-fold excess of **X2**, **X5**, **Y2**, **Y5**, **Z2**, or **Z5**. The right-most lane depicts recognition of **DH1** using 100-fold excess of **B2**:**B5**. For experimental conditions, see Figure 3-4.

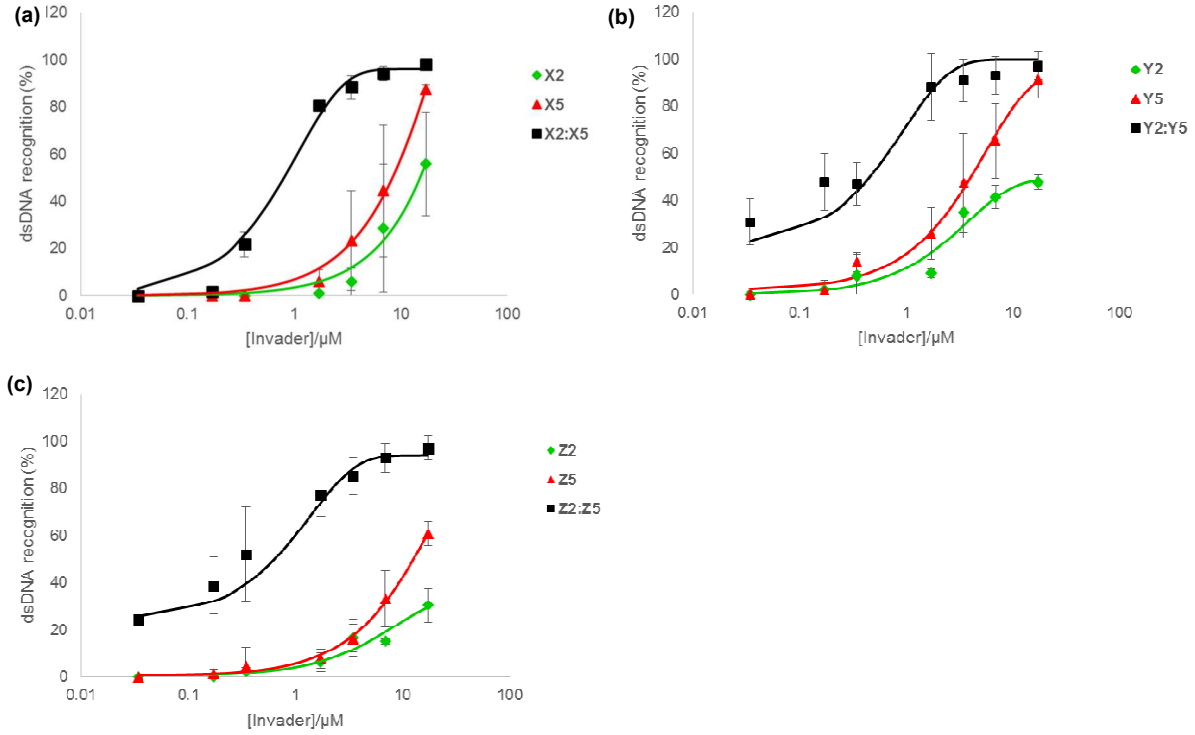


Figure 3-S9: Dose-response curves (average of three independent experiments, error bars represent standard deviation) for recognition of **DH1** using single-stranded ONs compared to double-stranded Invader probes. (a) **X2**: C_{50} = 15.1 μM; **X5**: C_{50} = 8.3 μM; (b) **Y2**: C_{50} >17.2 μM; **Y5**: C_{50} = 4.0 μM; (c) **Z2**: C_{50} = >17.2 μM; **Z5**: C_{50} = 12.5 μM. For experimental conditions, see Figure 3-4.

3.6 References

1) Besch, R.; Giovannangeli, C.; Degitz, K. *Curr. Drug Targets* **2004**, 5, 691 – 703.

2) Rogers, F. A.; Lloyd, J. A.; Glazer, P. M. *Curr. Med. Chem.*: *Anti-Cancer Agents* **2005**, 5, 319 – 326.

3) Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. J. *Mol. BioSyst.* **2006**, 2, 551 – 560.

4) Nielsen. P. E. *Chem*. *Biodiv*. **2010**, 7, 786 – 804.

5) Mukherjee, A.; Vasquez, K. M. *Biochimie* **2011**, 93, 1197 – 1208.

6) Aiba, Y.; Sumaoka, J.; Komiyama, M. *Chem*. *Soc*. *Rev*. **2011**, 40, 5657 – 5668.

7) Vaijayanthi, T.; Bando, T.; Pandian, G. N.; Sugiyama, H. *ChemBioChem* **2012**, 13, 2170 – 2185.

8) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, 36, 5123 – 5138.

9) Kaihatsu, K.; Janowski, B. A.; Corey, D. R. *Chem. Biol*. **2004**, 11, 749 – 758.

10) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol*. **2003**, 13, 284 – 299.

11) Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem*. **2013**, 21, 6101 – 6114.

12) Gaj, T.; Gersbach, C. A.; Barbas III, C. F. *Trends Biotechnol*. **2013**, 31, 397 – 405.

13) Tse, W. C.; Boger, D. L. *Chem. Biol*. **2004**, 11, 1607 – 1617.

14) Hamilton, P. L.; Arya, D. P. *Nat. Prod. Rep*. **2012**, 29, 134 – 143.

15) Bentin, T.; Larsen, H. J.; Nielsen, P. E. *Biochemistry* **2003**, 42, 13987 – 13995.

16) Kaihatsu, K.; Shah, R. H.; Zhao, X.; Corey, D. R. *Biochemistry* **2003**, 42, 13996 – 14003.

17) Moreno, P. M. D.; Geny, S.; Pabon, Y. V.; Bergquist, H.; Zaghloul, E. M.; Rocha, C. S. J.; Oprea, I. I.; Bestas, B.; Andaloussi, S. E. L.; Jørgensen, P. T.; Pedersen, E. B.; Lundin, K. E.; Zain, R.; Wengel, J.; Smith, C. I. E. *Nucleic Acids Res*. **2013**, 41, 3257 – 3273.

18) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc*. **1990**, 112, 2435 – 2437.

19) Filichev, V. V.; Nielsen, C. M.; Bomholt, N.; Jessen, C. H.; Pedersen, E. B. *Angew. Chem. Int. Ed*. **2006**, 45, 5311 – 5315.

20) Rusling, D. A.; Powers, V. E. C.; Ranasinghe, R. T.; Wang, Y.; Osborne, S. D.; Brown, T.; Fox, K. *Nucleic Acids Res*. **2005**, 33, 3025 – 3032.

21) Hari, Y.; Obika, S.; Imanishi, T. *Eur*. *J. Org. Chem*. **2012**, 2875 – 2887.

22) Kutyavin, I. V.; Rhinehart, R. L.; Lukhtanov, E. A.; Gorn, V. V.; Meyer Jr., R. B.; Gamper Jr., H. B. *Biochemistry* **1996**, 35, 11170 – 11176.

23) Filichev, V. V.; Vester, B.; Hansen, L. H.; Pedersen, E. B. *Nucleic Acids Res*. **2005**, 33, 7129 – 7137.

24) Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc*. *Natl*. *Acad*. *Sci*. *U.S.A*. **1999**, 96, 11804 – 11808.

25) Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corradini, R.; Sforza, S.; Komiyama, M. *Nucleic Acids Res*. **2008**, 36, 1464 – 1471.

26) Demidov, V. V.; Protozanova, E.; Izvolsky, K. I.; Price, C.; Nielsen, P. E.; Frank-Kamenetskii, M. D. *Proc*. *Natl*. *Acad*. *Sci*. *U.S.A*. **2002**, 99, 5953 – 5958.

27) Sumaoka, J.; Komiyama, M. *Chem*. *Lett*. **2014**, 43, 1581 – 1583.

28) Rapireddy, S.; Bahal, R.; Ly, D. H. *Biochemistry* **2011**, 50, 3913 – 3918.

29) Bahal, R.; Sahu, B.; Rapireddy, S.; Lee, C-M.; Ly, D. H. *ChemBioChem* **2012**, 13, 56 – 60.

30) Sander, J. D.; Joung, J. K. *Nat*. *Biotechnol*. **2014**, 32, 347 – 355.

31) Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org*. *Biomol*. *Chem*. **2010**, 8, 2028 – 2036.

32) Sau, S. P.; Madsen, A. S.; Podbevsek, P.; Andersen, N. K.; Kumar, T. S.; Andersen, S.; Rathje, R. L.; Anderson, B. A.; Guenther, D. C.; Karmakar, S.; Kumar, P.; Plavec, J.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem*. **2013**, 78, 9560 – 9570.

33) Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *J. Org. Chem*. **2013**, 78, 12040 – 12048.

34) Denn, B.; Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *Chem. Commun*. **2013**, 49, 9851 – 9853.

35) Didion, B. A.; Karmakar, S.; Guenther, D. C.; Sau, S.; Verstegen, J. P.; Hrdlicka, P. J. *ChemBioChem* **2013**, 14, 1534 – 1538.

36) Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J. *Org. Biomol. Chem*. **2014**, 12, 7758 – 7773.

37) Crothers, D. M. *Biopolymers* **1968**, 6, 575 – 584.

38) Kumar, T. S.; Madsen, A. S.; Østergaard, M. E.; Sau, S. P.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem*. **2009**, 74, 1070 – 1081.

39) Andersen, N. K.; Anderson, B. A.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem*. **2013**, 78, 12690 – 12702.

40) Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem*. **2011**, 76, 7119 – 7131.

41) Matray, T. J.; Kool, E. T. Selective and stable DNA base pairing without hydrogen bonds. *J. Am. Chem. Soc*. **1998**, *1*20, 6191 – 6192.

42) Kalra, N.; Babu, B. R.; Parmar, V. S.; Wengel, J. *Org. Biomol. Chem*. **2004**, 2, 2885 – 2887.

43) McGee, D. P. C.; Vaughn-Settle, A.; Vargeese, C.; Zhai, Y. *J. Org. Chem*. **1996**, 61, 781 – 785.

44) Abdel-Magid, A. F.; Carson, K. G.; Harris, B. D.; Maryanoff, C. A.; Shah, R. D. *J. Org. Chem.* **1996**, 61, 3849 – 3862.

45) Yamana, K.; Iwase, R.; Furutani, S.; Tsuchida, H.; Zako, H.; Yamaoka, T.; Murakami. A. *Nucleic Acids Res*. **1999**, 27, 2387 – 2392.

46) Nakamura, M.; Fukunaga, Y.; Sasa, K.; Ohtoshi, Y.; Kanaori, K.; Hayashi, H.; Nakano, H.; Yamana, K. *Nucleic Acids Res*. **2005**, 33, 5887 – 5895.

47) Christensen, U. B.; Pedersen, E. B. *Nucleic Acids Res*. **2002**, 30, 4918 – 4925.

48) Bryld, T.; Højland, T.; Wengel, J. *Chem*. *Commun*. **2004**, 1064 – 1065.

49) Marin, V.; Hansen, H. F.; Koch, T. R.; Armitage, B. A. *J*. *Biomol*. *Struct*. *Dyn*. **2004**, 21, 841 – 8520.

50) Dougherty, G.; Pilbrow, J. R. *Int. J. Biochem*. **1984**, 16, 1179 – 1192.

51) Asanuma, H.; Fujii, T.; Kato, T.; Kashida, H. *J. Photochem. Photobiol., C*. **2012**, 13, 124 – 135.

52) Zinger, D.; Geacintov, N. E. *Photochem*. *Photobiol*. **1988**, *47*, 181-188.

53) Manoharan, M.; Tivel, K. L.; Zhao, M.; Nafisi, K.; Netzel, T. L. *J*. *Phys*. *Chem*. **1995**, *99*, 17461-17472.

54) Seo, Y. J.; Ryu, J. H.; Kim, B. H. *Org*. *Lett*. **2005**, *7*, 4931-4933.

55) Østergaard, M. E.; Kumar, P.; Baral, B.; Guenther, D. C.; Anderson, B. A.; Ytreberg, F. M.; Deobald, L.; Paszczynski, A. J.; Sharma, P. K.; Hrdlicka, P. J. *Chem*. *Eur*. *J*. **2011**, *17*, 3157-3165.

56) Wilson, J. N.; Cho, Y.; Tan, S.; Cuppoletti, A.; Kool, E. T. *ChemBioChem* **2008**, *9*, 279-285.

57) Bag, S. S.; Saito, Y.; Hanawa, K.; Kodate, S.; Suzuka, I.; Saito, I. *Bioorg*. *Med*. *Chem*. *Lett*. **2006**, 16, 6338 – 6341.

58) Asanuma, H.; Akahane, M.; Kondo, N.; Osawa, T.; Kato, T.; Kashida, H. *Chem*. *Sci*., **2012**, 3, 3165 – 3169.

59) Gupta, P. Langkjær N.; Wengel, J. *Bioconjugate Chem*. **2010**, 21, 513 – 520.

60) Saswata Karmakar, Ph.D-dissertation, University of Idaho, January 2014.

61) Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, 13, 515 – 537.

62) DNA hairpins were chosen as targets since alternatives, such as PCR fragments, cannot be readily used in electrophoretic mobility shift assays due to insufficient mobility differences between recognition complexes and unreacted targets (results not shown).

63) Asseline, U.; Cheng, E. *Tetrahedron Lett*. **2001**, 42, 9005 – 9010.

64) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J*. *Org*. *Chem*. **1997**, 62, 7512 – 7515.

65) Dale, T. J.; Rebek Jr., J. *J. Am. Chem. Soc.* **2006**, 128, 4500 – 4501.

66) Dioubankova, N. N.; Malakhov, A. D.; Stetsenko, D. A.; Gait, M. J.; Volynsky, P. E.; Efremov, R. G.; Korshun, V. A. *ChemBioChem* **2003**, 4, 841 – 847.

67) Astakhova, I. V.; Korshun, V. A.; Jahn, K.; Kjems, J.; Wengel, J. *Bioconjugate Chem*. **2008**, 19, 1995 – 2007.

CHAPTER 4: Double-stranded DNA Recognition by Enzymatically Stable Phosphorothioate Invader Probes

Brooke A. Anderson, Saswata Karmakar and Patrick J. Hrdlicka

Department of Chemistry, University of Idaho, Moscow, Idaho 83844, United States

*Manuscript in preparation

Abstract

Oligonucleotides capable of interrupting the flow of genetic information continue to attract considerable attention due to their successful use within antisense, antigene, and other gene-targeting strategies. Oligonucleotides modified with 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-aminouridine and 2′-*O*-(pyren-1-yl)methyluridine display greatly increased affinity toward complementary DNA and improved binding specificity as a result of intercalation of the pyrene residue. Double-stranded probes with +1 interstrand zipper arrangements of these monomers display efficient recognition of DNA hairpin targets in cell-free environments due to the energetic nature of the zipper duplex and the high stability of the probe-target duplexes. Applications in cellular contexts will require these probes to be stable to endogenous nucleases. Oligonucleotides with phosphorothioate backbones are known to display enhanced enzymatic stability. Here we describe the synthesis of a series of phosphorothioate (PS) oligonucleotides containing modifications of 2′-N-(pyren-1-yl)methyl-2′-N-methyl-2′-aminouridine and 2′-O-(pyren-1-yl)methyluridine. The results demonstrate that alterations in the backbone chemistry of Invader-modified oligonucleotides results in 1) increased affinity toward mixed-sequence

DNA targets, 2) efficient recognition of DNA hairpin targets and (3) greatly enhanced stability toward nucleases. These characteristics render PS-Invaders based on 2′-N-(pyren-1-yl)methyl-2′-N-methyl-2′-aminouridine and 2′-O-(pyren-1-yl)methyluridine monomers of considerable interest for DNA targeting applications in cellular contexts.

4.1 Introduction

Development of efficient strategies for sequence-unrestricted recognition of double-stranded DNA (dsDNA) is a long-standing goal in biological chemistry, which is inspired by the promise for tools that can detect, regulate, and modify genes.[1-6] The demand for effective dsDNA targeting methodologies has resulted in a well-established toolbox of chemically modified ligands capable of dsDNA recognition, including triplex-forming oligonucleotides (TFOs)[7], peptide nucleic acids (PNAs)[8], and minor-groove binding polyamides.[9] However, these pioneering dsDNA targeting approaches display limitations that have restricted their widespread use. More specifically, TFOs and standard PNAs require polypurine target sites, which limits target availability. Furthermore, most PNA-based dsDNA invasion approaches require non-physiological saline conditions, unmodified TFOs generally require non-physiological pH conditions, and polyamides can typically only recognize short target sequences, which renders unique targeting problematic. Thus, probes that efficiently recognize mixed-sequence DNA under physiologically relevant conditions remain tenuous.

The drawbacks of the current dsDNA-recognition methodologies have spurred development of alternative strategies for mixed-sequence targeting of dsDNA such as pseudocomplementary DNA (pcDNA)[10], pcPNA[11], bis-PNAs[12], γ-PNAs[13], TFOs with engineered nucleobases[14],

engineered proteins[15], among other approaches[16]. Nonetheless, while these strategies are effective, each probe comes with its own set of limitations and there remains a need for probes that enable rapid, potent, and specific targeting of mixed-sequence dsDNA under physically relevant conditions, which are inexpensive, compatible with delivery agents, and enzymatically stable under physiological environments.

In 2005, Hrdlicka and coworkers introduced Invader probes as an alternative approach toward mixed-sequence recognition of dsDNA.[17] These probes are energetically activated duplexes modified with intercalator-functionalized nucleotides which are capable of recognizing dsDNA (for an illustration, see Figure 4-1). Probe activation arises from the intercalator organization inside the duplex, termed the +1 zipper arrangement, which results in duplex destabilization. Presumably, probe activation is a result of the intercalators being forced into the same region of the duplex which is in violation of the 'nearest neighbor exclusion principle' which states that intercalators, at most, bind to every second base pair of a DNA duplex or local duplex unwinding and the formation of an 'energetic hotspot' can take place.[17-20] On the other hand, the two strands that constitute the Invader probe display exceptional affinity toward complementary DNA (cDNA) as they are able to site-specifically position intercalators into the duplex core resulting in strongly stabilizing stacking interactions with neighboring base pairs. We have used the difference in thermostabilities between Invader probes and probe-target duplexes to drive recognition of short iso-sequential mixed-sequence DNA hairpins[19-22] and chromosomal DNA targets.[23]

The original Invader probes were based on activated duplexes modified with 2′-N-(pyren-1-yl)-2′-amino-α-L-LNA (Locked Nucleic Acid).[17,19] Due to the challenging synthesis of 2′-amino-α-L-LNA monomers, we recently identified 2′-N-(pyren-1-yl)methyl-2′-N-methyl-2′-

aminouridine and 2′-O-(pyren-1-yl)methyluridine monomers **X** and **Y** as synthetically more readily accessible structural and functional mimics of the original Invader probes.[24] The resulting Invader probes display similar dsDNA-recognition efficiency[21] but were significantly easier to synthesize.[25,26] Identification of simpler scaffolds has enabled us to perform systematic structure-property relationship studies with the goal of optimizing the recognition efficiency of Invader duplexes, such as the influence of intercalator density, intercalator-linker length and composition, and the number and location of energetic hotspots.[19,24]

In the present study, we set out to develop enzymatically stable Invader probes for recognition of dsDNA targets by synthesizing **X**- and **Y**-modified Invaders with a phosphorothioate backbone (PS-Invaders). Phosphorothioate groups are known to alter the susceptibility of nucleic acids to nucleases without major influence on the nucleic acid conformation.[27] We hypothesized that the enzymatic stability of the phosphorothioate backbone modification[28] would confer protection from nucleases for Invader probes and facilitate their use in mammalian cell culture studies. However, PS-modifications are known to decrease the thermodynamics of duplexes and therefore it was necessary to determine the impact PS-modifications have on dsDNA recognition efficiency of Invaders. These PS-Invaders were characterized by means of thermal denaturation experiments, UV-vis absorption and fluorescence spectroscopy, DNA recognition experiments, and enzymatic stability experiments.

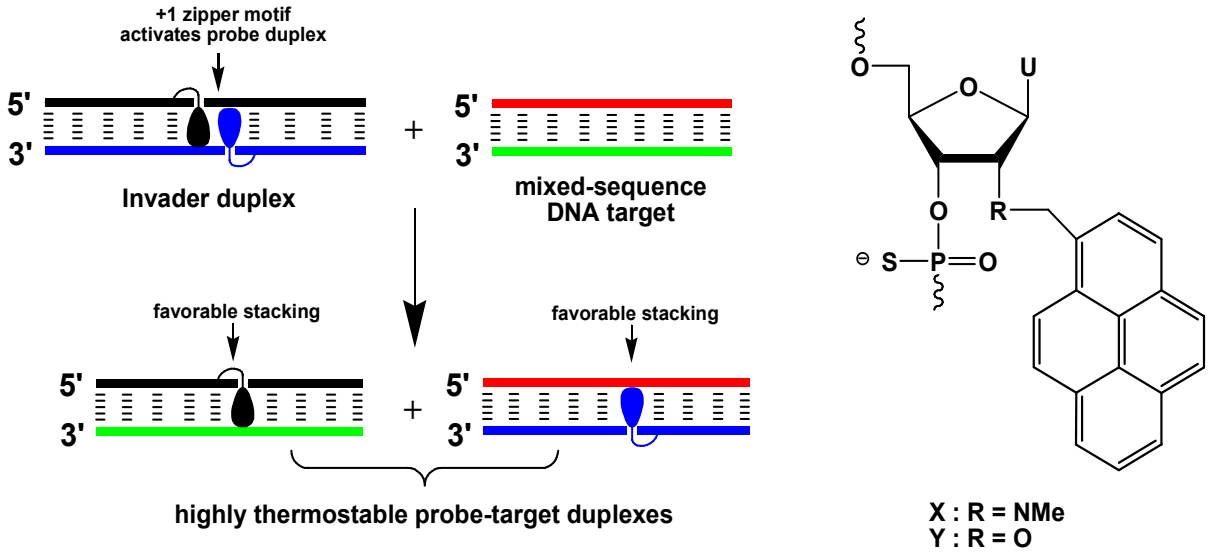


Figure 4-1: Illustration of the Invader approach for recognition of mixed-sequence DNA and structures of monomers used herein. Droplets denote intercalating pyrene moieties.

4.2 Results and Discussion

4.2.1 Synthesis of Modified ONs and Experimental Design

The corresponding phosphoramidites of monomers **X** and **Y** were synthesized as previously described[22,24] and incorporated into phosphorothioate DNA (PS-DNA) strands via machine-assisted DNA synthesis. The following hand-coupling conditions were used for incorporation of monomers **X** and **Y** (coupling time; activator; coupling yield): **X** (15 min; 5-[3,5-bis(trifluoromethyl)phenyl]-1*H*-tetrazole; ~99%) and **Y** (15 min; 4,5-dicyanoimidazole; ~99%). Sulfurization was realized using 3-((*N*,*N*-dimethylaminomethylidene)amino)-3*H*-1,2,4-dithiazole-5-thione. The identity and purity of all modified PS-DNA was established via MALDI-TOF (Table 4-S1) and ion-pair reverse phase HPLC (>80% purity), respectively.

PS-DNA strands modified with monomers **X** and **Y** were studied in the same 9-mer and 13-mer sequence contexts that were previously used to study **X**- and **Y**-modified phosphodiester DNA (PO-DNA).[24,29] PS-DNA strands with a single incorporation of monomer **X** or **Y** in the 5′-G**B**G ATA TGC context are denoted **X1** and **Y1**. Similar conventions apply for the **B2**-**B6** series of 9-mer PS-DNA and the **B7**-**B14** series of 13-mer PS-DNA (Table 4-1).

4.2.2 Thermal Denaturation Properties of X-/Y-modified PS-DNA

Thermal denaturation temperatures (T_m's) of duplexes between **X**- or **Y**-modified PS-DNA and complementary DNA or RNA (cDNA/cRNA) were determined in medium salt phosphate buffer ([Na^+] = 110 mM). The resulting denaturation curves display monophasic sigmoidal transitions unless otherwise noted. (Figure 4-S1 in the Supporting Information).

Incorporation of **X** or **Y** monomers into PS-DNA strands results in markedly increased stability of duplexes with cDNA (ΔT_m per modification between +2.0 and +13.0 °C, Table 4-1), yet significantly decreased stability with cRNA (ΔT_m per modification between -8.0 to +3.0 °C, Table 4-1) relative to duplexes between PS-DNA and cDNA/cRNA. **X**-modified PS-DNA strands display slightly higher affinity toward cDNA than **Y**-modified strands. PS-DNA strands in which the pyrene-functionalized monomer is flanked by a 3′-purine results in greater duplex stabilization than strands with 3′-flanking pyrimidines (e.g., compare ΔT_m's for **B1** and **B4**, Table 4-1). This sequence-specific trend, along with the observed prominent DNA selectivity (Table 4-S2), is characteristic of oligonucleotides (ONs) modified with intercalating pyrene moieties.[30,31] All of these trends closely mirror our observations with **X**-/**Y**-modified PO-DNA strands[22,24,29], except that the relative stabilizing effects of the **X** and **Y** monomers generally are slightly lower when incorporated in PS-DNA strands (i.e., ΔT_m per modification is ~2 °C

lower). The absolute T_m's of the duplexes between with **X**-/**Y**-modified PS-DNA and cDNA are ~12 °C (**B1**-**B6**) and ~10 °C (**B7**-**B14**) lower than for the corresponding duplexes with **X**-/**Y**-modified PO-DNA.[24,29] Presumably, the decrease is largely due to the well-known destabilizing nature of the phosphorothioate backbone (typically 1.0-1.5 °C decrease in T_m per phosphorothioate modification).[27,32] For a discussion of binding specificities of **X**-/**Y**-modified PS-DNA, the Reader is referred to the Supporting Information (Tables 4-S2 and 4-S3).

Table 4-1: Change in thermal denaturation temperature (ΔT_m) of duplexes between **X**-/**Y**-modified PS-DNA **B1**-**B14** and cDNA/cRNA relative to duplexes between unmodified PS-DNA and cDNA/cRNA.[a]

			ΔT_m (°C)			
			+cDNA		+cRNA	
ON	PS-DNA sequence	**B =**	**X**	**Y**	**X**	**Y**
B1	5′-G**B**G ATA TGC		+5.0	+4.5	-4.0	-4.0
B2	5′-GTG A**B**A TGC		+13.0	+12.0	±0.0	±0.0
B3	5′-GTG ATA **B**GC		+8.0	+5.5	-4.0	-4.0
B4	3′-CAC **B**AT ACG		+4.0	+2.0	-5.0	-5.5
B5	3′-CAC TA**B** ACG		+12.5	+11.0	+3.0	+1.5
B6	3′-CAC **B**A**B** ACG		+11.0	+10.5	<-8.0	-6.5
B7	5′-GG**B** ATA TAT AGG C		+6.0	+5.5	-	-
B8	3′-CCA **B**AT ATA TCC G		+9.5	+9.0	-	-
B9	5′-GG**B** A**B**A TAT AGG C		+12.0	+12.5	-	-
B10	3′-CCA **B**A **B**ATA TCC G		+16.0	+16.0	-	-
B11	5′- GGT A**B**A **B**AT AGG C		+16.5	+16.0	-	-
B12	3′- CCA TA**B** A**B**A TCC G		+18.0	+17.5	-	-
B13	5′-GG**B** ATA TA**B** AGG C		+15.5	+13.5	-	-
B14	3′-CCA **B**AT ATA **B**CC G		+18.0	+16.0	-	-

[a] ΔT_m = change in T_m relative to duplexes between unmodified PS-DNA and cDNA/cRNA. Thus, **B1**-**B3** are measured relative to 5′-[PS-DNA]-GTGATATGC (T_m = 19.5 °C and 18.0 °C with cDNA and cRNA, respectively); **B4**-**B6** are measured relative to 3′-[PS-DNA]-CACTATACG (T_m = 19.5 °C and 18.0 °C with cDNA and cRNA, respectively); **B7**/**B9**/**B11**/**B13** are measured relative to 5′-[PS-DNA]-GGTATATATAGGC (T_m = 29.0 °C with cDNA) and **B8**/**B10**/**B12**/**B14** are measured relative to 3′-[PS-DNA]- CCATATATATCCG (T_m = 26.5 °C with cDNA). T_m's are determined as the maximum of the first derivative of melting curves (A_{260} vs T) recorded in medium salt phosphate buffer ([Na^+] = 110 mM, [Cl^-] = 100 mM, pH 7.0 (NaH_2PO_4/Na_2HPO_4)), using 1.0 μM of each strand. Reported T_m's are averages of at least two measurements within 1.0 °C; A = adenin-9-yl DNA monomer, C = cytosin-1-yl DNA monomer, G = guanin-9-yl DNA monomer, T = thymin-1-yl DNA monomer. For structures of monomers **X** and **Y**, see Figure 4-1. '-' denotes not determined

4.2.3 Photophysical Properties of X-/Y-modified PS-DNA

UV-Vis absorption and steady-state fluorescence emission spectra of **X**- and **Y**-modified 9-mer ONs were recorded in the absence or presence of cDNA/cRNA to further ascertain the intercalative binding mode of the attached pyrene. Indeed, hybridization of the modified PS-DNA with cDNA results in prominent hypochromic and bathochromic shifts of the pyrene absorption maxima ($\Delta\lambda_{max}$ = 0-5 nm, Figures 4-2, 4-S2 and 4-S3; Table 4-S5), which is indicative of ground-state electronic interactions between pyrenes and nucleobases[33] and intercalation. Slightly smaller bathochromic shifts are observed upon hybridization with cRNA, most likely since intercalation into the more compressed and RNA-like duplexes is less favorable.[34] Slightly greater bathochromic shifts are generally observed for **X**-modified PS-DNA than for **Y**-modified PS-DNA, which also was observed with the corresponding **X**-/**Y**-modified PO-DNA.[24] We speculate that this is due to particularly efficient intercalation of the pyrene moiety of monomer **X**.

Steady-state fluorescence emission spectra (λ_{ex} = 350 nm, T = 10 °C) of these duplexes display two vibronic bands at λ_{em} = 381±1 nm and 400±1 nm, respectively, as well as a small shoulder at ~420 nm (Figures 4-2 and 4-S4). Hybridization with cDNA/cRNA is accompanied by 1.4- to 4.6-fold decreases in fluorescence intensity, with greater decreases being observed upon

duplex formation with cDNA. Hybridization-induced decreases in fluorescence intensity are often observed for intercalating pyrene moieties as the fluorescence is quenched by neighboring nucleobases.[33b]

The above results strongly suggest that the pyrene moieties of the **X** and **Y** monomers maintain their intercalative binding modes when incorporated in PS-DNA strands, which is a prerequisite for their use as structural elements of Invader probes.

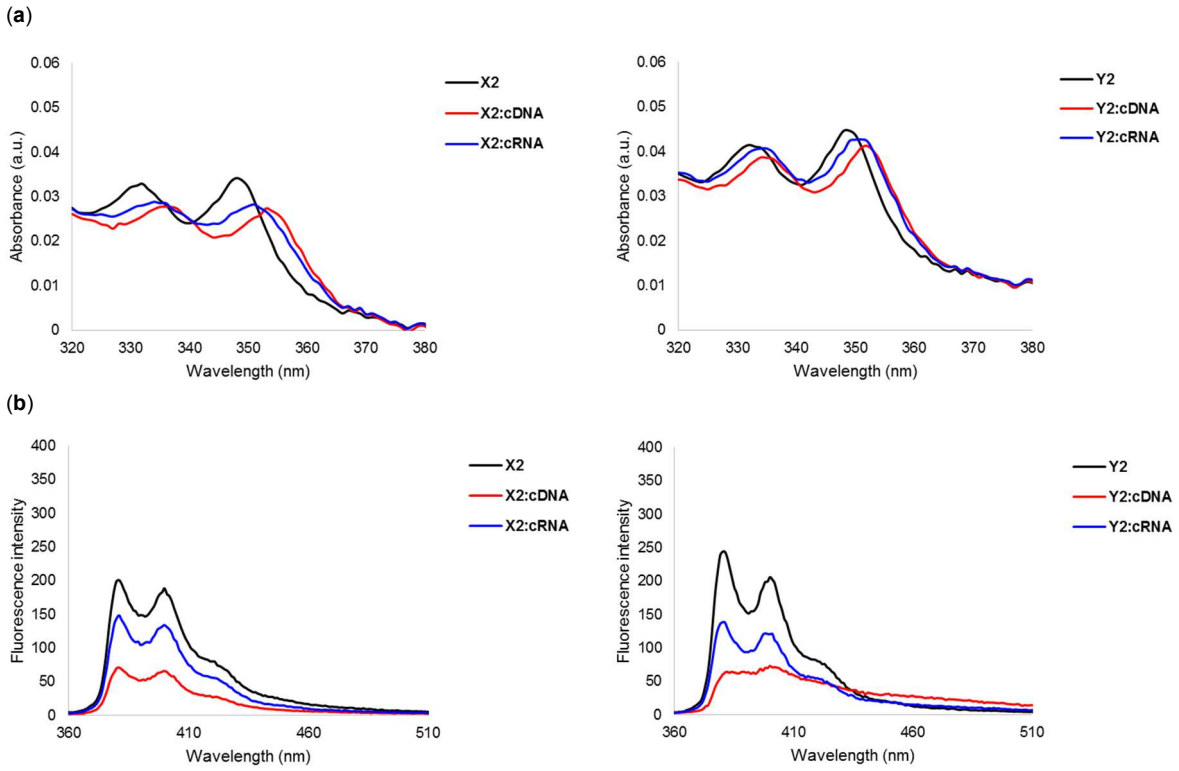


Figure 4-2: (a) Absorbance spectra of single-stranded **X2** and **Y2** and the corresponding duplexes with cDNA/cRNA. Spectra were recorded at T = 5 °C. (b) Steady-state fluorescence emission spectra of **X2** and **Y2** and the corresponding duplexes with cDNA/cRNA. Spectra were recorded at 10 °C using λ_{ex} = 350 nm. Each strand was used at 1.0 μM concentration in T_m buffer.

4.2.4 Thermal Denaturation Properties of PS-DNA Duplexes with Interstrand Zipper Arrangements of X or Y Monomers

Next, we set out to determine the T_m's of PS-DNA duplexes with different interstrand zipper arrangements of monomers **X** and **Y** (Table 4-2 - for a definition of the zipper nomenclature, see Experimental Section). In line with our observations in the corresponding PO-DNA series[21], PS-DNA duplexes with +4 or -1 interstrand zipper arrangements of **X** and **Y** monomers are very stable relative to reference duplexes due to stabilizing contributions from both monomers. In contrast, PS-DNA duplexes with +2 and +1 interstrand zippers are very unstable (no transitions above 10 °C were observed). Presumably, these interstrand zipper arrangements of **X** or **Y** monomers result in destabilization of PS-DNA duplexes due to violation of the previously discussed nearest neighbor exclusion principle. 13-mer PS-DNA duplexes with one +1 interstrand zipper of **X** or **Y** monomers or two sequential +1 interstrand zippers of **X** monomers are also very unstable (note ΔT_m for **X7**:**X8**, **X9**:**X10**, **X11**:**X12** and **Y7**:**Y8**, Table 4-2), while PS-DNA duplexes with two +1 interstrand zipper arrangements of **Y** monomers are slightly more stable, mirroring the trends observed for the corresponding PO-DNA series.[29] Accordingly, **X**-modified PS-DNA Invaders are therefore more strongly activated for dsDNA recognition than the corresponding **Y**-modified Invaders. In agreement with our main hypothesis, PS-DNA Invaders display much lower absolute T_m's than the corresponding PO-DNA Invaders (PS-DNA T_m = 12.5-26.0 °C, Table 4-2 vs PO-DNA T_m = 33.5-45.0 °C, reference 29), which is expected to translate into more facile probe dissociation during dsDNA recognition.

Table 4-2: T_m and ΔT_m values for PS-DNA duplexes with different interstrand zipper arrangements of **X** and **Y** monomers.[a]

ON	ZP	PS-DNA duplex **B** =	T_m (ΔT_m) /°C **X**	**Y**
B1 **B5**	+4	5′-G**B**G ATA TGC 3′-CAC TA**B** ACG	29.5 (>+19.5)	29.5 (>+19.5)
B1 **B4**	+2	5′-G**B**G ATA TGC 3′-CAC **B**AT ACG	<10.0	<10.0
B2 **B5**	+1	5′-GTG A**B**A TGC 3′-CAC TA**B** ACG	<10.0	<10.0
B2 **B4**	-1	5′-GTG A**B**A TGC 3′-CAC **B**AT ACG	22.0 (>+12.0)	19.0 (>+9.0)
B7 **B8**	+1	5′-GG**B** ATA TAT AGG C 3′-CCA **B**AT ATA TCC G	13.5 (-4.0)	12.5 (-5.0)
B9 **B10**	+1	5′-GG**B** A**B**A TAT AGG C 3′-CCA **B**A **B**ATA TCC G	17.5 (±0.0)	23.5 (+6.0)
B11 **B12**	+1	5′- GGT A**B**A **B**AT AGG C 3′- CCA TA**B** A**B**A TCC G	18.0 (+0.5)	26.0 (+8.5)
B13 **B14**	+1	5′-GG**B** ATA TA**B** AGG C 3′-CCA **B**AT ATA **B**CC G	24.5 (+7.0)	24.0 (+6.5)

[a] ΔT_m = change in T_m relative to reference duplexes 5′-[PS-DNA]-GTGATATGC:3′-[PS-DNA]-CACTATACG (T_m < 10.0 °C) and 5′-[PS-DNA]-GGTATATATAGGC:3′-[PS-DNA]-CCATATATATCCG (T_m = 17.5 °C). See Table 4-1 for experimental conditions. ZP = zipper.

4.2.5 Recognition of DNA Hairpins using Energetically Activated Probe Duplexes

Next, we studied the dsDNA-targeting properties of 9-mer and 13-mer **X**-/**Y**-modified PS-DNA Invaders. For this we used two different 3′-digoxigenin (DIG) labeled DNA hairpins (DH) as model dsDNA targets, i.e., **DH1** and **DH2**, which are comprised of 9-mer and 13-mer double-stranded stems, respectively, that are linked by a T_{10} loop (Figures 4-3 and 4-S5).

Surprisingly, room-temperature incubation of **DH1** with 9-mer PS-DNA Invaders **X2**:**X5** or **Y2**:**Y5** did not result in the formation of slower-migrating recognition complexes on non-

denaturing PAGE gels even at 500-fold probe excess (Figure 4-S5b). Control experiments, in which **DH1** was annealed with excess **X2**:**X5** or **Y2**:**Y5** followed by room temperature incubation, also failed to produce recognition complexes (Figure 4-S5c), suggesting that the recognition complexes are not sufficiently stable at these experimental conditions (note that T_m's for duplexes between **X2**/**X5**/**Y2**/**Y5** and cDNA only are 30.5-32.5 °C, Table 4-1).

Room temperature incubation of **DH2** with 13-mer PS-DNA Invaders containing one +1 interstrand zipper arrangement of **X** or **Y** monomers also does not result in dsDNA recognition (Figure 4-3b). In contrast, dose-dependent formation of a slower moving band is observed when PS-DNA Invaders with two energetic hotspots are used (Figure 4-3b). Thus, between 24-56 % dsDNA-recognition is observed when these Invaders are used at 200-fold excess. The **X**-modified PS-DNA Invaders recognize **DH2** slightly more efficiently than **Y**-modified Invaders (Figure 4-3c), which is consistent with the former being energetically more strongly activated for dsDNA recognition (i.e., larger stability differences between probe-target duplexes vs probe duplexes are observed). PS-DNA Invaders **X11**:**X12** and **X13**:**X14** recognize **DH2** most efficiently in this series, with C_{50} values of 2.6 μM and 2.9 μM, respectively. PS-DNA Invaders generally display slightly lower dsDNA recognition efficiencies than the corresponding PO-DNA Invaders (Figure 4-S6), which suggests that high probe-target stability is a more important factor for efficient dsDNA-recognition than Invader thermolability.

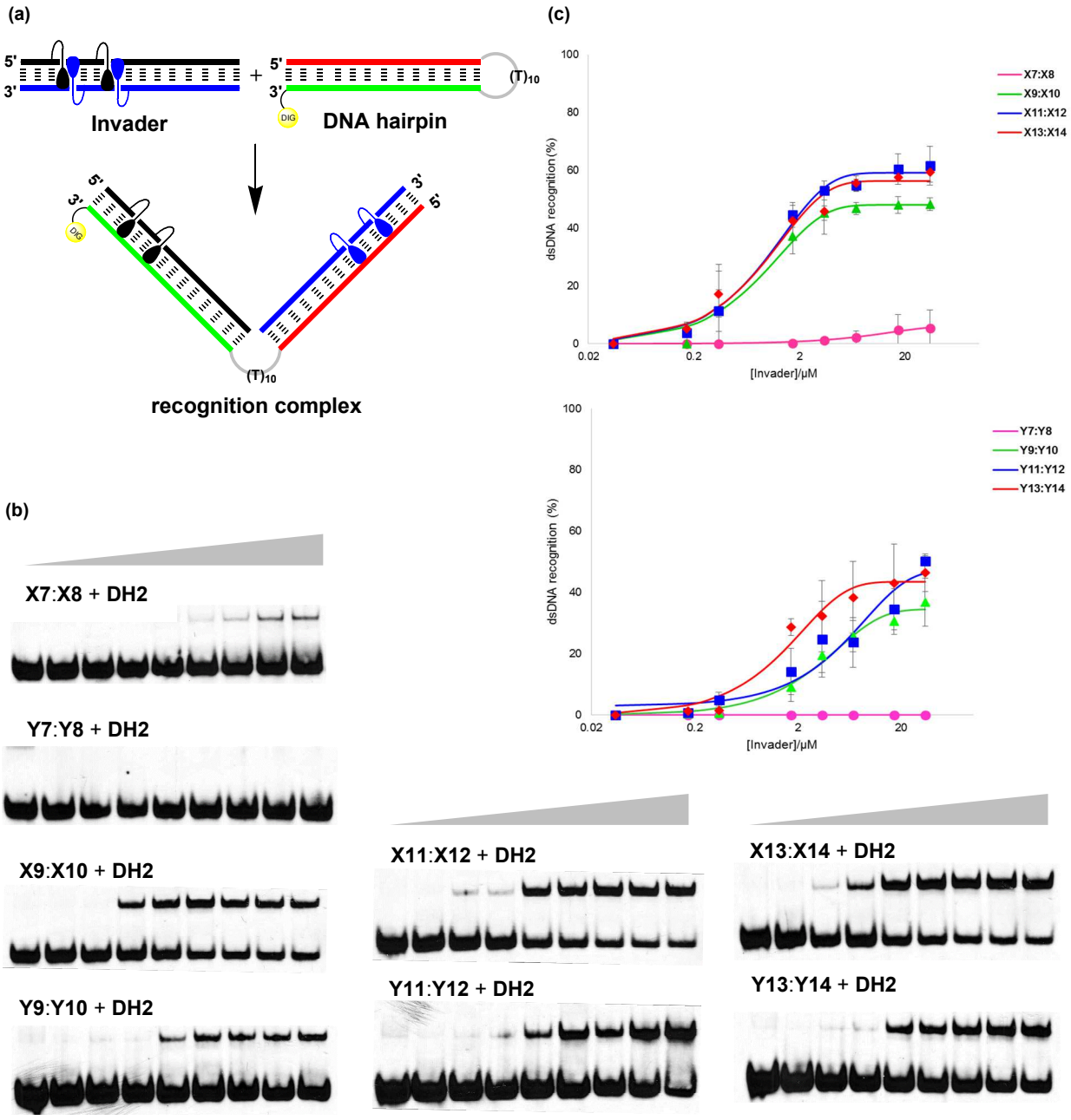


Figure 4-3: Recognition of dsDNA model target **DH2** using different Invader probes. (a) Illustration of recognition process; (b) representative electrophoretograms for recognition of **DH2** using 1-/5-/10-/50-/100-/200-/500-/1000-fold excess of **X** and **Y**-modified 13-mer invaders; (c) dose-response curves (average of at least three independent experiments, error bars represent standard deviation). Experimental conditions for electrophoretic mobility shift

assay: separately pre-annealed targets (34.4 nM) and probes (variable concentrations) were incubated for 12-16 h at room temperature in 1X HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 10% sucrose, 1.4 mM spermine tetrahydrochloride, pH 7.2) and then run on 16% non-denaturing PAGE (70V, 2.5h, ~4 °C) using 0.5x TBE as a running buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA); DIG: digoxigenin.

4.2.6 Enzymatic Stability of Individual Invader Strands

Finally, we studied the stability of individual Invader strands against snake venom phosphordiesterase (SVPDE - a 3′-exonuclease), as a significant proportion of PS-DNA Invaders may be dissociated at biologically relevant experimental temperatures. As expected, **X**- and **Y**-modified PS-DNA **X13** and **Y13**, as well as the PS-DNA reference strand, are completely stable against SVPDE (Figure 4-4). In contrast, the PO-DNA analog of **Y13** confers minimal protection relative to the unmodified PO-DNA reference sequence (95% degradation after ~50 min and ~15 min, respectively, Figure 4-4), while the PO-DNA analog of **X13** is moderately stable (~95% degradation after ~21 h). The latter observation suggests that the 2′-*N*-methyl-2′-amino-DNA skeleton of monomer **X** provides better protection against 3′-exonucleases than the RNA skeleton of monomer **Y**, rendering them as particularly interesting building blocks for Invader-based DNA-targeting applications (for DNAse I stability, see Figure 4-S7 in the Supporting Information).

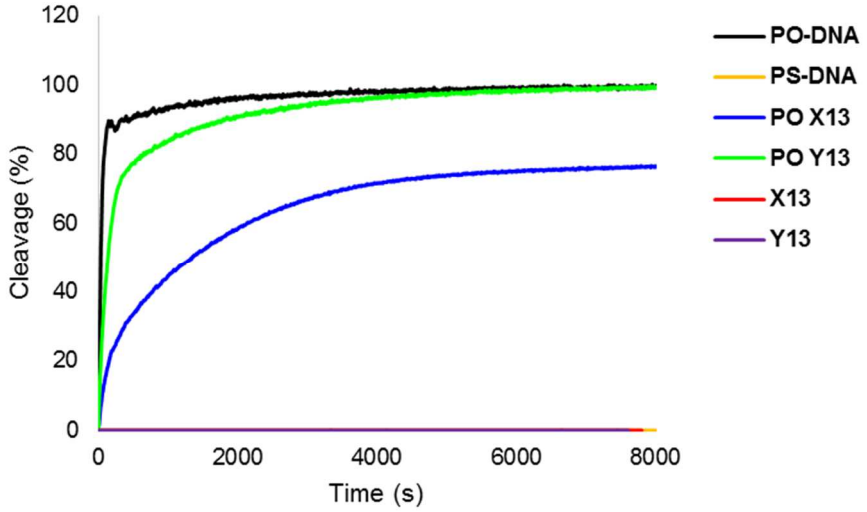


Figure 4-4: 3′-Exonuclease degradation of individual Invader and reference strands. Experiments were performed at 37 °C in magnesium buffer (200 µL solution, 50 mM Tris-HCl, 10 mM Mg^{2+}, pH 9.0) using [ON] = 3.3 µM and 0.03 U of snake venom phosphodiesterase.

4.3 Conclusions

We have developed a series of 2′-N-(pyren-1-yl)methyl-2′-N-methyl-2′-aminouridine and 2′-O-(pyren-1-yl)methyluridine modified phosphorothioate oligonucleotides. Thermal denaturation studies reveal that incorporation of **X**/**Y** into PS-DNA results in increased affinity toward cDNA relative to unmodified PS-DNA strands. However, the absolute T_m of duplexes between **X**- and **Y**-modified PS-DNA and cDNA were 10-12 °C lower than the corresponding duplexes between PO-DNA and cDNA. PS-Invader probes with two +1 zipper arrangements of monomers **X** or **Y** were found to recognize model dsDNA targets. Additionally, **X**- and **Y**-modified PS-DNA strands are much more inert toward enzymatic digestion by 3′-exonuclease (SVPDE) than **X**- and **Y**-modified PO-DNA strands. These results suggest that PS-Invaders exhibit promising properties for DNA targeting applications in cellular environments.

4.4 Experimental Section

4.4.1 Protocol - Synthesis and Purification of ONs

The individual PS-Invader strands were synthesized on a 0.2 μmol scale using a DNA synthesizer, succinyl linked LCAA-CPG (long chain alkyl amine controlled pore glass) columns with a pore size of 500Å, and standard protocols for incorporation of A^{Bz}, C^{Bz}, G^{iBu} and T DNA phosphoramidites. The following hand-coupling conditions were used for incorporation of monomers **X** and **Y** (coupling time; activator; coupling yield): **X** (15 min; 5-[3,5-bis(trifluoromethyl)phenyl]-1*H*-tetrazole; ~99%) and **Y** (15 min; 4,5-dicyanoimidazole; ~99%). Modified phosphoramidites were used at 50-fold molar excess and 0.05 M concentration in anhydrous CH_3CN. Sulfurization was carried out using 3-((*N*,*N*-dimethylaminomethylidene)amino)-3*H*-1,2,4-dithiazole-5-thione (Sulfurizing Reagent II, Glen Research) as per the manufacturer's specifications (~2 min). Cleavage from solid support and removal of protecting groups was accomplished using 32% aq. ammonia (55 °C, 16-24 h). ONs were purified in the DMT-on mode via ion-pair reverse phase HPLC (C18 column) using a 0.05 M triethylammonium acetate - water/acetonitrile gradient. This was followed by detritylation (80% aq. AcOH) and precipitation ($NaOAc/NaClO_4$/acetone, -18 °C for 12-16 h). The identity of synthesized ONs was established through MALDI-MS analysis (Table 4-S1) recorded in positive ions mode on a quadrupole time-of-flight tandem mass spectrometer equipped with a MALDI source using anthranilic acid, 3-hydroxypicolinic acid (3-HPA) or 2′,4′,6′-trihydroxyacetophenone (THAP) as matricies. Purity was verified by ion-pair reverse phase HPLC running in analytical mode (>80%).

4.4.2 Protocol – Thermal Denaturation Studies

ON concentrations were estimated using the following extinction coefficients for DNA (OD/μmol): G (12.01), A (15.20), T (8.40), C (7.05); RNA (OD/μmol): G (13.70), A (15.40), U (10.00), C (9.00); pyrene (22.4)[35]. Strands were thoroughly mixed and denatured by heating to 70-85 °C, followed by cooling to the starting temperature of the experiment. Quartz optical cells with a path length of 1.0 cm were used. T_m's of duplexes (1.0 μM final concentration of each strand) were measured using a UV/Vis spectrophotometer equipped with a Peltier temperature controller and determined as the maximum of the first derivative of thermal denaturation curves (A_{260} vs. T) recorded in medium salt phosphate buffer (T_m buffer: 100 mM NaCl, 0.1 mM EDTA and pH 7.0 adjusted with 10 mM Na_2HPO_4 and 5 mM Na_2HPO_4). The temperature of the denaturation experiments ranged from at least 15 °C below T_m to 20 °C above T_m (although not below 2 °C). A temperature ramp of 0.5 °C/min was used in all experiments. Reported T_m's are averages of two experiments within ± 1.0 °C.

4.4.3 Protocol - Absorption Spectra

UV-vis absorption spectra (range 200-600 nm) were recorded at 5 °C using the same samples and instrumentation as in the thermal denaturation experiments.

4.4.4 Protocol - Steady-state Fluorescence Emission Spectra

Steady-state fluorescence emission spectra of **X**- or **Y**-modified ONs and the corresponding duplexes with complementary DNA/RNA targets, were recorded in non-deoxygenated thermal denaturation buffer (each strand at 1.0 μM concentration) and obtained as an average of five scans using an excitation wavelength of λ_{ex} = 350 nm. Excitation and emission slits of 5.0 nm

and 2.5 nm, respectively, were used along with a scan speed of 600 nm/min. Experiments were conducted at 10 °C.

4.4.5 Protocol - Electrophoretic Mobility Shift Assay

This assay was performed essentially as previously described.[23] Unmodified DNA hairpins **DH1** and **DH2** were obtained from commercial sources and used without further purification. The DNA hairpins were 3′-DIG-labeled using the 2nd generation DIG Gel Shift Kit (Roche Applied Bioscience) following the manufacturer's recommendation. DIG-labeled ONs obtained in this manner were diluted and used without further purification in the recognition experiments. Pre-annealed probes (85 °C for 10 min, cooled to room temperature over 15 min) and DIG-labeled DNA hairpins (34.4 nM) were mixed and incubated in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 10% sucrose, 1.44 mM spermine tetrahydrochloride, pH 7.2) overnight (12-16 h) at ambient temperature. The reaction mixtures were then diluted with 6x DNA loading dye (Fermentas) and loaded onto a 16% non-denaturing polyacrylamide gel. Electrophoresis was performed using a constant voltage of 70 V for 2.5 h at ~4 °C using 0.5x TBE as a running buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA). Gels were blotted onto positively charged nylon membranes (Roche Applied Bioscience) using constant voltage with external cooling (100V, ~4 °C). The membranes were exposed to anti-digoxigenin-AP F_{ab} fragments as recommended by the manufacturer of the DIG Gel Shift Kit, transferred to a hybridization jacket, and incubated with the substrate (CSPD) in detection buffer for 10 min at 37 °C. The chemiluminescence of the formed product was captured on X-ray film, which was developed using an X-Omatic 1000A X-ray film developer (Kodak). The resulting bands were quantified using Image J software. Invasion efficiency was determined as the intensity ratio

between the recognition complex band and the total lane. An average of three independent experiments is reported along with standard deviations. Non-linear regression was used to fit data points from dose-response experiments, using a script written for the "Solver" module in Microsoft Office Excel.[36]

4.4.6 Protocol – 3′-Exonuclease Stability Assay

The stability of ONs against SVPDE (snake venom phosphodiesterase, Worthington Biochemical Corporation) was determined by monitoring the increase in absorbance at 260 nm as a function of time. SVPDE dissolved in H_2O (1.27 μL, 0.52 μg, 0.03 U) was added to a solution of ON (3.3 μM) in magnesium buffer (200 μL, 50 mM Tris·HCl, 10 mM $MgCl_2$, pH 9.0) at 37 °C.

4.4.7 Definition of Zipper Nomenclature

The following nomenclature describes the relative arrangement between two pyrene-functionalized monomers positioned on opposing strands in a duplex. The number *n* describes the distance measured in number of base pairs and has a positive value if a monomer is shifted toward the 5′-side of its own strand relative to a second reference monomer on the other strand. Conversely, *n* has a negative value if a monomer is shifted toward the 3′-side of its own strand relative to a second reference monomer on the other strand.

4.5 Supporting Information

4.5.1 Additional Tables, Figures, and Discussion

Table 4-S1: MALDI-MS of modified ONs.[a]

ON	PS-DNA sequence	Calc. *m/z* $[M+H]^+$	Found *m/z* $[M+H]^+$
X1	5′-G**X**G ATA TGC	3110.4	3110.3
X2	5′-GTG A**X**A TGC	3110.4	3110.4
X3	5′-GTG ATA **X**GC	3110.4	3110.2
X4	3′-CAC **X**AT ACG	3039.4	3039.2
X5	3′-CAC TA**X** ACG	3039.4	3039.5
X6	3′-CAC **X**A**X** ACG	3211.5	3212.0
Y1	5′-G**Y**G ATA TGC	3097.5	3097.2
Y2	5′-GTG A**Y**A TGC	3097.5	3097.4
Y3	5′-GTG ATA **Y**GC	3097.5	3097.2
Y4	3′-CAC **Y**AT ACG	3026.5	3026.3
Y5	3′-CAC TA**Y** ACG	3026.5	3026.4
Y6	3′-CAC **Y**A**Y** ACG	3242.5	3242.3
X7	5′-GG**X** ATA TAT AGG C	4433.5	4433.4
X8	3′-CCA **X**AT ATA TCC G	4313.5	4313.4
X9	5′-GG**X** A**X**A TAT AGG C	4662.6	4662.4
X10	3′-CCA **X**A**X** ATA TCC G	4542.6	4542.4
X11	5′- GGT A**X**A **X**AT AGG C	4662.6	4662.6
X12	3′-CCA TA**X** A**X**A TCC G	4542.6	4542.4
X13	5′-GG**X** ATA TA**X** AGG C	4662.6	4662.4
X14	3′-CCA **X**AT ATA **X**CC G	4542.6	4542.5
Y7	5′-GG**Y** ATA TAT AGG C	4420.5	4420.5
Y8	3′-CCA **Y**AT ATA TCC G	4300.5	4300.5
Y9	5′-GG**Y** A**Y**A TAT AGG C	4636.6	4636.6
Y10	3′-CCA **Y**A**Y** ATA TCC G	4516.5	4516.6
Y11	5′- GGT A**Y**A **Y**AT AGG C	4636.6	4636.6
Y12	3′-CCA TA**Y** A**Y**A TCC G	4516.5	4516.5
Y13	5′-GG**Y** ATA TA**Y** AGG C	4636.6	4636.8
Y14	3′-CCA **Y**AT ATA **Y**CC G	4516.5	4516.7

[a] For structure of monomers **X** and **Y** see Figure 4-1 in the main manuscript.

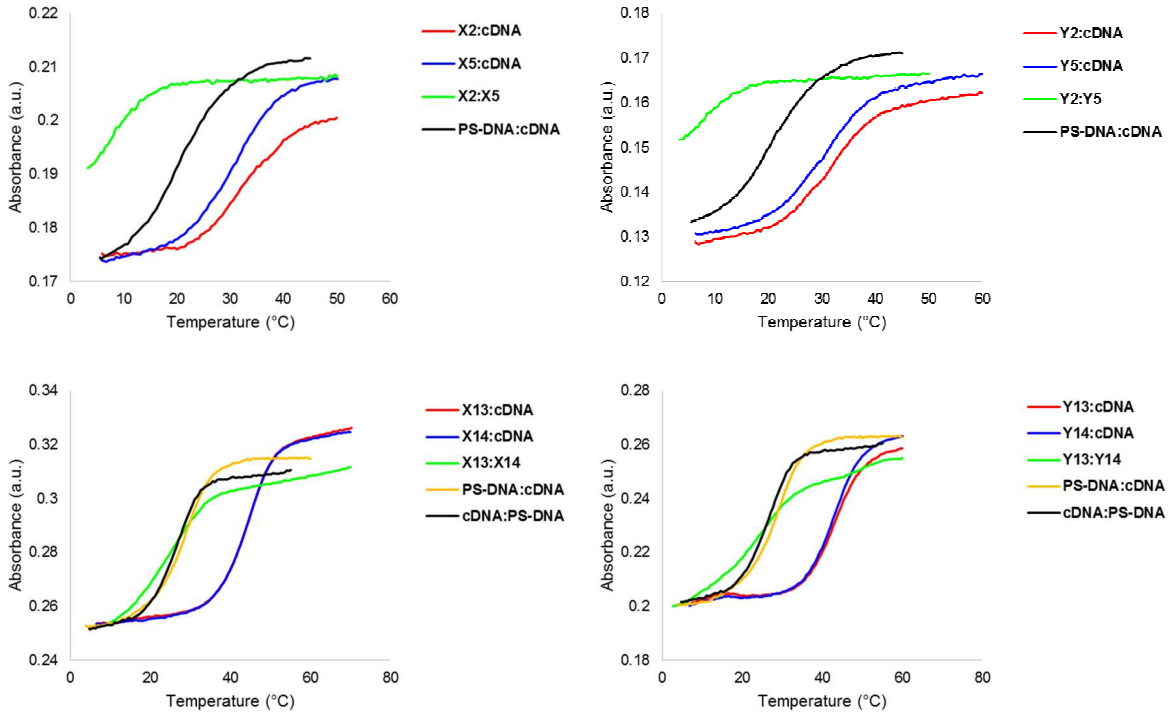


Figure 4-S1: Representative thermal denaturation profiles of duplexes between **X**-/**Y**-modified PS-DNA strands and cDNA. For experimental conditions, see Table 4-1.

Table 4-S2: DNA selectivity of **X**- and **Y**-modified PS-DNA.[a]

ON	PS-DNA sequence	$\Delta\Delta T_m$ (DNA-RNA) (°C)
X1	5′-G**X**G ATA TGC	+9.0
X2	5′-GTG A**X**A TGC	+13.0
X3	5′-GTG ATA **X**GC	+12.0
X4	3′-CAC **X**AT ACG	+9.0
X5	3′-CAC TA**X** ACG	+9.5
X6	3′-CAC **X**A**X** ACG	> +19.0
Y1	5′-G**Y**G ATA TGC	+8.5
Y2	5′-GTG A**Y**A TGC	+12.0
Y3	5′-GTG ATA **Y**GC	+9.5
Y4	3′-CAC **Y**AT ACG	+7.5
Y5	3′-CAC TA**Y** ACG	+9.5
Y6	3′-CAC **Y**A**Y** ACG	+17.0

[a] DNA selectivity defined as $\Delta\Delta T_m$ (DNA-RNA) = ΔT_m (vs cDNA) - ΔT_m (vs cRNA).

Binding specificity of X-/Y-modified PS-DNA.

The binding specificities of centrally modified 9-mer PS-DNA strands were studied using DNA targets with mismatched nucleotides opposite to the modification (Table 4-S3). Excellent discrimination of the C-mismatched target is observed, while discrimination of G- or T-mismatched targets is much less efficient. In contrast, doubly modified 9-mer PS-DNA very efficiently discriminate DNA targets with a single mismatched nucleotide opposite of the central 2′-deoxyriboadenosine (Table 4-S4). These trends mirror our observations with **X**-/**Y**-modified PO-DNA strands, further substantiating intercalative binding modes of the pyrene moeities.[22,24]

Table 4-S3: Discrimination of mismatched DNA targets by **X2** and **Y2**.[a]

			DNA: 3′-CAC T**B**T ACG			
			T_m (°C)	ΔT_m (°C)		
ON	PS-DNA sequence	**B** =	**A**	C	G	T
X2	5′-GTG A**X**A TGC		**32.5**	-16.0	-2.5	-9.0
Y2	5′-GTG A**Y**A TGC		**31.5**	-13.0	-5.5	-7.5

[a] For conditions of thermal denaturation experiments, see Table 4-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched DNA:DNA duplex.

Table 4-S4: Discrimination of mismatched DNA targets by **X6** and **Y6**.[a]

			DNA: 5′-GTG A**B**A TGC			
			T_m (°C)	ΔT_m (°C)		
ON	PS-DNA sequence	**B** =	**T**	A	C	G
X6	3′-CAC **X**A**X** ACG		**30.5**	<-20.5	-13.0	-14.0
Y6	3′-CAC **Y**A**Y** ACG		**30.0**	-18.5	-16.0	-13.5

[a] For conditions of thermal denaturation experiments, see Table 4-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched DNA:DNA duplex.

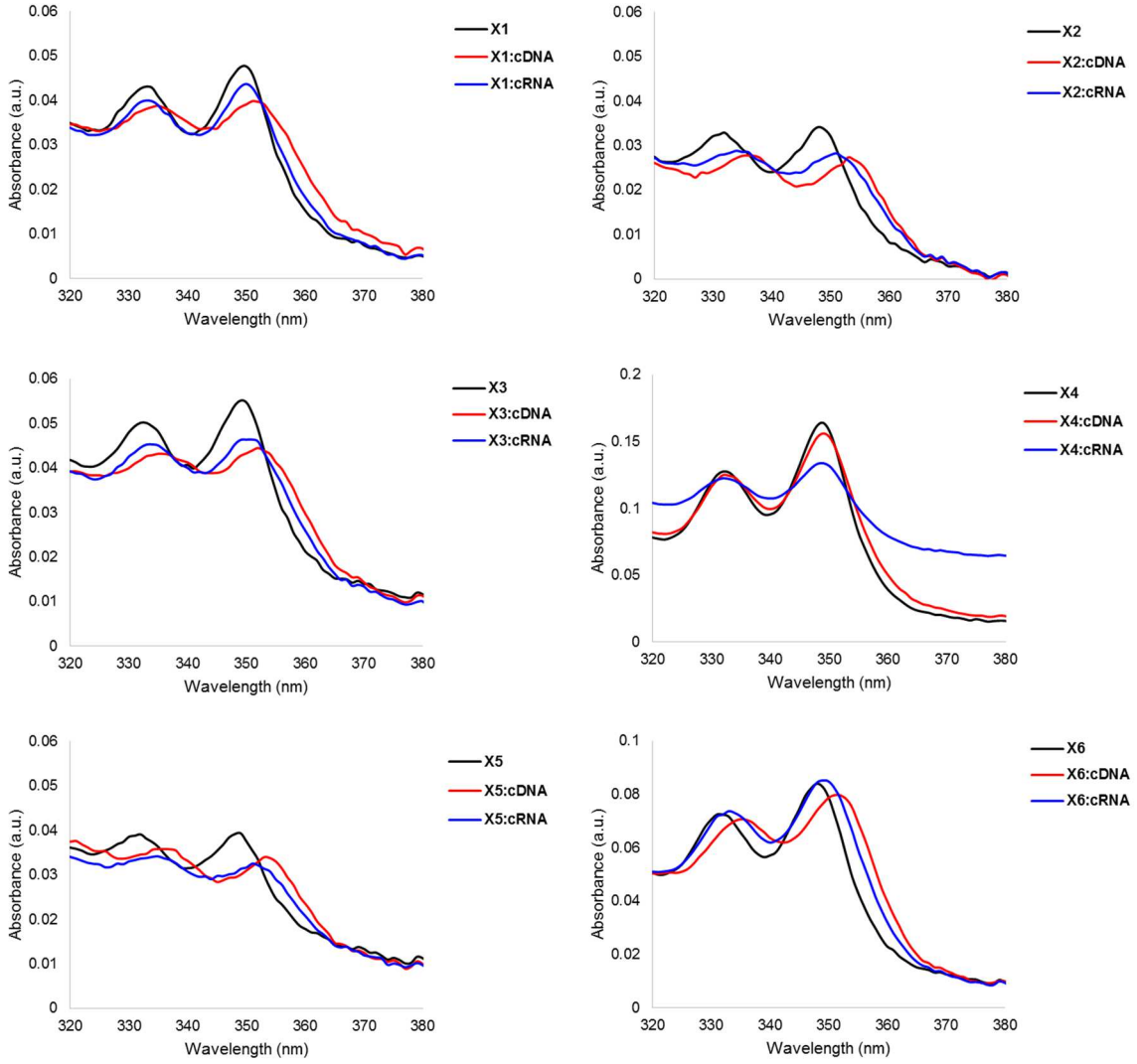


Figure 4-S2: Absorption spectra of single-stranded **X**-modified PS-DNA and the corresponding duplexes with cDNA/cRNA. Spectra were recorded at $T = 5$ °C using each strand at 1.0 μM concentration in T_m buffer. Different axis scales are used.

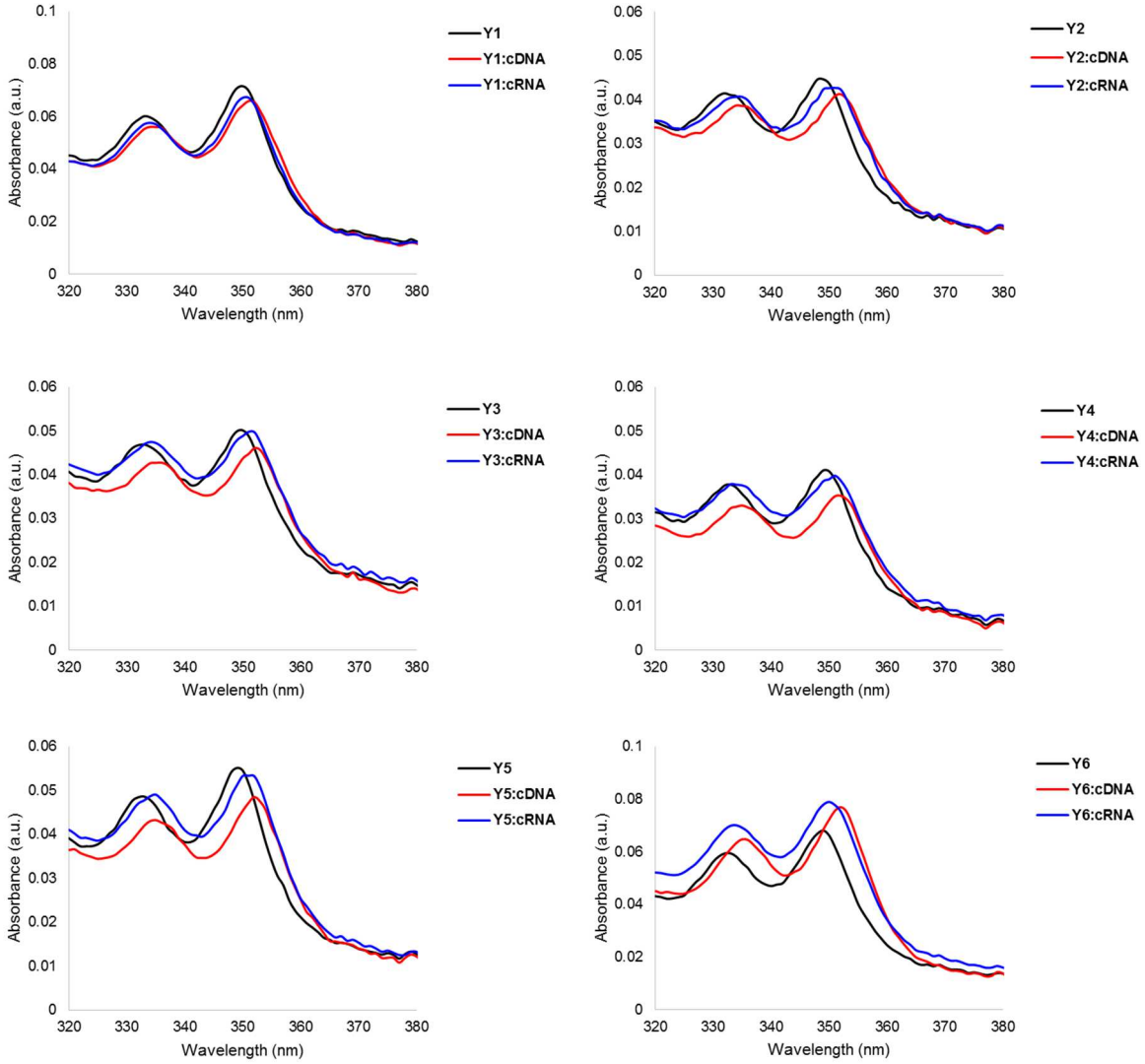


Figure 4-S3: Absorption spectra of single-stranded **Y**-modified PS-DNA and the corresponding duplexes with cDNA/cRNA. Spectra were recorded at $T = 5$ °C using each strand at 1.0 μM concentration in T_m buffer. Different axis scales are used.

Table 4-S5: Absorption maxima in the 300-500 nm region for **X**- and **Y**-modified PS-DNA and the corresponding duplexes with cDNA or cRNA.[a]

		λ_{max} [$\Delta\lambda_{max}$]/nm		
ON	PS-DNA sequence	SSP	+cDNA	+cRNA
X1	5′-G**X**G ATA TGC	350	351 [+1]	350 [±0]
X2	5′-GTG A**X**A TGC	348	353 [+5]	351 [+3]
X3	5′-GTG ATA **X**GC	349	352 [+3]	351 [+2]
X4	3′-CAC **X**AT ACG	349	349 [±0]	349 [±0]
X5	3′-CAC TA**X** ACG	349	353 [+4]	351 [+2]
X6	3′-CAC **X**A**X** ACG	348	351 [+3]	349 [+1]
Y1	5′-G**Y**G ATA TGC	350	351 [+1]	351 [+1]
Y2	5′-GTG A**Y**A TGC	349	352 [+3]	351 [+2]
Y3	5′-GTG ATA **Y**GC	350	352 [+2]	351 [+1]
Y4	3′-CAC **Y**AT ACG	349	352 [+3]	351 [+2]
Y5	3′-CAC TA**Y** ACG	349	352 [+3]	351 [+2]
Y6	3′-CAC **Y**A**Y** ACG	349	352 [+3]	350 [+1]

[a] SSP = single-stranded probe. Measurements were performed at 5 °C using a spectrophotometer and quartz optical cells with 1.0 cm path lengths. For buffer conditions, see Table 4-1.

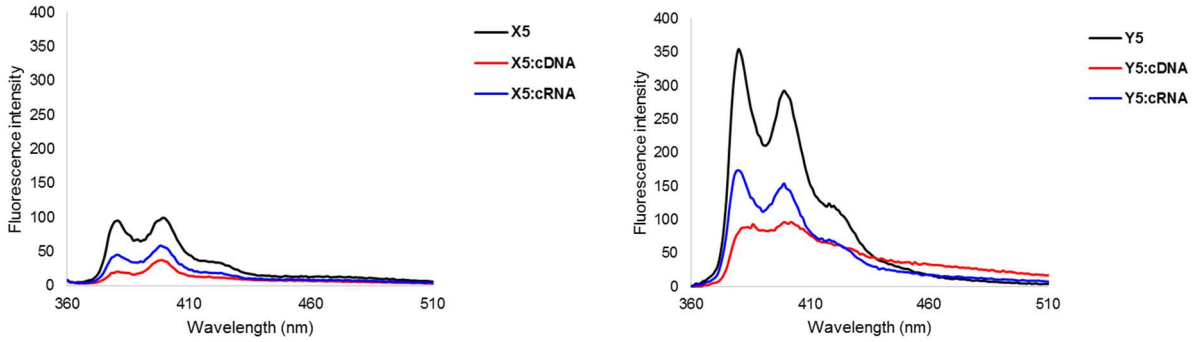


Figure 4-S4: Representative steady-state fluorescence emission spectra of single-stranded **X5** and **Y5** ONs and the corresponding duplexes with cDNA/cRNA. Spectra were recorded at T = 10 °C using λ_{ex} = 350 nm. Each strand was used at 1.0 μM concentration in T_m buffer.

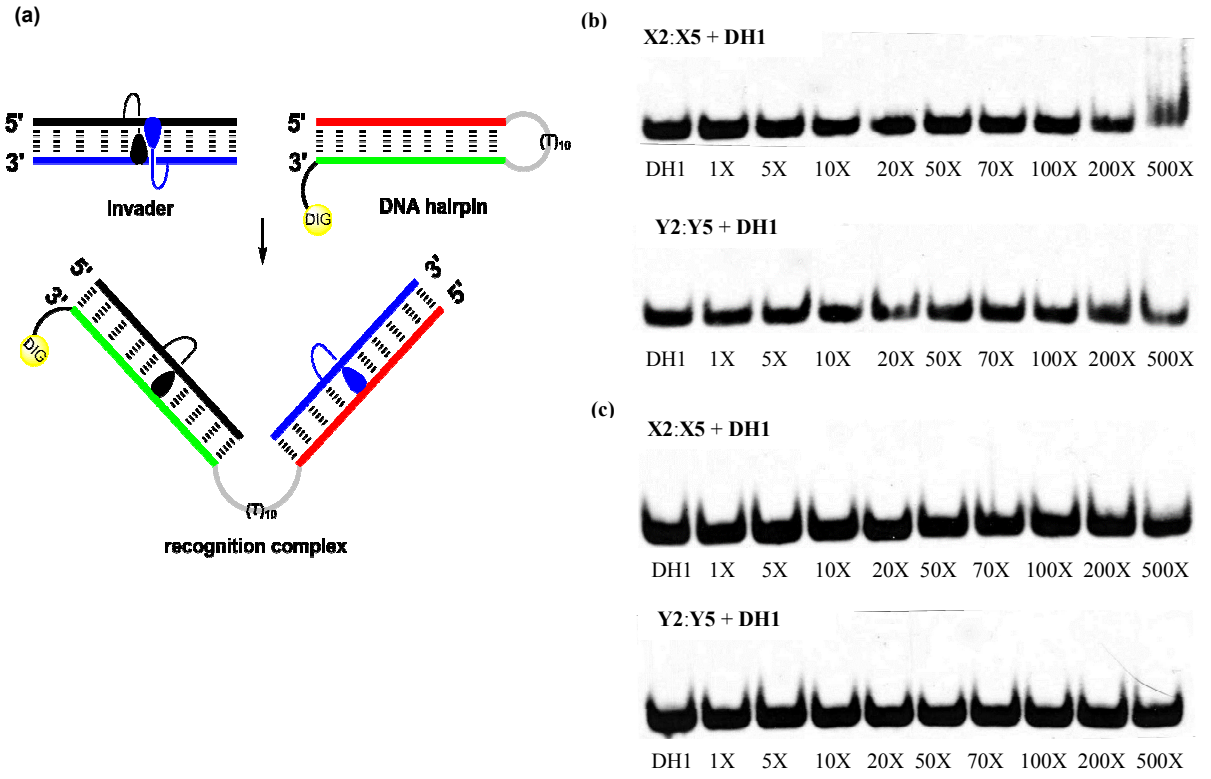


Figure 4-S5: Recognition of dsDNA model target **DH1** using **X2**:**X5** or **Y2**:**Y5**. (a) Illustration of recognition process. Sequence of **DH1**: 5′-GTGATATGC-(T_{10})-GCTTATCAC-DIG-3′; (b) representative gel electrophoretograms from experiments in which **DH1** (34.4 nM) was incubated with **X2**:**X5** or **Y2**:**Y5** at ambient temperature for 12-16 h; (c) representative gel electrophoretograms from experiments in which **DH1** (34.4 nM) was annealed in the presence of **X2**:**X5** or **Y2**:**Y5** at 85 °C for 15 min, followed by cooling to room temperature over ~30 min and incubation at ambient temperature for 12-16 h. Experiments were conducted in 1X HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 10% sucrose, 1.4 mM spermine tetrahydrochloride, pH 7.2) and then run on 16% non-denaturing PAGE (performed at 70V, 2.5h, ~4 °C) using 0.5x TBE as a running buffer (45 mM Tris, 45 mM boric acid, 1mM EDTA); DIG: digoxigenin.

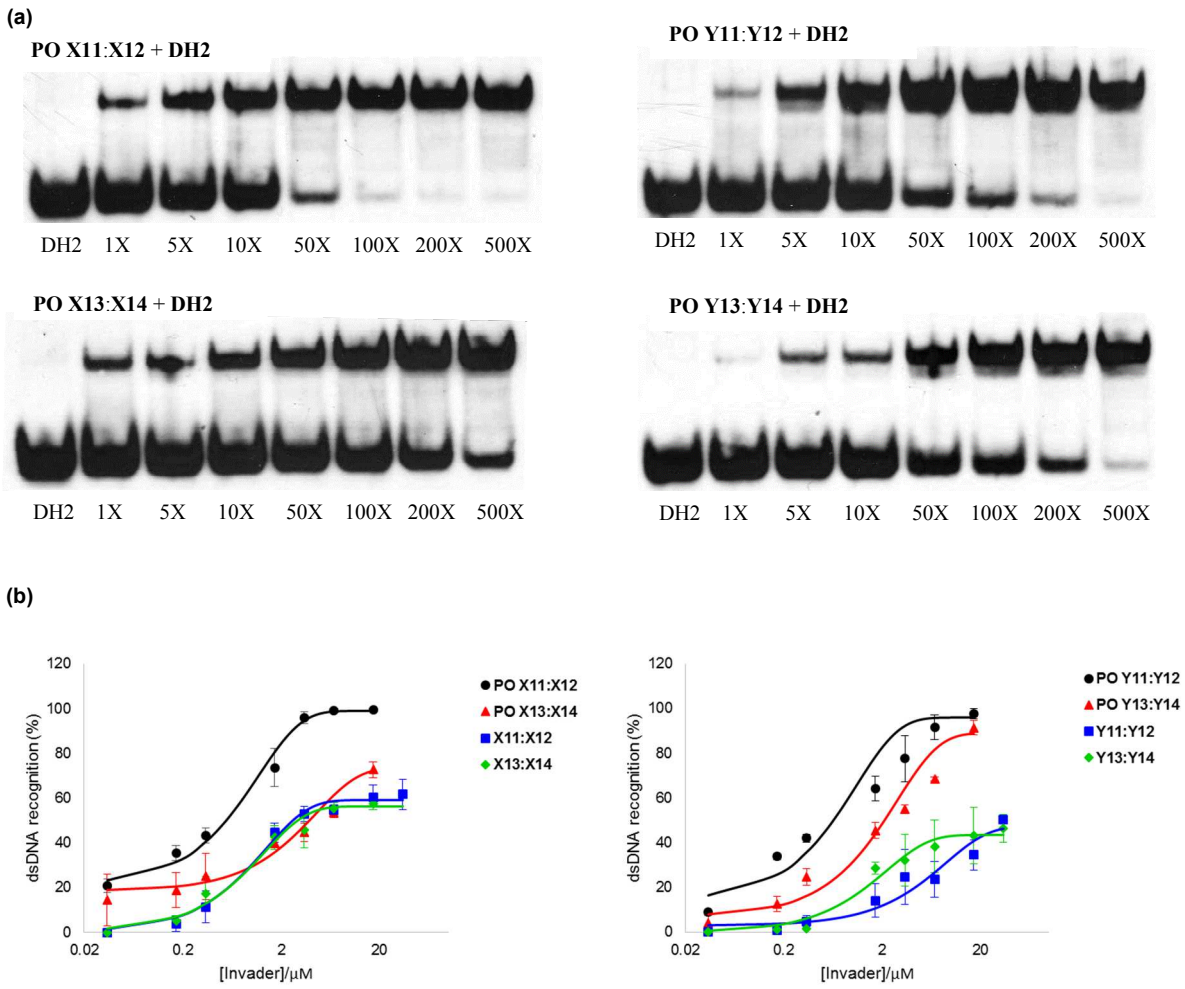


Figure 4-S6: Recognition of dsDNA model target **DH2** using different Invader probes. (a) Representative electrophoretograms for recognition of **DH2** using 1- to 500-fold excess of the PO-DNA versions of **X11**:**X12**, **X13**:**X14**, **Y11**:**Y12**, **Y13**:**Y14**; (b) dose-response curves (average of at least three independent experiments, error bars represent standard deviation) relative to the corresponding PS-DNA Invaders. The sequence of DNA hairpin **DH2** and experimental are given in Figure 4-3.

Stability of Invader Probes Against DNase 1 - Protocol

An aqueous solution of DNase I (Worthington Biochemical Corporation - 0.61 μL of a 2 μg/mL solution) was added to a 6.3 μM solution of pre-annealed Invader in TE buffer (100 μL, 10 mM Tris·HCl, 0.1 mM EDTA, 10 mM $MgCl_2$, pH 8.0) and the mixture was incubated at 20 °C in a water bath. Aliquots (10 μL) were removed at specific times (0.5, 1, 2, 5, 10, 20 and 30 min) and degradation was quenched by addition of ethidium bromide buffer (2.0 mL, 5 mM Tris·HCl, 0.5 mM EDTA, 0.5 μg/mL EtBr, pH 8.0). The fluorescence intensity of the solution was measured (λ_{ex} = 525 nm; λ_{em} = 600 nm) using the same instrumentation employed for the steady-state fluorescence experiments. Intensities were averaged over 15 minutes. Experiments were performed in duplicates and representative trends are shown.

Stability of Invader Probes Against DNase 1 - Discussion

The stability of Invader probes against DNase I was evaluated using an ethidium bromide based assay. In this assay, high levels of fluorescence are observed when the studied duplex is intact due to intercalated ethidium bromide, while low levels of fluorescence are expected if a duplex has been degraded to single strands (or shortened to a level where the duplex dissociates).[37]

Similar assays have been used to show that long PS-DNA duplexes exhibit excellent stability against DNase I.[28b,28c] Figure 4-S7 shows the fluorescence intensity profiles of representative 13-mer Invaders in the presence of DNase I. The PO-DNA analogs of **X13**:**X14** and **Y13**:**Y14** are moderately resistant to DNase I degradation with half-lives of 15 min and >30 min, respectively, whereas the unmodified PO-DNA duplex is rapidly degraded.

PS-DNA Invaders **X13**:**X14** and **Y13**:**Y14** did not show any change in fluorescence relative to background EtBr buffer (data not shown), presumably because the probe duplexes are

dissociated into single strands at the experimental conditions used for this assay (T_m of **X13**:**X14** and **Y13**:**Y14** are <25 °C, Table 4-3). Longer and more thermostable PS-DNA Invaders will need to be synthesized in order to assess their DNAse I stability via this assay.

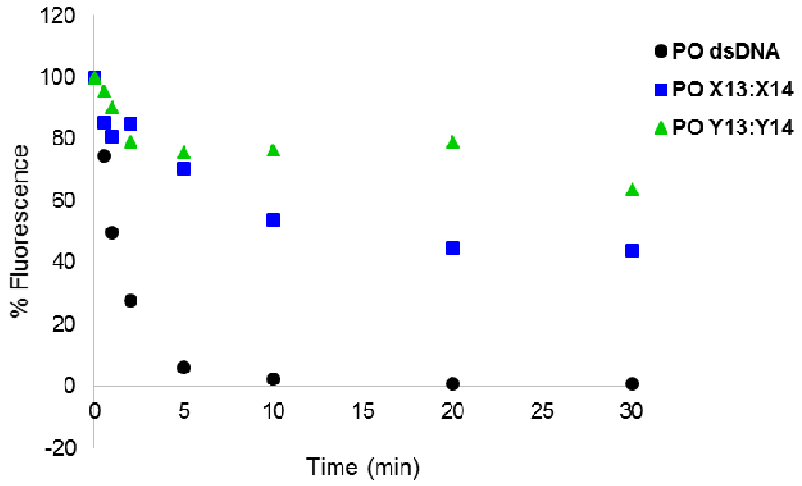


Figure 4-S7: DNase I stability of PO-DNA Invaders **X13**:**X14**, **Y13**:**Y14** and the corresponding unmodified PO-DNA duplex, as assessed by ethidium bromide assay (for details, see protocols). Curves are average of two experiments.

Acknowledgements

We thank Dale C. Guenther and Grace H. Anderson for the use of PO-Invaders and data from a previous study, for this manuscript.

4.6 References

(1) Rogers, F. A.; Lloyd, J. A.; Glazer, P. M. *Curr. Med. Chem.: Anti- Cancer Agents* **2005**, 5, 319 − 326.

(2) Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. *J. Mol. BioSyst*. **2006**, 2, 551 − 560.

(3) Nielsen, P. E. *ChemBioChem* **2010**, 11, 2073 − 2076.

(4) Mukherjee, A.; Vasquez, K. M. *Biochimie* **2011**, 93, 1197 − 1208.

(5) Aiba, Y.; Sumaoka, J.; Komiyama, M. *Chem. Soc. Rev.* **2011**, 40, 5657 − 5668.

(6) Vaijayanthi, T.; Bando, T.; Pandian, G. N.; Sugiyama, H. *ChemBioChem* **2012**, 13, 2170 − 2185.

(7) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, 36, 5123 – 5138.

(8) a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, 254, 1497 – 1500. (b) Bentin, T.; Nielsen, P. E. *Biochemistry* **1996**, 35, 8863 – 8869. (c) Kaihatsu, K.; Braasch, D. A.; Cansizoglu, A.; Corey, D. R. *Biochemistry* **2002**, 41, 11118 – 11125. (d) Kaihatsu, K.; Janowski, B. A.; Corey, D. R. *Chem. Biol*. **2004**, 11, 749 – 758.

(9) (a) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, 114, 8783 – 8794. (b) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol*. **2003**, 13, 284 – 299. (c) Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem*. **2013**, 21, 6101 – 6114.

(10) (a) Kutyavin, I. V.; Rhinehart, R. L.; Lukhtanov, E. A.; Gorn, V. V.; Meyer, R. B., Jr.; Gamper, H. B., Jr. *Biochemistry* **1996**, 35, 11170 − 11176. (b) Woo, J.; Meyer, R. B.; Gamper, H. B. *Nucleic Acids Res.* **1996**, 24, 2470 – 2475.

(11) (a) Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 11804 – 11808. (b) Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corradini, R.; Sforza, S.; Komiyama, M. *Nucleic Acids Res*. **2008**, 36, 1464 − 1471.

(12) (a) Griffith, M. C.; Risen, L. M.; Greig, M. J.; Lesnik, E. A.; Sprankle, K. G.; Griffey, R. H.; Kiely, J. S.; Freier, S. M. *J. Am. Chem. Soc.* **1995**, 117, 831 – 832. (b) Bentin, T.; Larsen,

H. J.; Nielsen, P. E. *Biochemistry* **2003**, 42, 13987 – 13995. (c) Kaihatsu, K.; Shah, R. H.; Zhao, X.; Corey, D.; *Biochemistry*, **2003**, 42, 13996 – 14003.

(13) (a) Rapireddy, S.; Bahal, R.; Ly, D. H. *Biochemistry* **2011**, 50, 3913 – 3918. (b) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, 128, 10258 – 10267.

(14) (a) Rusling, D. A.; Powers, V. E. C.; Ranasinghe, R. T.; Wang, Y.; Osborne, S. D.; Brown, T.; Fox, K. *Nucleic Acids Res.* **2005**, 33, 3025 −3032. (b) Hari, Y.; Obika, S.; Imanishi, T. *Eur. J. Org. Chem.* **2012**, 2875 – 2887.

(15) (a) Bogdanove, A. J.; Voytas, D. F. *Science* **2011**, 333, 1843 − 1846. (b) Gaj, T.; Gersbach, C. A.; Barbas, C. F., III. *Trends Biotechnol.* **2013**, 31, 397 – 405.

(16) Moreno, P. M. D.; Geny, S.; Pabon, Y. V.; Bergquist, H.; Zaghloul, E. M.; Rocha, C. S. J.; Oprea, I. I.; Bestas, B.; El-Andaloussi, S.; Jørgensen, P. T.; Pedersen, E. B.; Lundin, K. E.; Zain, R.; Wengel, J.; Smith, C. I. E. *Nucleic Acids Res.* **2013**, 41, 3257 – 3273. (b) Zaghloul, E. M.; Madsen, A. S.; Moreno, P. M. D.; Oprea, I. I.; El-Andaloussi, S.; Bestas, B.; Gupta, P.; Pedersen, E. B.; Lundin, K. E.; Wengel, J.; Smith, C. I. E. *Nucleic Acids Res.* **2011**, 39, 1142 – 1154. (c) Ge, R.; Heinonen, J. E.; Svahn, M. G.; Mohamed, A. J.; Lundin, K. E.; Smith, C. I. E. *FASEB J.* **2007**, 21, 1902 − 1914. (d) Filichev, V. V.; Vester, B.; Hansen, L. H.; Pedersen, E. B. *Nucleic Acids Res*. **2005**, 33, 7129 − 7137.

(17) Hrdlicka, P. J.; Kumar, T. S.; Wengel, J. *Chem. Commun*. **2005**, 4279 − 4281.

(18) Crothers, D. M. *Biopolymers* **1968**, 6, 575 – 584.

(19) Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2010**, 8, 2028 − 2036.

(20) Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J.; *Org. Biomol. Chem.* **2014**, 12, 7758 – 7773.

(21) Sau, S. P.; Madsen, A. S.; Podbevsek, P.; Andersen, N. K.; Kumar, T. S.; Andersen, S.; Rathje, R. L.; Anderson, B. A.; Guenther, D. C.; Karmakar, S.; Kumar, P.; Plavec, J.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, 78, 9560 – 9570.

(22) Anderson, B. A.; Onley, J. J.; Hrdlicka, P. J. *Manuscript in progress*. See Chapter 3.

(23) Didion, B. A.; Karmakar, S.; Guenther, D. C.; Sau, S. P.; Verstegen, J. P.; Hrdlicka, P. J. *ChemBioChem* **2013**, 4, 3447 − 3454.

(24) Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem*. **2011**, 76, 7119 − 7131.

(25) Kumar, T. S.; Madsen, A. S.; Østergaard, M. E.; Sau, S. P.; Wengel, J.; Hrdlicka, P. *J. Org. Chem.* **2009**, 74, 1070 − 1081.

(26) Andersen, N. K.; Anderson, B. A.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, 78, 12690 – 12702.

(27) (a) Suggs, J. W.; Taylor, D. A. *Nucleic Acids Res.* **1985**, 13, 5707 – 5716. (b) Laplanche, L. A.; James, T. L.; Powell, C.; Wilson, W. D.; Uznanski, B.; Stec, W. J.; Summers, M. F.; Zon, G. *Nucleic Acids Res.* **1986**, 14, 9081 – 9093.

(28) (a) Morgan, A. R.; Evans, D. H.; Lee, J. S.; Pulleyblank, D. E. *Nucleic Acids Res.* **1979**, 7, 571 – 593. (b) Braun, R. P.; Lee, J. S. *J. Immun.* **1988**, 141, 2084 – 2089. (c) Latimer, L. J. P.; Hampel, K.; Lee, J. S. *Nucleic Acids Res.* **1989**, 17, 1549 – 1561.

(29) Guenther, D. C.; Anderson, G. H.; Anderson, B. A.; Karmakar, S.; Hrdlicka, P. J. *Manuscript in Progress.*

(30) Nakamura, M.; Fukunaga, Y.; Sasa, K.; Ohtoshi, Y.; Kanaori, K.; Hayashi, H.; Nakano, H.; Yamana, K. *Nucl. Acids Res.* **2005**, 33, 5887 − 5895.

(31) Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, 78, 12040 – 12048.

(32) Stein, C. A.; Subasinghe, C.; Shinozuka, K.; Cohen, J. S. *Nucleic Acids Res.* **1988**, 16, 3209 – 3221.

(33) Dougherty, G.; Pilbrow, J. R. *Int. J. Biochem*. **1984**, 16, 1179 – 1192; (b) Asanuma, H.; Fujii, T.; Kato, T.; Kashida, H. *J. Photochem. Photobiol., C* **2012**, 13, 124 – 135.

(34) Marin, V.; Hansen, H. F.; Koch, T. R.; Armitage, B. A. *J. Biomol. Struc. Dyn.*, **2004**, 21, 841 – 850.

(35) Dioubankova, N. N.; Malakhov, A. D.; Stetsenko, D. A.; Gait, M. J.; Volynsky, P. E.; Efremov, R. G.; Korshun, V. A. *ChemBioChem* **2003**, *4*, 841 – 847.

(36) Brown, A. M. *Comput. Meth. Prog. Biomed.* **2001**, 65, 181 – 200.

(37) (a) Morgan, A. R.; Lee, J. S.; Pulleyback, D. E.; Murray, N. L.; Evans, D. H. *Nucleic Acids Res.* **1979**, 7, 547 – 570. (b) Morgan, A.R.; Evans, D. H.; Lee, J. S.; Pulleyback, D. E. *Nucleic Acids Res.* **1979**, 7, 571 – 594.

CHAPTER 5: 2′-Atom Substitutions and Their Affect on Invader-Mediated DNA Recognition Efficiency

Brooke A. Anderson and Patrick J. Hrdlicka

Department of Chemistry, University of Idaho, Moscow

Abstract

Invader probes have proven to be promising tools for DNA diagnostic applications. These double-stranded probes are energetically activated for mixed-sequence recognition of biological DNA through modification with "+1 interstrand zippers" of monomers such 2'-amino-α-L-LNA, RNA and 2′-*N*-methyl-2′-amino-DNA that are modified at the 2′-position with intercalating pyrene moieties. In order to further optimize the design of Invader probes, we set out to study the impact of the electronegativity of the atom at the 2′-position on Invader-mediated DNA recognition efficiency. Toward this end, we synthesized 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine phosphoramidite **4** and incorporated it into oligodeoxyribonucleotides (ONs). Pseudorotational analysis of 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine shows the furanose ring adopts a predominantly *South*-type conformation, whereas 2′-*O*-(pyren-1-yl)methyluridine adopts an equal mixture of *North*- and *South*- conformations and 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-aminouridine adopts an exclusively *South*-type conformation. Thermal denaturation studies, analysis of thermodynamic parameters, and DNA recognition experiments demonstrate that Invaders based on 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine display less efficient DNA recognition than Invaders based on 2′-*O*-(pyren-1-yl)methyluridine or 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-aminouridine.

5.1 Introduction

Nucleic acids are one of the most important classes of biomolecules within an organism's cells. DNA plays key roles in cellular processes, such as carrying the genetic information that encodes for the proteins that perform the various catalytic and regulatory functions within the cell. Development of synthetically modified ligands that can recognize naturally occurring nucleic acids has attracted considerable attention due to their potential as biological, diagnostic, and therapeutic probes.[1-6] Probes that recognize double-stranded DNA (dsDNA) via one of the grooves is a well-established methodology which includes probe technologies such as triplex-forming oligonucleotides (TFOs)[7], Peptide Nucleic Acids (PNAs)[8], pyrrole-imidazole polyamides[9], and engineered proteins.[1,10] However, these technologies have limitations such as requirements for polypurine tracts, non-physiological pH and salt conditions, and/or short targets. Alternative methodologies that recognize mixed-sequence dsDNA via duplex-invasion techniques, which alleviate some of the sequence restrictions of the aforementioned approaches, such as pseudo-complementary DNA (pcDNA)[11], pcPNA[12], and γ-PNAs[13], among other approaches[14-17] are particularly appealing due to the predictability of the Watson-Crick base-pairing rules. However, many of the current DNA duplex-invasion techniques fall to the shortcomings of sequence and concentration limitations, self-inhibitory effects, and low salinity requirements. Thus, non-protein based probes that efficiently recognize mixed-sequence DNA at physiologically relevant conditions remain elusive.

Our laboratory has introduced Invader probes as an alternative approach toward mixed-sequence recognition of dsDNA.[18] These probes are energetically activated duplexes with +1 interstrand zipper arrangements of intercalator-functionalized nucleotides (for an illustration, see Fig. 5-1). This particular zipper arrangement results in a locally perturbed and destabilized

region of the probe duplex, most likely since the intercalators are forced into the same region which is in violation of the 'nearest neighbor exclusion principle.' The principle states that two sites directly neighboring an occupied intercalation site must remain unoccupied, or intercalation becomes anti-cooperative due to duplex expandibility.[19] Thus, two intercalators in one site is unfavorable, resulting in local duplex unwinding and the formation of an 'energetic hotspot'.[18,20] On the other hand, the two strands that constitute the Invader probe are able to site-specifically position intercalators into the duplex core upon hybridization with complementary DNA (cDNA), leading to exceptionally stabilized duplexes. We have used the difference in thermostabilities between Invader probes and probe-target duplexes to drive recognition of short iso-sequential mixed-sequence DNA hairpins[18,20-22] and chromosomal DNA targets.[23]

We are interested in evaluating chemical variants for optimization of the Invader probe design, such as the influence of linker length, chemical composition and orientation between the intercalator and sugar skeleton.[21,24] These studies indicate a methyl linker at the 1- or 2-position of pyrene result in the most efficient dsDNA recognition. In this report, we compare the dsDNA recognition characteristics of Invaders that are modified with 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine as the key activating component relative to Invader probes based on current generation building blocks, i.e., 2′-*O*-(pyren-1-yl)methyl-ribonucleotides and 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-aminouridine (Figure 5-1).[24] It is well-known that the nature of the 2′-substituent influences the conformation of furanose sugars via stereoelectronic effects.[25,26] At the onset of this study, we hypothesized that a less electronegative element would weaken the gauche effect between O4′ and the 2′-substituent, leading to increased propensity for C_2′-endo (*South*-type) conformations, more favorable pre-organization for intercalation of the

attached pyrene moiety, and even higher affinity toward cDNA than current Invader probes. Herein, we describe i) a short synthetic route to 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine phosphoramidite **4** and its incorporation in oligodeoxyribonucleotides (ONs), and the results from ii) coupling constant analysis providing insights into conformational preferences, iii) thermal denaturation experiments and thermodynamic parameter analysis, iv) UV-Vis absorption and fluorescence emission experiments, and v) comparative recognition experiments using model dsDNA targets relative to Invader based on 2′-O-(pyren-1-yl)methyluridine and 2′-N-(pyren-1-yl)methyl-2′-N-methyl-2′-aminouridine monomers **O** and **N**, respectively.

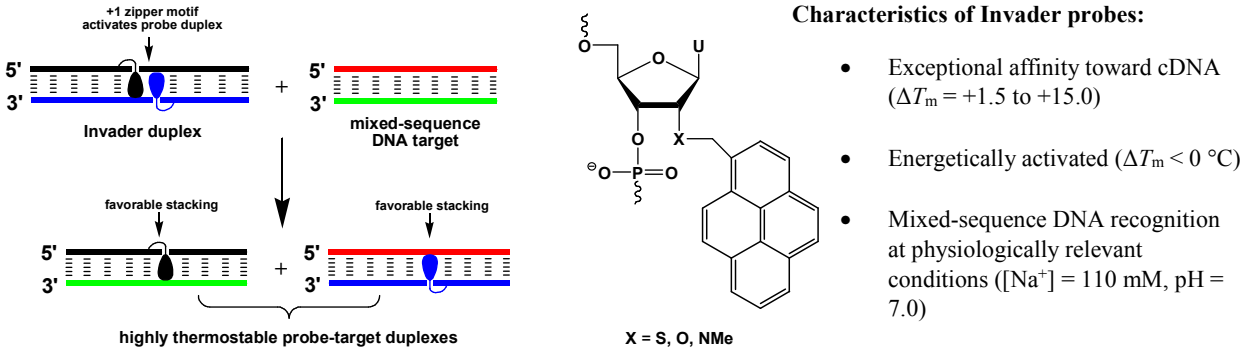


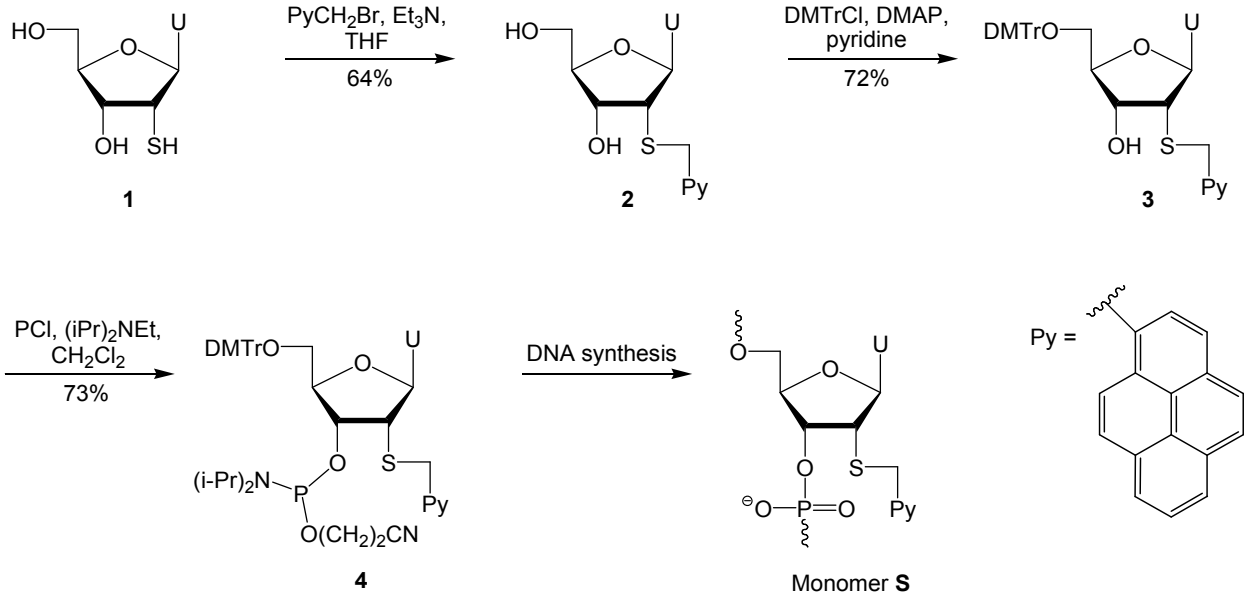
Figure 5-1: Illustration and characteristics of the Invader approach. Droplets denote the intercalating moiety.

5.2 Results and Discussion

5.2.1 Synthesis of 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine Phosphoramidite

2′-Deoxy-2′-thiouridine **1** - prepared from uridine in ~50% yield over three steps as described in the literature[27] - was used as the starting material for the synthesis of phosphoramidite **4**

(Scheme 5-1). Nucleoside **1** was alkylated at the S2′-position using pyren-1-ylmethyl chloride under mildly basic conditions,[28] to afford nucleoside **2** in 64% yield. Similar yields were obtained when pyren-1-ylmethyl bromide was used as the alkylating agent (results not shown). Standard O5′-DMT protection afforded nucleoside **3** in 72% yield, which was treated with 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite and Hünig's base to give target phosphoramidite **4** in 73% yield.



Scheme 5-1: Synthesis of target nucleoside **4**. DMTr = 4,4′-dimethoxytrityl; DMAP = 4-dimethylaminopyridine; PCl reagent = 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite.

5.2.2 Structural Analysis of Nucleosides

To test the hypothesis that the lower electronegativity of the 2′-thio atom of 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine (monomer **S)** induces a greater proportion of *South* type furanose conformations relative to monomers **N** and **O**, we performed a coupling constant analysis. Thus,

$^{3}J_{HH}$ scalar coupling constants for the endocyclic sugar protons of nucleoside **3** were used as input in a Matlab-based pseudorotational analysis program[29] which facilitates determination of pseudorotation phase angles (P) and puckering amplitudes (ϕ_m) for five-membered ring systems, by solving modified Karplus-Diez-Donders equations (Table 5-S1 in the Supporting Information).[29-31]

Nucleoside **3** is predicted to predominantly adopt a *South* type conformation (P = 143°, ϕ_m = 38°, %S = 61%), while the corresponding nucleoside of monomer **O** is predicted to have two equally populated main conformations, i.e., a *North* type conformation (P = 11°, ϕ_m = 38°) and a *South* type conformation (P = 130°, ϕ_m = 33°, %S = 49%). Somewhat surprisingly, the corresponding nucleoside of monomer **N** is predicted to exclusively adopt *South* type conformation (P = 145°, ϕ_m = 38°, main conformer (61% population) and P = 122°, ϕ_m = 27°). These observations suggest that i) the furanose ring of monomer **S** indeed may be more likely to occupy *South* type conformations than the furanose ring of monomer **O**, and ii) the furanose ring of monomer **N** is very prone to adopt *South* type conformations, presumably due to the additional steric bulk of the 2′-*N*-methyl group.

5.2.3 Synthesis of Modified ONs

Phosphoramidite **4** was used in automated solid phase DNA synthesis to incorporate monomer **S** into oligodeoxyribonucleotides (ONs) using extended hand-coupling times (15 min) and 4,5-dicyanoimidazole as an activator, which resulted in stepwise coupling yields of >95%. The identity and purity of the modified ONs was established through MALDI-ToF (Table 5-S2) and ion-pair reverse phase HPLC (>90% purity).

5.2.4 Thermal Denaturation Properties

The thermal denaturation temperatures (T_m's) of duplexes between **S**-modified ONs and complementary DNA or RNA (cDNA/cRNA) were determined from thermal denaturation curves recorded in medium salt buffer ($[Na^+]$ = 110 mM) and were compared relative to unmodified and corresponding **N**- or **O**-modified duplexes. Duplexes between **S**-modified ONs and cDNA are significantly more stable than the corresponding unmodified reference duplex (ΔT_m between ±0.0 and +8.0 °C, Table 5-1), while duplexes with cRNA are destabilized (ΔT_m between -10.0 to -2.0 °C, Table 5-1). ONs in which monomer **S** is flanked by 3′-purines induce particularly prominent duplex stabilization (compare ΔT_m's for **B2**- and **B4**- series, Table 5-1), which – along with the observed DNA selectivity (Table 5-S3) – are characteristics that often are observed for ONs modified with intercalating pyrene moieties.[22,32-35] Surprisingly, **S**-modified ONs display significantly lower affinity toward cDNA/cRNA than **N**- and **O**-modified ONs (T_m's between 1.5 and 7.0 °C lower against cDNA, and between 4.0 and 6.0 °C lower against cRNA, Table 5-1). Clearly, pyrene intercalation is associated with an energetic penalty. Given the relatively similar conformational preferences of these monomers, we speculate that the decreased cDNA/cRNA affinity of **S**-modified ONs either is due to steric interference of the larger 2′-thio group or disruption of hydration layers at the brim of the minor groove.

Table 5-1: ΔT_m's of duplexes between **B1-B6** and cDNA/cRNA.[a]

			ΔT_m [°C]					
			+ cDNA			+ cRNA		
ON	Sequence	**B** =	**S**	**O**[b]	**N**[c]	**S**	**O**[b]	**N**[c]
B1	5′-G**B**G ATA TGC		+2.0	+5.0	+5.0	-7.0	-2.0	-2.0
B2	5′-GTG A**B**A TGC		+8.0	+12.5	+15.0	-2.0	+4.0	+3.0
B3	5′-GTG ATA **B**GC		+5.0	+8.0	+9.0	-4.5	±0.0	-0.5
B4	3′-CAC **B**AT ACG		±0.0	+3.5	+1.5	-10.0	-4.5	-6.5
B5	3′-CAC TA**B** ACG		+8.0	+11.5	+15.0	-2.0	+2.5	+3.0
B6	3′-CAC **B**A**B** ACG		+8.0	+14.0	+14.0	-10.5	-1.0	-3.0

[a]ΔT_m = change in T_m relative to reference duplexes **D1**:**D4** (T_m = 29.5 °C), **D1**:**R4** (T_m = 27.5 °C) or **R1**:**D4** (T_m = 27.5 °C), where **D1**: 5′-GTG ATA TGC, **D4**: 3′-CAC TAT ACG, **R1**: 5′-GUG AUA UGC and **R4**: 3′-CAC UAU ACG. T_m's are determined as the maximum of the first derivative of melting curves (A_{260} vs T) recorded in medium salt buffer ($[Na^+]$ = 110 mM, $[Cl^-]$ = 100 mM, pH 7.0 (NaH_2PO_4/Na_2HPO_4)), using 1.0 µM of each strand. T_m's are averages of at least two measurements within 1.0 °C. A = adenin-9-yl DNA monomer, C = cytosin-1-yl DNA monomer, G = guanin-9-yl DNA monomer and T = thymin-1-yl DNA monomer. For structures of monomers **S, O, N** see Figure 5-1.

[b]From reference 24

[c]From reference 35.

5.2.5 Binding Specificity

The binding specificity of centrally modified ONs (**B2**-series) was studied using DNA strands with mismatched nucleotides opposite to the pyrene-functionalized monomer (Table 5-2). As expected for intercalator-modified ONs,[36,37] the **B2**-series discriminates mismatched DNA targets less efficiently than reference strand **D1**. While **S2** and **O2** display similar binding fidelity, significantly better mismatch discrimination is observed for **N2**. Additional binding specificity data are discussed in the Supporting Information (Tables 5-S4 and 5-S5).

Table 5-2: Discrimination of mismatched DNA targets by **S2**/**O2**/**N2** and reference strands.[a]

			DNA: 3′-CAC T**B**T ACG			
			T_m [°C]	ΔT_m [°C]		
ON	**Sequence**	**B=**	**A**	C	G	T
D1	5′-GTG ATA TGC		**29.5**	-16.5	-9.5	-17.0
S2	5′-GTG A**S**A TGC		**37.5**	-15.0	-3.0	-7.0
O2[b]	5′-GTG A**O**A TGC		**42.0**	-13.0	-5.0	-6.5
N2[b]	5 -GTG A**N**A TGC		**44.5**	-23.0	-3.5	-13.0

[a] For conditions of thermal denaturation experiments, see Table 5-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched duplex

[b] From reference 24.

5.2.6 Photophysical Properties of S-modified ONs

UV-vis absorption and steady-state fluorescence emission spectra of **S**-modified ONs in the presence or absence of complementary DNA/RNA targets were recorded to further ascertain the binding mode of the attached pyrene moiety, as intercalation is known to induce bathochromic shifts of pyrene absorption due to ground-state electronic interactions with nucleobases,[38] along with nucleobase-mediated quenching of pyrene fluorescence.[39] Indeed, hybridization of **S**-modified ONs with cDNA and cRNA generally results in minor bathochromic shifts of the pyrene absorption maxima (Table 5-3 and Figure 5-S2). Smaller bathochromic shifts than with **O-** or **N-**modified ONs are observed, which indicates weaker electronic interactions with nucleobases and, hence, less pronounced intercalation.

Table 5-3: Absorption maxima in the 300-400 nm region for **S**/**O**/**N**-modified ONs and the corresponding duplexes with cDNA/cRNA.[a]

		λ_{max} [$\Delta\lambda_{max}$] /nm								
	B=	**S**			**O**[b]			**N**[b]		
ON	Sequence	SSP	+cDNA	+cRNA	SSP	+cDNA	+cRNA	SSP	+cDNA	+cRNA
B1	5′-G**B**G ATA TGC	355	354 [±0]	354 [-1]	350	353 [+3]	352 [+2]	349	353 [+4]	351 [+2]
B2	5′-GTG A**B**A TGC	353	354 [+1]	357 [+4]	348	353 [+5]	352 [+4]	348	353 [+5]	351 [+3]
B3	5′-GTG ATA **B**GC	353	354 [+1]	356 [+3]	350	353 [+3]	352 [+2]	349	353 [+4]	354 [+5]
B4	3′-CAC **B**AT ACG	353	357 [+4]	356 [+3]	350	352 [+2]	352 [+2]	349	354 [+5]	349 [±0]
B5	3′-CAC TA**B** ACG	352	354 [+2]	356 [+4]	349	352 [+2]	352 [+3]	348	354 [+6]	352 [+4]
B6	3′-CAC **B**A**B** ACG	353	357 [+4]	355 [+2]	---	---	---	348	352 [+4]	347 [-1]

[a]Measurements were performed at 5 °C (O, N) or 10 °C (S) using a spectrophotometer and quartz optical cells with 1.0 cm path lengths. Buffer conditions are as in thermal denaturation experiments.

[b]From reference 24 and 35.

Steady-state fluorescence emission spectra of **S**-modified ONs and the corresponding duplexes with cDNA/cRNA display two vibronic bands at λ_{em} = 383±1 nm and 401±2 nm, respectively, as well as a small shoulder at ~420 nm. As expected for intercalating pyrene moieties, the fluorescence intensity decreases upon hybridization with DNA/RNA targets (Figure 5-2).

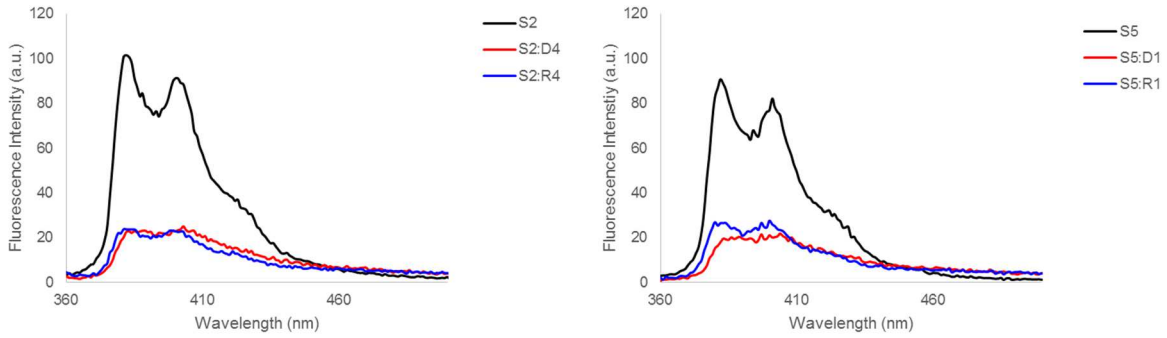


Figure 5-2: Steady-state fluorescence emission spectra of representative **S**-modified ONs and the corresponding duplexes with cDNA/cRNA. Spectra were recorded at T = 10 °C using λ_{ex} = 350 nm. Each strand was used at 1.0 μM concentration in T_m buffer.

5.2.7 Biophysical Properties of Duplexes with Interstrand Zippers of Monomer S

Next, we set out to study DNA duplexes with different interstrand arrangements of **S** monomers to estimate their potential for dsDNA recognition via dual invasion. In line with our previous observations with intercalator-functionalized probes,[20-22,35] duplexes with +1 interstrand zippers[†] of **S** monomers are more thermolabile and energetically activated than probes with other zipper configurations (compare T_m's for **S2**:**S5** relative to other probe duplexes, Table 5-4). The activated nature of **S2**:**S5** was verified through analysis of thermodynamic parameters, which were derived from denaturation curves via line fitting.[40] Thus, formation of **S2**:**S5** is considerably less favorable than formation of reference duplexes or probe duplexes with other **S**-zippers (compare $\Delta\Delta G^{293}$ values in third ΔG^{293} column, Table 5-4). The energetic nature of **S2**:**S5**, which is weakly enthalpic in origin ($\Delta\Delta H$ = +11 kJ/mol, Table 5-S6), presumably

[†]For an explanation of zipper nomenclature, refer to the Experimental Section

reflects the fact that the nearest neighbor exclusion principle is violated in duplexes with +1 interstrand zippers of **S** monomers.

The activated nature of **S2**:**S5** becomes even clearer when estimating the binding energy for recognition of iso-sequential dsDNA targets as ΔG_{rec}^{293} (ON_A:ON_B) = ΔG^{293} (ON_A:cDNA) + ΔG^{293} (cDNA:ON_B) - ΔG^{293} (ON_A:ON_B) - ΔG^{293} (dsDNA) where ON_A:ON_B is a duplex with an interstrand zipper arrangement of monomers. Probes that are energetically activated for recognition of dsDNA via the process depicted in Figure 5-1, will display negative ΔG_{rec}^{293} values since the products of the recognition process (i.e., probe-target duplexes) are more stable than the reactants (i.e., double-stranded probes and target duplexes). Indeed, much lower ΔG_{rec}^{293} values are observed for **S2**:**S5** than for other probe duplexes (ΔG_{rec}^{293} trend: **S2**:**S5**<<**S2**:**S4**≤**S1**:**S2**<**S1**:**S5**, Table 5-4).

Contrary to our initial hypothesis, **S2**:**S5** is less activated for dsDNA recognition than **O2:O5** (ΔG_{rec}^{293}= -29 kJ/mol)[21] or **N2:N5** (ΔG_{rec}^{293} = -40 kJ/mol)[35], in large part because the probe-target duplexes are less stable (ΔG^{293} for **S2**:cDNA and **S5**:cDNA = -10 kJ/mol, Table 5-4, compared to ΔG^{293} for **O2**:cDNA, **O5**:cDNA, **N2**:cDNA and **N5**:cDNA = -14, -12, -20 and -19 kJ/mol, respectively[21,35]).

Table 5-4: Biophysical properties of **S**-modified probe duplexes.[a]

ON	ZP	Sequence	T_m (°C)	$\Delta G^{293}[\Delta\Delta G^{293}]$ (kJ/mol)			ΔG^{293}_{rec} (kJ/mol)	λ_{max} (nm)
				upper ON vs cDNA	lower ON vs cDNA	probe duplex		
S1 **S5**	**+4**	5′-G**S**G ATA TGC 3′-CAC TA**S** ACG	40.5	-49±0 [-4]	-55±1 [-10]	-59±1 [-14]	±0	353
S1 **S4**	**+2**	5′-G**S**G ATA TGC 3′-CAC **S**AT ACG	30.5	-49±0 [-4]	-48±0 [-3]	-46±0 [-1]	-6	353
S2 **S5**	**+1**	5′-GTG A**S**A TGC 3′-CAC TA**S** ACG	**25.5**	**-55±1 [-10]**	**-55±1 [-10]**	**-44±1 [+1]**	**-21**	**351**
S2 **S4**	**-1**	5′-GTG A**S**A TGC 3′-CAC **S**AT ACG	33.5	-55±1 [-10]	-48±0 [-3]	-50±0 [-5]	-8	355

[a] ZP = zipper. For conditions of thermal denaturation and absorption experiments, see Table 5-1 and Table 5-3, respectively. $\Delta\Delta G^{293}$ is measured relative to ΔG^{293} for **D1**:**D4** = -45 kJ/mol. ΔG^{293}_{rec} (**ON$_A$**:**ON$_B$**) = ΔG^{293} (**ON$_A$**:cDNA) + ΔG^{293} (cDNA:**ON$_B$**) - ΔG^{293} (**ON$_A$**:**ON$_B$**) - ΔG^{293} (dsDNA). "±" denotes standard deviation.

Another unique characteristic of DNA duplexes with +1 interstrand arrangements of intercalator-modified nucleotide monomers[21,34,35] that also is observed for **S2**:**S5**, is the blue-shifted pyrene absorption, which is indicative of reduced pyrene-nucleobase interactions due to a locally perturbed duplex geometry (compare λ_{max} for **S2**:**S5** with λ_{max} for other probe duplexes (Table 5-4 and Figure 5-S4) or probe-target duplexes (Table 5-3)). Moreover, steady-state fluorescence emission spectra of **S2**:**S5** (and of +2 zipper probe **S1**:**S4**) exhibit prominent and unstructured emission at λ_{em} ~490 nm, which is consistent with the formation of pyrene-pyrene excimers. Probe duplexes with other zipper arrangements do not exhibit emission at λ_{em} ~490 nm (Figure 5-3).[18,20,22,34] Based on our previously published molecular modeling structures of **O2**:**O5**,[21] we hypothesize that the two pyrene moieties of **S2**:**S5** co-stack inside the duplex core leading to excimer formation, while the excimer signal of **S1**:**S4** most likely arises because of pyrene stacking in the major groove as suggested for other probes with +2 zipper arrangements of intercalator-modified nucleotides.[41]

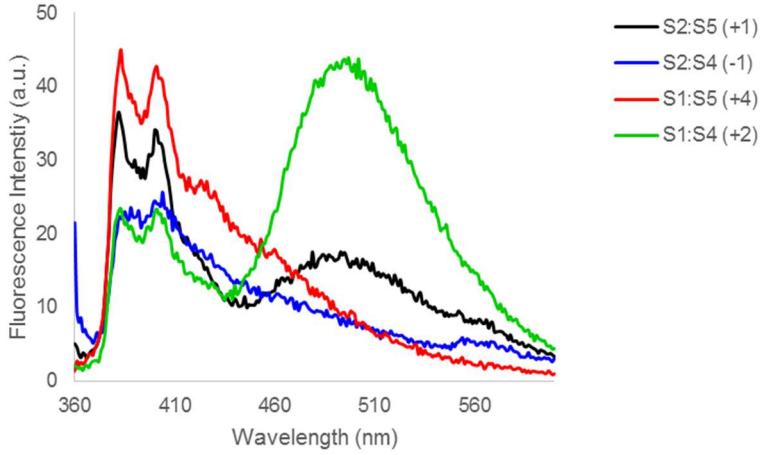


Figure 5-3: Steady-state fluorescence emission spectra of DNA duplexes with different interstrand monomer arrangements of **S**. For experimental conditions, see Figure 5-2.

5.2.8 Recognition of DNA using Activated Probe Duplexes

Our previous studies have shown that efficient dsDNA recognition requires probes that are strongly energetically activated ($\Delta G_{rec}^{293} << 0$ kJ/mol).[20,21,34,35] We therefore set out to evaluate the dsDNA-targeting efficiency of **S2**:**S5** using a 3′-digoxigenin (DIG) labeled DNA hairpin (DH) as a model dsDNA target, which is comprised of a 9-mer double-stranded mixed sequence stem linked by a T_{10} loop (Figure 5-4). However, incubation of **DH1** with **S2**:**S5** in a HEPES buffer for 12-16 hours at ambient temperature did not result in formation of slower-migrating recognition complexes on non-denaturing PAGE gels, even at the highest tested probe concentration (Figure 5-4). This contrasts our results with **O2**:**O5**[21] and **N2**:**N5**[35], which result in ~50% dsDNA recognition when used at ~20-fold excess, but is in line with the relatively low dsDNA-targeting potential of **S2**:**S5** as judged by the ΔG_{rec}^{293} values. A similar outcome was observed when **DH1** was annealed in the presence of **S2**:**S5** followed by room temperature

incubation (Figure 5-S5), which suggests that the recognition complex is not sufficiently stable at these experimental conditions. Thus, the results strongly suggest that Invaders based on 2'-*O*-(pyren-1-yl)methyluridine monomer **O** or 2′-amino-2′-deoxy-2′-*N*-(pyren-1-ylmethyl)-2′-*N*-methyl-uridine monomer **N**, but not 2′-thio-2′-deoxy-2′-*S*-(pyren-1-ylmethyl)uridine monomer **S**, are suitable for dsDNA recognition.

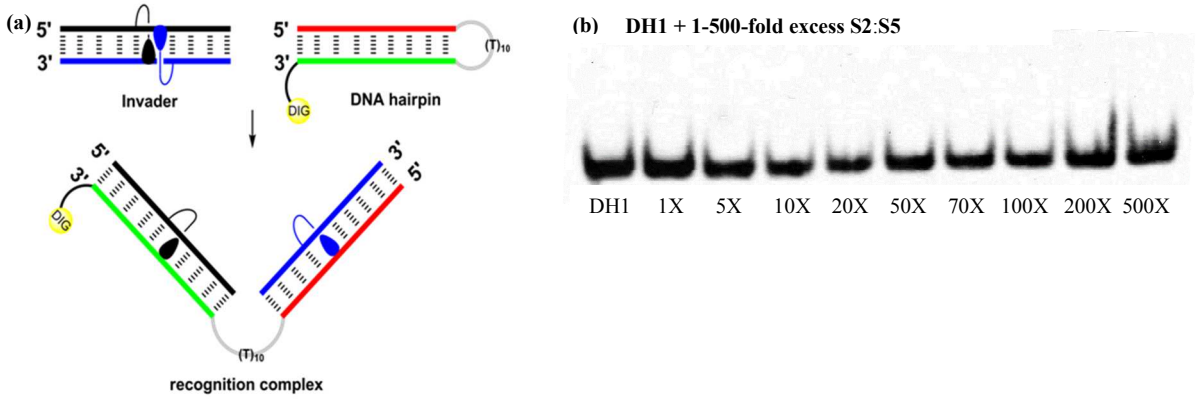


Figure 5-4: Attempted recognition of model dsDNA target **DH1** using Invader **S2**:**S5**. (a) Illustration of recognition process. Sequence of **DH1**: 5′-GTGATATGC-(T_{10})-GCTTATCAC-DIG-3′. (b) Representative electrophoretogram from recognition of **DH1** using 1-500 fold excess of **S2**:**S5** incubated. Experimental conditions for electrophoretic mobility shift assay: separately pre-annealed targets (34.4 nM) and **S2**:**S5** (variable concentration) were incubated at ambient temperature for 12-16h in 1X HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 10% sucrose, 1.4 mM spermine tetrahydrochloride, pH 7.2) and then run on 16% non-denaturing PAGE (performed at 70V, 2.5h, ~4 °C) using 0.5x TBE as a running buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA); DIG: digoxigenin.

5.3 Conclusions

A short, high yielding synthetic route to 2′-thio-2′-deoxy-2′-*S*-(pyren-1-yl)methyluridine has been developed. Pseudorotational analysis of these nucleosides indicates that the furanose ring predominantly adopts a *South*-type conformation, whereas the furanose ring of the corresponding 2'-*O*-(pyren-1-yl)methyluridine (nucleoside **O**) displays more equal distributions of *North*- and *South*- conformations; the furanose ring of 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-aminouridine (nucleoside **N**) exclusively adopts *South*-type conformation, suggesting that the N2′-methyl substituent overrides the stereoelectronic influence from the decreased electronegativity of the 2′-substituent. Thermal denaturation experiments show that **S**-modified ONs display moderately increased affinity toward cDNA compared to reference strands, but less so than **O**- and **N**- modified ONs. The lower cDNA affinity is most likely a result of steric interference of the large 2′-thio group or perturbation of the hydration spine in the minor groove. The moderately increased cDNA affinity of **S**- modified ONs is a result of pyrene intercalation as indicated by bathochromic pyrene absorption shifts, reduced fluorescence emission, and thermodynamic stabilization upon duplex formation. Although **S**-modified double-stranded probes with +1 zipper arrangements are energetically activated (ΔG_{rec}^{293} = -21 kJ/mol), these probes were not able to recognize dsDNA model targets, indicating 2′-thio-2′-deoxy-2′-*S*-(pyren-1-ylmethyl)uridine monomer **S** is not suitable for dsDNA recognition via the Invader approach.

5.4 Experimental Section

5.4.1 Synthesis of 2′-deoxy-5′-*O*-(4,4′-dimethoxytrityl)-2′-thio-2′-*S*-(pyren-1-yl)methyl uridine Phosphoramidite

2′-Deoxy-2′-*S*-(pyren-1-yl)methyl-2′-thiouridine (2). To a flame-dried round-bottomed flask flushed with argon, was added nucleoside **1**[27a] (250 mg, 0.96 mmol), 1-pyrenylmethyl chloride (360 mg, 1.44 mmol), anhydrous DMF (6 mL), and anhydrous Et_3N (200 μL, 1.44 mmol). The reaction mixture was stirred at room temperature under an argon atmosphere for 17h whereupon it was diluted with EtOAc (100 mL) and washed with brine (2 x 30 mL), saturated aqueous $NaHCO_3$ (30 mL) and H_2O (30 mL). The organic layer was dried (Na_2SO_4) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-7 % MeOH in CH_2Cl_2, v/v) to afford nucleoside **2** as a pale yellow solid (290 mg, 64%). R_f = 0.3 (10 % MeOH in CH_2Cl_2); MALDI-HRMS 497.1148 ($[M+Na]^+$, $C_{26}H_{22}N_2O_5S{\cdot}Na^+$, Calcd. 497.1142); ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.40 (br s, 1H, ex, NH), 8.37-8.40 (d, 1H, J = 9.3 Hz, Py), 8.28-8.31 (2d, 2H, J = 7.8 Hz, 7.5 Hz, Py), 8.21-8.24 (d, 1H, J = 9.3 Hz, Py), 8.18-8.21 (d, 1H, J = 7.8 Hz, Py), 8.15-8.18 (d, 1H, J = 9.1 Hz, Py), 8.12-8.15 (d, 1H, J = 9.1 Hz, Py), 8.06-8.10 (t, 1H, J = 7.5 Hz, Py), 7.94-7.97 (d, 1H, J = 7.8 Hz, Py), 7.79 (d, 1H, J = 8.3 Hz, H6), 6.21 (d, 1H, J = 8.5 Hz, H1′), 5.72 (d, 1H, ex, J = 5.5 Hz, 3′-OH), 5.53 (d, 1H, J = 8.5 Hz, H5), 5.10 (t, 1H, ex, J = 5.2 Hz, 5′-OH), 4.52-4.55 (d, 1H, J = 12.0 Hz, CH_2), 4.48-4.51 (d, 1H, J = 12.2 Hz, CH_2), 4.25-4.29 (m, 1H, H3′), 3.91-3.94 (m, 1H, H4′), 3.56-3.66 (m, 3H, H2′, H5′); ^{13}C NMR (75.5 MHz, DMSO-d_6) δ 162.7, 150.8, 140.1 (C6), 131.6, 130.7, 130.3, 130.2, 128.4, 127.9 (Py), 127.4 (Py), 127.3 (Py), 127.1 (Py), 126.3 (Py), 125.25 (Py), 125.17 (Py), 124.7 (Py),

124.2, 123.8, 123.5 (Py), 102.3 (C5), 87.6 (C1′), 86.6 (C4′), 71.9 (C3′), 61.3 (C5′), 52.5 (C2′), 32.6 (CH_2Py).

2′-Deoxy-5′-*O*-(4,4′-dimethoxytrityl)-2′-*S*-(pyren-1-yl)methyl-2′-thiouridine (3). Nucleoside **2** (200 mg, 0.42 mmol) was coevaporated with anhydrous 1,2-dichloroethane (2 × 5 mL) and re-dissolved in anhydrous pyridine (5 mL). To this was added 4,4′-dimethoxytritylchloride (DMTrCl, 186 mg, 0.55 mmol) and a catalytic amount of 4-dimethylaminopyridine (DMAP, ~5 crystals). The reaction mixture was stirred at room temperature under an argon atmosphere for 19 h, when it was diluted with CH_2Cl_2 (40 mL) and washed with saturated aqueous $NaHCO_3$ (30 mL). The aqueous layer was back-extracted with CH_2Cl_2 (3 × 15 mL) and the combined organic layers were dried (Na_2SO_4), evaporated to dryness and co-evaporated with absolute EtOH and toluene (2:1 v/v, 3 × 20 mL). The resulting residue was purified by silica gel column chromatography (0-6% MeOH in CH_2Cl_2, v/v) to yield nucleoside **3** as a tan-colored foam (236 mg, 72%). R_f = 0.7 (10% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m/z* 799.2474 ($[M+Na]^+$, $C_{47}H_{40}N_2O_7S{\cdot}Na^+$, Calcd. 799.2448); ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.48 (d, 1H, ex, J = 2.0 Hz, NH), 8.38-8.41 (d, 1H, J = 9.0 Hz, Py), 8.28-8.31 (d, 2H, J = 7.5 Hz, Py), 8.21-8.24 (d, 1H, J = 9.0 Hz, Py), 8.11-8.18 (m, 3H, Py), 8.06-8.10 (t, 1H, J = 7.5 Hz, Py), 7.95-7.98 (d, 1H, J = 8.0 Hz, Py), 7.64 (d, 1H, J = 8.0 Hz, H6), 7.35-7.38 (m, 2H, DMTr), 7.29-7.33 (m, 2H, DMTr), 7.21-7.26 (m, 5H, DMTr), 6.85-6.88 (m, 4H, DMTr), 6.21 (d, 1H, J = 7.5 Hz, H1′), 5.77 (d, 1H, ex, J = 5.5 Hz, 3′-OH), 5.42 (dd, 1H, J = 8.0 Hz, 2.0 Hz, H5), 4.57-4.60 (d, 1H, J = 12.0 Hz, CH_2), 4.50-4.54 (d, 1H, J = 12.0 Hz, CH_2), 4.21-4.25 (m, 1H, H3′), 3.98-4.02 (m, 1H, H4′), 3.66-3.75 (m, 7H, H2′, CH_3O), 3.26-3.31 (m, 1H, H5′), 3.20-3.24 (m, 1H, H5′); ^{13}C NMR (125 MHz, DMSO-d_6) δ 162.7, 158.1,

150.6, 144.6, 140.1 (C6), 135.3, 135.2, 131.5, 130.7, 130.3, 130.2, 129.70 (DMTr), 129.68 (DMTr), 128.4, 127.84 (DMTr), 127.76 (Py), 127.6 (DMTr), 127.4 (Py), 127.3 (Py), 127.2 (Py), 126.7 (DMTr), 126.3 (Py), 125.3 (Py), 125.2 (Py), 124.7 (Py), 124.2, 123.8, 123.5 (Py), 113.20 (DMTr), 113.19 (DMTr), 102.2 (C5), 88.1 (C1′), 85.9, 84.5 (C4′), 71.2 (C3′), 63.5 (C5′), 55.0 (CH_3O), 52.3 (C2′), 32.6 (CH_2Py).

2′-Deoxy-3′-*O*-(*N*,*N*-diisopropylamino-2-cyanoethoxyphosphinyl)-5′-*O*-(4,4′-dimethoxytrityl)-2′-*S*-(pyren-1-yl)methyl-2′-thiouridine (4). Nucleoside **3** (160 mg, 0.21 mmol) was coevaporated with anhydrous 1,2-dichloroethane (2 × 3 mL) and re-dissolved in anhydrous CH_2Cl_2 (3 mL). To this was added anhydrous *N*,*N*-diisopropylethylamine (DIPEA, 145 μL, 0.82 mmol) followed by dropwise addition of 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (PCl reagent, 92 μL, 0.41 mmol). The reaction mixture was stirred under an argon atmosphere at room temperature for 2 h when absolute EtOH (0.5 mL) was added. Solvents were evaporated off and the resulting residue was purified by silica gel column chromatography (0-60% EtOAc in petroleum ether, v/v) followed by precipitation from CH_2Cl_2 and petroleum ether to afford nucleoside **4** as a tan-colored foam (146 mg, 73%). R_f = 0.6 (3% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m/z* 999.3487 ($[M+Na]^+$, $C_{56}H_{57}N_3O_8S{\cdot}Na^+$, Calcd. 999.3527); ^{31}P NMR (121 MHz, $CDCl_3$) δ 150.9, 149.9.

5.4.2 Protocol – Pseudorotational Analysis of Nucleosides

A Matlab-based pseudorotational analysis program[29] was used to determine *P* and ϕ_m for the furanose ring of nucleoside **3** from $^3J_{HH}$ scalar coupling constants of the endocyclic sugar protons. Appropriate electronegativity (ε_{en}) values for the β-D-ribose ring and the C2′-down

substituents were selected from the electronegativity editor interface using DMSO solvent values: OR (ε_{en} = 1.400), NR_2 (ε_{en} = 1.200) and SR (ε_{en} = 0.785). Coupling constants for $^3J_{H1'H2'}$, $^3J_{H2'H3'}$, and $^3J_{H3'H4'}$ were obtained from ^{1}H NMR (500 MHz, DMSO-d_6, 25 °C) data analysis. Two prototypical *North* and *South* conformations (P_A = 18°, P_B = 153.6°, ϕ_m = 38°, proportion of conformer A = 30-50%) were chosen as seed structures for iterative fitting of the modified Karplus-Diez-Donders equations to the experimental $^3J_{HH}$ coupling constants. The key features of the optimized conformers and their relative proportions are reported in Table 5-S1.

5.4.3 Protocol - Synthesis and Purification of ONs

Modified ONs were synthesized on a 0.2 μmol scale using a DNA synthesizer and succinyl linked LCAA-CPG (long chain alkyl amine controlled pore glass) columns with a pore size of 500Å. Standard protocols for incorporation of DNA monomers were used. Extended hand-coupling conditions (15 min) were used for incorporation of monomer **S** using 4,5-dicyanoimidazole as an activator with >95% coupling yield. Extended oxidation (45s) was used. Cleavage from solid support and removal of protecting groups was accomplished using 32% aqueous ammonia (55 °C, 12 h). ONs were purified in the DMTr-on mode via ion-pair reverse phase HPLC (C18 column) using a 0.05 M triethylammonium acetate - water/acetonitrile gradient. This was followed by detritylation (80% aq. AcOH) and precipitation (NaOAc/$NaClO_4$/acetone, -18 °C for 12-16 h). The identity of synthesized ONs was established through MALDI-MS analysis recorded in positive ion mode on a quadrupole time-of-flight tandem mass spectrometer equipped with a MALDI source using anthranilic acid, 3-hydroxypicolinic acid (3-HPA) as a matrix (Table 5-S2). Purity was verified by ion-pair reverse phase HPLC running in analytical mode (>90%).

5.4.4 Protocol - Thermal Denaturation Studies

ON concentrations were estimated using the following extinction coefficients for DNA (OD/μmol): G (12.01), A (15.20), T (8.40), C (7.05); RNA (OD/μmol): G (13.70), A (15.40), U (10.00), C (9.00); and pyrene (22.4)[42]. Strands were thoroughly mixed and denatured by heating to 70 °C, followed by cooling to the starting temperature of the experiment. Quartz optical cells with a path length of 1.0 cm were used. Thermal denaturation temperatures of duplexes (1.0 μM final concentration of each strand) were measured using a UV/vis spectrophotometer equipped with a 12-cell Peltier temperature controller and determined as the maximum of the first derivative of the thermal denaturation curve (A_{260} vs. T) recorded in medium salt phosphate buffer (T_m buffer: 100 mM NaCl, 0.1 mM EDTA; pH 7.0 adjusted with 10 mM Na_2HPO_4 and 5 mM Na_2HPO_4). The temperature of the denaturation experiments ranged from at least 15 °C below T_m to 20 °C above T_m (although not below 3 °C). A temperature ramp of 0.5 °C/min was used in all experiments. Reported T_m's are averages of two experiments within ± 1.0 °C.

5.4.5 Protocol - Determination of Thermodynamic Parameters

Thermodynamic parameters for duplex formation were determined through baseline fitting of denaturation curves using software provided with the UV/vis spectrometer. Bimolecular reactions, two-state melting behavior, and a heat capacity change of $\Delta C_p = 0$ upon hybridization were assumed.[40] A minimum of two experimental denaturation curves were each analyzed at least three times to minimize errors arising from baseline choice. Averages and standard deviations are listed.

5.4.6 Protocol - Absorption Spectra

UV-vis absorption spectra (range 200-600 nm) were recorded at 10 °C using the same samples and instrumentation as in the thermal denaturation experiments.

5.4.7 Protocol - Steady-state Fluorescence Emission Spectra

Steady-state fluorescence emission spectra of **S**-modified ONs modified and the corresponding duplexes with cDNA/cRNA, were recorded in non-deoxygenated thermal denaturation buffer (each strand at 1.0 μM concentration) and obtained as an average of five scans using an excitation wavelength of λ_{ex} = 350 nm. Excitation and emission slits of 5.0 nm and 2.5 nm, respectively were used along with a scan speed of 600 nm/min. Experiments were determined at 10 °C with N_2 flow to ascertain maximal hybridization of probes to DNA/RNA targets and minimize condensation, respectively.

5.4.8 Protocol - Electrophoretic Mobility Shift Assay

This assay was performed essentially as previously described.[23] Unmodified DNA hairpin **DH1** was obtained from commercial sources and used without further purification. The DNA hairpin was 3′-DIG-labeled using the 2nd generation DIG Gel Shift Kit (Roche Applied Bioscience) as recommended by the manufacturer. The DIG-labeled ON obtained in this manner was diluted and used without further purification in the recognition experiments. Pre-annealed probe (85 °C for 10 min, cooled to room temperature over 15 min) and DIG-labeled DNA hairpins (34.4 nM) were mixed and incubated in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 10% sucrose, 1.44 mM spermine tetrahydrochloride, pH 7.2) for the specified time at ambient temperature (~21±3 °C). The reaction mixtures were then diluted with 6x DNA loading dye

(Fermentas) and loaded onto a 16% non-denaturing polyacrylamide gel. Electrophoresis was performed using a constant voltage of 70 V for 2.5 h at ~4 °C using 0.5x TBE as a running buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA). Gels were blotted onto positively charged nylon membranes (Roche Applied Bioscience) using constant voltage with external cooling (100V, ~4 °C). The membranes were exposed to anti-digoxigenin-AP F_{ab} fragments as recommended by the manufacturer of the DIG Gel Shift Kit, transferred to a hybridization jacket, and incubated with the substrate (CSPD) in detection buffer for 10 min at 37 °C. The chemiluminescence of the formed product was captured on X-ray film, which was developed using an X-Omatic 1000A X-ray film developer (Kodak). The resulting bands were quantified using Image J software. Invasion efficiency was determined as the intensity ratio between the recognition complex band and the total lane. An average of three independent experiments is reported.

5.4.9 Explanation of Zipper Nomenclature

The following nomenclature describes the relative arrangement between two monomers positioned on opposing strands in a duplex. The number *n* describes the distance measured in number of base pairs and has a positive value if a monomer is shifted toward the 5′-side of its own strand relative to a second reference monomer on the other strand. Conversely, *n* has a negative value if a monomer is shifted toward the 3′-side of its own strand relative to a second reference monomer on the other strand.

5.5 Supporting Information

5.5.1 General Experimental Section

Reagents and solvents were commercially available, of analytical grade and used without further purification. Petroleum ether of the distillation range 60-80 °C was used. Solvents were dried over activated molecular sieves: *N*,*N*-dimethylformamide (3Å); pyridine, CH_2Cl_2, 1,2-dichloroethane, triethylamine, and *N,N*-diisopropylethylamine (4Å). Water content of anhydrous solvents was verified on Karl-Fisher apparatus. Reactions were monitored by TLC using silica gel coated plates with a fluorescence indicator (SiO_2-60, F-254) which were visualized a) under UV light and/or b) by dipping in 5% conc. H_2SO_4 in absolute ethanol (v/v) followed by heating. Silica gel column chromatography was performed with Silica gel 60 (particle size 0.040–0.063 mm) using moderate pressure (pressure ball). Evaporation of solvents was carried out under reduced pressure at temperatures below 45 °C. After column chromatography, appropriate fractions were pooled, evaporated and dried at high vacuum for at least 12 h to give the obtained products in high purity (>95%) as ascertained by 1D NMR techniques. Exchangeable (ex) protons were detected by disappearance of signals upon D_2O addition. Assignments of NMR spectra are based on 2D spectra (COSY, HSQC) and DEPT spectra. Quaternary carbons are not assigned in ^{13}C NMR but verified from HSQC and DEPT spectra (absence of signals). MALDI-HRMS spectra of compounds were recorded on a mass spectrometer using 2,5-dihydroxybenzoic acid as a matrix and polyethylene glycol (PEG 600) as an internal calibration standard.

5.5.2 Additional Tables, Figures, and Discussion

Table 5-S1: $^{3}J_{HH}$ scalar coupling constants, pseudorotation phase angles (P) and puckering amplitudes (ϕ_m) for nucleoside **3** (monomer **S**) and the corresponding nucleosides for monomer **O** and **N**.

Monomer	**S**	**O**	**N**
$J_{1'-2'}$	7.5	4.1[a]	8.3[a]
$J_{2'-3'}$	6.0	4.9[a]	5.7[a]
$J_{3'-4'}$	3.9	5.8[a]	2.3[a]
Conformation I			
P	56.5	11.0	121.6‡
ϕ_m	26.9	38.4	27.4
Conformation II			
P	143.0	130.3	145.1‡
ϕ_m	37.5	33.4	37.8
% Conformation II	61	49	61

‡The Matlab script was unable to fit the data for this nucleoside to a *North* type conformations, despite different input conditions. Invariably, two *South* type conformations were predicted.

[a]Data from reference 24.

Table 5-S2: MALDI-MS of ONs modified with monomer **S**.[a]

ONs	Sequence	Calc. *m/z* $[M+H]^+$	Found *m/z* $[M+H]^+$
S1	5′-G**S**G ATA TGC	2985.5	2985.8
S2	5′-GTG A**S**A TGC	2985.5	2985.8
S3	5′-GTG ATA **S**GC	2985.5	2984.7
S4	3′-CAC **S**AT ACG	2914.5	2914.8
S5	3′-CAC TA**S** ACG	2914.5	2914.8
S6	3′-CAC **S**A**S** ACG	3146.6	3146.9

[a]For structure of monomer **S** see Figure 5-1 in the main manuscript.

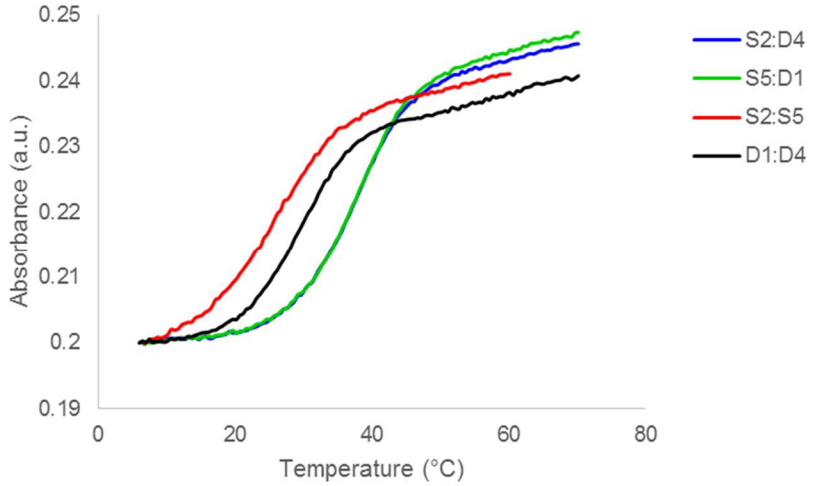


Figure 5-S1: Representative thermal denaturation curves of **S**-modified duplexes and reference duplex **D1**:**D4**. For experimental conditions, see Table 5-1.

Additional Discussion Regarding DNA-selectivity

The DNA selectivity, defined as $\Delta\Delta T_m$ (DNA-RNA) = ΔT_m (vs cDNA) - ΔT_m (vs cRNA) > 0 °C of **S**-modified ONs ($\Delta\Delta T_m$ from +9.0 to +16.5 °C, Table 5-S3) is very similar to the DNA selectivity observed for **O**- and **N**-modified ONs, indicating that the pyrene moiety of **S**-modified ONs likely is situated in a similar environment as in **O**- and **N**-modified ONs.

Table 5-S3: DNA selectivity of **B1-B6**.[a]

			$\Delta\Delta T_m$/mod (DNA-RNA) [°C]		
ON	**Duplex**	**B =**	**S**	**O**[b]	**N**[b]
B1	5′-G**B**G ATA TGC		+9.0	+7.0	+9.5
B2	5′-GTG A**B**A TGC		+10.0	+8.5	+12.5
B3	5′-GTG ATA **B**GC		+9.5	+8.0	+8.5
B4	3′-CAC **B**AT ACG		+10.0	+8.0	+8.5
B5	3′-CAC TA**B** ACG		+10.0	+9.0	+12.5
B6	3′-CAC **B**A**B** ACG		+16.5	+15.0	+18.5

[a] DNA selectivity defined as $\Delta\Delta T_m$ (DNA-RNA) = ΔT_m (vs cDNA) - ΔT_m (vs cRNA).
[b] From reference 24.

Additional Discussion of DNA/RNA Binding Specificity

Reference strand **D1** displays the expected specificity profile against mismatched RNA targets, i.e., a) mismatched duplexes have lower T_m's than matched duplexes, and b) T:rG is the least efficiently discriminated mismatched base pair (Table 5-S4). Singly modified **S2** and benchmark ONs **O2** and **N2** display comparable binding specificity except for the T:rG mismatched base pair.

The binding specificities of ONs with two next-nearest neighbor modifications (**B6**-series) were determined using DNA targets with a mismatched nucleotide opposite of the central 2'-deoxyriboadenosine (Table 5-S5). We have previously shown that **O6** and **N6** display surprisingly good binding fidelity compared to the poor binding specificities of **O2** and **N2** (Table 5-2). Interestingly, **S6** discriminates the mismatched DNA targets less efficiently than **O6** and **N6**, which suggests that there are limitations to this strategy for improving binding specificity.

Table 5-S4: Discrimination of mismatched RNA targets by **B2**-series and reference strands.[a]

			RNA: 3′-CAC U**B**U ACG			
			T_m[°C]	ΔT_m[°C]		
ON	**Sequence**	**B**=	**A**	C	G	U
D1	5′ – GTG ATA TGC		**27.5**	<-17.5	-4.5	<-17.5
S2	5′ – GTG A**S**A TGC		**25.5**	<-15.5	-5.5	-8.5
O2[b]	5′ – GTG A**O**A TGC		**31.5**	-17.5	-3.5	-9.5
N2[b]	5′ – GTG A**N**A TGC		**30.5**	-16.5	-0.5	-13.0

[a] For conditions of thermal denaturation experiments, see Table 1. T_m-values of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched DNA:RNA duplex.

[b] From reference 24.

Table 5-S5: Discrimination of mismatched DNA targets by **S6**/**O6**/**N6** and reference strands.[a]

			DNA: 5′-GTG A**B**A TGC			
			T_m[°C]	ΔT_m[°C]		
ON	**Sequence**	**B**=	**T**	A	C	G
D4	3′-CAC TAT ACG		**29.5**	-17.0	-15.5	-9.0
S6	3′-CAC **S**A**S** ACG		**37.5**	-12.5	-12.0	-12.0
O2[b]	3′-CAC **O**A**O** ACG		**43.5**	-24.0	-17.0	-14.0
N2[b]	3′-CAC **N**A**N** ACG		**43.5**	-21.5	-10.5	-13.5

[a]For conditions of thermal denaturation experiments, see Table 5-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched duplex

[b]From reference 24.

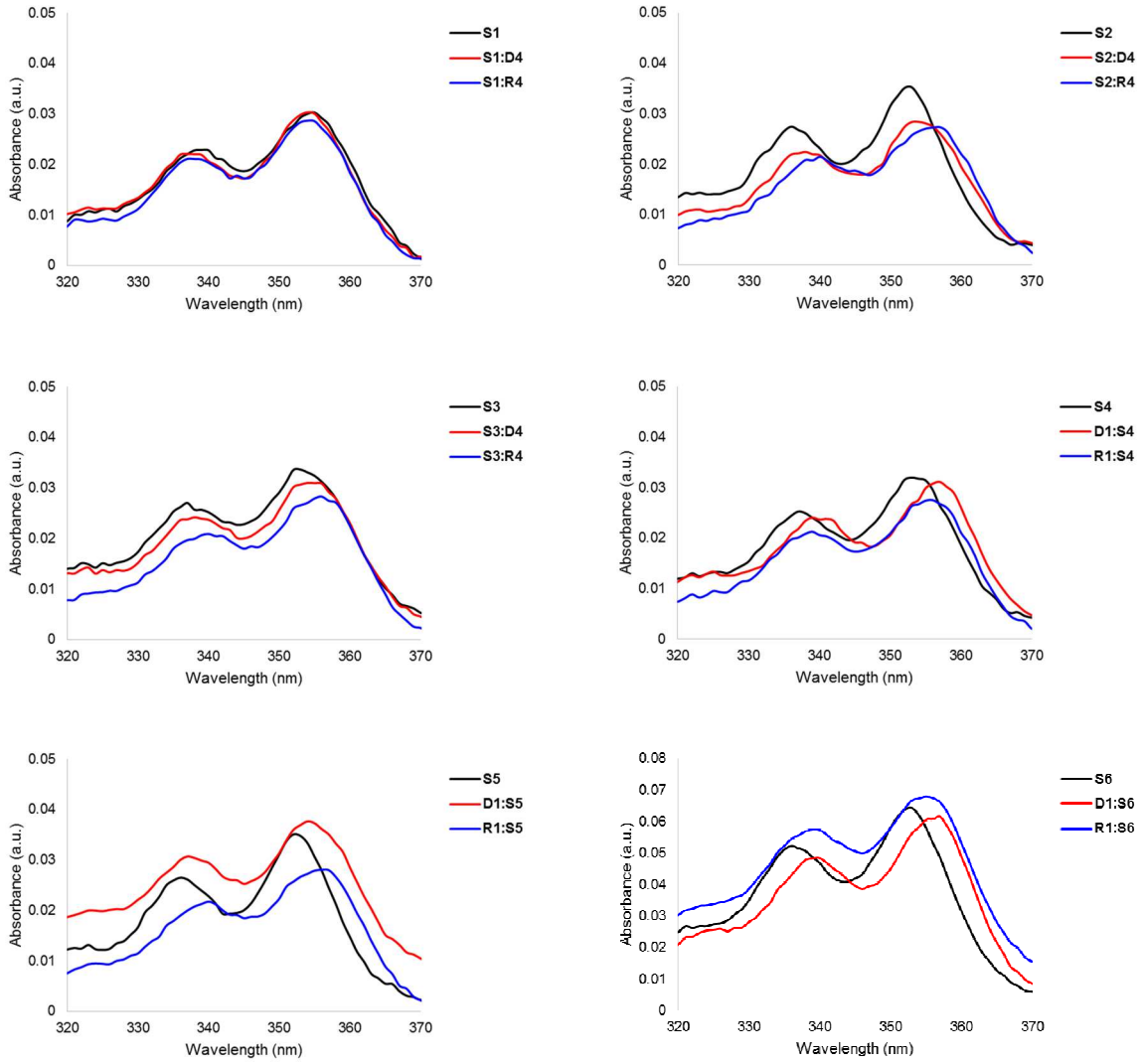


Figure 5-S2: Absorption spectra of single-stranded **S1**-**S6** and the corresponding duplexes with cDNA/cRNA targets. Spectra were recorded at T = 10 °C using each strand at 1.0 μM concentration in T_m buffer. Note that a different Y-axis scale is used for **S6**.

Discrimination of Mismatched DNA Targets via Fluorescence

Neither singly or doubly **S**-modified ONs are useful for fluorescent discrimination of DNA targets with mismatched nucleotides opposite of monomer **S** (Figure 5-S3). Thus, **S**-modified ONs have limited potential as probes for fluorescent discrimination single nucleotide polymorphisms (SNPs).

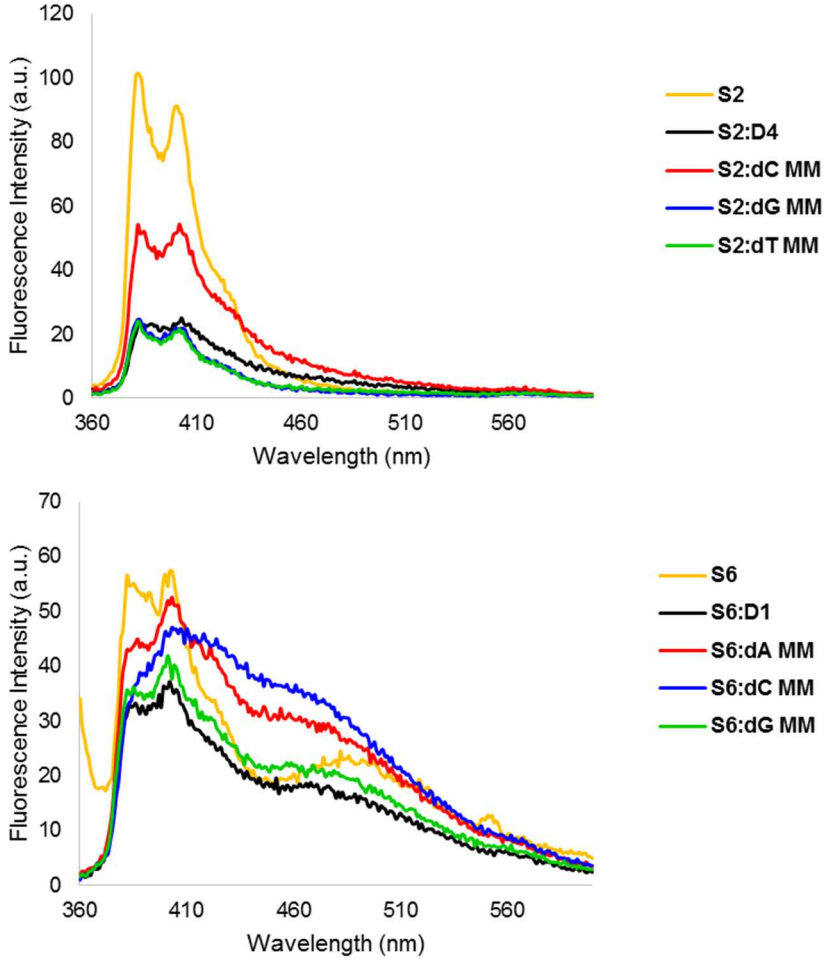


Figure 5-S3: Steady-state fluorescence emission spectra of **S2** and **S6** against matched and mismatched DNA targets. Spectra were recorded at T = 10 °C using λ_{ex} = 350 nm and each strand at 1.0 μM concentration in T_m buffer.

Table 5-S6: Change in enthalpy upon duplex formation (ΔH) and reaction enthalpy during recognition of iso-sequential dsDNA target **D1**:**D4** (ΔH_{rec}).[a]

ON	ZP	Sequence	ΔH [$\Delta\Delta H$] (kJ/mol) upper ON vs cDNA	lower ON vs cDNA	probe duplex	ΔH_{rec} (kJ/mol)
S1 **S5**	**+4**	5′-G**S**G ATA TGC 3′-CAC TA**S** ACG	-308±3 [-22]	-331±9 [-45]	-356±8 [-70]	+3
S1 **S4**	**+2**	5′-G**S**G ATA TGC 3′-CAC **S**AT ACG	-308±3 [-22]	-314±4 [-28]	-260±6 [+26]	-76
S2 **S5**	**+1**	5′-GTG A**S**A TGC 3′-CAC TA**S** ACG	**-320±6 [-34]**	**-331±9 [-45]**	**-275±10 [+11]**	**-90**
S2 **S4**	**-1**	5′-GTG A**S**A TGC 3′-CAC **S**AT ACG	-320±6 [-34]	-314±4 [-28]	-316±4 [-30]	-32

[a] $\Delta\Delta H$ is measured relative to ΔH for **D1**:**D4** = -286 kJ/mol. $\Delta H_{rec} = \Delta H$ (upper ON vs cDNA) + ΔH (lower ON vs cDNA) - ΔH (probe duplex) - ΔH (**D1**:**D4**). "ZP" and "±" denotes zipper and standard deviation, respectively.

Table 5-S7: Change in entropy at 293K upon duplex formation ($-T^{293}\Delta S$) and reaction entropy during recognition of iso-sequential dsDNA target **D1**:**D4** ($-T^{293}\Delta S_{rec}$).[a]

Duplex	ZP	Sequence	$-T^{293}\Delta S$ [$\Delta(T^{293}\Delta S)$] (kJ/mol) upper ON vs cDNA	lower ON vs cDNA	probe duplex	$-T^{293}\Delta S_{rec}$ (kJ/mol)
S1 **S5**	**+4**	5'-G**S**G ATA TGC 3'-CAC TA**S** ACG	259±3 [+19]	275±9 [+35]	297±8 [+57]	-3
S1 **S4**	**+2**	5'-G**S**G ATA TGC 3'-CAC **S**AT ACG	259±3 [+19]	266±4 [+26]	214±5 [-26]	+71
S2 **S5**	**+1**	5'-GTG A**S**A TGC 3'-CAC TA**S** ACG	**265±6 [+25]**	**275±9 [+35]**	**231±9 [-9]**	**+69**
S2 **S4**	**-1**	5'-GTG A**S**A TGC 3'-CAC **S**AT ACG	265±6 [+25]	266±4 [+26]	265±4 [+25]	+26

[a] $\Delta(T^{293}\Delta S)$ is measured relative to $-T^{293}\Delta S$ for **D1**:**D4** = 240 kJ/mol. $-T^{293}\Delta S_{rec} = \Delta(T^{293}\Delta S)$ (upper ON vs cDNA) + $\Delta(T^{293}\Delta S)$ (lower ON vs cDNA) - $\Delta(T^{293}\Delta S)$ (probe duplex).

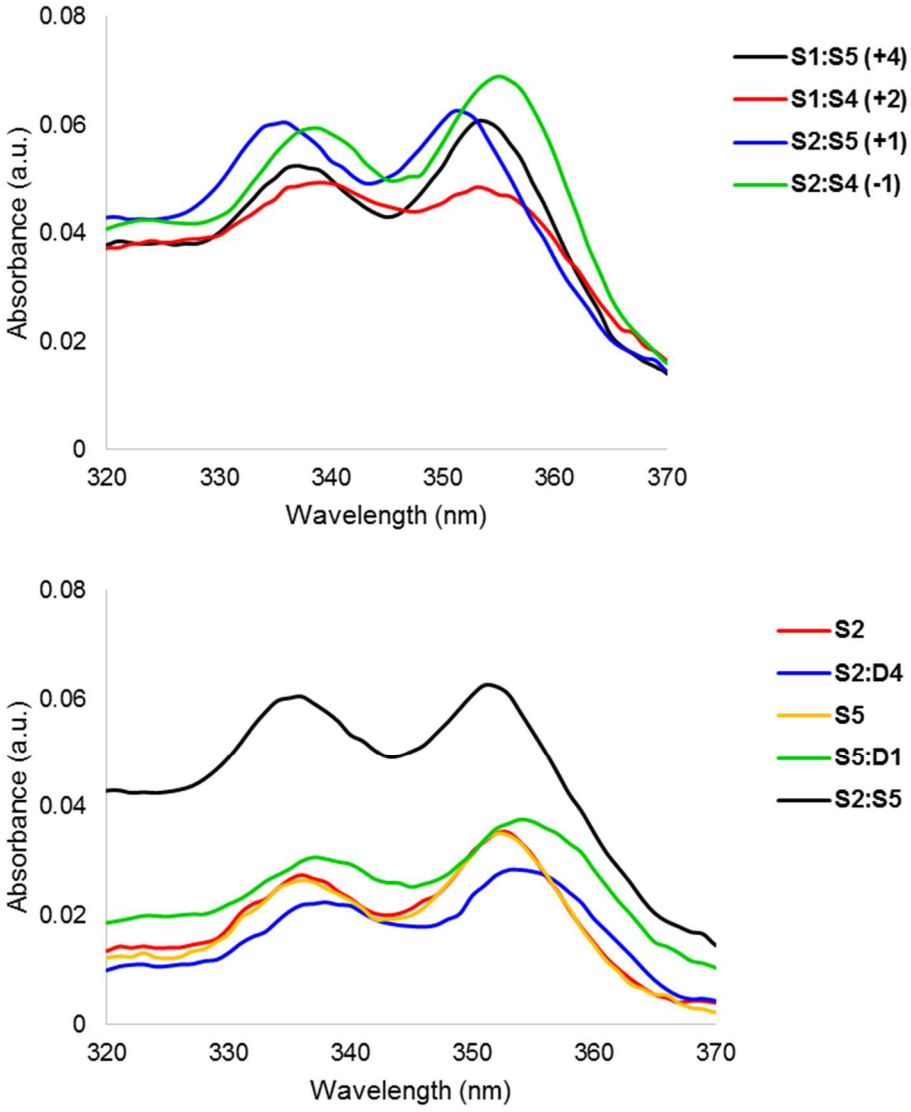


Figure 5-S4: Absorption spectra of representative Invaders, duplexes between probe strands and cDNA, and single-stranded probes. Recorded at T = 10 °C using each strand at 1.0 μM concentration in T_m buffer.

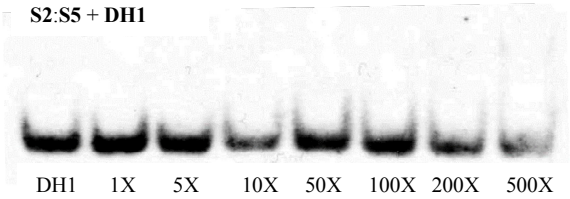


Figure 5-S5: Representative gel electrophoretogram from recognition experiments in which target **DH1** (34.4 nM) was annealed together with **S2**:**S5** (variable concentrations) at 85 °C for 15 min followed by slow cooling to room temperature over ~30 min and incubation at ambient temperature for 12-16 h. Experimental conditions otherwise as specified in Figure 5-4.

5.6 References

(1) Rogers, F. A.; Lloyd, J. A.; Glazer, P. M. *Curr. Med. Chem.: Anti-Cancer Agents* **2005**, 5, 319 − 326.

(2) Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. *J. Mol. BioSyst*. **2006**, 2, 551 − 560.

(3) Nielsen, P. E. *ChemBioChem* **2010**, 11, 2073 − 2076.

(4) Mukherjee, A.; Vasquez, K. M. *Biochimie* **2011**, 93, 1197 − 1208.

(5) Aiba, Y.; Sumaoka, J.; Komiyama, M. *Chem. Soc. Rev.* **2011**, 40, 5657 − 5668.

(6) Vaijayanthi, T.; Bando, T.; Pandian, G. N.; Sugiyama, H. *ChemBioChem* **2012**, 13, 2170 − 2185.

(7) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, 36, 5123 – 5138.

(8) (a) Nielsen, P. E., Egholm, M., Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497 – 1500.. (b) Bentin, T.; Nielsen, P. E. *Biochemistry* **1996**, *35*, 8863 – 8869. (c) Kaihatsu, K.;

Braasch, D. A.; Cansizoglu, A.; Corey, D. R. *Biochemistry* **2002**, *41*, 11118 – 11125. (d) Kaihatsu, K.; Janowski, B. A.; Corey, D. R. *Chem. Biol*. **2004**, 11, 749 - 758.

(9) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol*. **2003**, 13, 284 – 299; (b) Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem*. **2013**, 21, 6101 – 6114.

(10) (a) Bogdanove, A. J.; Voytas, D. F. *Science* **2011**, 333, 1843 − 1846. (b) Gaj, T.; Gersbach, C. A.; Barbas, C. F., III. *Trends Biotechnol.* **2013**, 31, 397 – 405.

(11) Kutyavin, I. V.; Rhinehart, R. L.; Lukhtanov, E. A.; Gorn, V. V.; Meyer, R. B., Jr.; Gamper, H. B., Jr. *Biochemistry* **1996**, 35, 11170 − 11176.

(12) (a) Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 11804 − 11808. (b) Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corradini, R.; Sforza, S.; Komiyama, M. *Nucleic Acids Res*. **2008**, 36, 1464 − 1471.

(13) (a) Rapireddy, S.; Bahal, R.; Ly, D. H. *Biochemistry* **2011**, 50, 3913 – 3918. (b) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, 128, 10258 – 10267.

(14) (a) Griffith, M. C., Risen, L. M., Greig, M. J., Lesnik, E. A., Sprankle, K. G., Griffey, R. H., Kiely, J. S. & Freier, S. M. *J. Am. Chem. Soc.* **1995**, 117, 831 – 832. (b) Bentin, T.; Larsen, H. J.; Nielsen, P. E. *Biochemistry* **2003**, 42, 13987 – 13995. (c) Kaihatsu, K.; Shah, R. H.; Zhao, X.; Corey, D.; *Biochemistry* **2003**, 42, 13996 – 14003.

(15) (a) Rusling, D. A.; Powers, V. E. C.; Ranasinghe, R. T.; Wang, Y.; Osborne, S. D.; Brown, T.; Fox, K. *Nucleic Acids Res.* **2005**, 33, 3025 −3032. (b) Hari, Y.; Obika, S.; Imanishi, T. *Eur. J. Org. Chem.* **2012**, 2875 −2887.

(16) Moreno, P. M. D.; Geny, S.; Pabon, Y. V.; Bergquist, H.; Zaghloul, E. M.; Rocha, C. S. J.; Oprea, I. I.; Bestas, B.; El-Andaloussi, S.; Jørgensen, P. T.; Pedersen, E. B.; Lundin, K. E.; Zain, R.; Wengel, J.; Smith, C. I. E. *Nucleic Acids Res.* **2013**, 41, 3257 – 3273.

(17) (a) Zaghloul, E. M.; Madsen, A. S.; Moreno, P. M. D.; Oprea, I. I.; El-Andaloussi, S.; Bestas, B.; Gupta, P.; Pedersen, E. B.; Lundin, K. E.; Wengel, J.; Smith, C. I. E. *Nucleic Acids Res.* **2011**, 39, 1142 – 1154. (b) Ge, R.; Heinonen, J. E.; Svahn, M. G.; Mohamed, A. J.; Lundin, K. E.; Smith, C. I. E. *FASEB J.* **2007**, 21, 1902 − 1914.

(18) Hrdlicka, P. J.; Kumar, T. S.; Wengel, J. *Chem. Commun*. **2005**, 4279 − 4281.

(19) Crothers, D. M. *Biopolymers* **1968**, 6, 575 – 584.

(20) Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2010**, 8, 2028 − 2036.

(21) Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J.; *Org. Biomol. Chem.* **2014**, 12, 7758 – 7773.

(22) Sau, S. P.; Madsen, A. S.; Podbevsek, P.; Andersen, N. K.; Kumar, T. S.; Andersen, S.; Rathje, R. L.; Anderson, B. A.; Guenther, D. C.; Karmakar, S.; Kumar, P.; Plavec, J.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, 78, 9560 – 9570.

(23) Didion, B. A.; Karmakar, S.; Guenther, D. C.; Sau, S. P.; Verstegen, J. P.; Hrdlicka, P. J. *ChemBioChem* **2013**, 4, 3447 − 3454.

(24) Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem*. **2011**, 76, 7119 − 7131.

(25) Saenger, W. *Principles of Nucleic Acid Structure*, Springer Advanced Texts in Chemistry **1984**, Springer-Verlag, New York.

(26) Guschlbauer, W.; Jankowski, K. *Nucleic Acids Res.* **1980**, 8, 1421 – 1433.

(27) (a) Divakar, K. J.; Mottah, A.; Reese, C. B.; Sanghvi, Y. S. *J. Chem. Soc. Perkin Trans. 1* **1990**, 969 – 974. (b) Divakar, K. J.; Reese, C. B. *J. Chem. Soc. Perkin Trans. 1* **1982**, 1625 – 1628.

(28) Ozaki, H.; Momiyama, S.; Yokotsuka, K.; Sawai, H. *Tet. Lett.* **2001**, 42, 677 – 680.

(29) Hendrickx, P. M. S.; Martins, J. C. *Chem. Cent. J.* **2008**, 2:20.

(30) Donders, L. A.; Deleeuw, FAAM.; Altona, C. *Magn. Res. Chem.* **1989**, 27, 556 – 563.

(31) (a) Diez, E.; Esteban, A.; Guilleme, J.; Bermejo, F. *J. Mol. Struct.* **1981**, 70, 61 – 64. (b) Diez, E.; Esteban, A.; Bermejo, F.; Altona, C.; de Leeuw, F.A.A.M. *J. Mol. Struct.* **1984**, 125, 49 – 65. (c) de Leeuw, F.A.A.M.; van Kampen, P.; Altona, C.; Diez, E.; Esteban, A. *J. Mol. Struct.* **1984**, 125, 67 – 88.

(32) Nakamura, M.; Fukunaga, Y.; Sasa, K.; Ohtoshi, Y.; Kanaori, K.; Hayashi, H.; Nakano, H.; Yamana, K. *Nucleic Acids Res.* **2005**, 33, 5887 − 5895.

(33) Marin, V.; Hansen, H. F.; Koch, T. R.; Armitage, B. A. *J. Biomol. Struc. Dyn.* **2004**, 21, 841 – 850.

(34) Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, 78, 12040 – 12048.

(35) Anderson, B. A.; Onley, J. J.; Hrdlicka, P. J. *Manuscript in preparation.* See Chapter 3.

(36) Kumar, T. S.; Madsen, A. S.; Østergaard, M. E.; Sau, S. P.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem*. **2009**, 74, 1070 – 1081.

(37) (a) Korshun, V. A.; Stetsenko, D. A.; Gait, M. J. *J. Chem. Soc., Perkin Trans. 1* **2002**, 1092 – 1104. (b) Dohno, C.; Saito, I. *ChemBioChem* **2005**, *6*, 1075 – 1081.

(38) Asanuma, H.; Fujii, T.; Kato, T.; Kashida, H. *J. Photochem. Photobiol., C* **2012**, 13, 124 – 135.

(39) Dougherty, G.; Pilbrow, J. R. *Int. J. Biochem*. **1984**, 16, 1179 – 1192.

(40) Mergny, J.L.; Lacroix, L. *Oligonucleotides* **2003**, 13, 515 – 537.

(41) (a) Dioubankova, N. N.; Malakhov, A. D.; Stetsenko, D. A.; Gait, M. J.; Volynsky, P. E.; Efremov, R. G.; Korshun, V. A. *ChemBioChem* **2003**, 4, 841 – 847. (b) Astakhova, I. V.; Malakhov, A. D.; Stepanova, I. A.; Ustinov, A. V.; Bondarev, S. L.; Paramonov, A. S.; Korshun, V. A. *Bioconj. Chem*. **2007**, 18, 1972 – 1980. (c) Astakhova, I. V.; Ustinov, A. V.; Korshun, V. A.; Wengel, J. *Bioconj. Chem*., **2011**, 22, 533 – 539.

(42) Dioubankova, N. N.; Malakhov, A. D.; Stetsenko, D. A.; Gait, M. J.; Volynsky, P. E.; Efremov, R. G.; Korshun, V. A. *ChemBioChem* **2003**, 4, 841 – 847.

CHAPTER 6: The Merger of Two Strategies for Recognition of Double-stranded DNA: Pseudocomplementary Invaders

Brooke A. Anderson and Patrick J. Hrdlicka

Department of Chemistry, University of Idaho, Moscow, Idaho 83844, United States

*Manuscript in prepartion

Abstract

The recognition of chromosomal DNA by synthetic ligands is a versatile strategy for the detection and modulation of genes. Despite the progress with minor-groove binding polyamides, triplex-forming oligonucleotides and peptide nucleic acids, there remains a need for probes that can sequence-specifically target unique mixed-sequence DNA under physiologically relevant conditions. Invader probes comprise a fundamentally different approach toward this goal which entails the use of energetically activated DNA duplexes. Here, two different types of pseudocomplementary Invaders are introduced as a novel probe technology for mixed-sequence recognition of double-stranded DNA targets. The first design consists of 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-2-thiouridine modified Invader probes with pseudocomplementary hotspots. The second chimeric double-stranded probes consist of +1 interstrand zipper arrangements of 2′-*N*-(pyren-1-ylmethyl)-2′-*N*-methyl-2′-aminouridine monomers and pseudocomplementary base pairs between 2-aminoadenosine and 2-thiouridine. In both designs, the component ONs bind less strongly to each other than to complementary DNA targets, which drives the energetically activated duplex toward dsDNA recognition via formation of highly stabilized probe-target duplexes. Experiments with model

DNA hairpins demonstrate these probes recognize dsDNA efficiently and proceed with excellent fidelity. The synthesis of the necessary building blocks will be described along with detailed biophysical characterization of ONs and probe duplexes modified with these units.

6.1 Introduction

The development of agents capable of sequence-specific and efficient recognition of mixed-sequence double-stranded DNA (dsDNA) is a long-standing goal in nucleic acid research due to their tremendous potential as tools for fundamental gene function studies, diagnostics, and therapeutics.[1,2] This has prompted extensive exploration into chemical modifications that allow facile tuning of the stability and functional properties of oligonucleotides (ONs) to enhance their performance for DNA-targeting applications.[3] Significant progress toward this end however has proven highly challenging as the Watson-Crick base pairs, where the sequence information is stored, are buried deeply within the duplex core and are not readily accessible to exogenous agents, hence a robust method for recognizing mixed-sequence dsDNA remains elusive despite the efforts of many investigators.

Many types of DNA targeting methodologies have been described. Some of the most current DNA recognition methods utilize chemically modified ligands to recognize chemical features that are accessible via one of the grooves– i.e., pyrrole-imidazole polyamides,[4,5] engineered proteins[6-8] and triplex-forming oligonucleotides[9] or peptide nucleic acids (PNAs),[10] among other approaches.[11] While major progress has been made with these approaches, many limitations persist. For example, polyamides have binding site size limitations, triplex-based and regular PNA-based approaches have polypurine tract limitations and non-physiological pH requirements, and most PNA-based approaches also require non-physiological salt

concentrations (i.e. low saline conditions). To circumvent some of the sequence limitations of regular PNA-based approaches, tail-clamp PNA and γ-PNA modifications were introduced. Tail-clamp PNAs, have a mixed-sequence extension adjacent to a homopurine run and bind DNA via a combined duplex/triplex strand invasion mode.[12] However, they still require a short polypurine tract for invasion. γ-PNA are single-stranded probes that are fully modified with conformationally pre-organized PNA building blocks capable of recognizing internal regions of mixed-sequence dsDNA regardless of GC-content; however, they also require non-physiological salinity for optimal strand invasion.[13]

Probes that recognize mixed-sequence dsDNA via duplex invasion have gained considerable attention due to the predictability of the Watson-Crick base-pairing rules. However, these approaches must overcome a steep energetic penalty since the pre-existing base pairs of DNA must be broken prior to probe binding. Pseudo-complementary (pc) base pairs are unnatural nucleobase modifications that consist of 2-aminoadenosine (**D**) and 2-thiothymine (**S**) in place of adenosine (A) and thymine (T), respectively (See Figure 6-1b for structures of **D** and **S**). Where an A:T base pair consists of two hydrogen bonds forming a natural base pair, a **D**:**S** is a mismatched base pair due to a steric clash between the 2-amino group of **D** and the 2-thio group of **S**; however, **D**:T and **S**:A are stable base pairs, with the **D**:T base pair being particularly stable due to the addition of a third hydrogen bond relative to the natural A:T base pair.[14] Therefore, complementary pc base pairs have reduced affinity for one another but high affinity for unmodified complementary DNA or RNA. These pseudo-complementary base pairs can be inserted into DNA strands in place of natural base-pairs (pcDNAs) and have been shown to invade homologous duplexes under conditions where unmodified natural DNAs failed.[14] Likewise, these unnatural base pairs have been explored in PNAs (pcPNAs[15]) and have been

shown to have efficient double-duplex invasion into internal regions of mixed-sequence dsDNA.[16] However, self-inhibitory effects observed at high pcPNA concentrations and the requirement for low salinity potentially impose limitations on pcPNA-mediated strand invasion in biological media.[17] Thus, probes that efficiently recognize mixed-sequence dsDNA targets at physiological conditions remain elusive.

Our laboratory has recently introduced Invader probes as an alternative approach toward mixed-sequence recognition of dsDNA via a duplex invasion mechanism.[18] These probes are energetically activated for dsDNA recognition through modification with intercalator-functionalized nucleotides (for an illustration, see Figure 6-1). The activation of the probe arises from the intercalator organization inside the duplex, termed the +1 zipper arrangement (for a definition of zipper nomenclature, see Experimental Section), which results in duplex destabilization. This motif forces two intercalators into the same region of the duplex, which is in violation of the 'nearest neighbor exclusion principle' according to which intercalators, at most, bind to every second base pair of a DNA duplex as local duplex unwinding and the formation of an 'energetic hotspot' otherwise occurs.[18-20] On the other hand, the two strands that constitute an Invader probe displays exceptional affinity toward complementary DNA (cDNA) since the intercalators interact strongly with neighboring base pairs. We have used the difference in thermostabilities of Invader probes and probe-target duplexes to drive recognition of short iso-sequential mixed-sequence DNA hairpins[18,20-22] and chromosomal DNA targets.[23]

The original Invader probes utilized 2′-*N*-(pyren-1-yl)-2′-amino-α-L-LNA (Locked Nucleic Acid) as the key activating components.[18,20] The challenging synthesis of 2′-amino-α-L-LNA monomers resulted in a study that identified simpler structural and functional mimics, leading to the discovery of 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-aminouridine monomer as a next-

generation Invader modification.[24] The resulting Invader probe displayed similar dsDNA-recognition efficiency[21] but was significantly easier to synthesize.[25-26] Identification of a simpler scaffold enables us to perform systematic structure-property relationship studies with the goal of optimizing the recognition efficiency of Invader duplexes.

In the present study, we set out to maximize the thermodynamic dsDNA recognition potential of Invader probes by synthesizing Invaders that are modified with pseudocomplementary base pairs (pcInvaders). We set out to increase the change in Gibbs free energy (ΔG_{rxn}) for dsDNA recognition by developing even more strongly energetically activated probes by incorporating both energetic hotspot elements (i.e., +1 zippers of intercalator-functionalized nucleotides) and pseudo-complementary base pairs (Figure 6-1). These pseudocomplementary Invaders are characterized by means of thermal denaturation experiments, analysis of thermodynamic parameters, UV-vis absorption and fluorescence spectroscopy, and DNA recognition experiments.

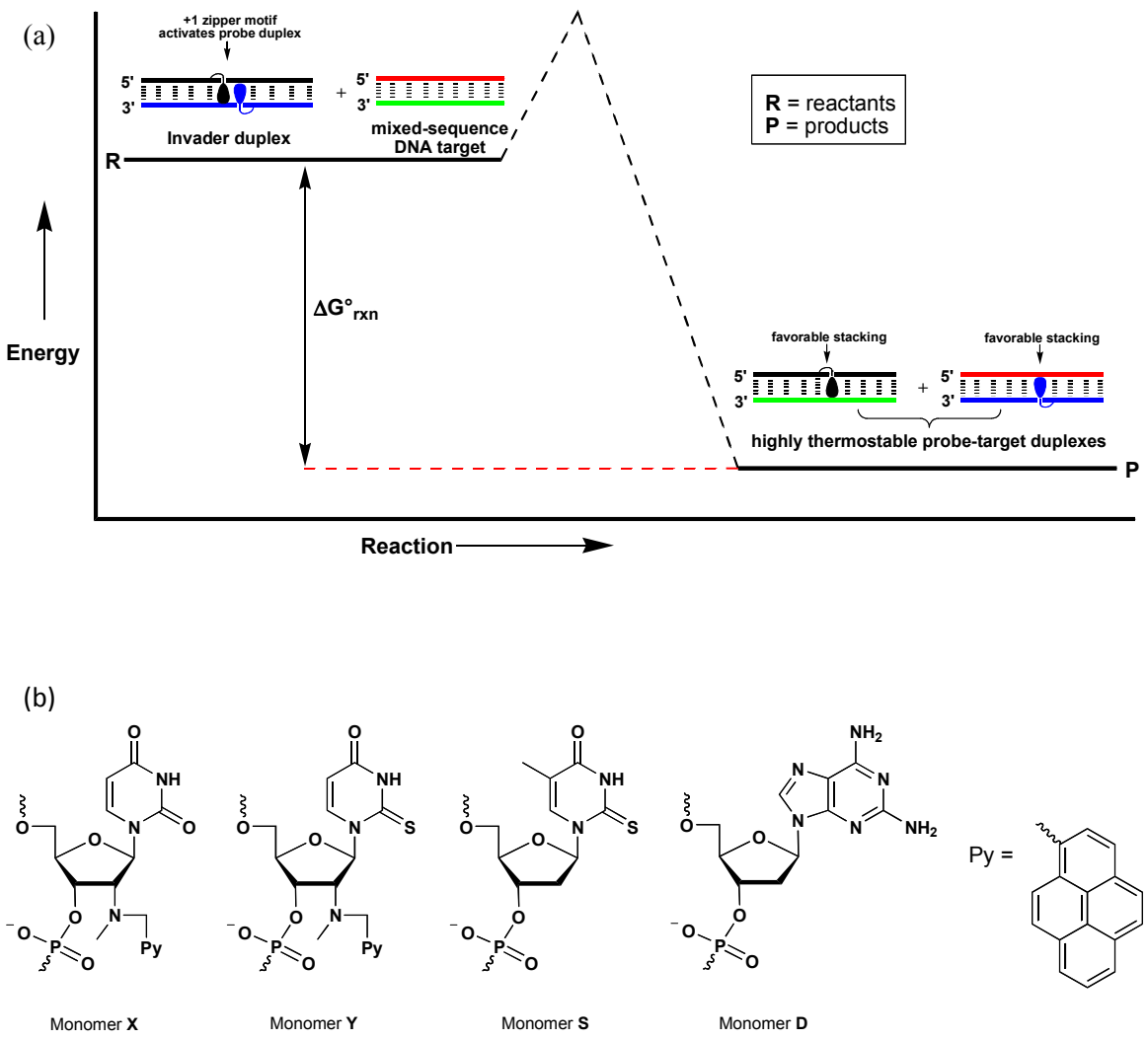
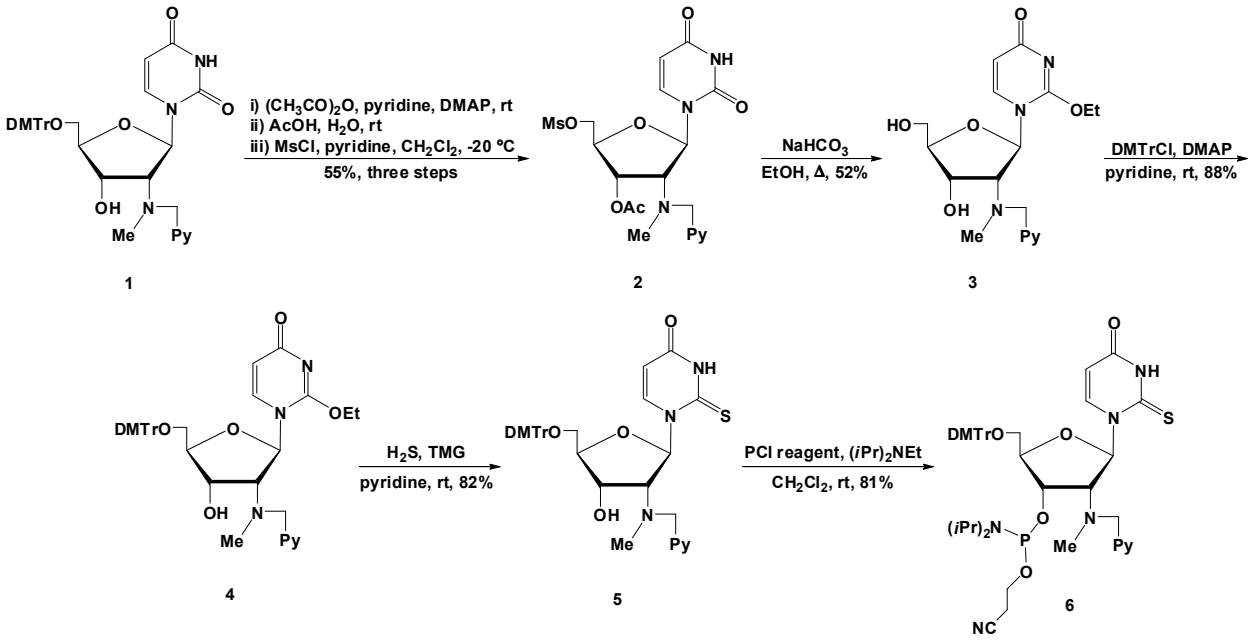


Figure 6-1: (a) Energy diagram illustration of the Invader approach for recognition of mixed-sequence DNA and (b) structures of monomers used herein.

6.2 Results and Discussion

6.2.1 Synthesis of N2′-pyrene-functionalized 2-thiouridine Phosphoramidite

Target phosphoramidite **6** was obtained from known nucleoside **1** by following the same general strategy that was used for the synthesis of 2′-*O*-[2-(methoxy)ethyl]-2-thiothymidine (Scheme 6-1).[27] Thus, nucleoside **1** – obtained in 61% yield over six steps from uridine[28] – was first subjected to a sequence of protecting group manipulations, i.e., 3′-*O*-acetylation, 5′-*O*-detritylation and 5′-*O*-methanesulfonylation, to afford nucleoside **2** in 55% yield over three steps. Prolonged refluxing in anhydrous ethanol in the presence of sodium bicarbonate,[27] results in the formation of 2-*O*-ethyluridine derivative **3** in 52% yield, presumably via nucleophilic opening of an O2,O5′-anhydrouridine intermediate. Subsequent O5′-dimethoxytritylation of **3** using standard conditions affords nucleoside **4** in 88% yield, which upon treatment with H_2S-saturated pyridine in the presence of 1,1,3,3-tetramethylguanidine[29] provides 2-thiouridine derivative **5** in 82% yield. Treatment of nucleoside **5** with 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite and *N*,*N*-diisopropylethylamine affords target phosphoramidite **6** in 81% yield, corresponding to an overall yield of ~17% from nucleoside **1**.



Scheme 6-1: Synthesis of 2′-*N*-(pyren-1-yl)methyl-2′-amino-2′-deoxy-2′-*N*-methyl-2-thiouridine phosphoramidite **6**. DMTr = 4,4′-dimethoxytrityl; Py = pyren-1-yl; Ms = methanesulfonyl; TMG = 1,1,3,3-tetramethylguanidine; PCl reagent = 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite.

6.2.2 Synthesis of Modified ONs

Phosphoramidite **6** was used to incorporate monomer **Y** into oligodeoxyribonucleotides (ONs) via automated solid-phase DNA synthesis. Extended hand-coupling (15 min) and the use of 4,5-dicyanoimidazole as an activator resulted in stepwise coupling yields of ~95%. ONs modified with monomer **X** were synthesized as previously described (15 min coupling, 5-[3,5-bis(trifluoromethyl)phenyl]-1*H*-tetrazole as activator, ~99% coupling yield).[24] The corresponding phosphoramidites of monomers **S** and **D** were obtained from commercial sources and incorporated into ONs using the conditions for incorporation of monomer **Y** (stepwise

coupling yields >95%). To prevent desulfurization in ONs modified with 2-thiouracil monomers **Y** or **S**, nucleotide phosphite to phosphate oxidation was performed using *tert*-butylhydroperoxide/CH_3CN/H_2O (10 min) rather than the standard aqueous iodine solution.[30] The identity and purity of the modified ONs was established through MALDI-TOF (Table 6-S1 in Supporting Information) and ion-pair reverse-phase HPLC (>80% purity), respectively.

6.2.3 Thermal Denaturation Properties

Pyrene-functionalized 2-thiouracil monomer **Y** was incorporated into the same 9-mer mixed-sequence ONs that we previously used for evaluation of benchmark monomer **X**.[28] Thermal denaturation temperatures (T_m's) of duplexes between **Y**-modified ONs and complementary DNA or RNA (cDNA/cRNA) were determined in medium salt phosphate buffer ($[Na^+]$ = 110 mM) and compared relative to unmodified and **X**-modified duplexes. The resulting denaturation curves display the expected monophasic sigmoidal transitions (Figure 6-S1). Duplexes between **Y1**-**Y4** and cDNA are significantly more stable than unmodified reference duplexes (ΔT_m between +2.5 and +11.5 °C, Table 6-1), while heteroduplexes with cRNA are far less stable (ΔT_m between -6.5 to +2.5 °C, Table 6-S2). ONs in which **Y** monomers are flanked by 3′-purines induce greater duplex stabilization than ONs with 3′-flanking pyrimidines (e.g., compare ΔT_m's for **Y1** and **Y3**, Table 6-1). This sequence dependence, along with the prominent DNA selectivity (Table 6-S3), are typical characteristics of ONs modified with intercalating pyrene moieties.[31,21] Somewhat surprisingly, **Y**-modified ONs form slightly less stable duplexes with cDNA than their **X**-modified counterparts (compare T_m's of **X1**-**X4** and **Y1**-**Y4**, Table 6-1), which suggests that the mechanisms underlying the stabilizing effects of the pyrene and 2-thiouracil moieties are not fully compatible.

In order to generate Invader probes with 'pseudocomplementary energetic hotspots', we also synthesized ONs in which 2-amino-2′-deoxyadenosine monomers flank monomer **Y** (i.e., the **DY**-series). Replacing regular 2′-deoxyadenosines with **D** monomers increases the cDNA/cRNA affinity of **Y**-modified ONs by 1-5 °C, presumably due to the extra hydrogen bond in **D**:T base pairs relative to canonical A:T base pairs (e.g., compare ΔT_m for **Y1** and **DY1**, Table 6-1 and Table 6-S2). Similar relative increases in cDNA affinity were observed upon replacing 2′-deoxyadenosines with **D** monomers in otherwise unmodified ONs (see ΔT_m for **D1**-**D4** vs cDNA, Table 6-1), which is in line with literature reports.[32,33]

.

Table 6-1: Thermal denaturation temperatures of duplexes between **X**-, **Y**- or **DY**-modified ONs and cDNA.[a]

ON	Sequence	ΔT_m (°C)
X1[b]	5′-GTG A**X**A TGC	+15.0
X2[b]	3′-CAC **X**AT ACG	+1.5
X3[b]	3′-CAC TA**X** ACG	+15.0
X4[b]	3′-CAC **X**A**X** ACG	+14.0
Y1	5′-GTG A**Y**A TGC	+11.5
Y2	3′-CAC **Y**AT ACG	+2.5
Y3	3′-CAC TA**Y** ACG	+11.0
Y4	3′-CAC **Y**A**Y** ACG	+10.0
DY1	5′-GTG A**YD** TGC	+16.0
DY2	5′-GTG **DY**A TGC	+13.0
DY3	3′-CAC **YD**T ACG	+3.5
DY4	3′-CAC T**DY** ACG	+16.0
D1	5′-GTG **D**TA TGC	+1.0
D2	5′-GTG AT**D** TGC	+2.0
D3	3′-CAC T**D**T ACG	+3.0
D4	5′-GTG **D**T**D** TGC	+5.0

[a]ΔT_m = change in T_m relative to reference duplex **DNA1**:**DNA2** ($T_m \equiv 29.5$ °C), where **DNA1**: 5′-GTG ATA TGC and **DNA2**: 3′-CAC TAT ACG. T_m's are determined as the maximum of the first derivative of melting curves (A_{260} vs T) recorded in medium salt phosphate buffer ($[Na^+] = 110$ mM, $[Cl^-] = 100$ mM, pH 7.0 (NaH_2PO_4/Na_2HPO_4)), using 1.0 μM of each strand. Reported T_m's are averages of at least two measurements within 1.0 °C; A = adenin-9-yl DNA monomer, C = cytosin-1-yl DNA monomer, G = guanin-9-yl DNA monomer and T = thymin-1-yl DNA monomer. For structures of monomers **X, Y** and **D**, see Figure 6-1.

[b]Data previously reported in reference 28.

The binding specificity of centrally modified ONs was examined using DNA targets with mismatched nucleotides opposite to the pyrene-functionalized nucleotide (Table 6-2). We have previously shown that **X1** discriminates C- and T-mismatched DNA targets with similar or better efficiency than the corresponding unmodified ON, while G-mismatched DNA targets are poorly discriminated (Table 6-2).[28] Centrally modified **Y1** displays greatly improved

discrimination of G-mismatch targets (compare ΔT_m's for **Y1** and **X1**, Table 6-2), presumably due to the steric hindrance and weaker H-bonding ability of sulfur relative to oxygen, causing the 2-thiouracil moiety to destabilize the wobble base pair with G.[34] Interestingly, **DY1** discriminates all mismatched DNA targets rather efficiently, indicating that incorporation of flanking 2-amino-2′-deoxyadenosine is a strategy toward improving binding fidelities of **Y**-modified ONs. Additional mismatch discrimination data are shown in Table 6-S4 in the Supporting Information.

Table 6-2: Discrimination of mismatched DNA targets by **X1**/**Y1**/**DY1** and reference ONs.[a]

		DNA: 3′-CAC T**B**T ACG			
		T_m [°C]	ΔT_m [°C]		
ON	**Sequence** **B** =	**A**	C	G	T
DNA1	5′-GTG ATA TGC	**29.5**	-16.5	-9.5	-17.0
X1[b]	5′-GTG A**X**A TGC	**44.5**	-23.0	-3.5	-13.0
Y1	5′-GTG A**Y**A TGC	**41.0**	-19.5	-12.5	-8.5
DY1	5′-GTG A**YD** TGC	**45.5**	-23.5	-16.5	-11.5

[a] For conditions of thermal denaturation experiments, see Table 6-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched DNA:DNA duplex.

[b] Data from reference 24.

6.2.4 Photophysical Properties of Y- and DY-modified ONs and Duplexes with Complementary DNA/RNA

UV-Vis absorption and steady-state fluorescence emission spectra of **Y**- and **DY**-modified ONs were recorded in the absence or presence of cDNA/cRNA to gain additional insight into the binding mode of the pyrene moiety. Hybridization of **Y**- and **DY**-modified ONs with cDNA/cRNA results in prominent hypochromic and bathochromic shifts of the pyrene

absorption maxima ($\Delta\lambda_{max}$ = 3-7 nm, Table 6-S5, Figures 6-2 and 6-S2), which is indicative of ground-state electronic interactions between pyrenes and nucleobases[35] and intercalation. Steady-state fluorescence emission spectra (λ_{ex} = 350 nm, T = 5 °C) of these duplexes display two vibronic bands at λ_{em} = 381±1 nm and 400±1 nm, respectively, as well as a small shoulder at ~420 nm (Figures 6-2 and 6-S3). Hybridization with cDNA/cRNA is accompanied by minor changes in fluorescence intensity (0.5- to 5.4-fold); more pronounced quenching of fluorescence is observed for **DY**-modified ONs.

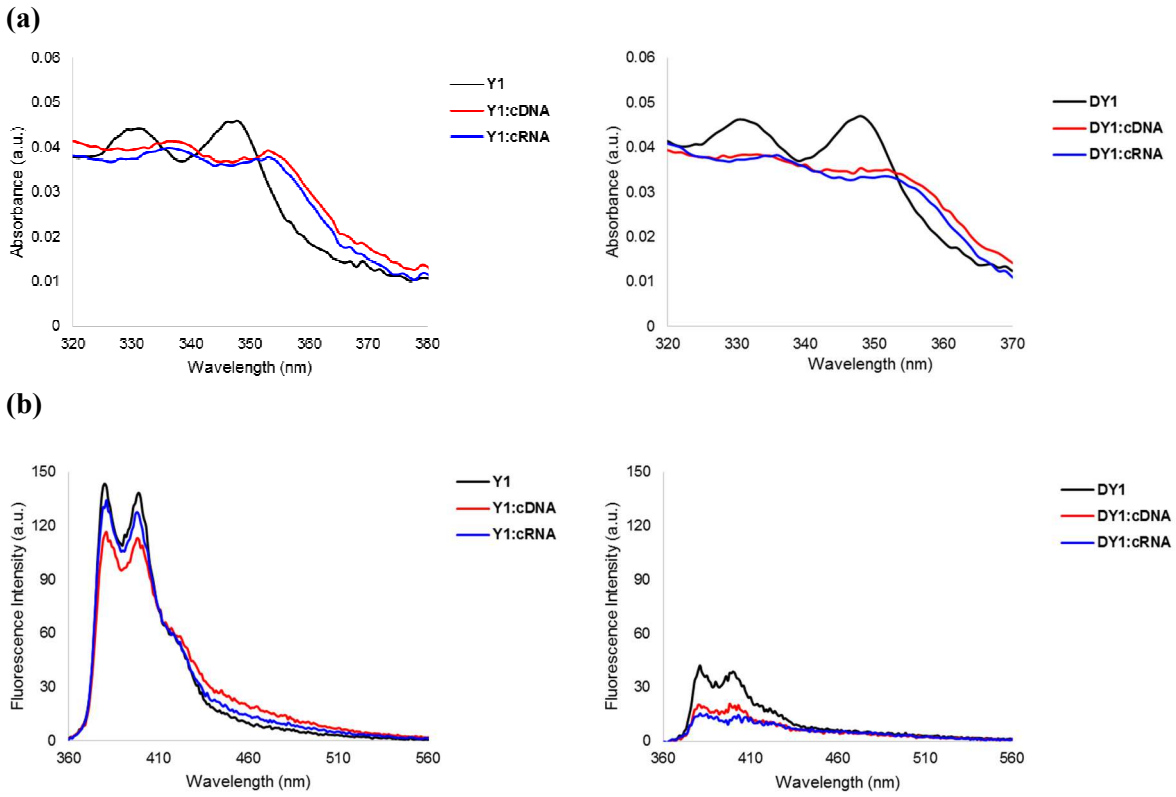


Figure 6-2: (a) Absorption spectra of single-stranded **Y1** and **DY1** and the corresponding duplexes with cDNA/cRNA targets. Spectra were recorded at T = 10 °C. (b) Steady-state fluorescence emission spectra of **Y1** and **DY1** and the corresponding duplexes with cDNA/cRNA targets. Spectra were recorded at 5 °C using λ_{ex} = 350 nm. Each strand was used at 1.0 μM concentration in T_m buffer.

6.2.5 Biophysical Properties of DNA Duplexes with Interstrand Zippers of Y-monomers

Next, we set out to study DNA duplexes with different interstrand arrangements of **Y** monomers and **YD** segments to identify probe architectures that are more strongly energetically activated for dsDNA-recognition than regular Invader probes. As expected from our previous studies on Invader constructs[20-22], **Y1**:**Y3**, which features a +1 interstrand arrangement of **Y** monomers, is much more thermolabile than **Y1**:**Y2**, featuring a -1 interstrand zipper of **Y** monomers (Table 6-3), or than duplexes between **Y**-modified ONs and cDNA (Table 6-1). As observed with other Invader modifications.[20-22] This destabilization is likely due to violation of the nearest neighbor exclusion principle (i.e., the local intercalator density is too high leading to duplex perturbation). In line with our hypothesis, introduction of a 2,6-diaminopurine **D** monomer opposite of the pyrene-functionalized 2-thiouracil **Y** monomer, decreases the T_m's of the duplexes (e.g., compare ΔT_m of **DY2**:**DY3** and **Y1**:**Y2**, Table 6-3). Interestingly, the destabilizing effect of the pseudocomplementary base pairs is more pronounced when the **Y** monomers are in a -1 zipper orientation (compare drop in T_m of 9.5 °C between **DY2**:**DY3** and **Y1**:**Y2** relative to the drop of 2.0 °C between **DY1**:**DY4** and **Y1**:**Y3**, Table 6-3). This unfortunately indicates that the pseudocomplementary energetic hotspot architecture of **DY1**:**DY4** does not allow for full harnessing of the destabilizing effects from both structural elements.

The above T_m-based conclusions are supported by the Gibbs free energies associated with duplex formation, which were derived from denaturation curves via line fitting.[36] Thus: i) Duplexes between **Y**-modified ONs and cDNA are much more stable than unmodified reference duplexes ($\Delta\Delta G^{293}$ between -16 and -7 kJ/mol, first and second ΔG^{293} columns, Table 6-3) due to highly favorable enthalpy ($\Delta\Delta H$ between -76 and -44 kJ/mol, Table 6-S6 in the Supporting

Information). ii) In comparison, the corresponding **X**-modified duplexes are slightly more stable, while the **D**-modified duplexes are much less stable (e.g., compare $\Delta\Delta G^{293}$ for **Y1**:cDNA, **X1**:cDNA and **D1**:cDNA, Table 6-3). iii) Duplexes between ONs with **YD**-motifs and cDNA are generally less stable than the corresponding **Y**-modified duplexes (e.g., compare $\Delta\Delta G^{293}$ for **Y1**:cDNA and **DY2**:cDNA, Table 6-3). This contradicts the T_m-based conclusions (Table 6-1), but probably is a ramification of the different entropies of these duplexes (Table 6-S7), leading to different temperature dependencies of the Gibbs free energies. iv) Formation of **Y1**:**Y3** (+1 zipper) is less thermodynamically favorable than **Y1**:**Y2** (-1 zipper) (third ΔG^{293} column, Table 6-3). v) Duplexes with pseudocomplementary energetic hotspots are marginally less stable than the corresponding duplexes featuring only regular energetic hotspots (e.g., compare ΔG^{293} for **DY1**:**DY4** and **Y1**:**Y3**, Table 6-3). The energetic nature of **DY1**:**DY4** is the result of particularly unfavorable enthalpy ($\Delta\Delta H$ = +95 kJ/mol, Table 6-S6), which most likely reflects violation of the nearest neighbor exclusion principle and concomitant formation of destabilizing pseudocomplementary base pairs. vi) The stability of duplexes with pseudocomplementary base pairs but no energetic hotspots, ranges between slightly more stable to slightly more unstable than unmodified duplexes (e.g., see $\Delta\Delta G^{293}$ for **D1**:**Y2**, Table 6-3).

As a result of these stability trends, +1 zipper duplex **Y1**:**Y3** is much more energetically activated for recognition of isosequential dsDNA targets than -1 zipper duplex **Y1**:**Y2**, as gauged by ΔG_{rec}^{293} (ON_A:ON_B) = ΔG^{293} (ON_A:cDNA) + ΔG^{293} (cDNA:ON_B) - ΔG^{293} (ON_A:ON_B) - ΔG^{293} (dsDNA), where ON_A:ON_B is a double-stranded probe and 'dsDNA' is the isosequential dsDNA target for ON_A:ON_B (compare ΔG_{rec}^{293} for **Y1**:**Y2** and **Y1**:**Y3**, Table 6-3). **DY1**:**DY4**, featuring a pseudocomplementary energetic hotspot, is more energetically activated for dsDNA recognition than **Y1**:**Y3** (compare ΔG_{rec}^{293} for **DY1**:**DY4** and **Y1**:**Y3**, Table 6-3). As expected,

DY2:**DY3**, which also features two pseudocomplementary base pairs but has an -1 interstrand zipper arrangement of **Y** monomers, is far less activated for dsDNA-recognition (compare ΔG_{rec}^{293} for **DY2**:**DY3** and **DY1**:**DY4**, Table 6-3). Similarly, double-stranded probes with pseudocomplementary base pairs but no energetic hotspots are much less activated for dsDNA-recognition than **DY1**:**DY4** (e.g., compare ΔG_{rec}^{293} of **Y1**:**D3** and **DY1**:**DY4**, Table 6-3).

Contrary to our initial hypothesis, **DY1**:**DY4** has lower dsDNA-targeting potential than the parent Invader **X1**:**X3** (compare ΔG_{rec}^{293} for **DY1**:**DY4** and **X1**:**X3**, Table 6-3). This is due to a combination of two factors: i) pyrene intercalation perturbs the local duplex geometry, which presumably weakens the normally stabilizing base pairs between 2,6-diaminopurine:thymine (**D**:T) and 2-thiouracil:adenine (**Y**:A), leading to lower-than-expected stability of **DY**:cDNA duplexes (e.g., compare ΔG^{293} for **X1**:cDNA, **Y1**:cDNA and **DY1**:cDNA, Table 6-3), and ii) +1 interstrand zipper arrangements of nucleotide monomers with intercalating pyrene moieties perturb local duplex geometries,[22] which, in turn, may prevent the steric clashes between 2,6-diaminopurine and 2-thiouracil that normally occur in pseudocomplementary base pairs, resulting in less pronounced probe destabilization. The inverse argument is also possible, i.e., the pseudocomplementary base pairs of **DY1**:**DY4** interfere with the destabilizing forced intercalation of the two pyrene moieties. Support for the latter argument is obtained from UV-Vis absorption spectra of **DY1**:**DY4**. Normally, DNA duplexes with +1 interstrand zipper motifs of pyrene-functionalized monomers exhibit markedly blue-shifted pyrene absorption relative to DNA duplexes with other interstrand zipper motifs[22,28], which is indicative of reduced pyrene-nucleobase interactions[35b] due to a locally perturbed duplex geometry (e.g., compare λ_{max} for **X1**:**X3** relative to **X1**:**X2**, Table 6-3). The same trend is observed for **Y1**:**Y3**

(compare λ_{max} for **Y1**:**Y3** relative to **Y1**:**Y2**, Table 6-3), but the trend is less pronounced for **DY1**:**DY4** (compare λ_{max} for **DY1**:**DY4** relative to **DY2**:**DY3**, Table 6-3).

Table 6-3: Biophysical properties of **X**-, **Y**- or **DY**-modified DNA duplexes.[a]

ON	ZP	Sequence	T_m (°C)	$\Delta G^{293}[\Delta\Delta G^{293}]$ (kJ/mol)			ΔG_{rec}^{293} (kJ/mol)	λ_{max} (nm)
				upper ON vs cDNA	lower ON vs cDNA	probe duplex		
X1 **X2**	**-1**	5′-GTG A**X**A TGC 3′-CAC **X**AT ACG	42.5	-65±1 [-20]	-48±1 [-3]	-54±1 [-9]	-14	352
X1 **X3**	**+1**	5′-GTG A**X**A TGC 3′-CAC TA**X** ACG	28.5	-65±1 [-20]	-64±1 [-19]	-44±0 [+1]	-40	345
Y1 **Y2**	**-1**	5′-GTG A**Y**A TGC 3′-CAC **Y**AT ACG	39.5	-61±1 [-16]	-52±1 [-7]	-54±0 [-9]	-14	353
Y1 **Y3**	**+1**	5′-GTG A**Y**A TGC 3′-CAC TA**Y** ACG	28.5	-61±1 [-16]	-59±1 [-14]	-46±0 [-1]	-29	347
DY2 **DY3**	**-1**	5′-GTG **DY**A TGC 3′-CAC **YD**T ACG	30.0	-56±1 [-11]	-48±1 [-3]	-46±1 [-1]	-13	352
DY1 **DY4**	**+1**	5′-GTG A**YD** TGC 3′-CAC T**DY** ACG	26.5	-60±1 [-15]	-61±2 [-16]	-42±1 [+3]	-34	350
D1 **Y2**	**-**	5′-GTG **D**TA TGC 3′-CAC **Y**AT ACG	24.0	-47±1 [-2]	-52±1 [-7]	-43±1 [+3]	-11	352
Y1 **D3**	**-**	5′-GTG A**Y**A TGC 3′-CAC T**D**T ACG	32.5	-61±1 [-16]	-49±0 [-4]	-51±1 [-6]	-14	352
D2 **Y3**	**-**	5′-GTG AT**D** TGC 3′-CAC TA**Y** ACG	33.5	-48±0 [-3]	-59±1 [-14]	-51±0 [-6]	-11	353
D4 **Y4**	**-**	5′-GTG **D**T**D** TGC 3′-CAC **Y**A**Y** ACG	24.5	-51±1 [-6]	-58±0 [-13]	-44±0 [+1]	-20	353

[a]ZP = zipper. For conditions of thermal denaturation and absorption experiments, see Table 3-1 and Figure 3-2, respectively. $\Delta\Delta G^{293}$ is measured relative to ΔG^{293} for **DNA1**:**DNA2** = -45 kJ/mol. ΔG_{rec}^{293} (**ON$_A$**:**ON$_B$**) = ΔG^{293} (**ON$_A$**:cDNA) + ΔG^{293} (cDNA:**ON$_B$**) - ΔG^{293} (**ON$_A$**:**ON$_B$**) - ΔG^{293} (dsDNA). "±" denotes standard deviation. Data for **X1**:**X2** and **X1**:**X3** are from reference 28 and are included to facilitate comparison.

6.2.6 Recognition of DNA Hairpins using Energetically Activated Probe Duplexes

Based on the observed ΔG_{rec}^{293} values, we decided to experimentally test the dsDNA-recognition properties of **Y1**:**Y3** and **DY1**:**DY4** relative to benchmark Invader **X1**:**X3** using an electrophoretic mobility shift assay from our earlier studies.[21] Thus, a digoxigenin (DIG)

labeled DNA hairpin (DH) – comprised of a 9-mer double-stranded mixed-sequence stem, which is linked by a T_{10} loop – was used as a model dsDNA target (Figure 6-3a and 6-3b). Room temperature incubation of **DH1** with **Y1**:**Y3**, **DY1**:**DY4** or **X1**:**X3** in a HEPES buffer results in dose-dependent formation of a ternary recognition complex as evidenced by the emergence of a slower migrating band on non-denaturing PAGE gels (Figure 6-3c). Fitting of the corresponding dose-response curves via non-linear regression reveals that **X1**:**X3**, **Y1**:**Y3**, and **DY1**:**DY4** display C_{50} values of ~0.8 μM, ~2.8 μM, and ~1.5 μM, respectively (Figure 6-3d), which is in agreement with the observed trend in ΔG_{rec}^{293} values (Table 6-3).

The binding specificities of **Y1**:**Y3**, **DY1**:**DY4** and **X1**:**X3** were examined by incubating a 200-fold molar excess of the probes with DNA hairpins **DH2**-**DH7**, which deviate in the nucleotide sequence at one position relative to the Invader probes (Figure 6-3b). Importantly, no recognition was observed (Figure 6-3e), demonstrating that Invader-mediated dsDNA-recognition proceeds with single nucleotide fidelity.

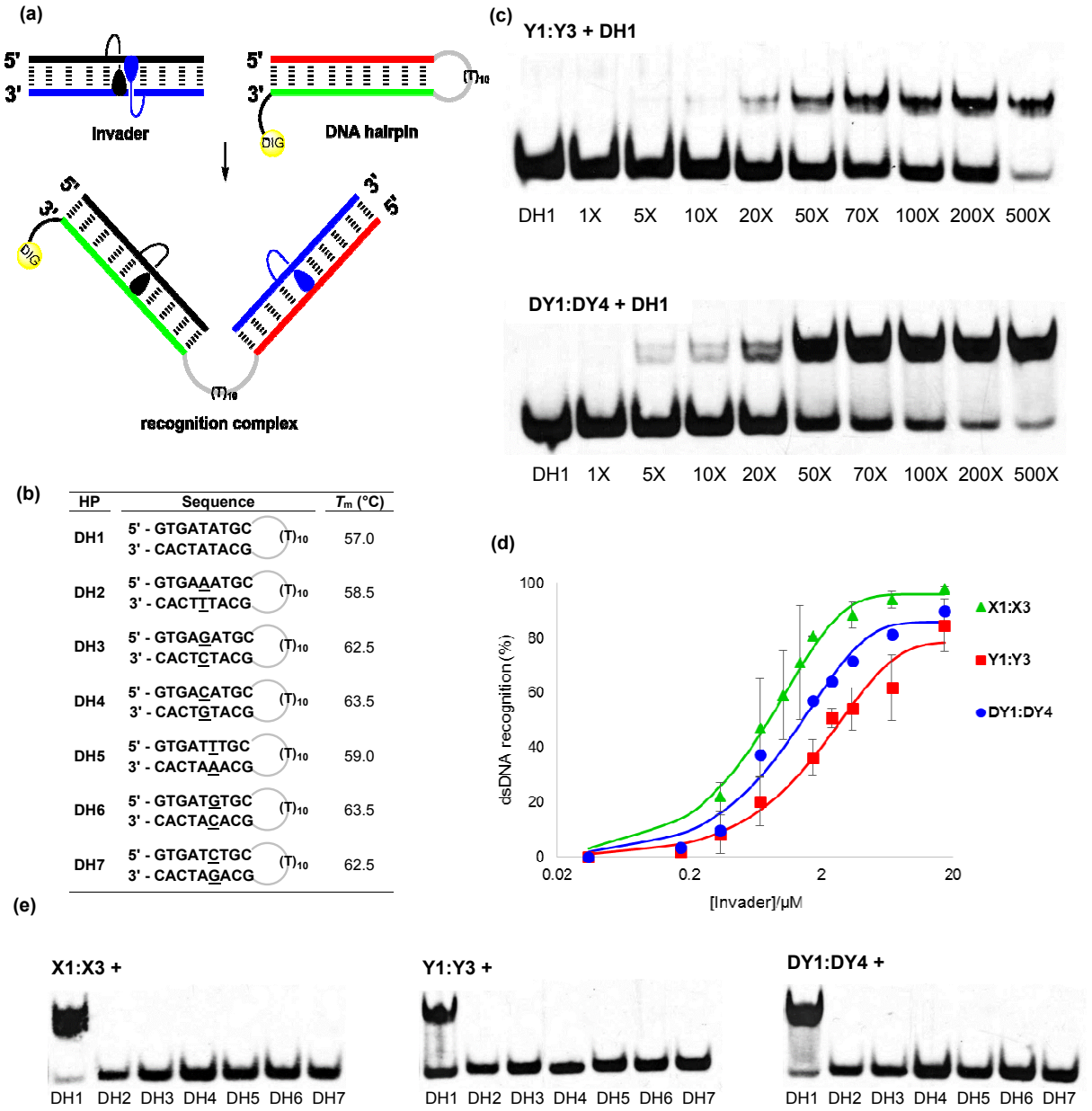


Figure 6-3: Recognition of DNA hairpins using energetically activated double-stranded probes. (a) Illustration of recognition process; (b) sequences of DNA hairpins with isosequential (**DH1**) or non-isosequential stems (**DH2-DH7**) – underlined nucleotides indicate sequence deviations relative to probes; (c) representative electrophoretograms of recognition of **DH1** using 1- to 500-fold excess of **Y1**:**Y3** or **DY1**:**DY4**; (d) dose-response curves (average of at

least three independent experiments; error bars represent standard deviation); (e) electrophoretograms illustrating incubation of **DH1**-**DH7** with 200-fold molar excess of **X1**:**X3**, **Y1**:**Y3**, or **DY1**:**DY4**. Experimental conditions for electrophoretic mobility shift assay: separately pre-annealed targets (34.4 nM) and probes (variable concentrations) were incubated 12-16h at ambient temperature in 1X HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 10% sucrose, 1.4 mM spermine tetrahydrochloride, pH 7.2) and then run on 16% non-denaturing PAGE (70V, 2.5h, ~4 °C) using 0.5x TBE as a running buffer (45 mM Tris, 45 mM boric acid, 1mM EDTA); DIG: digoxigenin.

6.2.7 A Change in Strategy – pcDNA with Energetic Hotspots

The above results suggest that incorporation of pseudocomplementary energetic hotspots (i.e., **DY** segments) does not result in major additional activation of the Invader probe and/or additional stabilization of probe-target duplexes as initially hypothesized. Based on the hypothesis that the full benefits of pseudocomplementary energetic hotspots are best harnessed by separating the two key structural features into distinct components, we set out to evaluate **DSX**-modified Invader probes, i.e., double-stranded probes featuring energetic hotspots comprised of the parent **X** monomers, as well as, regular pseudocomplementary base pairs between 2,6-diaminopurine **D** and 2-thiouracil **S** monomers (Table 6-4 and Figure 6-1). Toward this end, we synthesized 13-mer ONs allowing for the construction of: i) two different **DSX**-modified Invader probes, in which the energetic hotspot either is next to or one nucleotide away from two pseudocomplementary base pairs (**DSX1**:**DSX2** and **DSX3**:**DSX4**), ii) an Invader comprised of regular **X** monomers (**X5**:**X6**), iii) an Invader with a pseudocomplementary

energetic hotspot (**DY5**:**DY6**), and iv) three pcDNA, each containing two differentially spaced, regular pseudocomplementary base pairs (**SD1**:**SD2**, **SD3**:**SD4** and **SD5**:**SD6**).

First, T_m's for duplexes between individual probe strands and cDNA (Table 6-4, first two T_m columns) or cRNA (Table 6-S8) were determined. In line with our observations in the 9-mer series, **X**-modified ONs form highly thermostable duplexes with cDNA (ΔT_m = 11-12 °C) and less thermostable duplexes with cRNA (ΔT_m = 3 °C). By comparison, ONs with **DY**-motifs display slightly higher cRNA affinity (ΔT_m for **DY5**/**DY6** = 4.5-5.0 °C, Table 6-S8) but slightly lower cDNA affinity (ΔT_m = 10 °C, Table 6-4), again suggesting that the normally stabilizing base pairing between 2,6-diaminoadenosine:thymine and 2-thiouracil:adenine is weakened by pyrene intercalation. Interestingly, **DSX**-modified ONs display significantly increased cDNA and cRNA affinity relative to **X**-modified ONs ($\Delta T_{m,DNA}$ = 14-15 °C, Table 6-4; $\Delta T_{m,RNA}$ = 7.0-8.5 °C, Table 6-S8). Regular pcDNA strands form much less stabilized duplexes with cDNA ($\Delta T_{m,DNA}$ = 2.5-4.0 °C, Table 6-4).

As expected from the results in the 9-mer series, benchmark Invader probe **X5**:**X6** is relatively thermolabile (ΔT_m = -0.5 °C, Table 6-4). Also, Invader **DY5**:**DY6**, which features a pseudocomplementary energetic hotspot, is slightly less stable (ΔT_m = -1.5 °C for **DY5**:**DY6**, Table 6-4). Invaders with separated pseudocomplementary base pairs and energetic hotspots are more thermolabile still (ΔT_m = -2.0 °C and -3.0 °C for **DSX1**:**DSX2** and **DSX3**:**DSX4**, respectively, Table 6-4). Nonetheless, comparison with regular pcDNA reveals that the full destabilizing effect of the pseudocomplementary base pairs is not realized with the studied **DSX** probe architectures (e.g., compare ΔT_m of -8.0, -0.5 and -3.0 for **SD5**:**SD6**, **X5**:**X6** and **DSX3**:**DSX4**, respectively, Table 6-4).

The above T_m based conclusions are substantiated by the Gibbs free energies for formation of duplexes (Table 6-4). As a result of these stability trends, Invaders with separated energetic hotspots and pseudocomplementary base pairs are significantly more thermodynamically activated for dsDNA recognition than benchmark Invader **X5**:**X6**, **DY5**:**DY6** featuring a pseudocomplementary energetic hotspot or any of the regular pcDNA (trend in ΔG^{293}_{rec} values: **DSX3**:**DSX4**<**DSX1**:**DSX2**<<**X5**:**X6**<**DY5**:**DY6**<<**SD5:SD6**<**SD3:SD4**<**SD1:SD2**, Table 6-4).

Table 6-4: Thermal denaturation and thermodynamic properties of **X**-, **DY**-, **DSX**- and **DS**-modified duplexes.[a]

ON	Sequence	ΔT_m (°C)			ΔG^{293}[$\Delta\Delta G^{293}$] (kJ/mol)			ΔG^{293}_{rec} (kJ/mol)
		upper ON vs cDNA	lower ON vs cDNA	probe duplex	upper ON vs cDNA	lower ON vs cDNA	probe duplex	
DSX1 **DSX2**	5'-GGTA T**DX**A**S**A GGC 3'-CCAT A**S**A**XD**T CCG	+14.0	+14.0	-2.0	-76±1 [-15]	-78±1 [-17]	-52±0 [+9]	-41
DSX3 **DSX4**	5'-GGTA **S**A**X**AT**D** GGC 3'-CCAT **D**TA**X**A**S** CCG	+15.0	+14.0	-3.0	-80±1 [-19]	-77±2 [-16]	-52±0 [+9]	-44
DY5 **DY6**	5'-GGTA TA**YD**TA GGC 3'-CCAT AT**DY**AT CCG	+10.0	+10.0	-1.5	-74±2 [-13]	-73±1 [-12]	-55±1 [+6]	-31
X5 **X6**	5'-GGTA TA**X**ATA GGC 3'-CCAT ATA**X**AT CCG	+11.0	+12.0	-0.5	-76±2 [-15]	-78±1 [-17]	-59±0 [+2]	-34
SD1 **SD2**	5'-GGTA TA**SD**TA GGC 3'-CCAT AT**DS**AT CCG	+3.5	+2.5	-8.0	-64±1 [-3]	-64±1 [-3]	-50±1 [+11]	-17
SD3 **SD4**	5'-GGTA T**D**TA**S**A GGC 3'-CCAT A**S**AT**D**T CCG	+3.0	+2.5	-9.0	-65±2 [-4]	-64±1 [-3]	-50±1 [+11]	-18
SD5 **SD6**	5'-GGTA **S**ATAT**D** GGC 3'-CCAT **D**TATA**S** CCG	+4.0	+3.0	-8.0	-67±1 [-6]	-65±0 [-4]	-51±0 [+10]	-20

[a]ΔT_m = change in T_m relative to reference duplexes **DNA3**:**DNA4** ($T_m \equiv 37.5$ °C), where **DNA3**: 5′-GGTA TATATA GGC, **DNA4**: 3′-CCAT ATATAT CCG. $\Delta\Delta G^{293}$ is measured relative to ΔG^{293} for **DNA3**:**DNA4** = -61 kJ/mol. For definition of ΔG^{293}_{rec} see Table 6-3. "±" denotes standard deviation. For experimental conditions, see Table 6-1. For structures of monomers **X, Y**, **D** and **S**, see Figure 6-1.

6.2.8 Recognition of DNA Hairpins using Energetically Activated 13-mer Probe Duplexes

Based on the observed ΔG_{rec}^{293} values, we decided to experimentally test the dsDNA-targeting efficiency of **DSX1**:**DSX2**, **DSX3**:**DSX4**, **X5**:**X6** and **DY5**:**DY6** using a similar electrophoretic mobility shift assay as used in the 9-mer series. Hence, a DIG-labeled DNA hairpin (**DH8**) – comprised of a 13-mer double-stranded mixed-sequence stem, linked by a T_{10} loop – was used as a model target (Figure 6-4a). Incubation of **DH8** with the 13-mer Invaders results in dose-dependent formation of a slower moving band on non-denaturing PAGE gels (Figure 6-4b). As expected from the preliminary 9-mer studies, the parent Invader **X5**:**X6** recognizes **DH8** more efficiently at low probe concentrations than **DY5**:**DY6**, which has a pseudocomplementary energetic hotspot (Figure 6-4c). Gratifyingly, Invaders with separated pseudocomplementary base pairs and energetic hotspots display similar or improved recognition efficiency (see curves for **DSX1**:**DSX2** and **DSX3**:**DSX4**, respectively, Figure 6-4c), which follows the observed trend in ΔG_{rec}^{293} values. It is noteworthy that as little as 1.0 molar equivalents of **DSX3**:**DSX4** results in ~20% recognition of **DH8**, especially when considering that optimal Invader design normally calls for incorporation of multiple energetic hotspots.[37] This suggests that spatial separation of pseudocomplementary base pairs and energetic hotspots is a promising design principle for efficient dsDNA-targeting Invader probes.

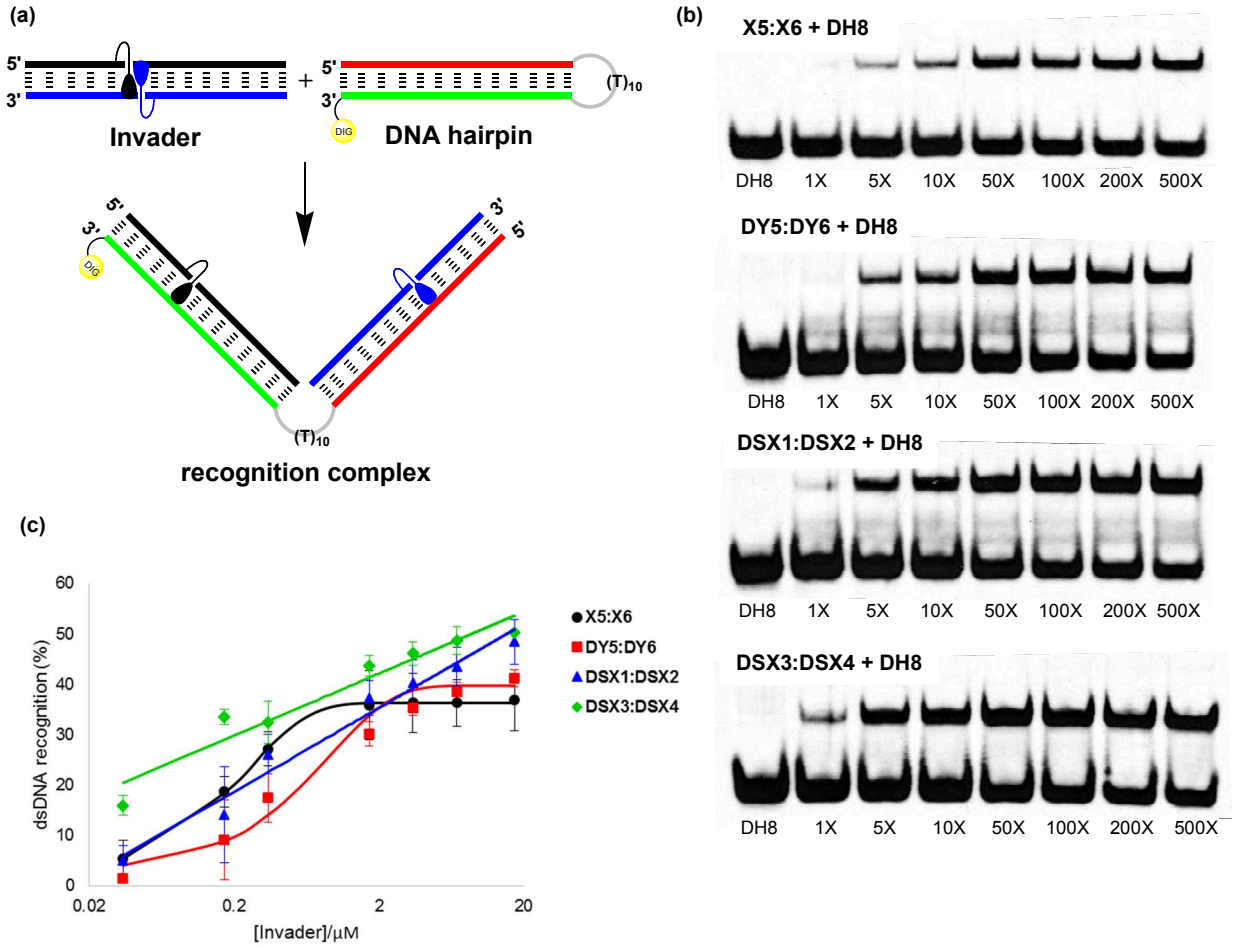


Figure 6-4: Recognition of dsDNA model target **DH8** using different Invader probes. (a) Illustration of recognition process; (b) representative electrophoretograms for recognition of **DH8** using 1- to 500-fold excess of **X5**:**X6**, **DY5**:**DY6** **DSX1**:**DSX2** or **DSX3**:**DSX4**; (c) dose-response curves (average of at least three independent experiments, error bars represent standard deviation). The sequence of DNA hairpin **DH8** is shown in Figure 6-5. For experimental conditions, see Figure 6-3.

Lastly, the binding specificities of these Invader probes were studied through incubation with DNA hairpins **DH9**-**DH14**, which differ in the nucleotide sequence at one position relative to

the probes (Figure 6-5a). Remarkably, the singly mismatched DNA hairpins were not recognized, even when using a 200-fold molar excess of **X5**:**X6**, **DY5**:**DY6** or **DSX1**:**DSX2** (Figure 6-5b). High-affinity Invader **DSX3**:**DSX4** shows trace recognition of **DH10** and **DH14** (<5% recognition). To understand why these mismatched targets are recognized, we performed a detailed binding specificity analysis of the individual strands of **X5**:**X6**, **DSX1**:**DSX2** and **DSX3**:**DSX4** (Tables 6-S9, 6-S10, and 6-S11 in the Supporting Information). Indeed, **DSX3** and **DSX4** display poor discrimination of mismatches in these positions (T_m's only reduced by 5-7 °C, relative to matched duplexes, Table 6-S11), rendering recognition of **DH10** and **DH14** possible at large probe excess. For additional discussion see the Supporting Information.

(a)

HP	Sequence	T_m (°C)
DH8	5' - GGTATATATAGGC 3' - CCATATATATCCG (T)$_{10}$	58.0
DH9	5' - GGTATTTATAGGC 3' - CCATAAATATCCG (T)$_{10}$	60.5
DH10	5' - GGTATGTATAGGC 3' - CCATACATATCCG (T)$_{10}$	63.5
DH11	5' - GGTATCTATAGGC 3' - CCATAGATATCCG (T)$_{10}$	63.0
DH12	5' - GGTATATAAAGGC 3' - CCATATATTTCCG (T)$_{10}$	60.0
DH13	5' - GGTATATAGAGGC 3' - CCATATATCTCCG (T)$_{10}$	62.5
DH14	5' - GGTATATACAGGC 3' - CCATATATGTCCG (T)$_{10}$	62.5

(b)

Figure 6-5. Recognition of mismatched DNA hairpins using activated double-stranded probes (a) sequences of DNA hairpins with isosequential (**DH8**) or non-isosequential stems (**DH9**-**DH14**); underlined nucleotides denote sequence deviations relative to Invader probes. (b) representative gel electrophoretograms from incubation of **DH8**-**DH14** with 200-fold molar excess of **X5**:**X6**, **DY5**:**DY6**, **DSX1**:**DSX2**, or **DSX3**:**DSX4**. For experimental conditions, see Figure 6-3.

6.3 Conclusions

A short synthetic route to a 2′-*N*-(pyren-1-yl)methyl-2′-amino-2′-deoxy-2′-*N*-methyl-2-thiouridine phosphoramidite was developed along with suitable conditions for introduction into ONs. ONs modified with these monomers display high affinity toward cDNA. Incorporation

of 2-aminoadenosine monomers next to 2′-*N*-(pyren-1-yl)methyl-2′-amino-2′-deoxy-2′-*N*-methyl-2-thiouridine resulted in additional increases in affinity toward cDNA. Invaders with pseudocomplementary hotspots (i.e., **DY**-probes) are thermolabile and inherently activated for recognition of dsDNA, but less so than Invaders with regular hotspots composed of 2′-*N*-(pyren-1-yl)methyl-2′-amino-2′-deoxy-2′-*N*-methyl-2-uridine. This is a result of the close proximity of the two destabilizing structural motifs (i.e., intercalation-mediated violation of the nearest-neighbor principle and pseudocomplementary base pairs) apparently not being fully compatible. Spatial separation of the two destabilizing structural motifs (i.e., **DSX**-probes) results in more strongly activated Invader probes, which allow for more efficient recognition of model dsDNA targets with excellent specificity. Thus, incorporation of regular energetic hotspots into pcDNA probes is a promising strategy to improve the dsDNA recognition potential of pcDNA and Invader probes. Ongoing recognition experiments in molecular biological contexts will clarify their potential as agents for regulation of gene expression and application in nucleic acid diagnostics.

6.4 Experimental Section

6.4.1 Synthesis of 2′-*N*-(pyren-1-yl)methyl-2′-amino-2′-deoxy-2′-*N*-methyl-2-thiouridine Phosphoramidite

3′-*O*-Acetyl-2′-amino-2′-deoxy-5′-*O*-methanesulfonyl-2′-*N*-methyl-2′-*N*-(pyren-1-ylmethyl)uridine (2). Nucleoside **1**[28] (3.40 g, 4.39 mmol) was coevaporated with anhydrous 1,2-dichloroethane (2 × 30 mL), redissolved in anhydrous pyridine (55 mL) and cooled to 0 °C. To this was added 4-dimethylaminopyridine (DMAP, 55 mg, 0.44 mmol) and acetic anhydride

(1.25 mL, 13.18 mmol). After stirring at ambient temperature for 12 h, the reaction mixture was diluted with EtOAc (150 mL) and washed with H_2O (80 mL) and saturated aqueous $NaHCO_3$ (80 mL). The organic layer was evaporated to near dryness and coevaporated with absolute EtOH:toluene (2:1 v/v, 3 × 30 mL). The resulting residue (~3.5 g) was suspended in AcOH:H_2O (4:1 v/v, 55 mL) and the reaction mixture was stirred overnight at room temperature. Solvents were evaporated off and the resulting residue was purified by silica gel column chromatography (0-5% MeOH/CH_2Cl_2, v/v) to afford O5′-hydroxy derivative (~1.8 g) as a white solid material. This material was coevaporated with anhydrous pyridine:CH_2Cl_2 (1:1 v/v, 2 × 30 mL), redissolved in anhydrous pyridine:CH_2Cl_2 (1:1 v/v, 44 mL) and cooled to -20 °C (ice/salt). Methanesulfonyl chloride (0.75 mL, 9.62 mmol) was added dropwise over 30 min and the reaction mixture was stirred at -20 °C for 2 h more. The reaction mixture was diluted with CH_2Cl_2 (80 mL) and washed with saturated aqueous $NaHCO_3$ (50 mL). The aqueous layer was back-extracted with CH_2Cl_2 (3 × 15 mL) and the organic layers were combined and evaporated to dryness, followed by coevaporation with absolute EtOH:toluene (2:1 v/v, 3 × 30 mL). The resulting residue was purified by silica gel column chromatography (0-3% MeOH/CH_2Cl_2, v/v) to afford nucleoside **2** (1.43 g, 55%) as a white foam. R_f = 0.4 (5% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m*/*z* 614.1600 ($[M+Na]^+$, $C_{30}H_{29}N_3O_8S{\cdot}Na^+$, Calc. 614.1568); ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.47 (d, 1H, ex, J = 2.2 Hz, NH), 8.33-8.36 (d, 1H, J = 9.0 Hz, Py), 8.25-8.29 (ap t, 2H, J = 7.6 Hz, Py), 8.21 (d, 1H, J = 7.5 Hz, Py), 8.15 (br s, 2H, Py), 8.10-8.12 (d, 1H, J = 9.0 Hz, Py), 8.04-8.08 (t, 1H, J = 7.6 Hz, Py), 7.92 (d, 1H, J = 7.5 Hz, Py), 7.73 (d, 1H, J = 8.2 Hz, H6), 6.40 (d, 1H, J = 7.4 Hz, H1′), 5.67 (dd, 1H, J = 8.2 Hz, 2.2 Hz, H5), 5.39 (dd, 1H, J = 6.6 Hz, 3.6 Hz, H3′), 4.44-4.48 (m, 2H, H5′), 4.34-4.43 (m, 3H, CH_2Py, H4′), 3.84 (ap t, 1H, J ~ 7.5 Hz, H2′), 3.20 (s, 3H, CH_3SO_2), 2.34 (s, 3H, CH_3N), 2.13 (s, 3H, CH_3CO); ^{13}C

NMR (125 MHz, DMSO-d_6) δ 169.7, 162.7, 150.5, 140.6 (C6), 131.9, 130.7, 130.3, 130.2, 129.1, 127.7 (Py), 127.3 (Py), 127.0 (Py), 126.8 (Py), 126.1 (Py), 125.1 (Py), 124.4 (Py), 124.2, 123.8, 123.7 (Py), 102.6 (C5), 83.7 (C1′), 80.0 (C4′), 71.7 (C3′), 69.1 (C5′), 64.9 (C2′), 57.5 (CH_2Py), 37.7 (CH_3N), 36.8 (CH_3SO_2), 20.9 (CH_3CO).

2′-Amino-2′-deoxy-2-*O*-ethoxy-2′-*N*-methyl-2′-*N*-(pyren-1-ylmethyl)uridine (3). Nucleoside **2** (0.77 g, 1.30 mmol) was dried through coevaporation with anhydrous 1,2-dichloroethane (3 × 6 mL) and suspended in anhydrous EtOH (25 mL). To this was added $NaHCO_3$ (275 mg, 3.25 mmol) and the reaction mixture was refluxed for 4 days. CH_2Cl_2 (100 mL) was added and the precipitate was filtered off and washed with CH_2Cl_2. The combined organic layers were evaporated to dryness and the resulting residue was purified via silica gel column chromatography (0-7% MeOH in CH_2Cl_2, v/v) to afford nucleoside **3** (0.34 g, 52%) as a white foam. R_f = 0.3 (10% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m*/*z* 522.1985 ($[M+Na]^+$, $C_{29}H_{29}N_3O_5{\cdot}Na^+$, Calc. 522.1999); ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.39-8.42 (d, 1H, J = 9.3 Hz, Py), 8.24-8.28 (m, 2H, Py), 8.19-8.22 (d, 1H, J = 8.0 Hz, Py), 8.14 (br s, 2H, Py), 8.03-8.10 (d+t, 2H, Py), 7.96-7.99 (d, 1H, J = 8.0 Hz, Py), 7.93 (d, 1H, J = 7.7 Hz, H6), 6.35 (d, 1H, J = 8.5 Hz, H1′), 5.83 (d, 1H, J = 7.7 Hz, H5), 5.53 (d, 1H, ex, J = 4.9 Hz, 3′-OH), 5.12 (t, 1H, ex, J = 5.2 Hz, 5′-OH), 4.43-4.51 (m, 3H, CH_2Py, H3′), 4.12-4.27 (m, 2H, O$\underline{CH_2}CH_3$), 3.99-4.01 (m, 1H, H4′), 3.58-3.62 (m, 2H, H5′), 3.47 (dd, 1H, J = 8.5 Hz, 5.2 Hz, H2′), 2.36 (s, 3H, CH_3N), 1.06 (t, 3H, J = 7.1 Hz, $\underline{CH_3}CH_2O$); ^{13}C NMR (125 MHz, DMSO-d_6) δ 169.3, 155.0, 138.1 (C6), 132.7, 130.7, 130.3, 130.1, 129.0, 127.5 (Py), 127.3 (Py), 126.9 (Py), 126.8 (Py), 126.1 (Py), 125.02 (Py), 124.99 (Py), 124.4 (Py), 124.1, 123.8, 123.6 (Py), 108.4 (C5), 87.6

(C4′), 85.0 (C1′), 71.2 (C3′), 68.8 (C2′), 64.2 (O<u>CH$_2$</u>CH$_3$), 61.7 (C5′), 57.1 (CH_2Py), 39.0 (CH_3N – overlap with DMSO-d_6 signal), 13.6 (<u>CH$_3$</u>CH_2O).

2′-Amino-2′-deoxy-5′-*O*-(4,4′-dimethoxytrityl)-2-*O*-ethoxy-2′-*N*-methyl-2′-*N*-(pyren-1-ylmethyl)uridine (4). Nucleoside **3** (0.32 g, 0.63 mmol) was coevaporated with anhydrous 1,2-dichloroethane (2 × 5 mL) and redissolved in anhydrous pyridine (6 mL). To this was added DMTrCl (0.26 g, 0.76 mmol) and 4-dimethylaminopyridine (DMAP, 8 mg, 0.06 mmol) and the reaction mixture was stirred at ambient temperature for 14 h at which point it was diluted with $CHCl_3$ (80 mL) and washed with saturated aqueous $NaHCO_3$ (30 mL) and H_2O (30 mL). The aqueous layer was back-extracted with $CHCl_3$ (3 × 15 mL) and the combined organic layers were dried (Na_2SO_4) and evaporated to dryness. The resulting residue was purified by silica gel column chromatography (0-2.5 % MeOH in $CHCl_3$, v/v) to afford nucleoside **4** (0.45 g, 88%) as a pale orange foam. R_f = 0.7 (10% MeOH in CH_2Cl_2, v/v); MALDI-HRMS *m*/*z* 824.3319 ($[M+Na]^+$, $C_{50}H_{47}N_3O_7Na^+$, Calc. 824.3306); ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.38-8.41 (d, 1H, J = 9.3 Hz, Py), 8.25-8.29 (m, 2H, Py), 8.17-8.19 (d, 1H, J = 7.8 Hz, Py), 8.11-8.15 (2d, 2H, J = 9.1 Hz, 9.1 Hz, Py), 8.04-8.09 (m, 2H, Py), 7.98-8.00 (d, J = 7.8 Hz, Py), 7.69 (d, 1H, J = 7.8 Hz, H6), 7.19-7.38 (m, 9H, DMTr), 6.83-6.88 (m, 4H, DMTr), 6.32 (d, 1H, J = 7.9 Hz, H1′), 5.60-5.64 (m, 2d, 1 ex, J = 7.8 Hz, 5.0 Hz, H5, 3′-OH), 4.48-4.54 (m, 2H, CH_2Py), 4.44-4.48 (m, 1H, H3′), 4.10-4.24 (m, 3H, H4′, O<u>CH$_2$</u>CH_3), 3.71 (s, 3H, CH_3O), 3.70 (s, 3H, CH_3O), 3.52-3.58 (dd, 1H, J = 7.9 Hz, 5.8 Hz, H2′), 3.31-3.35 (dd, 1H, J = 10.5 Hz, 5.0 Hz, H5′ - partial overlap with H_2O), 3.15-3.20 (dd, 1H, J = 10.5 Hz, 3.5 Hz, H5′), 2.43 (s, 3H, CH_3N), 1.05 (t, 3H, J = 7.0 Hz, <u>CH$_3$</u>CH_2O); ^{13}C NMR (125 MHz, DMSO-d_6) δ 169.2, 158.09, 158.08, 154.9, 144.5, 138.1 (C6), 135.3, 135.0, 132.6, 130.7, 130.2, 130.1, 129.7 (DMTr), 129.6 (DMTr),

129.0, 127.8 (DMTr), 127.6 (DMTr), 127.5 (Py), 127.3 (Py), 126.91 (Py), 126.86 (Py), 126.7 (DMTr), 126.1 (Py), 125.0 (Py), 124.4 (Py), 124.1, 123.9, 123.6 (Py), 113.20 (DMTr), 113.17 (DMTr), 108.1 (C5), 85.9, 85.7 (C4′), 85.3 (C1′), 71.0 (C3′), 68.1 (C2′), 64.2 ($O\underline{CH_2}CH_3$), 64.0 (C5′), 57.1 (CH_2Py), 55.0 (CH_3O), 38.8 (CH_3N), 13.6 ($\underline{CH_3}CH_2O$). A trace impurity of $CHCl_3$ was identified in the ^{13}C NMR at 79.1 ppm.

2′-Amino-2′-deoxy-5′-*O*-(4,4′-dimethoxytrityl)-2′-*N*-methyl-2′-*N*-(pyren-1-ylmethyl)-2-thiouridine (5). An ice-cold solution of anhydrous 1,1,3,3-tetramethylguanidine (TMG, 0.68 mL, 5.42 mmol) in anhydrous pyridine (10 mL) was saturated with hydrogen sulfide gas for 1 h while maintaining the temperature at 0 °C. The solution was transferred, using an argon-flushed syringe, to a pre-cooled flask containing nucleoside **4** (0.44 g, 0.54 mmol) and the reaction mixture was allowed to reach room temperature. After stirring under an argon atmosphere for 72 h, EtOAc (100 mL) was added and the organic layer was washed with saturated aqueous $NaHCO_3$ (50 mL) and H_2O (50 mL). The aqueous layer was back-extracted with CH_2Cl_2 (3 × 20 mL) and the combined organic layers were evaporated to dryness and coevaporated with absolute EtOH:toluene (2:1 v/v, 3 × 15 mL). The resulting residue was purified by silica gel column chromatography (0-70% EtOAc in petroleum ether, v/v) to afford nucleoside **5** (0.35 g, 82%) as a white foam. R_f = 0.8 (80% EtOAc in petroleum ether, v/v); MALDI-HRMS *m/z* 812.2765 ($[M+Na]^+$, $C_{48}H_{43}N_3O_6S{\cdot}Na^+$, Calc. 812.2797); ^1H NMR (500 MHz, DMSO-d_6) δ 12.77 (br s, 1H, ex, NH), 8.49-8.52 (d, 1H, J = 9.3 Hz, Py), 8.24-8.29 (m, 2H, Py), 8.17-8.20 (d, 1H, J = 8.0 Hz, Py), 8.12-8.16 (2d, 2H, J = 9.1 Hz, 9.1 Hz, Py), 8.02-8.11 (m, 3H, Py), 7.75 (d, 1H, J = 8.2 Hz, H6), 7.20-7.39 (m, 10H, DMTr, H1′), 6.84-6.89 (m, 4H, DMTr), 5.62 (dd, 1H, J = 8.2 Hz, 1.7 Hz, H5), 5.55 (d, 1H, ex, J = 4.9 Hz, 3′-OH), 4.53-

4.57 (d, 1H, J = 13.0 Hz, CH_2Py), 4.43-4.51 (m, 2H, CH_2Py, H3′), 4.08-4.12 (m, 1H, H4′), 3.71 (s, 3H, CH_3O), 3.70 (s, 3H, CH_3O), 3.48-3.52 (m, 1H, H2′), 3.33-3.37 (dd, 1H, J = 10.5 Hz, 5.0 Hz, H5′), 3.18-3.22 (dd, 1H, J = 10.5 Hz, 3.0 Hz, H5′), 2.44 (s, 3H, CH_3N); ^{13}C NMR (125 MHz, DMSO-d_6) δ 176.6, 159.0, 158.11, 158.10, 144.5, 140.8 (C6), 135.3, 135.0, 132.6, 130.7, 130.3, 130.2, 129.73 (DMTr), 129.68 (DMTr), 129.2, 128.0 (Py), 127.9 (DMTr), 127.6 (DMTr), 127.3 (Py), 127.0 (Py), 126.8 (Py), 126.7 (DMTr), 126.1 (Py), 125.02 (Py), 125.00 (Py), 124.4 (Py), 124.1, 124.0 (Py), 123.9, 113.3 (DMTr), 113.2 (DMTr), 106.9 (C5), 88.0 (C1′), 86.0, 85.4 (C4′), 71.0 (C3′), 68.6 (C2′), 63.9 (C5′), 57.6 (CH_2Py), 55.0 (CH_3O), 39.2 (CH_3N – overlap with DMSO-d_6).

2′-Amino-2′-deoxy-3′-*O*-(*N*,*N*-diisopropylamino-2-cyanoethoxyphosphinyl)-5′-*O*-(4,4′-dimethoxytrityl)-2′-*N*-methyl-2′-*N*-(pyren-1-ylmethyl)-2-thiouridine (6). To a flame-dried round-bottomed flask containing nucleoside **5** (150 mg, 0.19 mmol) was added anhydrous CH_2Cl_2 (2 mL), anhydrous *N*,*N*-diisopropylethylamine (DIPEA, 165 μL, 0.95 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (PCl reagent, 85 μL, 0.38 mmol). The reaction mixture was stirred at room temperature for 3.5 h at which point ice-cold EtOH (1.0 mL) was added. The reaction mixture was evaporated to dryness and the resulting residue was purified by silica gel column chromatography (0-55% EtOAc in petroleum ether, v/v) followed by precipitation from cold petroleum ether to afford nucleoside **6** (153 mg, 81%) as a white foam. R_f = 0.6 (50% EtOAc in petroleum ether, v/v); MALDI-HRMS *m/z* 1012.3843 ($[M+Na]^+$, $C_{57}H_{60}N_5O_7PS{\cdot}Na^+$, Calc. 1012.3887); ^{31}P NMR (121 MHz, $CDCl_3$) δ 150.9, 149.6.

6.4.2 Protocol - Synthesis and Purification of ONs

Modified ONs were synthesized on a 0.2 μmol scale using a DNA synthesizer, succinyl linked LCAA-CPG (long chain alkyl amine controlled pore glass) columns with a pore size of 500Å, and standard protocols for incorporation of A^{Bz}, C^{Bz}, G^{iBu} and T DNA phosphoramidites. The following hand-coupling conditions were used for incorporation of the corresponding phosphoramidites of monomers **X**, **Y**, **S** (N3/O4-toluoyl protected) and **D** (bis(diisobutylaminomethylidene) protected) (coupling time; activator; coupling yield): **X** (15 min; 5-[3,5-bis(trifluoromethyl)phenyl]-1*H*-tetrazole; ~99%), **Y** (15 min; 4,5-dicyanoimidazole; ~95%) and **S/D** (15 min; 4,5-dicyanoimidazole; ~99%). Modified phosphoramidites were used at 50-fold molar excess and 0.05 M concentration in CH_3CN. Extended oxidation (45s) with standard 0.05 M aqueous iodine was used for **D1**-**D4** and **X1**-**X6**. Extended oxidation (2 × 5 min oxidation with an acetonitrile wash between oxidations) using a *tert*-butylhydroperoxide/CH_3CN/H_2O solution (10/87/3, v/v/v) was used for all ONs containing **S** and **Y** modifications to prevent desulfurization.[30] Cleavage from solid support and removal of protecting groups was accomplished upon treatment with 32% aq. ammonia (55 °C, 16-24 h). ONs were purified in the DMT-on mode via ion-pair reverse phase HPLC (C_{18} column) using a 0.05 M triethylammonium acetate - water/acetonitrile gradient. This was followed by detritylation (80% aq. AcOH) and precipitation (NaOAc/$NaClO_4$/acetone, -18 °C for 12-16 h). The identity of synthesized ONs was established through MALDI-MS analysis (Table 6-S1) recorded in positive ions mode on a quadrupole time-of-flight tandem mass spectrometer equipped with a MALDI source using anthranilic acid, 3-hydroxypicolinic acid (3-HPA) or 2′,4′,6′-trihydroxyacetophenone (THAP) as matricies. Purity was verified by ion-pair reverse phase HPLC running in analytical mode (>80% unless otherwise mentioned).

6.4.3 Protocol - Thermal Denaturation Studies

ON concentrations were estimated using the following extinction coefficients for DNA (OD/μmol): G (12.01), A (15.20), T (8.40), C (7.05); RNA (OD/μmol): G (13.70), A (15.40), U (10.00), C (9.00); pyrene (22.4)[38], **D** (8.5)‡, **S** (10.0)‡ and **Y** (32.4) ‡. Strands were thoroughly mixed and denatured by heating to 70-85 °C, followed by cooling to the starting temperature of the experiment. Quartz optical cells with a path length of 1.0 cm were used. Thermal denaturation temperatures (T_m's) of duplexes (1.0 μM final concentration of each strand) were measured using a UV/Vis spectrophotometer equipped with a 12-cell Peltier temperature controller and determined as the maximum of the first derivative of thermal denaturation curves (A_{260} vs. T) recorded in medium salt phosphate buffer (T_m buffer: 100 mM NaCl, 0.1 mM EDTA and pH 7.0 adjusted with 10 mM Na_2HPO_4 and 5 mM Na_2HPO_4). The temperature of the denaturation experiments ranged from at least 15 °C below T_m to 20 °C above T_m (although not below 3 °C). A temperature ramp of 0.5 °C/min was used in all experiments. Reported T_m's are averages of two experiments within ± 1.0 °C.

6.4.4 Protocol - Determination of Thermodynamic Parameters

Thermodynamic parameters for duplex formation were determined through fitting of baselines of denaturation curves (van't Hoff analysis) using software provided with the UV/Vis spectrometer. Bimolecular reactions, two-state melting behavior, and a heat capacity change of $\Delta C_p = 0$ upon hybridization were assumed.[36] A minimum of two experimental denaturation

‡Extinction coefficients for **D** and **S** obtained from Glen Research, whereas **Y** is the sum of **S** and Py.

curves were each analyzed at least three times to minimize errors arising from baseline choice. Averages and standard deviations are listed.

6.4.5 Protocol - Absorption Spectra

UV-vis absorption spectra (range 200-600 nm) were recorded at 10 °C using the same samples and instrumentation as in the thermal denaturation experiments.

6.4.6 Protocol - Steady-state Fluorescence Emission Spectra

Steady-state fluorescence emission spectra of **Y**- or **DY**-modified ONs and the corresponding duplexes with complementary DNA/RNA targets, were recorded in non-deoxygenated thermal denaturation buffer (each strand at 1.0 μM concentration) and obtained as an average of five scans using an excitation wavelength of λ_{ex} = 350 nm. Excitation and emission slits of 5.0 nm and 2.5 nm, respectively, were used along with a scan speed of 600 nm/min. Experiments were determined at 5 °C (to ascertain maximal hybridization of probes to DNA/RNA targets) and under N_2 flow (to prevent condensation).

6.4.7 Protocol - Electrophoretic Mobility Shift Assay

This assay was performed essentially as previously described.[23] Unmodified DNA hairpins **DH1**-**DH14** were obtained from commercial sources and used without further purification. The DNA hairpins were 3′-DIG-labeled using the 2nd generation DIG Gel Shift Kit (Roche Applied Bioscience) following the manufacturer's recommendation. DIG-labeled ONs obtained in this manner were diluted and used without further purification in the recognition experiments. Pre-annealed probes (85 °C for 10 min, cooled to room temperature over 15 min) and DIG-labeled

DNA hairpins (34.4 nM) were mixed and incubated in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 10% sucrose, 1.44 mM spermine tetrahydrochloride, pH 7.2) for the specified time at ambient temperature (~21±3 °C). The reaction mixtures were then diluted with 6x DNA loading dye (Fermentas) and loaded onto a 16% non-denaturing polyacrylamide gel. Electrophoresis was performed using a constant voltage of 70 V for 2.5 h at ~4 °C using 0.5x TBE as a running buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA). Gels were blotted onto positively charged nylon membranes (Roche Applied Bioscience) using constant voltage with external cooling (100V, ~4 °C). The membranes were exposed to anti-digoxigenin-AP F_{ab} fragments as recommended by the manufacturer of the DIG Gel Shift Kit, transferred to a hybridization jacket, and incubated with the substrate (CSPD) in detection buffer for 10 min at 37 °C. The chemiluminescence of the formed product was captured on X-ray film, which was developed using an X-Omatic 1000A X-ray film developer (Kodak). The resulting bands were quantified using Image J software. Invasion efficiency was determined as the intensity ratio between the recognition complex band and the total lane. An average of three independent measurements is reported along with standard deviations. Non-linear regression was used to fit data points from dose-response experiments, using a script written for the "Solver" module in Microsoft Office Excel.[39]

6.4.8 Definition of Zipper Nomenclature

The following nomenclature describes the relative arrangement between two pyrene-functionalized monomers positioned on opposing strands in a duplex. The number *n* describes the distance measured in number of base pairs and has a positive value if a monomer is shifted toward the 5′-side of its own strand relative to a second reference monomer on the other strand. Conversely, *n* has a negative value if a monomer is shifted toward the 3′-side of its own strand relative to a second reference monomer on the other strand.

6.5 Supporting Information

6.5.1 General Experimental Section

Analytical grade reagents and solvents were obtained from commercial sources and used without further purification. Petroleum ether of the distillation range 60-80 °C was used. Solvents were dried over activated molecular sieves: CH_3CN, EtOH (3Å); CH_2Cl_2, 1,2-dichloroethane, pyridine, 1,1,3,3-tetramethylguanidine and *N,N*-diisopropylethylamine (4Å). The water content of anhydrous solvents was verified on Karl-Fisher apparatus. Reactions were conducted under an argon atmosphere whenever anhydrous solvents were used. Reactions were monitored by TLC using silica gel coated plates with a fluorescence indicator (SiO_2-60, F-254) which were visualized a) under UV light and/or b) by dipping in 5% conc. H_2SO_4 in absolute ethanol (v/v) followed by heating. Silica gel column chromatography was performed with Silica gel 60 (particle size 0.040–0.063 mm) using moderate pressure (pressure ball). Evaporation of solvents was carried *in vacuo* at temperatures below 45 °C. After column chromatography,

appropriate fractions were pooled, evaporated and dried at high vacuum for at least 12 h to provide products in high purity (>95%) as ascertained by 1D NMR techniques. Exchangeable (ex) protons were detected by disappearance of signals upon D_2O addition. Assignments of NMR spectra are based on 2D spectra (COSY, HSQC) and DEPT-spectra. Quaternary carbons are not assigned in ^{13}C NMR but verified from HSQC and DEPT spectra (absence of signals). MALDI-HRMS spectra of compounds were recorded on a mass spectrometer using 2,5-dihydroxybenzoic acid as a matrix and polyethylene glycol (PEG 600) as an internal calibration standard.

6.5.2 Additional Tables, Figures, and Discussion

Table 6-S1: MALDI-MS of modified ONs.[a]

ONs	Sequence	Calc. *m/z* [M+H]	Found *m/z* [M+H]
Y1	5′-GTG A**Y**A TGC	2999	2999
Y2	3′-CAC **Y**AT ACG	2928	2928
Y3	3′-CAC TA**Y** ACG	2928	2928
Y4	3′-CAC **Y**A**Y** ACG	3172	3172
DY1	5′-GTG A**YD** TGC	3014	3015
DY2	5′-GTG **DY**A TGC	3014	3014
DY3	3′-CAC **YD**T ACG	2943	2943
DY4[b]	3′-CAC T**DY** ACG	2943	2943
D1	5′-GTG **D**TA TGC	2769	2769
D2	5′-GTG AT**D** TGC	2769	2770
D3	3′-CAC T**D**T ACG	2698	2699
D4	5′-GTG **D**T**D** TGC	2784	2785
DY5[b]	5′-GGT ATA **YD**T AGG C	4273	4273
DY6[b]	3′-CCA TAT **DY**A TCC G	4153	4153
X5	5′-GGT ATA **X**AT AGG C	4241	4242
X6	3′-CCA TAT A**X**A TCC G	4121	4122
DSX1	5′-GGT AT**D** **X**A**S** AGG C	4273	4273
DSX2	3′-CCA TA**S** A**XD** TCC G	4153	4153
DSX3	5′-GGT A**S**A **X**AT **D**GG C	4273	4273
DSX4	3′-CCA T**D**T A**X**A **S**CC G	4153	4153
SD1	5′-GGT ATA **SD**T AGG C	4044	4044
SD2	3′-CCA TAT **DS**A TCC G	3924	3924
SD3	5′-GGT AT**D** TA**S** AGG C	4044	4044
SD4	3′-CCA TA**S** AT**D** TCC G	3924	3924
SD5	5′-GGT A**S**A TAT **D**GG C	4044	4044
SD6	3′-CCA T**D**T ATA **S**CC G	3924	3924

[a]For structures of monomers **X**, **Y**, **S,** and **D** see Figure 6-1 in the main manuscript. **X1**-**X4** were made in a previous study.[24]

[b]Synthesis of ONs with 5′-**YD**-3′ segments generally proved to be challenging, resulting in low overall yields and purity (**DY4**, **DY5** and **DY6** obtained in 75%, 60% and 70% purity), presumably due to desulfurization.

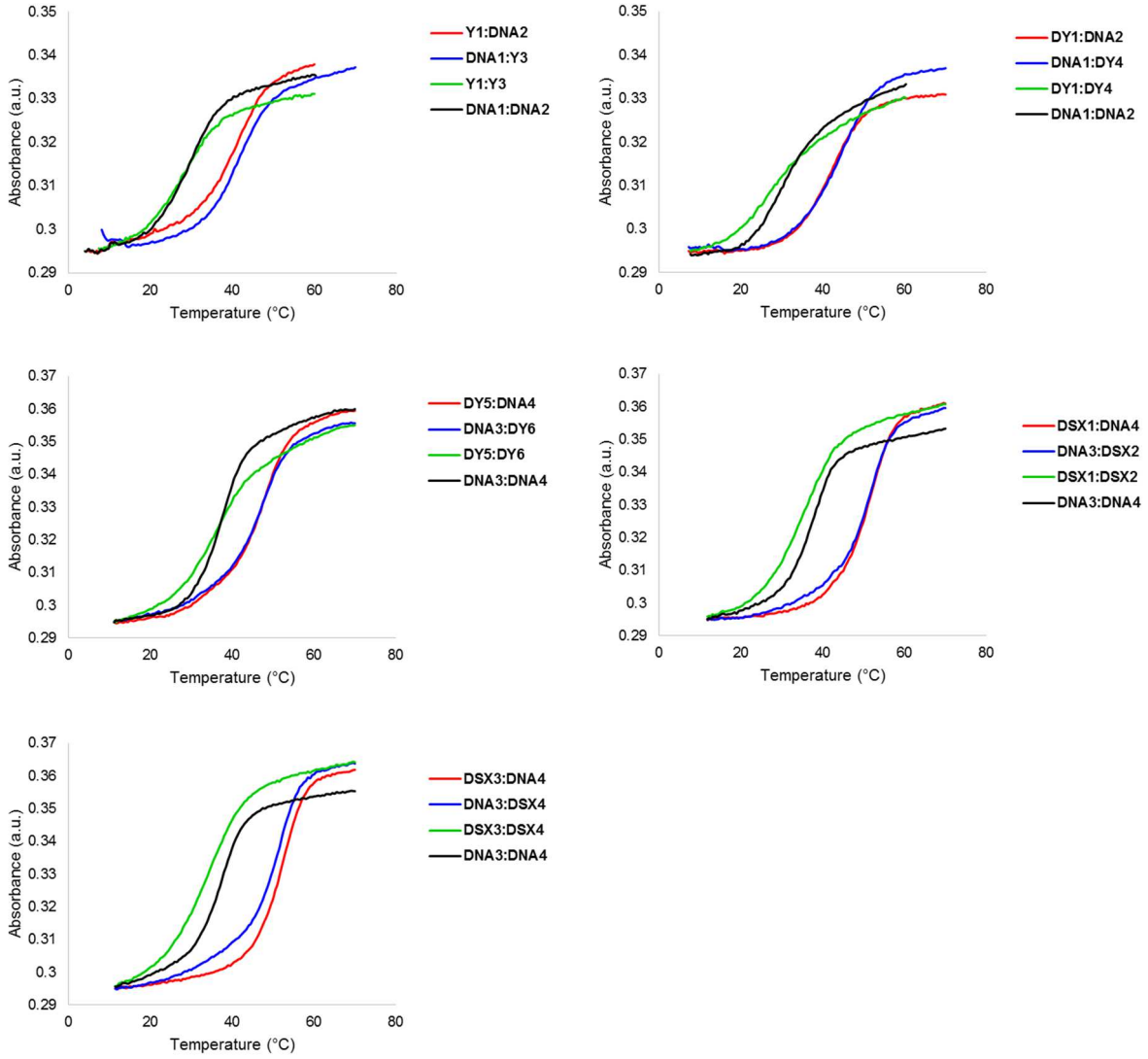


Figure 6-S1: Representative thermal denaturation curves of **Y**-, **DY**- and **DSX**-modified duplexes and reference duplexes. For experimental conditions, see Table 6-1.

Table 6-S2: Thermal denaturation temperatures of duplexes between **X**-, **Y**- or **DY**-modified ONs and cRNA.[a]

ON	Sequence	ΔT_m (°C)
X1[b]	5′-GTG A**X**A TGC	+3.0
X2[b]	3′-CAC **X**AT ACG	-6.5
X3[b]	3′-CAC TA**X** ACG	+3.0
X4[b]	3′-CAC **X**A**X** ACG	-3.0
Y1	5′-GTG A**Y**A TGC	+2.0
Y2	3′-CAC **Y**AT ACG	-6.5
Y3	3′-CAC TA**Y** ACG	+2.5
Y4	3′-CAC **Y**A**Y** ACG	-6.5
DY1	5′-GTG A**YD** TGC	+3.0
DY2	5′-GTG **DY**A TGC	+4.5
DY3	3′-CAC **Y**DT ACG	-3.5
DY4	3′-CAC TD**Y** ACG	+5.5

[a]ΔT_m = change in T_m relative to reference duplexes **DNA1**:**RNA2** ($T_m \equiv 27.5$ °C) and **RNA1**:**DNA2** ($T_m \equiv 27.5$ °C), where **DNA1**: 5′-GTG ATA TGC, **DNA2**: 3′-CAC TAT ACG, **RNA1**: 5′-GUG AUA UGC and **RNA2**: 3′-CAC UAU ACG; T_m's are determined as the maximum of the first derivative of melting curves (A_{260} vs T) recorded in medium salt phosphate buffer ($[Na^+] = 110$ mM, $[Cl^-] = 100$ mM, pH 7.0 (NaH_2PO_4/Na_2HPO_4)), using 1.0 µM of each strand. Reported T_m's are averages of at least two measurements within 1.0 °C; A = adenin-9-yl DNA monomer, C = cytosin-1-yl DNA monomer, G = guanin-9-yl DNA monomer, T = thymin-1-yl DNA monomer. For structures of monomers **X, Y**, **S, D**, see Figure 6-1.

[b]Data previously reported in reference 28.

Table 6-S3: DNA selectivity of **X**-, **Y**- or **DY**-modified ONs.[a]

ON	Sequence	$\Delta\Delta T_m$ (DNA-RNA) (°C)
X1[b]	5′-GTG A**X**A TGC	+12.0
X2[b]	3′-CAC **X**AT ACG	+8.0
X3[b]	3′-CAC TA**X** ACG	+12.0
X4[b]	3′-CAC **X**A**X** ACG	+17.0
Y1	5′-GTG A**Y**A TGC	+9.5
Y2	3′-CAC **Y**AT ACG	+9.0
Y3	3′-CAC TA**Y** ACG	+8.5
Y4	3′-CAC **Y**A**Y** ACG	+16.5
DY1	5′-GTG A**YD** TGC	+13.0
DY2	5′-GTG **DY**A TGC	+8.5
DY3	3′-CAC **YD**T ACG	+7.0
DY4	3′-CAC T**DY** ACG	+10.5

[a] DNA selectivity defined as $\Delta\Delta T_m$ (DNA-RNA) = ΔT_m (vs cDNA) - ΔT_m (vs cRNA).

[b]Data previously reported in reference 24; included to facilitate comparison.

Table 6-S4: Discrimination of mismatched DNA targets by **X4**, **Y4** and reference ONs.[a]

			DNA: 5′-GTG A**B**A TGC			
			T_m [°C]	ΔT_m [°C]		
ON	**Sequence**	**B** =	**T**	A	C	G
DNA2	3′-CAC TAT ACG		**29.5**	-17.0	-15.5	-9.0
X4[b]	3′-CAC **X**A**X** ACG		**43.5**	-21.5	-10.5	-13.5
Y4	3′-CAC **Y**A**Y** ACG		**39.5**	-22.0	-12.0	-11.0

[a]For conditions of thermal denaturation experiments, see Table 6-1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched DNA:DNA duplex.

[b]Data from reference 24.

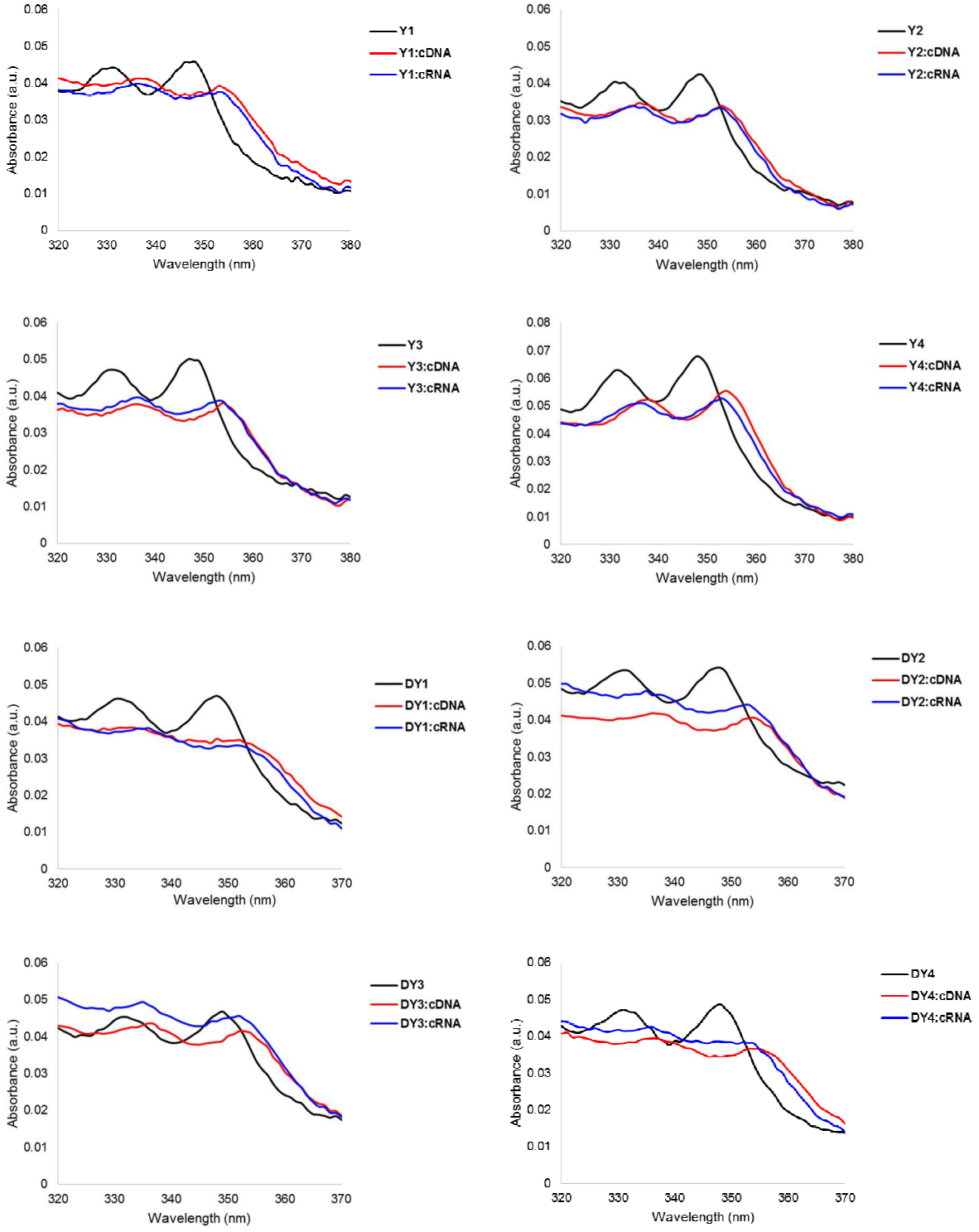


Figure 6-S2: Absorption spectra of single-stranded **Y**- or **DY**-modified ONs and the corresponding duplexes with cDNA/cRNA targets. Spectra were recorded at T = 10 °C using

each strand at 1.0 μM concentration in T_m buffer. Different axis scales are used. For sequences see Table 6-S5.

Table 6-S5: Absorption maxima in the 300-500 nm region for **Y**- and **DY**-modified ONs and the corresponding duplexes with complementary DNA or RNA.[a]

		λ_{max} [$\Delta\lambda_{max}$]/nm		
ON	Sequence	SSP	+cDNA	+cRNA
Y1	5′-GTG A**Y**A TGC	348	353 [+5]	353 [+5]
Y2	3′-CAC **Y**AT ACG	347	354 [+7]	353 [+6]
Y3	3′-CAC TA**Y** ACG	348	353 [+5]	353 [+5]
Y4	3′-CAC **Y**A**Y** ACG	348	354 [+6]	353 [+5]
DY1	5′-GTG A**YD** TGC	348	352 [+4]	352 [+4]
DY2	5′-GTG **DY**A TGC	348	354 [+6]	353 [+5]
DY3	3′-CAC **YD**T ACG	349	352 [+3]	352 [+3]
DY4	3′-CAC T**DY** ACG	348	353 [+5]	352 [+4]

[a] SSP = single-stranded probe. Measurements were performed at 10 °C using a spectrophotometer and quartz optical cells with 1.0 cm path lengths. For buffer conditions, see Table 6-1.

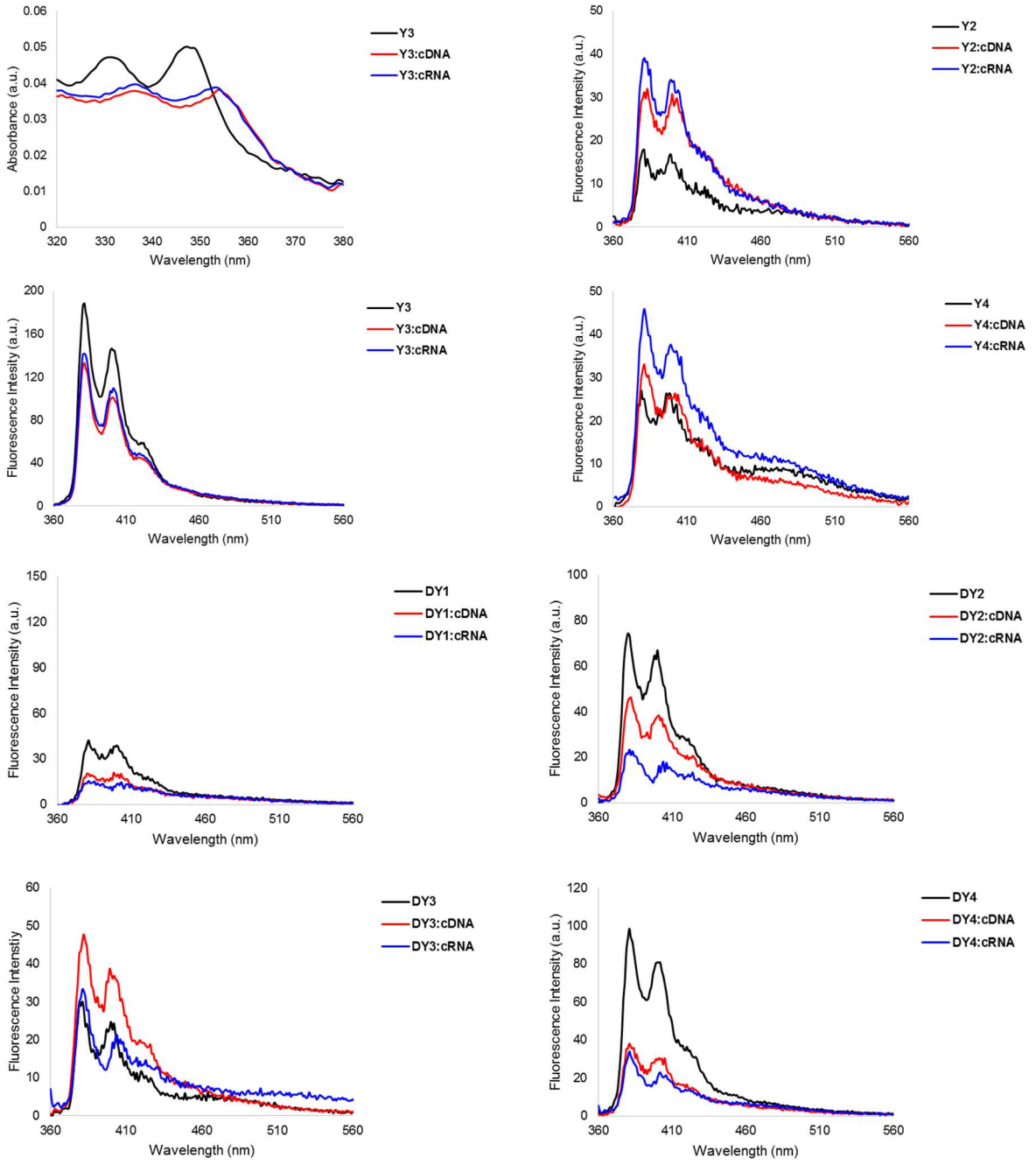


Figure 6-S3: Steady-state fluorescence emission spectra of **Y**- or **DY**-modified ONs and the corresponding duplexes with cDNA/cRNA targets. Spectra were recorded at 5 °C using λ_{ex} = 350 nm. Each strand was used at 1.0 μM concentration in T_m buffer. Note that different axis scales are used. For sequences see Table 6-S5.

Table 6-S6: Change in enthalpy upon duplex formation (ΔH) and reaction enthalpy for recognition of iso-sequential dsDNA target **DNA1**:**DNA2** (ΔH_{rec}).[a]

ON	ZP	Sequence	ΔH [$\Delta\Delta H$] (kJ/mol) upper ON vs cDNA	lower ON vs cDNA	probe duplex	ΔH_{rec} (kJ/mol)
Y1 **Y2**	**-1**	5′-GTG A**Y**A TGC 3′-CAC **Y**AT ACG	-362±11 [-76]	-350±9 [-64]	-296±4 [-10]	-130
Y1 **Y3**	**+1**	5′-GTG A**Y**A TGC 3′-CAC TA**Y** ACG	-362±11 [-76]	-330±10 [-44]	-296±2 [-10]	-110
DY2 **DY3**	**-1**	5′-GTG **DY**A TGC 3′-CAC **YD**T ACG	-296±6 [-10]	-278±13 [+8]	-292±14 [-6]	+4
DY1 **DY4**	**+1**	5′-GTG A**YD** TGC 3′-CAC T**DY** ACG	-343±12 [-57]	-336±19 [-50]	-191±11 [+95]	-202

[a] $\Delta\Delta H$ is measured relative to ΔH for **DNA1**:**DNA2** = -286 kJ/mol. ΔH_{rec} = ΔH (upper ON vs cDNA) + ΔH (lower ON vs cDNA) - ΔH (probe duplex) - ΔH (**DNA1**:**DNA2**). "ZP" and "±" denotes zipper and standard deviation, respectively.

Table 6-S7: Change in entropy at 293K upon duplex formation ($-T^{293}\Delta S$) and reaction entropy for recognition of iso-sequential dsDNA target **DNA1**:**DNA2** ($-T^{293}\Delta S_{rec}$).[a]

Duplex	ZP	Sequence	$-T^{293}\Delta S$ [$\Delta(T^{293}\Delta S)$] (kJ/mol) upper ON vs cDNA	lower ON vs cDNA	probe duplex	$-T^{293}\Delta S_{rec}$ (kJ/mol)
Y1 **Y2**	**-1**	5′-GTG A**Y**A TGC 3′-CAC **Y**AT ACG	301±10 [+61]	298±14 [+58]	242±4 [+2]	+117
Y1 **Y3**	**+1**	5′-GTG A**Y**A TGC 3′-CAC TA**Y** ACG	301±10 [+61]	271±9 [+31]	250±2 [+10]	+82
DY2 **DY3**	**-1**	5′-GTG **DY**A TGC 3′-CAC **YD**T ACG	240±5 [±0]	230±10 [-10]	246±14 [+6]	-16
DY1 **DY4**	**+1**	5′-GTG A**YD** TGC 3′-CAC T**DY** ACG	283±12 [+43]	274±18 [+33]	149±10 [-91]	+168

[a] $\Delta(T^{293}\Delta S)$ is measured relative to $-T^{293}\Delta S$ for **DNA1**:**DNA2** = 240 kJ/mol. $-T^{293}\Delta S_{rec}$ = $T^{293}\Delta S$ (upper ON vs cDNA) + $T^{293}\Delta S$ (lower ON vs cDNA) - $T^{293}\Delta S$ (probe duplex) - $T^{293}\Delta S$ (**DNA1**:**DNA2**).

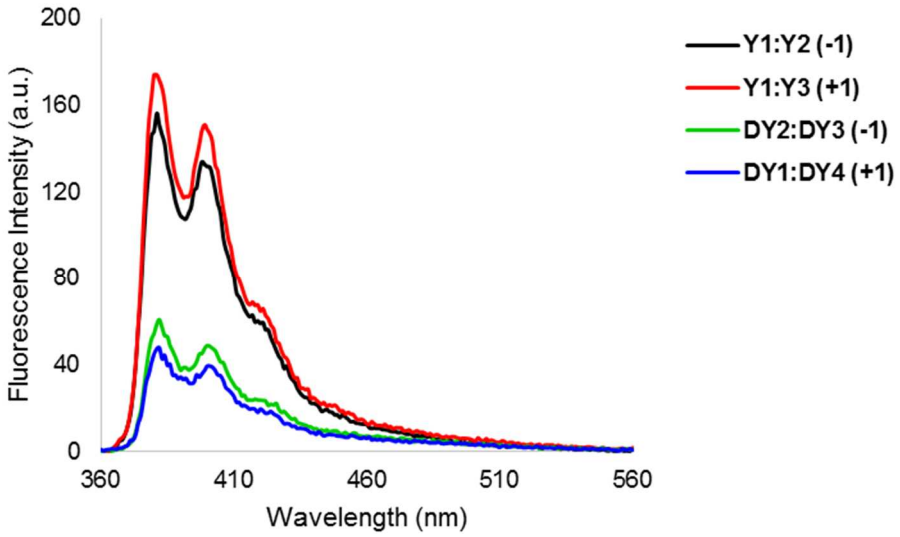


Figure 6-S4: Steady-state fluorescence emission spectra of DNA duplexes with interstrand zipper motifs of **Y** monomers. Flanking **D** monomers quench the fluorescence of **Y** units. For experimental conditions see Figure 6-S3.

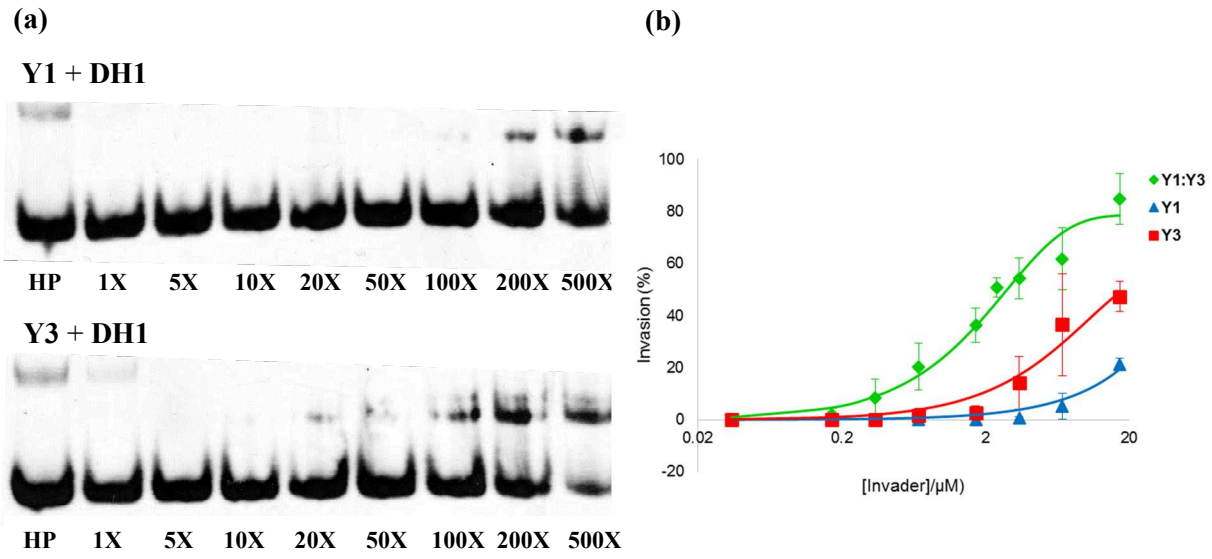


Figure 6-S5: Control experiments – incubation of **DH1** with single-stranded ONs **Y1** or **Y3**. (a) Representative gel electrophoretograms illustrating recognition of **DH1** using 1- to 500-fold excess of **Y1** or **Y3**. (b) Dose-response curves for recognition of **DH1** using **Y1**, **Y3** or **Y1**:**Y3** (average of three independent experiments; error bars denote standard deviation). **Y1**: C_{50} > 17.2 μM; **Y3**: C_{50} > 17.2 μM. Clearly, both strands comprising an Invader probe are necessary for efficient dsDNA-recognition. For experimental conditions, see Figure 6-3.

Table 6-S8: Thermal denaturation temperatures of duplexes between 13-mer probes and cRNA. Also shown is DNA selectivity.

ON	Sequence	+cRNA T_m (ΔT_m) [°C]	$\Delta\Delta T_m$ (DNA-RNA) [°C]
DSX1	5′-GGT AT**D** **X**A**S** AGG C	43.5 (+7.0)	+7.0
DSX2	3′-CCA TA**S** A**XD** TCC G	47.5 (+7.0)	+7.0
DSX3	5′-GGT A**S**A **X**AT **D**GG C	45.0 (+8.5)	+6.5
DSX4	3′-CCA T**D**T A**X**A **S**CC G	49.0 (+8.5)	+5.5
DY5	5′-GGT ATA **YD**T AGG C	41.5 (+5.0)	+6.0
DY6	3′-CCA TAT **DY**A TCC G	45.0 (+4.5)	+5.5
X5	5′-GGT ATA **X**AT AGG C	39.5 (+3.0)	+8.0
X6	3′-CCA TAT A**X**A TCC G	43.5 (+3.0)	+9.0
SD1	5′-GGT ATA **SD**T AGG C	42.5 (+6.0)	-2.5
SD2	3′-CCA TAT **DS**A TCC G	45.5 (+5.0)	-2.5
SD3	5′-GGT AT**D** TA**S** AGG C	41.5 (+5.0)	-2.0
SD4	3′-CCA TA**S** AT**D** TCC G	44.5 (+4.0)	-1.5
SD5	5′-GGT A**S**A TAT **D**GG C	41.5 (+5.0)	-1.0
SD6	3′-CCA T**D**T ATA **S**CC G	45.5 (+5.0)	-2.0

[a]ΔT_m = change in T_m relative to reference duplexes **DNA3:RNA4** ($T_m \equiv 36.5$ °C) or **RNA3**:**DNA4** ($T_m \equiv 40.5$ °C), where **DNA3**: 5′-GGT ATA TAT AGG C, **DNA4**: 3′-CCA TAT ATA TCC G, **RNA3**: 5′-GGU AUA UAU AGG C and **RNA4**: 3′-CCA UAU AUA UCC G; DNA selectivity defined as $\Delta\Delta T_m$ (DNA-RNA) = ΔT_m (vs cDNA) - ΔT_m (vs cRNA).

Table 6-S9: T_m's for duplexes between **X5** or **X6** and mismatched DNA targets, and thermal advantage (*TA*) values for recognition of mismatched dsDNA targets using **X5**:**X6**.[a]

ON	Sequence	T_m [ΔT_m] /°C			*TA*/°C
		Upper strand vs **X6**	Lower strand vs **X5**	dsDNA duplex	
DNA3 **DNA4**	5′-GGT ATA TAT AGG C 3′-CCA TAT ATA TCC G	49.5	48.5	37.5	-22.5
DNA5 **DNA6**	5′-GGT AT*T* TAT AGG C 3′-CCA AT*A* ATA TCC G	38.0 [-11.5]	36.5 [-12.0]	38.5	+2.0
DNA7 **DNA8**	5′-GGT AT*G* TAT AGG C 3′-CCA TA*C* ATA TCC G	43.5 [-6.0]	36.5 [-12.0]	42.0	±0.0
DNA9 **DNA10**	5′-GGT AT*C* TAT AGG C 3′-CCA TA*G* ATA TCC G	36.0 [-13.5]	41.5 [-7.0]	41.0	+1.5
DNA11 **DNA12**	5′-GGT ATA TA*A* AGG C 3′-CCA TAT AT*T* TCC G	36.5 [-13.0]	37.5 [-11.0]	38.0	+2.0
DNA13 **DNA14**	5′-GGT ATA TA*G* AGG C 3′-CCA TAT AT*C* TCC G	41.5 [-8.0]	35.5 [-13.0]	41.5	+2.5
DNA15 **DNA16**	5′-GGT ATA TA*C* AGG C 3′-CCA TAT AT*G* TCC G	36.5 [-13.0]	43.0 [-5.5]	42.5	+1.0

[a] T_m of **X5**:**X6** = 38.0 °C. ΔT_m relative to the fully complementary duplexes **DNA3**:**X6** or **X5**:**DNA4**. $TA_{ONA:ONB}$ = [T_m (ON_A:ON_B) + T_m (**X5**:**X6**)] – [ΔT_m (ON_A:**X6**) + ΔT_m(**X5**:ON_B)]. Thermal advantage (*TA*) is a T_m-based term that is analogous to ΔG^{293}_{rec}, serving as an estimate for dsDNA-recognition potential of a double stranded probe. Probes with strongly negative TA values are likely to be activated for recognition of a given dsDNA target, since the products of the recognition process (i.e., probe-target duplexes) are more thermostable than the reactants (i.e., probe duplexes and target duplexes).

Table 6-S10: T_m's for duplexes between **DSX1** or **DSX2** and mismatched DNA targets, and thermal advantage (*TA*) values for recognition of mismatched dsDNA targets using **DSX1**:**DSX2**.

ON	Sequence	T_m [ΔT_m] /°C[a]			*TA*/°C
		Upper strand vs **DSX2**	Lower strand vs **DSX1**	dsDNA duplex	
DNA3 **DNA4**	5′-GGT ATA TAT AGG C 3′-CCA TAT ATA TCC G	51.5	51.5	37.5	-30.0
DNA5 **DNA6**	5′-GGT AT*T* TAT AGG C 3′-CCA AT*A* ATA TCC G	41.5 [-10.0]	37.5 [-14.0]	38.5	-5.0
DNA7 **DNA8**	5′-GGT AT*G* TAT AGG C 3′-CCA TA*C* ATA TCC G	42.0 [-9.5]	41.5 [-10.0]	42.0	-6.0
DNA9 **DNA10**	5′-GGT AT*C* TAT AGG C 3′-CCA TA*G* ATA TCC G	35.5 [-16.0]	39.5 [-12.0]	41.0	+1.5
DNA11 **DNA12**	5′-GGT ATA TA*A* AGG C 3′-CCA TAT AT*T* TCC G	39.5 [-12.0]	39.5 [-12.0]	38.0	-5.5
DNA13 **DNA14**	5′-GGT ATA TA*G* AGG C 3′-CCA TAT AT*C* TCC G	42.0 [-9.5]	35.0 [-16.5]	41.5	±0.0
DNA15 **DNA16**	5′-GGT ATA TA*C* AGG C 3′-CCA TAT AT*G* TCC G	43.5 [-8.0]	39.5 [-12.0]	42.5	-5.0

[a] T_m of **DSX1**:**DSX2** = 35.5°C. ΔT_m relative to fully complementary duplexes **DNA3**:**DSX2** or **DSX1**:**DNA4**. $TA_{ONA:ONB}$ = [T_m (ON_A:ON_B) + T_m (**DSX1**:**DSX2**)] – [ΔT_m (ON_A:**DSX2**) + ΔT_m(**DSX1**:ON_B)]. For additional discussion of *TA*, see Table 6-S9.

Table 6-S11: T_m's for duplexes between **DSX3** or **DSX4** and mismatched DNA targets, and thermal advantage (*TA*) values for recognition of mismatched dsDNA targets using **DSX3**:**DSX4**.

ON	Sequence	T_m [ΔT_m] /°C[a] Upper strand vs **DSX4**	Lower strand vs **DSX3**	dsDNA duplex	*TA*/°C
DNA3 **DNA4**	5′-GGT ATA TAT AGG C 3′-CCA TAT ATA TCC G	51.5	52.5	37.5	-32.0
DNA5 **DNA6**	5′-GGT AT*T* TAT AGG C 3′-CCA AT*A* ATA TCC G	39.5 [-12.0]	38.0 [-14.5]	38.5	-4.5
DNA7 **DNA8**	5′-GGT AT*G* TAT AGG C 3′-CCA TA*C* ATA TCC G	46.5 [-5.0]	39.0 [-13.5]	42.0	-9.0
DNA9 **DNA10**	5′-GGT AT*C* TAT AGG C 3′-CCA TA*G* ATA TCC G	37.5 [-14.0]	43.5 [-9.0]	41.0	-5.5
DNA11 **DNA12**	5′-GGT ATA TA*A* AGG C 3′-CCA TAT AT*T* TCC G	38.5 [-13.0]	39.0 [-13.5]	38.0	-5.0
DNA13 **DNA14**	5′-GGT ATA TA*G* AGG C 3′-CCA TAT AT*C* TCC G	44.5 [-7.0]	37.5 [-15.0]	41.5	-6.0
DNA15 **DNA16**	5′-GGT ATA TA*C* AGG C 3′-CCA TAT AT*G* TCC G	39.5 [-12.0]	46.0 [-6.5]	42.5	-8.5

[a] T_m **DSX3**:**DSX4** = 34.5 °C. ΔT_m relative to fully complementary duplexes **DNA3**:**DSX4** or **DSX3**:**DNA4**. $TA_{ONA:ONB}$ = [T_m (ON_A:ON_B) + T_m (**DSX3**:**DSX4**)] – [ΔT_m (ON_A:**DSX4**) + ΔT_m(**DSX3**:ON_B)]. For additional discussion of *TA*, see Table 6-S9.

Additional discussion regarding the binding specificity of X5:X6, DSX1:DSX2 and DSX3:DSX4.

As expected, **X5:X6**, **DSX1:DSX2** and **DSX3:DSX4** are strongly energetically activated for recognition of the corresponding isosequential target **DNA3**:**DNA4**, as evidenced by the highly negative *TA* values (-22.5, -30.0 and -32.0 °C, respectively, Tables 6-S9, 6-S10, and 6-S11 – see footnote Table 6-S9 for definition of *TA*). In contrast, *TA* values greater than -6.0 °C are generally observed for recognition of mismatched DNA targets suggesting that recognition of mismatched targets is unlikely. The two exceptions hereto concerns recognition of mismatched

targets **DNA7**:**DNA8** and **DNA15**:**DNA16** using **DSX3**:**DSX4** (*TA* = -9.0 °C and -8.5 °C, respectively, Table 6-S11), which is in excellent agreement with the observed partial recognition of **DH10** and **DH14** when using **DSX3**:**DSX4** at large excess (Figure 6-5c).

Acknowledgements

This study was supported by Award Number GM088697 from the National Institute of General Medical Sciences, National Institutes of Health. Additionally, the project upon which this publication is based was partially supported by Grant No. 12658-I from the University of Idaho Student Grant Program. Such support does not constitute endorsement of SGP of the views expressed in this publication. We thank Dr. Alex Blumenfeld (Dept. Chemistry, Univ. Idaho) and Dr. Lee Deobald (EBI Murdock Mass Spectrometry Center, Univ. Idaho) for assistance with NMR and mass spectrometric analysis, and Prof. Carolyn Bohach (Food Science, Univ. Idaho) for access to gel documentation stations.

6.6 References

(1) (a) Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. *J. Mol. BioSyst*. **2006**, 2, 551 − 560. (b) Bell, N. M.; Micklefield, J. *ChemBioChem* **2009**, 10, 2691 – 2703.

(2) Mukherjee, A.; Vasquez, K. M. *Biochimie* **2011**, 93, 1197 − 1208.

(3) Bergstrom, D. E. *Current protocols of Nucleic Acid Chemistry*, **2001***,* Wiley, New York.

(4) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol*. **2003**, 13, 284 – 299.

(5) Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem*. **2003**, 21, 6101 – 6114.

(6) Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. J. *Mol. BioSyst*. **2006**, 2, 551 – 560.

(7) Bogdanove, A. J.; Voytas, D. F. *Science* **2011**, 333, 1843 – 1846.

(8) Gaj, T.; Gersbach, C. A.; Barbas III, C. F. *Trends Biotechnol*. **2013**, 31, 397 – 405.

(9) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, 36, 5123 – 5138.

(10) Nielsen, P. E., Egholm, M., Berg, R. H.; Buchardt, O. *Science* **1991**, 254, 1497 – 1500.

(b) Bentin, T.; Nielsen, P. E. *Biochemistry* **1996**, 35, 8863 – 8869. (c) Kaihatsu, K.; Braasch, D. A.; Cansizoglu, A.; Corey, D. R. *Biochemistry* **2002**, 41, 11118 – 11125. (d) Kaihatsu, K.; Janowski, B. A.; Corey, D. R. *Chem. Biol*. **2004**, 11, 749 – 758.

(11) For other groove-binding strategies see e.g., (a) Tse, W. C.; Boger, D. L. *Chem. Biol*. **2004**, 11, 1607 – 1617. (b) Hamilton, P. L.; Arya, D. P. *Nat. Prod. Rep*. **2012**, *29*, 134 – 143.

(12) (a) Kaihatsu, K.; Shah, R. H.; Zhao, X; Corey, D. R. *Biochemistry* **2003**, 42, 13996 – 14003. (b) Bentin, T.; Larsen, H. J.; Nielsen, P. E. *Biochemistry* **2003**, 42, 13987 – 13995.

(13) (a) Rapireddy, S.; Bahal, R.; Ly, D. H. *Biochemistry* **2011**, 50, 3913 – 3918.

(b) Bahal, R.; Sahu, B.; Rapireddy, S.; Lee, C-M.; Ly, D. H. *ChemBioChem* **2012**, 13, 56 – 60.

(14) Kutyavin, I.V., Rhinehart, R.L., Lukhtanov, E.A., Gorn, V.V., Meyer, R.B.Jr., and Gamper, H.B. *Biochemistry* **1996**, 35, 11170 – 11176.

(15) Lohse, J.; Dahl, O; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 11804 – 11808.

(16) Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corradini, R.; Sforza, S.; Komiyama, M. *Nucleic Acids Res.* (2008) 36, 1464 – 1471.

(17) Demidov, V. V.; Protozanova, E.; Izvolsky, K. I.; Price, C.; Nielsen, P. E.; Frank-Kamenetskii, M. D. *Proc*. *Natl*. *Acad*. *Sci*. *U.S.A.* **2002**, 99, 5953 – 5958.

(18) Hrdlicka, P. J.; Kumar, T. S.; Wengel, J. *Chem. Commun*. **2005**, 4279 − 4281.

(19) Crothers, D. M. *Biopolymers* **1968**, 6, 575 – 584.

(20) Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2010**, 8, 2028 − 2036.

(21) Sau, S. P.; Madsen, A. S.; Podbevsek, P.; Andersen, N. K.; Kumar, T. S.; Andersen, S.; Rathje, R. L.; Anderson, B. A.; Guenther, D. C.; Karmakar, S.; Kumar, P.; Plavec, J.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, 78, 9560 – 9570.

(22) Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2014**, 12, 7758 – 7773.

(23) Didion, B. A.; Karmakar, S.; Guenther, D. C.; Sau, S. P.; Verstegen, J. P.; Hrdlicka, P. J. *ChemBioChem* **2013**, 4, 3447 – 3454.

(24) Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem*. **2011**, 76, 7119 − 7131.

(25) Kumar, T. S.; Madsen, A. S.; Østergaard, M. E.; Sau, S. P.; Wengel, J.; Hrdlicka, P. *J. Org. Chem.* **2009**, 74, 1070 − 1081.

(26) Andersen, N. K.; Anderson, B. A.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, 78, 12690 – 12702.

(27) Rajeev, K. G.; Prakash, T. P.; Manoharan, M. *Org. Lett.* **2003**, 5, 3005 – 3008.

(28) Anderson, B. A.; Onley, J. J.; Hrdlicka, P. J. *Manuscript in Preparation*. Refer to Chapter 3.

(29) Reese, C. B.; Varaprasad, C. V. N. S. *J. Chem. Soc., Perkin Trans. 1* **1994**, 189 – 195.

(30) Kumar, R. K.; Davis, D. R. *J. Org. Chem.* **1995**, 60, 7726 – 7727.

(31) Nakamura, M.; Fukunaga, Y.; Sasa, K.; Ohtoshi, Y.; Kanaori, K.; Hayashi, H.; Nakano, H.; Yamana, K. *Nucleic Acids Res.* **2005**, 33, 5887 − 5895.

(32) Azhikina, T.; Veselovskaya, S.; Myasnikov, V.; Potapov, V.; Ermolayeva, O.; & Sverdlov, E.; *Proc. Natl. Acad. Sci. U.S.A.* **1993**, 90, 11460 – 11462.

(33) Sproat, B. S.; Lamond, A. I. *Antisense Research and Applications* **1993** CRC Press, Boca Raton, FL.

(34) Testa, S.M; Disney, M.D.; Turner, D.H.; Kierzek, R. *Biochemistry* **1999**, 38, 16655 – 16662.

(35) (a) Dougherty, G.; Pilbrow, J. R. *Int. J. Biochem*. **1984**, 16, 1179 – 1192; (b) Asanuma, H.; Fujii, T.; Kato, T.; Kashida, H. *J. Photochem. Photobiol., C* **2012**, 13, 124 – 135.

(36) Mergny, J.L.; Lacroix, L. *Oligonucleotides* **2003**, 13, 515 – 537.

(37) Guenther, D. C.; Anderson, G. H.; Anderson, B. A.; Karmakar, S.; Hrdlicka, P. J. *Manuscript in Preparation.*

(38) Dioubankova, N. N.; Malakhov, A. D.; Stetsenko, D. A.; Gait, M. J.; Volynsky, P. E.; Efremov, R. G.; Korshun, V. A. *ChemBioChem* **2003**, *4*, 841 – 847.

(39) Brown, A. M. *Computer Methods and Programs in Biomedicine* **2001**, 65, 181 – 200.

CHAPTER 7: Summary and Conclusions

Ten years of work utilizing Invader probes for recognition of double-stranded DNA has demonstrated that a variety of intercalator-modified nucleotides can readily invade mixed-sequence dsDNA with high efficiency and specificity under physiological saline conditions in cell-free hairpin assays. Synthesis of more readily accessible Invader monomers, i.e. 2′-*O*-(pyren-1-yl)methylribonucleotides and N2′-intercalator-functionalized 2′-*N*-methyl-2′-aminouridine, has enabled extensive structure-property relationship studies to be performed.

We initially hypothesized that the synthesis of N2′-pyrene-functionalized 2′-amino-α-L-LNA adenine monomers would be less labor intensive than the thymine counterparts, providing more easy access to building blocks for further characterization of this approach. This, in fact, was not the case, as the synthesis of these monomers proved very challenging. Moreover, Invader probes modified with the N2′-pyrene-functionalized 2′-amino-α-L-LNA adenine result in less efficient recognition of model dsDNA targets than Invaders based on the corresponding thymine monomers.

The discovery of N2′-intercalator-functionalized 2′-*N*-methyl-2′-amino-DNA as functional and structural mimics of the original 2'-*N*-pyrene-functionalized-2'-amino-α-L-LNA, allowed us to conduct structure-property relationship studies with the aim of optimizing Invader probes for more efficient dsDNA recognition.

Analysis of the influence of intercalator size using N2′-pyrene-/perylene-/coronone-functionalized 2′-*N*-methyl-2′-aminouridines revealed that attachment of larger intercalators results in higher invasion efficiency, with Invaders based on N2′-(perylen-3-yl)methyl-2′-*N*-

methyl-2′-amino-2′-deoxyuridines resulting in the most efficient recognition of dsDNA model targets in this series.

Invader probes that are based on a phosphorothioate sugar-phosphate backbone and which are modified with 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-2′-deoxyuridine or 2′-*O*-(pyren-1-yl)methyluridine monomers recognized model dsDNA targets less efficiently than corresponding Invaders based on regular phosphordiester backbones. However, phosphorothioate-based Invaders strands are much more stable to degradation by 3′-exonucleases than regular Invaders, suggesting that the former exhibit promising properties for DNA targeting applications in cellular environments.

Invaders based on 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine were compared to Invaders based on 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-2′-deoxyuridines and 2′-*O*-(pyren-1-yl)methyluridine in order to study the influence of electronegativity of the 2′-sugar atom on dsDNA recognition. Although pseudorotational analysis of 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine intermediates indicated a greater preference for *South*-type conformations than 2′-*O*-(pyren-1-yl)methyluridine, ONs modified with the former display lower cDNA affinity than ONs modified with either 2′-*O*-(pyren-1-yl)methyluridine or 2′-*N*–(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-2′-deoxyuridines. Accordingly, Invader probes modified with 2′-*S*-(pyren-1-yl)methyl-2′-thiouridine showed very low dsDNA recognition potential as judged by ΔG_{rec}^{293} values and did not result in recognition of model dsDNA targets.

To increase the thermodynamic driving force for Invader-mediated dsDNA recognition even further, pseudo-complementary Invaders (pcInvaders) were developed. In the first design, incorporation of 2-aminoadenosine monomers opposite of 2′-*N*-(pyren-1-yl)methyl-2′-amino-

2′-deoxy-2′-*N*-methyl-2-thiouridine monomers afforded pcInvader probes, which were less activated than regular Invaders based on 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-2′-deoxyuridines, since the close proximity of two destabilizing structural motifs (i.e., intercalation-mediated violation of the nearest-neighbor principle and pseudocomplementary base pairs) are not fully compatible with each other. In the second design, separation of the two destabilizing structural motifs (i.e. pseudocomplementary base pairs and +1 interstrand zippers of 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-2′-deoxyuridines) resulted in much more efficient dsDNA recognition indicating that incorporation of regular pseudocomplementary base pairs into Invader probes is therefore a very promising strategy to improve the dsDNA recognition efficiency.

The studies reported herein have identified many promising chemistries for Invader-based recognition of dsDNA. Our results suggest that recognition of dsDNA can be extended to mixed-sequence regions. Future work should include: i) synthesis and characterization of all four canonical nucleobases of next-generation N2′-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-DNAs to determine any inherent limitations of this approach, ii) synthesis and characterization of next-generation Invader probes with longer sequences containing multiple hotspots for high affinity and specificity targeting, specifically 2′-*N*-(pyren-1-yl)methyl-2′-*N*-methyl-2′-amino-2′-deoxyuridines, 2′-*N*-(perylen-3-yl)methyl-2′-*N*-methyl-2′-amino-2′-deoxyuridines and pseudocomplementary Invaders, and (iii) use of next-generation Invaders to regulate gene expression in molecular biological contexts, e.g., plasmids and cells.

Appendix A: Copyright Permission

RightsLink®

Home | Account Info

Title: The triple helix: 50 years later, the outcome:

Author: Maria Duca, Pierre Vekhoff, Kahina Oussedik, Ludovic Halby, Paola B. Arimondo

Publication: Nucleic Acids Research

Publisher: Oxford University Press

Date: 09/01/2008

Logged in as:

Brooke Anderson
University of Idaho

Account #:
3000890261

LOGOUT

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Oxford University Press ("Oxford University Press"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

Get the printable license.

License Number	3575100395502
License date	Feb 23, 2015
Licensed content publisher	Oxford University Press
Licensed content publication	Nucleic Acids Research
Licensed content title	The triple helix: 50 years later, the outcome:
Licensed content author	Maria Duca, Pierre Vekhoff, Kahina Oussedik, Ludovic Halby, Paola B. Arimondo
Licensed content date	09/01/2008
Volume number	36
Issue number	16
Type of Use	Thesis/Dissertation
Requestor type	Academic/Educational institute
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	2
Will you be translating?	No

Author of this OUP article	No
Order reference number	3
Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015
Estimated size(pages)	360
Publisher VAT ID	GB 125 5067 30
Total	0.0 USD

Title: High Thermal Stability of 5′-5′-Linked Alternate Hoogsteen Triplexes at Physiological pH

Author: Vyacheslav V. Filichev,Mads C. Nielsen,Niels Bomholt,Carsten H. Jessen,Erik B. Pedersen

Publication: Angewandte Chemie International Edition

Publisher: John Wiley and Sons

Date: Jul 17, 2006

Order Completed

Thank you for your order.

This Agreement between University of Idaho -- Brooke Anderson ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

Get the printable license.

License Number	3575150468704
License date	Feb 24, 2015
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Angewandte Chemie International Edition
Licensed Content Title	High Thermal Stability of 5′-5′-Linked Alternate Hoogsteen Triplexes at Physiological pH
Licensed Content Author	Vyacheslav V. Filichev,Mads C. Nielsen,Niels Bomholt,Carsten H. Jessen,Erik B. Pedersen
Licensed Content Date	Jul 17, 2006
Licensed Content Pages	5
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 1
Will you be translating?	No
Order reference number	61
Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA

Expected completion date	Mar 2015
Expected size (number of pages)	360
Requestor Location	University of Idaho 875 Perimeter Dr, MS 2343 MOSCOW, ID 83844 United States Attn: Brooke A Anderson
Billing Type	Invoice
Billing address	University of Idaho 875 Perimeter Dr, MS 2343 Moscow, ID 83844 United States Attn: Brooke A Anderson
Total	0.00 USD

Title: Programmable DNA-binding small molecules

Publication: Bioorganic & Medicinal Chemistry

Publisher: Elsevier

Date: 15 October 2013

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Elsevier ("Elsevier"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Get the printable license.

License Number	3575151287806
License date	Feb 24, 2015
Licensed content publisher	Elsevier
Licensed content publication	Bioorganic & Medicinal Chemistry
Licensed content title	Programmable DNA-binding small molecules
Licensed content author	None
Licensed content date	15 October 2013
Licensed content volume number	21
Licensed content issue number	20
Number of pages	14
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	64
Original figure numbers	Figure 1 and Figure 2
Title of your thesis/dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA

Expected completion date	Mar 2015
Estimated size (number of pages)	360
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD

Title: Recognition of the DNA minor groove by pyrrole-imidazole polyamides

Author: Peter B Dervan,Benjamin S Edelson

Publication: Current Opinion in Structural Biology

Publisher: Elsevier

Date: June 2003

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Elsevier ("Elsevier"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Get the printable license.

License Number	3575160790443
License date	Feb 24, 2015
Licensed content publisher	Elsevier
Licensed content publication	Current Opinion in Structural Biology
Licensed content title	Recognition of the DNA minor groove by pyrrole-imidazole polyamides
Licensed content author	Peter B Dervan,Benjamin S Edelson
Licensed content date	June 2003
Licensed content volume number	13
Licensed content issue number	3
Number of pages	16
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	68
Original figure numbers	Figure 1a and Figure 1b
Title of your thesis/dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015
Estimated size (number of pages)	360
Elsevier VAT number	GB 494 6272 12

Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD

Title: Peptide nucleic acids as therapeutic agents
Author: Peter E Nielsen
Publication: Current Opinion in Structural Biology
Publisher: Elsevier
Date: June 1999

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Elsevier ("Elsevier"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Get the printable license.

License Number	3575170163208
License date	Feb 24, 2015
Licensed content publisher	Elsevier
Licensed content publication	Current Opinion in Structural Biology
Licensed content title	Peptide nucleic acids as therapeutic agents
Licensed content author	Peter E Nielsen
Licensed content date	June 1999
Licensed content volume number	9
Licensed content issue number	3
Number of pages	5
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	82
Original figure numbers	Figure 1
Title of your thesis/dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015

Estimated size (number of pages)	360
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Oxford University Press ("Oxford University Press"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

Get the printable license.

License Number	3575111039074
License date	Feb 23, 2015
Licensed content publisher	Oxford University Press
Licensed content publication	Nucleic Acids Research
Licensed content title	Efficient pH-independent sequence-specific DNA binding by pseudoisocytosine-containing bis-PNA
Licensed content author	Michael Egholm, Leif Christensen, Kim L. Deuholm, Ole Buchardt, James Coull, Peter E. Nielsen
Licensed content date	January 25, 1995
Volume number	23
Type of Use	Thesis/Dissertation
Requestor type	Academic/Educational institute
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No
Author of this OUP article	No
Order reference number	85
Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015

Estimated size(pages)	360
Publisher VAT ID	GB 125 5067 30
Total	0.00 USD

Title: A Simple Cytosine to G-Clamp Nucleobase Substitution Enables Chiral γ-PNAs to Invade Mixed-Sequence Double-Helical B-form DNA

Author: Venugopal Chenna,Srinivas Rapireddy,Bichismita Sahu,Cristina Ausin,Enrique Pedroso,Danith H. Ly

Publication: ChemBioChem

Publisher: John Wiley and Sons

Date: Sep 24, 2008

Order Completed

Thank you for your order.

This Agreement between University of Idaho -- Brooke Anderson ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

Get the printable license.

License Number	3575170772808
License date	Feb 24, 2015
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	ChemBioChem
Licensed Content Title	A Simple Cytosine to G-Clamp Nucleobase Substitution Enables Chiral γ-PNAs to Invade Mixed-Sequence Double-Helical B-form DNA
Licensed Content Author	Venugopal Chenna,Srinivas Rapireddy,Bichismita Sahu,Cristina Ausin,Enrique Pedroso,Danith H. Ly
Licensed Content Date	Sep 24, 2008
Licensed Content Pages	4
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 1

Will you be translating?	No
Order reference number	104
Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015
Expected size (number of pages)	360
Requestor Location	University of Idaho 875 Perimeter Dr, MS 2343 MOSCOW, ID 83844 United States Attn: Brooke A Anderson
Billing Type	Invoice
Billing address	University of Idaho 875 Perimeter Dr, MS 2343 Moscow, ID 83844 United States Attn: Brooke A Anderson
Total	0.00 USD

Title: Structural isomers of bis-PNA bound to a target in duplex DNA11Edited by I. Tinoco

Publication: Journal of Molecular Biology

Publisher: Elsevier

Date: 16 March 2001

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Elsevier ("Elsevier"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Get the printable license.

License Number	3575181095337
License date	Feb 24, 2015
Licensed content publisher	Elsevier
Licensed content publication	Journal of Molecular Biology
Licensed content title	Structural isomers of bis-PNA bound to a target in duplex DNA11Edited by I. Tinoco
Licensed content author	None
Licensed content date	16 March 2001
Licensed content volume number	307
Licensed content issue number	1
Number of pages	8
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	86
Original figure numbers	Figure 1c and Figure 5
Title of your thesis/dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015

Estimated size (number of pages)	360
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Oxford University Press ("Oxford University Press"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

Get the printable license.

License Number	3575120151960
License date	Feb 23, 2015
Licensed content publisher	Oxford University Press
Licensed content publication	Nucleic Acids Research
Licensed content title	Targeted correction of a thalassemia-associated β-globin mutation induced by pseudo-complementary peptide nucleic acids:
Licensed content author	Pallavi Lonkar, Ki-Hyun Kim, Jean Y. Kuan, Joanna Y. Chin, Faye A. Rogers, Melissa P. Knauert, Ryszard Kole, Peter E. Nielsen, Peter M. Glazer
Licensed content date	06/01/2009
Volume number	37
Issue number	11
Type of Use	Thesis/Dissertation
Requestor type	Academic/Educational institute
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No
Author of this OUP article	No
Order reference number	111

Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015
Estimated size(pages)	360
Publisher VAT ID	GB 125 5067 30
Total	0.00 USD

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Oxford University Press ("Oxford University Press"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

Get the printable license.

License Number	3575120640363
License date	Feb 23, 2015
Licensed content publisher	Oxford University Press
Licensed content publication	Nucleic Acids Research
Licensed content title	Development of bis-locked nucleic acid (bisLNA) oligonucleotides for efficient invasion of supercoiled duplex DNA:
Licensed content author	Pedro M. D. Moreno, Sylvain Geny, Y. Vladimir Pabon, Helen Bergquist, Eman M. Zaghloul, Cristina S. J. Rocha, Iulian I. Oprea, Burcu Bestas, Samir EL Andaloussi, Per T. Jørgensen, Erik B. Pedersen, Karin E. Lundin, Rula Zain, Jesper Wengel, C. I. Edvard Smith
Licensed content date	03/01/2013
Volume number	41
Issue number	5
Type of Use	Thesis/Dissertation
Requestor type	Academic/Educational institute
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No

Author of this OUP article	No
Order reference number	121
Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015
Estimated size(pages)	360
Publisher VAT ID	GB 125 5067 30
Total	0.00 USD

Title: ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering

Author: Thomas Gaj,Charles A. Gersbach,Carlos F. Barbas

Publication: Trends in Biotechnology

Publisher: Elsevier

Date: July 2013

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Elsevier ("Elsevier"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Get the printable license.

License Number	3575200986006
License date	Feb 24, 2015
Licensed content publisher	Elsevier
Licensed content publication	Trends in Biotechnology
Licensed content title	ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering
Licensed content author	Thomas Gaj,Charles A. Gersbach,Carlos F. Barbas
Licensed content date	July 2013
Licensed content volume number	31
Licensed content issue number	7
Number of pages	9
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	121
Original figure numbers	Figure 1a and Figure 1b
Title of your thesis/dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015
Estimated size (number of pages)	360
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD

VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Nature Publishing Group ("Nature Publishing Group"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

Get the printable license.

License Number	3575210436914
License date	Feb 24, 2015
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature
Licensed content title	Biotechnology: Rewriting a genome
Licensed content author	Emmanuelle Charpentier, Jennifer A. Doudna
Licensed content date	Mar 6, 2013
Type of Use	reuse in a dissertation / thesis
Volume number	495
Issue number	7439
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 1
Author of this NPG article	no
Your reference number	144
Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA
Expected completion date	Mar 2015

Estimated size (number of pages)	360
Total	0.00 USD

Title: Easily denaturing nucleic acids derived from intercalating nucleic acids: thermal stability studies, dual duplex invasion and inhibition of transcription start:

Author: Vyacheslav V. Filichev, Birte Vester, Lykke H. Hansen, Erik B. Pedersen

Publication: Nucleic Acids Research

Publisher: Oxford University Press

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Oxford University Press ("Oxford University Press"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

Get the printable license.

License Number	3575121105104
License date	Feb 23, 2015
Licensed content publisher	Oxford University Press
Licensed content publication	Nucleic Acids Research
Licensed content title	Easily denaturing nucleic acids derived from intercalating nucleic acids: thermal stability studies, dual duplex invasion and inhibition of transcription start:
Licensed content author	Vyacheslav V. Filichev, Birte Vester, Lykke H. Hansen, Erik B. Pedersen
Licensed content date	None
Volume number	33
Issue number	22
Type of Use	Thesis/Dissertation
Requestor type	Academic/Educational institute
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	2
Will you be translating?	No
Author of this OUP article	No
Order reference number	155
Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA

Expected completion date	Mar 2015
Estimated size(pages)	360
Publisher VAT ID	GB 125 5067 30
Total	0.0 USD 1.0

Title: Invaders: Recognition of Double-Stranded DNA by Using Duplexes Modified with Interstrand Zippers of 2′-O-(Pyren-1-yl)methyl-ribonucleotides

Author: Bradley A. Didion,Saswata Karmakar,Dale C. Guenther,Sujay P. Sau,John P. Verstegen,Patrick J. Hrdlicka

Publication: ChemBioChem

Publisher: John Wiley and Sons

Date: Aug 23, 2013

Order Completed

Thank you for your order.

This Agreement between University of Idaho -- Brooke Anderson ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

Get the printable license.

License Number	3575180095350
License date	Feb 24, 2015
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	ChemBioChem
Licensed Content Title	Invaders: Recognition of Double-Stranded DNA by Using Duplexes Modified with Interstrand Zippers of 2′-O-(Pyren-1-yl)methyl-ribonucleotides
Licensed Content Author	Bradley A. Didion,Saswata Karmakar,Dale C. Guenther,Sujay P. Sau,John P. Verstegen,Patrick J. Hrdlicka
Licensed Content Date	Aug 23, 2013
Licensed Content Pages	5
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 3
Will you be translating?	No
Order reference number	151
Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA

Expected completion date	Mar 2015
Expected size (number of pages)	360
Requestor Location	University of Idaho 875 Perimeter Dr, MS 2343 MOSCOW, ID 83844 United States Attn: Brooke A Anderson
Billing Type	Invoice
Billing address	University of Idaho 875 Perimeter Dr, MS 2343 Moscow, ID 83844 United States Attn: Brooke A Anderson
Total	0.00 USD

Order Completed

Thank you very much for your order.

This is a License Agreement between University of Idaho -- Brooke Anderson ("You") and Oxford University Press ("Oxford University Press"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

Get the printable license.

License Number	3575121452464
License date	Feb 23, 2015
Licensed content publisher	Oxford University Press
Licensed content publication	Nucleic Acids Research
Licensed content title	Intercalating nucleic acids containing insertions of 1-O-(1-pyrenylmethyl)glycerol: stabilisation of dsDNA and discrimination of DNA over RNA:
Licensed content author	Ulf B. Christensen, Erik B. Pedersen
Licensed content date	11/15/2002
Volume number	30
Issue number	22
Type of Use	Thesis/Dissertation
Requestor type	Academic/Educational institute
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No
Author of this OUP article	No
Order reference number	156
Title of your thesis / dissertation	Synthesis and Characterization of Energetically Activated Duplexes for Sequence-Unrestricted Recognition of Double-Stranded DNA

Expected completion date	Mar 2015
Estimated size(pages)	360
Publisher VAT ID	GB 125 5067 30
Total	0.00 USD