

Exploiting the Energy of Intercalation for Sequence-Specific 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

Dale C. Guenther

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

Committee Members: Eric B. Brauns, Ph.D.; James J. Nagler, Ph.D.;

Deborah Stenkamp, Ph.D.

Department Administrator: Ray von Wandruszka, Ph.D.

December 2015

**Authorization to Submit Dissertation**

This dissertation of Dale C. Guenther, submitted for the degree of Doctorate of Philosophy with Major in Chemistry and titled "Exploiting the Energy of Intercalation for Sequence-Specific 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: \_\_\_\_\_ Date: \_\_\_\_\_  
Patrick J. Hrdlicka, Ph.D.

Committee Members: \_\_\_\_\_ Date: \_\_\_\_\_  
Eric B. Brauns, Ph.D.

\_\_\_\_\_ Date: \_\_\_\_\_  
James J. Nagler, Ph.D.

\_\_\_\_\_ Date: \_\_\_\_\_  
Deborah Stenkamp, Ph.D.

Department Administrator: \_\_\_\_\_ Date: \_\_\_\_\_  
Ray von Wandruszka, Ph.D.

## Abstract

There has been great interest in developing probes capable of recognizing specific regions of double-stranded DNA (dsDNA) due to the vast potential of applications in the fields of molecular biology, genomic diagnostics, and DNA nanotechnology. Early efforts focused on targeting the accessible grooves of the DNA helix by means of triplex forming oligonucleotides (TFOs) or pyrrole/imidazole polyamides. However, such attempts have been limited by sequence restrictions and/or requirement for non-physiological conditions. A conceptually more elegant, but also more challenging, approach is to target the buried base pairs of the duplex and utilize the intrinsic Watson-Crick binding rules of nucleic acids. Single-stranded peptide nucleic acids (PNA) have a neutral backbone and bind to complementary DNA with very high affinity, facilitating duplex invasion, albeit at low ionic strengths.

We have introduced Invader probes as an alternative approach for recognition of dsDNA. A mini-review of this approach is presented in CHAPTER 1. These probes are modified double-stranded oligonucleotides that are energetically activated for dsDNA recognition due to the low stability of the probe duplex and high affinity duplexes of each individual probe strand towards complementary dsDNA. The driving force for dsDNA recognition is the result of 2'-intercalator-functionalized nucleotides positioned in +1 interstrand zipper arrangements, termed 'energetic hotspots'. Major efforts in the laboratory have been focused on optimizing the chemistry of the monomers, while my projects primarily focused on the optimization of the probe architecture and applications of Invader probes. Initially, I set out to determine the influence of the number, location, and distance between the energetic hotspots. Optimal probe designs were then used for sequence-

specific recognition of chromosomal DNA in a non-denaturing fluorescence in situ hybridization assay (CHAPTER 2). I then introduce bulged Invader probes, which feature an alkyl bulge in the double-stranded probe, which induces fraying and localized perturbation, resulting in faster and more efficient recognition of dsDNA (CHAPTER 3 and 4). Invader probes capable of photoactivated interstrand cross-linking are presented in APPENDIX A, which results in a covalent bond between the probe and DNA strands, resulting in more persistent binding. Finally, I evaluated Invader probes for inhibition of in vitro transcription (APPENDIX B).

## Vita

### EDUCATION

**University of Idaho**, Moscow, ID  
*Doctor of Philosophy*, Chemistry; Emphasis: Nucleic Acid Chemistry Fall 2015  
*Bachelor of Science*, Chemistry (Professional Option) with minor  
 in Mathematics May 2009

### RESEARCH EXPERIENCE

**University of Idaho, Department of Chemistry**, Moscow, ID  
*Graduate Research Assistant* 2010-present  
*Advisor*: Professor Patrick J. Hrdlicka  
*Undergraduate Research Assistant* 2007-2009

### TEACHING EXPERIENCE

**University of Idaho, Department of Chemistry**, Moscow, ID  
*Teaching assistant*: Chem 278/374 - Organic Chemistry I/II  
 laboratory courses 2008, 2014-present

### HONORS/AWARDS

Dr. Alan M. Gewirtz Memorial Scholarship (Oligonucleotide  
 Therapeutics Society) 2014  
 Larry L. Nash Award 2013  
 Brian and Gayle Hill Undergraduate Research Fellowship 2008-2009  
 Idaho INBRE Undergraduate Fellowship 2008  
 Student Grant Program - awarded for 'Development of Synthetic  
 Methodologies for Functionalized Nucleosides' 2008

### PUBLICATIONS

Guenther, D. C.; Karmakar, Saswata; Hrdlicka, P. J. Bulged Invader probes: Activated duplexes for mixed-sequence dsDNA recognition with improved thermodynamic and kinetic profiles. *Chem. Comm.*, **2015**, *51*, 15051-15054.

Guenther, D. C.; Anderson, G. H.; Karmakar, S.; Anderson, B. A.; Didion, B. A.; Guo, W.; Verstegen, J. P.; Hrdlicka, P. J. Invader probes: Harnessing the energy of intercalation to facilitate recognition of chromosomal DNA for diagnostic applications. *Chem. Sci.*, **2015**, *6*, 5006-5015.

Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J. Recognition of double-stranded DNA using energetically activated duplexes with interstrand zippers of 1-, 2- or 4-pyrenyl-functionalized O2'-alkylated RNA monomers. *Org. Biomol. Chem.*, **2014**, *12*, 7758-7773.

Guenther, D. C.; Kumar, P.; Anderson, B. A.; Hrdlicka, P. J. C5-Amino acid functionalized LNA: Positively poised for antisense applications. *Chem. Commun.*, **2014**, *50*, 9007-9009.

Kaura, M.; Guenther, D. C.; Hrdlicka, P. J. Carbohydrate-functionalized Locked Nucleic Acids: Oligonucleotides with extraordinary binding affinity, target specificity, and enzymatic stability. *Org. Lett.*, **2014**, *16*, 3308-3311.

Kumar, P.; Østergaard, M. E.; Baral, B.; Anderson, B. A.; Guenther, D. C.; Kaura, M.; Raible, D. J.; Sharma, P. K.; Hrdlicka, P. J. Synthesis and biophysical properties of C5-functionalized LNA (Locked Nucleic Acid). *J. Org. Chem.*, **2014**, *79*, 5047-5061.

Kumar, P.; Baral, B.; Anderson, B. A.; Guenther, D. C.; Østergaard, M.E.; Sharma, P. K.; Hrdlicka, P. J. C5-Alkynyl-functionalized alpha-L-LNA: Synthesis, thermal denaturation experiments and enzymatic stability. *J. Org. Chem.*, **2014**, *79*, 5062-5073.

Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. Recognition of mixed-sequence DNA duplexes: Design guidelines for Invaders based on 2'-O-(pyren-1-yl)methyl-RNA monomers. *J. Org. Chem.*, **2013**, *78*, 12040-12048.

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. 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.

Didion, B. A.\*; Karmakar, S.\*; Guenther, D. C.\*; Sau, S. P.; Versteegen, J. P.; Hrdlicka, P. J. Invaders: Recognition of double-stranded DNA by using duplexes modified with interstrand zippers of 2'-O-(pyren-1-yl)methyl-ribonucleotides. *ChemBioChem*, **2013**, *14*, 1534-1538. \*Joint first co-authors.

Denn, B.; Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. Sandwich assay for mixed-sequence recognition of double-stranded DNA: Invader-based detection of targets specific to foodborne pathogens. *Chem. Commun.*, **2013**, *49*, 9851-9853.

Østergaard, M. E.; Kumar, P.; Baral, B.; Guenther, D. C.; Anderson, B. A.; Ytreberg, F. M.; Deobald, L.; Paszczyński, A. J.; Sharma, P. K.; Hrdlicka, P. J. C5-Functionalized DNA, LNA, and alpha-L-LNA: Positional control of polarity sensitive fluorophores leads to improved SNP-typing. *Chem. Eur. J.*, **2011**, *17*, 3157-3165.

Østergaard, M. E.; Guenther, D. C.; Kumar, P.; Baral, B.; Deobald, L.; Paszczyński, A. J.; Sharma, P. K.; Hrdlicka, P. J. Pyrene-functionalized triazole-linked 2'-deoxyuridines-probes for discrimination of single nucleotide polymorphisms (SNPs). *Chem. Comm.*, **2010**, *46*, 4929-4931.

Østergaard, M. E.; Kumar, P.; Baral, B.; Raible, D. J.; Kumar, T. S.; Anderson, B. A.; Guenther, D. C.; Deobald, L.; Paszczyński, A. J.; Sharma, P. K.; Hrdlicka, P. J. C5-Functionalized LNA: Unparalleled hybridization properties and enzymatic stability. *ChemBioChem*, **2009**, *10*, 2740-2743.

Rigoli, J. W.; Østergaard, M. E.; Canady, K. M.; Guenther, D. C.; Hrdlicka, P. J. Selective deacylation of peracylated ribonucleosides. *Tetrahedron Lett.*, **2009**, *50*, 1751-1753.

## **Acknowledgements**

I sincerely appreciate Dr. Patrick Hrdlicka for helping me realize my potential as a researcher and scientist. His unwavering support and encouragement has persisted since I first started in his lab as an undergraduate and led to my decision to join his group as a graduate student. He has pushed me to dream big and pursue my goals, goals that I didn't even know I had, and for that I am very grateful for.

I appreciate all my former and current lab mates for helping me with my projects by providing starting material and stimulating discussions. I would especially like to thank Michael Oestergaard, Saswata Karmakar, and Brooke Anderson for making the lab a friendly and collaborative working environment.

I thank my committee members, Eric Brauns, James Nagler, and Deborah Stenkamp for participating in my proposal and dissertation defense.

**Dedication**

*To my friends and family for believing I could do it.*



## Table of Contents

<b>Authorization to Submit Dissertation</b> .....	ii
<b>Abstract</b> .....	iii
<b>Vita</b> .....	v
<b>Acknowledgements</b> .....	vii
<b>Dedication</b> .....	viii
<b>Table of Contents</b> .....	ix
<b>List of Figures</b> .....	xiii
<b>List of Tables</b> .....	xxxii
<b>List of Abbreviations</b> .....	xxxix
<b>CHAPTER 1: A survey of strategies to improve the efficiency of Invader probes for recognition of mixed-sequence double-stranded DNA</b> .....	1
1.1 Introduction to DNA targeting probes .....	1
1.2 Introduction to the Invader probe approach .....	7
1.3 Optimization of 2'-intercalator-functionalized nucleotide monomers for use in Invader probes .....	10
1.4 Optimization of Invader probe architectures .....	21
1.5 Applications of Invader probes .....	35
1.6 Conclusions .....	37
1.7 References .....	39
<b>CHAPTER 2: Invader probes: Harnessing the energy of intercalation to facilitate recognition of chromosomal DNA for diagnostic applications</b> .....	43
2.1 Introduction .....	44

2.2 Results and discussion .....	48
2.3 Conclusions .....	66
2.4 Acknowledgements .....	67
2.5 Supplementary data .....	67
2.6 References .....	95
<b>CHAPTER 3: Bulged Invader probes: activated duplexes for mixed-sequence dsDNA recognition with improved thermodynamic and kinetic profiles .....</b>	<b>99</b>
3.1 Introduction .....	100
3.2 Results and discussion .....	103
3.3 Conclusions .....	113
3.4 Acknowledgements .....	113
3.5 Supplementary data .....	113
3.6 References and notes .....	123
<b>CHAPTER 4: Optimization and application of bulged Invader probes for detection of mixed-sequence dsDNA .....</b>	<b>126</b>
4.1 Introduction .....	127
4.2 Results and discussion .....	130
4.3 Conclusions .....	142
4.4 Supplementary data .....	143
4.5 References .....	164
<b>CHAPTER 5: Summary and conclusions .....</b>	<b>166</b>
<b>APPENDIX A: Psoralen-modified Invader probes for sequence-specific recognition and photoactivated interstrand cross-linking of dsDNA .....</b>	<b>169</b>

A.1 Introduction .....	169
A.2 Results and discussion .....	171
A.3 Conclusions .....	180
A.4 Supporting information .....	181
A.5 References .....	188
<b>APPENDIX B: Inhibition of in vitro transcription by Invader probes .....</b>	<b>190</b>
B.1 Introduction .....	190
B.2 Results and discussion .....	190
B.3 Conclusions .....	205
B.4 Supporting information .....	205
B.5 References .....	214
<b>APPENDIX C: Auxiliary projects .....</b>	<b>215</b>
C.1 Recognition of double-stranded DNA using energetically activated duplexes with interstrand zippers of 1-, 2- or 4-pyrenyl-functionalized O2'- alkylated RNA monomers .....	215
C.2 Invaders: Recognition of double-stranded DNA by using duplexes modified with interstrand zippers of 2'-O-(pyren-1-yl)methyl-ribonucleotides .....	217
C.3 Identification and characterization of second-generation Invader Locked Nucleic Acids (LNAs) for mixed-sequence recognition of double-stranded DNA .....	218
C.4 Recognition of mixed-sequence DNA duplexes: Design guidelines for	

Invaders based on 2'- <i>O</i> -(pyren-1-yl)methyl-RNA monomers .....	220
C.5 Sandwich assay for mixed-sequence recognition of double-stranded DNA: Invader-based detection of targets specific to foodborne pathogens .....	222
<b>APPENDIX D: Copyright Permissions .....</b>	<b>223</b>

## List of Figures

<b>Figure 1.1-1.</b> Current approaches for targeting dsDNA: triplex forming oligonucleotides, pyrrole/imidazole polyamides, peptide nucleic acids, pseudocomplementary DNA/PNA and protein-mediated recognition. ....	3
<b>Figure 1.1-2.</b> Illustration of binding to DNA grooves by TFO and polyamides .....	4
<b>Figure 1.1-3.</b> Structure of PNA and $\gamma$ -PNA.....	6
<b>Figure 1.1-4.</b> Pseudocomplementary base pairing. ....	7
<b>Figure 1.2-1.</b> The principle of the Invader approach for recognition of dsDNA.....	7
<b>Figure 1.2-2.</b> Depictions of the intercalative binding mode of the 2'-intercalator-functionalized nucleotides. Color code: intercalators (blue), sugar phosphate backbone (red), and nucleobases (green). Adapted from Ref. 35 with permission from the Royal Society of Chemistry. ....	9
<b>Figure 1.3-1.</b> Regions of 2'-intercalator-functionalized nucleotides optimized for use as building blocks for Invader probes .....	10
<b>Figure 1.3-2.</b> Different monomers explored to facilitate efficient intercalation upon hybridization. ....	11
<b>Figure 1.3-3.</b> Recognition of the stem regions in DNA hairpins by Invader probes. ....	16

**Figure 1.3-4.** Various 2'-intercalator-functionalized nucleotide monomers used as components of Invader probes. Phe = phenanthren-2/3-yl, 1-Py = pyren-1-yl, 2-Py = pyren-2-yl, 4-Py = pyren-4-yl, Tri = triphenylen-2-yl, Per = perylen-3-yl, Cor = coronen-1-yl.....17

**Figure 1.3-5.** Invader probe building blocks based on 2'-alkylated RNA with substituted pyrene as intercalators. Note: Py-Br, Py-Me, and Py-Et were used as inseparable mixtures of isomers.....19

**Figure 1.4-1.** Sequence dependence of the energetic hotspot. Bolded and underlined letters represent **Y1(Py)**-modified nucleotides with the corresponding nucleobase. ....22

**Figure 1.4-2.** Illustration of assay used to determine dissociation kinetics of recognition complex. ....26

**Figure 1.4-3.** Invader probes containing non-nucleosidic bulges for increased recognition of dsDNA.....28

**Figure 1.4-4.** Representative examples of Invader probes with intra- or interstrand arrangements of bulges of C<sub>9</sub> monomers. ....30

**Figure 1.4-5.** The principle of pseudocomplementary Invader probes.....32

**Figure 1.5-1.** Sandwich-based assay for detection and quantification of dsDNA (left) and non-denaturing fluorescence in situ hybridization (FISH) assay for detection of chromosomal DNA (right).....36

**Figure 2.1-1.** (a) Illustration of the Invader approach for recognition of dsDNA. (b) Structure of Invader monomers discussed herein. ....48

**Figure 2.2-2.** (a) Available Gibbs free energy at 293K ( $\Delta G_{rec}^{293}$ ) for Invader-mediated recognition of isosequential dsDNA targets, and (b-d) change in Gibbs free energy upon formation of **X**- or **Y**-modified DNA duplexes. The  $\Delta G_{293}$  for the dsDNA reference is shown as a dotted line at -57 kJ/mol. See Tables 2.5-6 and 2.5-7 for tabulated data. .... 53

**Figure 2.2-3.** (a) Illustration of the electrophoretic mobility shift assay (EMSA) used to evaluate dsDNA recognition. (b) dsDNA recognition by Invader probes. DIG-labeled **DH1** (34.4 nM) was incubated with 6.88  $\mu$ M of a pre-annealed Invader probe in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) for 17 h at room temperature. For representative electrophoretograms see Figure 2.5-4. ....55

**Figure 2.2-4.** Dose-response curves for recognition of **DH1** using a) **M1:M2**, b) **M9:M10**, or c) **M19:M20**. For experimental conditions, see Figure 2.2-3. .... 56

**Figure 2.2-5.** Kinetic profile of **DH1** recognition using 200-fold molar excess of a) **M1:M2**, b) **M9:M10**, or c) **M19:M20**. For incubation conditions, see Figure 2.2-3.

Aliquots were taken at specific time points, flash frozen in liquid N<sub>2</sub>, and stored at -76 °C until analysis. ....57

**Figure 2.2-6.** Temperature dependence of dsDNA recognition kinetics using a) **X9:X10** or b) **Y9:Y10** having two consecutive energetic hotspots. Experiments were performed as described in Figure 2.2-3 with the exception of different incubation temperatures. ....58

**Figure 2.2-7.** Dissociation kinetics of recognition complexes between DNA hairpins and Invader probes. (a) Illustration of competition assay. (b) Representative gel electrophoretograms for dissociation reactions. Recognition complexes were formed (incubation of 34.4 nM **DH1** with 200-fold molar excess of **M1:M2** or **M9:M10** for 17 h at room temperature), followed by addition of 2,000-fold molar excess of 5'-GGTATATATAGGC:3'-CCATATATATCCG. Incubation conditions are as described in Figure 2.2-3. Reactions were quenched at specific times points as described in Figure 2.2-5. .... 60

**Figure 2.2-8.** Discrimination of non-complementary DNA hairpins (**DH2-DH7**) using 1000-, 500- and 50-fold excess of **M1:M2**, **M9:M10** and **M19:M20**, respectively. For experimental conditions, see Figure 2.2-3.  $T_m$ 's of **DH1-DH7** are between 58.5–63.5 °C (Table 2.2-15). .... 61

**Figure 2.2-9.** Images from fluorescence in situ hybridization using Y-chromosome specific Invader probes under non-denaturing conditions. Invader probe **INV4** was



added to nuclei from male bovine kidney cells in interphase (upper panel) or metaphase (middle panel), or to nuclei from female bovine fibroblast cells (lower panel). Images viewed using Cy3 (left column) or DAPI (middle column) filter settings; overlays are shown in the right column. Incubation: 3 h at 38.5 °C in 10 mM Tris-Cl, pH 8.0 and 1 mM EDTA. Cells were visualized at 400x magnification using a Zeiss AxioSkop 40 fluorescence microscope and images captured using a Zeiss AxioCam MRc5 camera. For additional images using probes **INV1-INV4**, see Figure

2.5-10..... 65

**Figure 2.2-10.** Illustration of hypothetical recognition mechanism..... 66

**Figure 2.5-1.** Representative thermal denaturation curves of Invader probes, corresponding duplexes between individual probe strands and cDNA, and unmodified reference duplex. For experimental conditions, see Table 2.2-1..... 76

**Figure 2.5-2.** (a) Change in reaction enthalpy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{\text{rec}}$ ), and (b-d) change in enthalpy ( $\Delta H$ ) upon formation of **X**- or **Y**-modified DNA duplexes. The  $\Delta H$  for the dsDNA reference is shown as a dotted line at -382 kJ/mol. See Table 2.2-1 for experimental conditions. See Tables 2.5-6 and 2.5-7 for tabulated data.....84

**Figure 2.5-3.** (a) Change in reaction entropy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $-T^{293}\Delta S_{\text{rec}}$ ), and (b-d) change in entropy ( $-T^{293}\Delta S$ ) upon formation of **X**- or **Y**-modified DNA duplexes. The  $-T^{293}\Delta S$  for

the dsDNA reference is shown as a dotted line at 325 kJ/mol. See Table 2.2-1 for experimental conditions. See Tables 2.5-10 and 2.5-11 for tabulated data. ....87

**Figure 2.5-4.** Representative electrophoretograms of data shown in Figure 2.2-1 of main text: dsDNA recognition by 200-fold molar excess of Invader probes modified either monomer **X** (top) or **Y** (bottom). For sequences and experimental conditions see Table 2.2-1 and Figure 2.2-3, respectively. .... 88

**Figure 2.5-5.** (a) Recognition of **DH1** or **DH8** using Invader probes **X1:X2–X19:X20** and (b) representative gel images. See Figure 2.2-1 for experimental conditions. ....88

**Figure 2.5-6.** Plots of  $-\ln(1-C)$  versus time, where  $C$  is the ratio of recognition complex with respect to the total hairpin concentration, used to determine pseudo-first order rate constants for initial phases of **DH1** recognition using a) **M1:M2**, b) **M9:M10**, or c) **M19:M20** at 200-fold molar excess. See Figure 2.2-1 for incubation conditions. The linearity of the plots suggests that the reaction obeys pseudo-first order kinetics. ....90

**Figure 2.5-7.** Recognition of **DH1** using individual probe strands. Single- stranded probes (and double-stranded Invader controls) were used at 1000-, 500-, or 50-fold molar excess for **M1:M2**, **M9:M10**, or **M19:M20**, respectively. For experimental conditions, see Figure 2.2-1. ....90

**Figure 2.5-8.** Gel electrophoretogram of gender-specific PCR amplicons. Zfx and Zfy bands represent X- and Y-chromosome specific amplicons, respectively. Lane “M” contains a molecular weight standard (Lo DNA Marker, Bionexus). PCR amplicons from female (lanes 1 and 3) and male cell lines (lanes 2 and 4). 2% agarose gels were used. ....92

**Figure 2.5-9.** Images from FISH experiments conducted at non-denaturing conditions using nuclei from male bovine kidney cells incubated with the single-stranded Cy3-labeled PNA probe 5'-Cy3-OO-AGCCCTGTGCCCTG. Images viewed using Cy3 (left column) or DAPI (middle column) filter settings. Incubation: 3h at 38.5 °C in 10 mM Tris-Cl, pH 8.0 and 1 mM EDTA. Cells were visualized at 400x magnification using a Zeiss AxioSkop 40 fluorescent microscope and images captured using a Zeiss AxioCam MRc5 camera. “O” denotes a 9-atom ethylene glycol linker (i.e., “O-linker”)...... 93

**Figure 2.5-10.** Additional images from FISH experiments using nuclei from male bovine kidney cells incubated with the following Y-chromosome specific Invaders: **INV1** (top panel; nuclei in metaphase prior to cell division), **INV2** (second panel; nuclei in interphase), **INV2** (third panel; inter- and metaphase nuclei), **INV3** (fourth panel; interphase nuclei) and **INV4** (bottom panel; interphase nuclei). Images viewed using Cy3 (left column) or DAPI (middle column) filter settings; overlay images are shown in the right column. Sequences are shown in Table 2.2-1. For experimental details, see legend of Figure 2.2-9..... 94

**Figure 2.5-11.** Images from FISH experiments involving nuclei from male bovine kidney cells that were treated with proteinase K (upper panel), RNase (middle panel) and DNase (lower panel) prior to incubation with Invader **INV4**. Note the continued presence of Cy3-signals in proteinase K or RNase pre-treated samples, and the absence of Cy3-signals in DNase pre-treated samples. For experimental details, see legend of Figure 2.2-9..... 95

**Figure 3.2-1.** Schematic representation of dsDNA recognition by Invader probes containing non-nucleosidic bulges and the chemical modifications utilized for this approach..... 104

**Figure 3.2-2.** (a) Schematic representation of the electrophoretic mobility shift assay used to evaluate dsDNA recognition of Invader probes. (b) Representative electrophoretograms for recognition of model dsDNA target **DH1** (34.4  $\mu\text{M}$ ) by different Invader probes (6.88  $\mu\text{M}$ ) at 8 °C. (c) Histogram showing the average of three experiments; error bars represent standard deviation. DIG-labeled **DH1** (5'-GGTATATATAGGC-T<sub>10</sub>-GCCTATATATACC-3') was incubated with pre-annealed Invader probe in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) for 17 h. .... 109

**Figure 3.2-3.** Dose-response curves for recognition of dsDNA by Invader probes **ON3:ON2**, **ON3:ON4**, **ON7:ON2**, and **ON1:ON2** at (a) 8 °C or (b) 22 °C. Experimental conditions otherwise as in Figure 3.2-2..... 110

**Figure 3.2-4.** a) Assays used to determine kinetic parameters for dsDNA recognition using representative Invader probes. b) *Left*: Kinetics of recognition complex formation at 22 °C using 100-fold molar excess of Invader probes. *Right*: Competitive dissociation kinetics of recognition complexes between DNA hairpins and Invader probes (for representative gel electrophoretograms, see Figure 3.5-6). 100-fold molar excess of Invader probes (3.44  $\mu\text{M}$ ) was incubated with **DH1** for 24 h, followed by addition of a 1000-fold molar excess of linear competitor dsDNA target (34.4  $\mu\text{M}$  – sequence: 5'-GGTATATATAGGC:3'-CCATATATATCCG).  $T = 22\text{ }^\circ\text{C}$ . ..... 112

**Figure 3.5-1.** Representative thermal denaturation curves of Invader probes, duplexes between individual probe strands and cDNA, and unmodified reference duplexes. For experimental conditions, see Table 3.2-1. .... 118

**Figure 3.5-2.** . The structure of recognition complexes formed upon incubation of different Invaders strands/probes and DNA hairpin **DH1** at 22 °C. Invader probes used at 100-fold excess (3.44  $\mu\text{M}$ ). Experimental conditions are otherwise as described in Figure 3.2-2. The greater mobility of the recognition complex between **ON7:ON2** and **DH1** relative to the complexes involving **ON1:ON2**, **ON3:ON2** and **ON5:ON2**, and the similar mobility relative to the complex involving **ON2**, strongly suggests that a binary complex is formed. Nonetheless, recognition of **DH1** is more efficient with **ON7:ON2** than **ON2**, indicating that **ON7** – despite its low cDNA affinity – plays a role in the recognition process, possibly through weak/transient binding to the binary complex, thus preventing re-formation of **DH1**. ..... 120

**Figure 3.5-3.** Recognition of model dsDNA target **DH1** using 200-fold molar excess (6.88  $\mu\text{M}$ ) of different Invader probes at 8 °C. Image is a composite of two electrophoretograms (lanes 1-3 and lanes 4-7). Experiments were performed in duplicate. Very similar results (not shown) were observed at 22 °C..... 122

**Figure 3.5-4.** Recognition of **DH1** by individual Invader probe strands. Each lane includes **DH1** incubated with 200-fold molar excess (6.88  $\mu\text{M}$ ) of single-stranded probes or conventional Invader probe **ON1:ON2** at 22 °C. Experiments were performed in duplicate. Single-stranded probes, **ON1** and **ON2**, result in ~50% recognition of **DH1**, while **ON1:ON2** results in ~70% recognition, emphasizing the need for both strands for maximal recognition..... 122

**Figure 3.5-5.** Pseudo-first order rate plots for dsDNA recognition by Invader probes at 22 °C. Raw time profiles shown in Figure 3.2-4.....123

**Figure 3.5-6.** Representative electrophoretograms from the competitive dissociation assay shown in Figure 3.2-4. .... 123

**Figure 3.5-7.** Specificity of Invader-mediated dsDNA recognition. Fully base-paired DNA hairpins containing a single base-pair mismatch (red) relative to Invader probes **ON1:ON2** (left) and **ON3:ON2** (right) were used. Base-pairs above electrophoretograms correspond to **B:B'** in: 5'-GGTATBTATAGGC-T<sub>10</sub>-GCCTATABB'ATACC. A 200-fold molar excess of Invader probes (6.88  $\mu\text{M}$ ) were used. Experiments were performed in duplicate at 22 °C (shown) or 37 °C (results not

shown – identical observations). Excellent discrimination of the singly mismatched DNA hairpins was observed with both Invader probes. .... 124

**Figure 4.2-1.** Illustration of bulged Invader probes and modifications used in the present study ..... 130

**Figure 4.2-2.** Change in Gibbs free energy of hybridization at 25 °C for a) individual probe strands and complementary DNA or b) Invader probe duplexes. c) Gibb's free energy for recognition of isosequential dsDNA ( $\Delta G_{\text{rec}} = (\Delta G_{5\text{'-ON:cDNA}} + \Delta G_{3\text{'-ON:cDNA}}) - (\Delta G_{\text{probe duplex}} + \Delta G_{\text{dsDNA}})$ ). Dotted line represents  $\Delta G$  of unmodified DNA = -97 kJ/mol. For tabulated data, see Table 4.4-2..... 137

**Figure 4.2-3.** Recognition of hairpin DNA target by Invader probes monitored via an electrophoretic mobility shift assay. Illustration of assay (upper left), Invader probes in which only one (right panel) or both strands contain bulges (lower left). Measurements were obtained after incubation of 200-fold molar excess of pre-annealed Invader probe (6.88  $\mu\text{M}$ ) with DIG-labeled DNA hairpin (34.4 nM) for 17 h in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM  $\text{MgCl}_2$ , pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride). For tabulated data, see Table 4.4-8..... 140

**Figure 4.2-4.** Characterization of DNA recognition by Invader probes at 37 °C. (a) Dose-dependence of recognition for selected probes and (b) specificity of recognition using 400-fold molar excess Invader probe (13.8  $\mu\text{M}$ ). See Figure 4.2-3 for other

experimental conditions. For full sequence of hairpin DNA and corresponding  $T_m$ 's, see Table 4.4-9. .... 142

**Figure 4.2-5.** Bulged Invader probes used in nd-FISH for sequence-specific detection of chromosomal DNA of interphase or metaphase nuclei spreads. Invader probes (100 nM) with the conventional probe construct (a), or bulged probes where  $\underline{\mathbf{X}} = \mathbf{4}$  (b) or  $\underline{\mathbf{X}} = \mathbf{9}$  (c) were incubated for 3 h with isolated nuclei at 37.5 °C in Tris buffer ([Tris-HCl] = 10 mM, [EDTA] = 1 mM, [KCl] = 50 mM, pH 8.3). Targeting probe sequence: 5'-Cy3-AGCCCUGTGCCCTG:3'-TCGGGACACXGGGAC-Cy3. d) Triply mismatched probe (5'-Cy3-AGCGCUGAGGCCTG:3'-TCGCGACTC4CGGAC-Cy3) was used; italicized bases depict mismatches with respect to the target sequence.  $T_m$ 's and  $TA$  of Invader probes shown in Table 4.4-11 ... 144

**Figure 4.4-1.** Representative thermal denaturation curves of Invader probes, duplexes consisting of individual probe strands and cDNA, and unmodified reference cDNA duplex. For experimental conditions, see Table 4.2-1. .... 151

**Figure 4.4-2.** Fluorescence-based assay used to monitor strand exchange process between Invader probes and isosequential dsDNA targets. Py = pyrene. Data shown in Table 4.4-7. .... 158

**Figure 4.4-3.** Steady-state fluorescence emission spectra of Invader probes in the absence or presence of an isosequential dsDNA target. Measurements were performed in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2,



10% sucrose, 1.44 mM spermine tetrahydrochloride) at 22 °C and recorded after incubating pre-annealed Invader probe (1.25  $\mu$ M) with or without addition of pre-annealed linear dsDNA (1.25  $\mu$ M) for ~24 h..... 160

**Figure 4.4-4.** Representative electrophoretograms for targeting DNA hairpin (**DH1**) by Invader probes for data shown in Figure 4.2-2..... 162

**Figure A.1-1.** Structure of psoralen and the adducts formed with thymine after irradiation with UVA light..... 163

**Figure A.1-2.** Illustration of psoralen-modified Invader probes and ICL formation after UV irradiation..... 173

**Figure A.2-1.** Recognition of hairpin DNA (34.4 nM) by psoralen-modified Invader probes. Screen of probes at 100-fold molar excess (a) and concentration dependence (b-d) of recognition after incubation with DNA hairpin for 17 h at 22 °C. dsDNA recognition is determined from densitometry measurements as the ratio of recognition complex/DNA hairpin for a given lane. Representative electrophoretograms shown in Figure A.4-2. .... 177

**Figure A.2-2.** Illustration of DNA hairpin recognition using psoralen-modified Invader probe, followed by irradiation to form stable thymine adducts that covalently cross-links the probe and DNA hairpin (a). Representative electrophoretograms (b) of denaturing PAGE after incubation of psoralen-modified Invader probes (200-fold molar excess) with DNA hairpin followed by UV-irradiation (3 h) and denaturing

PAGE (7M urea). For incubation conditions see A.2-1. Note: the above electrophoretogram was compiled from two images.....179

**Figure A.2-3.** Cross-linking between individual probe strands annealed with single-stranded 33 mer DNA (D1 or D2), demonstrated using denaturing PAGE and visualized by Sybr gold stain (a). Psoralen-modified Invader probes incubated with DIG-labeled 33 bp dsDNA (b and c) for 17 h, followed by irradiation with 365 nm light (+) or exposed to ambient light (-) for 1 h. Electrophoretograms were captured on x-ray film after processing for chemiluminescent signal with anti-DIG alkaline phosphatase. For incubation conditions see Figure A.2-1. \*Invader probe annealed with the dsDNA (90 °C for 10 min and cooled to RT over 15 min) followed by irradiation with UV-light.

D1: 5'-  
AAGCTGCACAGGTATATATAGGCCGCATATGCA,  
D2: 3'-  
TTCGACGTGTCCATATATATCCGGCGTATACGT. Underlined region corresponds to embedded target sequence.....181

**Figure A.4-1.** Representative thermal denaturation curves of individual probe strands with complementary DNA (left) and Invader probe duplexes (right)..... 186

**Figure A.4-2.** Representative electrophoretograms for data shown in Figure A.2-1..... 190

**Figure A.4-3.** Thermal denaturation curve of **ON4**:cDNA with increasing lengths of irradiation with UV-light..... 190

**Figure B.2-1.** Architecture of plasmid 14971 (~4.4 kb), which contains a human c-Myc insert (~1.4 kb) in the pBluescript SK vector at the EcoRI restriction site (a). Red bars indicate Invader probe binding sites associated with either the T7 template or T3 template. (b) Possible mechanisms of Invader probe-mediated inhibition of in vitro transcription..... 194

**Figure B.2-2.** Inhibition of in vitro transcription by Invader probes. Linearized plasmid DNA (0.5  $\mu$ g, SfoI digest of plasmid 14971) incubated with Invader probes (1000-fold molar excess) for 17 h at 37 °C in transcription buffer (contains Mg<sup>2+</sup> and DTT, concentration and other components not published- Ambion) followed by addition of rNTPs (0.25 mM each) and T7 RNAP (15 U) and incubation for 1 h at 37 °C. Separation of fragments was achieved using 5% denaturing-PAGE and visualized with Sybr Gold stain. Quantification of signal was determined using ImageJ software.. 198

**Figure B.2-3.** Linearized plasmid DNA (SfoI and SpeI double-digest of plasmid 14971; 0.5  $\mu$ g) incubated with Invader probes (1000-fold molar excess, 17 h, 37 °C, in transcription buffer) followed by addition of rNTPs (0.25 mM each), T3 and T7 RNAP (30 or 15 U, respectively; lanes 1 - 7) and incubation at 37 °C for 1 h. Invader probes expected to inhibit T7 (lane 2 – 5) or T3 (lane 6 and 7) catalyzed RNA synthesis. Lane 8: ssRNA ladder. Separation achieved using 5% denaturing-PAGE and visualized with Sybr gold stain. Note that lack of band ~90 nt in lane 1 is likely due to loading differences, as other experiments showed this band. These experiments were repeated in a HEPES buffer with similar results..... 200

**Figure B.2-4.** Controls for observed inhibition of in vitro transcription of template plasmid DNA (SfoI and SpeI digest of plasmid 14971). Addition of T3 RNAP (lane 1), T7 RNAP (lane 2) or both T3 and T7 RNAP (lane 3 – 9). Invader probe or corresponding unmodified dsDNA (500-fold for lane 4 - 7; 100-fold or 10-fold for lane 8 and 9, respectively) was added to template DNA (0.25  $\mu$ g) and incubated for 17 h at 37 °C in transcription buffer followed by addition of rNTPs (0.42 mM) and RNAP (7.5 U) and incubation for 1 h at 37 °C. Lane 10: ssRNA ladder. Separation of fragments was achieved using 5% denaturing-PAGE and visualized with Sybr stain. ... 202

**Figure B.2-5.** Evaluation of sequence-specificity of Invader probes for inhibition of in vitro transcription. (a) Electrophoretogram of RNA transcripts after incubation of plasmid 14971 (SfoI digest; 0.25  $\mu$ g) with increasing concentrations of Invader probe (10-, 100-, or 1000-fold molar excess) incubated at 37 °C for 17 h in transcription buffer, followed by addition of rNTP (0.25 mM each) and T7 RNAP (7.5 U) and incubation for 1 h at 37 °C. Separation achieved using 5% denaturing PAGE and visualized with Sybr stain. (b) Colocalization of increasing concentrations (10-, 100-, or 1000-fold molar excess) of Inv1 and Inv1MM with plasmid 14791 (SfoI digest) incubated for 17 h at 37 °C in CutSmart buffer (1.43x: 71.5 mM potassium acetate, 28.6 mM tris-acetate, 14.3 mM magnesium acetate, 143  $\mu$ g/mL BSA, pH 7.9). Plasmid DNA visualized using Sybr stain (upper panel) after electrophoresis (4% nd-PAGE), which was subsequently processed for chemiluminescent detection of DIG-labeled probe (middle panel). The lower panel is a merger of these two images, which represents colocalization of Invader probe and plasmid DNA. ....204

**Figure B.2-6.** Illustration of TH plasmid construct. Synthetic 500 bp insert with 7 identical target sites was cloned into the pUC57 vector. Digestion with HpaI or SpeI yields templates appropriate for in vitro transcription under T7 control with 1 or 6 binding sites, respectively. Electrophoretogram shows complete digestion of plasmid DNA with HpaI or SpeI after separation using 1% agarose gel and ethidium bromide staining. Lane 1: 1 kb ladder. A control plasmid (TH control plasmid) was also cloned, where the target sites were excluded but otherwise the sequence is identical (325 bp insert). .....205

**Figure B.2-7.** Evaluation of Invader probes for inhibition of in vitro transcription using HpaI digests of TH plasmid as DNA template, which contains one copy of the Invader target sequence under T7 RNAP control, and comparison to TH control template, which lacks the target sequence (a). Electrophoretogram of RNA transcript after incubation with template (0.25  $\mu$ g) with Invader probe (1000-fold molar excess) incubated at 37 °C for 17 h in transcription buffer, followed by addition of rNTP (0.25 mM each) and T7 RNAP (3.75 U) and incubation for 4 min at 37 °C. (b) Colocalization of Inv6 and Inv7 (1000-fold molar excess) with supercoiled (lane 2 and 3) or linearized TH plasmid (7 binding sites; lanes 4 and 5) or TH control plasmid (no binding sites; lanes 6 and 7) incubated for 17 h at 37 °C after annealing for 2 min at 95 °C in CutSmart buffer (1.43x: 71.5 mM potassium acetate, 28.6 mM tris-acetate, 14.3 mM magnesium acetate, 143  $\mu$ g/mL BSA, pH 7.9). ..... 206

**Figure B.4-1.** (a) Digestion of plasmid 14971 (~4.3 kb) by SfoI (lane 3) and EcoRI (lane 4), visualized by ethidium bromide after electrophoresis with 1% agarose gel (70 V, 1 h). Lane 1: 1 kb ladder. Lane 5: 100 bp ladder. SfoI (1 cut - 852 bp) and EcoRI (2 cuts - ~1400 bp). (b) Double-digest of plasmid 14971 with SfoI and SpeI (lane 3). Lane 4: 1 kb ladder and lane 5: 100 bp ladder..... 212

**Figure B.4-2.** Evaluation of inhibition of in vitro transcription after incubation of Invader probe (1000-fold molar excess) with plasmid DNA (0.5 $\mu$ g, SfoI digest of plasmid 14971) for 17 h at 37 °C in transcription buffer. Prior to addition of rNTPs (0.25 mM each) and T7 RNAP (15 U), excess probes were removed from the reaction mixture using a spin column (isolates species >100 bp). The reaction mixture was then incubated for 1 h at 37 °C, followed by separation of fragments using 5% denaturing PAGE, and visualized with Sybr Gold stain..... 212

**Figure B.4-3.** (a) Colocalization of Inv1 and Inv1MM (1000-fold molar excess) with plasmid 14791 (SfoI digest) incubated for 17 h at specified temperature in CutSmart buffer (1.43x: 71.5 mM potassium acetate, 28.6 mM tris-acetate, 14.3 mM magnesium acetate, 143  $\mu$ g/mL BSA, pH 7.9). (b) Invader probes (1000-fold molar excess) incubated with linearized plasmid 14791 in different buffers (tx: Ambion transcription buffer; HEPES: 71.5 mM HEPES, 143 mM NaCl, 7.15 mM MgCl<sub>2</sub>, 14.3% sucrose, pH 7.2, 2.06 mM spermine). Plasmid DNA visualized using Sybr stain (upper panel) after electrophoresis (4% nd-PAGE), which was subsequently processed for chemiluminescent detection of DIG-labeled probe (middle panel). The

lower panel is a merge of these two images, which represents colocalization of Invader probe and plasmid DNA. For additional data see Figure B.2-3. ....213

**Figure B.4-4.** Expected binding by Invader probes to the promoter regions for T7 and T3 RNAPs.....215

## List of Tables

<b>Table 1.3-1.</b> Properties of Invader probes modified with various 2'-pyrene-functionalized nucleotides with different sugar skeletons and linkers between pyrene and the sugar. ....	14
<b>Table 1.3-2.</b> Properties of Invader probes modified with various 2'-intercalator functionalized nucleotides. ....	18
<b>Table 1.3-3.</b> Properties of Invader probes modified with substituted pyrene building blocks. ....	20
<b>Table 1.4-1.</b> Properties of various 13-mer Invader probe constructs. ....	24
<b>Table 1.4-2.</b> Thermal properties and dsDNA recognition of Bulge Invader probe constructs. ....	31
<b>Table 1.4-3.</b> Properties of pc-Invader probes. ....	34
<b>Table 2.2-1.</b> Thermal denaturation temperatures ( $T_m$ 's) and thermal advantages ( $TA$ 's) of X- and Y-modified DNA duplexes. ....	51
<b>Table 2.2-2.</b> Summary of dsDNA-recognition characteristics of selected Invader probes. ....	58
<b>Table 2.2-3.</b> $T_m$ and $\Delta G_{rec}^{310}$ values of Invader probes used in the nd-FISH study. ....	63



<b>Table 2.5-1.</b> MALDI-MS of ONs modified with 2'- <i>O</i> -(pyren-1-yl)methyl-RNA monomers.....	74
<b>Table 2.5-2.</b> MALDI-MS of ONs modified with 2'- <i>N</i> -(pyren-1-ylmethyl)-2'- <i>N</i> -methylamino-uridine monomers. ....	75
<b>Table 2.5-3.</b> Thermal denaturation temperatures ( $T_m$ 's) and thermal advantage ( $TA$ ) of DNA duplexes modified with 2'- <i>O</i> -(pyren-1-yl)methyluridine monomers. ....	77
<b>Table 2.5-4.</b> Thermal denaturation temperatures ( $T_m$ 's) and thermal advantage ( $TA$ ) of DNA duplexes modified with 2'- <i>N</i> -(pyren-1-ylmethyl)-2'- <i>N</i> -methylaminouridine monomers. ....	78
<b>Table 2.5-5.</b> Thermal denaturation temperatures ( $T_m$ 's) and thermal advantage ( $TA$ ) of DNA duplexes modified with 0-zippers of 2'- <i>O</i> -(pyren-1-yl)methyl-RNA monomers. ...	79
<b>Table 2.5-6.</b> Change in Gibbs free energy ( $\Delta G$ ) at 293K upon formation of X-modified DNA duplexes and change in reaction free energy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{293}$ ). ....	80
<b>Table 2.5-7.</b> Change in Gibbs free energy ( $\Delta G$ ) at 293K upon formation of Y-modified DNA duplexes and change in reaction free energy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{293}$ ). ....	81

**Table 2.5-8.** Change in enthalpy ( $\Delta H$ ) upon formation of **X**-modified duplexes and change in reaction enthalpy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{rec}$ ). .....82

**Table 2.5-9.** Change in enthalpy ( $\Delta H$ ) upon formation of **Y**-modified duplexes and change in reaction enthalpy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{rec}$ ). .....83

**Table 2.5-10.** Change in entropy at 293K ( $-T^{293}\Delta S$ ) upon formation of **X**-modified duplexes and change in reaction entropy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $-T^{293}\Delta S_{rec}$ ). .....85

**Table 2.5-11.** Change in entropy at 293K ( $-T^{293}\Delta S$ ) upon formation of **Y**-modified duplexes and change in reaction entropy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $-T^{293}\Delta S_{rec}$ ). .....86

**Table 2.5-12.** Thermal denaturation temperatures of DNA hairpin targets used in this study. ....89

**Table 2.5-13.** Change in Gibbs free energy ( $\Delta G$ ) at 310K upon formation of **X**-modified DNA duplexes and change in reaction free energy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{310}$ ). .....91

<b>Table 2.5-14.</b> Change in enthalpy ( $\Delta H$ ) upon formation of <b>X</b> -modified duplexes and change in reaction enthalpy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{\text{rec}}$ ). .....	91
<b>Table 2.5-15.</b> Change in entropy at 310K ( $-T^{310}\Delta S$ ) upon formation of <b>Y</b> -modified duplexes and change in reaction entropy upon Invader-mediated recognition of isosequential dsDNA targets ( $-T^{310}\Delta S_{\text{rec}}$ ). .....	92
<b>Table 3.2-1.</b> Thermal denaturation temperatures ( $T_m$ 's) and thermal advantages ( $TA$ 's) for modified DNA duplexes. ....	106
<b>Table 3.2-2.</b> Summary of parameters for dsDNA recognition by representative Invader probes.....	113
<b>Table 3.5-1.</b> MALDI-MS of modified ONs. ....	117
<b>Table 3.5-2.</b> Degree of recognition of DNA hairpin <b>DH1</b> using a 200-fold molar excess of different Invader probes at two different temperatures.....	119
<b>Table 3.5-3.</b> Thermal denaturation temperatures ( $T_m$ 's) of additional Invader probes. .	121
<b>Table 4.2-1.</b> Thermal denaturation temperatures ( $T_m$ 's) of Invader probe duplexes and individual probe strands hybridized with complementary DNA. ....	133
<b>Table 4.2-2.</b> Thermal denaturation temperatures of singly mismatched DNA duplexes. ....	135

**Table 4.4-1.** MALDI-MS of modified ONs ..... 150

**Table 4.4-2.** Change in Gibbs free energy ( $\Delta G$ ) at 298 K for formation of probe duplexes (dsInv), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{298}$ ).... 152

**Table 4.4-3.** Change in Gibbs free energy ( $\Delta G$ ) at 310 K for formation of probe duplexes (dsInv), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{310}$ ).....153

**Table 4.4-4.** Change in enthalpy ( $\Delta H$ ) for formation of probe duplexes (dsInv), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{rec}$ ) .....154

**Table 4.4-5.** Change in entropy at 298 K ( $-T^{298}\Delta S$ ) for formation of probe duplexes (dsInv), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $-T^{298}\Delta S_{rec}$ ) .....155

**Table 4.4-6.** Change in entropy at 310 K ( $-T^{310}\Delta S$ ) for formation of probe duplexes (dsInv), and duplexes between individual probe strands and complementary DNA (5'-

Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $T^{310}\Delta S_{rec}$ ) .....	156
<b>Table 4.4-7.</b> Recognition of linear dsDNA targets by Invader probes .....	159
<b>Table 4.4-8.</b> Degree of DNA recognition using 200-fold molar excess of the different Invader probes. ....	161
<b>Table 4.4-9.</b> Thermal denaturation temperatures of DNA hairpins. ....	163
<b>Table 4.4-10.</b> MALDI-MS of ONs targeting a unique region of the <i>DYZ-1</i> satellite region of the Y-chromosome of <i>bos Taurus</i> . ....	164
<b>Table 4.4-11.</b> Thermal denaturation temperatures of Invader probes that are derivatives of 'Inv7:8' targeting <i>DYZ-1</i> satellite region of the Y-chromosome of <i>bos Taurus</i> . ....	165
<b>Table A.2-1.</b> Thermal hybridization properties of psoralen modified Invader probes. ...	166
<b>Table A.4-1.</b> MALDI-MS of modified ONs .....	186
<b>Table A.4-2.</b> . Change in Gibbs free energy ( $\Delta G$ ) at 298 K for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{298}$ ).....	187

**Table A.4-3.** Change in enthalpy ( $\Delta H$ ) at 298 K for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in enthalpy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{rec}^{298}$ ). .....177

**Table A.4-4.** Change in entropy ( $-T^{298}\Delta S$ ) at 298 K for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in entropy upon Invader-mediated recognition of isosequential dsDNA targets ( $-T^{298}\Delta S_{rec}^{298}$ ). .....178

**Table B.2-1.** Thermal denaturation temperatures ( $T_m$ 's) of Invader probes evaluated for inhibition of in vitro transcription. .... 196

**Table B.4-1.** MALDI-MS of ONs modified with 2'-*O*-(pyren-1-yl)methyl-RNA monomers. .... 214

**List of Abbreviations**

A	Adenin-9-yl
C	Cytosin-1-yl
Cas	CRISPR Associated
cDNA	Complementary DNA
CRISPR	Clustered Regularly Interspaced Short Palendromic Repeat
cRNA	Complementary RNA
Cor	Coronene
DNA	Deoxyribonucleic Acid
dsDNA	Double-stranded DNA
EDTA	N, N, N', N'- ethylenediamine tetraacetic acid
EtBr	Ethidium Bromide
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High Pressure Liquid Chromatography
ICL	Interstrand cross-link
LNA	Locked Nucleic Acid
MALDI	Matrix Assisted Laser Desorption Ionization
MS	Mass Spectrometry
nt	Nucleotide
ON	Oligodeoxyribonucleotide
PAGE	Polyacrylamide Gel Electrophoresis
pc	Pseudocomplementary

Pery	Perylene
PNA	Peptide Nucleic Acid
PO	Phosphodiester
PS	Phosphorothioate
Py	Pyrene
RNA	Ribonucleic Acid
rt	Room Temperature
ss	Single-stranded
SVPDE	Snake Venom Phosphodiesterase Enzyme
T	Thymin-1-y
TA	Thermal Advantage
TBE	Tris/Borate/EDTATBM Tris/Borate/Magnesium
Tm	Thermal Denaturation Temperature
TFO	Triplex Forming Oligonucleotide
U	Uracil-1-yl
UV	Ultra Violet



## **CHAPTER 1: A survey of strategies to improve the efficiency of Invader probes for recognition of mixed-sequence double-stranded DNA**

Dale C. Guenther and Patrick J. Hrdlicka

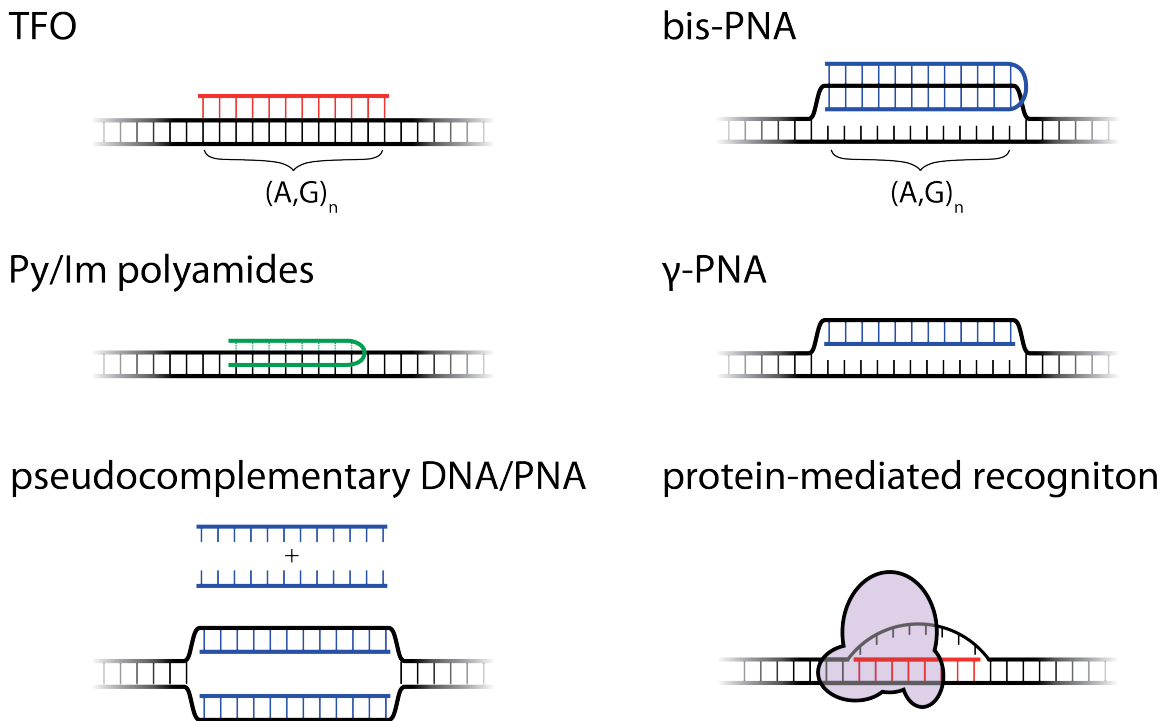
### **1.1 Introduction to DNA targeting probes**

DNA is commonly referred to as the molecule of life, because the information necessary for protein synthesis, cell regulation and function is stored within chromosomal DNA as repeating units of four nucleotides: adenine, thymine, guanine and cytosine. The genetic code for protein synthesis is clearly defined in the form of codons, i.e., trinucleotides that code for a specific amino acid during translation. However, the mechanisms of gene regulation and function of non-coding DNA are less well understood. Development of ligands capable of recognizing double-stranded DNA (dsDNA) has been a long-standing goal, motivated by the ability to detect, query, or manipulate site-specific regions of the genome in order to gain functional and structural information.<sup>1-4</sup>

Recently, a new class of dsDNA targeting agents have been developed that involves RNA-guided nucleases, referred to as clustered regulatory interspaced short palindromic repeats (CRISPR) with CRISPR associated proteins (CRISPR/Cas). These nucleases are present in bacteria and most archaea and are believed to have evolved as a form of acquired DNA-encoded immunity that protects host cells from invasive species.<sup>5</sup> The protein (Cas) can be engineered to elicit double-strand breaks in mammalian cells,<sup>6</sup> and thereby inducing homologous recombination, or designed as a fusion with a fluorescent protein that lacks endonuclease activity, for sequence-specific detection of dsDNA.<sup>7</sup> The specificity of binding is thought to be determined by the guide RNA strand, however the

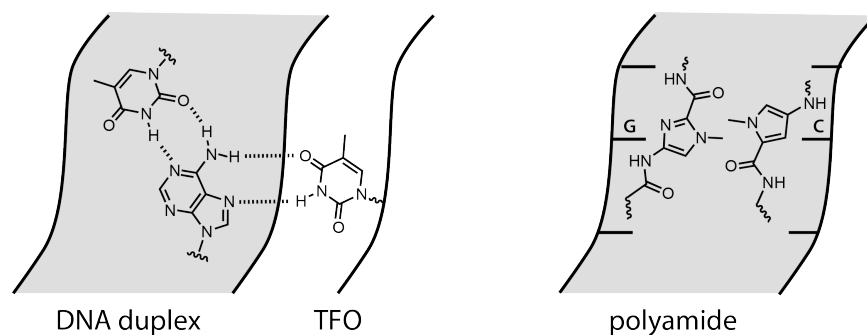
associated protein (Cas) can lead to off-target effects.<sup>6</sup> Additionally, the use of this technology is dependent on endogenous cellular machinery for formation of the RNA and protein components of CRISPR/Cas, which limits potential applications to metabolically competent cellular systems.

Hybridization-based probes, on the other hand, offer the advantage that binding interactions are governed by Watson-Crick base pairing (adenine:thymine/uracil and cytosine:guanine) or other recognition schemes in the major or minor grooves (Figure 1.1-1). They are simple to design and predictably bind to target sequences with high affinity. Early examples of DNA targeting agents bind to the more accessible grooves, and include triplex forming oligonucleotides (TFO)<sup>8,9</sup> and pyrrole/imidazole (Py/Im) polyamides.<sup>10,11</sup> Binding of TFOs occurs through formation of Hoogsteen hydrogen bonds in the major groove (Figure 1.1-2), and necessitates a long polypurine stretch in one of the target strands. Additionally, low pH ensures protonation of cytosine and formation of C<sup>+</sup>GC base triplets. Many chemical modifications have been developed to increase the binding affinity of TFOs at physiological conditions, e.g., by reducing charge repulsion of the strands in the triplex (e.g., 2'-*O*-(2-aminoethyl)-ribonucleotides and peptide nucleic acids, PNAs), entropic penalties (e.g., locked nucleic acids, LNAs), and the requirement for low pH (e.g., 5-methylcytosine).<sup>8,10</sup> It has been proposed that TFO target sites (>15 bases with at least 50% G-content) are overrepresented in the human genome, with ~98% of annotated genes having at least one target site within the promoter/transcribed region.<sup>12</sup> However, these sites may not be accessible and the use of TFOs still faces significant sequence restrictions and recognition of mixed sequence regions is not possible.



**Figure 1.1-1.** Current approaches for targeting dsDNA: triplex forming oligonucleotides, pyrrole/imidazole polyamides, peptide nucleic acids, pseudocomplementary DNA/PNA and protein-mediated recognition.

The development of polyamides for targeting dsDNA was inspired by the natural products, netropsin and distamycin A, which are crescent shaped ligands that fit into the minor groove of dsDNA, and bind via hydrogen bonding interactions to exposed hydrogen bond donors and shape-specific van der Waals interactions.<sup>13</sup> Py/Im polyamides consist of repeating *N*-methylpyrrole and *N*-methylimidazole units, which bind through base-pair specific contacts in the minor groove (Figure 1.1-2). However, they are only capable of recognizing short regions (<8 bp) of dsDNA, which precludes recognition of unique genomic DNA regions.<sup>14</sup>

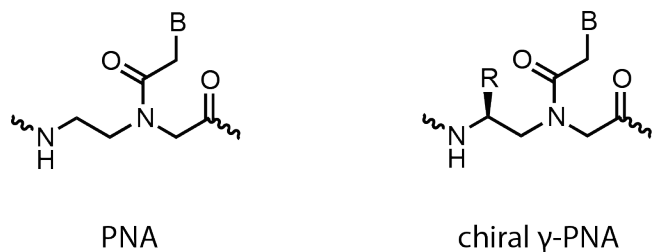


**Figure 1.1-2.** Illustration of binding to DNA grooves by TFO and polyamides.

Peptide nucleic acids (PNAs) are repeating units of canonical nucleobases on an achiral, uncharged *N*-(2-aminoethyl)-glycine amino backbone (Figure 1.1-3). PNA has extraordinarily high affinity towards complementary DNA (cDNA) due to reduced electrostatic repulsion, and can bind through both Watson-Crick and Hoogsteen base pairing. Initial efforts focused on triplex-based strategies of homopyrimidine PNAs. However, rather than forming regular triplexes it was noted that a 2:1 PNA:DNA triplex is formed, in which one strand binds via Hoogsteen base pairing while a second PNA strand invades the target duplex to bind via Watson-Crick base pairing. Two PNA strands can be joined via a linker to reduce the entropic penalty. These so-called bis-PNAs (Figure 1.1-1) result in enhanced strand invasion process.<sup>15</sup> To reduce the sequence limitations of bis-PNA, tail-clamp-PNAs were introduced.<sup>16,17</sup> These probes are closely related to bis-PNA, but have an extended mixed-sequence invasion arm, which allows for a reduction in the length of the polypurine target region, thus increasing the number of potential target sites. Tail-clamp PNA probes have been used to inhibit transcription,<sup>17</sup> however, the inherent sequence limitation of triplex-based approaches still applies.

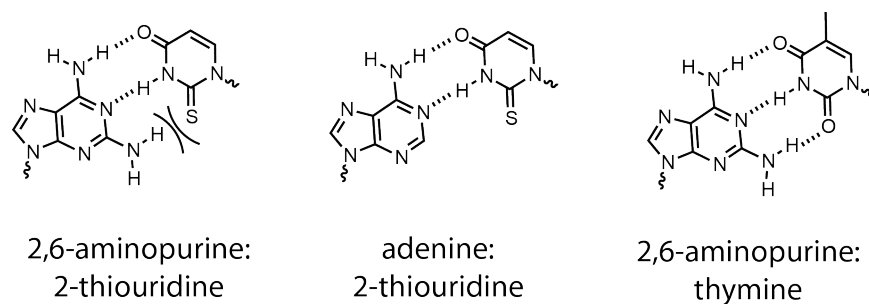
Probes that rely on invasion of Watson-Crick base pairs are conceptually more challenging than probes targeting features presented in the accessible grooves, because the recognition elements are buried within the duplex. However, such probes offer the advantage of targeting mixed sequence regions and utilizing predictable Watson-Crick base pairing rules. In order to stably bind to the buried Watson-Crick face of a dsDNA target region, a probe needs to bind with significantly stronger affinity than the two target strands have towards themselves. Single-stranded PNAs have been shown to enable recognition of mixed-sequence dsDNA regions. However, recognition via strand invasion was only efficient when targeting supercoiled plasmids with inverted repeats within AT-rich regions or when targeting single-stranded region in the open complex of transcriptionally active genes.<sup>18,19</sup>

The use of chiral and suitable conformationally restricted PNA backbone,  $\gamma$ -PNA (Figure 1.1-3) dramatically increases the affinity towards cDNA. This is likely due to preorganization of the  $\gamma$ -PNA strand for hybridization and concomitantly reduced entropic penalty upon DNA binding.<sup>20</sup> A fully modified, single-stranded probe allows for recognition of mixed-sequence dsDNA regions at low ionic strengths.<sup>21,22</sup> Additional chemical modifications have been developed to improve the poor solubility (e.g., incorporation of terminal lysine units and short diethylene glycol),<sup>22</sup> and improve the binding kinetics (e.g., conjugation with polycationic peptides),<sup>23</sup> but non-specific aggregation and poor cellular uptake issues remain that need to be addressed.



**Figure 1.1-3.** Structure of PNA and  $\gamma$ -PNA.

The use of probes capable of binding to both strands of DNA targets is a desirable approach because of the potential for increased stability arising from additional Watson-Crick bonds relative to single-stranded probes that leave one target strand unhybridized (D-loop). One way to realize this is to use double-stranded probes, in which the two probe strands have low affinity towards each other, but high affinity towards complementary DNA targets. The use of pseudocomplementary nucleotides is one such approach to achieve a favorable energetic gradient for dsDNA. The best example of a pseudocomplementary base pair is 2-thiothymine and 2,6-aminopurine (Figure 1.1-4).<sup>24</sup> In the double-stranded probe, these base pairs are destabilized due to steric clashes between the sulfur atom and exocyclic amine. However, when hybridized to cDNA the steric clashes are alleviated, and 2,6-aminopurine forms three hydrogen bonds with thymine, thus resulting in a stable duplex.



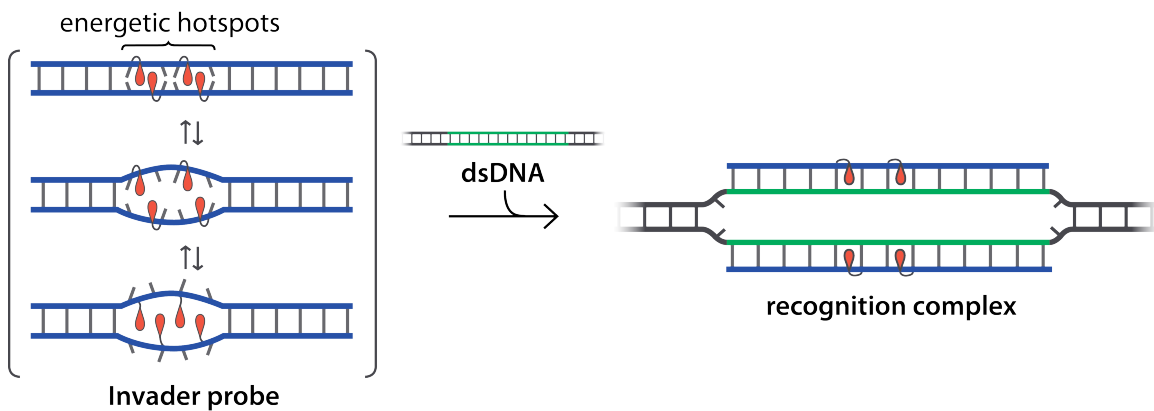
**Figure 1.1-4.** Pseudocomplementary base pairing.

This concept has been extended to PNA (pcPNA), which were shown to recognize mixed-sequence dsDNA regions via a double-duplex strand invasion mechanism<sup>25</sup> and induce homologous recombination in mammalian cells.<sup>26</sup> However, the binding of pcPNA *ex vivo* has been limited to conditions with low ionic strength and AT-rich target sites. Chemical optimization of pcPNA probes by introduction of chiral positive charges to the pcPNA backbone (e.g., *N*-(2-aminoethyl)-D-lysine)) partially alleviates these restrictions.<sup>25,27</sup> Recently, it has been suggested that the activity of regular pcPNA is efficient at physiologically relevant ionic strengths when conducted at ‘molecular crowding’ conditions (i.e. high concentrations of poly(ethylene glycol)),<sup>28</sup> which mimic nuclear environments, and is supported by the observed intracellular activity of pcPNAs.<sup>26</sup>

## 1.2 Introduction to the Invader probe approach

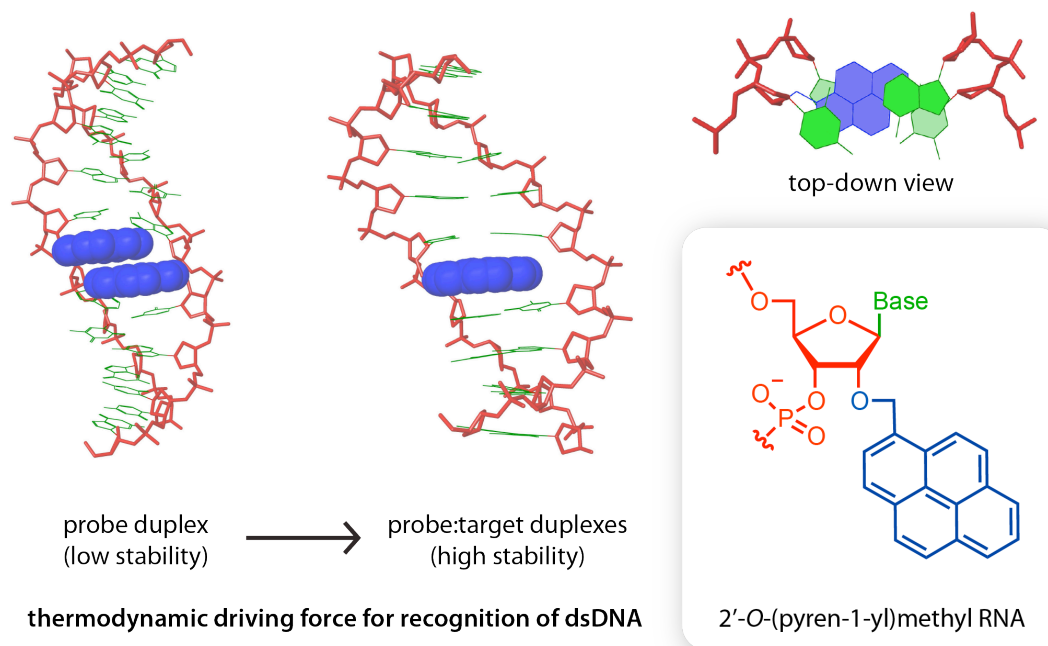
Our laboratory has introduced an alternative approach for recognition of dsDNA, which is based on double-stranded DNA probes that are activated for recognition of dsDNA due to stability differences between the probe duplexes (low stability) and probe:target duplexes (high stability; Figure 1.2-1).<sup>29-31</sup> The novelty of the Invader approach resides in

the probe duplex, in which traditionally stabilizing intercalators are arranged in such a manner that they perturb the duplex, resulting in destabilization.<sup>29,31,32</sup> This energetic hotspot, or region of destabilization, is accomplished by arranging 2'-intercalator-functionalized nucleotides in +1 interstrand zippers, which forces the intercalators to occupy the same region within the duplex (Figure 1.2-2). The observed destabilization of the probe duplex occurs because the 'nearest neighbor exclusion' principle (NNEP) is violated, which states that, at most, intercalators bind every second base pair due to limits in local helix expandability.<sup>33,34</sup> In contrast, the individual probe strands have very high affinity toward cDNA, as the lower density of intercalators no longer violates the NNEP. The difference in stability between the probe and probe:target duplexes generates a favorable energetic gradient for recognition of mixed sequence dsDNA.



**Figure 1.2-1.** The principle of the Invader approach for recognition of dsDNA.



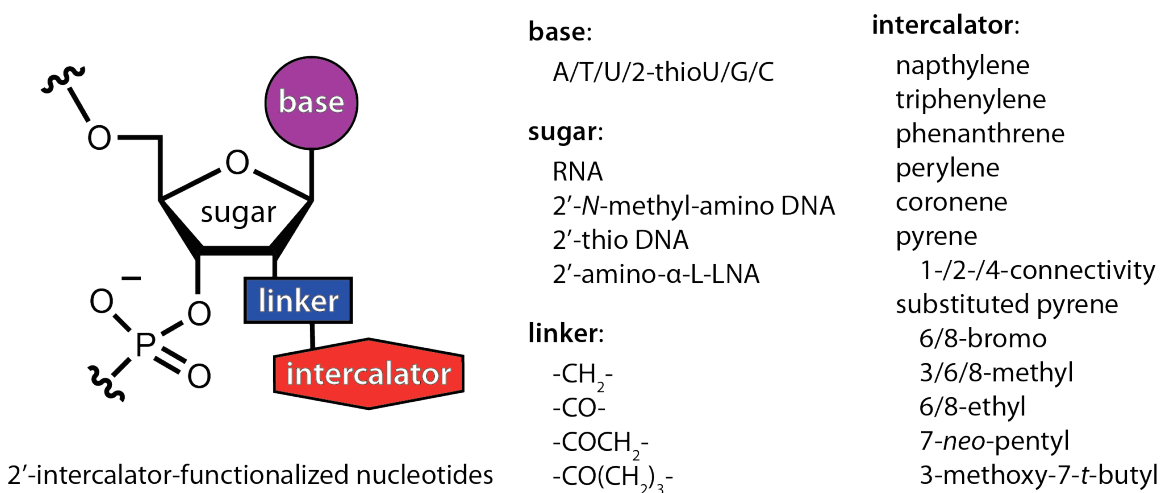


**Figure 1.2-2.** Depictions of the intercalative binding mode of the 2'-intercalator-functionalized nucleotides. Color code: intercalators (blue), sugar phosphate backbone (red), and nucleobases (green). Adapted from Ref. 35 with permission from the Royal Society of Chemistry.

The suggested double-intercalation mechanism for destabilization of Invader probes is supported by NMR studies,<sup>31</sup> as local perturbation of the Watson-Crick base pairs flanking the intercalator-modified monomers is observed. Additionally, double-stranded probes with +1 interstrand zipper arrangements of 2'-intercalator-functionalized nucleotides exhibit unique fluorescence spectra compared to probes with other zipper arrangements.<sup>31</sup> A prominent feature of 2'-pyrene-functionalized nucleotides arranged in this manner is the formation of a pyrene-pyrene excimer emission ( $\lambda_{\text{max}} \sim 495$  nm), which is indicative of a coplanar arrangement of two pyrene moieties with an interplanar distance of  $\sim 3.4$  Å.<sup>36</sup> Neither excimer formation, nor probe duplex destabilization, is observed with other zipper motifs,<sup>31,32</sup> which highlights the importance of this particular motif.

### 1.3 Optimization of 2'-intercalator-functionalized nucleotide monomers as building blocks for Invader probes

We set out to study how the structure of intercalator-modified nucleotides influence the properties Invader probes by systematically varying i) the sugar skeleton of the monomers and ii) the size, connectivity and substitution pattern of the intercalator (Figure 1.3-1).



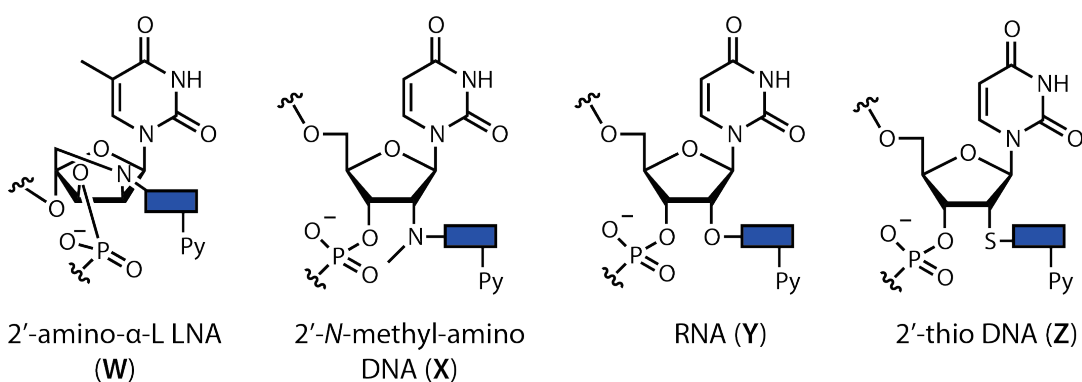
**Figure 1.3-1.** Regions of 2'-intercalator-functionalized nucleotides optimized for use as building blocks for Invader probes.

Precise positioning of the intercalators is a pre-requisite for the successful use of Invader probes for dsDNA recognition. The first building blocks that were used as components of Invader probes, were N2'-functionalized 2'-amino- $\alpha$ -L LNAs (**W**), which are 'locked' into a C2'-*endo* sugar conformation.<sup>29</sup> Intercalators at this position are preorganized to intercalate, as the strength of  $\pi$ - $\pi$  stacking between the nucleobase and intercalator directs it towards the duplex core.<sup>30</sup> However, the cumbersome synthesis of these monomers (<1% yield over 18 steps) motivated us to explore other nucleotide chemistries that exhibit similar properties based on earlier NMR studies based on DNA duplexes

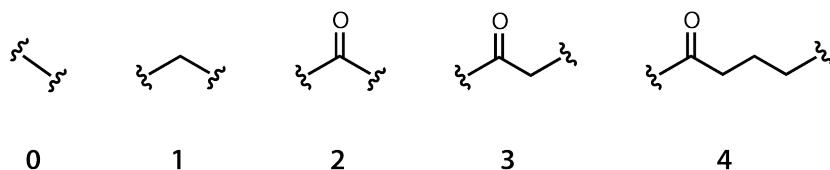
modified with the more flexible monomer 2'-*O*-(pyren-1-yl)methyl RNA monomers.<sup>37</sup> Intercalative binding modes are also supported by molecular dynamics simulations (Figure 1.2-2),<sup>35</sup> as well as photophysical and hybridization properties.<sup>31</sup>

A series of 2'-intercalator-functionalized monomers based on: 2'-*N*-methyl-amino DNA (**X**), RNA (**Y**), and thio-DNA (**Z**) skeletons were therefore explored (Figure 1.3-2). These sugar skeletons allow for functionalization with a variety of intercalators and linker chemistries in as few as four steps.<sup>38-40</sup> The nomenclature to describe the many modifications that were evaluated is as follows: the first letter denotes sugar followed by the number corresponding to the linker, and finally the intercalator is listed within parenthesis.

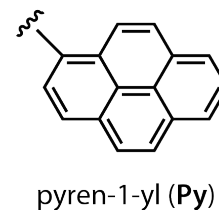
### Sugars



### Linkers



### Intercalator



**Figure 1.3-2.** Different monomers explored to facilitate efficient intercalation upon hybridization.

<sup>1</sup>H NMR spectra of nucleosides of 2'-N-methyl-amino DNA monomer **X1(Py)** suggests that it exclusively exists in *South* type conformations, presumably because of unfavorable interactions between the 2'-N-methyl and the 3'-oxygen in the *North* conformations. However, nucleosides of **Y1-** and **Z1(Py)**, however, are more equally populated between *North* and *South* type conformations.<sup>40</sup> Although the nucleoside sugar confirmation may not necessarily predict the adopted conformation in oligonucleotides or duplexes, it is of note that ONs that are modified with **X1-**, **Y1-**, or **Z1(Py)** display high affinity towards cDNA (Table 1.3-1), consistent with intercalative binding modes. The corresponding 2'-thio-DNA monomer **Z1(Py)**, however, results in only moderate stabilization of duplexes, likely due to steric interference of the larger 2'-sulfur atom and/or perturbation of the solvation layer.<sup>40</sup> As expected, double-stranded probes with +1 interstrand zipper arrangements of the monomers are relatively unstable. The origin of this destabilization is dominated by unfavorable enthalpy, caused by perturbed base pairing. The energetic difference in thermal stability between the probe duplex and probe:cDNA duplexes represents the thermodynamic driving force for dsDNA targeting of Invader probes and can be parameterized as the thermal advantage ( $TA = T_m(\text{probe:cDNA duplexes}) - T_m(\text{probe duplex} + \text{DNA target duplex})$ ), or – more accurately – change in free energy of recognition ( $\Delta G_{\text{rec}} = \Delta G(\text{probe:cDNA duplexes}) - \Delta G(\text{probe duplex} + \text{DNA target duplex})$ ). In general, all the probes are energetically activated for dsDNA recognition (compare TAs in Table 1.3-1). Additionally, the studied modified oligonucleotides display a strong preference for binding to DNA strands (Table 1.3-1 - DNA selectivity). This is consistent with an intercalative binding mode, of the pyrene as intercalating moieties' favor the less compressed B-type helix geometry of DNA duplexes.<sup>41,42</sup>

In summary, the sugar skeletons (**W/X/Y/Z**) provide a scaffold that facilitates precise intercalation of intercalators attached to the 2'-heteroatom. Invader probes that are based on **X1(Py)** and **Y1(Py)** monomers represent second-generation sugar skeletons that result in energetically activated probes for recognition of dsDNA and mimic the hybridization properties of the original **W1(Py)** monomer.

The length and flexibility of the linker that connects the intercalator to the sugar skeleton (Figure 1.3-2) has also been explored and optimized for conditions that favor intercalation. In the 2'-amino- $\alpha$ -L-LNA series (**W**) we demonstrated that incorporation of monomers with shorter linkers result in greater duplex stabilization than monomers with longer linkers (**2>3>>4**; Table 1.3-1), presumably due to the entropic penalty associated with longer linkers.<sup>30</sup> Moreover, monomers in which the pyrene moiety is linked via an amide bond induce greater increases in cDNA affinity than monomers linked via an amine bond (e.g., **W2>W1**). Conversely, for the 2'-*N*-methyl-amino DNA monomers, amine linkages are strongly favored over the more rigid amide linkers (e.g., **X1>X2**). This could reflect a difference of where the intercalator is positioned based on the sugar conformation; **W**-based monomers are more precisely positioned for intercalation with the rigid linker, but the other sugars necessitate a more flexible linker to achieve intercalation.

**Table 1.3-1.** Properties of Invader probes modified with various 2'-pyrene-functionalized nucleotides with different sugar skeletons and linkers between pyrene and the sugar.<sup>a</sup>

modification	thermal stability				TA	dsDNA recognition
	Inv:cDNA	DNA-selectivity	MM	probe duplex		
W1(Py)	++	++	-	+	+++	+++
W2(Py)	+++	++	---	-	+++	+++
W3(Py)	++	+	---	+	+++	++
W4(Py)	+	-	---	--	++	nd
X1(Py)	++	---	--	-	+++	+++
X2(Py)	-	-	-	---	+	-
X3(Py)	+	-	-	+	+	-
Y0(Py)	-	--	--	nd	nd	nd
Y1(Py)	++	+	--	-	+++	+++
Z1(Py)	+	++	---	--	++	-

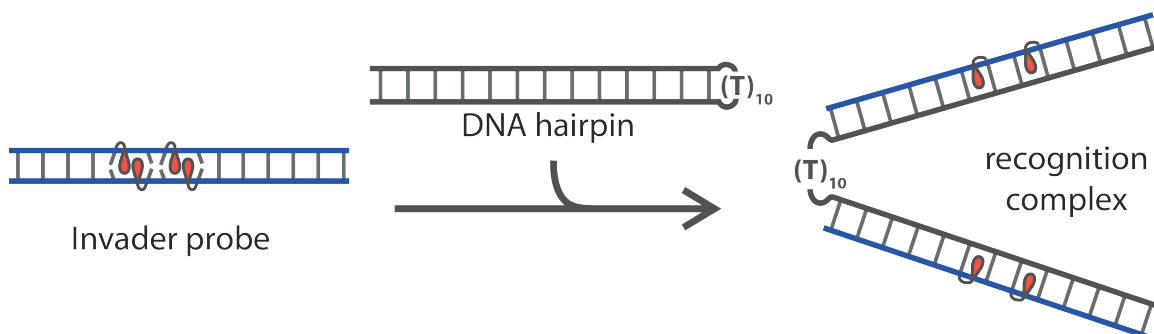
<sup>a</sup> For thermal stability: (+) indicates thermal stability greater than unmodified DNA; (-) indicates a lower thermal stability than unmodified DNA. MM indicates thermal discrimination of DNA containing a single nucleotide mismatch opposite the modified monomer; increasing number of (-) represents greater thermal discrimination. TA = thermal advantage (see text for description); increasing number of (+) indicates greater potential for recognizing isosequential dsDNA. dsDNA recognition efficiency is determined from EMSAs using DNA hairpins as targets; (-) no recognition observed. Number of symbols depicts the extent of the effect for the given parameter (e.g., +++ indicates effect is in the upper third greatest difference for that parameter compared to the other probes). nd = not determined. Compiled from data in Ref. 31, 39, and 40.

The activation of Invader probes for dsDNA-recognition was predicted by calculating the TA, or difference in thermal denaturation temperatures between product duplexes and reactant duplexes involved in the recognition of isosequential cDNA. All probes are significantly activated for recognition of isosequential dsDNA targets. In particular, Invader probes based on amino- $\alpha$ -L-LNA skeleton (**W**) are strongly activated due to very high stability of the cDNA duplexes and low stability of the probe duplex. The greatest activation is observed for probes in which the monomers have short linkers connecting the pyrene moiety and sugar skeleton (**W2**>**W1**; Table 1.3-1). Invader probes based on

second-generation monomer chemistries, 2'-*N*-methyl-amino DNA (**X1**) and 2'-functionalized RNA (**Y1**) display similar potential for dsDNA-recognition as **W1** and only slightly lower than **W2**.<sup>31</sup>

An electrophoretic mobility shift assay (EMSA) has been used to determine the ability of Invader probes to recognize the double-stranded stem region of a DNA hairpin (DH); pre-annealed Invader probes were incubated with the DNA hairpin target at physiological-like salinity and pH. The reaction mixture was then analyzed using non-denaturing polyacrylamide gel electrophoresis (nd-PAGE), where recognition of the DNA hairpin results in the formation of a ternary complex with significantly reduced mobility (Figure 1.3-3). This assay allows for quantification of dsDNA-recognition by Invader probes. The probes that were predicted to be the most strongly activated, indeed, resulted in the most efficient dsDNA-recognition of hairpin DNA (Table 1.3-1). It appears the thermal stability of the newly formed duplexes is a predominant parameter governing the formation of stable recognition complexes.

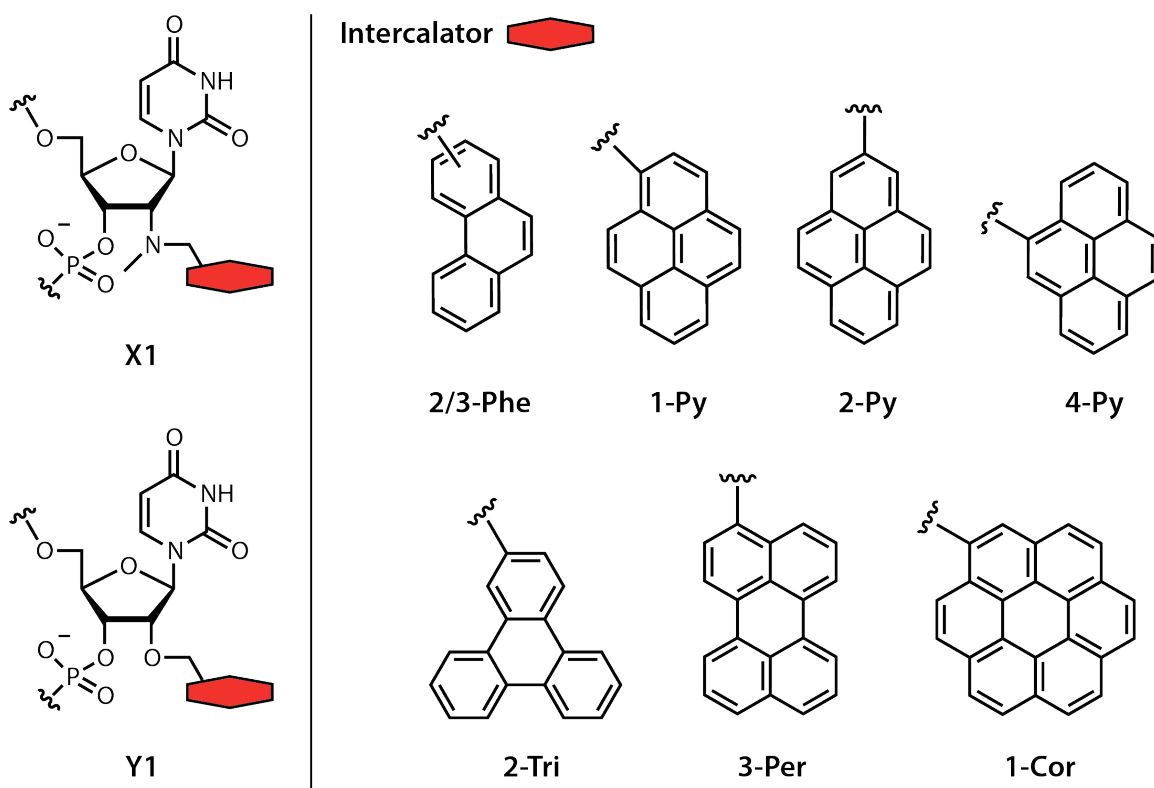
In summary, the linker connecting the intercalator to the sugar skeleton influences the hybridization properties of Invader probes and subsequent activation for dsDNA recognition. Invader probes were revealed to containing **W2(Py)**, **W1(Py)**, **X1(Py)** or **Y1(Py)** monomers, were found to be thermodynamically activated and allow for recognition of model dsDNA targets.<sup>31</sup>



**Figure 1.3-3.** Recognition of the stem regions in DNA hairpins by Invader probes.

Next, the influence of the intercalator was evaluated using 2'-*N*-methyl amino DNA (**X1**) or RNA (**Y1**) sugar skeletons (Figure 1.3-4).<sup>35,39</sup> Increasing the aromatic surface area of the intercalator was anticipated to increase the potential stacking interactions within the duplex core.  $T_m$ 's indicate this to be the case (stability of Inv:cDNA duplexes **Phe**<**Py**≤**Tri**<**Per**<**Cor**; Table 1.3-2). Even with the increased affinity towards cDNA, these probes strands maintain good thermal discrimination of singly mismatched DNA targets.





**Figure 1.3-4.** Various 2'-intercalator-functionalized nucleotide monomers used as components of Invader probes. Phe = phenanthren-2/3-yl, 1-Py = pyren-1-yl, 2-Py = pyren-2-yl, 4-Py = pyren-4-yl, Tri = triphenyl-2-yl, Per = perylene-3-yl, Cor = coronen-1-yl.

In general, double-stranded Invader probes modified with larger intercalators also display higher thermal stability (**Py**<**Per**<**Cor**). Although this is not desired, the energetic gradient is still maintained, due to the significantly increased thermal stability of the probe:cDNA duplexes.<sup>38</sup> This is reflected in the *TA* values, where Invader probes that are modified with intercalators with a surface area greater than phenanthrene (e.g., **Py/Per/Tri/Cor**) display prominent thermodynamic potential for dsDNA recognition. This suggests that a minimum aromatic surface is required for the intercalative effects to be realized. These probes also demonstrate efficient and specific recognition of DNA hairpins, with relatively little difference between the intercalators for the **X1**-series (Table 1.3-2).

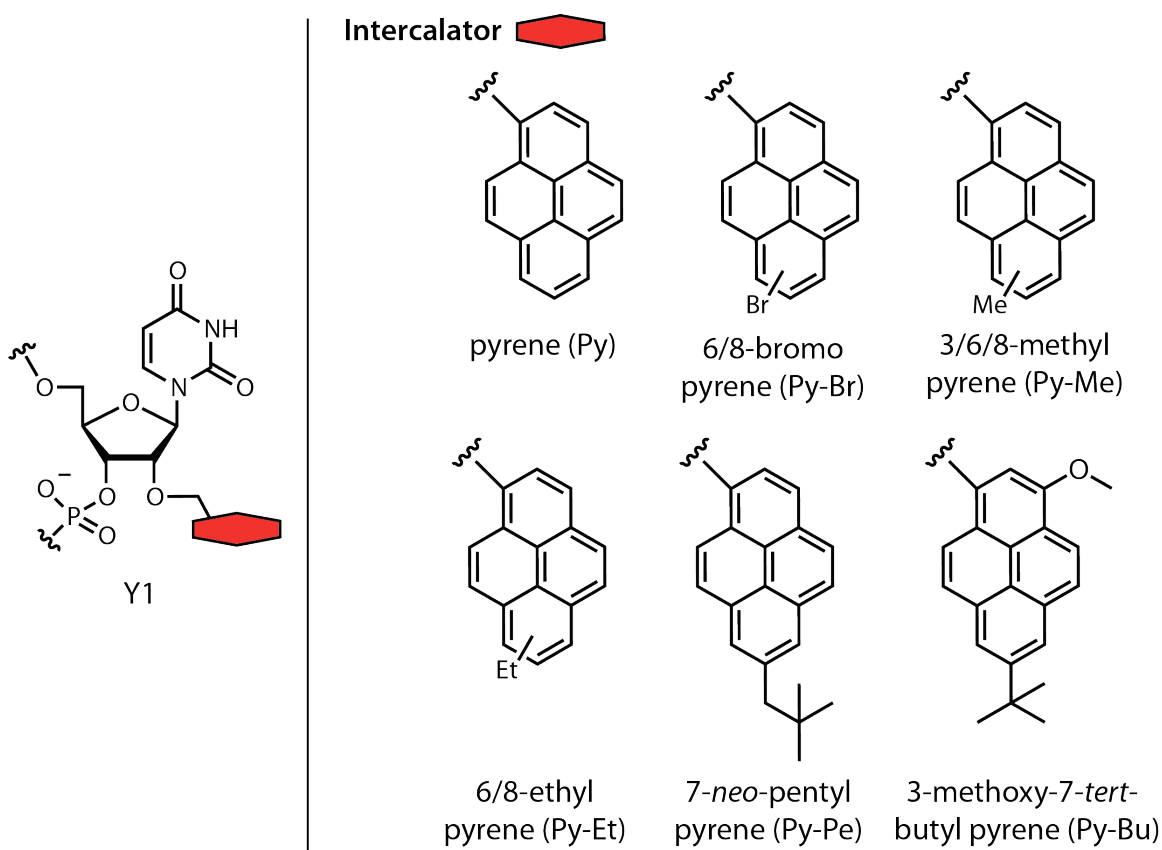
**Table 1.3-2.** Properties of Invader probes modified with various 2'-intercalator functionalized nucleotides.<sup>a</sup>

sugar and linker	intercalator	thermal stability				TA	dsDNA recognition
		Inv:cDNA	DNA selectivity	MM	probe duplex		
X1	1-Py	++	---	--	-	+++	+++
	Per	+++	+	---	+	+++	+++
	Cor	+++	+++	--	++	+++	+++
Y1	Phe	-	---	-	--	+	nd
	1-Py	++	+	--	-	+++	+++
	2-Py	++	++	--	-	+++	+++
	4-Py	+	-	--	---	++	-
	Tri	++	+	-	-	++	++
	Cor	+++	++	--	+++	++	++

<sup>a</sup> See Table 1.3-1 for column descriptions nd = not determined. Compiled from data in Ref. 35, 38, and 39.

The influence of the relative orientation between the pyrene moiety and the sugar skeleton was evaluated by varying the connectivity of the linker at the 1-, 2- or 4-linked pyrene derivatives (Figure 1.3-4).<sup>35</sup> Probes modified with **Y1(1-Py)** and **Y1(2-Py)** show very similar hybridization profiles and thermodynamic potential for dsDNA-recognition. In contrast, probes containing **Y1(4-Py)** display reduced thermal affinity towards cDNA and consequently lower potential for dsDNA recognition. Molecular modeling of duplexes between individual probe strands and cDNA suggests these trends to be due to differential stacking interactions between the pyrene and the flanking nucleobases. For example, the pyrene of **Y1(2-Py)** overlaps with all four nucleobases, while **Y1(1-Py)** only stacks with three of the four neighboring nucleobases. Only two nucleobases stack with the pyrene of **Y1(4-Py)**, which likely explains the low cDNA affinity. Probes bearing **Y1(4-Py)** failed to recognize the model hairpin DNA target, while efficient recognition by probes with **Y1(1-Py)** and **Y1(2-Py)** is observed.<sup>35</sup>

In an effort to further increase the available energetic gradient for dsDNA recognition, substituted pyrenes were investigated as building blocks for Invader probes based on with the hypothesis that steric bulk near the intercalator will result in additional destabilization in the probe duplex, while maintaining the high stability of duplexes between individual probe strand and cDNA. Five Invader monomers were synthesized with either small alkyl/halide (3/6/8-methyl, 6/8-ethyl, or 6/8-bromo) or bulky (7-*neo*-pentyl or 3-methoxy-7-*tert*-butyl) substituents on the pyrene of the **Y1** skeleton (Figure 1.3-5; unpublished results).



**Figure 1.3-5.** Invader probe building blocks based on 2'-alkylated RNA with substituted pyrene as intercalators. Note: Py-Br, Py-Me, and Py-Et were used as inseparable mixtures of isomers.

The substituted pyrene monomers are generally well tolerated in duplexes with cDNA. Attachment of smaller substituents at the 6/8-position induces even greater thermal stability of duplexes with cDNA than unmodified pyrene, whereas bulky substituents at the 7-position result in decreased stability presumably to do limited space within the duplex core (Table 1.3-3). The individual probe strands containing substituted pyrenes as intercalators maintain similar thermal mismatch discrimination of singly mismatched DNA targets as with the parent **Y1(Py)** probe.

**Table 1.3-3.** Properties of Invader probes modified with substituted pyrene building blocks.<sup>a</sup>

sugar and linker	intercalator	thermal stability				TA	dsDNA recognition
		Inv: cDNA	DNA-selectivity	MM	probe duplex		
 Y1	Py	++	+	--	-	+++	+++
	Py-Br	+++	-	--	-	+++	nd
	Py-Me	+++	++	--	--	+++	+++
	Py-Et	++	-	--	-	++	nd
	Py-Pe	+	+	--	--	++	nd
	Py-Bu	+	++	--	---	++	-

<sup>a</sup> See Table 1.3-1 for column descriptions. nd = not determined. Unpublished data.

Invader probes based on the additionally modified pyrene monomers display decreased thermal stability relative to probes based on unmodified pyrenes (**Py**). Invader probes modified with monomers featuring large substituents at the 7-position (**Py-Pe** and **Py-Bu**) result in especially destabilized probe duplexes.

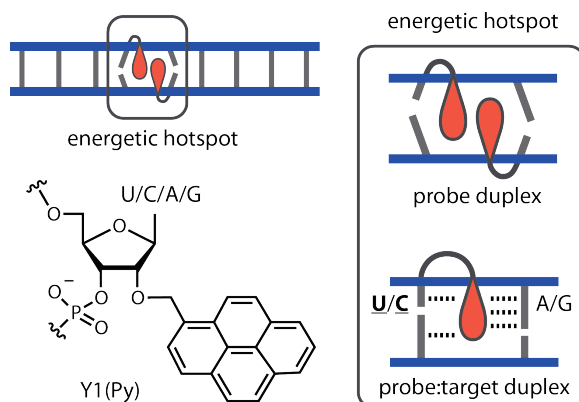
Comparison of the *TA* values for Invader probes based on substituted pyrene monomers, suggests that introduction of small substituents at the 6/8-position slightly increase the driving force for recognition of isosequential dsDNA, whereas substituents at the 7-

position of pyrene moiety reduces the targeting potential. This was confirmed in the EMSA, where Invader probes, based on the **Y1(Py-Me)** monomer, display similar or slightly better dsDNA recognition efficiency as **Y1(Py)**-based Invader probes. Probes based on the **Y1(Py-Bu)** monomer show no recognition of dsDNA presumably due to the low cDNA affinity.

#### 1.4 Optimization of Invader probe architectures

##### *Sequence preference of energetic hotspot*

The influence that the nucleobase moieties of the 2'-*O*-(pyren-1-yl)methyl RNA monomers, **Y1(Py)**, have on the dsDNA recognition efficiency of Invader duplexes was studied to identify any sequence preferences (Figure 1.4-1). Toward this end, the four canonical 2'-*O*-(pyren-1-yl)methyl-RNA A/C/G/U monomers were prepared and incorporated into ONs.<sup>43</sup> In general, the individual probe strands display greater affinity toward cDNA than unmodified DNA, although the G-monomer has a much less pronounced stabilizing effect than the A/C/U monomers. This appears to be due to weakening of the G:C base pair rather than non-intercalative binding modes of the pyrene, since bathochromic shifts of the pyrene absorption maxima, consistent with an intercalative binding mode,<sup>44</sup> are observed. The stabilizing effect of these monomers is much more pronounced when flanked by a 3'-purine rather than a 3'-pyrimidines. These observations are consistent with a 3'-directed intercalative binding mode, which has also been proposed from NMR analysis of the U-monomer.<sup>37</sup>



**Figure 1.4-1.** Sequence dependence of the energetic hotspot. Bolded and underlined letters represent **Y1(Py)**-modified nucleotides with the corresponding nucleobase.












Invader probes with all possible permutations of nucleobases within the energetic hotspot have been evaluated and consistently found to be less stable than the corresponding unmodified dsDNA. Invader probes are energetically activated regardless of which of the 2'-*O*-(pyren-1-yl)methyl-RNA **A/C/G/U** monomers are used for construction of the energetic hotspot. However, the thermodynamic potential for dsDNA recognition is greatest for Invader probes that have hotspots comprised of **U** or **C** monomers. This arrangement results in efficient stacking between the pyrene and the 3'-flanking purine and concomitantly increased stability of duplexes with cDNA. The ability of Invader probes containing one energetic hotspot to recognize model DNA hairpin targets closely correlates with the available thermodynamic driving force (**UA:AU**>**CG:GC**>**UT:AA**>>**GC:CG**), defined as  $\Delta G_{\text{rec}} = \Delta G (\text{probe duplex} + \text{dsDNA}) - \Delta G (\text{probe:cDNA duplexes})$ . Following this insight, a series of probes with three energetic hotspots were made that were predicted to display a range of dsDNA affinities. The experimentally determined dsDNA recognition efficiencies were indeed in line with expectations. Interestingly, even Invader probes that contained several weakly activated

hotspots resulted in observable dsDNA recognition, demonstrating minimal sequence limitations and flexibility in the design of Invader probes.

*Optimizing the number and location of energetic hotspots on dsDNA recognition by Invader probes.*

Different probe architectures were evaluated in an effort to increase the dsDNA-recognition efficiency of Invader probes. Thus, a library of 13-mer Invader probes with different numbers and locations of energetic hotspots comprised of either **W1(Py)**, **X1(Py)**, or **Z1(Py)** monomers were studied (Table 1.4-1).<sup>32,45</sup>

**Table 1.4-1.** Properties of various 13-mer Invader probe constructs.<sup>a</sup>

	Probe construct	duplex stability ( $\Delta G$ )			dsDNA recognition of:	
		Inv:cDNA	probe duplex	$\Delta G_{\text{rec}}$	PM	MM
W1(Py)		+	--	+	nd	nd
		++	-	++	nd	nd
		++	-	+++	nd	nd
X1(Py)		+	--	+	++	-
		++	-	++	+++	-
		++	+	++	++	-
		+++	-	+++	+++	+++
Y1(Py)		+	--	++	++	-
		++	-	+++	++	-
		++	+	+++	++	-
		+++	+	+++	+++	+++

<sup>a</sup> For duplex stability: (+) indicates affinity of duplex greater than unmodified dsDNA; (-) indicates a lower thermal stability than unmodified dsDNA.  $\Delta G_{\text{rec}}$  = free energy of recognition (see text for description); increasing number of (+) indicates greater potential for recognizing isosequential dsDNA. dsDNA recognition efficiency is determined from EMSAs using DNA hairpins as targets that are either perfectly matched (PM) or contain a single base pair mismatch with respect to the probe sequence (MM) but otherwise fully complementary. (-) no recognition observed. Number of symbols depicts the extent of the effect for the given parameter compared to other probes. nd = not determined. Compiled from data in ref. 32,45.

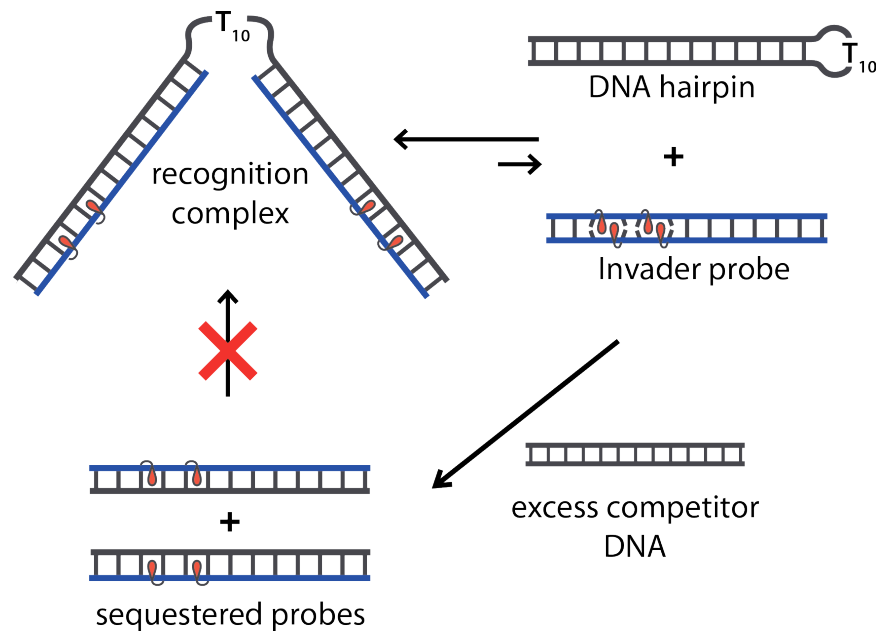
The cDNA affinity of individual Invader probe strands increases with higher level of modification (1<2<4 hotspots), as the intercalative effect of the pyrenes is maximized. The destabilizing effect of additional energetic hotspots in the double-stranded probe is not additive, possibly because the duplex already is perturbed after installment of the first hotspot. The thermodynamic driving force for dsDNA-recognition increases with additional energetic hotspots primarily because of the increased cDNA affinity, which



increases the energetic gradient between probe:cDNA duplexes and the double-stranded probe. Invader probes based on the first generation monomer chemistry, **W1(Py)**, were evaluated for mixed-sequence dsDNA recognition by incubation with pre-annealed isosequential linear dsDNA. Probes with either one or two energetic hotspots result in efficient recognition of the corresponding isosequential dsDNA target at equimolar quantities at a variety of ionic strengths up to 710 mM Na<sup>+</sup>.<sup>32</sup> Recognition occurs quickly, 50% within ~30 min, as inferred by the decreasing pyrene-pyrene excimer signal upon addition of the dsDNA target. The kinetics of recognition are temperature dependent, as increasing the temperature from 4 to 37 °C increases the rate of formation.

A library of Invader probes containing one, two, or four energetic hotspots were constructed using the second-generation monomers, **X1(Py)** and **Y1(Py)** were evaluated. All of these probes recognized the model DNA hairpin target, albeit to varying degrees (Table 11.4-1). In general, increasing the number of hotspots results in more efficient recognition as demonstrated by lower C<sub>50</sub>-values (concentration needed to achieve 50% recognition). Probes with one or two hotspots based on 2'-N-(pyren-1-yl)methyl-amino DNA **X1(Py)** result in slightly more efficient dsDNA-recognition than corresponding probes based on 2'-O-(pyren-1-yl)methyl RNA **Y1(Py)** monomers, while little monomer-specific differences were observed for probes with four hotspots. Increasing the number of hotspots also results in faster recognition kinetics, with **Y1(Py)**-modified probes displaying slightly faster kinetics than **X1(Py)**-modified probes. Additionally, increasing the number of energetic hotspots results in slower dissociation as determined by a competitive dissociation assay, where the recognition complex was challenged with a large excess of linear competitor dsDNA (Figure 1.4-2). The linear isosequential dsDNA

sequesters any dissociating probes, preventing reformation of the recognition complex. Recognition complexes formed with Invader probes containing a single hotspot dissociate within 24 h for probes with **X1(Py)**- and 8 h for **Y1(Py)**-based probes, whereas probes containing two hotspots are much more stable with >60% remaining after 24 h.<sup>45</sup>



**Figure 1.4-2.** Illustration of assay used to determine dissociation kinetics of recognition complex.

The binding specificity of these probes was evaluated using DNA hairpins in which the sequence of the stem differs in one position relative to the Invader probe. Probes with one or two hotspots display excellent discrimination of single-mismatched DNA hairpins, at conditions where complete recognition of matched DNA hairpins occurs. However, only partial discrimination is observed for the highly modified probes, presumably due to the very high cDNA affinity of the individual probe strands (Table 1.4-1).

Changing the number, location, and chemistry of the energetic hotspots within the Invader probe representing a modular approach for tuning the energetics and specificity of dsDNA.

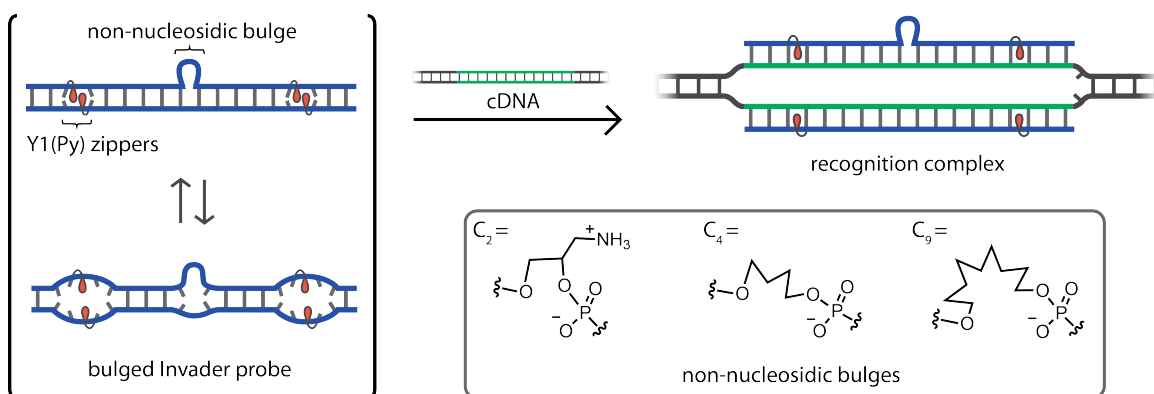
*Enzymatically stable phosphorothioate Invader probes*

To further facilitate biological applications of Invader probes, we set out to design probes that are enzymatically stable in cellular environments by replacing the native phosphodiester (PO) backbone, which is susceptible to nuclease degradation, with a phosphorothioate (PS) backbone (Figure 1.4-3).<sup>46</sup> Bathochromic shifts of pyrene absorption maxima and reduced fluorescence emission upon hybridization with cDNA suggest that probe strands with PS-backbones bind to cDNA in a similar manner as PO-based probe strands, thus ensuring that intercalation of the pyrene moiety still ensues. While 9 mer PS-based Invader probes display a moderately favorable thermodynamic gradient for recognition of isosequential dsDNA, it is not sufficient to drive recognition of model DNA hairpin targets. Longer probes (13 mers) with at least two energetic hotspots had to be used to ensure recognition of their DNA hairpin targets.

Incorporation of Invader monomers **Y1(Py)** into a PO backbone confers moderate protection against 3'-exonuclease (snake venom phosphodiesterase) of individual probe strands relative to unmodified DNA with 95 % degradation occurring in 50 min or 15 min, respectively. Probes with **X1(Py)** monomers are more stable yet with 95% degradation occurring within ~21 h. PS-Invader probe strands, on the other hand, are inert to exonuclease degradation and offer a promising backbone modification to enable applications within cellular environments.

*Bulged Invader probes: development of more labile duplexes for increased recognition of dsDNA*

In order to develop Invader probes capable of efficient and specific recognition of longer target sequences, we developed probes that contain non-nucleosidic bulges based on the hypothesis that these bulges will generate more labile double-stranded probe duplexes, as stacking is perturbed, leading to accelerated nucleation with, and invasion of dsDNA targets. A library of Invader probes featuring a centrally located bulge on either one strand or both probe strands were synthesized (Figure 1.4-3). The influence of the bulge length on the thermal hybridization profiles and dsDNA recognition characteristics was evaluated using three different linker chemistries ( $C_2$ ,  $C_4$ , or  $C_9$ ), which were incorporated once or three consecutive times (unpublished results).



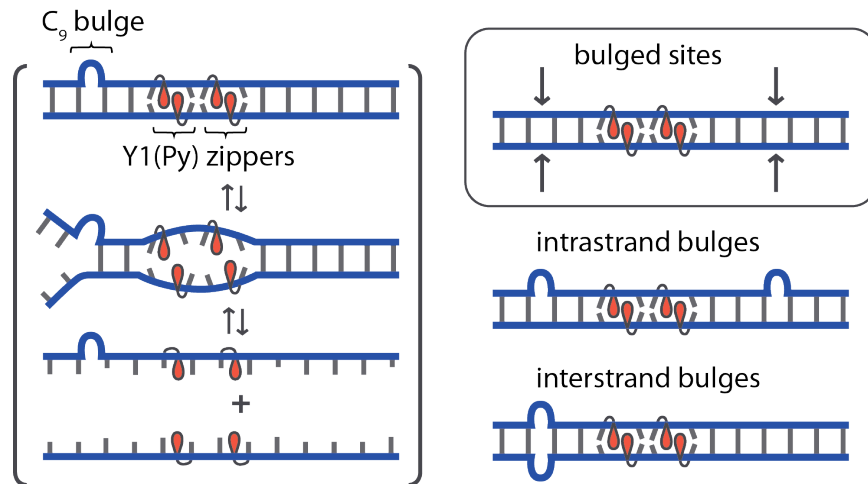
**Figure 1.4-3.** Invader probes containing non-nucleosidic bulges for increased recognition of dsDNA.

Introduction of a bulge into Invader probe strands invariably reduces the cDNA affinity of the individual probe strands relative to conventional probe strands. Incorporation of bulges into one strand of the Invader probes results in length dependent destabilization of the probe duplex ( $C_9 > C_4 > C_2$ ). Extending the length of the small bulges through additional incorporations ( $3 \times C_2$  or  $3 \times C_4$ ) further destabilizes the probe duplex. However, further

increases in the length of the  $C_9$  bulge ( $C_9 \times 3$ ) do not result in additional destabilization. This suggests that a single  $C_9$  bulge is sufficient to break the cooperativity of base stacking in the duplex core. Bulged Invader probes can also be effectively destabilized through modification of both strands with the shorter bulges ( $3 \times C_2$  or  $C_4$ ). Importantly, the difference in stability between the double-stranded probe and probe:cDNA duplexes is maintained or increased. Invader probes with a bulge in just one strand generally result in more efficient recognition of model DNA hairpin targets than conventional Invader probes, while similar dsDNA-recognition is observed if bulges of  $C_2$  or  $C_4$  monomers, as single or triple incorporations, are located on both probe strands. However, probes with bulges of  $C_9$  on both strands do not recognize DNA hairpin targets. In the latter design any kinetic advantage is likely outweighed by the reduced stability of the probe:target duplexes. Bulged Invader probes show excellent binding specificity for cDNA, as DNA hairpins whose stem region differs in one position relative to the Invader probes are readily discriminated.

Bulged Invader probes containing a single  $C_9$  bulge are particularly efficient at recognizing dsDNA and this bulge chemistry was therefore used to optimize the architecture of Invader probes further. Probes with  $C_9$ -bulges located near one or both the termini were evaluated, based on the hypothesis that the non-nucleosidic bulge will promote end-fraying, thus revealing the Watson-Crick face of the probe strands and increasing the probability of nucleation events for improved rates of dsDNA recognition. Bulged Invader probes containing either one or two bulges in either interstrand (on opposite strands of the probe duplex) or intrastrand (on the same strand of the probe

duplex) arrangements (Figure 1.4-4) were synthesized and characterized with respect to thermal hybridization properties and recognition of dsDNA.<sup>47</sup>







**Figure 1.4-4.** Representative examples of Invader probes with intra- or interstrand arrangements of bulges of  $C_9$  monomers.

Individual probe strands containing bulges display reduced affinity toward cDNA relative to the conventional probe strands, with strands containing two bulges resulting in the greatest decreases in affinity. Double-stranded probes containing a single bulge are destabilized with respect to the conventional Invader probe, but probes with two intra- or interstrand arrangements of bulges result in particularly destabilized probes. Bulged Invader probes with one or two  $C_9$  bulges have similar or greater thermodynamic potential for dsDNA recognition relative to the conventional probe, since the probe duplex is much more unstable or labile. Increasing the number of bulges in the Invader probes results in more efficient recognition of the model DNA hairpin targets, as long as at least one of the resulting probe:cDNA duplexes is thermodynamically stable. Invader probes with interstrand arrangements of bulges recognize dsDNA more at lower temperatures (8 °C versus 22 °C). This likely reflects easier denaturation of the

particularly labile probe duplexes, rendering the Watson-Crick faces of the probe strands more accessible for nucleation than conventional Invader probes.

**Table 1.4-2.** Thermal properties and dsDNA recognition of Bulge Invader probe constructs.<sup>a</sup>

Invader probe	TA	recognition of dsDNA at:		rate of formation	dissociation kinetics
		8 °C	22 °C		
	++	+	+	+	++
	++	++	+++	++	++
	+++	+++	+++	++	+
	++	+++	++	+++	+++

<sup>a</sup> TA = thermal advantage for recognition of isosequential dsDNA (see text for description); increasing number of (+) indicates greater effect. dsDNA recognition efficiency is determined from EMSAs using DNA hairpin targets. Dissociation kinetics were determined using competitive dissociation assay; increasing number of (+) denotes slower off-rates. Number of symbols depicts the extent of the effect for the given parameter compared to other probes. nd = not determined. Data compiled from ref. 47.

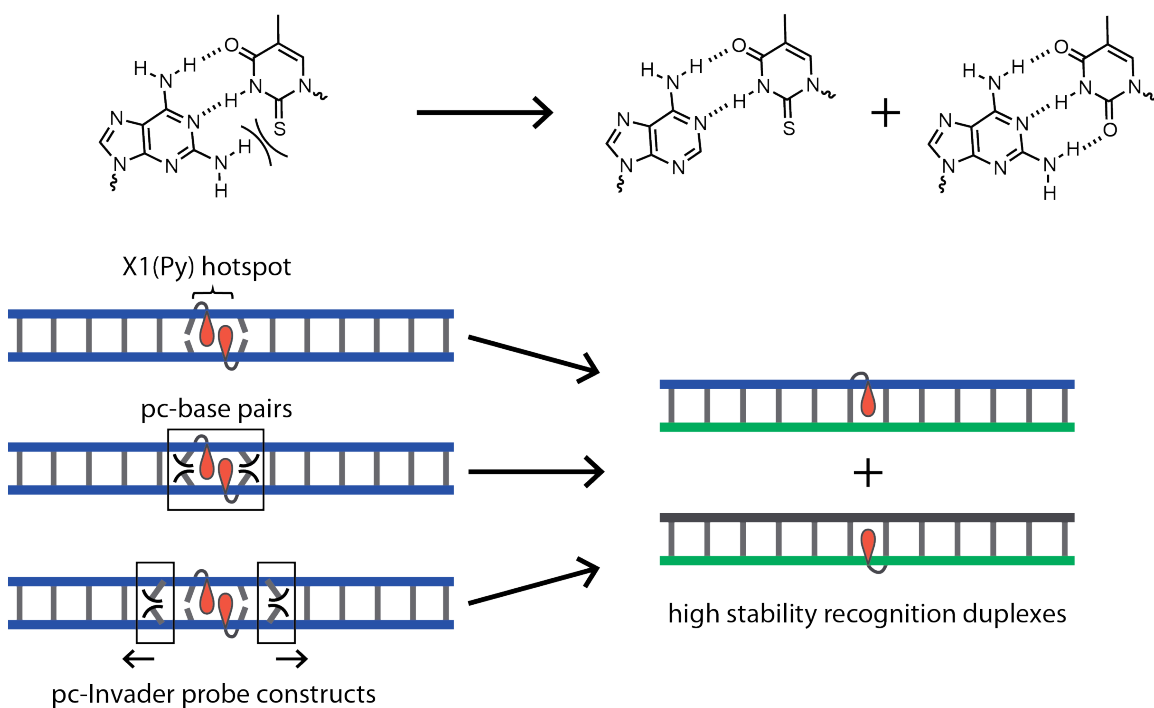
Introduction of a near terminal C<sub>9</sub> bulge increases the rate of recognition as much as 4-fold relative to the conventional probe that does not contain a bulge. This is likely due to kinetic advantage of more labile probe duplexes. The dissociation of the recognition complex was evaluated using the competition assay shown in Figure 1.4-2. In general, the use of Invader probes containing near-terminal bulges results in faster dissociation of the recognition complexes (>90% dissociated within 6 h) compared to conventional probes. This is not unexpected considering the lower cDNA affinity of the bulge-containing probe strand. Surprisingly, recognition complexes formed with Invader probes containing two bulges in intrastrand arrangements persists for >24 h. In this construct, the bulged strand does not stably bind to the recognition complex and may be playing a role in slowing down dissociation by transiently binding to the recognition complex and/or weak

affinity towards the competitor DNA strand. In summary, bulged Invader probes improve the rate and lower the concentration needed for dsDNA recognition relative to conventional probes.

### *Pseudocomplementary Invader probes*

In an alternative approach toward improving the kinetics and thermodynamics of dsDNA recognition, the concept of pseudocomplementary base pairs was merged with the underlying principle of Invader probes. The Invader probe approach is similar to the use of pseudocomplementary DNA, in as much it relies on a thermodynamic gradient arising from labile probe duplexes but more stable probe:cDNA duplexes. We hypothesized that merging these two strategies would result in synergistically improved recognition of dsDNA.

pseudocomplementary base pairing







**Figure 1.4-5.** The principle of pseudocomplementary Invader probes.



Pseudocomplementary Invader probes were synthesized to contain a centrally located energetic hotspot composed of **X1(Py)** monomers and pseudocomplementary base pairs either within the energetic hotspot, which required synthesis of the novel intercalator-functionalized monomer, or on neighboring base pairs that were incorporated using commercially available phosphoramidites (Figure 1.4-5).

Surprisingly, when the energetic hotspot is composed of pseudocomplementary base pairs, the thermodynamic gradient for dsDNA recognition is lower than for conventional Invader probes (Table 1.4-3). This is a result of lower cDNA affinity without additional destabilization of probe duplex. A likely explanation why this is the case is that the steric clashes between the pseudocomplementary nucleobases do not occur in the already perturbed energetic hotspot. However, when the pseudocomplementary base pairs are moved away from the energetic hotspot, both structural elements contribute favorably by decreasing the stability of the probe duplex while maintaining the high stability of probe:cDNA duplexes, thus resulting in more favorable energetics for dsDNA recognition. Thus, pc-Invader probes represent an alternative approach for achieving the necessary energetic gradient for recognition of dsDNA by modification of the conventional Invader probe construct (unpublished results).

**Table 1.4-3.** Properties of pc-Invader probes.

probe construct	duplex stability ( $\Delta G$ )			$\Delta G_{\text{rec}}$	dsDNA recognition
	Inv:cDNA	probe duplex			
	++	-	++	++	
	+	--	++	++	
	++	---	+++	+++	
	++	---	+++	+++	

<sup>a</sup> For duplex affinity: (+) indicates affinity of duplex greater than unmodified dsDNA; (-) indicates a lower thermal stability than unmodified dsDNA.  $\Delta G_{\text{rec}}$  = free energy of recognition (see text for description); increasing number of (+) indicated greater potential for recognizing isosequential dsDNA. dsDNA recognition is determined for targeting DNA hairpins using EMSA. Number of symbols depicts the extent of the effect for the given parameter compared to other probes.

#### *Toe-hold Invader probes*

Invader probes with single-stranded overhangs (toe-holds) and two centrally located energetic hotspots of **Y1(Py)** monomers were evaluated for recognition of a linear 33 mer dsDNA where an internal target region. These probes were developed with the hypothesis that the single-stranded regions of the probe strand can readily bind to the transiently exposed Watson-Crick face, or breathing, of target DNA, thus increasing nucleation events and subsequent strand invasion. Probes were synthesized with 3 or 6 nt overhangs on the conventional probe architecture and evaluated using a mobility shift assay, where invasion of the target dsDNA results in the formation of two binary recognition complexes (individual probe strands hybridized with 33 mer target strands) with higher mobility on gel electrophoresis than the dsDNA target alone. Blunt-ended probes and probes with 3 nt overhangs failed to recognize the linear dsDNA target. Toe-hold Invader probes with 6 nt efficiently recognize the internal mixed-sequence dsDNA target at non-

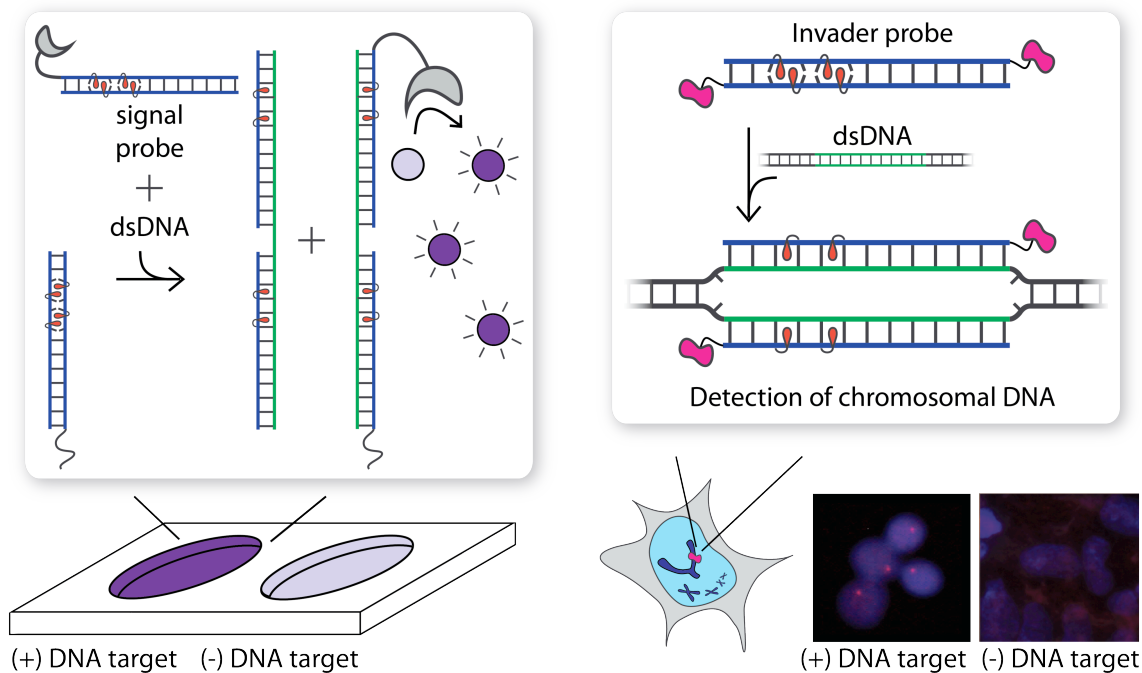
denaturing conditions. Invasion efficiency is improved further by incorporating two **Y1(Py)** monomers in each of the toe-hold regions, presumably due to the increased stability of the nucleation complex, thus leading to a higher probability of successful duplex invasion.

Linear dsDNA containing one, two, or three base pairs differing in sequence relative to toe-hold Invader probes was used to evaluate the specificity of recognition. Probes with unmodified toe-hold regions efficiently discriminate these non-complementary targets. However, some binding of toe-hold Invader probes with **Y1(Py)** monomers in the toe-hold region occurs with the mismatched target sequences, though to a lesser extent than the matched target.

### **1.5 Applications of Invader probes**

Hybridization-based probes capable of sequence-specific recognition of dsDNA are of interest for diagnostic applications in food safety where rapid, specific, and sensitive detection of bacterial contamination is desirable. We developed a sandwich assay utilizing Invader probes for the detection of dsDNA specific to pathogenic bacteria. This assay was designed for use with a 96-well plate fluorescence plate reader, where a double-stranded capture probe – an Invader probe that is modified with **Y1(Py)**-based energetic hotspots – is immobilized on the plate. It is activated for recognition of a dsDNA region specific to food pathogens such as *Salmonella enterica*, *Campylobacter jejuni*, or *Escherichia coli*. In the presence of the specific dsDNA and a biotin-labeled signaling Invader probe two ternary complexes are formed, one in solution and one that is attached to the microplate (Figure 1.5-1, left). Excess signaling probe is then removed

followed by addition of streptavidin-horseradish peroxidase conjugate and substrate resulting in the formation of a fluorescent signal that can be used for quantification of specific pathogenic dsDNA.<sup>48</sup> Detection of common bacterial contamination in food preparation was demonstrated using this sandwich-based assay. The assay is conducted at isothermal conditions, avoids laborious sample preparation, and is highly sensitive (down to 20 pM). Any mixed-sequence targets can, in principle, be detected via this approach, which renders Invader probes as a promising approach for diagnostic and biotechnological applications.



**Figure 1.5-1.** Sandwich-based assay for detection and quantification of dsDNA (left) and non-denaturing fluorescence in situ hybridization (FISH) assay for detection of chromosomal DNA (right).

Invader probes were developed for use in non-denaturing fluorescent in situ hybridization (nd-FISH) assays, where Cy3-labelled Invader probes allowed for visualization of specific DNA loci using fluorescent microscopy.<sup>45</sup> Invader probes, based on the

optimized 2'-*O*-(pyren-1-yl)methyl RNA **Y1(Py)** monomers and probe architectures, were designed to target a unique region within the *DYZ-1* satellite region ( $\sim 6 \times 10^4$  tandem repeats) of the bovine Y-chromosome. Gratifyingly, incubation of these probes with fixed interphase nuclei and metaphasic spreads from a male bovine cell line result in a punctate fluorescent signal (Figure 1.5-1, left), with excellent labeling coverage ( $\sim 90\%$ ). Specific binding was inferred from a series of control experiments that did not produce any signals, i.e., Y-chromosome specific probes with nuclei isolated from female bovine fibroblast cells lacking the target, incubation of probes differing in sequences at three positions with respect to the target sequence, and pre-treatment of cells with DNase. Experiments where the nuclei were pre-treated with RNase or proteinase still showed punctate signals, suggesting that dsDNA is the molecular target of Invader probes. Invader probes recognize specific mixed-sequence regions of chromosomal DNA efficiently and specifically under non-denaturation conditions.

## 1.6 Conclusions

The thermodynamic driving force of dsDNA recognition by Invader probes depends on the ability to precisely position intercalators with a duplex. The original monomer that facilitated intercalation utilized the 2'-amino- $\alpha$ -L LNA sugar skeleton, however, synthetic challenges prompted us to explore other building blocks for use in Invader probes. Monomers have been optimized with respect to the i) sugar skeleton of 2'-intercalator-functionalized nucleotide and ii) the nature of the intercalator: surface area, connectivity and substitution patterns. Second-generation monomers, 2'-*O*-(pyren-1-yl)methyl RNA and 2'-*N*-(pyren-1-yl)methyl-amino DNA, emerged as synthetically accessible functional and structural mimics of the original monomer. These sugar

skeletons were then used to optimize the intercalator, where pyrene was shown to be the minimal aromatic surface area necessary to achieve the energetic gradient necessary for dsDNA recognition.

Optimal dsDNA recognition properties occur with 1- or 2-linked pyrene derivatives. Invader probes recognize mixed-sequence dsDNA targets, however there is a preference for a 3'-purine flanking the 2'-intercalator-functionalized monomer. This sequence preference of the modified region can be partially alleviated by incorporation of several energetic hotspots of any sequence composition. In order to achieve efficient recognition of dsDNA, while still maintaining specificity there should be approximately one energetic hotspot for every 4-5 base pairs of the Invader probe.

Different probe architectures were explored that promote more labile probe duplexes to improve the thermodynamics and kinetics of recognition. Introduction of non-nucleosidic bulges into Invader probes dramatically improves the rate of dsDNA recognition. A single bulge of ~12 atoms in length is enough to achieve this rate enhancement. Invader probes with pseudocomplementary base pairs improve the efficiency of invasion when they are located away from energetic hotspots.

We have demonstrated that Invader probes recognize dsDNA efficiently and specifically in cell-free assays at physiological-like conditions without the need for denaturation or annealing. Fluorophore-labeled Invader probes were shown to detect a Y-chromosome specific region in isolated nuclei under non-denaturing conditions. In order to use these probes for live-cell imaging, the target dsDNA region must be accessible for successful duplex invasion to occur (e.g., transcriptionally active regions). Additionally, probes

must be trafficked to the nucleus, likely necessitating further modification with, e.g., nuclear-localization signals.

Invader probes offer a promising tool to enable mixed-sequence recognition of dsDNA at non-denaturing conditions. Proof-of-concept experiments demonstrate the potential applications of Invader probes for detection of dsDNA in food safety and fluorescent in situ hybridization assays, and hint at much broader applications involving fundamental molecular biology, disease diagnostics, and DNA nanotechnologies.

## 1.7 References

1. Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. J. *Mol. BioSyst.* **2006**, *2*, 551.
2. Nielsen, P. E. *Chem. Biodiv.* **2010**, *7*, 786.
3. Aiba, Y.; Sumaoka, J.; Komiyama, M. *Chem. Soc. Rev.* **2011**, *40*, 5657.
4. Vaijayanthi, T.; Bando, T.; Pandian, G. N.; Sugiyama, H. *ChemBioChem* **2012**, *13*, 2170.
5. Marraffini, L. A.; Sontheimer, F. J. *Nat. Rev. Genet.* **2010**, *11*, 181.
6. Turitz Cox, D. B.; Platt, R. J.; Zhang, F. *Nat. Med.* **2015**, *21*, 121.
7. Chen, B.; Gilbert, L. A.; Cimini, B. A.; Schnitzbauer, J.; Zhang, W.; Li, G.-W.; Park, J.; Blackburn, E. H.; Wessman, J. S.; Qi, L. S.; Huang, B. *Cell* **2013**, *155*, 1479.
8. Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, *36*, 5123.
9. Buchini, S; Leumann, C. J. *Curr. Opin. Chem. Biol.* **2003**, *7*, 717.
10. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284.

11. Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem.* **2013**, *21*, 6101.
12. Mukherjee, A.; Vasquez, K. M. *Biochimie* **2011**, *93*, 1197.
13. Bailly, C.; Chaires, J. B. *Bioconj. Chem.* **1998**, *9*, 513.
14. Praseuth, D; Guieysse, A. L.; Helene, C. *Biochim. Biophys. Acta* **1999**, *1489*, 181.
15. 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.
16. Bentin, T.; Larsen, H. J.; Nielsen, P. E. *Biochemistry* **2003**, *42*, 13987.
17. Kaihatsu, K.; Shah, R. H.; Zhao, X.; Corey, D. R. *Biochemistry* **2003**, *42*, 13996.
18. Janowski, B. A.; Kaihatsu, K.; Huffman, K. E.; Schwartz, J. C.; Ram, R.; Hardy,  
D.; Mendelson, C. R.; Corey, D. R. *Nat. Chem. Biol.* **2005**, *1*, 210.
19. Zhang, X.; Ishihara, T.; Corey, D. R. *Nucleic Acids Res.* **2000**, *28*, 3332.
20. Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.;  
Ly, D. H. *J. Am. Chem. Soc.*, **2006**, *128*, 10258.
21. Rapireddy, S.; Bahal, R.; Ly, D. H. *Biochemistry* **2011**, *50*, 3913.
22. Bahal, R.; Sahu, B.; Rapireddy, S.; Lee, C-M.; Ly, D. H. *ChemBioChem* **2012**, *13*,  
56.
23. Ishihara, T.; Corey, D. R. *J. Am. Chem. Soc.* **1999**, *121*, 2012.
24. Kutuyavin, I. V.; Rhinehart, R. L.; Lukhtanov, E. A.; Gorn, V. V.; Meyer Jr., R. B.;  
Gamper Jr., H. B. *Biochemistry* **1996**, *35*, 11170.
25. Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11804.
26. 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.



27. Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corrandini, R.; Sforza, S.; Komiyama, M. *Nucleic Acids Res.* **2008**, *36*, 1464.
28. Sumaoka, J.; Komiyama, M. *Chem. Lett.*, **2014**, *43*, 1581.
29. Hrdlicka, P. J.; Kumar, T. S.; Wengel, J. *Chem. Commun.* **2005**, 4279.
30. Kumar, T. S.; Madsen, A. S.; Østergaard, M. E.; Sau, S. P.; Wengel, J.; Hrdlicka, P. J. *J. Org. Chem.* **2009**, *74*, 1070.
31. 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.
32. Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2010**, *8*, 2028.
33. Crothers, D. M. *Biopolymers* **1968**, *6*, 575.
34. Persil, O.; Hud, N. V. *Trends Biotechnol.* **2007**, *25*, 433.
35. Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2014**, *12*, 7758.
36. Winnik, F. M. *Chem. Rev.* **1993**, *93*, 587.
37. Nakamura, M.; Fukunaga, Y.; Sasa, K.; Ohtoshi, Y.; Kanaori, K.; Hayashi, H.; Nakano, H.; Yamana, K. *Nucleic Acids Res.* **2005**, *33*, 5887.
38. Anderson, B. A.; Onley, J. J.; Hrdlicka, P. J. *J. Org. Chem.*, **2015**, *80*, 5395.
39. Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T. B.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem.* **2011**, *76*, 7119.
40. Anderson, B. A.; Hrdlicka, P. J. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 3999.
41. Korshun, V. A.; Stetsenko, D. A.; Gait, M. J. *J. Chem Soc., Perkin Trans. 1*, **2002**, 1092.

42. Dohno, C.; Saito, I. *ChemBioChem* **2005**, *6*, 1075.
43. Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, *78*, 12040.
44. Asanuma, H.; Fujii, T.; Kato, T.; Kashida, H. *J. Photochem. Photobiol., C* **2012**, *13*, 124.
45. Guenther, D. C.; Anderson, G. H.; Karmakar, S.; Anderson, B. A.; Didion, B. A.; Guo, W.; Verstegen, J. P.; Hrdlicka, P. J. *Chem. Sci.*, **2015**, *6*, 5006.
46. Anderson, B. A.; Karmakar, S.; Hrdlicka, P. J. *Molecules* **2015**, *20*, 13780.
47. Guenther, D. C.; Karmakar, S.; Hrdlicka, P. J. *Chem. Comm.* **2015**, *151*, 15051.
48. Denn, B.; Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *Chem. Commun.* **2013**, *49*, 9851.

**CHAPTER 2: Invader probes: Harnessing the energy of intercalation to facilitate recognition of chromosomal DNA for diagnostic applications**

Dale C. Guenther,<sup>a</sup> Grace H. Anderson,<sup>a,b</sup> Saswata Karmakar,<sup>a</sup> Brooke A. Anderson,<sup>a</sup>  
Bradley A. Didion,<sup>c</sup> Wei Guo,<sup>c</sup> John P. Versteegen<sup>c</sup> and Patrick J. Hrdlicka.<sup>a,\*</sup>

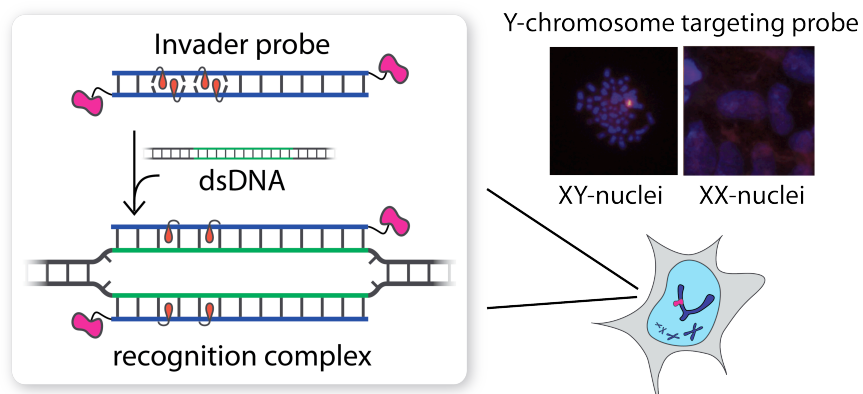
<sup>a</sup> Department of Chemistry, University of Idaho, 875 Perimeter Dr, Moscow, ID 83844-2343, USA.

<sup>b</sup> Department of Biological Sciences, Montana Tech of the University of Montana, 1300 W Park St, Butte, MT 59701-8997, USA.

<sup>c</sup> MoFA, PO Box 930187, 419 Venture Ct., Verona, WI 53593, USA.

Published in: *Chem. Sci.* **2015**, *6*, 5006-5015

Published by The Royal Society of Chemistry.



## Abstract

Development of probes capable of recognizing specific regions of chromosomal DNA has been a long-standing goal for chemical biologists. Current strategies such as PNA, triplex-forming oligonucleotides, and polyamides are subject to target choice limitations and/or necessitate non-physiological conditions, leaving a need for alternative approaches. Toward this end, we have recently introduced double-stranded oligonucleotide probes that are energetically activated for DNA recognition through modification with +1 interstrand zippers of intercalator-functionalized nucleotide monomers. Here, probes with different chemistries and architectures – varying in the position, number, and distance between the intercalator zippers – are studied with respect to hybridization energetics and DNA-targeting properties. Experiments with model DNA targets demonstrate that optimized probes enable efficient ( $C_{50} < 1 \mu\text{M}$ ), fast ( $t_{50} < 3\text{h}$ ), kinetically stable ( $> 24\text{h}$ ), and single nucleotide specific recognition of DNA targets at physiologically relevant ionic strengths. Optimized probes were used in non-denaturing fluorescence in situ hybridization experiments for detection of gender-specific mixed-sequence chromosomal DNA target regions. These probes present themselves as a promising strategy for recognition of chromosomal DNA, which will enable development of new tools for applications in molecular biology, genomic engineering and nanotechnology.

## 2.1 Introduction

There is an unmet need for chemical probes capable of recognizing biological DNA for identification, regulation, and manipulation of genes.<sup>1-7</sup> Considerable progress has been

made towards this end with triplex forming oligonucleotides (TFOs),<sup>8</sup> polyamides<sup>9,10</sup> peptide nucleic acids (PNA),<sup>11,12</sup> and – more recently – engineered proteins,<sup>3,13,14</sup> though significant limitations exist with all of these approaches. For example, TFOs only form Hoogsteen base pairs in the major groove of DNA duplexes containing long purine tracts, which reduces the number of suitable target sites within a genome.<sup>8,12</sup> Pyrrole-imidazole minor groove of DNA duplexes, but typically only recognize short target regions (<8 bp), which may impede recognition of unique genomic sites.<sup>9,10,15</sup> This is less of a concern with engineered nucleases but their construction requires the use of advanced molecular cloning techniques,<sup>14</sup> and there are mounting concerns regarding the high frequency of off-target effects.<sup>16</sup> PNAs, in which canonical nucleobases are attached to an *N*-(2-aminoethyl)glycine backbone, display strong affinity towards complementary single-stranded DNA (ssDNA), allowing for strand invasion of double-stranded DNA (dsDNA) through simultaneous Watson-Crick and Hoogsteen base pairing. However, recognition of dsDNA using regular PNAs is typically subject to similar sequence limitations as TFOs,<sup>11,12</sup> although alternative PNA-based strategies with more relaxed sequence requirements have been developed.<sup>17,18</sup> The use of a conformationally restricted PNA backbone,  $\gamma$ -PNA, substantially increases the binding affinity towards ssDNA, presumably due to strand preorganization and reduced entropic penalties. Single-stranded  $\gamma$ -PNAs have been shown to recognize mixed-sequence dsDNA target regions (150-300 bp) at low ionic strengths via duplex invasion, resulting in the formation of a D-loop, in which a segment of one of the DNA target strands is unhybridized.<sup>19,20</sup> Nonetheless, invasion is inefficient at physiological-like ionic strength.

Double-duplex invasion mechanisms, in which a double-stranded probe binds to both strands of a DNA target, are desirable due to the potential stability gain from having both target strands engaged in Watson-Crick base-pairing. In order for a double-stranded probe to invade the buried Watson-Crick face of a dsDNA target, both probe strands must have substantially higher affinity toward their targets than they have towards themselves. Introduction of pseudocomplementary base pairs is one approach to achieve a favorable energetic gradient for recognition of dsDNA. This is accomplished through the use of modified base pairs such as 2-thiothymine and 2-aminoadenine, which are destabilized due to steric interactions between the sulfur atom and the additional exocyclic amino group, but maintain adequate affinities towards canonical nucleotide binding partners.<sup>21</sup> This concept has been used with PNA backbones, and these pseudocomplementary PNA were shown to recognize internal regions of mixed-sequence dsDNA.<sup>22-24</sup> Although the requirement of low salt conditions remains in order for a stable recognition complex to form, it may be partially overcome under conditions that mimic molecular crowding in the nucleus.<sup>25</sup>

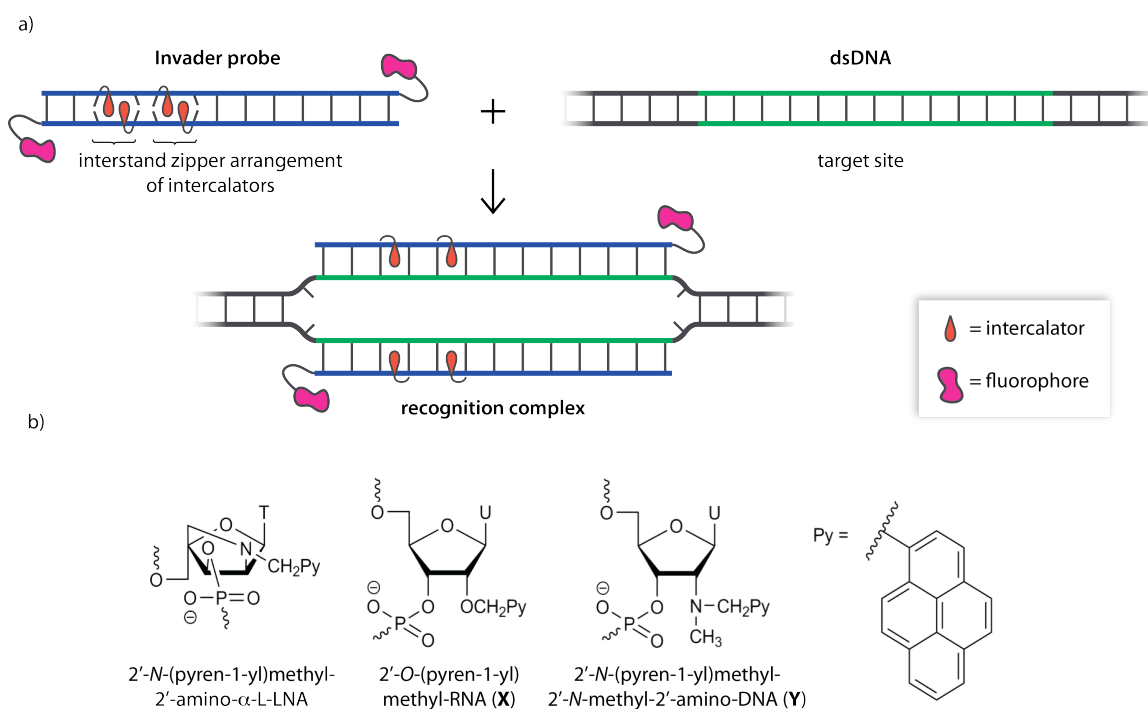
We have recently introduced a fundamentally different strategy for recognition of dsDNA, which is based on double-stranded oligodeoxyribonucleotide (ON) probes that are energetically activated through modification with +1 interstrand zippers of intercalator-functionalized nucleotides (Figure 2.1-1; for a description of the zipper nomenclature, see section 2.5).<sup>26,27</sup> This particular motif forces the intercalators into the same region of the probe duplex resulting in unwinding and destabilization, as the nearest neighbor exclusion principle<sup>28</sup> is violated, which is why we have coined the term *energetic hotspot* to describe this motif. According to this principle, the two sites

neighboring a bound intercalator will remain unoccupied due to limitations in local helix expandability (every intercalation event unwinds the duplex by  $\sim 3.4 \text{ \AA}$ )<sup>29</sup> and/or to avoid disruption of highly stable stacking interactions between nucleobases and the first bound intercalator.<sup>30</sup> In contrast, each of the two strands comprising the energetically activated probes, display very high affinity toward cDNA, since duplex formation is accompanied by strongly stabilizing stacking interactions between intercalators and nucleobases (Figure 2.1-1). The energy difference between the reactants (i.e., the double-stranded probe and DNA target) and products (i.e., the two probe-target duplexes formed as part of the recognition complex), provides the driving force for dsDNA-recognition (Figure 2.1-1).

We initially used 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA monomers as the key activating components of these *Invader probes*, but recently discovered that ONs modified with the simpler 2'-*O*-(pyren-1-yl)methyl-RNA and 2'-*N*-(pyren-1-yl)methyl-2'-*N*-methyl-2'-amino-DNA monomers display very similar hybridization properties (Figure 2.1-1).<sup>27</sup> We have utilized *Invader probes* based on 2'-*O*-(pyren-1-yl)methyl-RNA monomers for diagnostic proof-of-concept applications. For example, we have developed a colorimetric sandwich assay based on *Invader capture/signaling probes* for recognition of 28-mer mixed-sequence dsDNA fragments specific to important food pathogens. Targets are detected at concentrations down to 20 pM with excellent binding specificity.<sup>31</sup> In another study, we demonstrated that *Invader probes* can detect chromosomal DNA target regions in fixed interphase or metaphase nuclei under non-denaturing conditions.<sup>32</sup>

In the present study, *Invader probes* with different architectures – varying in the position, number and distance between energetic hotspots – based on either 2'-*O*-(pyren-1-

yl)methyluridine monomer **X** or 2'-*N*-(pyren-1-yl)methyl-2'-*N*-methyl-2'-aminouridine monomer **Y**, are characterized with respect to denaturation and thermodynamic properties, and dsDNA recognition efficiency, kinetics, and specificity. Informed by these insights, optimized Invader probes were used in non-denaturing fluorescence in situ hybridization (nd-FISH) experiments for detection of gender-specific chromosomal DNA target regions at near physiological conditions.



**Figure 2.1-1.** (a) Illustration of the Invader approach for recognition of dsDNA. (b) Structure of Invader monomers discussed herein.

## 2.2 Results and discussion

### *Thermal denaturation properties of Invader probes*

A library of twenty different 13-mer **X**- or **Y**-modified Invader probes was synthesized, in which the position, number, and distance between energetic hotspots were systematically varied (Table 2.2-1). Thermal denaturation temperatures ( $T_m$ 's) were



determined for each component of the Invader-mediated dsDNA-recognition process, i.e., the double-stranded Invader probe, the corresponding dsDNA target region, and the two probe-target duplexes (i.e., 5'-Inv:cDNA and 3'-Inv:cDNA). The  $T_m$ 's were used to calculate the *thermal advantage* [ $TA = T_m(5\text{'-Inv:cDNA}) + T_m(3\text{'-Inv:cDNA}) - T_m(\text{Invader}) - T_m(\text{dsDNA target region})$ ], which 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 thermodynamic dsDNA recognition potential.

Singly modified ONs display greatly increased affinity towards cDNA relative to unmodified ONs, with **Y**-modified ONs forming slightly more stable duplexes ( $\Delta T_m = 7.0 - 11.0$  °C vs  $8.0 - 13.5$  °C, for **X1-X8** and **Y1-Y8**, respectively, Table 2.2-1). Incorporation of a second pyrene-functionalized nucleotide in the same strand results in further stabilization ( $\Delta T_m$  for **M9-M18** =  $14.0-21.5$  °C). However, the  $T_m$  increases are less than additive, suggesting that intercalation of the first pyrene moiety negatively impacts the energetics of the second intercalation event (e.g., compare  $\Delta T_m$  of **M1**:cDNA and **M3**:cDNA relative to **M9**:cDNA).

Invader duplexes with one +1 interstrand zipper arrangement – or energetic hotspot – of **X** or **Y** monomers display low  $T_m$ 's ( $\Delta T_m = -1.0$  to  $+3.0$  °C) and are, accordingly, activated for recognition of dsDNA targets ( $TA \gg 0$  °C for **M1:M2-M7:M8**, Table 2.2-1). In accordance with previous results,<sup>27</sup> double-stranded probes with other interstrand zipper arrangements of **X**- or **Y**-monomers are not activated for dsDNA-recognition (compare  $TA$  values, Table 2.2-1 and Tables 2.5-3 to 2.5-5). This is because the intercalators only are forced to occupy the same region – leading to violation of the

nearest neighbor principle – when the corresponding monomers are placed in +1 interstrand zipper arrangements.<sup>33</sup> DNA duplexes with two energetic hotspots are moderately stabilized ( $\Delta T_m$  for **M9:M10** - **M17:M18** = 1.0 – 11.5 °C), with higher  $T_m$ 's being observed for **X**-modified probes and probes with two non-consecutive energetic hotspots. These trends mirror our results for Invader probes modified with 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA.<sup>26</sup> All of the double hotspot Invader probes are strongly activated for dsDNA-recognition due to the very high cDNA affinity of the individual strands, with **Y**-modified probes generally being more strongly activated ( $TA$  for **M9:M10** - **M17:M18** = 21.5 – 35.0 °C). The results with **X19:X20** and **Y19:Y20**, having four consecutive intercalator zippers, underscore the above conclusions.

**Table 2.2-1.** Thermal denaturation temperatures ( $T_m$ 's) and thermal advantages ( $TA$ 's) of X- and Y-modified DNA duplexes.<sup>a</sup>

ON	Sequence	M = Monomer X			Monomer Y		
		$T_m$ [ $\Delta T_m$ ] (°C)			$T_m$ [ $\Delta T_m$ ] (°C)		
		Probe duplex	5'-Inv:cDNA 3'-Inv:cDNA	TA (°C)	Probe duplex	5'-Inv:cDNA 3'-Inv:cDNA	TA (°C)
<b>M1</b>	5'-GG <b>M</b> ATATATAGGC	36.5	44.5 [+7.0]	+18.0	33.5	45.5 [+8.0]	+22.0
<b>M2</b>	3'-CC <b>A</b> MATATATCCG	[-1.0]	47.5 [+10.0]		[-4.0]	47.5 [+10.0]	
<b>M3</b>	5'-GGT <b>A</b> MATATAGGC	36.5	47.5 [+10.0]	+22.0	40.5	48.5 [+11.0]	+21.5
<b>M4</b>	3'-CC <b>A</b> T <b>A</b> MATATCC G	[-1.0]	48.5 [+11.0]		[+3.0]	51.0 [+13.5]	
<b>M5</b>	5'-GGT <b>A</b> T <b>A</b> MATAGGC	36.5	48.5 [+11.0]	+22.0	38.5	49.5 [+12.0]	+22.5
<b>M6</b>	3'-CC <b>A</b> T <b>A</b> T <b>A</b> MATCCG	[-1.0]	47.5 [+10.0]		[+1.0]	49.0 [+11.5]	
<b>M7</b>	5'-GGT <b>A</b> T <b>A</b> T <b>A</b> MAGGC	35.5	47.5 [+10.0]	+21.0	36.5	48.0 [+10.5]	+22.0
<b>M8</b>	3'-CC <b>A</b> T <b>A</b> T <b>A</b> T <b>A</b> MCCG	[-2.0]	46.5 [+9.0]		[-1.0]	48.0 [+10.5]	
<b>M9</b>	5'-GG <b>M</b> <b>A</b> MATATAGGC	40.0	51.5 [+14.0]	+29.5	43.5	51.5 [+14.0]	+25.0
<b>M10</b>	3'-CC <b>A</b> <b>M</b> <b>A</b> MATATCCG	[+2.5]	55.5 [+18.0]		[+6.0]	54.5 [+17.0]	
<b>M11</b>	5'-GG <b>M</b> <b>A</b> T <b>A</b> MATAGGC	49.0	53.5 [+16.0]	+23.5	48.0	56.0 [+18.5]	+29.0
<b>M12</b>	3'-CC <b>A</b> <b>M</b> <b>A</b> T <b>A</b> MATCCG	[+11.5]	56.5 [+19.0]		[+10.5]	58.5 [+21.0]	
<b>M13</b>	5'-GG <b>M</b> <b>A</b> T <b>A</b> T <b>A</b> MAGGC	49.0	52.5 [+15.0]	+21.5	45.0	55.5 [+18.0]	+32.0
<b>M14</b>	3'-CC <b>A</b> <b>M</b> <b>A</b> T <b>A</b> T <b>A</b> MCCG	[+11.5]	55.5 [+18.0]		[+7.5]	59.0 [+21.5]	
<b>M15</b>	5'-GGT <b>A</b> <b>M</b> <b>A</b> MATAGGC	45.0	55.5 [+18.0]	+28.5	38.5	55.5 [+18.0]	+35.0
<b>M16</b>	3'-CC <b>A</b> T <b>A</b> <b>M</b> <b>A</b> MATCCG	[+7.5]	55.5 [+18.0]		[+1.0]	55.5 [+18.0]	
<b>M17</b>	5'-GGT <b>A</b> <b>M</b> <b>A</b> <b>M</b> AGGC	47.5	54.5 [+17.0]	+24.0	46.5	56.0 [+18.5]	+28.0
<b>M18</b>	3'-CC <b>A</b> T <b>A</b> T <b>A</b> <b>M</b> <b>A</b> MCCG	[+10.0]	54.5 [+17.0]		[+9.0]	56.0 [+18.5]	
<b>M19</b>	5'-GG <b>M</b> <b>A</b> <b>M</b> <b>A</b> <b>M</b> AGGC	50.0	65.5 [+28.0]	+45.5	39.5	66.5 [+29.0]	+56.5
<b>M20</b>	3'-CC <b>A</b> <b>M</b> <b>A</b> <b>M</b> <b>A</b> <b>M</b> MCCG	[+12.5]	67.5 [+30.0]		[+2.0]	67.0 [+29.5]	

<sup>a</sup>  $\Delta T_m$  = change in  $T_m$  relative to unmodified dsDNA ( $T_m = 37.5$  °C); thermal denaturation curves were recorded in medium salt buffer ( $[Na^+] = 110$  mM,  $[Cl^-] = 100$  mM, pH 7.0 ( $NaH_2PO_4/Na_2HPO_4$ ),  $[EDTA] = 0.2$  mM) and  $[ON] = 1.0$   $\mu$ M; see main text for definition of  $TA$ . A = adenin-9-yl DNA monomer, C = cytosin-1-yl DNA monomer, G = guanin-9-yl DNA monomer, T = thymin-1-yl DNA monomer

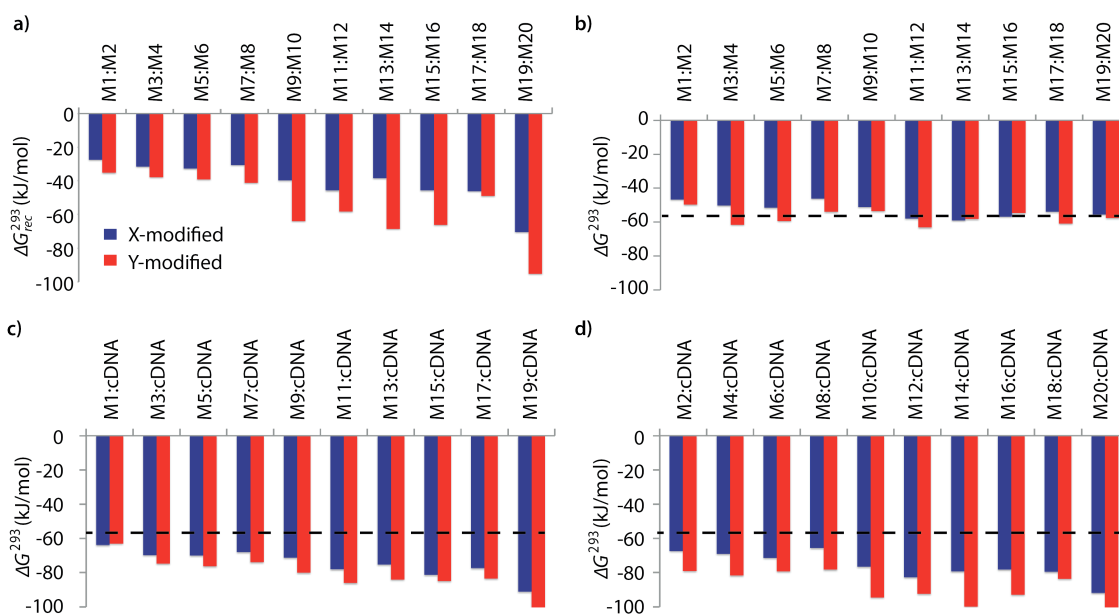
#### *Thermodynamic parameters for duplex formation*

The available free energy for the prototypical dsDNA recognition process can also be

parameterized as  $\Delta G_{rec}^{293} = \Delta G^{293}(5'-Inv:cDNA) + \Delta G^{293}(3'-Inv:cDNA) - \Delta G^{293}(\text{Invader}) -$

$\Delta G^{293}$  (dsDNA target region). Thermodynamic parameters for duplex formation were obtained from thermal denaturation curves via the van't Hoff method (Tables 2.5-6 to 2.5-11). Consistent with the  $T_m$ -based conclusions, Invader probes are strongly activated for dsDNA-recognition (i.e.,  $\Delta G_{rec}^{293} \ll 0$  kJ/mol, Figure 2.2-2a) due to the low stability of the probe duplexes (i.e.,  $\Delta \Delta G^{293}$  between -6 and +11 kJ/mol, Figure 2.2-2b) and the high stability of the probe-target duplexes (i.e.,  $\Delta \Delta G^{293}$  between -52 and -6 kJ/mol, Figure 2.2-2c and 2.2-2d). Recognition of dsDNA is very strongly enthalpically favored ( $\Delta H_{rec} \ll 0$  kJ/mol, Tables 2.5-83-11 and 2.5-9 and Figure 2.5-2), further underlining that forced intercalation is the main driving force (stabilizing in probe-target duplexes and destabilizing in Invader probes).

Invader probes with multiple energetic hotspots display more favorable dsDNA-recognition thermodynamics than single hotspot probes (compare  $\Delta G_{rec}^{293}$  for **M1:M2-M7:M8** vs **M9:M10-M19:M20**, Figure 2.2-2a), due to the exceptionally high cDNA affinity of the individual strands (note the highly negative  $\Delta G^{293}$  values for duplexes between **M9-M20** and cDNA, Figure 2.2-2c and 2.2-2d). **Y**-modified Invader probes are more strongly activated for dsDNA-recognition than corresponding **X**-modified probes ( $\Delta G_{rec}^{293}$  more favorable by 3-31 kJ/mol, Figure 2.2-2a), due to the higher stability of **Y**-modified probe-target duplexes (compare blue and red bars in Figures 2.2-2c and 2.2-2d). Interestingly, **X**-/**Y**-modified Invader probes are more strongly activated for dsDNA recognition than isosequential probes based on the original 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA monomers ( $\Delta G_{rec}^{293}$  more favorable by 1-29 kJ/mol).<sup>26</sup>



**Figure 2.2-2.** (a) Available Gibbs free energy at 293K ( $\Delta G_{rec}^{293}$ ) for Invader-mediated recognition of isosequential dsDNA targets, and (b-d) change in Gibbs free energy upon formation of X- or Y-modified DNA duplexes. The  $\Delta G^{293}$  for the dsDNA reference is shown as a dotted line at -57 kJ/mol. See Tables 2.5-6 and 2.5-7 for tabulated data.

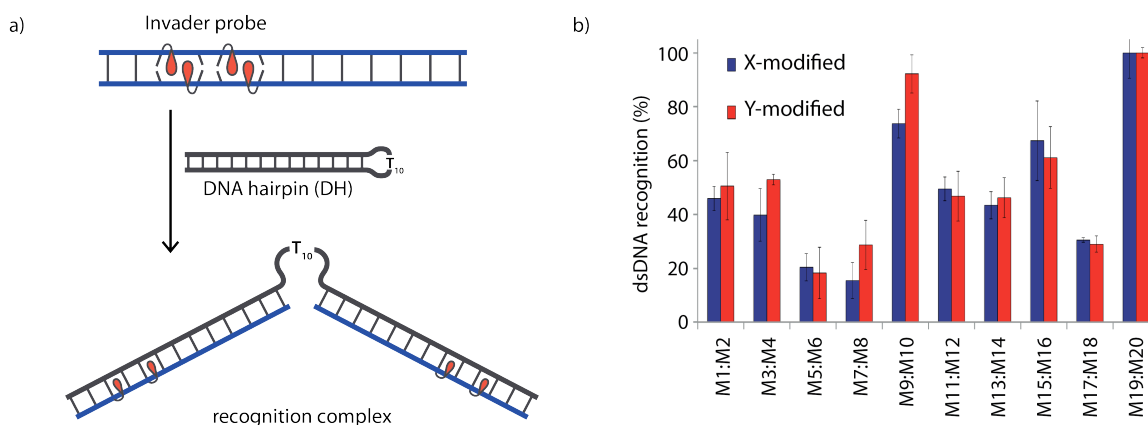
### *Recognition of model dsDNA targets*

The dsDNA recognition characteristics of the Invader probes were evaluated using an electrophoretic mobility shift assay (EMSA) that we developed in our preliminary studies (Figure 2.2-3a).<sup>32</sup> A 3'-digoxigenin (DIG) labeled DNA hairpin (DH), comprised of a 13-mer isosequential double-stranded stem that is connected on one side by a T<sub>10</sub> linker, was used as a model dsDNA target. Successful recognition of the stem by an Invader probe is expected to result in the formation of a ternary complex with decreased mobility during non-denaturing polyacrylamide gel electrophoresis (nd-PAGE). All twenty Invader probes were screened at a concentration of 6.88  $\mu$ M (i.e., 200-fold molar excess with respect to **DH1**) to identify probe architectures and monomer chemistries that result in efficient dsDNA recognition (Figure 2.2-3b). All of the Invader probes result in

recognition of the mixed-sequence stem when incubated at room temperature for 17h (Figure 2.5-4). As expected, dsDNA recognition efficiency increases with more highly modified probes, with **M19:M20** resulting in virtually complete recognition of **DH1** (compare e.g. **M1:M2** < **M9:M10** < **M19:M20**, Figure 3b).

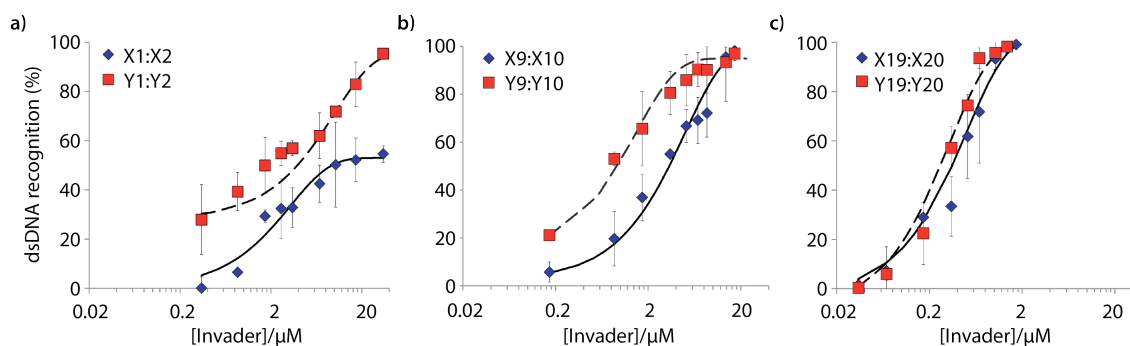
Closer inspection of the results reveals that single hotspot Invader probes **M1:M2** and **M3:M4**, in which the hotspot is located toward the ‘left’ terminus, recognize **DH1** more efficiently than **M5:M6** and **M7:M8**, in which the hotspot is located toward the ‘right’ terminus, despite having similar  $\Delta G_{\text{fold}}^{\text{dsDNA}}$  values (Figure 2.2-3b). To determine if these trends are due to fraying at the ‘left’ terminus of the target, **X1:X2–X7:X8** were incubated with **DH8** in which the ‘left’ side of the stem is connected via a  $T_{10}$  loop instead (Figure 2.5-5). Indeed, **DH8** is recognized more efficiently by **X5:X6** and **X7:X8** but the trend is not fully reversed, suggesting that additional factors, such as the higher GC-content at the ‘right’ end, also impact recognition efficiency. Along similar lines, Invader probe **X9:X10**, featuring two consecutive energetic hotspots near the ‘left’ terminus, results in more efficient recognition of **DH1** than **X17:X18** where two consecutive energetic hotspots are located near the ‘right’ terminus (Figure 2.2-3b). This trend is partially reversed when these Invaders are incubated with **DH8** (Figure 2.5-5). On the other hand, Invader probe **X15:X16**, featuring two central consecutive hotspots, recognizes **DH1** and **DH8** with similar efficiency, while Invader probes **M11:M12** and **M13:M14**, having two separated energetic hotspots, result in slightly less efficient dsDNA recognition (Figure 2.2-3b and Figure 2.5-5). In summary, these results suggest that Invader probes with multiple hotspots - irrespective of the substitution pattern -

enable recognition of dsDNA model targets, although targets with minimally fraying termini are more challenging.



**Figure 2.2-3.** (a) Illustration of the electrophoretic mobility shift assay (EMSA) used to evaluate dsDNA recognition. (b) dsDNA recognition by Invader probes. DIG-labeled **DH1** (34.4 nM) was incubated with 6.88  $\mu$ M of a pre-annealed Invader probe in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) for 17 h at room temperature. For representative electrophoretograms see Figure 2.5-4.

Six Invader probes – featuring one, two or four consecutive hotspots based on either monomer **X** or **Y** – were selected from this initial screen for more thorough characterization. Dose response experiments were performed to determine  $C_{50}$  values, i.e., the probe concentration that results in 50% recognition of **DH1** (Figure 2.2-4 and Table 2.2-2). Increasing the number of energetic hotspots progressively decreases the  $C_{50}$  values from single digit micromolar to submicromolar ranges. Probes based on 2'-*N*-(pyren-1-yl)methyl-2'-*N*-methyl-2'-aminouridine monomer **Y** display lower  $C_{50}$  values than probes based on 2'-*O*-(pyren-1-yl)methyluridine monomer **X**, which is in line with the observed  $\Delta G_{\text{bind}}^{\text{293}}$  values.



**Figure 2.2-4.** Dose-response curves for recognition of **DH1** using a) **M1:M2**, b) **M9:M10**, or c) **M19:M20**. For experimental conditions, see Figure 2.2-3.

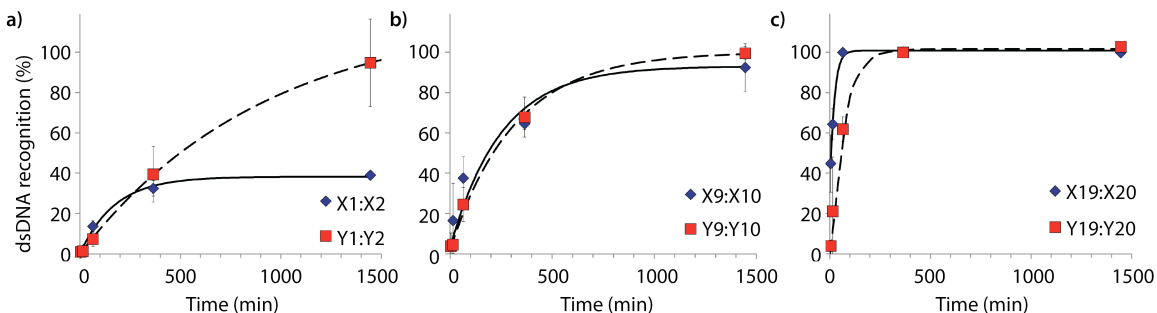
Recognition of **DH1** using individual probe strands was also examined. Incubation of **DH1** with 1000-fold molar excess of **M1** or **M2** or 500-fold molar excess of **M9** or **M10** results only in trace formation of recognition complexes, demonstrating that both strands of an Invader probe normally are necessary for efficient dsDNA-recognition (Figure 2.5-7). However, the use of 50-fold molar excess of **X19**, **X20** or **Y20** results in complete recognition of **DH1**. Clearly, the cDNA affinity of these strands is able to overcome the entropic penalty associated with opening the Watson-Crick base-pairs of the stem region and leaving one hairpin arm unhybridized.

#### *Kinetics of dsDNA-recognition*

Time-course experiments were performed, in which Invader probes (200-fold molar excess) were incubated with model target **DH1** and quenched at specific time-points to elucidate dsDNA-recognition kinetics (Figure 2.2-5 and Table 2.2-2). Recognition of **DH1** proceeds incrementally faster using more highly modified Invader probes, with 50% recognition ( $t_{50}$ ) being attained within  $\sim 3$ h with double hotspot probe **M9:M10** and within 10-50 min with quadruple hotspot probe **M19:M20**. It is noteworthy that **X**-modified probes have faster recognition kinetics than the corresponding **Y**-modified probes despite

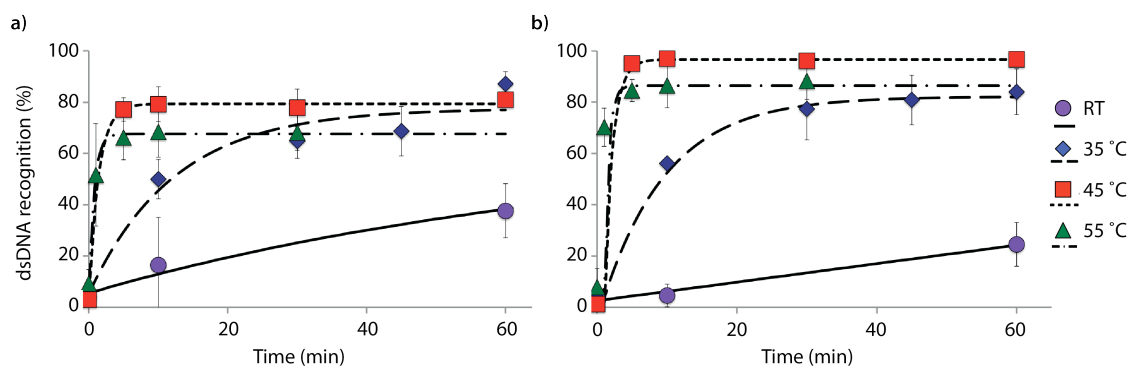


less favorable  $C_{50}$  values. Similar conclusions are reached based on pseudo-first order rate constants (Figure 2.5-6, Table 2.2-2).



**Figure 2.2-5.** Kinetic profile of **DH1** recognition using 200-fold molar excess of a) **M1:M2**, b) **M9:M10**, or c) **M19:M20**. For incubation conditions, see Figure 2.2-3. Aliquots were taken at specific time points, flash frozen in liquid  $N_2$ , and stored at  $-76\text{ }^\circ\text{C}$  until analysis.

The reaction kinetics are highly dependent on the incubation temperature (Figure 2.2-6). Thus, incubation of DNA hairpin **DH1** with probes **X9:X10** or **Y9:Y10** at  $\sim 8\text{ }^\circ\text{C}$  fails to result in any dsDNA recognition (data not shown), while incubation at  $35\text{ }^\circ\text{C}$  or  $45\text{ }^\circ\text{C}$  results in major rate enhancements relative to room temperature incubation (e.g.,  $\sim 20\%$ ,  $50\%$  and  $80\%$  recognition after 10 min using **X9:X10** at room temperature,  $35\text{ }^\circ\text{C}$  and  $45\text{ }^\circ\text{C}$ , respectively). The rate enhancements are, most likely, due to increased denaturation of the probes ( $T_m$ 's of **X9:X10** or **Y9:Y10** =  $40.0$  and  $43.5\text{ }^\circ\text{C}$ , respectively, Table 2.2-1) rather than denaturation of **DH1** ( $T_m = 58.5\text{ }^\circ\text{C}$ , Table 2.5-12). Incubation at  $55\text{ }^\circ\text{C}$  results in fast, but less pronounced, product formation, presumably because the recognition



**Figure 2.2-6.** Temperature dependence of dsDNA recognition kinetics using a) **X9:X10** or b) **Y9:Y10** having two consecutive energetic hotspots. Experiments were performed as described in Figure 2.2-3 with the exception of different incubation temperatures.

complex is partially denatured at this temperature ( $T_m$  of **X9/X10/Y9/Y10** vs cDNA = 51.5 – 55.5 °C, Table 2.2-1). These results demonstrate that dsDNA recognition can be accelerated by increasing experimental temperatures to 5 - 10 °C below the  $T_m$  of the probe-target recognition duplexes, when practically possible.

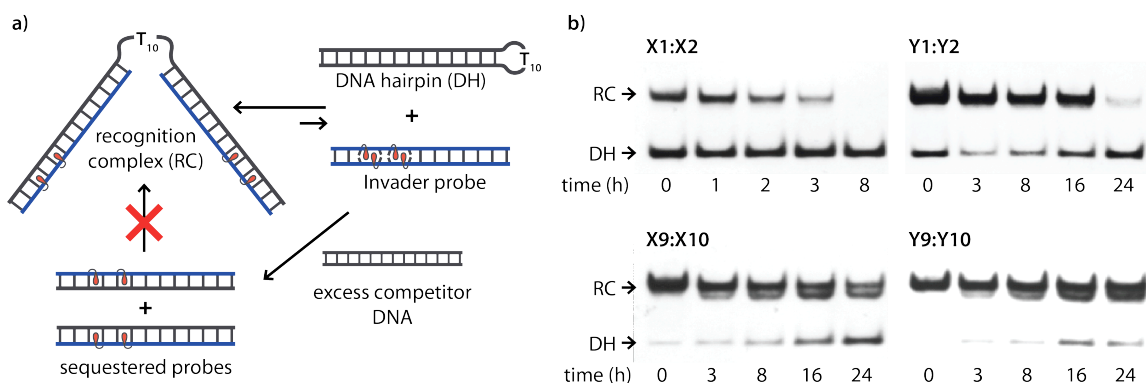
**Table 2.2-2.** Summary of dsDNA-recognition characteristics of selected Invader probes.<sup>a</sup>

Invader probe	X-modified Invader					Y-modified Invader				
	Rec <sub>200</sub> (%)	$C_{50}$ ( $\mu$ M)	$t_{50}$ (min)	$k_{obs}$ ( $\text{min}^{-1}$ )	$k_{rel}$	Rec <sub>200</sub> (%)	$C_{50}$ ( $\mu$ M)	$t_{50}$ (min)	$k_{obs}$ ( $\text{min}^{-1}$ )	$k_{rel}$
<b>M1</b> <b>M2</b>	46	9.4	ND	$1.1 \times 10^{-3}$	1	51	3.9	498	$1.4 \times 10^{-3}$	1.3
<b>M9</b> <b>M10</b>	74	3.4	185	$7.5 \times 10^{-3}$	6.8	92	0.9	211	$4.3 \times 10^{-3}$	3.9
<b>M19</b> <b>M20</b>	>95	0.4	7	$9.7 \times 10^{-2}$	88	>95	0.3	47	$1.5 \times 10^{-2}$	13.5

<sup>a</sup> Rec<sub>200</sub> denotes degree of recognition when using Invader at a 200-fold molar excess.  $C_{50}$ ,  $t_{50}$  and  $k_{obs}$  values obtained from Figures 2.2-4, 2.2-5, and 2.5-6, respectively.  $k_{rel}$  are calculated relative to the pseudo-first order rate constant for **X1:X2**.

*Stability of recognition complexes*

The kinetic stability of the recognition complexes with **DH1** was studied next. The recognition complex was allowed to form and was then incubated with a large excess of linear competitor dsDNA target that rapidly sequesters any dissociating Invader strands,<sup>26</sup> preventing their re-association with **DH1** (Figure 2.2-7a). Indeed gradual disappearance of the recognition complex is observed with time (Figure 2.2-7b). The recognition complex between **DH1** and **X1:X2** dissociates within 8 h, while the complex between **DH1** and **Y1:Y2** requires ~24 h for complete dissociation. In contrast, the recognition complexes between **DH1** and **M9:M10** are very stable, as evidenced by the small amounts of **DH1** formed after 24 h (~60% and ~85% of the recognition complexes with **X9:X10** and **Y9:Y10**, respectively, remaining). It is noteworthy that the collapse of the recognition complexes between **DH1** and **M9:M10**, in all likelihood, proceeds via a binary complex in which only one of the two probe-target recognition duplexes has dehybridized (notice the band immediately below the recognition complex band, Figure 2.2-7b). Thus, although single-stranded **M9** or **M10** cannot overcome the activation energy of the recognition process and invade the stem of **DH1** (Figure 2.5-9), it appears that an Invader strand can remain bound to the hairpin after dissociation of the other strand.

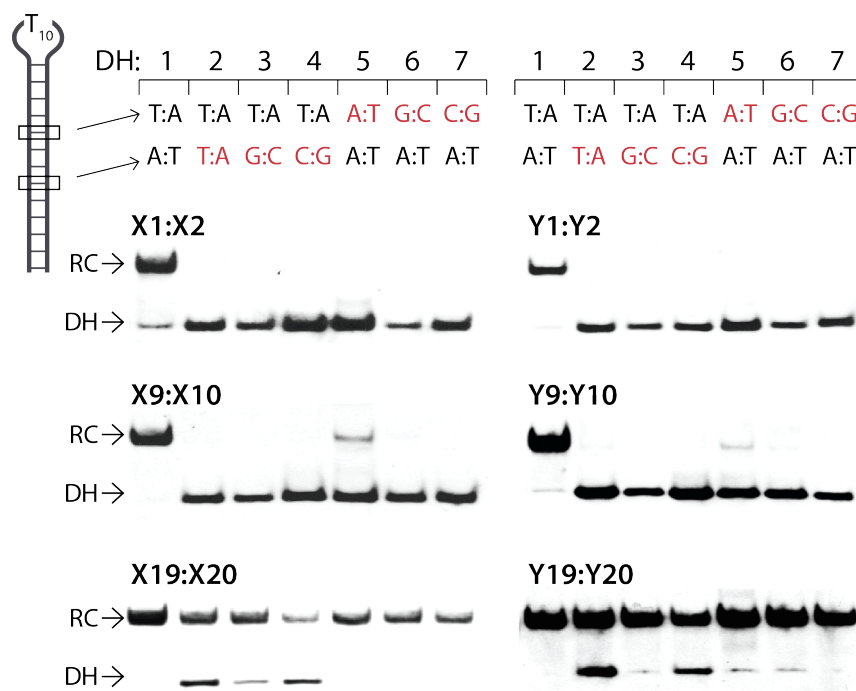


**Figure 2.2-7.** Dissociation kinetics of recognition complexes between DNA hairpins and Invader probes. (a) Illustration of competition assay. (b) Representative gel electrophoretograms for dissociation reactions. Recognition complexes were formed (incubation of 34.4 nM **DH1** with 200-fold molar excess of **M1:M2** or **M9:M10** for 17 h at room temperature), followed by addition of 2,000-fold molar excess of 5'-GGTATATATAGGC:3'-CCATATATATCCG. Incubation conditions are as described in Figure 2.2-3. Reactions were quenched at specific time points as described in Figure 2.2-5.

#### *Binding specificity of dsDNA recognition*

The above results show that it is possible to design energetically activated duplexes that enable efficient ( $C_{50} < 1 \mu\text{M}$ ), fast ( $t_{50} < 3 \text{ h}$ ), and kinetically stable ( $> 24 \text{ h}$ ) recognition of mixed-sequence dsDNA targets under physiologically relevant buffer conditions. To assess the specificity of the recognition process, the six selected Invader probes were incubated with DNA hairpins **DH2–DH7** (Figure 2.2-8), which have stem regions that differ in the nucleotide sequence relative to the probes at either the 6- or 8-position of the stem. Probes with one or two hotspots generally display excellent discrimination of the mismatched targets ( $< 10\%$  recognition of **DH5** with **M9:M10**) at conditions resulting in very efficient recognition of **DH1** using sequence-matched probes (1000- and 500-fold molar excess of **M1:M2** and **M9:M10**, respectively). In contrast, the very high dsDNA affinity of the constructs containing four energetic hotspots compromises the specificity

of the recognition process. Only **X19:X20** exhibits partial discrimination when the non-complementary base pairs are centrally positioned in the target region (55–80% recognition of **DH2-DH4**), which indicates that the initial nucleation site is distal to the loop region.



**Figure 2.2-8.** Discrimination of non-complementary DNA hairpins (**DH2-DH7**) using 1000-, 500- and 50-fold excess of **M1:M2**, **M9:M10** and **M19:M20**, respectively. For experimental conditions, see Figure 2.2-3.  $T_m$ 's of **DH1-DH7** are between 58.5–63.5 °C (Table 2.2-15).

### *Detection of chromosomal DNA*

Encouraged by these results, we set out to examine Invader probes based on 2'-*O*-(pyren-1-yl)methyl-RNA monomers as FISH probes for recognition of chromosomal DNA using non-denaturing conditions. Unlike conventional FISH assays, which require denaturation of chromosomal DNA by heat and/or formamide treatment,<sup>34</sup> nd-FISH approaches enable

mapping of chromosomal loci at mild conditions, thus offering the prospect of in vivo imaging. The majority of previously reported nd-FISH approaches are based either on classic dsDNA-targeting agents (i.e., TFOs, PNAs and polyamides)<sup>35-40</sup> or the presence of uniquely accessible DNA regions,<sup>41,42</sup> which has limited the widespread use of these approaches.

Four Cy3-labeled Invader probes, varying in length as well as number and position of energetic hotspots (**INV1-INV4**, probe lengths: 11-14 nt, 2-3 hotspots, Table 2.2-3), were designed against a target region within the *DYZ-1* satellite region ( $\sim 6 \times 10^4$  tandem repeats) of the bovine (*Bos taurus*) Y chromosome (NCBI code: M26067; target site: 562-575).<sup>43</sup> We have previously used PNA FISH probes targeting this site to determine the gender of bovine somatic cells, spermatozoa, and embryos.<sup>44-46</sup> However, single-stranded fluorophore-labeled PNAs fail to produce signals under non-denaturing conditions (Figure 2.5-9).

As expected, the designed Invader probes display low stability relative to isosequential dsDNA target regions, while the individual strands form very stable duplexes with cDNA, resulting in prominent dsDNA-targeting potential for the probes (see  $T_m$  and  $\Delta G_{\text{inv}}^{\text{ds}}$  values, Table 2.2-3 – for additional thermodynamic parameters, see Tables 2.5-13 to 2.5-15).

**Table 2.2-3.**  $T_m$  and  $\Delta G_{ref}^{310}$  values of Invader probes used in the nd-FISH study.<sup>a</sup>

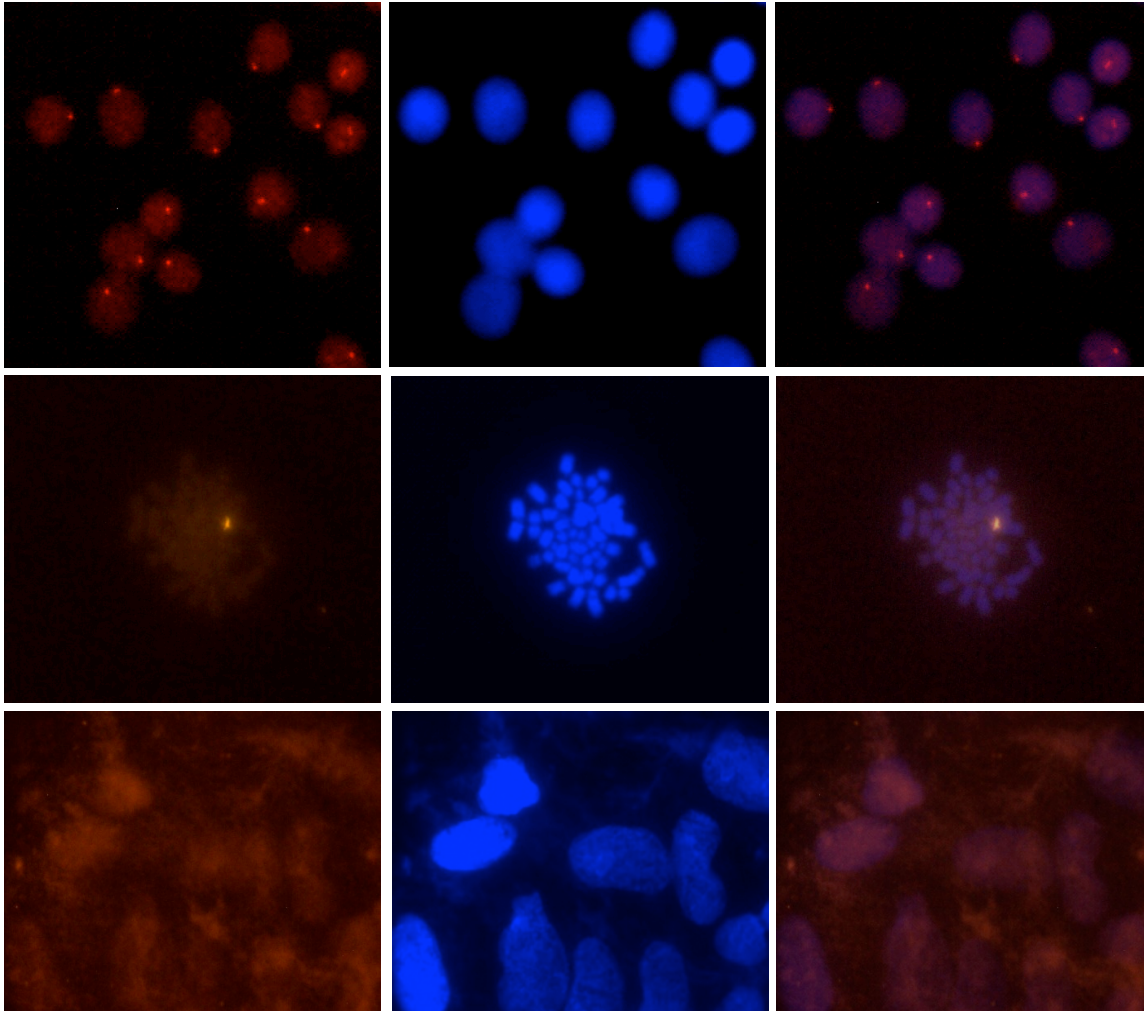
Invader Probe	Sequence	$T_m$ [ $\Delta T_m$ ] (°C)			$\Delta G_{ref}^{310}$ (kJ/mol)
		Probe duplex	5'-Inv: cDNA	3'-Inv: cDNA	
<b>INV1</b>	5'-Cy3-AGCC <u>CU</u> UGCCCTG 3'-TCGGGAC <u>CA</u> CGGGAC-Cy3	58.0 [-2.5]	69.5 [+9.0]	74.5 [+14.0]	60.5 -46
<b>INV2</b>	5'-Cy3-CC <u>U</u> UGCCCTG 3'-GGAC <u>CA</u> CGGGAC-Cy3	48.0 [-2.5]	59.5 [+9.0]	65.5 [+15.0]	50.5 -30
<b>INV3</b>	5'-Cy3-CC <u>U</u> GTGCCCTG 3'-GGAC <u>CA</u> CGGG <u>A</u> C-Cy3	47.0 [-3.5]	59.0 [+8.5]	64.0 [+13.5]	50.5 -24
<b>INV4</b>	5'-Cy3- <u>A</u> GCC <u>CU</u> GTGCCCTG 3'-TCGGGAC <u>CA</u> CGGG <u>A</u> C-Cy3	61.5 [+1.0]	69.5 [+9.0]	75.5 [+15.0]	60.5 -29

<sup>a</sup>  $\Delta T_m$  = change in  $T_m$  values relative to corresponding unmodified and unlabeled reference duplex. For experimental details see Table 2.2-1. **A**, **C** and **U** denote 2'-*O*-(pyren-1-yl)methyladenosine,<sup>47</sup> 2'-*O*-(pyren-1-yl)methylcytidine<sup>47</sup> and 2'-*O*-(pyren-1-yl)methyluridine (monomer **X**), respectively.

Gratifyingly, incubation of these Cy3-labeled Invader probes with fixed interphase nuclei from a male bovine kidney cell line (CCL-22, MDBK) at non-denaturing conditions (3h, 38.5 °C, 10 mM Tris-Cl, pH 8.0 and 1 mM EDTA), produces a single fluorescent signal that localizes to the heterochromatic region, consistent with the expected localization of the *DYZ-1* satellite target (Figure 2.2-9 upper panel and Figure 2.5-10). The high labeling coverage (~90%), i.e., the proportion of nuclei with localized signals, is noteworthy. Localized signals are also observed when Cy3-labeled Invader probes are incubated with nuclei captured in metaphase (Figure 2.2-9 middle panel and Figure 2.5-10). All of the probes result in robust Cy3-signals, with little variation between the different probes (Figure 2.5-10). Nuclei that were pre-treated with RNase A or proteinase K prior to incubation with Invader **INV4** display similar signals as nuclei without pretreatment, while nuclei pre-treated with DNase I are devoid of localized signals (Figure 2.5-11), which verifies that DNA is the molecular target of the Invader probes. As expected, there

is also an absence of localized signals when nuclei from a female bovine fibroblast cell line are incubated with a Y-chromosome specific Invader probe (Figure 2.2-9 lower panel). This, along with observations from our initial studies showing lack of signal formation when triply mismatched Invaders are incubated with nuclei from the male bovine kidney cell line,<sup>32</sup> strongly suggests that the Invader probes specifically bind to their intended targets.

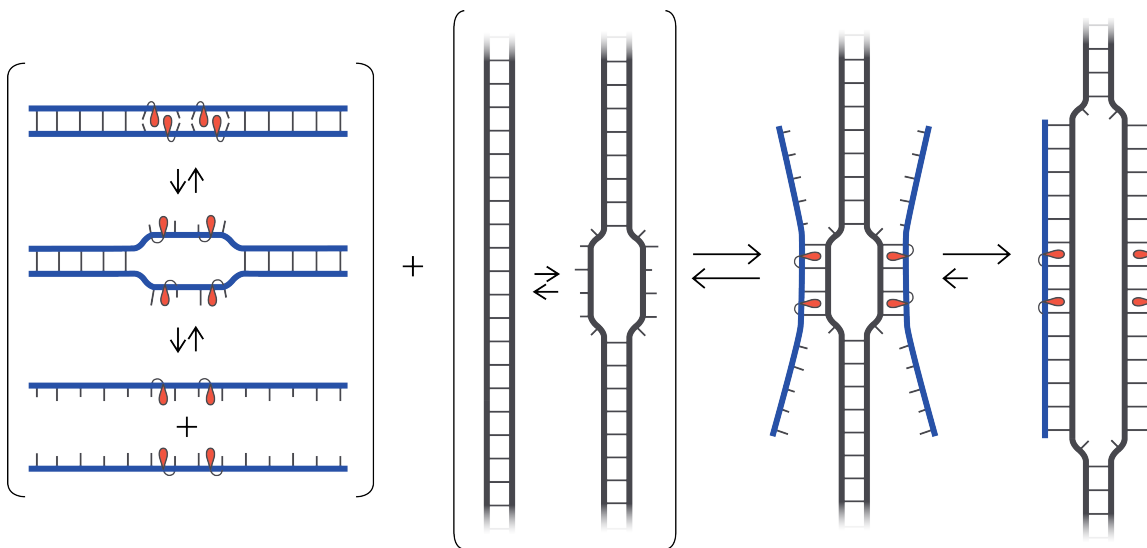




**Figure 2.2-9.** Images from fluorescence in situ hybridization using Y-chromosome specific Invader probes under non-denaturing conditions. Invader probe **INV4** was added to nuclei from male bovine kidney cells in interphase (upper panel) or metaphase (middle panel), or to nuclei from female bovine fibroblast cells (lower panel). Images viewed using Cy3 (left column) or DAPI (middle column) filter settings; overlays are shown in the right column. Incubation: 3 h at 38.5 °C in 10 mM Tris-Cl, pH 8.0 and 1 mM EDTA. Cells were visualized at 400x magnification using a Zeiss AxioSkop 40 fluorescence microscope and images captured using a Zeiss AxioCam MRc5 camera. For additional images using probes **INV1-INV4**, see Figure 2.5-10.

While our understanding of the recognition mechanism remains incomplete, we speculate that the distorted Invader probes trap regions of chromosomal DNA via a double duplex invasion mechanism analogous to that of pcPNAs (Figure 2.2-10).<sup>24</sup> Studies with  $\gamma$ -PNA<sup>19</sup> have suggested that DNA is sufficiently dynamic to enable strand invasion at 37 °C,

provided that the ligand has sufficient binding free energy. Presumably, nucleation of the Invader probes is initiated due to the exceptionally high cDNA affinity of the intercalator-functionalized nucleotides, and base-pairing proceeds until a stable double duplex invasion recognition complex is formed. Accordingly, many previously inaccessible mixed-sequence dsDNA target regions may become accessible to exogenous probes.



**Figure 2.2-10.** Illustration of hypothetical recognition mechanism.

### 2.3 Conclusions

This study demonstrates that Invader probes can be designed to display efficient ( $C_{50} < 1 \mu\text{M}$ ), fast ( $t_{50} < 3\text{h}$ ), kinetically stable ( $> 24\text{h}$ ) and single nucleotide specific recognition of mixed-sequence DNA targets. Probe duplexes with energetic hotspots comprised of +1 interstrand zippers of 2'-*O*-(pyren-1-yl)methyl-RNA or 2'-*N*-(pyren-1-yl)methyl-2'-*N*-methyl-2'-amino-DNA monomers are thermodynamically activated for DNA recognition, whereas probes with other zipper motifs are not, which underscores the unique properties of +1 intercalator zipper motifs. Recognition of DNA targets proceeds progressively

faster and with greater efficiency with every additional energetic hotspot that is incorporated into the Invader probes. These guidelines enabled the design of Invaders for successful detection of gender-specific mixed-sequence chromosomal DNA target regions in bovine kidney cells.

The insights gained from this study will enable the design of efficient Invader probes for DNA-targeting applications including gene regulation via transcriptional interference, in vivo imaging of chromosomal DNA targets, cell-sorting of genotype-specific cells, and development of artificial restriction ‘enzymes’.

## **2.4 Acknowledgements**

This study has enjoyed support from Award Number GM088697 from the National Institute of General Medical Sciences, National Institutes of Health; Awards IF13-001 and IF14-012 from the Higher Education Research Council, Idaho State Board of Education; The Office of Naval Research (N00014-10-1-0282); and INBRE Program, NIH grant no. P20 RR016454 (National Center for Research Resources) and P20 GM103408 (National Institute of General Medical Sciences). We thank Dr. Lee Deobald (Murdock Mass Spectrometry Center, Univ. Idaho) for assistance with mass spectrometric analysis and Dr. Carolyn J. Hovde (Food Science, Univ. Idaho) for access to gel documentation stations.

## **2.5 Supplementary data**

*Protocol - synthesis and purification of ONs*

**X/Y**-modified ONs were synthesized on an automated DNA synthesizer (0.2  $\mu$ mol scale) using a long chain alkyl amine controlled pore glass (LCAA-CPG) solid support with a pore size of 500 Å. The corresponding phosphoramidites of monomers **X** and **Y** were prepared as previously described<sup>48</sup> and incorporated into ONs via hand-couplings (0.05 M in acetonitrile, using 0.01 M 4,5-dicyanoimidazole or 5-ethylthio-1*H*-tetrazole as the activators (15 min) for A/C/**X**-modified or **Y**-modified ONs, respectively) with extended oxidation (45 s). Treatment with 32% ammonia (55 °C, 17 h) facilitated deprotection and cleavage from solid support. DMT-protected ONs were purified via ion-pair reverse phase HPLC (XTerra MS C18 column: 0.05 M triethyl ammonium acetate and acetonitrile gradient) followed by detritylation (80% acetic acid, 20 min) and precipitation (NaOAc, NaClO<sub>4</sub>, acetone, -18 °C, 16 h). The purity and identity of synthesized ONs were verified using analytical HPLC (>85% purity) and MALDI-MS analysis (Tables 2.5-1 and 2.5-2) recorded on a Quadrupole Time-of-Flight (Q-TOF) mass spectrometer with anthranilic acid or 3-hydroxypicolinic acid matrix for **X**- or **Y**-modified ONs, respectively.

Cy3-labeled **X**-modified ONs were synthesized as described above with the following modifications. After the incorporation of the last nucleotide, a C6-amino-modifier (Glen Research) was incorporated via hand coupling (4,5-dicyanoimidazole, 15 min, anhydrous CH<sub>3</sub>CN). The resulting ONs were worked up, purified, detritylated (glacial AcOH, 45 min, rt) and precipitated essentially as described above. The resulting amine-terminated ONs were dissolved in nanopure water, quantified, evaporated to dryness on a speedvac, dissolved in a minimum volume of water, and coupled with Cy3-*N*-hydroxysuccinimide ester (Lumiprobe, LLC) in DMSO as recommended by the vendor. Cy3-labeled ONs

were precipitated twice from ethanol (0 °C for ~3h) and purified by RP-HPLC as described above.

*Protocol - thermal denaturation experiments*

The concentrations of ONs were estimated using the following extinction coefficients ( $OD_{260}/\mu\text{mol}$ ): G (12.01), A (15.20), T (8.40), C (7.05), pyrene (22.4) and Cy3 (4.93).<sup>49</sup>

Thermal denaturation temperatures were calculated as the first-derivative maximum of the  $A_{260}$  vs.  $T$  curve. ONs (1.0  $\mu\text{M}$ ) were annealed (85 °C for 2 min) in medium salt buffer ( $[\text{Na}^+] = 110 \text{ mM}$ ,  $[\text{Cl}^-] = 100 \text{ mM}$ , pH 7.0 ( $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ ),  $[\text{EDTA}] = 0.2 \text{ mM}$ ) and subsequent cooling to the starting temperature. The experimental temperature ranged from at least 15 °C below  $T_m$  to 15 °C above  $T_m$ , with the  $T_m$  being determined as the average of two experiments within  $\pm 1.0$  °C.

*Protocol - determination of thermodynamic parameters*

Thermodynamic parameters for duplex formation were determined through baseline fitting of denaturation curves (van't Hoff method) using software with the UV-Vis spectrophotometer. Bimolecular reactions, two-state melting behavior, and constant heat capacity were assumed.<sup>50</sup> Two curves per experiment were analyzed at least three times to minimize errors arising from baseline choice.

*Protocol - electrophoretic mobility shift assay*

DNA hairpins (DH) were obtained from commercial sources and used without further purification. Hairpins were labeled using the 2<sup>nd</sup> generation DIG Gel Shift Kit (Roche Applied Bioscience). Briefly, 11-digoxigenin-ddUTP was incorporated at the 3'-end of

the hairpin (100 pmol) using a recombinant terminal transferase. The reaction mixture was quenched through addition of EDTA (0.05 M), diluted to 68.8 nM, and used without further processing. The recognition experiments were conducted essentially as previously reported.<sup>32</sup> Thus, Invader probes (variable concentration) were annealed (90 °C for 2 min, followed by cooling to room temperature) and subsequently incubated with DIG-labeled DNA hairpins (34.4 nM) in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) at room temperature (~21 °C) for a specified period. Loading dye (6X) was added and the reaction mixtures were loaded onto 12% non-denaturing TBE-PAGE (45 mM tris-borate, 1 mM EDTA; acrylamide:bisacrylamide (19:1)). Electrophoresis was performed using constant voltage (70 V) at ~4 °C for 1.5 h. Bands were blotted onto positively charged nylon membranes (100 V, 30 min, ~4 °C) and cross-linked through exposure to UV (254 nm, 5 × 15 watt bulbs, 3 min). Membranes were incubated with anti-digoxigenin-alkaline phosphatase F<sub>ab</sub> fragments as recommended by manufacturer, and transferred to a hybridization jacket. Membranes were incubated with the chemiluminescence substrate (CSPD) for 10 min at 37 °C, and chemiluminescence was captured on X-ray films. Digital images of developed X-ray films were obtained using a Fluor-S MultiImager and quantified using appropriate software (Quantity One). The percentage of dsDNA recognition was calculated 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 (±).

#### *Protocol - verification of gender via PCR*

The gender of the somatic cell lines was confirmed via PCR as previously reported,<sup>51</sup> using primers that specifically target bovine *ZFX* and *ZFY* gene sequences located on the

X and Y chromosomes, respectively. The presence of a single band is indicative of a female, while two bands signal a male (Figure S7). Multiplex PCR (MJ Research PTC-200 thermocycler) was performed using the following protocol: 95 °C pre-denaturation step (5 min), followed by 35 cycles of: denaturation at 95 °C (1 min), annealing at 60 °C (1 min) and extension at 72 °C (1 min). Following the last PCR cycle, additional extension at 72 °C (10 min) was performed to ensure complete elongation of any remaining single-stranded DNA. PCR reactions were performed in a total volume of 20 uL comprised of ~1 uL of genomic DNA (~100 ng), 1 uL of each primer (20 nM), and 0.5 uL of dNTPs (10 mM). Taq polymerase, 10X buffer and MgCl<sub>2</sub> was provided using the HotStar Taq DNA polymerase kit (Qiagen, Valencia, CA). The amplicons were resolved by gel electrophoresis on 2% agarose gels, which were run with a 2 kb molecular weight standard (Lo DNA Marker, Bionexus, Oakland CA). Images were captured using a gel imaging software system (Quantity One 1-D Analysis Software, Bio-Rad, Hercules, CA).

#### *Protocol - cell culture and nuclei preparation*

Male bovine kidney (MDBK, ATCC: CCL-22, Bethesda, MD) and female bovine fibroblast (Minitube, Verona, WI) cell lines were maintained in DMEM with GlutaMax (Gibco, 10569-010) and 10% fetal bovine serum (Invitrogen). The cell lines were cultured in separate 25 mL flasks at 38.5 °C in a 5% CO<sub>2</sub> atmosphere for 24-72 h to achieve confluent growth. Following this, the medium was treated with colcemid for 2h and then replaced with 0.25% trypsin-EDTA in DMEM. The cells were cultured for 20 min as stated above to detach adherent cells. The loosened cells were aspirated and transferred to 15 mL centrifuge tubes and centrifuged at 600x g for 4 min to pellet the

cells. The supernatant was aspirated off, replaced with 75 mM KCl (hypotonic buffer), and the cells incubated for 20 min at 37 °C. The cell suspension was centrifuged and the supernatant aspirated off. The pellet – containing somatic nuclei – was subsequently resuspended in methanol and glacial acetic acid (3:1, v/v) and stored at -20 °C until use.

*Protocol - fluorescence in situ hybridization*

An aliquot (3-5 uL) was taken from the abovementioned somatic nuclei suspension and dropped onto a plastic slide (Sex-Y™, Minitube, Verona, WI). The slide was briefly placed in an oven to fix nuclei and evaporate solvents (2 min, 60 °C) and subsequently cooled to room temperature. An aliquot (~250 uL) of the labeling buffer [~15 ng of a Cy3-labeled Invader probe in 500 uL TE buffer (10 mM Tris and 1 mM EDTA; pH 8.0)] was placed on top of the fixed nuclei. The slide was placed in a plastic culture dish, covered with a lid and moved to a 38.5 °C incubator for ~3h. Following incubation, the slide was rinsed 3 min in warm TE buffer (38.5 °C) and left to dry at room temperature. An aliquot (~3 uL) of Gold SlowFade plus DAPI (Invitrogen) was placed directly on the slide and a round coverslip was mounted for fluorescence imaging. A Zeiss AxioSkop 40 fluorescent microscope (50 W, HBO mercury lamp), equipped with Cy3 and DAPI filter sets, was used to visualize the nuclei at 400× magnification. Images of fluorescently labeled nuclei were captured using a Zeiss AxioCam MRc5 camera and processed with AxioVision (version 4.8) software. Excellent slide-to-slide reproducibility was observed using this protocol within a given batch of nuclei preparations.

Control experiments involving DNase/RNase/proteinase pre-treated MDBK nuclei included the following steps prior to incubation with Invader probes. DNase pre-



treatment: approximately 1  $\mu$ L DNase I (Sigma-Aldrich D7291) was diluted in 1X reaction buffer (20 mM Tris-HCl, 2 mM MgCl<sub>2</sub>, pH 8.3) and the solution was incubated with fixed nuclei for 20 min at 37 °C, followed by rinsing with TE buffer. RNase pre-treatment: approximately 1  $\mu$ L RNase A (Sigma-Aldrich R4642) was diluted in 100  $\mu$ L of 10 mM Tris-HCl (pH 6.5), and the solution incubated with fixed nuclei for 45 min at 37 °C, followed by rinsing with TE buffer. Proteinase pre-treatment: approximately 1  $\mu$ L of proteinase K (Sigma-Aldrich P6556) was diluted in 200  $\mu$ L of 10 mM Tris-HCl (pH 7.5) and the solution incubated with fixed nuclei for 1h at 37 °C, followed by rinsing with TE buffer.

*Definition - interstrand zipper arrangement*

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.

**Table 2.5-1.** MALDI-MS of ONs modified with 2'-*O*-(pyren-1-yl)methyl-RNA monomers.<sup>a</sup>

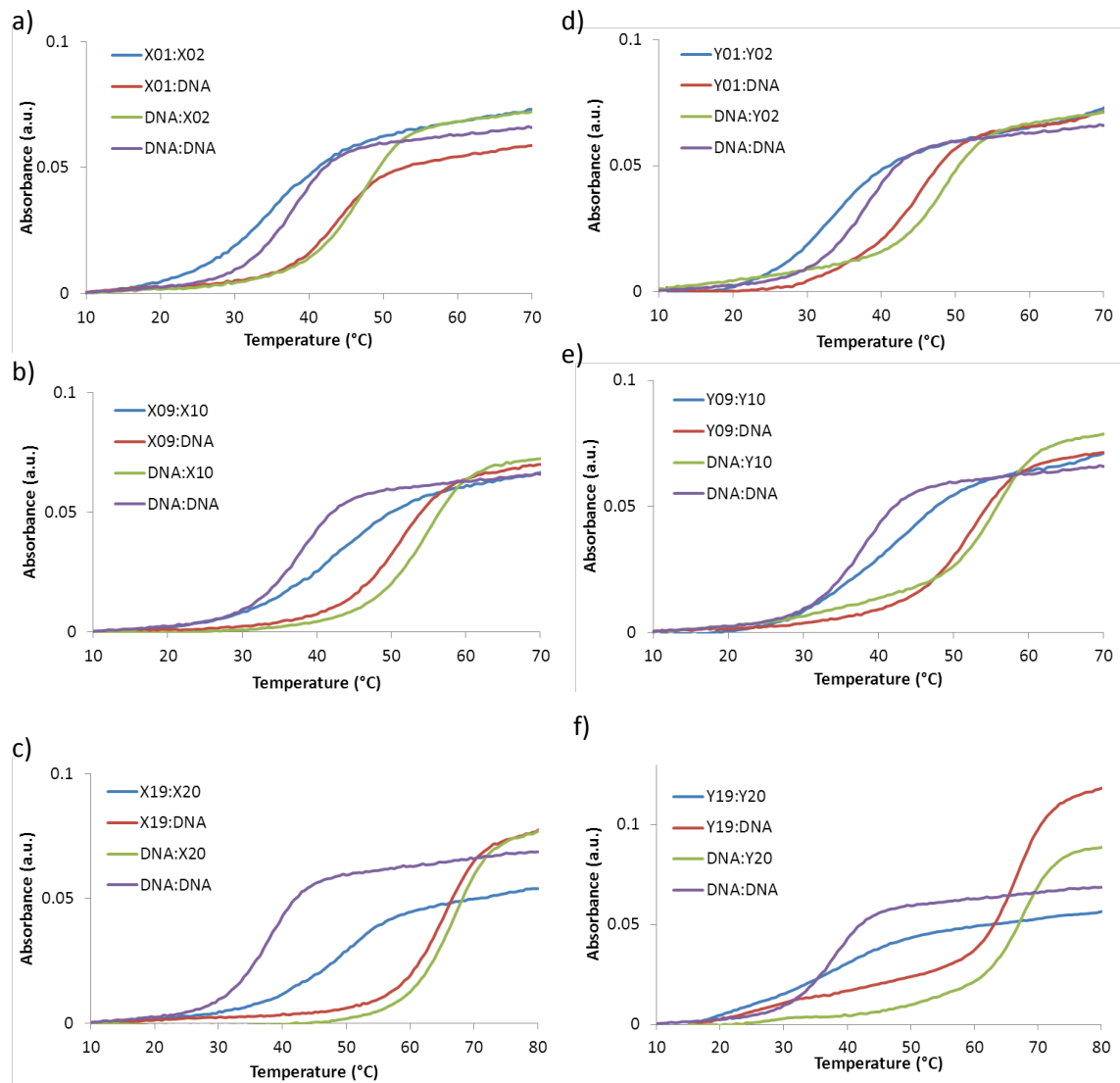
ON	Sequence	Observed <i>m/z</i> [M+H] <sup>+</sup>	Calculated <i>m/z</i> [M+H] <sup>+</sup>
X1*	5'-GG <u>X</u> ATA TAT AGG C	4228	4228
X2*	3'-CCA <u>X</u> AT ATA TCC G	4108	4108
X3	5'-GGT A <u>X</u> A TAT AGG C	4228	4228
X4	3'-CCA T <u>A</u> X ATA TCC G	4108	4108
X5*	5'-GGT ATA <u>X</u> AT AGG C	4228	4228
X6*	3'-CCA TAT A <u>X</u> A TCC G	4108	4108
X7*	5'-GGT ATA T <u>A</u> X AGG C	4228	4228
X8*	3'-CCA TAT ATA <u>X</u> CC G	4108	4108
X9*	5'-GG <u>X</u> A <u>X</u> A TAT AGG C	4444	4444
X10*	3'-CCA <u>X</u> A <u>X</u> ATA TCC G	4324	4324
X11	5'-GG <u>X</u> ATA <u>X</u> AT AGG C	4444	4444
X12	3'-CCA <u>X</u> AT A <u>X</u> A TCC G	4324	4324
X13	5'-GG <u>X</u> ATA T <u>A</u> X AGG C	4444	4444
X14	3'-CCA <u>X</u> AT ATA <u>X</u> CC G	4324	4324
X15	5'-GGT A <u>X</u> A <u>X</u> AT AGG C	4444	4444
X16	3'-CCA T <u>A</u> X A <u>X</u> A TCC G	4324	4324
X17	5'-GGT ATA <u>X</u> A <u>X</u> AGG C	4444	4444
X18	3'-CCA TAT A <u>X</u> A <u>X</u> CC G	4324	4324
X19	5'-GG <u>X</u> A <u>X</u> A <u>X</u> A <u>X</u> AGG C	4876	4876
X20	3'-CCA <u>X</u> A <u>X</u> A <u>X</u> A <u>X</u> CC G	4757	4756
A1	3'-CC <u>A</u> TAT ATA TCC G	4122	4122
A2	3'-CCA T <u>A</u> T ATA TCC G	4122	4122
A3	3'-CCA TAT <u>A</u> TA TCC G	4122	4122
A4	3'-CCA TAT AT <u>A</u> TCC G	4122	4122
A5	3'-CC <u>A</u> T <u>A</u> T ATA TCC G	4352	4352
A6	3'-CC <u>A</u> TAT <u>A</u> TA TCC G	4352	4352
A7	3'-CC <u>A</u> TAT AT <u>A</u> TCC G	4351	4352
A8	3'-CCA T <u>A</u> T <u>A</u> TA TCC G	4352	4352
A9	3'-CCA TAT <u>A</u> T <u>A</u> TCC G	4351	4352
A10	3'-CC <u>A</u> T <u>A</u> T <u>A</u> TA TCC G	4813	4813
INV1-u	5'-[Cy3]-AGCCC <u>U</u> GUGCCCTG	5265	5265
INV1-l	3'-TCGGGAC <u>A</u> CGGGAC-[Cy3]	5391	5391
INV2-u	5'-[Cy3]-CC <u>U</u> GUGCCCTG	4334	4334
INV2-l	3'-GGAC <u>A</u> CGGGAC-[Cy3]	4469	4469
INV3-u	5'-[Cy3]-CC <u>U</u> GTGCC <u>C</u> TG	4348	4348
INV3-l	3'-GGAC <u>A</u> CGGG <u>A</u> C-[Cy3]	4469	4469
INV4-u	5'-[Cy3]- <u>A</u> GCC <u>U</u> GTGCC <u>C</u> TG	5509	5509
INV4-l	3'-TCGGGAC <u>A</u> CGGG <u>A</u> C-[Cy3]	5621	5621

<sup>a</sup> **A**, **C** and **U** denote 2'-*O*-(pyren-1-yl)methyladenosine,<sup>47</sup> 2'-*O*-(pyren-1-yl)methylcytidine<sup>47</sup> and 2'-*O*-(pyren-1-yl)methyluridine (monomer **X**), respectively. Cy3 = cyanine 3. \* Previously published in reference 32.

**Table 2.5-2.** MALDI-MS of ONs modified with 2-*N*-(pyren-1-ylmethyl)-2'-*N*-methylamino-uridine monomers.

ON	Sequence	Observed <i>m/z</i> [M+H] <sup>+</sup>	Calculated <i>m/z</i> [M+H] <sup>+</sup>
<b>Y01</b>	5'-GGY ATA TAT AGG C	4243	4242
<b>Y02</b>	3'-CCA YAT ATA TCC G	4123	4122
<b>Y03</b>	5'-GGT AYA TAT AGG C	4240	4242
<b>Y04</b>	3'-CCA TAY ATA TCC G	4121	4122
<b>Y05</b>	5'-GGT ATA YAT AGG C	4242	4242
<b>Y06</b>	3'-CCA TAT AYA TCC G	4122	4122
<b>Y07</b>	5'-GGT ATA TAY AGG C	4241	4242
<b>Y08</b>	3'-CCA TAT ATA YCC G	4121	4122
<b>Y09</b>	5'-GGY AYA TAT AGG C	4472	4471
<b>Y10</b>	3'-CCA YAY ATA TCC G	4352	4351
<b>Y11</b>	5'-GGY ATA YAT AGG C	4471	4471
<b>Y12</b>	3'-CCA YAT AYA TCC G	4351	4351
<b>Y13</b>	5'-GGY ATA TAY AGG C	4471	4471
<b>Y14</b>	3'-CCA YAT ATA YCC G	4351	4351
<b>Y15</b>	5'-GGT AYA YAT AGG C	4472	4471
<b>Y16</b>	3'-CCA TAY AYA TCC G	4351	4351
<b>Y17</b>	5'-GGT ATA YAY AGG C	4471	4471
<b>Y18</b>	3'-CCA TAT AYA YCC G	4351	4351
<b>Y19</b>	5'-GGY AYA YAY AGG C	4930	4929
<b>Y20</b>	3'-CCA YAY AYA YCC G	4811	4809

<sup>a</sup> **Y** = 2-*N*-(pyren-1-ylmethyl)-2'-*N*-methylamino-uridine (monomer **Y**).



**Figure 2.5-1.** Representative thermal denaturation curves of Invader probes, corresponding duplexes between individual probe strands and cDNA, and unmodified reference duplex. For experimental conditions, see Table 2.2-1.

**Table 2.5-3.** Thermal denaturation temperatures ( $T_m$ 's) and thermal advantage ( $TA$ ) of DNA duplexes modified with 2'-*O*-(pyren-1-yl)methyluridine monomers.<sup>a</sup>

ON	Sequence	Zipper	$T_m$ [ $\Delta T_m$ ] (°C)		$TA$ (°C)
			Probe duplex	5'-ON:cDNA 3'-ON:cDNA	
X1	5'-GGXATATATAGGC	+7	53.0	44.5 [+7.0]	+0.5
X8	3'-CCATATATA $\underline{X}$ CCG		[+15.5]	46.5 [+9.0]	
X1	5'-GGXATATATAGGC	+5	55.5	44.5 [+7.0]	-1.0
X6	3'-CCATATA $\underline{X}$ ATCCG		[+18.0]	47.5 [+10.0]	
X3	5'-GGTAXATATAGGC	+5	56.0	47.5 [+10.0]	+0.5
X8	3'-CCATATATA $\underline{X}$ CCG		[+18.5]	46.5 [+9.0]	
X1	5'-GGXATATATAGGC	+3	54.0	44.5 [+7.0]	+1.5
X4	3'-CCATA $\underline{X}$ ATATCCG		[+16.5]	48.5 [+11.0]	
X3	5'-GGTAXATATAGGC	+3	57.0	47.5 [+10.0]	+0.5
X6	3'-CCATATA $\underline{X}$ ATCCG		[+19.5]	47.5 [+10.0]	
X5	5'-GGTATA $\underline{X}$ ATAGGC	+3	56.0	48.5 [+11.0]	+1.5
X8	3'-CCATATATA $\underline{X}$ CCG		[+18.5]	46.5 [+9.0]	
X3	5'-GGTAXATATAGGC	-1	53.0	47.5 [+10.0]	+4.5
X2	3'-CCA $\underline{X}$ ATATATCCG		[+15.5]	47.5 [+10.0]	
X5	5'-GGTATA $\underline{X}$ ATAGGC	-1	55.0	48.5 [+11.0]	+4.5
X4	3'-CCATA $\underline{X}$ ATATCCG		[+17.5]	48.5 [+11.0]	
X7	5'-GGTATATA $\underline{X}$ AGGC	-1	54.0	47.5 [+10.0]	+3.5
X6	3'-CCATATA $\underline{X}$ ATCCG		[+16.5]	47.5 [+10.0]	
X5	5'-GGTATA $\underline{X}$ ATAGGC	-3	57.0	48.5 [+11.0]	+1.5
X2	3'-CCA $\underline{X}$ ATATATCCG		[+19.5]	47.5 [+10.0]	
X7	5'-GGTATATA $\underline{X}$ AGGC	-3	57.0	47.5 [+10.0]	+1.5
X4	3'-CCATA $\underline{X}$ ATATCCG		[+19.5]	48.5 [+11.0]	
X7	5'-GGTATATA $\underline{X}$ AGGC	-5	57.0	47.5 [+10.0]	+0.5
X2	3'-CCA $\underline{X}$ ATATATCCG		[+19.5]	47.5 [+10.0]	

<sup>a</sup>  $\Delta T_m$  = change in  $T_m$  relative to corresponding unmodified DNA duplex ( $T_m = 37.5$  °C). For experimental conditions and definition of  $TA$ , see Table 2.2-1.

**Table 2.5-4.** Thermal denaturation temperatures ( $T_m$ 's) and thermal advantage ( $TA$ ) of DNA duplexes modified with 2'-*N*-(pyren-1-ylmethyl)-2'-*N*-methylaminouridine monomers.<sup>a</sup>

ON	Sequence	Zipper	$T_m$ [ $\Delta T_m$ ] (°C)		$TA$ (°C)
			Probe duplex	5'-ON:eDNA 3'-ON:eDNA	
Y1	5'-GGYATATATAGGC	+7	55.5	45.5 [+8.0]	+0.5
Y8	3'-CCATATATA $\underline{Y}$ CCG		[+18.0]	48.0 [+10.5]	
Y1	5'-GGYATATATAGGC	+5	56.0	45.5 [+8.0]	+1.0
Y6	3'-CCATATA $\underline{Y}$ ATCCG		[+18.5]	49.0 [+11.5]	
Y3	5'-GGTAYATATAGGC	+5	58.0	48.5 [+11.0]	+1.0
Y8	3'-CCATATATA $\underline{Y}$ CCG		[+20.5]	48.0 [+10.5]	
Y1	5'-GGYATATATAGGC	+3	56.0	45.5 [+8.0]	+3.0
Y4	3'-CCATA $\underline{Y}$ ATATCCG		[+18.5]	51.0 [+13.5]	
Y3	5'-GGTAYATATAGGC	+3	57.5	48.5 [+11.0]	+2.5
Y6	3'-CCATATA $\underline{Y}$ ATCCG		[+20.0]	49.0 [+11.5]	
Y5	5'-GGTATA $\underline{Y}$ ATAGGC	+3	57.0	49.5 [+12.0]	+3.0
Y8	3'-CCATATATA $\underline{Y}$ CCG		[+19.5]	48.0 [+10.5]	
Y3	5'-GGTAYATATAGGC	-1	54.5	48.5 [+11.0]	+4.0
Y2	3'-CCA $\underline{Y}$ ATATATCCG		[+17.0]	47.5 [+10.0]	
Y5	5'-GGTATA $\underline{Y}$ ATAGGC	-1	57.0	49.5 [+12.0]	+6.0
Y4	3'-CCATA $\underline{Y}$ ATATCCG		[+19.5]	51.0 [+13.5]	
Y7	5'-GGTATATA $\underline{Y}$ AGGC	-1	54.5	48.0 [+10.5]	+4.5
Y6	3'-CCATATA $\underline{Y}$ ATCCG		[+17.0]	49.0 [+11.5]	
Y5	5'-GGTATA $\underline{Y}$ ATAGGC	-3	58.5	49.5 [+12.0]	+1.0
Y2	3'-CCA $\underline{Y}$ ATATATCCG		[+21.0]	47.5 [+10.0]	
Y7	5'-GGTATATA $\underline{Y}$ AGGC	-3	59.5*	48.0 [+10.5]	+2.0
Y4	3'-CCATA $\underline{Y}$ ATATCCG		[+22.0]	51.0 [+13.5]	
Y7	5'-GGTATATA $\underline{Y}$ AGGC	-5	58.5	48.0 [+10.5]	-0.5
Y2	3'-CCA $\underline{Y}$ ATATATCCG		[+21.0]	47.5 [+10.0]	

<sup>a</sup>  $\Delta T_m$  = change in  $T_m$  relative to corresponding unmodified DNA duplex ( $T_m = 37.5$  °C). For experimental conditions and definition of  $TA$ , see Table 2.2-1.

\* Two transitions were observed ( $T_m = \sim 25$  °C and 59.5 °C) suggesting a complicated melting transition.

**Table 2.5-5.** Thermal denaturation temperatures ( $T_m$ 's) and thermal advantage ( $TA$ ) of DNA duplexes modified with 0-zippers of 2'-*O*-(pyren-1-yl)methyl-RNA monomers.<sup>a</sup>

ON	Sequence	Zipper	$T_m$ [ $\Delta T_m$ ] (°C)		$TA$ (°C)
			Probe duplex	5'-ON:cDNA 3'-ON:cDNA	
<b>X1</b> <b>A1</b>	5'-GG <u>X</u> ATATATAGGC 3'-CC <u>A</u> TATATATCCG	0	38.0 [+0.5]	44.5 [+7.0] 34.0 [-3.5]	+3.0
<b>X3</b> <b>A2</b>	5'-GGT <u>X</u> ATATATAGGC 3'-CCAT <u>A</u> TATATATCCG	0	40.0 [+2.5]	47.5 [+10.0] 31.0 [-6.5]	+1.0
<b>X5</b> <b>A3</b>	5'-GGTATA <u>X</u> ATAGGC 3'-CCATAT <u>A</u> TATCCG	0	40.0 [+2.5]	48.5 [+11.0] 31.0 [-6.5]	+2.0
<b>X7</b> <b>A4</b>	5'-GGTATATA <u>X</u> AGGC 3'-CCATATAT <u>A</u> TCCG	0	40.0 [+2.5]	47.5 [+10.0] 31.0 [-6.5]	+1.0
<b>X9</b> <b>A5</b>	5'-GG <u>X</u> <u>A</u> <u>X</u> ATATAGGC 3'-CC <u>A</u> <u>T</u> <u>A</u> TATATATCCG	0	41.0 [+3.5]	51.5 [+14.0] 27.0 [-10.5]	±0.0
<b>X11</b> <b>A6</b>	5'-GG <u>X</u> ATA <u>X</u> ATAGGC 3'-CC <u>A</u> TAT <u>A</u> TATATCCG	0	41.0 [+3.5]	53.5 [+16.0] 26.0 [-11.5]	+1.0
<b>X13</b> <b>A7</b>	5'-GG <u>X</u> ATATA <u>X</u> AGGC 3'-CC <u>A</u> TATAT <u>A</u> TCCG	0	42.0 [+4.5]	52.5 [+15.0] 26.0 [-11.5]	-1.0
<b>X15</b> <b>A8</b>	5'-GGT <u>X</u> <u>A</u> <u>X</u> ATAGGC 3'-CCAT <u>A</u> <u>T</u> <u>A</u> TATATCCG	0	42.0 [+4.5]	55.5 [+18.0] 26.0 [-11.5]	+2.0
<b>X17</b> <b>A9</b>	5'-GGTATA <u>X</u> <u>A</u> <u>X</u> AGGC 3'-CCATAT <u>A</u> <u>T</u> <u>A</u> TCCG	0	42.0 [+4.5]	54.5 [+17.0] 26.0 [-11.5]	+1.0
<b>X19</b> <b>A10</b>	5'-GG <u>X</u> <u>A</u> <u>X</u> <u>A</u> <u>X</u> AGGC 3'-CC <u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> TCCG	0	44.0 [+6.5]	65.5 [+28.0] 43.0 [+5.5]	+27.0

<sup>a</sup>  $\Delta T_m$  = change in  $T_m$  relative to corresponding unmodified DA duplex ( $T_m = 37.5$  °C). For experimental conditions and definition of  $TA$ , see Table 2.2-1. **A** = 2'-*O*-(pyren-1-yl)methyladenosine monomer.<sup>47</sup>

**Table 2.5-6.** Change in Gibbs free energy ( $\Delta G$ ) at 293K upon formation of X-modified DNA duplexes and change in reaction free energy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{293}$ ).<sup>a</sup>

Invader probe	Sequence	$\Delta G^{293}[\Delta\Delta G^{293}]$ (kJ/mol)			$\Delta G_{rec}^{293}$ (kJ/mol)
		Probe duplex	5'-Inv:cDNA	3'-Inv:cDNA	
<b>X1</b>	5'-GG <b>X</b> ATATATAGGC	-47 [+10]	-64 [-7]	-67 [-10]	-27
<b>X2</b>	3'-CC <b>A</b> XATATATCCG				
<b>X3</b>	5'-GGT <b>A</b> XATATAGGC	-50 [+7]	-70 [-13]	-69 [-12]	-32
<b>X4</b>	3'-CCAT <b>A</b> XATATCCG				
<b>X5</b>	5'-GGTAT <b>A</b> XATAGGC	-52 [+5]	-70 [-13]	-71 [-14]	-32
<b>X6</b>	3'-CCATAT <b>A</b> XATCCG				
<b>X7</b>	5'-GGTATAT <b>A</b> XAGGC	-46 [+11]	-68 [-11]	-66 [-9]	-31
<b>X8</b>	3'-CCATATAT <b>A</b> XCCG				
<b>X9</b>	5'-GG <b>X</b> <b>A</b> XATATAGGC	-51 [+6]	-71 [-14]	-77 [-20]	-40
<b>X10</b>	3'-CC <b>A</b> <b>X</b> <b>A</b> XATATCCG				
<b>X11</b>	5'-GG <b>X</b> AT <b>A</b> XATAGGC	-58 [-1]	-78 [-21]	-83 [-26]	-46
<b>X12</b>	3'-CC <b>A</b> <b>X</b> AT <b>A</b> XATCCG				
<b>X13</b>	5'-GG <b>X</b> ATAT <b>A</b> XAGGC	-59 [-2]	-75 [-18]	-79 [-22]	-38
<b>X14</b>	3'-CC <b>A</b> <b>X</b> ATAT <b>A</b> XCCG				
<b>X15</b>	5'-GGT <b>A</b> <b>X</b> <b>A</b> XATAGGC	-57 [ $\pm 0$ ]	-81 [-24]	-78 [-21]	-45
<b>X16</b>	3'-CCAT <b>A</b> <b>X</b> <b>A</b> XATCCG				
<b>X17</b>	5'-GGTAT <b>A</b> <b>X</b> <b>A</b> XAGGC	-54 [+3]	-77 [-20]	-80 [-23]	-46
<b>X18</b>	3'-CCATAT <b>A</b> <b>X</b> <b>A</b> XCCG				
<b>X19</b>	5'-GG <b>X</b> <b>A</b> <b>X</b> <b>A</b> <b>X</b> AGGC	-55 [+2]	-91 [-34]	-92 [-35]	-71
<b>X20</b>	3'-CC <b>A</b> <b>X</b> <b>A</b> <b>X</b> <b>A</b> XCCG				

<sup>a</sup>  $\Delta\Delta G^{293}$  is measured relative to the corresponding unmodified DNA duplex ( $\Delta G^{293} = -57$  kJ/mol).  $\Delta G_{rec}^{293} = \Delta G^{293}(5\text{-Inv:cDNA}) + \Delta G^{293}(3\text{-Inv:cDNA}) - \Delta G^{293}(\text{Invader}) - \Delta G^{293}(\text{dsDNA})$ .



**Table 2.5-7.** Change in Gibbs free energy ( $\Delta G$ ) at 293K upon formation of Y-modified DNA duplexes and change in reaction free energy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{293}$ ).<sup>a</sup>

Invader probe	Sequence	$\Delta G^{293}[\Delta\Delta G^{293}]$ (kJ/mol)			$\Delta G_{rec}^{293}$ (kJ/mol)
		Probe duplex	5'-Inv:cDNA	3'-Inv:cDNA	
<b>Y1</b>	5'-GGYATATATAGGC	-50 [+7]	-63 [-6]	-79 [-22]	-35
<b>Y2</b>	3'-CCAYATATATCCG				
<b>Y3</b>	5'-GGTAYATATAGGC	-62 [-5]	-75 [-18]	-82 [-25]	
<b>Y4</b>	3'-CCATAYATATCCG				
<b>Y5</b>	5'-GGTATAYATAGGC	-59 [-2]	-76 [-19]	-79 [-22]	
<b>Y6</b>	3'-CCATATAATCCG				
<b>Y7</b>	5'-GGTATATAYAGGC	-54 [+3]	-74 [-17]	-78 [-21]	
<b>Y8</b>	3'-CCATATATAYCCG				
<b>Y9</b>	5'-GGYAYATATAGGC	-54 [+3]	-80 [-23]	-94 [-37]	
<b>Y10</b>	3'-CCAYAYATATCCG				
<b>Y11</b>	5'-GGYATAYATAGGC	-63 [-6]	-86 [-29]	-92 [-35]	
<b>Y12</b>	3'-CCAYATAYATCCG				
<b>Y13</b>	5'-GGYATATAYAGGC	-58 [-1]	-84 [-27]	-100 [-43]	
<b>Y14</b>	3'-CCAYATATAYCCG				
<b>Y15</b>	5'-GGTAYAYATAGGC	-54 [+3]	-85 [-28]	-93 [-36]	
<b>Y16</b>	3'-CCATAYAYATCCG				
<b>Y17</b>	5'-GGTATAYAYAGGC	-61 [-4]	-83 [-26]	-84 [-27]	
<b>Y18</b>	3'-CCATATAYAYCCG				
<b>Y19</b>	5'-GGYAYAYAYAGGC	-58 [-1]	-109 [-52]	-101 [-44]	
<b>Y20</b>	3'-CCAYAYAYAYCCG				

<sup>a</sup>  $\Delta\Delta G^{293}$  is measured relative to the corresponding unmodified DNA duplex ( $\Delta G^{293} = -57$  kJ/mol).  $\Delta G_{rec}^{293} = \Delta G^{293}(5\text{'-Inv:cDNA}) + \Delta G^{293}(3\text{'-Inv:cDNA}) - \Delta G^{293}(\text{Invader}) - \Delta G^{293}(\text{dsDNA})$ .

**Table 2.5-8.** Change in enthalpy ( $\Delta H$ ) upon formation of **X**-modified duplexes and change in reaction enthalpy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{\text{rec}}$ ).<sup>a</sup>

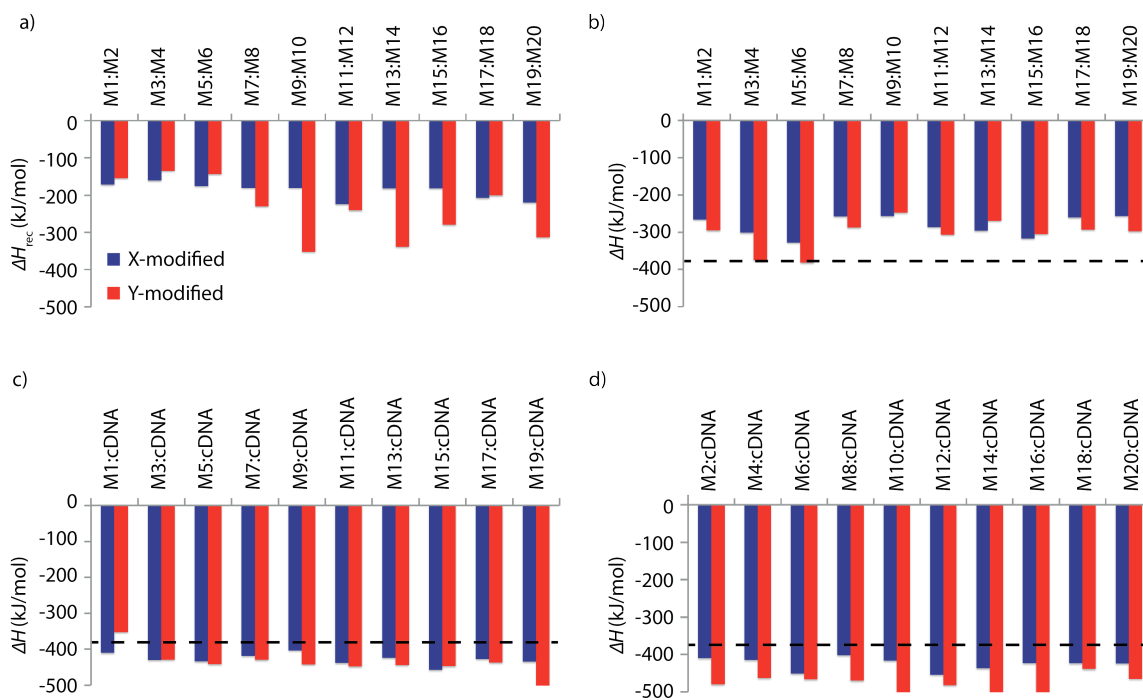
Invader probe	Sequence	$\Delta H$ [ $\Delta\Delta H$ ] (kJ/mol)			$\Delta H_{\text{rec}}$ (kJ/mol)
		Probe duplex	5'Inv:cDNA	3'-Inv:cDNA	
<b>X1</b>	5'-GG <b>X</b> ATATATAGGC	-266 [+116]	-410 [-28]	-410 [-28]	-172
<b>X2</b>	3'-CC <b>A</b> XATATATCCG				
<b>X3</b>	5'-GGT <b>A</b> XATATAGGC	-301 [+81]	-430 [-48]	-414 [-32]	-161
<b>X4</b>	3'-CC <b>A</b> T <b>A</b> XATATCCG				
<b>X5</b>	5'-GGT <b>A</b> T <b>A</b> XATAGGC	-328 [+54]	-434 [-52]	-451 [-69]	-175
<b>X6</b>	3'-CC <b>A</b> T <b>A</b> T <b>A</b> XATCCG				
<b>X7</b>	5'-GGT <b>A</b> T <b>A</b> T <b>A</b> XAGGC	-257 [+125]	-418 [-36]	-402 [-20]	-181
<b>X8</b>	3'-CC <b>A</b> T <b>A</b> T <b>A</b> T <b>A</b> XCCG				
<b>X9</b>	5'-GG <b>X</b> <b>A</b> XATATAGGC	-257 [+125]	-402 [-20]	-416 [-34]	-179
<b>X10</b>	3'-CC <b>A</b> X <b>A</b> XATATCCG				
<b>X11</b>	5'-GG <b>X</b> <b>A</b> T <b>A</b> XATAGGC	-286 [+96]	-438 [-56]	-454 [-72]	-224
<b>X12</b>	3'-CC <b>A</b> X <b>A</b> T <b>A</b> XATCCG				
<b>X13</b>	5'-GG <b>X</b> ATAT <b>A</b> XAGGC	-296 [+86]	-423 [-41]	-437 [-55]	-182
<b>X14</b>	3'-CC <b>A</b> XATAT <b>A</b> XCCG				
<b>X15</b>	5'-GGT <b>A</b> X <b>A</b> XATAGGC	-317 [+65]	-457 [-75]	-423 [-41]	-181
<b>X16</b>	3'-CC <b>A</b> T <b>A</b> X <b>A</b> XATCCG				
<b>X17</b>	5'-GGT <b>A</b> T <b>A</b> X <b>A</b> XAGGC	-260 [+122]	-427 [-45]	-424 [-42]	-209
<b>X18</b>	3'-CC <b>A</b> T <b>A</b> T <b>A</b> X <b>A</b> XCCG				
<b>X19</b>	5'-GG <b>X</b> <b>A</b> X <b>A</b> X <b>A</b> XAGGC	-256 [+126]	-434 [-52]	-424 [-42]	-220
<b>X20</b>	3'-CC <b>A</b> X <b>A</b> X <b>A</b> X <b>A</b> XCCG				

<sup>a</sup>  $\Delta\Delta H$  is measured relative to the corresponding unmodified DNA duplex ( $\Delta H = -382$  kJ/mol).  $\Delta H_{\text{rec}} = \Delta H(5'\text{-Inv:cDNA}) + \Delta H(3'\text{-Inv:cDNA}) - \Delta H(\text{Invader}) - \Delta H(\text{dsDNA})$ .

**Table 2.5-9.** Change in enthalpy ( $\Delta H$ ) upon formation of **Y**-modified duplexes and change in reaction enthalpy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{\text{rec}}$ ).<sup>a</sup>

Invader probe	Sequence	$\Delta H$ [ $\Delta\Delta H$ ] (kJ/mol)			$\Delta H_{\text{rec}}$ (kJ/mol)
		Probe duplex	5'Inv:cDNA	3'-Inv:cDNA	
<b>Y1</b>	5'-GG <b>Y</b> ATATATAGGC	-295 [+87]	-352 [+30]	-480 [-98]	-155
<b>Y2</b>	3'-CCA <b>Y</b> ATATATCCG				
<b>Y3</b>	5'-GGT <b>Y</b> ATATATAGGC	-375 [+7]	-429 [-47]	-463 [-81]	-135
<b>Y4</b>	3'-CCAT <b>Y</b> ATATATCCG				
<b>Y5</b>	5'-GGTAT <b>Y</b> ATATAGGC	-382 [ $\pm 0$ ]	-441 [-59]	-466 [-84]	-143
<b>Y6</b>	3'-CCATAT <b>Y</b> ATATCCG				
<b>Y7</b>	5'-GGTATAT <b>Y</b> AGGC	-287 [+95]	-429 [-47]	-470 [-88]	-230
<b>Y8</b>	3'-CCATATAT <b>Y</b> CCG				
<b>Y9</b>	5'-GG <b>Y</b> <b>Y</b> ATATATAGGC	-247 [+135]	-442 [-60]	-539 [-157]	-352
<b>Y10</b>	3'-CCA <b>Y</b> <b>Y</b> ATATATCCG				
<b>Y11</b>	5'-GG <b>Y</b> AT <b>Y</b> ATATAGGC	-307 [+75]	-448 [-66]	-482 [-100]	-241
<b>Y12</b>	3'-CCA <b>Y</b> AT <b>Y</b> ATATCCG				
<b>Y13</b>	5'-GG <b>Y</b> ATAT <b>Y</b> AGGC	-269 [+113]	-444 [-62]	-546 [-164]	-342
<b>Y14</b>	3'-CCA <b>Y</b> ATAT <b>Y</b> CCG				
<b>Y15</b>	5'-GGT <b>Y</b> <b>Y</b> ATATAGGC	-305 [+77]	-447 [-65]	-519 [-137]	-279
<b>Y16</b>	3'-CCAT <b>Y</b> <b>Y</b> ATATCCG				
<b>Y17</b>	5'-GGTAT <b>Y</b> <b>Y</b> AGGC	-293 [+89]	-437 [-55]	-439 [-57]	-201
<b>Y18</b>	3'-CCATAT <b>Y</b> <b>Y</b> CCG				
<b>Y19</b>	5'-GG <b>Y</b> <b>Y</b> <b>Y</b> ATATAGGC	-297 [+85]	-527 [-145]	-465 [-83]	-313
<b>Y20</b>	3'-CCA <b>Y</b> <b>Y</b> <b>Y</b> ATATCCG				

<sup>a</sup>  $\Delta\Delta H$  is measured relative to the corresponding unmodified DNA duplex ( $\Delta H = -382$  kJ/mol).  $\Delta H_{\text{rec}} = \Delta H(5'\text{-Inv:cDNA}) + \Delta H(3'\text{-Inv:cDNA}) - \Delta H(\text{Invader}) - \Delta H(\text{dsDNA})$ .



**Figure 2.5-2.** (a) Change in reaction enthalpy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{rec}$ ), and (b-d) change in enthalpy ( $\Delta H$ ) upon formation of X- or Y-modified DNA duplexes. The  $\Delta H$  for the dsDNA reference is shown as a dotted line at -382 kJ/mol. See Table 2.2-1 for experimental conditions. See Tables 2.5-6 and 2.5-7 for tabulated data.

**Table 2.5-10.** Change in entropy at 293K ( $-T^{293}\Delta S$ ) upon formation of X-modified duplexes and change in reaction entropy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $-T^{293}\Delta S_{rec}$ ).<sup>a</sup>

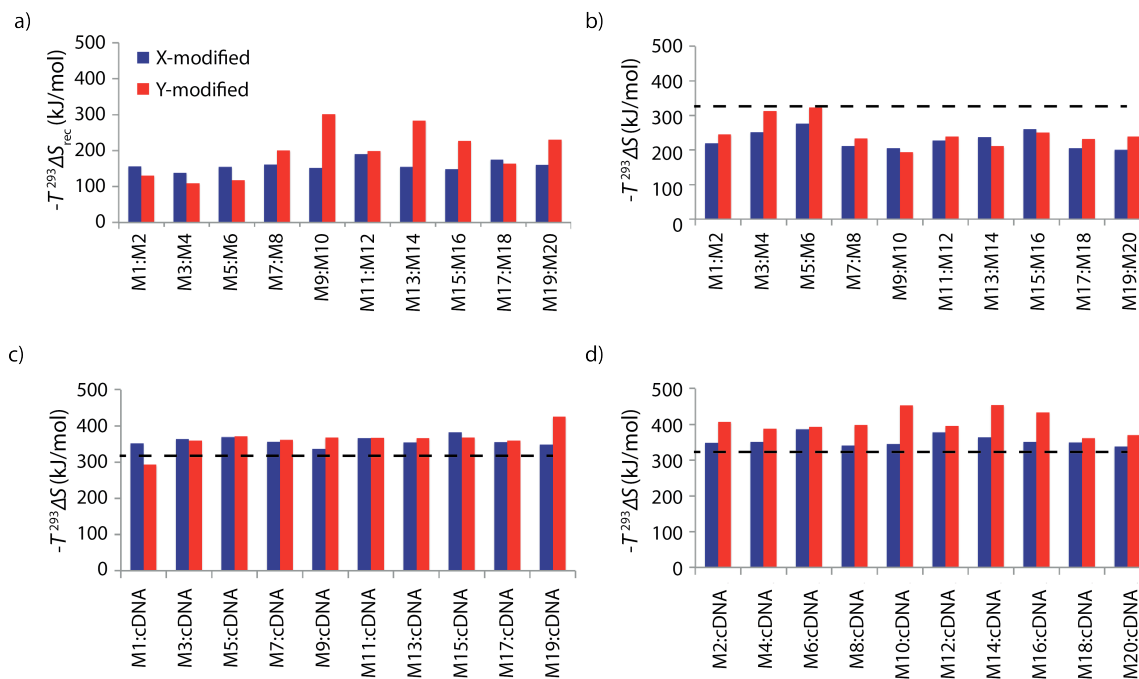
Invader probe	Sequence	$-T^{293}\Delta S$ [ $\Delta(T^{293}\Delta S)$ ] (kJ/mol)			$-T^{293}\Delta S_{rec}$ (kJ/mol)
		Probe duplex	5'-Inv:cDNA	3'-Inv:cDNA	
<b>X1</b>	5'-GG <b>X</b> ATATATAGGC	219 [-106]	346 [+21]	342 [+17]	144
<b>X2</b>	3'-CC <b>A</b> XATATATCCG				
<b>X3</b>	5'-GGT <b>A</b> XATATAGGC	251 [-74]	358 [+33]	345 [+20]	127
<b>X4</b>	3'-CC <b>A</b> T <b>A</b> XATATCCG				
<b>X5</b>	5'-GGT <b>A</b> T <b>A</b> XATAGGC	276 [-49]	363 [+38]	380 [+55]	142
<b>X6</b>	3'-CC <b>A</b> T <b>A</b> T <b>A</b> XATCCG				
<b>X7</b>	5'-GGT <b>A</b> T <b>A</b> T <b>A</b> XAGGC	211 [-114]	350 [+25]	336 [+11]	150
<b>X8</b>	3'-CC <b>A</b> T <b>A</b> T <b>A</b> T <b>A</b> XCCG				
<b>X9</b>	5'-GG <b>X</b> <b>A</b> XATATAGGC	205 [-120]	331 [+6]	339 [+14]	140
<b>X10</b>	3'-CC <b>A</b> X <b>A</b> XATATCCG				
<b>X11</b>	5'-GG <b>X</b> AT <b>A</b> XATAGGC	228 [-97]	360 [+35]	371 [+46]	178
<b>X12</b>	3'-CC <b>A</b> X <b>A</b> T <b>A</b> XATCCG				
<b>X13</b>	5'-GG <b>X</b> AT <b>A</b> T <b>A</b> XAGGC	237 [-88]	348 [+23]	357 [+32]	143
<b>X14</b>	3'-CC <b>A</b> X <b>A</b> T <b>A</b> T <b>A</b> XCCG				
<b>X15</b>	5'-GGT <b>A</b> X <b>A</b> XATAGGC	260 [-65]	376 [+51]	345 [+20]	136
<b>X16</b>	3'-CC <b>A</b> T <b>A</b> X <b>A</b> XATCCG				
<b>X17</b>	5'-GGT <b>A</b> T <b>A</b> X <b>A</b> XAGGC	205 [-120]	349 [+24]	344 [+19]	163
<b>X18</b>	3'-CC <b>A</b> T <b>A</b> T <b>A</b> X <b>A</b> XCCG				
<b>X19</b>	5'-GG <b>X</b> <b>A</b> X <b>A</b> X <b>A</b> XAGGC	201 [-124]	343 [+18]	332 [+7]	149
<b>X20</b>	3'-CC <b>A</b> X <b>A</b> X <b>A</b> X <b>A</b> XCCG				

<sup>a</sup>  $\Delta(T^{293}\Delta S)$  is measured relative to the corresponding unmodified DNA duplex ( $-T^{293}\Delta S = 325$  kJ/mol).  $-T^{293}\Delta S_{rec} = \Delta(T^{293}\Delta S)$  (5'-Inv:cDNA) +  $\Delta(T^{293}\Delta S)$  (3'-Inv:cDNA) -  $\Delta(T^{293}\Delta S)$  (Invader).

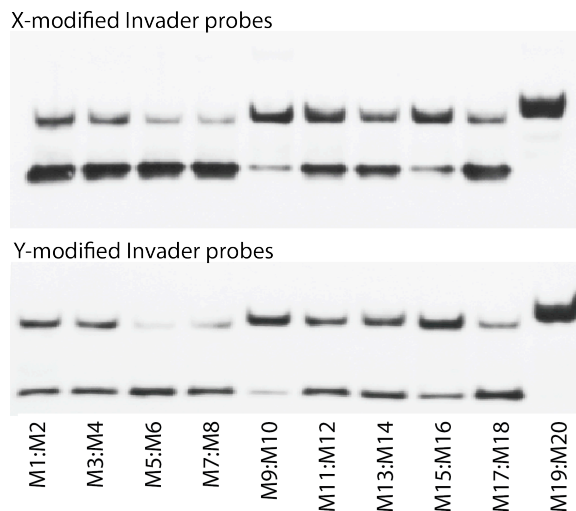
**Table 2.5-11.** Change in entropy at 293K ( $-T^{293}\Delta S$ ) upon formation of Y-modified duplexes and change in reaction entropy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $-T^{293}\Delta S_{rec}$ ).<sup>a</sup>

Invader probe	Sequence	$-T^{293}\Delta S$ [ $\Delta(T^{293}\Delta S)$ ] (kJ/mol)			$-T^{293}\Delta S_{rec}$ (kJ/mol)
		Probe duplex	5'-Inv:cDNA	3'-Inv:cDNA	
<b>Y1</b>	5'-GGYATATATAGGC	245 [-80]	289 [-36]	401 [+76]	120
<b>Y2</b>	3'-CCAYATATATCCG				
<b>Y3</b>	5'-GGTAYATATAGGC	313 [-12]	354 [+29]	381 [+56]	
<b>Y4</b>	3'-CCATAYATATCCG				
<b>Y5</b>	5'-GGTATAYATAGGC	322 [-3]	365 [+40]	387 [+62]	
<b>Y6</b>	3'-CCATATAYATCCG				
<b>Y7</b>	5'-GGTATATAYAGGC	233 [-92]	355 [+30]	392 [+67]	
<b>Y8</b>	3'-CCATATATAYCCG				
<b>Y9</b>	5'-GGYAYATATAGGC	194 [-131]	362 [+37]	445 [+120]	
<b>Y10</b>	3'-CCAYAYATATCCG				
<b>Y11</b>	5'-GGYATAYATAGGC	239 [-86]	361 [+36]	389 [+64]	
<b>Y12</b>	3'-CCAYATAYATCCG				
<b>Y13</b>	5'-GGYATATAYAGGC	211 [-114]	359 [+34]	446 [+121]	
<b>Y14</b>	3'-CCAYATATAYCCG				
<b>Y15</b>	5'-GGTAYAYATAGGC	250 [-75]	362 [+37]	426 [+101]	
<b>Y16</b>	3'-CCATAYAYATCCG				
<b>Y17</b>	5'-GGTATAYAYAGGC	232 [-93]	353 [+28]	355 [+30]	
<b>Y18</b>	3'-CCATATAYAYCCG				
<b>Y19</b>	5'-GGYAYAYAYAGGC	240 [-85]	418 [+93]	364 [+39]	
<b>Y20</b>	3'-CCAYAYAYAYCCG				

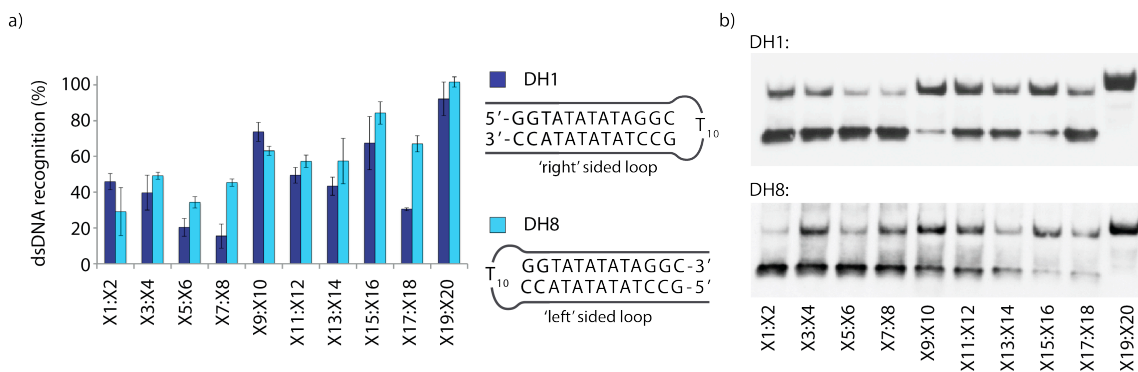
<sup>a</sup>  $\Delta(T^{293}\Delta S)$  is measured relative to the corresponding unmodified DNA duplex ( $-T^{293}\Delta S = 325$  kJ/mol).  $-T^{293}\Delta S_{rec} = \Delta(T^{293}\Delta S)$  (5'-Inv:cDNA) +  $\Delta(T^{293}\Delta S)$  (3'-Inv:cDNA) -  $\Delta(T^{293}\Delta S)$  (Invader).



**Figure 2.5-3.** (a) Change in reaction entropy upon hypothetical Invader-mediated recognition of isosequential dsDNA targets ( $-T^{293}\Delta S_{\text{rec}}$ ), and (b-d) change in entropy ( $-T^{293}\Delta S$ ) upon formation of X- or Y-modified DNA duplexes. The  $-T^{293}\Delta S$  for the dsDNA reference is shown as a dotted line at 325 kJ/mol. See Table 2.2-1 for experimental conditions. See Tables 2.5-10 and 2.5-11 for tabulated data.











**Figure 2.5-4.** Representative electrophoretograms of data shown in Figure 2.2-1 of main text: dsDNA recognition by 200-fold molar excess of Invader probes modified either monomer **X** (top) or **Y** (bottom). For sequences and experimental conditions see Table 2.2-1 and Figure 2.2-3, respectively.



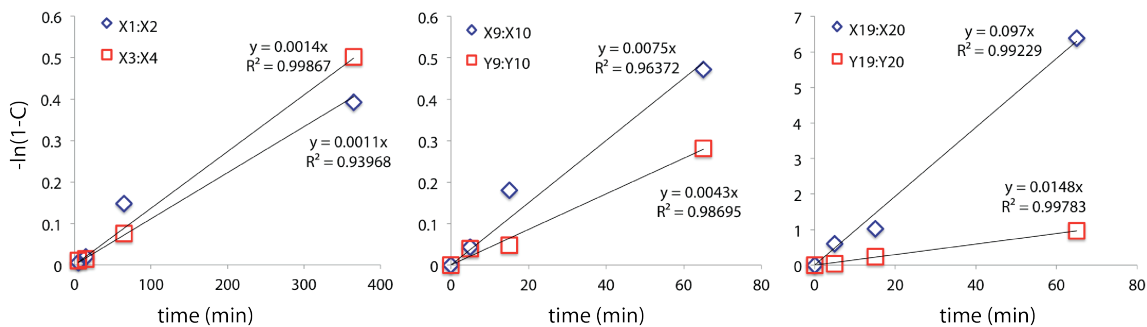
**Figure 2.5-5.** (a) Recognition of **DH1** or **DH8** using Invader probes **X1:X2–X19:X20** and (b) representative gel images. See Figure 2.2-1 for experimental conditions.



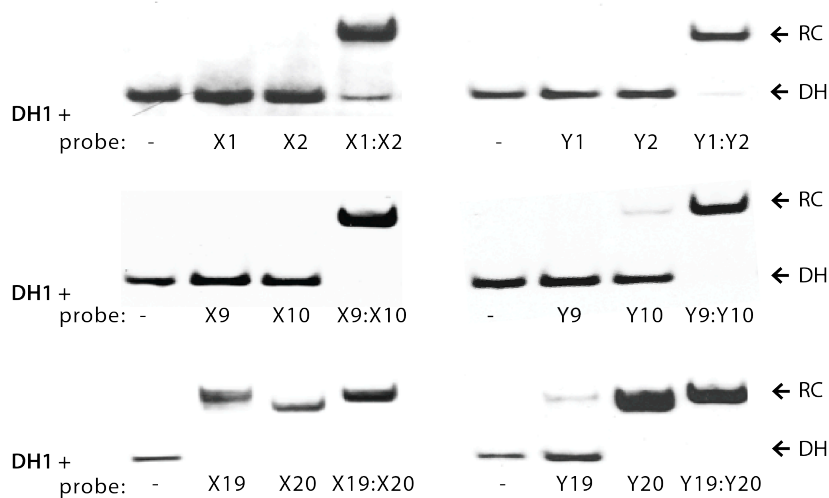
**Table 2.5-12.** Thermal denaturation temperatures of DNA hairpin targets used in this study.<sup>a</sup>

DH	Sequence	$T_m$ (°C)
1	 5'-GGTATATATAGGC 3'-CCATATATATCCG	58.5
2	 5'-GGTATTTATAGGC 3'-CCATAAATATCCG	60.5
3	 5'-GGTATGTATAGGC 3'-CCATACATATCCG	63.5
4	 5'-GGTATCTATAGGC 3'-CCATAGATATCCG	63.0
5	 5'-GGTATATAAGGC 3'-CCATATATCCG	60.0
6	 5'-GGTATATAGAGGC 3'-CCATATATCCG	62.5
7	 5'-GGTATATACAGGC 3'-CCATATATCCG	62.5
8	 GGTATATATAGGC-3' CCATATATATCCG-5'	59.5

<sup>a</sup>For experimental conditions, see Table 2.2-1.



**Figure 2.5-6.** Plots of  $-\ln(1-C)$  versus time, where  $C$  is the ratio of recognition complex with respect to the total hairpin concentration, used to determine pseudo-first order rate constants for initial phases of **DH1** recognition using a) **M1:M2**, b) **M9:M10**, or c) **M19:M20** at 200-fold molar excess. See Figure 2.2-1 for incubation conditions. The linearity of the plots suggests that the reaction obeys pseudo-first order kinetics.



**Figure 2.5-7.** Recognition of **DH1** using individual probe strands. Single-stranded probes (and double-stranded Invader controls) were used at 1000-, 500-, or 50-fold molar excess for **M1:M2**, **M9:M10**, or **M19:M20**, respectively. For experimental conditions, see Figure 2.2-1.

**Table 2.5-13.** Change in Gibbs free energy ( $\Delta G$ ) at 310K upon formation of **X**-modified DNA duplexes and change in reaction free energy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{310}$ ).<sup>a</sup>

INV	Sequence	$\Delta G^{310}[\Delta\Delta G^{310}]$ (kJ/mol)				$\Delta G_{rec}^{310}$ (kJ/mol)
		Probe duplex	5'-Inv: cDNA	3'-Inv: cDNA	dsDNA	
1	5'-Cy3-AGCCCUGUGCCCTG 3'-TCGGGACACGGGAC-Cy3	-61 [+11]	-83 [-11]	-96 [-24]	-72	-46
2	5'-Cy3-CCUGUGCCCTG 3'-GGACACGGGAC-Cy3	-48 [+5]	-60 [-7]	-71 [-18]	-53	-30
3	5'-Cy3-CCUGTGCCCTG 3'-GGACACGGGAC-Cy3	-49 [+4]	-61 [-8]	-65 [-12]	-53	-24
4	5'-Cy3-AGCCCUGUGCCCTG 3'-TCGGGACACGGGAC-Cy3	-54 [+18]	-65 [+7]	-90 [-18]	-72	-29

<sup>a</sup>  $\Delta\Delta G^{293}$  is measured relative to the corresponding unmodified DNA duplex.  $\Delta G_{rec}^{310} = \Delta G^{310}(5'-Inv:cDNA) + \Delta G^{310}(3'-Inv:cDNA) - \Delta G^{310}(\text{Invader}) - \Delta G^{310}(\text{dsDNA})$ .

**Table 2.5-14.** Change in enthalpy ( $\Delta H$ ) upon formation of **X**-modified duplexes and change in reaction enthalpy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{rec}$ ).<sup>a</sup>

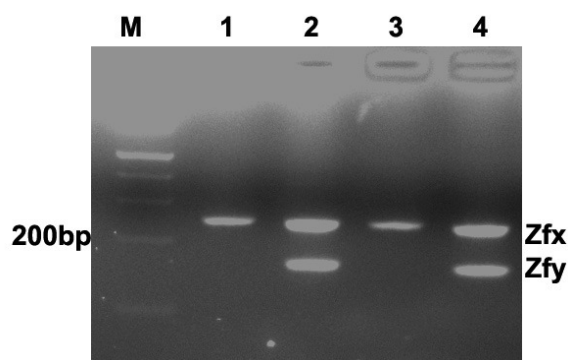
INV	Sequence	$\Delta H [\Delta\Delta H]$ (kJ/mol)				$\Delta H_{rec}$ (kJ/mol)
		Probe duplex	5'-Inv: cDNA	3'-Inv: cDNA	dsDNA	
1	5'-Cy3-AGCCCUGUGCCCTG 3'-TCGGGACACGGGAC-Cy3	-342 [+127]	-476 [-7]	-539 [-70]	-469	-204
2	5'-Cy3-CCUGUGCCCTG 3'-GGACACGGGAC-Cy3	-210 [+144]	-333 [+21]	-391 [-37]	-354	-160
3	5'-Cy3-CCUGTGCCCTG 3'-GGACACGGGAC-Cy3	-213 [+141]	-350 [+4]	-350 [+4]	-354	-133
4	5'-Cy3-AGCCCUGUGCCCTG 3'-TCGGGACACGGGAC-Cy3	-203 [+266]	-295 [+174]	-472 [-3]	-469	-95

<sup>a</sup>  $\Delta\Delta H$  is measured relative to the corresponding unmodified DNA duplex.  $\Delta H_{rec} = \Delta H(5'-Inv:cDNA) + \Delta H(3'-Inv:cDNA) - \Delta H(\text{Invader}) - \Delta H(\text{dsDNA})$ .

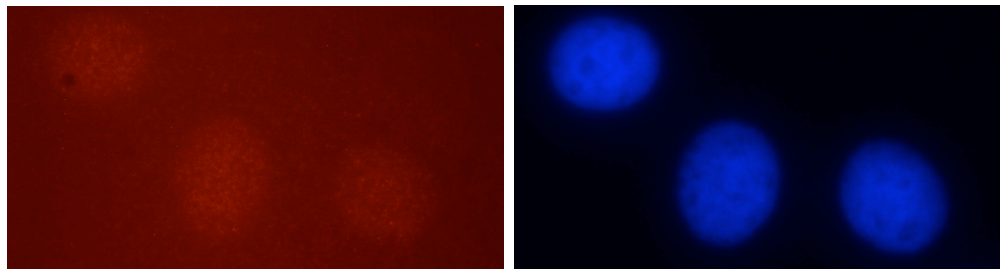
**Table 2.5-15.** Change in entropy at 310K ( $-T^{310}\Delta S$ ) upon formation of **Y**-modified duplexes and change in reaction entropy upon Invader-mediated recognition of isosequential dsDNA targets ( $-T^{310}\Delta S_{rec}$ ).<sup>a</sup>

INV	Sequence	$-T^{310}\Delta S$ [ $\Delta(T^{310}\Delta S)$ ] (kJ/mol)				$-T^{310}\Delta S_{rec}$ (kJ/mol)
		Probe duplex	5'-Inv:cDNA	3'-Inv:cDNA	dsDNA	
1	5'-Cy3-AGCCC <u>UGUG</u> CCCTG 3'-TCGGGAC <u>AC</u> CGGGAC-Cy3	281 [-116]	393 [-4]	443 [+46]	397	158
2	5'-Cy3-CC <u>UGUG</u> CCCTG 3'-GGAC <u>AC</u> CGGGAC-Cy3	162 [-139]	273 [-28]	320 [+19]	301	130
3	5'-Cy3-CC <u>UGTG</u> CCCTG 3'-GGAC <u>AC</u> CGGGAC-Cy3	164 [-137]	289 [-12]	285 [-16]	301	109
4	5'-Cy3-AGCCC <u>UGUG</u> CCCTG 3'- <u>TC</u> GGGAC <u>AC</u> CGGGAC-Cy3	148 [-249]	229 [-168]	382 [-15]	397	66

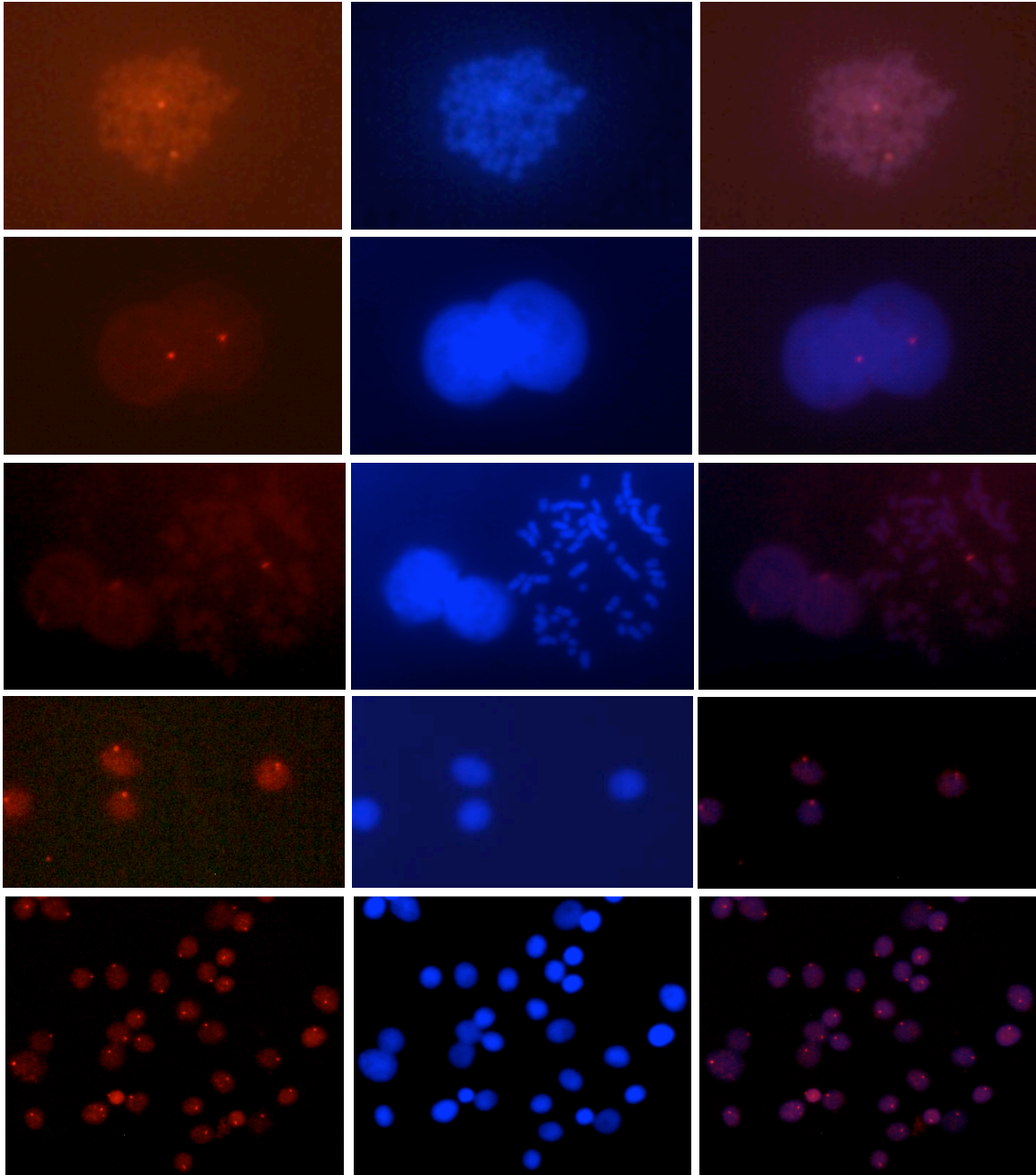
<sup>a</sup>  $\Delta(T^{293}\Delta S)$  is measured relative to the corresponding unmodified DNA duplex.  $-T^{310}\Delta S_{rec} = \Delta(T^{310}\Delta S)$  (5'-Inv:cDNA) +  $\Delta(T^{310}\Delta S)$  (3'-Inv:cDNA) -  $\Delta(T^{310}\Delta S)$  (Invader probe).



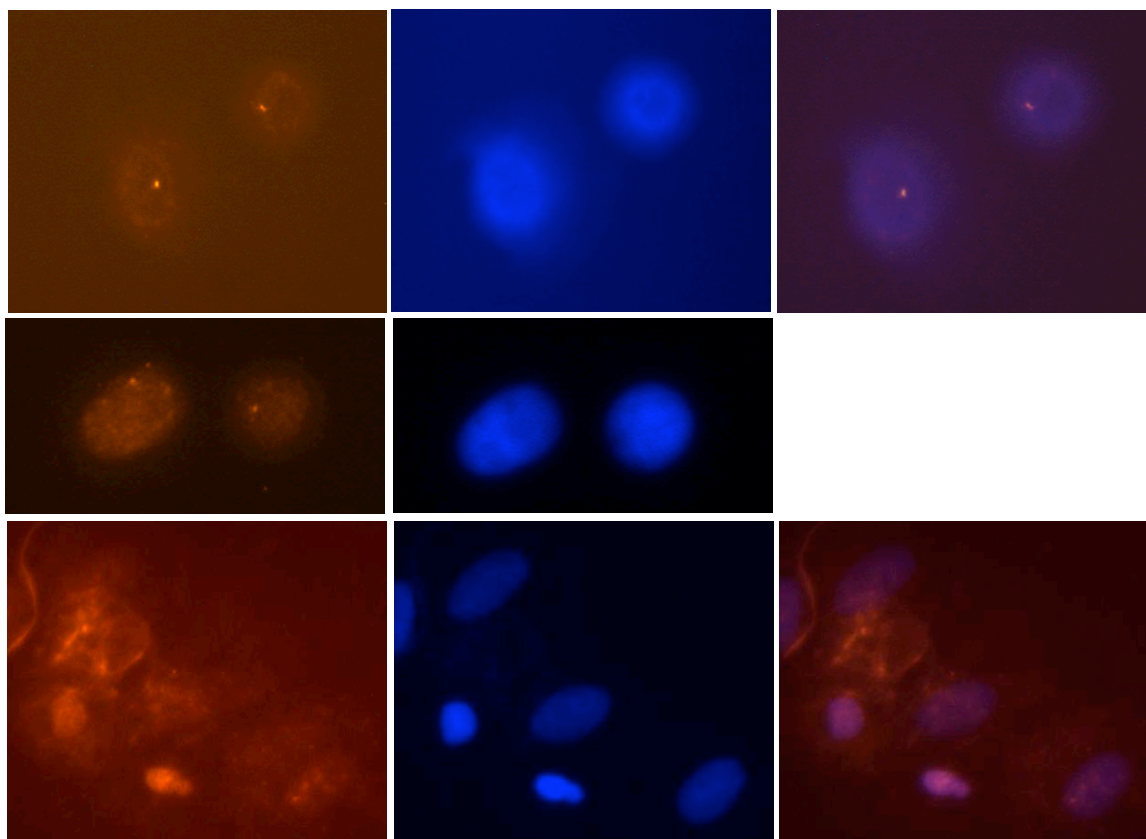
**Figure 2.5-8.** Gel electrophoretogram of gender-specific PCR amplicons. Zfx and Zfy bands represent **X**- and **Y**-chromosome specific amplicons, respectively. Lane “M” contains a molecular weight standard (Lo DNA Marker, Bionexus). PCR amplicons from female (lanes 1 and 3) and male cell lines (lanes 2 and 4). 2% agarose gels were used.



**Figure 2.5-9.** Images from FISH experiments conducted at non-denaturing conditions using nuclei from male bovine kidney cells incubated with the single-stranded Cy3-labeled PNA probe 5'-Cy3-OO-AGCCCTGTGCCCTG. Images viewed using Cy3 (left column) or DAPI (middle column) filter settings. Incubation: 3h at 38.5 °C in 10 mM Tris-Cl, pH 8.0 and 1 mM EDTA. Cells were visualized at 400x magnification using a Zeiss AxioSkop 40 fluorescent microscope and images captured using a Zeiss AxioCam MRc5 camera. "O" denotes a 9-atom ethylene glycol linker (i.e., "O-linker").



**Figure 2.5-10.** Additional images from FISH experiments using nuclei from male bovine kidney cells incubated with the following Y-chromosome specific Invaders: **INV1** (top panel; nuclei in metaphase prior to cell division), **INV2** (second panel; nuclei in interphase), **INV2** (third panel; inter- and metaphase nuclei), **INV3** (fourth panel; interphase nuclei) and **INV4** (bottom panel; interphase nuclei). Images viewed using Cy3 (left column) or DAPI (middle column) filter settings; overlay images are shown in the right column. Sequences are shown in Table 2.2-1. For experimental details, see legend of Figure 2.2-9.



**Figure 2.5-11.** Images from FISH experiments involving nuclei from male bovine kidney cells that were treated with proteinase K (upper panel), RNase (middle panel) and DNase (lower panel) prior to incubation with Invader **INV4**. Note the continued presence of Cy3-signals in proteinase K or RNase pre-treated samples, and the absence of Cy3-signals in DNase pre-treated samples. For experimental details, see legend of Figure 2.2-9.

### 3.6 References

1. Besch, R.; Giovannangeli, C.; Degitz, K. *Curr. Drug Targets* **2004**, *5*, 691.
2. Rogers, F. A.; Lloyd, J. A.; Glazer, P. M. *Curr. Med. Chem.: Anti-Cancer Agents* **2005**, *5*, 319.
3. Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. J. *Mol. BioSyst.* **2006**, *2*, 551.
4. Nielsen, P. E. *Chem. Biodiv.* **2010**, *7*, 786.
5. Mukherjee, A.; Vasquez, K. M. *Biochimie* **2011**, *93*, 1197.

6. Aiba, Y.; Sumaoka, J.; Komiyama, M. *Chem. Soc. Rev.* **2011**, *40*, 5657.
7. Vajjayanthi, T.; Bando, T.; Pandian, G. N.; Sugiyama, H. *ChemBioChem* **2012**, *13*, 2170.
8. Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, *36*, 5123.
9. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284.
10. Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem.* **2013**, *21*, 6101.
11. Nielsen, P. E., Egholm, M., Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
12. Kaihatsu, K.; Janowski, B. A.; Corey, D. R. *Chem. Biol.* **2004**, *11*, 749.
13. Gaj, T.; Gersbach, C. A.; Barbas III, C. F. *Trends Biotechnol.* **2013**, *31*, 397.
14. Sander, J. D.; Joung, J. K. *Nat. Biotechnol.* **2014**, *32*, 347.
15. Trauger, J. W.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 3534.
16. Fu, Y.; Foden, J. A.; Khayter, C.; Maeder, M. L.; Reyon, D.; Joung, J. K.; Sander, J. D. *Nat. Biotechnol.* **2013**, *31*, 822.
17. Bentin, T.; Larsen, H. J.; Nielsen, P. E. *Biochemistry* **2003**, *42*, 13987.
18. Kaihatsu, K.; Shah, R. H.; Zhao, X.; Corey, D. R. *Biochemistry* **2003**, *42*, 13996.
19. Rapireddy, S.; Bahal, R.; Ly, D. H. *Biochemistry* **2011**, *50*, 3913.
20. Bahal, R.; Sahu, B.; Rapireddy, S.; Lee, C-M.; Ly, D. H. *ChemBioChem* **2012**, *13*, 56.
21. Kutuyavin, I. V.; Rhinehart, R. L.; Lukhtanov, E. A.; Gorn, V. V.; Meyer Jr., R. B.; Gamper Jr., H. B. *Biochemistry* **1996**, *35*, 11170.
22. Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11804.



23. Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corradini, R.; Sforza, S.; Komiyama, M. *Nucleic Acids Res.* **2008**, *36*, 1464.
24. 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.
25. Sumaoka, J.; Komiyama, M. *Chem. Lett.* **2014**, *43*, 1581.
26. Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2010**, *8*, 2028.
27. 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.; Wendel, J.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, *78*, 9560.
28. Crothers, D. M. *Biopolymers* **1968**, *6*, 575.
29. Tsai, C.; Jain, S. C.; Sobell, H. M. *J. Mol. Biol.* **1977**, *114*, 301.
30. Williams, L. D.; Egli, M.; Gao, Q.; Rich, A. In: Structure and Function, Volume 1: Nucleic Acids, 1992, pp 107-125 (Eds. Sarma R. H.; Sarma, M. H.). Adenine press.
31. Denn, B.; Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *Chem. Commun.* **2013**, *49*, 9851.
32. Didion, B. A.; Karmakar, S.; Guenther, D. C.; Sau, S. P.; Verstegen, J. P.; Hrdlicka, P. J. *ChemBioChem* **2013**, *14*, 1534.
33. Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2014**, *12*, 7758.
34. Levsky, J. M.; Singer, R. H. *J. Cell Sci.* **2003**, *116*, 2833.
35. Johnson III, M. D.; Fresco, J. R. *Chromosoma* **1999**, *108*, 181.

36. Schmitt, E.; Schwarz-Finsterle, J.; Stein, S.; Mueller, P.; Mokhir, A.; Kraemer, R.; Cremer, C.; Hausmann, M. *Methods Mol. Biol.* **2010**, *659*, 185.
37. Molenaar, C.; Wiesmeijer, K.; Verwoerd, N. P.; Khazen, S.; Eils, R.; Tanke, H. J.; Dirks, R. W. *EMBO J.* **2003**, *22*, 6631.
38. Janssen, S.; Durussel, T.; Laemmli, U. K. *Mol. Cell* **2000**, *6*, 999.
39. Gygi, M. P.; Ferguson, M. D.; Mefford, H. C.; Lund, K. P.; O'Day, C.; Zhou, P.; Friedman, C.; van der Engh, G.; Stolowitz, M. L.; Trask, B. J. *Nucleic Acids Res.* **2002**, *30*, 2790.
40. Silahtaroglu, A. N.; Tommerup, N.; Vissing, H. *Mol. Cell. Probes* **2003**, *17*, 165.
41. Sugimura, K.; Takebayashi, S.; Ogata, S.; Taguchi, H.; Okumura, K. *Biosci. Biotechnol. Biochem.* **2007**, *71*, 627.
42. Cuadrado, N.; Jouve, N. *Chromosoma*, **2010**, *19*, 495.
43. Perret, J.; Shia, Y.; Fries, R.; Vassart, G.; Georges, M. *Genomics* **1990**, *6*, 482.
44. Bleher, R.; Erwin, W.; Paprocki, A. M.; Syverson, C. M.; Koppang, R.; Didion, B. A. *Reprod. Fertil. Dev. B* **2009**, *21*, 227.
45. Didion, B. A.; Bleher, R. *Reprod. Fertil. Dev.* **2008**, *21*, 229.
46. Didion, B. A.; Erwin, W.; Bleher, R. WO2009/079456A2.
47. Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, *78*, 12040.
48. Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem.* **2011**, *76*, 7119.
49. Morgan, M. A.; Okamoto, K.; Kahn, J. D.; English, D. S. *Biophys. J.* **2005**, *89*, 2588.
50. Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515.

51. Kirkpatrick, B.; Monson, R. *J. Reprod. Fertil.* **1993**, 98, 335-340.

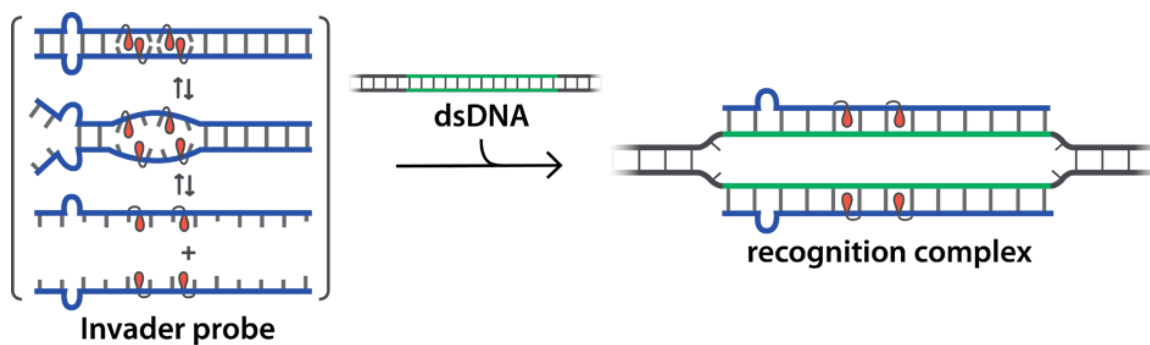
**CHAPTER 3: Bulged Invader probes: activated duplexes for mixed-sequence  
dsDNA recognition with improved thermodynamic and kinetic profiles**

Dale C. Guenther, Saswata Karmakar, and Patrick J. Hrdlicka.\*

\*Department of Chemistry, University of Idaho, 875 Perimeter Dr, Moscow, ID 83844-  
2343, USA.

Published in: *Chem. Commun.* **2015**, 51, 15051-15054.

Reproduced by permission of The Royal Society of Chemistry



## Abstract

Hybridization-based probes for recognition of double-stranded DNA (dsDNA) offer great promise as fundamental research tools, diagnostic agents, and components for applications in nanotechnology. Double-stranded oligonucleotides with +1 interstrand zipper arrangements of intercalator-functionalized nucleotides are energetically activated for recognition of mixed-sequence chromosomal DNA at non-denaturing conditions. Here, strategic incorporation of non-nucleosidic nonyl ( $C_9$ ) bulges into these probes is shown to effect higher affinity (>5-fold), faster (>4-fold) and longer lasting dsDNA recognition relative to conventional Invader probes.

## 3.1 Introduction

Chemical probes capable of sequence-specific recognition of double-stranded DNA (dsDNA) have tremendous potential as tools in genomic diagnostics, structural and functional elucidations, and nanotechnology.<sup>1-8</sup> Hybridization-based approaches are particularly interesting due to their predictable binding interactions and the resulting ease of design. To realize sequence-specific dsDNA recognition, the probes must either gain access to the Watson-Crick face or bind via extrahelical contacts, such as Hoogsteen base pairing. Triplex-forming oligonucleotides<sup>1,9</sup> and peptide nucleic acids (PNAs)<sup>5,10</sup> are examples of the latter approach, which have been used with some success. However, triplex-based approaches rely on the presence of long polypurine regions, which limits the number of suitable targets. Conformationally restricted  $\gamma$ -PNAs,<sup>11</sup> on the other hand, bind to complementary DNA (cDNA) with such high affinity that they can invade the Watson-Crick base-pairing of dsDNA target regions, albeit only at non-physiological

ionic strengths, resulting in the displacement of one of the target strands and the formation of a D-loop.

Double-stranded probes that can bind to both target strands offer the promise of even greater thermodynamic driving forces for dsDNA recognition. This approach also provides access to more specific probes, as binding to mismatched dsDNA targets would generate two mismatched duplexes, a thermodynamically very unfavorable process.<sup>12</sup> However, facile dissociation of the probe duplex is necessary for this approach to work. One approach to realize this has been through the use of pseudocomplementary (pc) base pairs such as 2,6-diaminopurine and 2-thiouracil, which form weak base-pairs with each other, while allowing for stable pairing with thymine and adenine in target strands.<sup>13</sup> The energy difference between the double-stranded probe and the resulting probe-target duplexes generates a thermodynamic gradient for dsDNA recognition. While only modest successes have been achieved with pcDNA,<sup>14</sup> pcPNA have been shown to recognize internal regions of mixed-sequence dsDNA, although only at low ionic strengths.<sup>15,16</sup>

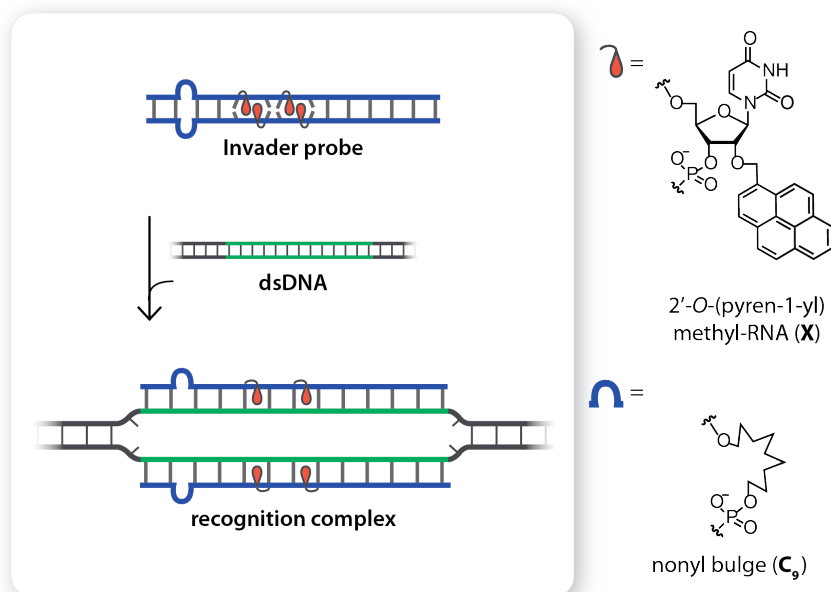
As part of our efforts aimed at developing probes for mixed-sequence dsDNA-recognition at physiological conditions, we recently introduced so-called Invader probes, which also rely on energy differences between probe duplexes and recognition complexes to drive dsDNA-recognition (Figure 3.2-1).<sup>17,18</sup> These double-stranded probes contain intercalator-functionalized nucleotides that are arranged in +1 interstrand zipper motifs, whereby two intercalators are forced to compete for the same inter-base-pair region, leading to violation of the nearest-neighbor exclusion principle<sup>19-21</sup> and probe destabilization.<sup>17,18,22-25</sup> In the recognition complex, where each probe strand is bound to a complementary DNA target region, the intercalators no longer compete for the same

space, leading to strong stabilization of the resulting duplex due to efficient  $\pi$ - $\pi$ -stacking interactions between the intercalator and the neighboring base-pairs. In previous studies, we have: i) identified synthetically more readily accessible analogs of the N2'-pyrene-functionalized 2'-amino- $\alpha$ -L-LNA (Locked Nucleic Acid) monomers originally used in Invader probes,<sup>18</sup> which include the 2'-*O*-(pyren-1-yl)methyl-RNA monomers shown in Figure 3.2-1, ii) studied the influence on dsDNA recognition efficiency of the intercalator, linker, and nucleobase of the key building blocks, as well as, the number and distance between the intercalator-functionalized nucleotides,<sup>17,18,22-25</sup> and iii) demonstrated recognition of chromosomal DNA targets at non-denaturing conditions.<sup>25</sup>

Herein, we describe improved dsDNA recognition using a novel Invader probe architecture that contains non-nucleosidic nonyl (C<sub>9</sub>) bulge inserts (Figure 3.2-1). This design was pursued based on the hypothesis that internal C<sub>9</sub> bulges would destabilize the probe duplex and promote local denaturation, thus revealing the Watson-Crick face of the probe and accelerate nucleation with, and invasion of, dsDNA targets. Bulges have been previously used to tune the hybridization properties of oligonucleotides.<sup>26,27</sup> Although they induce minimal perturbation of the global duplex conformation, they invariably destabilize duplexes.<sup>27</sup> By carefully adjusting the number and position of C<sub>9</sub> bulges, we hypothesized that we would be able to tune the thermodynamic properties in a manner that destabilizes probe duplexes more than probe-target duplexes, resulting in a more prominent thermodynamic driving force and faster dsDNA recognition.

### 3.2 Results and discussion

A library of Invader probes, containing two consecutive +1 interstrand zipper arrangements of 2'-*O*-(pyren-1-yl)methyl-RNA-U monomers at the center and one or two C<sub>9</sub> bulges at one or both termini, were synthesized (Table 3.2-1). Thermal denaturation temperatures ( $T_m$ 's) of these probes and the duplexes with cDNA were determined in medium salt phosphate buffer and compared to conventional Invader probes without C<sub>9</sub> bulges. As expected from our previous work, reference Invader strands **ON1** and **ON2** form very stable duplexes with cDNA ( $\Delta T_m = 18$  °C relative to unmodified ON).<sup>25</sup> The insertion of a single C<sub>9</sub> bulge into an Invader strand greatly reduces  $T_m$ 's (-9 to -12 °C) relative to **ON1** or **ON2**. Insertion of two C<sub>9</sub> bulges potentiates these trends ( $T_m < 15$  °C for **ON7** or **ON8** vs cDNA).



**Figure 3.2-1.** Schematic representation of dsDNA recognition by Invader probes containing non-nucleosidic bulges and the chemical modifications utilized for this approach.



The double-stranded Invader probes display significantly lower  $T_m$ 's than the corresponding duplexes between probe strands and cDNA, verifying our previous observations that +1 interstrand zipper arrangements of **X** monomers result in destabilization of DNA duplexes (e.g., compare  $T_m$  of **ON3:ON4** vs **ON3:cDNA** and **ON4:cDNA**). Invader probes, in which two  $C_9$  bulges either are present on the same strand or on two different strands but at the same terminus, are particularly destabilized.

The thermodynamic dsDNA recognition potential of a specific Invader probe can be estimated by the term *thermal advantage*, given as  $TA = T_m(5'\text{-Inv:cDNA}) + T_m(3'\text{-Inv:cDNA}) - T_m(\text{Invader probe}) - T_m(\text{dsDNA target})$ , with large positive values signifying a strongly activated probe. Invader probe **ON1:ON2**, which is based on a traditional probe architecture without bulges, has a prominent  $TA$  value of 28.5 °C due to the high  $T_m$ 's of the probe:cDNA duplexes and the low  $T_m$ 's of the probe duplex and dsDNA target (compare  $T_m$  of **ON1:cDNA**, **ON2:cDNA**, **ON1:ON2**, and the unmodified reference duplex, Table 3.2-1). Invader probes with a single  $C_9$  bulge (e.g., **ON3:ON2**) display similar or slightly higher  $TAs$  since the bulge destabilizes probe:cDNA and Invader probe duplexes to similar degrees. Probes **ON3:ON4** and **ON5:ON6**, which have two  $C_9$  bulges at one of the termini, display significantly increased dsDNA recognition potential ( $TAs > 35.5$  °C) because the probe duplexes are very strongly destabilized, while probe-target duplexes only are mildly destabilized; presumably, this is because two

**Table 3.2-1.** Thermal denaturation temperatures ( $T_m$ 's) and thermal advantages ( $TA$ 's) for modified DNA duplexes.

Invader probe	Sequence	$T_m$ [ $\Delta T_m$ ] ( $^{\circ}\text{C}$ )			
		Probe duplex	5'-Inv:cDNA	3'-Inv:cDNA	$TA$ ( $^{\circ}\text{C}$ )
1 2	5'-GGTAX AX ATAGGC 3'-CCATA XA XATCCG	45.0 [+7.5]	55.5 [+18.0]	55.5 [+18.0]	28.5
3 2	5'-GG <sup>9</sup> TAX AX ATAGGC 3'-CC ATA XA XATCCG	31.5 [-6.0]	44.0 [+6.5]	55.5 [+18.0]	30.5
5 2	5'-GGTAX AX ATAG <sup>9</sup> GC 3'-CCATA XA XATC CG	33.0 [-4.5]	44.5 [+7.0]	55.5 [+18.0]	29.5
1 4	5'-GG TAX AX ATAGGC 3'-CC <sup>9</sup> ATA XA XATCCG	35.0 [-2.5]	55.5 [+18.0]	46.5 [+9.0]	29.5
1 6	5'-GGTAX AX ATAG GC 3'-CCATA XA XATC <sup>9</sup> CG	28.5 [-9.0]	55.5 [+18.0]	43.5 [+6.0]	33.0
3 4	5'-GG <sup>9</sup> TAX AX ATAGGC 3'-CC <sup>9</sup> ATA XA XATCCG	<15	44.0 [+6.5]	46.5 [+9.0]	>38
5 6	5'-GGTAX AX ATAG <sup>9</sup> GC 3'-CCATA XA XATC <sup>9</sup> CG	<15	44.5 [+7.0]	43.5 [+6.0]	>35.5
7 2	5'-GG <sup>9</sup> TAX AX ATAG <sup>9</sup> GC 3'-CC ATA XA XATC CG	<15	<15	55.5 [+18.0]	-
1 8	5'-GG TAX AX ATAG GC 3'-CC <sup>9</sup> ATA XA XATC <sup>9</sup> CG	<15	55.5 [+18.0]	<15	-
3 6	5'-GG <sup>9</sup> TAX AX ATAG GC 3'-CC ATA XA XATC <sup>9</sup> CG	28.5 [-9.0]	44.0 [+6.5]	43.5 [+6.0]	21.5
5 4	5'-GG TAX AX ATAG <sup>9</sup> GC 3'-CC <sup>9</sup> ATA XA XATC CG	32.5 [-5.0]	44.5 [+7.0]	46.5 [+9.0]	21.0

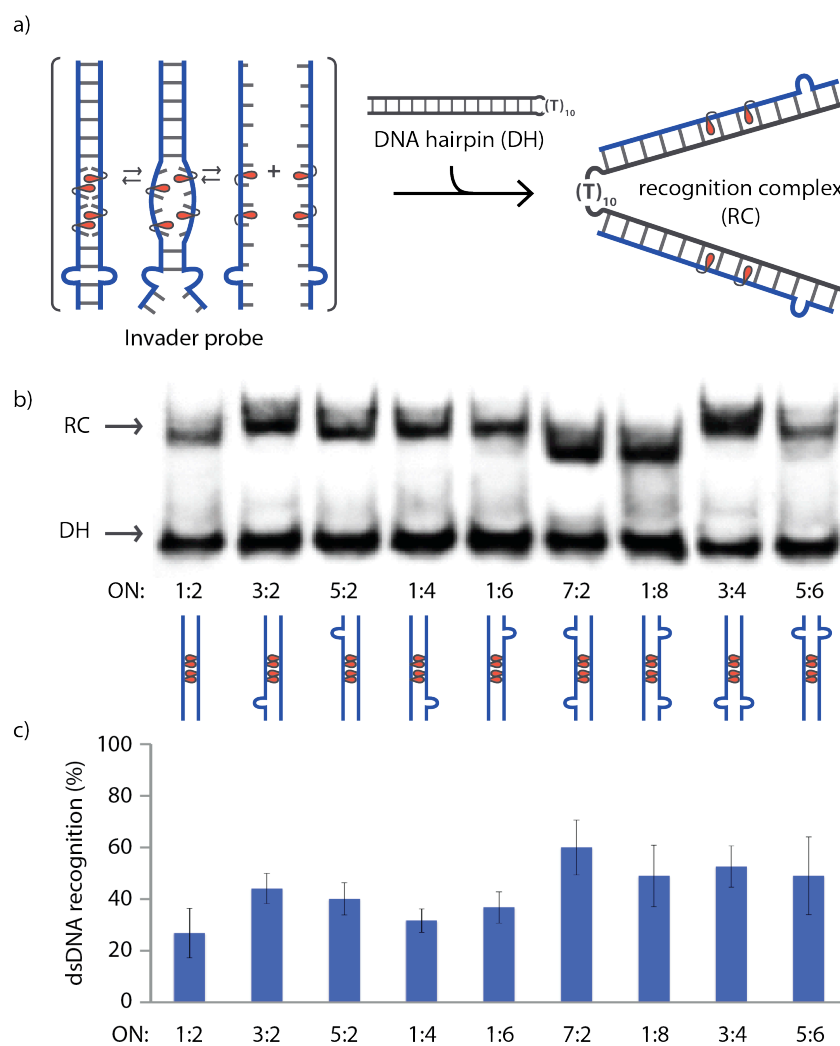
<sup>a</sup>  $\Delta T_m$  is calculated relative to the corresponding unmodified dsDNA ( $T_m = 37.5$   $^{\circ}\text{C}$ ; Thermal denaturation curves were recorded in medium salt phosphate buffer ( $[\text{Na}^+] = 110$  mM,  $[\text{Cl}^-] = 100$  mM, pH 7.0 ( $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ ),  $[\text{EDTA}] = 0.2$  mM) and each  $[\text{ON}] = 0.5$   $\mu\text{M}$ ; see main text for definition of  $TA$ . A = adenin-9-yl DNA monomer, C = cytosin-1-yl DNA monomer, G = guanin-9-yl DNA monomer, T = thymin-1-yl DNA monomer.

adjacent  $\text{C}_9$  bulges (as in probe duplexes) have a more detrimental effect on base-pairing cooperativity than two separate  $\text{C}_9$  bulges (as in probe-target duplexes). Along these

lines, Invader probes with two C<sub>9</sub> bulges on separate strands and termini (**ON3:ON6** and **ON5:ON4**) display lower dsDNA recognition potential because the probe duplexes are not destabilized as much. *T<sub>A</sub>* values for Invader probes with two C<sub>9</sub> bulges on one strand (**ON7:ON2** and **ON1:ON8**) could not be determined due to the low stability of probe-target duplexes. *T<sub>A</sub>* values provide an estimate for the thermodynamic dsDNA recognition potential of specific Invader probes.<sup>28</sup> However, other factors, such as the *T<sub>m</sub>*'s of the Invader probes relative to the experimental temperatures used, likely influence recognition efficiency and kinetics. To gain a deeper understanding of these factors, an electrophoretic mobility shift assay (EMSA) was performed. Pre-annealed Invader probes were incubated with DNA hairpin **DH1**, in which the double-stranded target region is linked via a decameric thymidine loop (Figure 3.2-2a). Recognition of this model target results in the formation of a recognition complex, which is observed as a slower moving band on non-denaturing polyacrylamide gel electrophoresis (Figure 3.2-2b). A 200-fold molar excess of Invader probes was incubated with **DH1** at 8 °C for 17 h. At these conditions, the conventional Invader probe **ON1:ON2** only results in ~22% recognition, whereas single bulge Invader probes effect more efficient recognition (30-42%) (Figures 3.2-2b and 2c and Table 3.5-2). Invader probes with two C<sub>9</sub> bulges at one terminus (**ON3:ON4** and **ON5:ON6**) or two C<sub>9</sub> bulges on the same strand (**ON1:ON8** and **ON7:ON2**) recognize the dsDNA target even more efficiently (41-55%). The recognition complexes formed with **ON1:ON8** and **ON7:ON2** have slightly greater electrophoretic mobilities than those formed with other Invader probes. This is almost certainly because binary, rather than ternary, recognition complexes are formed, as **ON7** and **ON8** have very low cDNA affinity (*T<sub>m</sub>* < 15 °C for **ON7/ON8:cDNA**, Table 3.2-1 – see also Figure

3.5-2). Invader probes with two C<sub>9</sub> bulges on separate strands and termini (**ON3:ON6** and **ON5:ON4**) do not result in detectable dsDNA recognition, suggesting that the process is energetically unfavorable (Figure 3.5-3). For similar reasons, Invader probes with three or four bulge insertions (Table 3.5-3) also do not result in detectable dsDNA recognition (Figure 3.5-3).

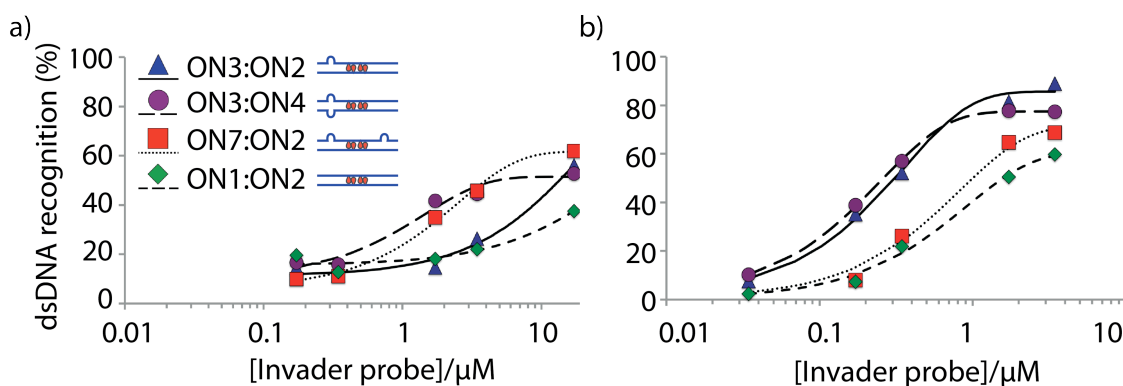
While conventional Invader strands **ON1** and **ON2** result in some recognition of **DH1** when used as single-stranded probes, none of the C<sub>9</sub>-containing single-stranded Invader probes result in significant recognition of **DH1** (Figure 3.5-4). Interestingly, **ON7:ON2** results in greater dsDNA recognition than single-stranded **ON2**, indicating that the presence **ON7** in **ON7:ON2** is advantageous despite its low cDNA affinity (Figure 3.5-2).



**Figure 3.2-2.** (a) Schematic representation of the electrophoretic mobility shift assay used to evaluate dsDNA recognition of Invader probes. (b) Representative electrophoretograms for recognition of model dsDNA target **DH1** ( $34.4 \mu\text{M}$ ) by different Invader probes ( $6.88 \mu\text{M}$ ) at  $8 \text{ }^\circ\text{C}$ . (c) Histogram showing the average of three experiments; error bars represent standard deviation. DIG-labeled **DH1** (5'-GGTATATATAGGC-T<sub>10</sub>-GCCTATATATAACC-3') was incubated with pre-annealed Invader probe in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) for 17 h.

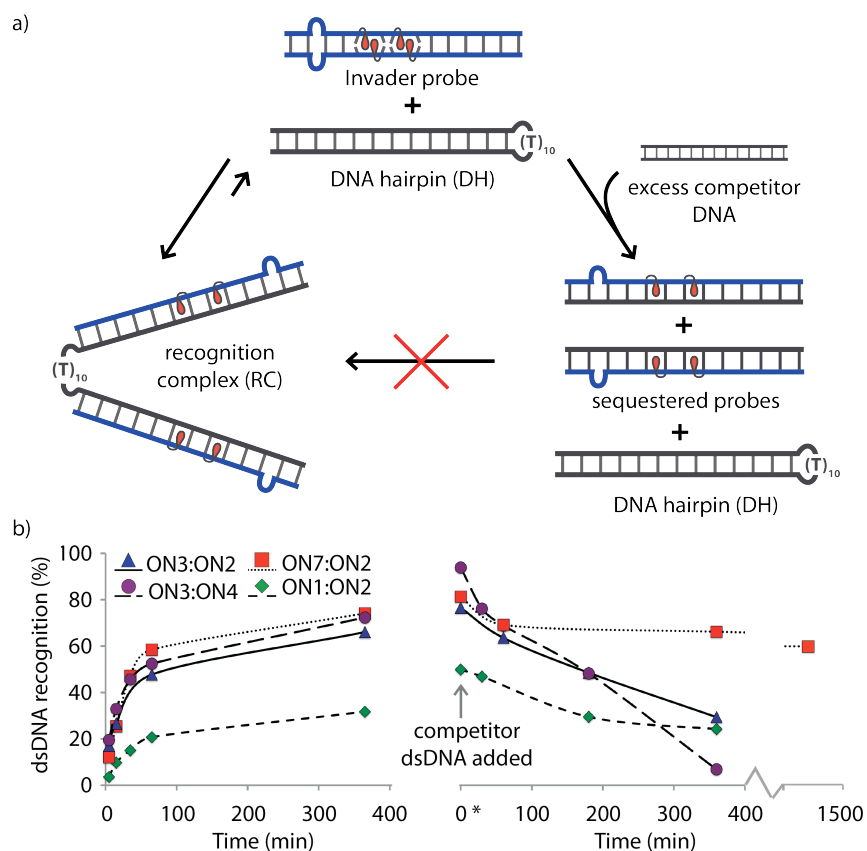
Next, dose-response EMSAs were performed at  $8 \text{ }^\circ\text{C}$  or ambient temperature ( $22 \text{ }^\circ\text{C}$ ) for representative Invader probes (Figure 3.2-3). At ambient temperature, single bulge Invader **ON3:ON2** and Invader **ON3:ON4**, which has two bulges at the same terminus,

display similar dose-response profiles and sub-micromolar  $C_{50}$  values (i.e. the probe concentration resulting in 50% recognition of **DH1**;  $\sim 0.3 \mu\text{M}$ , Table 3.2-2). Conventional Invader probe **ON1:ON2** has a significantly higher  $C_{50}$  value at  $\sim 1.6 \mu\text{M}$ , whereas Invader **ON7:ON2**, having two bulges on the same strand has an intermediate  $C_{50}$  value of  $\sim 1.0 \mu\text{M}$ . Incubation at  $8^\circ\text{C}$  results in slightly different dose-response trends (compare Figures 3.2-3a and 3b). Thus, double bulge Invaders **ON3:ON4** and **ON7:ON2** display lower  $C_{30}$  values, than single bulge Invader **ON3:ON2** or conventional Invader **ON1:ON2**. These observations suggest that probes with large thermodynamic driving forces result in more efficient dsDNA recognition at higher experimental temperatures, whereas probes with low  $T_m$ 's result in efficient dsDNA recognition at low experimental temperatures where breathing of base-pairs is minimal. Probes with low  $T_m$ 's are more likely to be partially or even fully dissociated at low experimental temperatures, thereby enabling the Watson-Crick face of the probe strands to be available for nucleation with DNA targets.



**Figure 3.2-3.** Dose-response curves for recognition of dsDNA by Invader probes **ON3:ON2**, **ON3:ON4**, **ON7:ON2**, and **ON1:ON2** at (a)  $8^\circ\text{C}$  or (b)  $22^\circ\text{C}$ . Experimental conditions otherwise as in Figure 3.2-2.

The kinetics of Invader-mediated dsDNA recognition were determined in experiments where a 100-fold molar probe excess was incubated with **DH1** at 22 °C (Figure 3.2-4). Interestingly, all of the bulge-containing Invaders display considerably faster recognition kinetics than the conventional Invader probe **ON1:ON2** (for pseudo-first order rate constants, see Table 3.2-2). Single bulge Invader **ON3:ON2** and Invader **ON3:ON4**, which has two bulges at the same terminus, exhibit 2.3 and 2.7-fold faster kinetics, respectively, while Invader **ON7:ON2**, which has two bulges in the same strand, displays 4.1-fold faster kinetics. One possible interpretation of these trends is that bulges promote partial or even full denaturation of the double-stranded Invader probes, thus revealing a Watson-Crick face for faster binding with the target DNA.



**Figure 3.2-4.** a) Assays used to determine kinetic parameters for dsDNA recognition using representative Invader probes. b) *Left:* Kinetics of recognition complex formation at 22 °C using 100-fold molar excess of Invader probes. *Right:* Competitive dissociation kinetics of recognition complexes between DNA hairpins and Invader probes (for representative gel electrophoretograms, see Figure 3.5-6). 100-fold molar excess of Invader probes (3.44  $\mu\text{M}$ ) was incubated with **DH1** for 24 h, followed by addition of a 1000-fold molar excess of linear competitor dsDNA target (34.4  $\mu\text{M}$  – sequence: 5'-GGTATATATAGGC:3'-CCATATATATCCG).  $T = 22$  °C.



**Table 3.2-2.** Summary of parameters for dsDNA recognition by representative Invader probes.

Probe	Sequence	$C_{50}^a$ 22 °C ( $\mu$ M)	$C_{30}^a$ 8 °C ( $\mu$ M)	$t_{50}^b$ (min)	$k_{obs}^c$ ( $\text{min}^{-1}$ )	$k_{rel}$
<b>ON1:ON2</b>	5'-GGTAX AX ATAGGC 3'-CCATA XA XATCCG	1.6	9.7	-	$3.1 \times 10^{-3}$	1
<b>ON3:ON2</b>	5'-GG <sup>9</sup> TAX AX ATAGGC 3'-CC <sup>9</sup> ATA XA XATCCG	0.3	5.2	110	$7.2 \times 10^{-3}$	2.3
<b>ON3:ON4</b>	5'-GG <sup>9</sup> TAX AX ATAGGC 3'-CC <sup>9</sup> ATA XA XATCCG	0.3	1.0	42	$8.4 \times 10^{-3}$	2.7
<b>ON7:ON2</b>	5'-GG <sup>9</sup> TAX AX ATAG <sup>9</sup> GC 3'-CC <sup>9</sup> ATA XA XATC <sup>9</sup> CG	1.0	1.5	41	$1.3 \times 10^{-2}$	4.1

<sup>a</sup> Calculated from dose-response curves shown in Figure 3.2-3. <sup>b</sup>  $t_{50}$  = time to reach 50% dsDNA recognition at 22 °C as calculated from time-course experiments shown in Figure 3.2-4. <sup>c</sup> Calculated from the pseudo-first order plots shown in Figure 3.5-5.

The dissociation kinetics of the different recognition complexes were evaluated using a competition assay,<sup>25</sup> in which preformed recognition complexes (24 h incubation at ambient temperature) were challenged with a 1000-fold excess of a linear competitor dsDNA target (Figure 3.2-4). Dissociating Invader strands bind to this competitor target,<sup>17</sup> resulting in formation of a faster moving band in non-denaturing gel electrophoresis consistent with re-formation of **DH1**. Approximately 25% of the recognition complexes between **DH1** and **ON1:ON2** or **ON3:ON2** remain intact after 6 h. The recognition complex between **DH1** and **ON3:ON4**, which contains two C<sub>9</sub> bulges at one terminus, undergoes rapid dissociation (>90% within 6 h), presumably due to the low cDNA affinity of **ON3** and **ON4**. Surprisingly, the recognition complex between **DH1** and **ON7:ON2**, featuring two C<sub>9</sub> bulges on one of the probe strands, is remarkably stable (~60% of complex intact after 24 h). This construct is unique, as only one probe strand (i.e., **ON2**) is firmly bound to the target in the recognition complex (Figure 3.5-2).

Given the slower dissociation of **DH1:(ON7):ON2** relative to **DH1:ON1:ON2**, it is clear that the unbound **ON7** plays a role in slowing down dissociation, possibly through weak/transient binding to the binary complex and/or very weak affinity toward the target competitor strand.

### **3.3 Conclusions**

In conclusion, we have demonstrated that Invader probes with appropriately positioned non-nucleosidic C<sub>9</sub> bulge insertions display significantly faster, more efficient, and longer-lasting recognition of mixed-sequence dsDNA targets than conventional Invader probes. The robustness, simplicity of design and synthesis, render these probes amenable to a variety of applications in molecular diagnostics and DNA nanotechnology.

### **3.4 Acknowledgements**

We appreciate financial support from award number GM088697 from the National Institute of General Medical Sciences, National Institutes of Health; awards IF13-001 and IF14-012 from the Higher Education Research Council, Idaho State Board of Education; and Minitube of America. We thank Dr. Lee Deobald (Murdock Mass Spectrometry Center, Univ. Idaho) for assistance with MS analysis and Dr. Carolyn J. Hovde (Food Science, Univ. Idaho) for access to gel documentation stations.

### **3.5. Supplementary data**

#### *Protocol - synthesis and purification of ONs*

Modified ONs were synthesized on an automated DNA synthesizer (0.2  $\mu$ mol scale) using a long chain alkyl amine controlled pore glass (LCAA-CPG) solid support with a

pore size of 500 Å. The corresponding phosphoramidite of monomer X was prepared as previously described<sup>29</sup> and incorporated into ONs via hand-couplings (0.05 M in acetonitrile, using 0.01 M 4,5-diccanoimidazole as the activators (15 min)) with extended oxidation (45 s). The nonyl linker was incorporated in a similar manner using the commercially available DMT-nonane diol phosphoramidite (ChemGenes). Treatment with 32% ammonia (55 °C, 17 h) facilitated deprotection and cleavage from solid support. DMT-protected ONs were purified via ion-pair reverse phase HPLC (XTerra MS C18 column: 0.05 M triethyl ammonium acetate and acetonitrile gradient) followed by detritylation (80% acetic acid, 20 min) and precipitation (NaOAc, NaClO<sub>4</sub>, acetone, -18 °C, 16 h). The purity and identity of synthesized ONs were verified using analytical HPLC (>85% purity) and MALDI-MS analysis (Tables 3.5-1) recorded on a Quadrupole Time-of-Flight (Q-TOF) mass spectrometer with anthranilic acid or 2,4,6-trihydroxyacetophenone matrix.

#### *Protocol - thermal denaturation experiments*

The concentrations of ONs were estimated using the following extinction coefficients (OD<sub>260</sub>/μmol): G (12.01), A (15.20), T (8.40), C (7.05) and pyrene (22.4). Thermal denaturation temperatures were calculated as the first-derivative maximum of A<sub>260</sub> vs T curves. ONs (0.5 μM) were annealed (85 °C, 2 min) in medium salt buffer ([Na<sup>+</sup>] = 110 mM, [Cl<sup>-</sup>] = 100 mM, pH 7.0 (NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>), [EDTA] = 0.2 mM) and subsequent cooling to the starting temperature. The experimental temperature ranged from 5 °C to at least 15 °C above the T<sub>m</sub>, with the T<sub>m</sub> determined as the average of two experiments within ±1.0 °C.

*Protocol - electrophoretic mobility shift assay*

DNA hairpins were obtained from commercial sources and used without further purification. Hairpins were labeled using the 2<sup>nd</sup> generation DIG Gel Shift Kit (Roche Applied Bioscience). Briefly, 11-digoxigenin-ddUTP was incorporated at the 3'-end of the hairpin (100 pmol) using a recombinant DNA terminal transferase. The reaction mixture was quenched through addition of EDTA (0.05 M), diluted to 68.8 nM, and used without further processing. The recognition experiments were conducted essentially as previously reported.<sup>29</sup> Thus, Invader probes (variable concentration) were annealed (90 °C for 2 min, followed by cooling to room temperature) and subsequently incubated with DIG-labeled DNA hairpins (34.4 nM final concentration) in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) at either 8 °C ± 2 °C or ambient temperature 22 °C ± 2 °C for a specified period. For time-course experiments (Figure 3.2-4 and 3.5-6), aliquots were taken at specific time points, flash frozen in liquid N<sub>2</sub>, and stored at -76 °C until analysis. Loading dye (6X) was added and the reaction mixtures were loaded onto 12% non-denaturing TBE-PAGE (45 mM tris-borate, 1 mM EDTA; acrylamide:bisacrylamide (19:1)). Electrophoresis was performed using constant voltage (70 V) at ~4 °C for 1.5 h. Bands were blotted onto positively charged nylon membranes (100 V, 30 min, ~4 °C) and cross-linked through exposure to UV light (254 nm, 5 × 15 watt bulbs, 3 min). Membranes were incubated with anti-digoxigenin-alkaline phosphatase F<sub>ab</sub> fragments as recommended by manufacturer, and transferred to a hybridization jacket. Membranes were incubated with the chemiluminescence substrate (CSPD) for 10 min at 37 °C, and chemiluminescence was captured on X-ray films. Digital images of developed X-ray

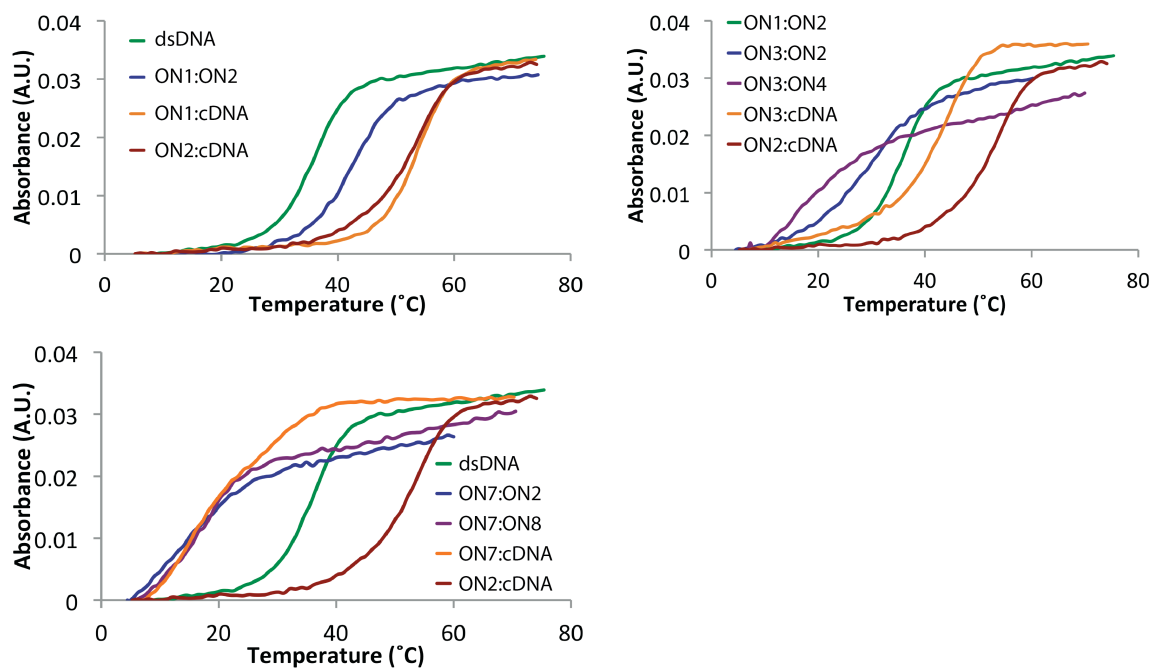
films were obtained using a Fluor-S MultiImager and quantified using appropriate software (Quantity One). The percentage of dsDNA recognition was calculated as the intensity ratio between the recognition complex band and the total lane. Unless otherwise noted, an average of three independent experiments is reported along with standard deviations ( $\pm$ ).

*Definition - interstrand zipper arrangement*

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.

**Table 3.5-1.** MALDI-MS of modified ONs.

ON	Sequence	Observed $m/z$ [M+H] <sup>+</sup>	Calculated $m/z$ [M+H] <sup>+</sup>
ON3	5'- GG <u>9</u> TATATATAGGC	4667.5	4667.6
ON4	3'- CC <u>9</u> ATATATATCCG	4547.4	4547.5
ON5	5'- GGTATATATAG <u>9</u> GC	4666.7	4667.6
ON6	3'- CCATATATATC <u>9</u> CG	4547.4	4547.5
ON7	5'- GG <u>9</u> TATATATAG <u>9</u> GC	4889.2	4889.7
ON8	3'- CC <u>9</u> ATATATATC <u>9</u> CG	4769.1	4769.6



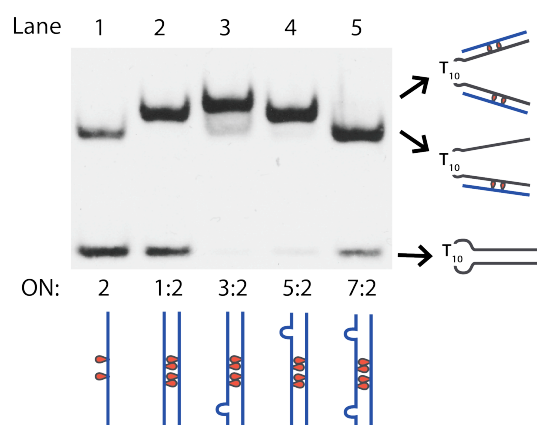
**Figure 3.5-1.** Representative thermal denaturation curves of Invader probes, duplexes between individual probe strands and cDNA, and unmodified reference duplexes. For experimental conditions, see Table 3.2-1.

**Table 3.5-2.** Degree of recognition of DNA hairpin **DH1** using a 200-fold molar excess of different Invader probes at two different temperatures.

ON	Sequence	Rec <sub>200</sub> (%)	
		8 °C <sup>a</sup>	22 °C <sup>b</sup>
1 2	5'-GGTAX AX ATAGGC 3'-CCATA XA XATCCG	22	>90
3 2	5'-GG <sup>9</sup> TAX AX ATAGGC 3'-CC <sup>9</sup> ATA XA XATCCG	38	>90
5 2	5'-GGTAX AX ATAG <sup>9</sup> GC 3'-CCATA XA XATC <sup>9</sup> CG	37	>90
1 4	5'-GG TAX AX ATAGGC 3'-CC <sup>9</sup> ATA XA XATCCG	30	>90
1 6	5'-GGTAX AX ATAG GC 3'-CCATA XA XATC <sup>9</sup> CG	34	>90
3 4	5'-GG <sup>9</sup> TAX AX ATAGGC 3'-CC <sup>9</sup> ATA XA XATCCG	51	>90
5 6	5'-GGTAX AX ATAG <sup>9</sup> GC 3'-CCATA XA XATC <sup>9</sup> CG	42	>90
7 2	5'-GG <sup>9</sup> TAX AX ATAG <sup>9</sup> GC 3'-CC <sup>9</sup> ATA XA XATC <sup>9</sup> CG	55	>90
1 8	5'-GG TAX AX ATAG GC 3'-CC <sup>9</sup> ATA XA XATC <sup>9</sup> CG	43	>90
3 6	5'-GG <sup>9</sup> TAX AX ATAG GC 3'-CC <sup>9</sup> ATA XA XATC <sup>9</sup> CG	<10	20
5 4	5'-GG TAX AX ATAG <sup>9</sup> GC 3'-CC <sup>9</sup> ATA XA XATC <sup>9</sup> CG	<10	43

<sup>a</sup> Data shown in Figure 3.2-2.

<sup>b</sup> Experiments performed in duplicate.



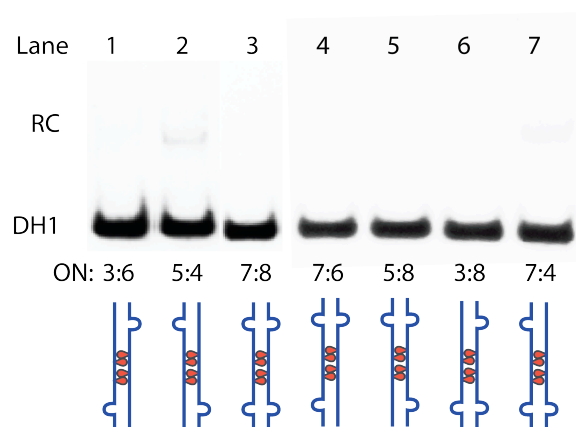
**Figure 3.5-2.** The structure of recognition complexes formed upon incubation of different Invaders strands/probes and DNA hairpin **DH1** at 22 °C. Invader probes used at 100-fold excess (3.44  $\mu$ M). Experimental conditions are otherwise as described in Figure 3.2-2. The greater mobility of the recognition complex between **ON7:ON2** and **DH1** relative to the complexes involving **ON1:ON2**, **ON3:ON2** and **ON5:ON2**, and the similar mobility relative to the complex involving **ON2**, strongly suggests that a binary complex is formed. Nonetheless, recognition of **DH1** is more efficient with **ON7:ON2** than **ON2**, indicating that **ON7** – despite its low cDNA affinity – plays a role in the recognition process, possibly through weak/transient binding to the binary complex, thus preventing re-formation of **DH1**.



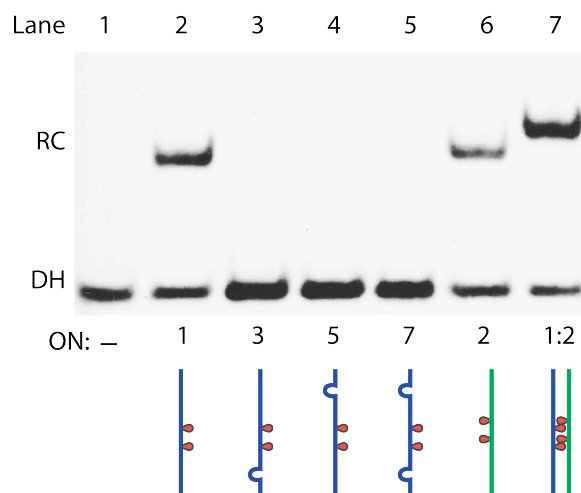
**Table 3.5-3.** Thermal denaturation temperatures ( $T_m$ 's) of additional Invader probes.<sup>a</sup>

Invader probe	Sequence	$T_m$ [ $\Delta T_m$ ] (°C)		
		Probe duplex	5'-Inv:cDNA	3'-Inv:cDNA
1 2	5'-GGTAXIAXIATAGGC 3'-CCATAIXAXATCCG	45.0	55.5	55.5
7 4	5'-GG <sup>9</sup> TAXIAXIATAG <sup>9</sup> GC 3'-CC <sup>9</sup> ATAIXAXATC <sup>9</sup> CG	<15	<15	46.5
7 6	5'-GG <sup>9</sup> TAXIAXIATAG <sup>9</sup> GC 3'-CC <sup>9</sup> ATAIXAXATC <sup>9</sup> CG	<15	<15	43.5
3 8	5'-GG <sup>9</sup> TAXIAXIATAG <sup>9</sup> GC 3'-CC <sup>9</sup> ATAIXAXATC <sup>9</sup> CG	<15	44.0	<15
5 8	5'-GG <sup>9</sup> TAXIAXIATAG <sup>9</sup> GC 3'-CC <sup>9</sup> ATAIXAXATC <sup>9</sup> CG	<15	44.5	<15
7 8	5'-GG <sup>9</sup> TAXIAXIATAG <sup>9</sup> GC 3'-CC <sup>9</sup> ATAIXAXATC <sup>9</sup> CG	<15	<15	<15

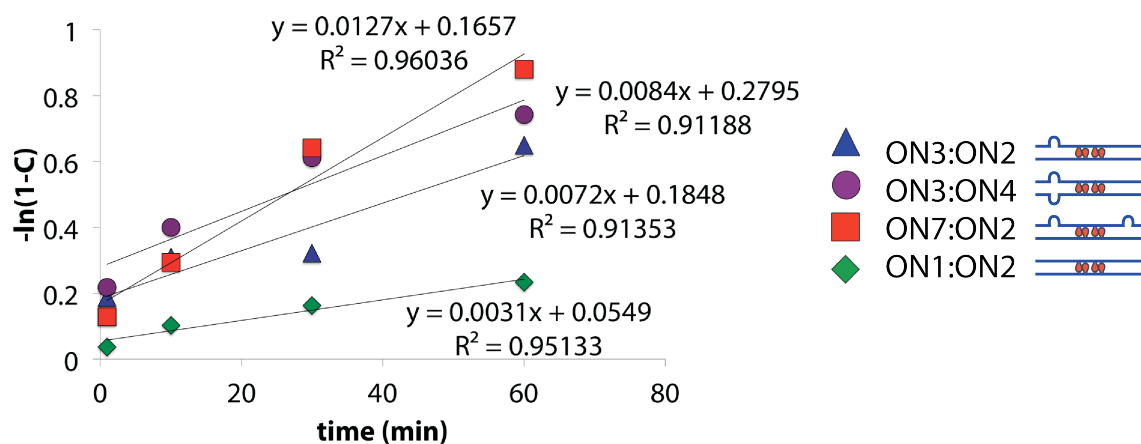
<sup>a</sup>The corresponding unmodified dsDNA has a  $T_m = 37.5$  °C.



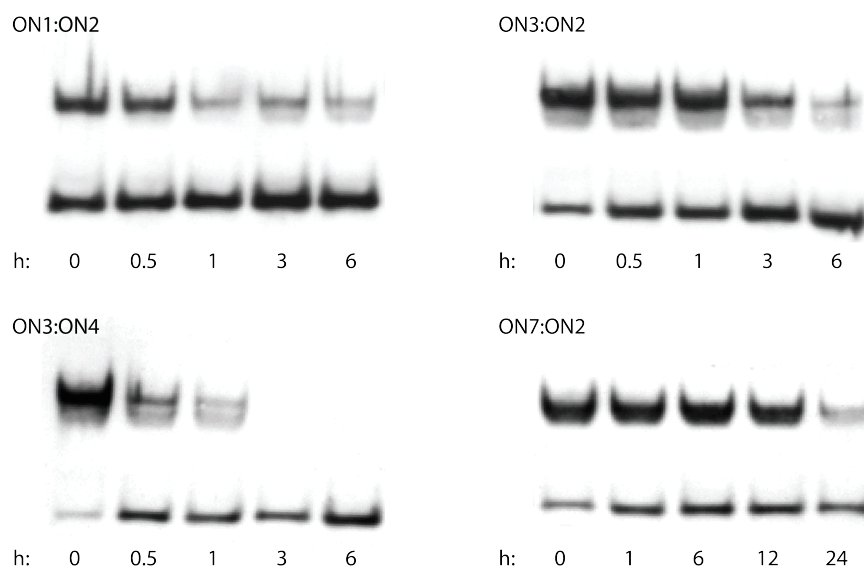
**Figure 3.5-3.** Recognition of model dsDNA target **DH1** using 200-fold molar excess ( $6.88 \mu\text{M}$ ) of different Invader probes at  $8^\circ\text{C}$ . Image is a composite of two electrophoretograms (lanes 1-3 and lanes 4-7). Experiments were performed in duplicate. Very similar results (not shown) were observed at  $22^\circ\text{C}$ .



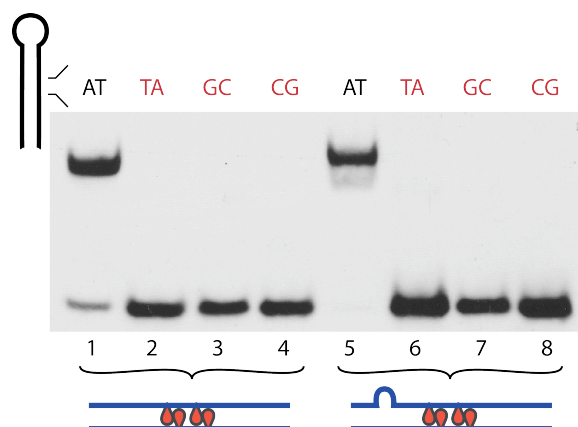
**Figure 3.5-4.** Recognition of **DH1** by individual Invader probe strands. Each lane includes **DH1** incubated with 200-fold molar excess ( $6.88 \mu\text{M}$ ) of single-stranded probes or conventional Invader probe **ON1:ON2** at  $22^\circ\text{C}$ . Experiments were performed in duplicate. Single-stranded probes, **ON1** and **ON2**, result in  $\sim 50\%$  recognition of **DH1**, while **ON1:ON2** results in  $\sim 70\%$  recognition, emphasizing the need for both strands for maximal recognition.



**Figure 3.5-5.** Pseudo-first order rate plots for dsDNA recognition by Invader probes at 22 °C. Raw time profiles shown in Figure 3.2-4.



**Figure 3.5-6.** Representative electrophoretograms from the competitive dissociation assay shown in Figure 3.2-4.



**Figure 3.5-7.** Specificity of Invader-mediated dsDNA recognition. Fully base-paired DNA hairpins containing a single base-pair mismatch (red) relative to Invader probes **ON1:ON2** (left) and **ON3:ON2** (right) were used. Base-pairs above electrophoretograms correspond to **B:B'** in: 5'-GGTATBTATAGGC-T<sub>10</sub>-GCCTATAB'B'ATACC. A 200-fold molar excess of Invader probes (6.88  $\mu$ M) were used. Experiments were performed in duplicate at 22 °C (shown) or 37 °C (results not shown – identical observations). Excellent discrimination of the singly mismatched DNA hairpins was observed with both Invader probes.

### 3.6. References and notes

1. Besch, R.; Giovannangeli, C.; Degitz, K. *Curr. Drug Targets* **2004**, *5*, 691.
2. Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. J. *Mol. BioSyst.* **2006**, *2*, 551.
3. Ackermann, D.; Famulok, M. *Nucleic Acids Res.* **2013**, *41*, 4729.
4. Aiba, Y.; Sumaoka, J.; Komiyama, M. *Chem. Soc. Rev.* **2011**, *40*, 5657.
5. Nielsen, P. E. *Chem. Biodiv.* **2010**, *7*, 786.
6. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.*, **2003**, *13*, 284.
7. Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem.*, **2013**, *21*, 6101.
8. Gai, T.; Gersbach, C. A.; Barbas III, C. F. *Trends Biotechnol.*, **2013**, *31*, 397.
9. Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, *36*, 5123.

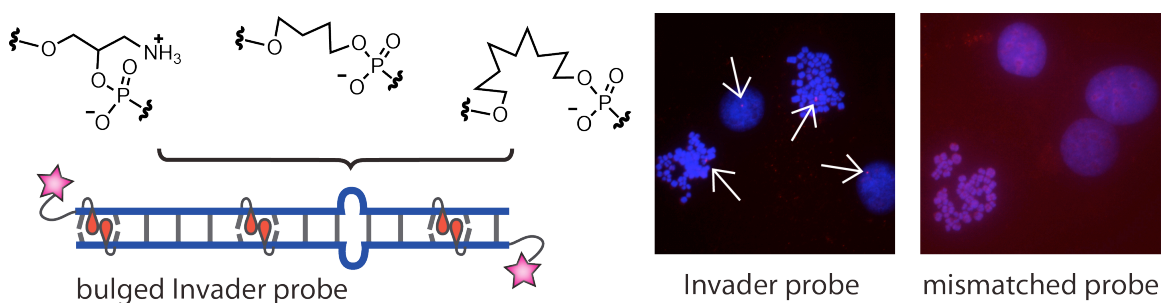
10. Nielsen, P. E., Egholm, M., Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
11. Bahal, R.; Sahu, B.; Rapireddy, S.; Lee, C-M.; Ly, D. H. *ChemBioChem* **2012**, *13*, 56.
12. Chen, S. X.; Zhange, D. Y.; Seelig, G. *Nature Chem.* **2013**, *5*, 782.
13. Kutuyavin, I. V.; Rhinehart, R. L.; Lukhtanov, E. A.; Gorn, V. V.; Meyer Jr., R. B.; Gamper Jr., H. B. *Biochemistry* **1996**, *35*, 11170.
14. Smolina, I. V.; Demidov, V. V. *Chem. Bio.* **2003**, *10*, 591.
15. Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11804.
16. Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corradini, R.; Sforza, S.; Komiyama, M. *Nucleic Acids Res.* **2008**, *36*, 1464.
17. Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2010**, *8*, 2028.
18. 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.
19. Crothers, D. M. *Biopolymers* **1968**, *6*, 575.
20. Tsai, C.; Jain, S. C.; Sobell, H. M. *J. Mol. Biol.* **1977**, *114*, 301.
21. Williams, L. D.; Egli, M.; Gao, Q.; Rich, A. In: *Structure and Function, Volume 1: Nucleic Acids*, 1992, pp 107-125 (Eds: Sarma R. H.; Sarma, M. H.). Adenine press.
22. Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, *78*, 12040.
23. Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2014**, *12*, 7758.
24. Anderson, B. A.; Onley, J. J.; Hrdlicka, P. J. *J. Org. Chem.* **2015**, *80*, 5395.

25. Guenther, D. C.; Anderson, G. H.; Karmakar, S.; Anderson, B. A.; Didion, B. A.; Guo, W.; Verstegen, J. P.; Hrdlicka, P. J. *Chem. Sci.* **2015**, *6*, 5006.
26. Braunlin, W.; Volker, J.; Plum, G. E.; Breslauer, K. J. *Biopolymers* **2013**, *99*, 408.
27. Pyshnyi, D. V.; Lomzov, A. A.; Pyshnaya, I. A.; Ivanova, E. M. *J. Biomol. Struct. Dyn.* **2006**, *23*, 567.
28. It was not possible to obtain thermodynamic data via the van't Hoff method, as denaturation curves were lacking clear base lines.
29. Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Jensen, T.; Nielsen, P.; Hrdlicka, P. *J. J. Org. Chem.*, **2011**, *76*, 7119.

## CHAPTER 4: Optimization and application of bulged Invader probes for detection of mixed-sequence dsDNA

Dale C. Guenther, Allison Rowley, Saswata Karmakar and Patrick J. Hrdlicka <sup>†,\*</sup>

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



### Abstract

Hybridization-based probes for selective recognition of double-stranded DNA (dsDNA) offer great promise as fundamental research tools, diagnostic agents, and components for applications in nanotechnology. Double-stranded DNA probes, modified with +1-interstrand motifs of O<sup>2</sup>'-pyrene-functionalized RNA monomers, are energetically activated for this process. Additional modification with non-nucleosidic bulge inserts generates labile probe duplexes for specific and efficient recognition of dsDNA, which is demonstrated by targeting linear, hairpin, and chromosomal DNA in nanomolar to low micromolar concentrations. Fluorescently labeled bulged Invader probes are used for gender-specific DNA detection in a non-denaturing fluorescent in situ hybridization assay, and hint at the potential for much broader diagnostic applications.

## 4.1 Introduction

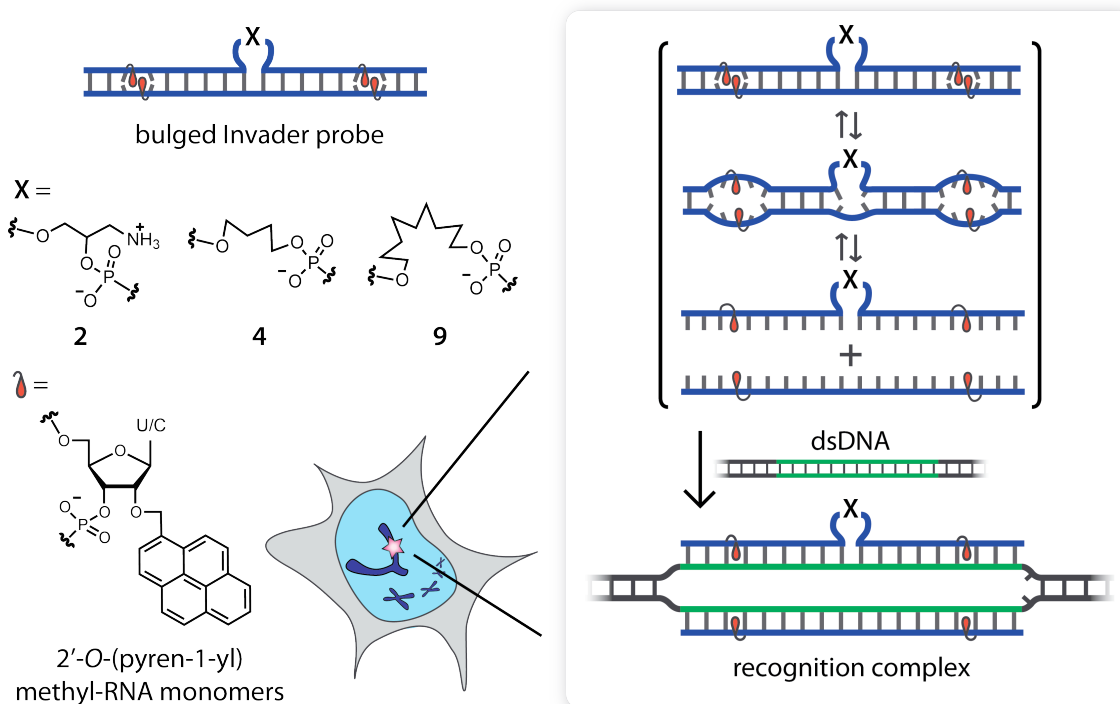
Development of modified oligonucleotide probes capable of recognizing double-strand DNA (dsDNA) has been a long-standing goal, which is motivated by the ability to detect, regulate, or manipulate DNA.<sup>1-5</sup> Probes must be able to either bind via extrahelical contacts or invade the target duplex in order to bind via Watson-Crick base pairs. Triplex forming oligonucleotides (TFOs),<sup>1,6</sup> peptide nucleic acids (PNAs)<sup>4,5,7</sup> or polyamides are examples of the former approach.<sup>8,9</sup> For stable binding, TFOs necessitate long polypurine/pyrimidine stretches, which restricts the number of potential target sites. Polyamides recognize short regions (<8 bp) of dsDNA, which precludes targeting of unique regions in genomic contexts. Conformationally restricted  $\gamma$ -PNAs, on the other hand, display very high affinity toward complementary DNA (cDNA) and are capable of accessing the Watson-Crick face of dsDNA target regions via a duplex invasion mechanism, albeit only at non-physiological ionic strengths.<sup>10</sup> Duplex invasion by a single-stranded probe, such as  $\gamma$ -PNA, results in displacement of one strand and formation of an unbound D-loop. In order for both DNA strands to be recognized via a double-duplex invasion mechanism, double-stranded probes must be used, in which the two complementary probe strands have low affinity towards each, while exhibiting very high affinity towards cDNA. An advantage of structured probes is that they are likely to display increased specificity, as two complementary duplexes would have to be dissociated in favor of two mismatched duplexes, a thermodynamically unfavorable process.<sup>11</sup> However, double-duplex invasion also potentially results in slower kinetics, as the recognition elements are buried within the two duplexes, thus limiting nucleation and strand exchange.



One strategy towards engineering energetically activated double-stranded probes has been through the use of pseudocomplementary (pc) base pairs, such as 2-thiouracil and 2,6-diaminopurine.<sup>12</sup> Base pairs between these modified nucleobases are weak due to a steric clash between the sulfur of 2-thiouracil and one of the exocyclic amine groups of 2,6-diaminopurine. In contrast, base pairing with canonical nucleobases (i.e. 2-thiouridine and adenine or 2,6-diaminopurine and thymine) is more stable than the natural base pair, resulting in the formation of a thermodynamic gradient for recognition of dsDNA. Pseudocomplementary PNA probes have been shown to recognize mixed-sequence dsDNA at low ionic strength or under molecular crowding conditions.<sup>13,14</sup>

We recently introduced Invader probes, which are double-stranded oligonucleotide probes that are modified with 2'-intercalator-functionalized nucleotides. These probes also rely on an energetic difference between the double-stranded probe and the duplexes formed by the individual probe strands with complementary DNA (cDNA), but the energetic gradient is generated in a different manner. Rather than using pseudocomplementary base pairs, intercalator-functionalized nucleotide monomers such as 2'-*O*-(pyren-1-yl)methyl-RNA monomers (Figure 4.1-1) are arranged in +1-interstrand zipper motifs in the double-stranded probe,<sup>15</sup> which forces the intercalators into the same region, leading to violation of the nearest-neighbor principle<sup>16</sup> and localized unwinding and destabilization of the probe duplex (Figure 4.1-1).<sup>15,17-21</sup> However, once the individual strands are hybridized to cDNA (e.g. after invasion of a dsDNA target), the intercalators are no longer vying for the same space, and instead stabilize the resulting duplex through  $\pi$ - $\pi$  interactions with neighboring base pairs. We have explored different monomer chemistries and arrangements thereof to optimize the thermodynamic gradient for dsDNA

recognition,<sup>15,17-21</sup> and have recently focused on using Invader probes with non-nucleosidic bulges to accelerate recognition kinetics.<sup>22</sup> As part of these efforts, we studied 13mer bulged Invader probes containing nine-atom alkyl bulges that have the potential to disrupt base stacking in the double-stranded probe.<sup>23</sup> We hypothesized that Invader probes modified with bulges and interstrand zippers of intercalator-functionalized monomers would be more energetically labile, denature more easily, and allow for more facile nucleation and faster reaction kinetics. Indeed, these bulged Invader probes resulted in more efficient dsDNA recognition (>5-fold), faster recognition (>4-fold) and more stable recognition complexes (>24 h) relative to conventional probes.<sup>22</sup>



**Figure 4.2-1.** Illustration of bulged Invader probes and modifications used in the present study.

To study the effect of the bulge chemistry on the efficiency of dsDNA-recognition in more detail, we herein evaluate a library of Invader probes containing three different

bulge monomers (i.e., two-, four-, or nine-carbon bulges), which were inserted as single or triple insertions at the center of one or both probe strands.

## 4.2 Results and discussion

Thermal denaturation temperatures ( $T_m$ 's) were determined for probe duplexes and duplexes between individual probe strands and cDNA (Table 4.2-1). The conventional probe construct (**ON1:ON2**), which does not contain a bulge, shows minimal stabilization ( $\Delta T_m = +1.5$  °C) relative to the corresponding unmodified reference DNA duplex. Invader probes containing a single bulge, are increasingly destabilized as the length of the bulge monomer increases, result in a length dependent destabilization (e.g., compare  $T_m$ 's for **ON3:ON2** < **ON7:ON2** < **ON11:ON2**). Incorporation of bulges in the 'bottom' strand (i.e., 3'-CCACCAGTTXGATAGACCT) results in greater destabilization, indicating that the effect of bulges on the stability of the probe is somewhat sequence-dependent (e.g., compare **ON3:ON2** and **ON1:ON4**). Extending the length of the small bulges (i.e., **X = 2** or **4**), by incorporating two additional monomers on the same strand (triple insertion), results in additionally destabilized probe duplexes (e.g., compare **ON3:ON2** and **ON5:ON2**). However, insertion of additional large bulge monomers (i.e., **X = 9**) does not seem to destabilize the probe duplex further (e.g., compare **ON11:ON2** with **ON13:ON2**). Invader probes with small bulges on opposing strands have additive destabilizing effects (e.g., compare **ON3:ON2** and **ON1:ON4** with **ON3:ON4**), whereas similar constructs involving large bulges have no additional destabilizing effect. Expanding the centrally located bulged region with triple insertions on both strands only has an additionally destabilizing effect with the small bulge monomers (**2**; to a lesser degree **4**), but has no additional effect with the larger

monomers. This is likely because the hydrophobic stacking interactions are already broken with a ~12 atom linker. In concert, these results demonstrate that the double-stranded probes can be effectively destabilized through modification of both strands with either **222 (ON5:ON6)** or **4 bulges (ON7:ON8)**, or through modification of one strand with a single **9 monomer (ON11:ON2 or ON1:ON12)**.

Incorporation of bulges into Invader probe strands invariably reduces cDNA affinity relative to conventional Invader probe strands (**ON1** or **ON2**). Probe strands containing large bulges reduce the cDNA affinity more than small bulges (e.g., compare **ON3:ON2** and **ON7:ON2** with **ON11:ON2**). Insertion of multiple small bulges into the probe strand progressively reduces cDNA affinity (e.g., compare **ON3:ON2** and **ON5:ON2**). However, insertion of additional **9 bulges** does not further reduce cDNA affinity of probe strands (e.g., compare **ON3:ON2** and **ON13:ON2**).

**Table 4.2-1.** Thermal denaturation temperatures ( $T_m$ 's) of Invader probe duplexes and individual probe strands hybridized with complementary DNA.<sup>a</sup>

Sequence					
5'-GGUGGTCAA X <sub>1</sub> CTATCUGGA 3'-CCACCAGTT X <sub>2</sub> GATAGACCT			$T_m$ [ $T_m$ ]/(°C)		
Probe (5'-ON:3'-ON)	X <sub>1</sub>	X <sub>2</sub>	probe duplex	5'-ON:cDNA	3'-ON:cDNA
1:2	-	-	60.5 [+1.5]	65.5 [+6.5]	68.5 [+9.5]
3:2	2	-	58.0 [-1.0]	63.0 [+4.0]	68.5 [+9.5]
1:4	-	2	54.0 [-5.0]	65.5 [+6.5]	64.5 [+5.5]
3:4	2	2	52.0 [-7.0]	63.0 [+4.0]	64.5 [+5.5]
5:2	222	-	52.0 [-7.0]	58.0 [-1.0]	68.5 [+9.5]
1:6	-	222	47.5 [-11.5]	65.5 [+6.5]	59.0 [0]
5:6	222	222	41.0 [-18.0]	58.0 [-1.0]	59.0 [0]
7:2	4	-	51.5 [-7.5]	58.0 [-1.0]	68.5 [+9.5]
1:8	-	4	47.5 [-11.5]	65.5 [+6.5]	60.0 [+1.0]
7:8	4	4	40.5 [-18.5]	58.0 [-1.0]	60.0 [+1.0]
9:2	444	-	44.0 [-15.0]	53.0 [-6.0]	68.5 [+9.5]
1:10	-	444	39.5 [-19.5]	65.5 [+6.5]	53.5 [-5.5]
9:10	444	444	37.0 [-22.0]	53.0 [-6.0]	53.5 [-5.5]
11:2	9	-	44.0 [-15.0]	53.5 [-5.5]	68.5 [+9.5]
1:12	-	9	42.0 [-17.0]	65.5 [+6.5]	55.5 [-3.5]
11:12	9	9	44.0 [-15.0]	53.5 [-5.5]	55.5 [-3.5]
13:2	999	-	45.0 [-14.0]	53.0 [-6.0]	68.5 [+9.5]
1:14	-	999	44.5 [-14.5]	65.5 [+6.5]	56.0 [-3.0]
13:14	999	999	45.0 [-14.0]	53.0 [-6.0]	56.0 [-3.0]

<sup>a</sup>Thermal denaturation curves were recorded in medium salt phosphate buffer ([Na<sup>+</sup>] = 110 mM, [Cl<sup>-</sup>] = 100 mM, pH 7.0 (NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>), [EDTA] = 0.2 mM) and each [ON] = 1.0 μM.  $\Delta T_m$  is calculated relative to unmodified DNA duplex ( $T_m$  = 59.0 °C). DNA monomers: A = adenin-9-yl, C = cytosin-1-yl, G = guanin-9-yl, T = thymin-1-yl. Invader monomers: U/C = 2'-O-(pyren-1-yl)methyl-RNA U or C monomers (U = uracil-1-yl; C = cytosin-1-yl). Bulge monomers (**X**): **2** = 1-amino-3-hydroxyprop-2-yl monomer, **4** = 4-hydroxybutyl monomer, **9** = 9-hydroxynonyl monomer.

The binding specificity of individual probe strands was evaluated using three different mismatched DNA targets, in which mismatched nucleotides were present right next to the bulge (**MM1** or **MM2**), in the unmodified region between the bulge and the pyrene-modified monomer (**MM3** or **MM4**), or near the pyrene-modified monomer (**MM5** or **MM6**) (Table 4.2-2). Conventional probe strands (**ON1** or **ON2**) display similar discrimination of mismatched targets as unmodified DNA (**ON15** or **ON16**), except when the mismatch is near the intercalator-modified monomer (**MM5** or **MM6**), in which case slightly reduced MM discrimination is observed. Bulged probe strands display similar or slightly improved discrimination of mismatched nucleotides, if these are in an unmodified region or opposite the intercalator-modified monomers. However, less efficient single nucleotide discrimination is observed if the mismatched nucleotide is located next to the bulged region (**MM1** or **MM2**).

**Table 4.2-2.** Thermal denaturation temperatures of singly mismatched DNA duplexes.<sup>a</sup>

ON	Sequence	$T_m$ [ $\Delta T_m$ ] (°C)		
		5'-ON:MM1 3'-ON:MM2	5'-ON:MM3 3'-ON:MM4	5'-ON:MM5 3'-ON:MM6
<b>15:16</b>	5'-GGTGGTCAACTATCTGGA 3'-CCACCAGTTGATAGACCT	50.5 [-8.5] 53.5 [-5.5]	55.5 [-3.5] 51.5 [-7.5]	50.5 [-8.5] 47.5 [-11.5]
<b>1:2</b>	5'-GGTGGTCAACTATCTGGA 3'-CCACCAGTTGATAGACCT	56.5 [-9.0] 63.0 [-5.5]	62.5 [-3.0] 61.5 [-7.0]	59.5 [-6.0] 59.5 [-9.0]
<b>3:4</b>	5'-GGTGGTCAA <b>2</b> CTATCTGGA 3'-CCACCAGTT <b>2</b> GATAGACCT	53.5 [-9.5] 59.5 [-5.0]	58.0 [-5.0] 56.5 [-8.0]	56.5 [-6.5] 54.5 [-10.0]
<b>5:6</b>	5'-GGTGGTCAA <b>222</b> CTATCTGGA 3'-CCACCAGTT <b>222</b> GATAGACCT	51.5 [-6.5] 55.0 [-4.0]	55.5 [-2.5] 51.0 [-8.0]	53.0 [-5.0] 48.0 [-11.0]
<b>7:8</b>	5'-GGTGGTCAA <b>4</b> CTATCTGGA 3'-CCACCAGTT <b>4</b> GATAGACCT	51.0 [-7.0] 55.0 [-5.0]	54.5 [-3.5] 51.0 [-9.0]	52.0 [-6.0] 49.0 [-11.0]
<b>9:10</b>	5'-GGTGGTCAA <b>444</b> CTATCTGGA 3'-CCACCAGTT <b>444</b> GATAGACCT	46.5 [-6.5] 51.0 [-2.5]	50.0 [-3.0] 44.0 [-9.5]	46.5 [-6.5] 43.0 [-10.5]
<b>11:12</b>	5'-GGTGGTCAA <b>9</b> CTATCTGGA 3'-CCACCAGTT <b>9</b> GATAGACCT	48.5 [-5.0] 53.5 [-2.0]	49.5 [-4.0] 47.0 [-8.5]	46.5 [-7.0] 44.0 [-11.5]
<b>13:14</b>	5'-GGTGGTCAA <b>999</b> CTATCTGGA 3'-CCACCAGTT <b>999</b> GATAGACCT	45.0 [-8.0] 54.0 [-2.0]	47.0 [-6.0] 47.5 [-8.5]	48.5 [-4.5] 47.0 [-9.0]

<sup>a</sup> For experimental conditions, see Table 4.2-1.  $\Delta T_m$  is calculated with respect to the  $T_m$  of the specific probe strand (ON) with cDNA. For values see Table 4.2-1. **MM1**: 5'-GGTGGTCAACTATCTGGA; **MM2**: 3'-CCACCAGTTGATAGACCT; **MM3**: 5'-GGTGGTCAACTATCTGGA; **MM4**: 3'-CCACCAGTTGATAGACCT; **MM5**: 5'-GGTGGTCAACTATCTGGA; **MM6**: 3'-CCACCAGTTGATAGACCT. Mismatched bases, with respect to the probe sequence, are shown in italics.

The thermodynamic parameters of hybridization for Invader probes and the individual probes strands with cDNA were obtained from melting curves via the van't Hoff method (Figure 4.2-2 and Tables 4.4-2 through 4.4-6). Trends in Gibbs free energy of duplex formation generally follow the previously discussed trends in  $T_m$ . The destabilization of double-stranded Invader probes containing a single bulge depends on the size of the

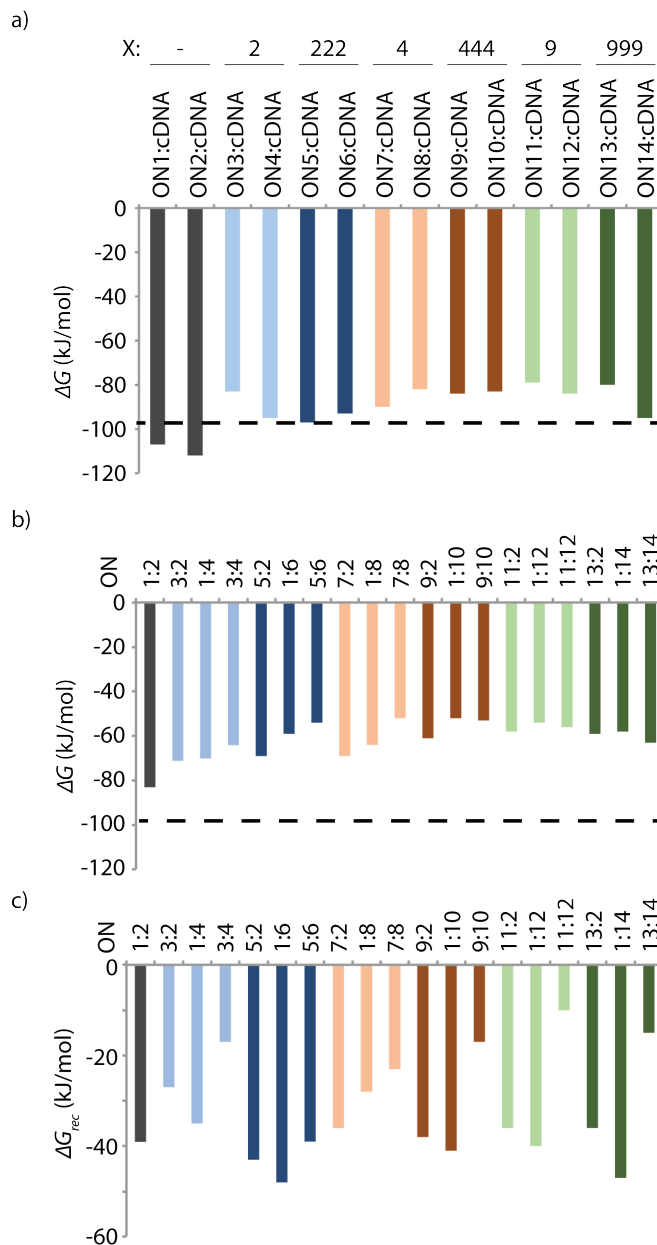
bulge (**2** < **4** < **9**). Increasing the number of monomers in the bulge results in further destabilization of the probe duplex when short bulge monomers **2** and **4** are used but not when monomer **9** is used. A similar trend is observed for probes with bulges on both strands. This supports the hypothesis that incorporation of a bulge with a minimum length (i.e., a single **9** monomer or triple incorporations of **2** or **4** monomers) is sufficient to interrupt the hydrophobic stacking interactions within the duplex core, and additional modification (e.g., triple incorporation of **9** monomer) does not result in increased destabilization.

Probe strands containing bulge inserts have lower cDNA affinity than the conventional probe. In general, probe strands with **4** or **9** monomers (single or triple incorporations) are more destabilized than probes with **2** or **222** monomers. The reason for this is not clear but could be due to differences in the distance between or orientation of the nucleotides flanking the bulge.

The thermodynamic driving force for recognition of isosequential dsDNA targets can be expressed as the energy difference between the 'product' duplexes (5'-ON:cDNA and 3'-ON:cDNA) and 'reactant' duplexes (Invader probe and dsDNA). To parameterize this, we have coined the term  $\Delta G_{\text{rec}}$ , defined as  $\Delta G_{\text{rec}} = (\Delta G_{5'\text{-ON:cDNA}} + \Delta G_{3'\text{-ON:cDNA}}) - (\Delta G_{\text{probe duplex}} + \Delta G_{\text{dsDNA}})$ , where more negative values signify a greater thermodynamic driving force for dsDNA-recognition. Invader probes featuring a short bulge on one strand (**2** or **4**) display lower thermodynamic potential for dsDNA-recognition than the conventional probe **ON1:ON2**. In contrast, probes with a longer bulge (**222**, **444**, **9**, or **999**) display similar or greater thermodynamic potential than the conventional probe. Probes with bulges located on both strands have significantly lower thermodynamic potential for recognition of



dsDNA, with the exception of probes with two **222** bulges where the higher cDNA affinity increases the energetic gradient similarly as the conventional probe.



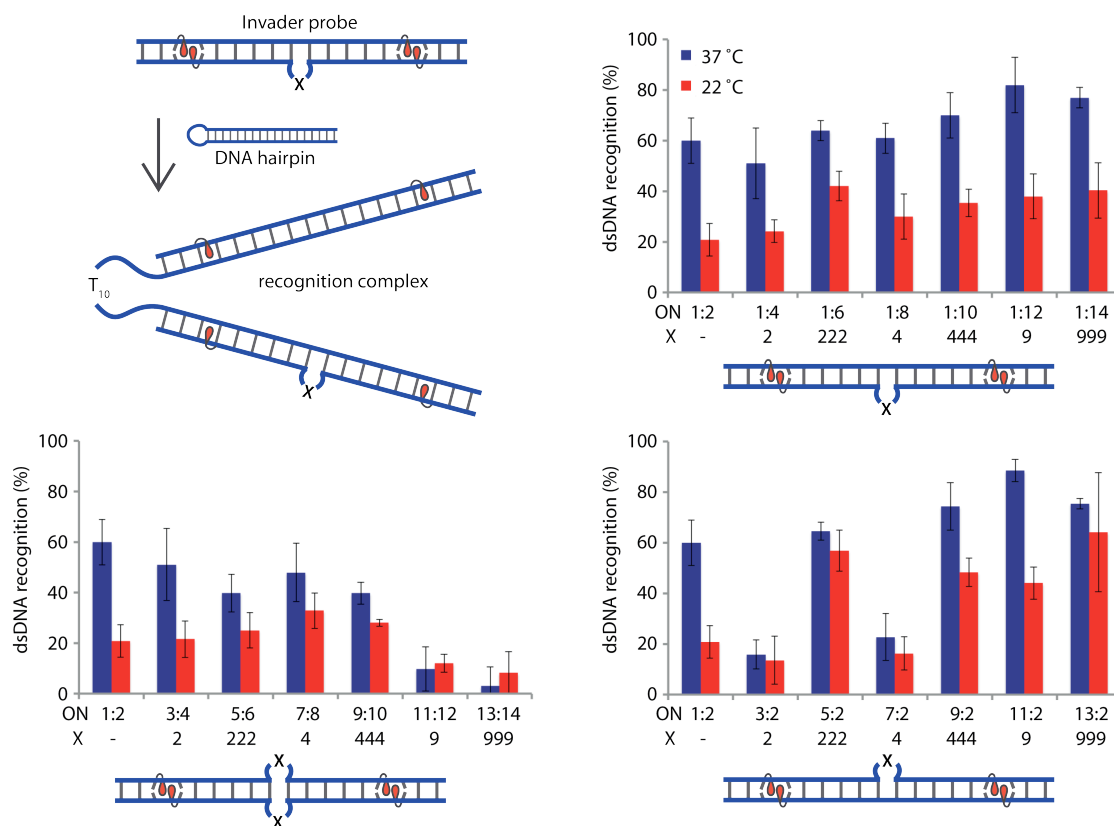
**Figure 4.2-2.** Change in Gibbs free energy of hybridization at 25 °C for a) individual probe strands and complementary DNA or b) Invader probe duplexes. c) Gibb's free energy for recognition of isosequential dsDNA ( $\Delta G_{rec} = (\Delta G_{5'-ON:cDNA} + \Delta G_{3'-ON:cDNA}) - (\Delta G_{probe\ duplex} + \Delta G_{dsDNA})$ ). Dotted line represents  $\Delta G$  of unmodified DNA = -97 kJ/mol. For tabulated data, see Table 4.4-2.

To assess the ability of bulged Invader probes to recognize dsDNA, we utilized a previously developed electrophoretic mobility shift assay (EMSA).<sup>17</sup> Pre-annealed Invader probes were incubated with a DNA hairpin (**D1**), in which the stem of the hairpin is complementary to the probe. Invasion of the double-stranded stem results in the formation of a recognition complex, which has a lower mobility on non-denaturing polyacrylamide gel electrophoresis (nd-PAGE) than the DNA hairpin.

In general, bulged Invader probes result in more efficient dsDNA recognition than the conventional probe at 22 °C (red bars; Figure 4.2-3). Probes with a single incorporation of a bulge monomer result in increased recognition with respect to the length of the monomer (e.g., **ON1:ON4** > **ON1:ON8** > **ON1:ON12**). Incorporation of three consecutive bulge monomers into Invader probes result in increased recognition for the short monomers (e.g., **ON1:ON6** and **ON1:ON10**) compared to probes with single monomer incorporations (e.g., **ON1:ON4** and **ON1:ON8**), but no additional recognition is observed for probes with triple insertions of the longer monomer **9**. Probes with bulges on both strands generally result in increased recognition at 22 °C, with the exception of **ON11:ON12** and **ON13:ON14**. These probes are also predicted to have the lowest thermodynamic potential (Table 4.4-2). In fact, Invader probes modified with bulges on both strands rarely possess an advantage over the single bulge design for a given bulge type at lower temperatures (i.e., 22 °C), which is also reflected by the  $\Delta G_{\text{rec}}$ -values (Table 4.4-2).

At higher temperatures (37 °C), recognition of dsDNA is significantly more efficient than at 22 °C, which is likely a kinetic effect, i.e., increased nucleation events as a consequence of increased base pair ‘breathing’ of duplexes, since dsDNA-recognition

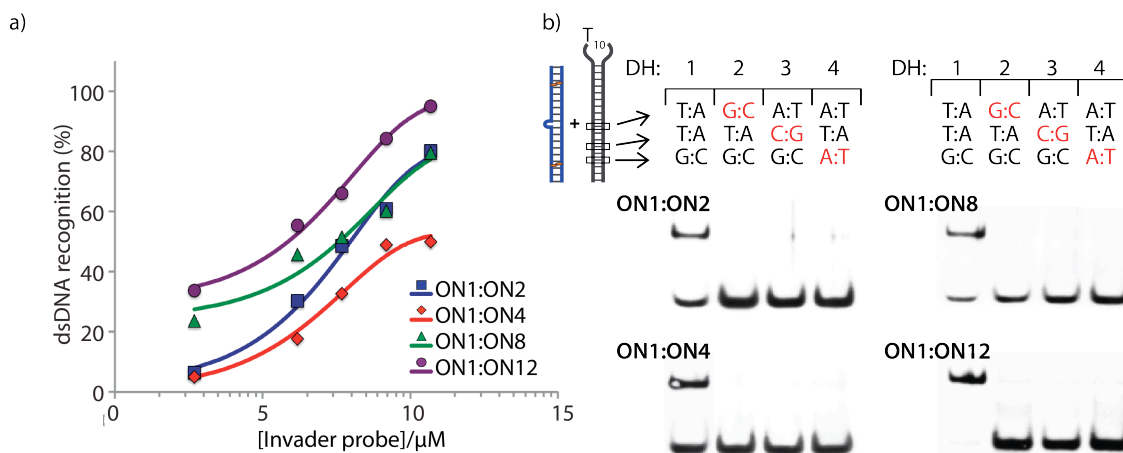
actually is slightly less thermodynamically favorable at higher temperatures (compare  $\Delta G_{\text{rec}}$ -values in Tables 4.4-2 and 4.4-3). Probes containing bulges on both strands generally result in less efficient dsDNA recognition compared to the conventional probe (**ON1:ON2**). Again, this is predicted by  $\Delta G_{\text{rec}}$  values (Table 4.4-3) and largely reflects the lower cDNA affinity of both probe strands. In general, bulged Invader probes, where only one strand contains a bulge, display more efficient dsDNA recognition than the convention probe, with the exception of **ON3:ON2** and **ON7:ON2**, which are not thermodynamically favorable for dsDNA recognition.



**Figure 4.2-3.** Recognition of hairpin DNA target by Invader probes monitored via an electrophoretic mobility shift assay. Illustration of assay (upper left), Invader probes in which only one (right panel) or both strands contain bulges (lower left). Measurements were obtained after incubation of 200-fold molar excess of pre-annealed Invader probe (6.88  $\mu\text{M}$ ) with DIG-labeled DNA hairpin (34.4 nM) for 17 h in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM  $\text{MgCl}_2$ , pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride). For tabulated data, see Table 4.4-8.

To understand the influence of the different bulge chemistries, the probe architecture with a single incorporation of each bulge monomer on the ‘lower’ probe strand was chosen for further evaluation. The concentration dependence of dsDNA recognition was evaluated using the electrophoretic mobility shift assay (Figure 4.2-4a). The conventional probe, **ON1:ON2**, has a  $C_{50}$ -value (the concentration at which 50% recognition occurs) of 3.9  $\mu\text{M}$ , while the probes featuring an alkyl bulge display significantly lower values (3.3 and 1.5  $\mu\text{M}$  for **ON1:ON8** and **ON1:ON12**, respectively). Interestingly, the dose-response for **ON1:ON2** and **ON1:ON8** is very similar at higher concentrations, however

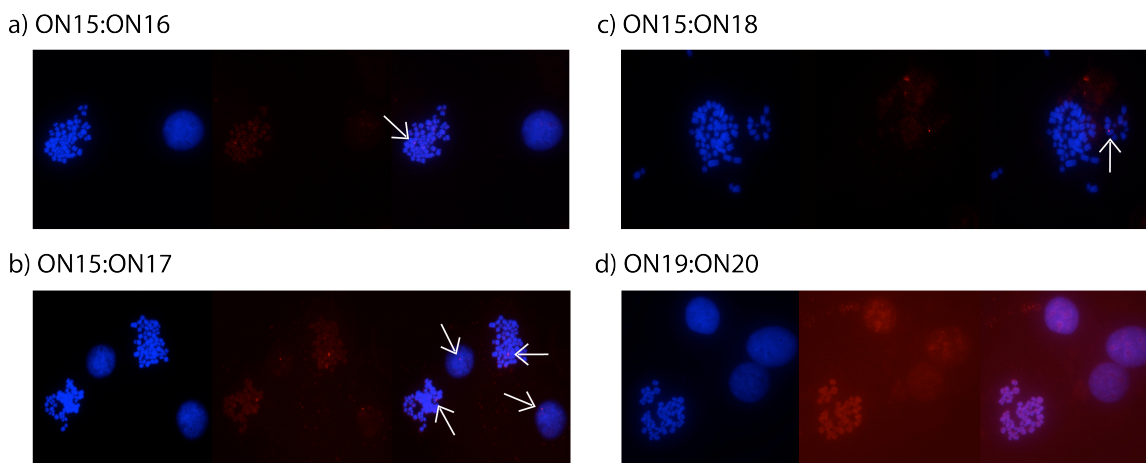
**ON1:ON8**, containing the four-atom bulge ( $X = 4$ ), results in more efficient recognition at lower concentrations (i.e., ~24% versus ~6% for **ON1:ON8** and **ON1:ON2**, respectively, at 10-fold molar probe excess). The Invader probe containing the two-atom bulge (**ON1:ON4**) must be used at higher concentrations than the convention probe to achieve recognition, with a  $C_{50}$ -value of  $10.2 \mu\text{M}$ .



**Figure 4.2-4.** Characterization of DNA recognition by Invader probes at 37 °C. (a) Dose-dependence of recognition for selected probes and (b) specificity of recognition using 400-fold molar excess Invader probe (13.8  $\mu\text{M}$ ). See Figure 4.2-3 for other experimental conditions. For full sequence of hairpin DNA and corresponding  $T_m$ 's, see Table 4.4-9.

The specificity of these bulged Invader probes was also investigated using the electrophoretic mobility shift assay. Probes were incubated with DNA hairpins that are fully complementary in the stem region, but contain a single base-pair mismatch with respect to the probe. If recognition complexes are formed, the mismatch would be located either adjacent to the bulge (**DH2**), in the unmodified region between the bulge and Invader monomers (**DH3**), or near the pyrene-functionalized monomers (**DH4**). Even at the highest concentration of probes examined (400-fold molar excess), the mismatched DNA hairpins are not recognized by any of the Invader probes (Figure 4.2-4b). This level of mismatch discrimination is remarkable considering that 17 of 18 base pairs are identical, and probably reflects the fact that the double duplex invasion is inherently more discriminatory than simple hybridization-based approaches, as both the structured DNA and structured probe must be disrupted to form two less stable complexes (compare Table 4.2-1 with Table 4.2-2).<sup>11</sup>

Encouraged by the superior recognition and excellent specificity, bulged Invader probes were used in a non-denaturing in situ fluorescence hybridization (nd-FISH) assay. Fluorophore-labeled probes with a single bulge of **4** or **9** monomers (**ON15:ON17** or **ON15:ON18**, respectively), were designed to target a gender-specific region of the *DYZ-1* satellite ( $\sim 6 \times 10^4$  tandem repeats) of the bovine (*Bos taurus*) Y-chromosome and compared to the conventional probe lacking a bulge (**ON15:ON16**).<sup>21</sup> Invader probes are simply incubated with isolated nuclei for 3 h, followed by removal of excess probe through rinsing and visualization using fluorescence microscopy. The emergence of a punctate signal in the Cy3 channel and colocalization with chromosomal DNA (DAPI stain), suggests sequence-specific dsDNA recognition by Invader probes (Figure 4.2-5). Bulged Invader probes appear to have more intense signal relative to the background, particularly probes with monomer **4**, than the conventional construct (compare Figure 4.2-5a with 4.2-5b and 4.2-5c). A mismatched probe (**ON19:ON20**) with the equivalent bulge architecture as **ON15:ON17**, but contains three base pair mismatches with respect to the target sequence, failed to yield any signal (Figure 4.2-5d), which lends further support that this is a sequence specific process.



**Figure 4.2-5.** Bulged Invader probes used in nd-FISH for sequence-specific detection of chromosomal DNA of interphase or metaphase nuclei spreads. Invader probes (100 nM) with the conventional probe construct (a), or bulged probes where  $\underline{\mathbf{X}} = 4$  (b) or  $\underline{\mathbf{X}} = 9$  (c) were incubated for 3 h with isolated nuclei at 37.5 °C in Tris buffer ([Tris-HCl] = 10 mM, [EDTA] = 1 mM, [KCl] = 50 mM, pH 8.3). Targeting probe sequence: 5'-Cy3-AGCCCUGTGCCCTG:3'-TCGGGACAC $\underline{\mathbf{X}}$ GGGAC-Cy3. d) Triply mismatched probe (5'-Cy3-AGCGCUGAGGCCTG:3'-TCGCGACTC4CGGAC-Cy3) was used; italicized bases depict mismatches with respect to the target sequence.  $T_m$ 's and  $TA$  of Invader probes shown in Table 4.4-11.

### 4.3 Conclusions

Bulged Invader probes represent a novel architecture capable of efficient recognition of mixed-sequence linear, hairpin, and chromosomal DNA at physiologic-like conditions. The specific chemistry of the bulge insert does not appear to play a critical role in activity. However, there does appear to be a dependence on the length of the bulge, with a preference for bulges greater than ~12 atoms in length based on thermodynamic analysis and recognition of hairpin DNA. Probes in which only one probe strand contains a bulge display greater thermodynamic potential for dsDNA recognition, which translates into increased recognition of a model hairpin DNA target. Bulged Invader probes can be used for efficient and specific detection of chromosomal DNA in nd-FISH and represent



an advancement for the development of hybridization-based probes for diagnostic applications at physiological conditions.

#### 4.4 Supplementary data

##### *Protocol - synthesis and purification of modified ONs*

ONs were synthesized via an automated DNA synthesizer (0.2  $\mu$ mol scale) following manufactures recommendations. Long chain alkyl amine controlled pore glass (LCAA-CPG, 500 Å pore size) pre-loaded with the 3'-nucleotide was used as a solid support. Modified nucleotides were incorporated via hand-coupling of the previously described phosphoramidite,<sup>26</sup> i.e., 0.01 M 4,5-dicyanoimidazole in acetonitrile for 15 min with extended oxidation (45 s) conditions. Non-nucleosidic linkers were incorporated in a similar manner using commercially available phosphoramidites, with the exception of monomer **2** where 5-ethylthio-1H-tetrazole was used as the activator and coupling was performed for 15 min. at 55 °C. ONs were cleaved from the solid support and base-labile groups were removed by treatment with 32% ammonia (55 °C, 17 h). DMT-protected ONs were purified using ion-pair reverse-phase HPLC (0.05 M triethylammonium acetate and acetonitrile gradient), followed by detritylation (80% acetic acid, 20 min), and precipitation (sodium acetate and sodium perchlorate in acetone, -18 °C, 16 h). The purity was determined via analytical HPLC (>85%) and identity was confirmed using MALDI-MS (2,4,6-trihydroxyacetophenone matrix).

ONs were labeled with cyanine 3 (Cy3) using a commercially available Cy3 phosphoramidite, which was coupled as described above, with the following exceptions:

coupling time was 5 min, detritylation was performed on the synthesizer, followed by subsequent deprotection in 32% ammonia (55 °C, 4 h) and purification.

*Protocol - thermal denaturation experiments*

The concentration of ON was estimated using the following extinction coefficients ( $OD_{260}/\mu\text{mol}$ ): G (12.01), A (15.20), T (8.40), C (7.05), pyrene (22.4), and Cy3 (4.93).<sup>25</sup>

Thermal denaturation temperatures ( $T_m$ 's) were determined as the average of the maximum of the first-derivative of the denaturing curves ( $A_{260}$  vs.  $T$ , rounded to the nearest 0.5 °C) from at least two experiments within 1.0 °C. ONs (0.5  $\mu\text{M}$ ) were annealed (85 °C for 2 min) in medium salt buffer ( $[\text{Na}^+] = 110 \text{ mM}$ ,  $[\text{Cl}^-] = 100 \text{ mM}$ , pH 7.0 ( $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ ),  $[\text{EDTA}] = 0.2 \text{ mM}$ ) and cooled down to the starting temperatures. Thermal denaturation curves were recorded from at least 15 °C below to 15 °C above the  $T_m$  using a temperature ramp of 1.0 °C/min.

*Protocol - thermodynamic parameters*

Thermodynamic parameters were obtained from the melting curves via baseline fitting (van't Hoff method). Bimolecular reactions, two-state melting behavior, and constant heat capacity were assumed.<sup>25</sup> Values are reported as the average of two different melting curves analyzed three times each.

*Definition - n interstrand zipper arrangement*

The following nomenclature describes the relative arrangement between two monomers positioned on opposing strands in the probe duplex. The number  $n$  describes the number of base pairs between the two monomers, where a positive value indicates that a

monomer is shifted toward the 5'-side of its own strand relative to the monomer on the other strand. Conversely, if  $n$  is negative, the monomer is shifted toward the 3'-side of the strand relative to the monomer on the other strand.

*Protocol – steady-state fluorescence emission experiments*

Linear dsDNA (5'-GGTGGTCAACTATCTGGA: 3'-CCACCAGTTGATAGACCT; both strands at a concentration of 1.25  $\mu$ M) was equilibrated in HEPES buffer (50 mM hepes, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) at 22 °C to which pre-annealed Invader probe (1.25  $\mu$ M) was added. The steady-state fluorescence spectrum was recorded (excitation = 350 nm, excitation slit width = 5 nm, emission slit width = 5 nm, PMT detector voltage = 800 V, quartz optical cell path length = 10 mm) after equilibration (~24 h). Percent dissociation – interpreted as percent recognition – was calculated from the ratio of fluorescence intensity from 450-650 nm, as an average of three scans, of Invader probe to Invader probe in the presence of linear complementary dsDNA.

*Protocol - electrophoretic mobility shift assay*

DNA hairpins, consisting of double-stranded stem region in which the strands are connected via T<sub>10</sub> linker, were purchased from a commercial vendor and used without further purification. The hairpins were annealed (95 °C, 2 min), and then enzymatically DIG-labeled according to the manufacturer's protocols. Briefly, digoxigenin-ddUTP was incubated with recombinant DNA 3'-terminal transferase and the DNA hairpin (15 min, 37 °C), the reaction was then quenched with EDTA (0.05 M) and the hairpin was used without further purification.

Invader probes were pre-annealed (95 °C, 2 min) and subsequently incubated with the DIG-labeled DNA hairpin (34.4 nM) in hepes buffer (50 mM hepes, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) at either 22 or 37° C (±2 °C). Following addition of loading dye, the reaction mixtures were loaded onto 12% non-denaturing polyacrylamide gels (45 mM tris-borate, 1 mM EDTA; acrylamide:bisacrylamide (19:1)). Following electrophoresis (~4 °C, 70 V, 2 h), the bands were electroblotted onto positively charged nylon membranes (100 V, 30 min, ~4 °C) and cross-linked (254 nm, 5 x 15 watt bulbs, 5 min). Membranes were incubated with anti-DIG alkaline phosphatase F<sub>ab</sub> fragments, according to the manufacturer's recommendations. After incubation with a chemiluminescent substrate (CSPD) at 37 °C for 10 min, images of the chemiluminescent bands were captured on x-ray film and quantified using densitometry software. The percentage of DNA recognition was calculated as the ratio of intensity of the recognition complex relative to the intensity of the total lane, and is reported as the average of three independent experiments.

#### *Protocol - cell culture and nuclei preparation*

Cells were shipped in the log phase to Cell Line Genetics, Inc. in Madison, WI for microscope slide preparation. Complete growth media was changed 24 h before cell harvesting, then 65 µL of colcemid was added and incubated at 37 °C and 5% CO<sub>2</sub> for 0.2 - 24h. Media was transferred to a conical tube and 1 mL of 0.05% Trypsin/EDTA was added to the culture flask to remove adhered cells. Another 1.5 mL of Trypsin/EDTA was added and incubated at 37 °C and 5% CO<sub>2</sub> for 6 - 8 min. Cells were transferred to the conical tube and centrifuged for 10 min at 1000 rpm. The supernatant was removed and 5 - 8 mL of 7.5 mM KCl hypotonic solution was added and let sit for 5 - 27 min. Ten drops

of Carnoy's fixative (3:1 methanol/glacial acetic acid) was added, gently mixed, and incubated at room temperature for 10m. Cells were centrifuged again for 10 min at 1000rpm and supernatant removed. More fixative was slowly added to a total of 5 - 8 mL, gently mixed, and incubated at room temperature for 30 min. Cells centrifuged again for another 10 min at 1000 rpm, and the supernatant removed. This process was repeated for a total of three fixative changes.

Cells were centrifuged for 10 min at 1000 rpm, the supernatant removed, and enough fixative added to make the suspension appear turbid or milky. A fresh microscope slide was dipped in distilled water, allowing the water to form a uniform sheet across the slide surface. Holding the slide at a 45° angle, a single drop of the cell suspension was placed onto the slide below the frosted end and allowed to run down the slide. The bottom of the slide was blotted and the slide was immediately placed in a Percival chamber (28 °C, 39% humidity) at a 20° angle and allowed to dry completely.

#### *Protocol - fluorescence in situ hybridization*

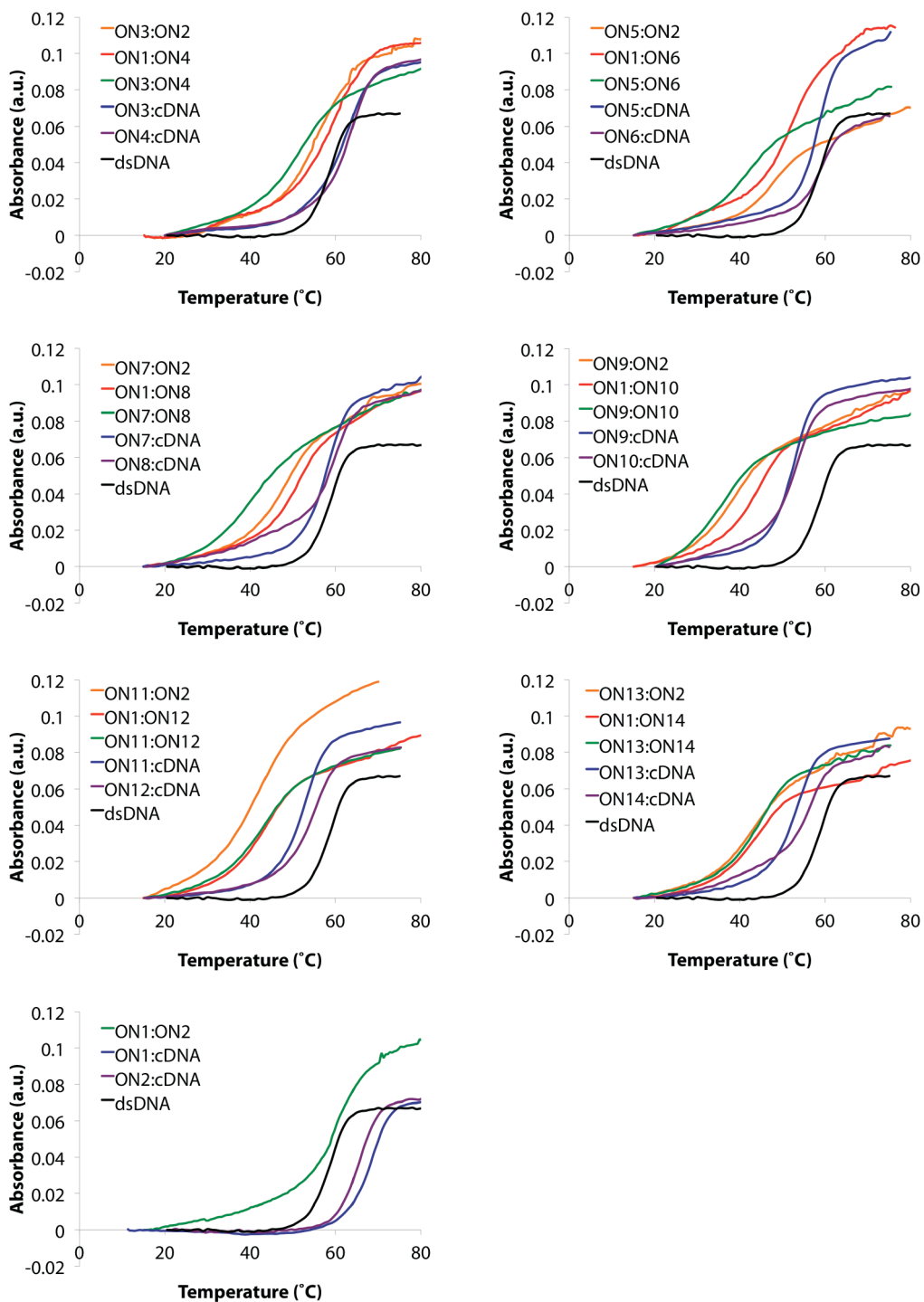
Slides were incubated with 100 µL of Cy3-labeled Invader probe solution ([Inv] = 20-30 nM) and 100 µL of PCR buffer ([Tris-HCl] = 10mM, [KCl] = 50 mM, pH 8.3) for 3 h at 38.5 °C. Slides were then washed with TE buffer ([Tris] = 10 mM, [EDTA] = 1 mM, pH 8.0) for 3 min by submersion and gentle pipetting. Slides were then rinsed with autoclaved water by pipetting. Slides were allowed to dry completely at room temperature. An aliquot of 3 µL of DAPI nuclear stain (Gold SlowFade Plus, Invitrogen) was placed on the slide. A glass, round coverslip was placed on the coverslip and sealed. The slides were then viewed on a Nikon Eclipse Ti-S/L100 fluorescence microscope

equipped with Cy3 and DAPI filter cubes at 60X magnification. Images were taken at 20 ms and 500 ms exposures for DAPI and Cy3, respectively. Images were then overlaid using the NIS-Elements imaging software.

**Table 4.4-1.** MALDI-MS of modified ONs.<sup>a</sup>

ON	Sequence	Observed	Calculated
		<i>m/z</i> [M+H] <sup>+</sup>	<i>m/z</i> [M+H] <sup>+</sup>
<b>1</b>	5'-GGTGGTCAACTATCTGGA	5984.2	5984.1
<b>2</b>	3'-CCACCAGTTGATAGACCT	5901.7	5901.1
<b>3</b>	5'-GGTGGTCAA <b>2</b> CTATCTGGA	6137.4	6137.1
<b>4</b>	3'-CCACCAGTT <b>2</b> GATAGACCT	6054.3	6054.1
<b>5</b>	5'-GGTGGTCAA <b>222</b> CTATCTGGA	6443.8	6443.1
<b>6</b>	3'-CCACCAGTT <b>222</b> GATAGACCT	6360.2	6360.1
<b>7</b>	5'-GGTGGTCAA <b>4</b> CTATCTGGA	6136.4	6136.1
<b>8</b>	3'-CCACCAGTT <b>4</b> GATAGACCT	6053.3	6053.1
<b>9</b>	5'-GGTGGTCAA <b>444</b> CTATCTGGA	6440.4	6440.1
<b>10</b>	3'-CCACCAGTT <b>444</b> GATAGACCT	6358.5	6357.2
<b>11</b>	5'-GGTGGTCAA <b>9</b> CTATCTGGA	6206.3	6206.2
<b>12</b>	3'-CCACCAGTT <b>9</b> GATAGACCT	6123.5	6123.2
<b>13</b>	5'-GGTGGTCAA <b>999</b> CTATCTGGA	6650.6	6650.4
<b>14</b>	3'-CCACCAGTT <b>999</b> GATAGACCT	6567.7	6567.4

<sup>a</sup> **U/C** = 2'-*O*-(pyren-1-yl)methyl-RNA U or C monomers (U = uracil-1-yl; C = cytosin-1-yl); **2** = 1-amino-3-hydroxyprop-2-yl monomer, **4** = 4-hydroxybutyl monomer, **9** = 9-hydroxynonyl monomer. 2,4,6-trihydroxyacetophenone was used as the matrix.



**Figure 4.4-1.** Representative thermal denaturation curves of Invader probes, duplexes consisting of individual probe strands and cDNA, and unmodified reference cDNA duplex. For experimental conditions, see Table 4.2-1.

**Table 4.4-2.** Change in Gibbs free energy ( $\Delta G$ ) at 298 K for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{298}$ ).<sup>a</sup>

Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGA 3'-CCACCAGTT X <sub>2</sub> GATAGACCT			$\Delta G^{298}$ (kJ/mol)			
ON	X <sub>1</sub>	X <sub>1</sub>	Invader probe	5'-Inv: cDNA	3'-Inv: cDNA	$\Delta G_{rec}^{298}$ (kJ/mol)
1:2	-	-	-83	-107	-112	-39
3:2	2	-	-71	-83	-112	-27
1:4	-	2	-70	-107	-95	-35
3:4	2	2	-64	-83	-95	-17
5:2	222	-	-69	-97	-112	-43
1:6	-	222	-59	-107	-93	-44
5:6	222	222	-54	-97	-93	-39
7:2	4	-	-69	-90	-112	-36
1:8	-	4	-64	-107	-82	-28
7:8	4	4	-52	-90	-82	-23
9:2	444	-	-61	-84	-112	-38
1:10	-	444	-52	-107	-83	-41
9:10	444	444	-53	-84	-83	-17
11:2	9	-	-58	-79	-112	-36
1:12	-	9	-54	-107	-84	-40
11:12	9	9	-56	-79	-84	-10
13:2	999	-	-59	-80	-112	-36
1:14	-	999	-58	-107	-95	-47
13:14	999	999	-63	-80	-95	-15

<sup>a</sup>  $\Delta G_{rec}^{298} = \Delta G^{298}(5\text{'-Inv:cDNA}) + \Delta G^{298}(3\text{'-Inv:cDNA}) - \Delta G^{298}(\text{Invader probe}) - \Delta G^{298}(\text{dsDNA})$ .  $\Delta G^{298}$  for unmodified DNA duplex ( $\Delta G^{298}(\text{dsDNA}) = -97$  kJ/mol).



**Table 4.4-3.** Change in Gibbs free energy ( $\Delta G$ ) at 310 K for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{310}$ ).<sup>a</sup>

Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGGA			$\Delta G^{310}$ (kJ/mol)			
3'-CCACCAGTT X <sub>2</sub> GATAGACCT			Invader probe	5'-Inv:cDNA	3'-Inv:cDNA	$\Delta G_{rec}^{310}$ (kJ/mol)
ON	X <sub>1</sub>	X <sub>2</sub>				
1:2	-	-	-69	-88	-93	-35
3:2	2	-	-60	-70	-93	-26
1:4	-	2	-58	-88	-78	-31
3:4	2	2	-53	-70	-78	-18
5:2	222	-	-56	-74	-93	-34
1:6	-	222	-49	-88	-77	-39
5:6	222	222	-42	-74	-77	-32
7:2	4	-	-55	-72	-93	-33
1:8	-	4	-51	-88	-68	-28
7:8	4	4	-41	-72	-68	-22
9:2	444	-	-47	-65	-93	-34
1:10	-	444	-41	-88	-65	-35
9:10	444	444	-38	-65	-65	-15
11:2	9	-	-46	-62	-93	-32
1:12	-	9	-43	-88	-66	-34
11:12	9	9	-45	-62	-66	-6
13:2	999	-	-47	-63	-93	-32
1:14	-	999	-46	-88	-75	-40
13:14	999	999	-49	-63	-75	-12

<sup>a</sup>  $\Delta G_{rec}^{310} = \Delta G^{310} (5'-Inv:cDNA) + \Delta G^{310} (3'-Inv:cDNA) - \Delta G^{310} (Invader\ probe) - \Delta G^{310} (dsDNA)$ .  $\Delta G^{310}$  for the corresponding unmodified DNA duplex is -77 kJ/mol.

**Table 4.4-4.** Change in enthalpy ( $\Delta H$ ) for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{rec}$ )<sup>a</sup>

Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGA 3'-CCACCAGTT X <sub>2</sub> GATAGACCT			$\Delta H$ (kJ/mol)			
ON	X <sub>1</sub>	X <sub>1</sub>	Invader probe	5'-Inv: cDNA	3'-Inv: cDNA	$\Delta H_{rec}$ (kJ/mol)
1:2	-	-	-436	-584	-587	-147
3:2	2	-	-345	-421	-587	-75
1:4	-	2	-361	-584	-511	-146
3:4	2	2	-338	-421	-511	-6
5:2	222	-	-372	-561	-587	-188
1:6	-	222	-322	-584	-592	-266
5:6	222	222	-341	-592	-561	-224
7:2	4	-	-406	-535	-587	-128
1:8	-	4	-381	-584	-431	-46
7:8	4	4	-312	-535	-431	-66
9:2	444	-	-410	-556	-587	-145
1:10	-	444	-328	-584	-536	-204
9:10	444	444	-422	-556	-536	-82
11:2	9	-	-354	-495	-587	-140
1:12	-	9	-317	-584	-513	-192
11:12	9	9	-337	-495	-513	-83
13:2	999	-	-361	-498	-587	-136
1:14	-	999	-356	-584	-600	-240
13:14	999	999	-415	-498	-600	-95

<sup>a</sup>  $\Delta H_{rec} = \Delta H(5'\text{-Inv:cDNA}) + \Delta H(3'\text{-Inv:cDNA}) - \Delta H(\text{Invader probe}) - \Delta H(\text{dsDNA})$ .  
 $\Delta H$  for unmodified DNA duplex ( $\Delta H(\text{dsDNA}) = -588$  kJ/mol).

**Table 4.4-5.** Change in entropy at 298 K ( $-T^{298}\Delta S$ ) for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon Invader-mediated recognition of isosequential dsDNA targets ( $-T^{298}\Delta S_{rec}$ ).<sup>a</sup>

Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGA 3'-CCACCAGTT X <sub>2</sub> GATAGACCT			$-T^{298}\Delta S$ (kJ/mol)			
ON	X <sub>1</sub>	X <sub>1</sub>	Invader probe	5'-Inv: cDNA	3'-Inv: cDNA	$-T^{298}\Delta S_{rec}$ (kJ/mol)
1:2	-	-	353	477	475	108
3:2	2	-	274	337	475	27
1:4	-	2	291	477	415	70
3:4	2	2	274	337	415	-33
5:2	222	-	304	468	475	148
1:6	-	222	263	477	495	218
5:6	222	222	287	495	468	185
7:2	4	-	337	445	475	92
1:8	-	4	317	477	349	18
7:8	4	4	260	445	349	43
9:2	444	-	349	472	475	107
1:10	-	444	275	477	453	164
9:10	444	444	369	472	453	65
11:2	9	-	297	416	475	103
1:12	-	9	263	477	429	152
11:12	9	9	280	416	429	74
13:2	999	-	302	418	475	100
1:14	-	999	298	477	505	193
13:14	999	999	352	418	505	80

<sup>a</sup>  $-T^{298}\Delta S_{rec} = -T^{298}\Delta S(5'-Inv:cDNA) + -T^{298}\Delta S(3'-Inv:cDNA) - (-T^{298}\Delta S(\text{Invader probe}) + -T^{298}\Delta S(\text{dsDNA}))$ .  $-T^{298}\Delta S$  for unmodified DNA duplex = 497 kJ/mol.

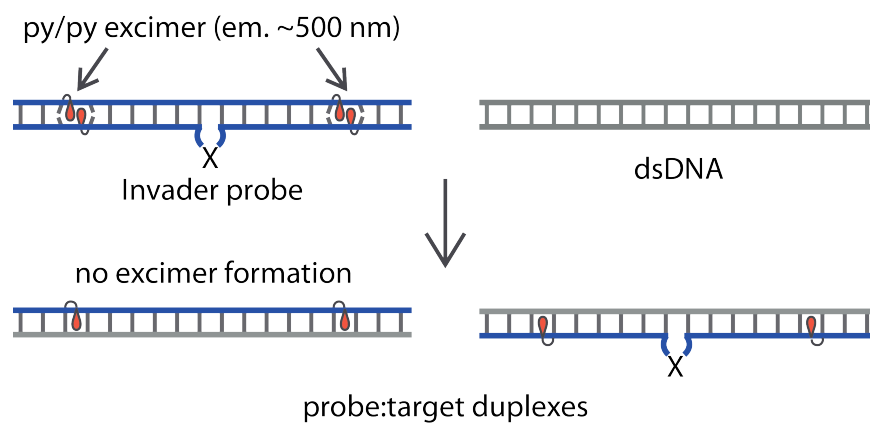
**Table 4.4-6.** Change in entropy at 310 K ( $-T^{310}\Delta S$ ) for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in reaction free energy upon Invader-mediated recognition of isosequential dsDNA targets ( $-T^{310}\Delta S_{rec}$ ).<sup>a</sup>

Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGA 3'-CCACCAGTT X <sub>2</sub> GATAGACCT			$-T^{310}\Delta S$ (kJ/mol)			
ON	X <sub>1</sub>	X <sub>1</sub>	Invader probe	5'-Inv: cDNA	3'-Inv: cDNA	$-T^{310}\Delta S_{rec}$ (kJ/mol)
1:2	-	-	367	496	494	112
3:2	2	-	285	351	494	49
1:4	-	2	303	496	432	114
3:4	2	2	286	351	432	-14
5:2	222	-	316	487	494	154
1:6	-	222	273	496	515	227
5:6	222	222	298	487	515	193
7:2	4	-	350	463	494	96
1:8	-	4	330	496	363	18
7:8	4	4	270	463	363	45
9:2	444	-	363	491	494	111
1:10	-	444	286	496	471	170
9:10	444	444	383	491	471	68
11:2	9	-	308	432	494	107
1:12	-	9	274	496	447	158
11:12	9	9	291	432	447	77
13:2	999	-	314	435	494	104
1:14	-	999	310	496	525	200
13:14	999	999	366	435	525	83

<sup>a</sup>  $-T^{310}\Delta S_{rec} = -T^{310}\Delta S(5'-Inv:cDNA) + -T^{310}\Delta S(3'-Inv:cDNA) - (-T^{310}\Delta S(\text{Invader probe}) + -T^{310}\Delta S(\text{dsDNA}))$ .  $-T^{310}\Delta S$  for the unmodified DNA duplex = 511 kJ/mol.

### *Recognition of linear dsDNA*

A fluorescence-based assay was used to monitor recognition of an isosequential dsDNA model target (Figure 4.4-2).<sup>14</sup> The double-stranded Invader probes exhibit prominent pyrene-pyrene excimer signals, which are centered at ~495 nm, as the +1 interstrand zipper arrangements of the 2'-*O*-(pyren-1-yl)methyl-RNA monomers force the pyrene moieties to intercalate and engage in  $\pi$ - $\pi$ -stacking (Figure 4.4-3). Following recognition of the dsDNA target, the pyrenes no longer  $\pi$ - $\pi$ -stack with each other resulting in predominant pyrene monomer emission (Figure 4.4-2). Probes were pre-annealed and equilibrated for 24 h in HEPES buffer, followed by addition of an equimolar quantity of the model dsDNA target. The ratios of the sum of signal intensities from 450 – 650 nm in the presence or absence of pre-annealed dsDNA are used to estimate the degree dsDNA recognition (Table 4.4-7). Conventional Invader probe **ON1:ON2** only results in ~13% dsDNA recognition under these conditions, while introduction of single nine-atom bulge on one strand more than doubles the strand invasion (25 and 35% for **ON11:ON2** and **ON1:ON12**, respectively). This is likely due to bulge-promoted breathing of the probe duplex, allowing for increased nucleation with the dsDNA target. In contrast, the probes with two bulges on opposite strands result in less efficient dsDNA-recognition, most likely as probe:target duplexes are not sufficiently stable under these conditions (Table 4.2-1). A similar trend is observed for probes with three consecutive bulge monomers (**X** = **999**), where the probe with bulges on both strand (**ON13:ON14**) results in similar invasion efficiency as **ON1:ON2**, but the constructs with a single bulge result in much greater recognition (30 and 44% for **ON1:ON14** and **ON13:ON2**, respectively).

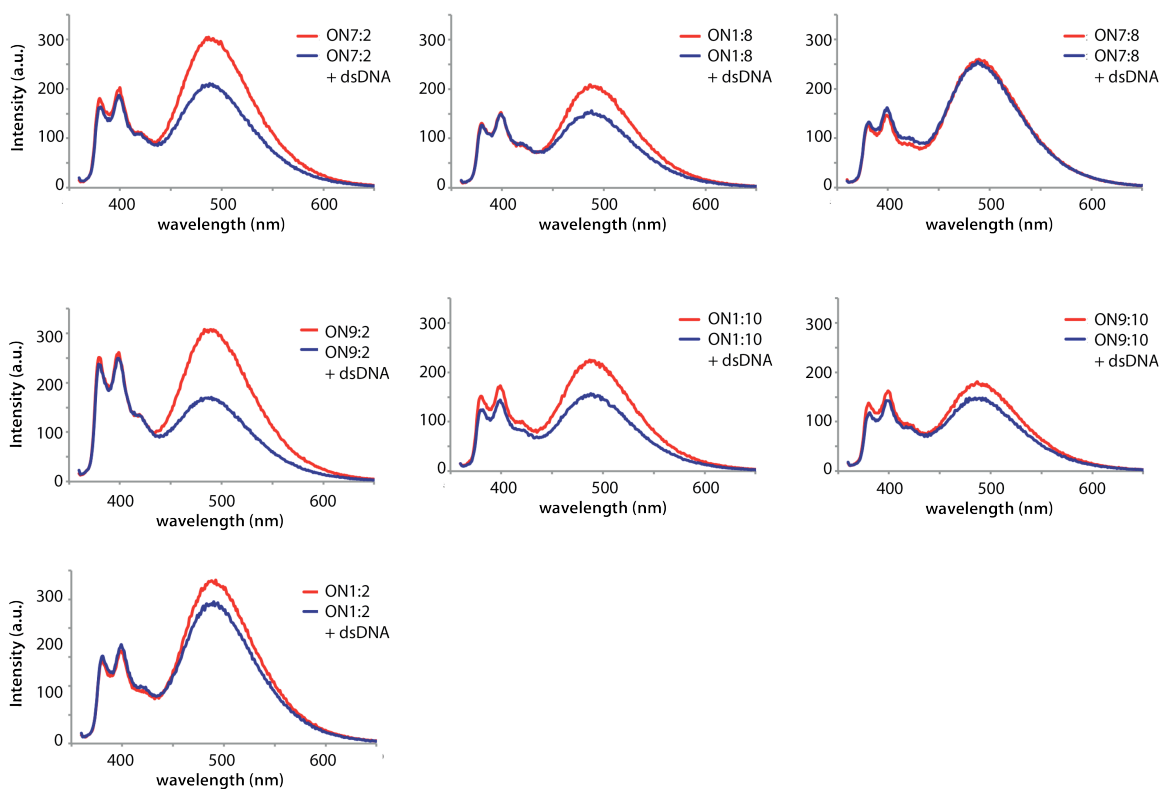


**Figure 4.4-2.** Fluorescence-based assay used to monitor strand exchange process between Invader probes and isosequential dsDNA targets. Py = pyrene. Data shown in Table 4.4-7.

**Table 4.4-7.** Recognition of linear dsDNA targets by Invader probes.<sup>a</sup>

ON	Sequence	Area from 450-650 nm (a.u.)		Strand invasion (%)
		Probe only	Probe + dsDNA	
1:2	5'-GGTGGTCAACTATCTGGA 3'-CCACCAGTTGATAGACCT	24221	21116	13
11:2	5'-GGTGGTCAA 9 CTATCTGGA 3'-CCACCAGTT GATAGACCT	26126	18177	30
1:12	5'-GGTGGTCAA CTATCTGGA 3'-CCACCAGTT 9 GATAGACCT	17871	13368	25
11:12	5'-GGTGGTCAA 9 CTATCTGGA 3'-CCACCAGTT 9 GATAGACCT	22296	21853	2
13:2	5'-GGTGGTCAA 999 CTATCTGGA 3'-CCACCAGTT GATAGACCT	26491	14958	44
1:14	5'-GGTGGTCAA CTATCTGGA 3'-CCACCAGTT 999 GATAGACCT	19428	13606	30
13:14	5'-GGTGGTCAA 999 CTATCTGGA 3'-CCACCAGTT 999 GATAGACCT	15539	12983	16

<sup>a</sup> Measurements were performed in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) at 22 °C and recorded after incubating pre-annealed Invader probe (1.25 μM) with or without addition of pre-annealed linear dsDNA (1.25 μM) for ~24 h. Strand invasion calculated from ratio of fluorescence emission from 450-650 nm of Invader probe in the presence or absence of dsDNA (100 - sum of signal (probe + cDNA)/sum of signal (probe only)\*100). Steady-state emission spectra shown in Figure 4.4-3.



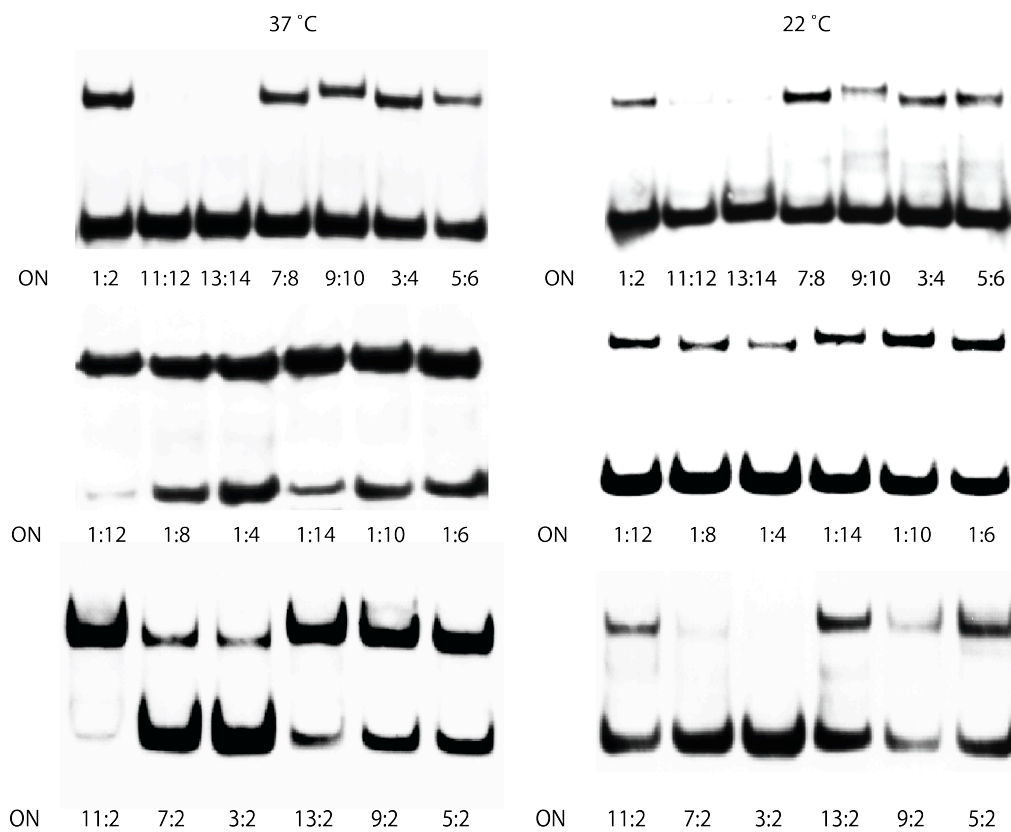
**Figure 4.4-3.** Steady-state fluorescence emission spectra of Invader probes in the absence or presence of an isosequential dsDNA target. Measurements were performed in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) at 22 °C and recorded after incubating pre-annealed Invader probe (1.25  $\mu$ M) with or without addition of pre-annealed linear dsDNA (1.25  $\mu$ M) for  $\sim$ 24 h.



**Table 4.4-8.** Degree of DNA recognition using 200-fold molar excess of the different Invader probes.<sup>a</sup>





ON	Sequence	dsDNA recognition (%)	
		22 °C	37 °C
1:2	5'-GGTGGTCAACTATCTGGA 3'-CCACCAGTTGATAGACCT	21 ± 6	60 ± 9
3:2	5'-GGTGGTCAA 2 CTATCTGGA 3'-CCACCAGTT GATAGACCT	14 ± 10	16 ± 6
1:4	5'-GGTGGTCAA CTATCTGGA 3'-CCACCAGTT 2 GATAGACCT	24 ± 5	51 ± 14
3:4	5'-GGTGGTCAA 2 CTATCTGGA 3'-CCACCAGTT 2 GATAGACCT	22 ± 6	51 ± 14
5:2	5'-GGTGGTCAA 222 CTATCTGGA 3'-CCACCAGTT GATAGACCT	57 ± 8	65 ± 4
1:6	5'-GGTGGTCAA CTATCTGGA 3'-CCACCAGTT 222 GATAGACCT	42 ± 6	64 ± 4
5:6	5'-GGTGGTCAA 222 CTATCTGGA 3'-CCACCAGTT 222 GATAGACCT	25 ± 7	40 ± 8
7:2	5'-GGTGGTCAA 4 CTATCTGGA 3'-CCACCAGTT GATAGACCT	16 ± 7	23 ± 9
1:8	5'-GGTGGTCAA CTATCTGGA 3'-CCACCAGTT 4 GATAGACCT	30 ± 9	61 ± 6
7:8	5'-GGTGGTCAA 4 CTATCTGGA 3'-CCACCAGTT 4 GATAGACCT	33 ± 7	48 ± 12
9:2	5'-GGTGGTCAA 444 CTATCTGGA 3'-CCACCAGTT GATAGACCT	48 ± 6	74 ± 9
1:10	5'-GGTGGTCAA CTATCTGGA 3'-CCACCAGTT 444 GATAGACCT	35 ± 5	70 ± 9
9:10	5'-GGTGGTCAA 444 CTATCTGGA 3'-CCACCAGTT 444 GATAGACCT	28 ± 1	40 ± 4
11:2	5'-GGTGGTCAA 9 CTATCTGGA 3'-CCACCAGTT GATAGACCT	44 ± 6	89 ± 4
1:12	5'-GGTGGTCAA CTATCTGGA 3'-CCACCAGTT 9 GATAGACCT	38 ± 9	82 ± 11
11:12	5'-GGTGGTCAA 9 CTATCTGGA 3'-CCACCAGTT 9 GATAGACCT	12 ± 4	10 ± 9
13:2	5'-GGTGGTCAA 999 CTATCTGGA 3'-CCACCAGTT GATAGACCT	64 ± 24	76 ± 2
1:14	5'-GGTGGTCAA CTATCTGGA 3'-CCACCAGTT 999 GATAGACCT	40 ± 11	77 ± 4
13:14	5'-GGTGGTCAA 999 CTATCTGGA 3'-CCACCAGTT 999 GATAGACCT	8 ± 8	3 ± 7

<sup>a</sup> Data shown in Figure 4.2-2.



**Figure 4.4-4.** Representative electrophoretograms for recognition of DNA hairpin **DH1** by Invader probes. Graphical representation of data shown in Figure 4.2-2.

**Table 4.4-9.** Thermal denaturation temperatures of DNA hairpins.<sup>a</sup>

DH	Sequence	$T_m$ (°C)
1	 5'-GGTGGTCAACTATCTGGA 3'-CCACCAGTTGATAGACCT	73.0
2	 5'-GGTGGTCAGCTATCTGGA 3'-CCACCAGTCGATAGACCT	74.5
3	 5'-GGTGGCCAACTATCTGGA 3'-CCACCGGTTGATAGACCT	74.5
4	 5'-GGTAGTCAACTATCTGGA 3'-CCATCAGTTGATAGACCT	68.5

<sup>a</sup> For experimental conditions, see Table 4.2-1.

**Table 4.4-10.** MALDI-MS of ONs targeting a unique region of the *DYZ-1* satellite region of the Y-chromosome of *bos Taurus*.<sup>a</sup>

ON	Sequence	Observed <i>m/z</i> [M+H] <sup>+</sup>	Calculated <i>m/z</i> [M+H] <sup>+</sup>
15	5'- <b>Cy3</b> <u>A</u> GCCC <u>U</u> GTGCC <u>C</u> TG	5397.5	5397.5
16	3'-TCGGGACACGGGAC <b>Cy3</b>	5508.6	5509.6
17	3'- TCG GGA <b>C</b> AC <b>4</b> GGG <b>A</b> C <b>Cy3</b>	5661.3	5661.6
18	3'- TCGGGAC <b>A</b> C <b>9</b> GGG <b>A</b> C <b>Cy3</b>	5731.2	5731.7
19	5'- <b>Cy3</b> <u>A</u> GCGC <u>U</u> GAGGC <u>C</u> TG	5486.0	5486.5
20	3'- TCGCGACTCCGGAC <b>Cy3</b>	5420.8	5420.5
21	3'- TCG CGA <b>C</b> TC <b>4</b> CGG <b>A</b> C <b>Cy3</b>	5572.0	5572.6

<sup>a</sup> **Cy3** = cyanine 3 phosphoramidite; **2** = 1-amino-3-hydroxyprop-2-yl monomer; **4** = 4-hydroxybutyl monomer; **9** = 9-hydroxynonyl monomer; **A/U/C** = 2'-*O*-(pyren-1-yl)methyl RNA monomers. 2,4,6-trihydroxyacetophenone was used as the matrix.

**Table 4.4-11.** Thermal denaturation temperatures of Invader probes that are derivatives of ‘Inv7:8’ targeting DYZ-1 satellite region of the Y-chromosome of *bos Taurus*.<sup>a</sup>

Invader probe	Sequence	Probe duplex	5'-Inv:cDNA 3'-Inv:cDNA	TA
ON15	5'- <b>Cy3</b> <u>AGCCCUGTGCCCTG</u>	66.0	69.5	+17.0
ON16	3'- <u>TCGGGACACGGGAC</u> <b>Cy3</b>		74.0	
ON15	5'- <b>Cy3</b> <u>AGCCCUGTG</u> <u>CCCTG</u>	41.5	69.5	+30.0
ON17	3'- <u>TCGGGACAC</u> <b>4</b> <u>GGGAC</u> <b>Cy3</b>		62.5	
ON15	5'- <b>Cy3</b> <u>AGCCCUGTG</u> <u>CCCTG</u>	32.5	69.5	+34.0
ON18	3'- <u>TCGGGACAC</u> <b>9</b> <u>GGGAC</u> <b>Cy3</b>		57.5	
ON19	5'- <b>Cy3</b> <u>AGCGCUGAG</u> <u>GCCTG</u>	43.5	78.5	+36.0
ON20	3'- <u>TCGCGACTC</u> <b>4</b> <u>CGGAC</u> <b>Cy3</b>		64.5	

<sup>a</sup>  $\Delta T_m$  = change in  $T_m$  relative to unmodified dsDNA ( $T_m$  = 60.5 °C or 63.5 °C in the case of MM probes); thermal denaturation curves were recorded in medium salt buffer ([Na<sup>+</sup>] = 110 mM, [Cl<sup>-</sup>] = 100 mM, pH 7.0 (NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>), [EDTA] = 0.2 mM) and [ON] = 1.0  $\mu$ M.  $TA = T_m$  (5'-Inv:cDNA + 3'-Inv:cDNA) –  $T_m$  (Invader probe + dsDNA). **Cy3** = cyanine 3 phosphoramidite; **2** = 1-amino-3-hydroxyprop-2-yl monomer; **4** = 4-hydroxybutyl monomer; **9** = 9-hydroxynonyl monomer; **A/U/C** = 2'-O-(pyren-1-yl)methyl RNA monomers. nt = no transition.

## 4.5 References

1. Besch, R.; Giovannangeli, C.; Degitz, K. *Curr. Drug Targets* **2004**, *5*, 691.
2. Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. J. *Mol. BioSyst.* **2006**, *2*, 551.
3. Ackermann, D.; Famulok, M. *Nucleic Acids Res.* **2013**, *41*, 4729.
4. Aiba, Y.; Sumaoka, J.; Komiyama, M. *Chem. Soc. Rev.* **2011**, *40*, 5657.
5. Nielsen, P. E. *Chem. Biodiv.* **2010**, *7*, 786.
6. Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, *36*, 5123.
7. Nielsen, P. E., Egholm, M., Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
8. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284.
9. Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem.* **2013**, *21*, 6101.
10. Bahal, R.; Sahu, B.; Rapireddy, S.; Lee, C-M.; Ly, D. H. *ChemBioChem* **2012**, *13*, 56.
11. Chen, S. X.; Zhange, D. Y.; Seelig, G. *Nature Chem.* **2013**, *5*, 782.
12. Kutuyavin, I. V.; Rhinehart, R. L.; Lukhtanov, E. A.; Gorn, V. V.; Meyer Jr., R. B.; Gamper Jr., H. B. *Biochemistry* **1996**, *35*, 11170.
13. Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11804.
14. Sumaoka, J.; Komiyama, M. *Chem. Lett.* **2014**, *43*, 1581.
15. Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2010**, *8*, 2028.
16. Crothers, D. M. *Biopolymers* **1968**, *6*, 575.
17. 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.

18. Karmakar, S.; Guenther, D. C.; Hrdlicka, P. J. *J. Org. Chem.* **2013**, *78*, 12040.
19. Karmakar, S.; Madsen, A. S.; Guenther, D. C.; Gibbons, B. C.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2014**, *12*, 7758.
20. Anderson, B. A.; Onley, J. J.; Hrdlicka, P. J. *J. Org. Chem.*, **2015**, *80*, 5395.
21. Guenther, D. C.; Anderson, G. H.; Karmakar, S.; Anderson, B. A.; Didion, B. A.; Guo, W.; Verstegen, J. P.; Hrdlicka, P. J. *Chem. Sci.*, **2015**, *6*, 5006.
22. Guenther, D. C.; Karmakar, S.; Hrdlicka, P. J. *Chem. Comm.* **2015**, *51*, 15051.
23. Pyshnyi, D. V.; Lomzov, A. A.; Pyshnaya, I. A.; Ivanova, E. M. *J. Biomol. Struct. Dyn.* **2006**, *23*, 567.
24. Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem.* **2011**, *76*, 7119.
25. Morgan, M. A.; Okamoto, K.; Kahn, J. D.; English, D. S. *Biophys. J.* **2005**, *89*, 2588.
26. Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515.

## CHAPTER 5: Summary and conclusions

The work presented in this dissertation has focused on the optimization of Invader probe architectures using 2'-*O*-(pyren-1-yl)methyl RNA and 2'-*N*-(pyren-1-yl)methyl-2'-*N*-2'-amino DNA monomers. A library of 13 mer probes with varying numbers and positions of energetic hotspots consisting of second-generation monomers was characterized with respect to thermal denaturation temperatures, thermodynamic parameters, and recognition of a model dsDNA hairpin target. The energetic gradient between the low affinity probe duplexes and high affinity duplexes of individual strands with complementary DNA (cDNA) increases with incorporation of more energetic hotspots. The thermodynamic driving force translates to more efficient recognition of DNA hairpin targets at non-denaturing conditions. The highly modified Invader probe with four energetic hotspots shows reduced discrimination of targets differing in sequence at one position relative to the target hairpin DNA. The less modified probes, however, display excellent discrimination of mismatched targets, which highlights the importance of balancing efficient but specific recognition by Invader probes. From this study, it is suggested there should be approximately one energetic hotspot for every 4-5 base pairs in the Invader probe.

Bulged Invader probes were subsequently pursued following the hypothesis that introduction of non-nucleosidic bulges within the probe will generate more labile probe duplexes, thus increasing the thermodynamic driving force for recognition of dsDNA. Additionally, the non-nucleosidic bulges likely result in locally perturbed regions where the Watson-Crick face is transiently exposed and can more readily participate in nucleation and subsequent duplex invasion of target DNA. A library of 18 mer Invader



probes with different bulge monomers located centrally was synthesized and evaluated with respect to thermal hybridization properties and recognition of model hairpin DNA targets. Probes with a single non-nucleosidic bulge of ~12 atoms is sufficient to break cooperativity of the probe duplex, which results in more efficient recognition of dsDNA than the conventional probes lacking bulges. To further optimize bulged Invader probes, a library of 13 mer probes with nonyl bulges located near the termini were synthesized and evaluated for thermodynamic and kinetic optimization of dsDNA recognition. This study demonstrated the superior rate of recognition of bulged Invader probes.

Optimized Invader probes were used in non-denaturing fluorescence in situ hybridization assays. Fluorophore-labeled probes targeting a specific region of the DYZ-1 satellite repeat were incubated with isolated interphase or metaphasic nuclei, resulting in punctate fluorescent signals corresponding to specific recognition of chromosomal dsDNA under non-denaturing conditions. Although these preliminary results are exciting and represent quite an accomplishment for Invader probes, some limitations still exist such as: currently limited to targeting highly repeated sequences to achieve the signal intensity necessary for visualization using fluorescence microscopy, it will be necessary to address cellular/nuclear uptake and enzymatic stability for live cell applications, and elucidate which regions of chromosomal DNA are accessibility by Invader probes. Chemical modifications such as incorporation of polycationic moieties or phosphorothioate may mediate some of these limitations.

DNA targeting agents have use as diagnostic probes but also offer the potential to interrupt gene expression. Invader probes were evaluated in an in vitro transcription assay, where the probe is designed to target a phage promoter or mid-transcript region of

linearized plasmid DNA. Preliminary results suggest that indeed Invader probes inhibit RNA synthesis by RNA polymerase. However, this does not seem to occur due to specific binding of the probe to the DNA target, as probes differ in three positions relative to the target sequence also inhibit RNA synthesis. A co-localization assay revealed that both the targeting probe and control probe that is not thermally activated to bind co-migrated with the plasmid DNA. To further the use of Invader probes for inhibition of in vitro transcription, a better understanding of the possible binding modes and control of specificity is needed.

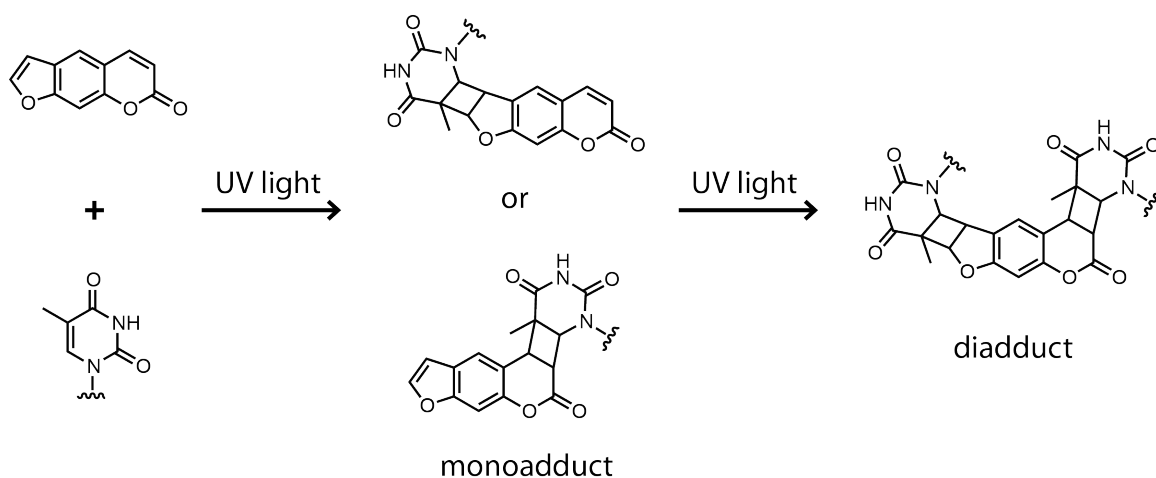
The Invader approach for recognition of dsDNA offers a promising tool to enable applications in fundamental molecular biology such as structural and functional elucidation of chromosomal DNA, antigene agents to modulate gene expression, or even to induce homologous recombination for gene editing.

## **APPENDIX A: Psoralen-modified Invader probes for sequence-specific recognition and photoactivated interstrand cross-linking of dsDNA**

Dale C. Guenther, Saswata Karmakar, Patrick J. Hrdlicka

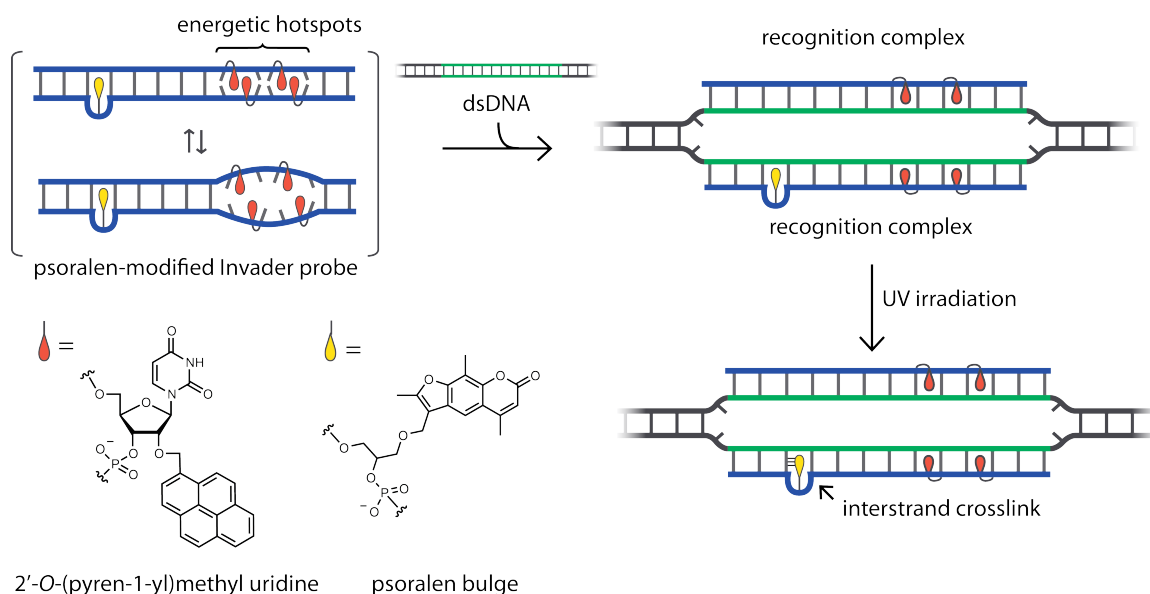
### **A.1 Introduction**

Chemical based probes capable of inducing site-specific interstrand cross-links (ICL) have been developed for applications such as nucleic acid detection,<sup>1</sup> fluorescent in situ hybridization assays (FISH),<sup>2</sup> and as a tool to induce sequence specific mutations.<sup>3</sup> Psoralen is one of the most important classes of photochemical reagents for inducing ICL.<sup>4</sup> Psoralen can intercalate into the core of the DNA duplex, and then upon UVA irradiation (365 nm), it first forms either a furan (FMA) or pyrone monoadduct (PMA) by reacting in a 2+2 cycloaddition with thymine (Figure A.1-1). A diadduct can then form from the reaction between the other photoreactive double bond (furan or pyrone) and a second thymine, thus cross-linking the two strands of the duplex.<sup>5</sup>



**Figure A.1-1.** Structure of psoralen and the adducts formed with thymine after irradiation with UVA light.<sup>5</sup>

We have developed Invader probes that enable specific recognition of mixed-sequence i) hairpin DNA, ii) linearized plasmid DNA, and iii) chromosomal DNA in isolated nuclei.<sup>6</sup> We recently introduced bulged Invader probes that improve the efficiency of dsDNA recognition by increasing the rate of recognition and lowering the concentration necessary to realize recognition.<sup>7</sup> Incorporation of psoralen bulges in Invader probes is particularly well-suited for applications involving sequence-specific cross-linking, since intercalation can potentially be controlled with high precision. In this manner, intercalation is hypothesized to occur with nucleobase resolution, thus allowing for precise positioning of the psoralen adjacent to a thymine and increasing the efficiency of cross-linking compared to terminus modification, where intercalation can occur between any base pairs within proximity of the long linker. Towards this end, we synthesized a series of psoralen-modified Invader probes where the psoralen was incorporated as either a bulge or as a replacement of the canonical 2'-deoxyadenosine, both of which are anticipated to position the psoralen moiety near thymines for efficient cross-linking.



**Figure A.1-2.** Illustration of psoralen-modified Invader probes and ICL formation after UV irradiation.

## A.2 Results and discussion

Thermal denaturation temperatures were determined for psoralen-modified Invader probes and duplexes between individual probes strands and complementary DNA (cDNA; Table A.2-1). Conventional probe strands (**ON1** or **ON2**) have very high affinity towards cDNA ( $\Delta T_m = +14.0$  to  $+18.0$  °C). Introduction of a psoralen bulge (**ON3** or **ON4**) only destabilizes the probe duplex by 2.0 to 2.5 °C relative to duplexes involving the conventional probe. This is likely because intercalation of the psoralen moiety increases stacking in the duplex, thereby largely counteracting the destabilizing effect of a non-nucleosidic bulge. Along these lines, it is interesting to note the more favorable enthalpic contributions (e.g., compare **ON1**:cDNA with **ON3**:cDNA, Table A.4-3). Probe strands featuring the psoralen moiety as a nucleobase replacement (**ON5** and **ON6**) display lower affinity towards cDNA than probe strands with psoralen bulges (**ON3** and

**ON4**). Of note, all examined probe:cDNA duplexes are significantly more stable than unmodified dsDNA ( $T_m = 37.5$  °C).

Surprisingly, Invader probes with psoralen bulges on both strands (e.g., **ON3:ON4** and **ON5:ON6**) result in very stable duplexes, which negatively impacts the energetic gradient for dsDNA recognition. The stabilization is of enthalpic nature (Table A.4-3), and is most likely a result of  $\pi$ - $\pi$  stacking interactions between the psoralen moieties and nearby nucleobases in the duplex core. Although it is desirable to have both strands capable of cross-linking (i.e. featuring a psoralen moiety), the thermal advantage ( $TA$ ), which is an estimate of the thermodynamic potential for dsDNA recognition as defined as  $T_m(5'\text{-Inv:cDNA} + 3'\text{-Inv:cDNA}) - T_m(\text{probe duplex} + \text{dsDNA})$ , is relatively low for these constructs (+17.0 and +7.0 °C for **ON3:ON4** and **ON5:ON6**, respectively). We therefore pursued a probe design that features only one psoralen unit. The double-stranded probes have similar or decreased stability compared to the conventional probe, which results in favorable energetics for recognition (e.g., compare  $TA$  for **ON3:ON2** vs **ON1:ON2**, Table A.2-1). Another design that breaks the stabilizing effect when psoralens are located across from each other, is to use a mixed motif, where one probe strand features the psoralen as a bulge while the other probe strand has the psoralen as a nucleobase surrogate (**ON5:ON4** and **ON3:ON6**). These probes are labile and thus possess a favorable energetic gradient for recognition.

**Table A.2-1.** Thermal hybridization properties of psoralen modified Invader probes.

Invader probe	Sequence	$T_m$ [ $\Delta T_m$ ] (°C)			TA (°C)
		Probe duplex	5'-Inv: cDNA	3'-Inv: cDNA	
1 2	5'-GGX AX ATATAGGC 3'-CCA XA XATATCCG	40.0 [+2.5]	51.5 [+14.0]	55.5 [+18.0]	29.5
3 4	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA <sup>P</sup> TCCG	49.0 [+11.5]	49.5 [+12.0]	54.0 [+16.5]	17.0
3 2	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA TCCG	40.0 [+2.5]	49.5 [+12.0]	55.5 [+18.0]	27.5
1 4	5'-GGX AX ATAT AGGC 3'-CCA XA XATA <sup>P</sup> TCCG	40.5 [+3.0]	51.5 [+14.0]	54.0 [+16.5]	27.5
5 6	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATA <sup>P</sup> CCG	49.5 [+12.0]	44.5 [+7.0]	49.5 [+12.0]	7.0
5 2	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATATCCG	36.5 [-1.0]	44.5 [+7.0]	55.5 [+18.0]	26.0
1 6	5'-GGX AX ATATAGGC 3'-CCA XA XATA <sup>P</sup> CCG	35.5 [-2.0]	51.5 [+14.0]	49.5 [+12.0]	28.0
5 4	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATA <sup>P</sup> TCCG	39.5 [+2.0]	44.5 [+7.0]	54.0 [+16.5]	21.5
3 6	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA <sup>P</sup> CCG	33.0 [-4.5]	49.5 [+12.0]	49.5 [+12.0]	28.5

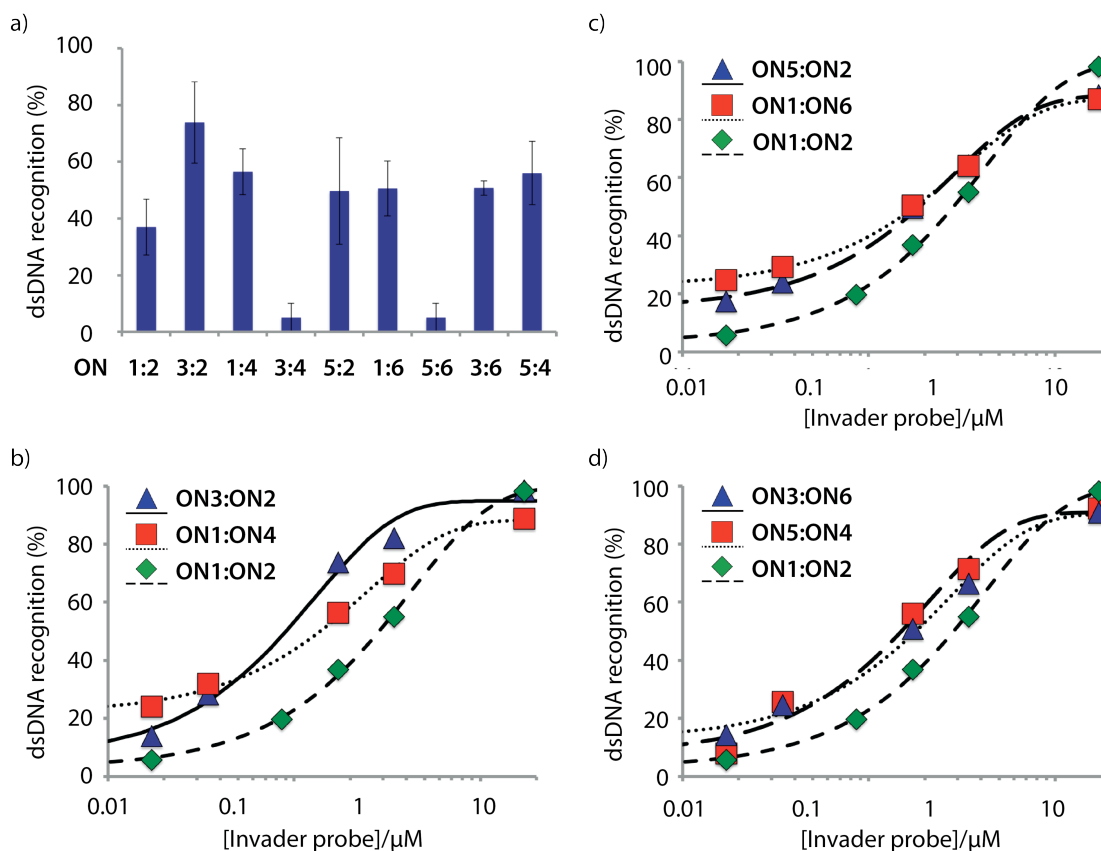
<sup>a</sup>  $\Delta T_m$  is calculated relative to the corresponding unmodified dsDNA ( $T_m = 37.5$  °C; Thermal denaturation curves were recorded in medium salt phosphate buffer ([Na<sup>+</sup>] = 110 mM, [Cl<sup>-</sup>] = 100 mM, pH 7.0 (NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>), [EDTA] = 0.2 mM) and each [ON] = 0.5  $\mu$ M. P = 1-methoxy-3-((2,5',8-trimethyl)-4'-psoralen)propan-2-yl monomer A = adenin-9-yl DNA monomer, C = cytosin-1-yl DNA monomer, G = guanin-9-yl DNA monomer, T = thymin-1-yl DNA monomer.

An electrophoretic mobility shift assay (EMSA) was used to validate dsDNA recognition of these probes. Briefly, Invader probes were incubated with a DIG-labeled DNA hairpin, where the stem region of the hairpin is complementary to the Invader probe. A successful recognition event results in opening of the stem, which is connect via a T<sub>10</sub> loop, and formation of a ternary complex. On non-denaturing gel electrophoresis, the DNA hairpin

and ternary recognition complex have significantly different mobilities, allowing for quantification of the extent of dsDNA recognition.

An initial screen of all constructs this assay revealed a pattern that follows the trends based on the thermodynamics. All probes save for **ON3:ON4** and **ON5:ON6**, invade the DNA hairpin. Interestingly, the concentration dependence of the probes reveals that the conventional probe (**ON1:ON2**) is the least efficient at recognition of dsDNA with a  $C_{50}$ -value, the concentration at which 50% recognition is achieved, of  $2.9 \mu\text{M}$ . The psoralen-modified probes have similar activity, with  $C_{50} = 0.9 - 1.9 \mu\text{M}$ .

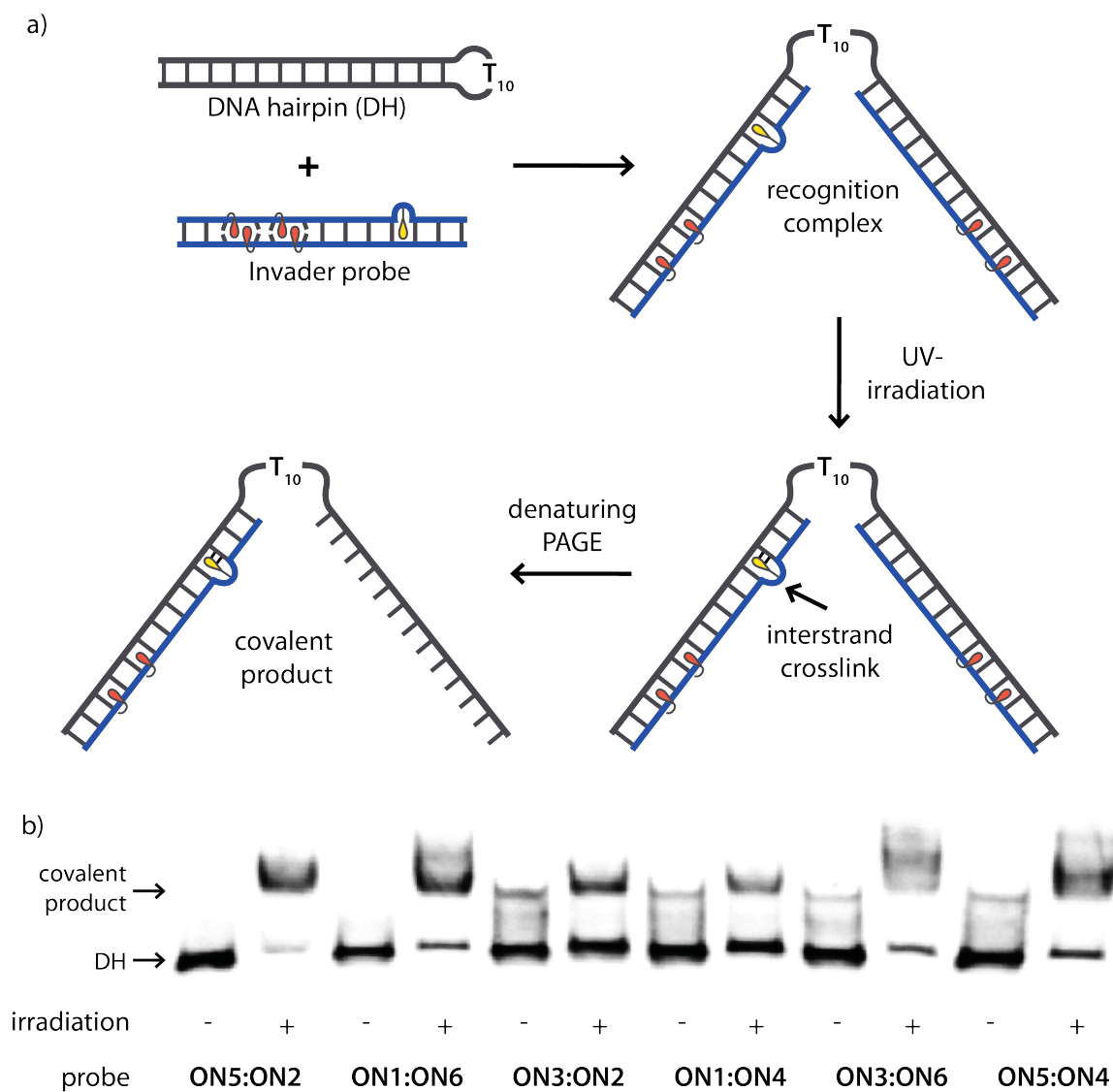




**Figure A.2-1.** Recognition of hairpin DNA (34.4 nM) by psoralen-modified Invader probes. Screen of probes at 100-fold molar excess (a) and concentration dependence (b-d) of recognition after incubation with DNA hairpin for 17 h at 22 °C. dsDNA recognition is determined from densitometry measurements as the ratio of recognition complex/DNA hairpin for a given lane. Representative electrophoretograms shown in Figure A.4-2.

The psoralen-modified Invader probes were then evaluated for interstrand cross-linking of the recognition complex by irradiation with UV-light (364 nm, 8 Watt for 1 h; Figure A.2-2). Denaturing PAGE reveals that covalently linked complexes with lower mobility are formed after irradiation (+). The small bands with the same mobility as a covalently linked recognition complex observed in the absence of irradiation, most likely represent incomplete denaturation of the recognition complex at 150 V (constant voltage) since subsequent experiments, in which constant power (40 W) was used during electrophoresis did not show evidence of this. These images are not shown because they

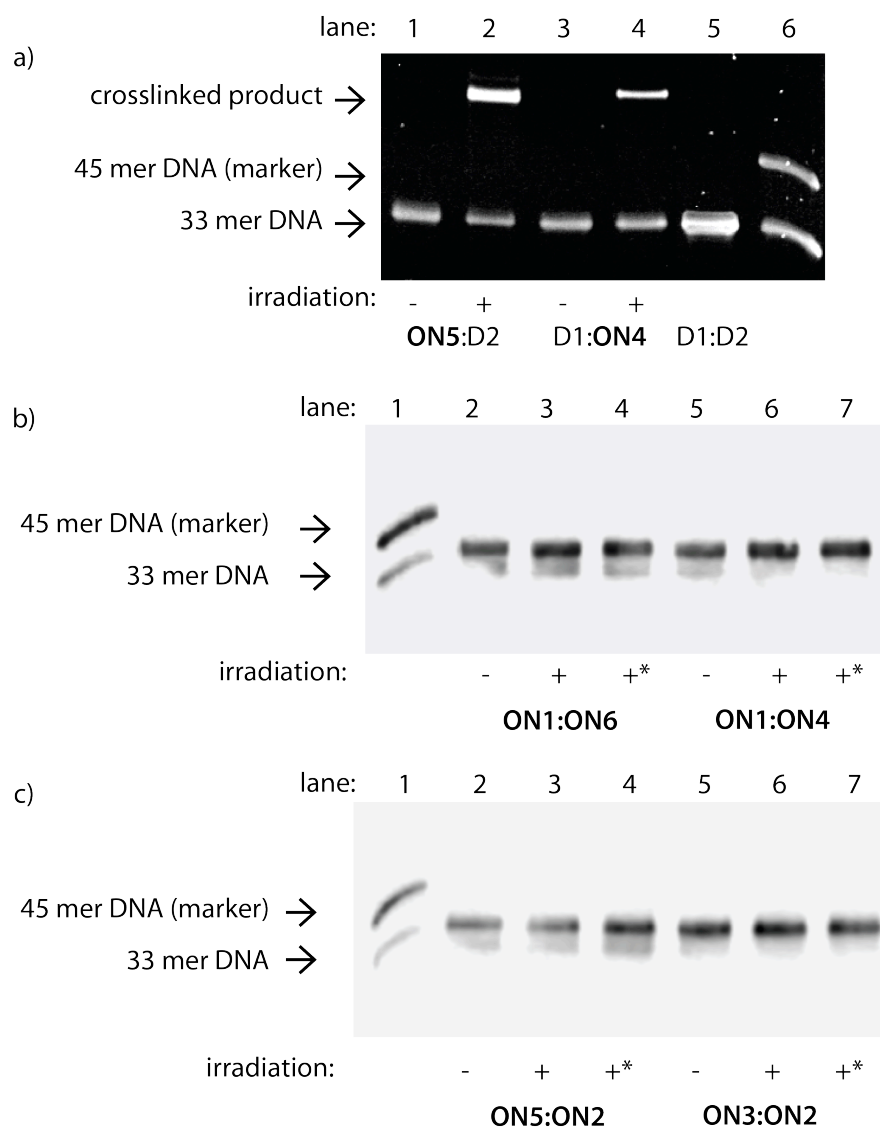
do not include good images for all of the Invader probes constructs under these conditions. The persistence of a slower mobility band after irradiation is consistent with a covalent interstrand cross-link between the psoralen on the Invader probe strand and the DNA hairpin. Further evidence supporting the formation of that interstrand cross-links was obtained from thermal denaturation curves of **ON4**:cDNA, in which the sigmoidal curve becomes less defined with increasing time of UV-light exposure (Figure A.4-3).



**Figure A.2-2.** Illustration of DNA hairpin recognition using psoralen-modified Invader probe, followed by irradiation to form stable thymine adducts that covalently cross-links the probe and DNA hairpin (a). Representative electrophoretograms (b) of denaturing PAGE after incubation of psoralen-modified Invader probes (200-fold molar excess) with DNA hairpin followed by UV-irradiation (3 h) and denaturing PAGE (7M urea). For incubation conditions see A.2-1. Note: the above electrophoretogram was compiled from two images.

In an effort to compare the cross-linking efficiency of Invader probes with psoralen as a bulge to probes with psoralen conjugated to the terminus and to demonstrate sequence-specific cross-linking of longer, linear dsDNA, we used a 33 bp dsDNA with an internal

target region. Initially, the cross-linking was verified by annealing the individual probe strands with the 33mer single-stranded target and irradiating the duplex with UV-light (365 nm), followed by staining with Sybr gold after denaturing PAGE (Figure A.2-3). Indeed, a lower mobility band is observed only when the duplexes were irradiated (lanes 2 and 4).



**Figure A.2-3.** Cross-linking between individual probe strands annealed with single-stranded 33 mer DNA (D1 or D2), demonstrated using denaturing PAGE and visualized by Sybr gold stain (a). Psoralen-modified Invader probes incubated with DIG-labeled 33 bp dsDNA (b and c) for 17 h, followed by irradiation with 365 nm light (+) or exposed to ambient light (-) for 1 h. Electrophoretograms were captured on x-ray film after processing for chemiluminescent signal with anti-DIG alkaline phosphatase. For incubation conditions see Figure A.2-1. \*Invader probe annealed with the dsDNA (90 °C for 10 min and cooled to RT over 15 min) followed by irradiation with UV-light. D1: 5'-AAGCTGCACAGGTATATATAGGCCGCATATGCA, D2: 3'-TTCGACGTGTCCATATATATCCGGCGTATACGT. Underlined region corresponds to embedded target sequence.

Encouraged by these results, we incubated psoralen-modified Invader probes with 33 bp DIG-labeled dsDNA containing the target sequence, irradiated the samples, followed by denaturing PAGE (Figure A.2-3b and c). Surprisingly, no low-mobility bands were observed (lanes 3 and 6) corresponding to either a lack of target recognition or lack of cross-linking. Annealing the reaction mixture by heating to 90 °C and cooling did not improve the results (lanes 4 and 7). A non-denaturing PAGE was performed to evaluate whether the recognition complex initially formed, and indeed no invasion of the longer target is observed (data not shown). This Invader probe core architecture appears does not have favorable enough thermodynamic potential for stable recognition of 33 bp linear dsDNA. This observation is also corroborated by an unrelated project, where Invader probe consisting of two zipper modifications centrally located also does not result in invasion of 33 bp dsDNA.

### **A.3 Conclusion**

Psoralen-modified Invader probes efficiently recognize DNA hairpins and form ICL upon UV-irradiation. They offer a unique opportunity for sequence-specific ICL with nucleobase resolution. To advance this project for recognition of linear dsDNA, a superior probe architecture is needed, such as Invader probes with toe-holds (single stranded overhangs), which preliminary data suggests would result in efficient invasion and allow for evaluation of precisely positioned psoralen moieties for sequence specific cross-linking with nucleobase resolution. Alternately, the 33 bp targets may not be optimal and a longer dsDNA (150 bp) may represent a better model target.

## A.4 Supporting information

### *Protocol - synthesis and purification of ONs*

Modified ONs were synthesized on an automated DNA synthesizer (0.2  $\mu$ mol scale) using a long chain alkyl amine controlled pore glass (LCAA-CPG) solid support with a pore size of 500 Å. The corresponding phosphoramidite of 2'-*O*-(pyren-1-yl)methyl uridine was prepared as previously described<sup>8</sup> and incorporated into ONs via hand-couplings (0.05 M in acetonitrile, using 0.01 M 4,5-diccanoimidazole as the activators (15 min)) with extended oxidation (45 s). Psoralen monomer, P, was incorporated in a similar manner using the commercially available psoralen phosphoramidite (ChemGenes). Treatment with 32% ammonia (55 °C, 17 h) facilitated deprotection and cleavage from solid support. DMT-protected ONs were purified via ion-pair reverse phase HPLC (XTerra MS C18 column: 0.05 M triethyl ammonium acetate and acetonitrile gradient) followed by detritylation (80% acetic acid, 20 min) and precipitation (NaOAc, NaClO<sub>4</sub>, acetone, -18 °C, 16 h). The purity and identity of synthesized ONs were verified using analytical HPLC (>85% purity) and MALDI-MS analysis (Table A.4-1) recorded on a Quadrupole Time-of-Flight (Q-TOF) mass spectrometer with 2,4,6-trihydroxyacetophenone as a matrix.

*Protocol - thermal denaturation experiments*

The concentrations of ONs were estimated using the following extinction coefficients ( $OD_{260}/\mu\text{mol}$ ): G (12.01), A (15.20), T (8.40), C (7.05), pyrene (22.4), and psoralen (16.50).<sup>9</sup> Thermal denaturation temperatures were calculated as the first-derivative maximum of  $A_{260}$  vs  $T$  curves. ONs (0.5  $\mu\text{M}$ ) were heated (85 °C, 2 min) in medium salt buffer ( $[\text{Na}^+] = 110 \text{ mM}$ ,  $[\text{Cl}^-] = 100 \text{ mM}$ , pH 7.0 ( $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ ),  $[\text{EDTA}] = 0.2 \text{ mM}$ ) and subsequently cooled to the starting temperature. The experimental temperature ranged from 5 °C to at least 15 °C above the  $T_m$ , with the  $T_m$  determined as the average of two experiments within  $\pm 1.0$  °C. Thermodynamic parameters were obtained from the denaturation curves via baseline fitting (van't Hoff method). Biomolecular reactions, two-state melting behavior, and constant heat capacity were assumed.<sup>10</sup> Values reported are the average of values obtained from two different melting curves analyzed three times each.

*Protocol - electrophoretic mobility shift assay*

Linear DNA strands and DNA hairpin were obtained from commercial sources and used without further purification. They were labeled using the 2<sup>nd</sup> generation DIG Gel Shift Kit (Roche Applied Bioscience). Briefly, 11-digoxigenin-ddUTP was incorporated at the 3'-end of the strand (100 pmol) using a recombinant DNA terminal transferase. The reaction mixture was quenched by addition of EDTA (0.05 M), diluted to 68.8 nM, and used without further processing. The recognition experiments were conducted essentially as previously reported.<sup>6</sup> Thus, Invader probes (variable concentrations) were annealed (90 °C for 2 min, followed by cooling to room temperature) and subsequently incubated with

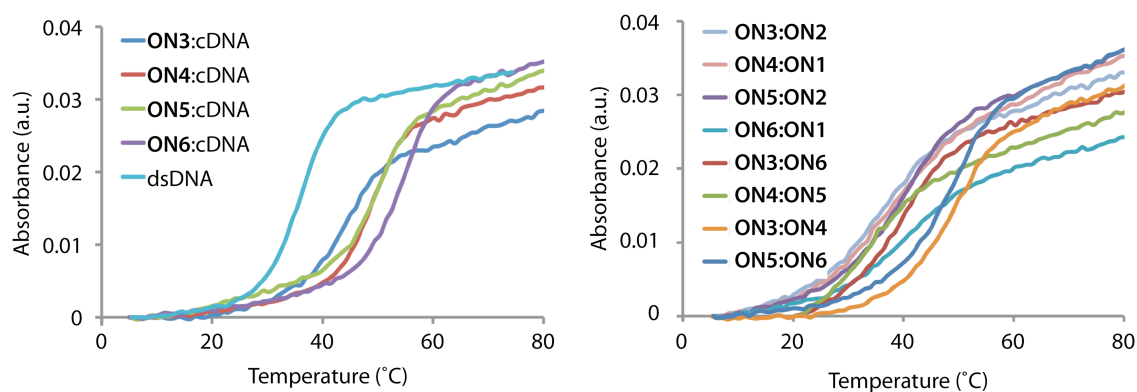


DIG-labeled DNA hairpins (34.4 nM final concentration) in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahydrochloride) at 22 °C ± 2 °C for a specified period. Select samples were irradiated for 3 h with a hand-held UV lamp (365 nm, 8-Watt). Loading dye (6X) was added and the reaction mixtures were loaded onto 12% non-denaturing TBE-PAGE (45 mM tris-borate, 1 mM EDTA; acrylamide:bisacrylamide (19:1); 70 V at ~4 °C for 1.5 h) or 16 - 20% denaturing PAGE (45 mM tris-borate, 1 mM EDTA, 7 M urea, (19:1); 150 V at RT for ~30 min). Gels were either stained (Sybr gold) and visualized directly or the bands were blotted onto positively charged nylon membranes (100 V, 30 min, ~4 °C) and cross-linked through exposure to UV light (254 nm, 5 × 15 watt bulbs, 3 min). Membranes were incubated with anti-digoxigenin-alkaline phosphatase F<sub>ab</sub> fragments as recommended by manufacturer, and transferred to a hybridization jacket. Membranes were incubated with the chemiluminescence substrate (CSPD) for 10 min at 37 °C, and chemiluminescence was captured on X-ray films. Digital images of developed X-ray films were obtained using a Fluor-S MultiImager and quantified using appropriate software (Quantity One). The percentage of dsDNA recognition was calculated as the intensity ratio between the recognition complex band and the total lane. Unless otherwise noted, an average of three independent experiments is reported along with standard deviations (±).

**Table A.4-1.** MALDI-MS of psoralen-modified ONs.<sup>a</sup>

ON	Sequence	Observed $m/z$ [M+H] <sup>+</sup>	Calculated $m/z$ [M+H] <sup>+</sup>
ON3	5'-GGTATATATPGGC	4525.5	4524.9
ON4	3'-CCATATATAPCCG	4414.4	4413.8
ON5	5'-GGTATATATPAGGC	4838.5	4837.9
ON6	3'-CCATATATAPTCCG	4718.5	4717.9

<sup>a</sup> For structure of modifications, see Figure A.1-2.



**Figure A.4-1.** Representative thermal denaturation curves of individual probe strands with complementary DNA (left) and Invader probe duplexes (right).

**Table A.4-2.** Change in Gibbs free energy ( $\Delta G$ ) at 298 K for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:cDNA and 3'-Inv:cDNA), and change in reaction free energy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta G_{rec}^{298}$ ).<sup>a</sup>

Invader probe	Sequence	$\Delta G^{298}$ (kJ/mol)			$\Delta G_{rec}^{298}$ (kJ/mol)
		Probe duplex	5'-Inv:cDNA	3'-Inv:cDNA	
1 2	5'-GGX AX ATATAGGC 3'-CCA XA XATATCCG	-51	-69	-75	-40
3 4	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA <sup>P</sup> TCCG	-60	-73	-77	-37
3 2	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA TCCG	-51	-73	-75	-44
1 4	5'-GGX AX ATAT AGGC 3'-CCA XA XATA <sup>P</sup> TCCG	-50	-69	-77	-43
5 6	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATA <sup>P</sup> CCG	-60	-60	-71	-18
5 2	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATATCGG	-47	-60	-75	-35
1 6	5'-GGX AX ATATAGGC 3'-CCA XA XATA <sup>P</sup> CCG	-47	-69	-71	-40
5 4	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATA <sup>P</sup> TCCG	-57	-60	-77	-27
3 6	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA <sup>P</sup> CCG	-44	-73	-71	-47

<sup>a</sup>  $\Delta G_{rec}^{298} = \Delta G^{298}(5'\text{-Inv:cDNA}) + \Delta G^{298}(3'\text{-Inv:cDNA}) - \Delta G^{298}(\text{Invader probe}) - \Delta G^{298}(\text{dsDNA})$ .  $\Delta G$  for unmodified DNA duplex = -53 kJ/mol. For experimental conditions see Table A.2-1.

**Table A.4-3.** Change in enthalpy ( $\Delta H$ ) at 298 K for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in enthalpy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{rec}$ ).<sup>a</sup>

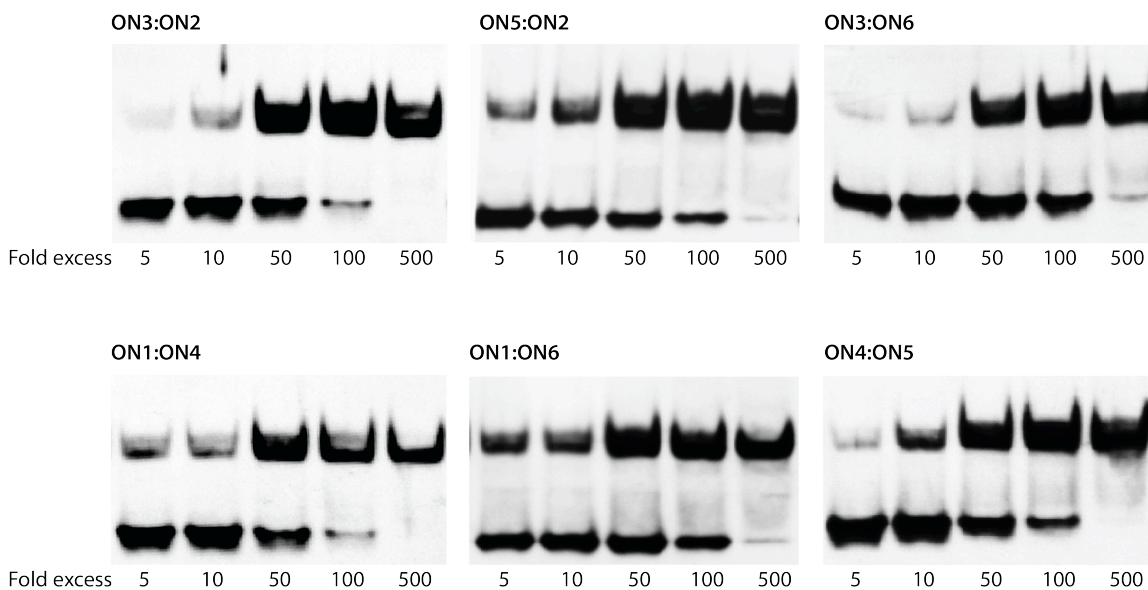
Invader probe	Sequence	$\Delta H$ (kJ/mol)			$\Delta H_{rec}$ (kJ/mol)
		Probe duplex	5'-Inv: cDNA	3'-Inv: cDNA	
1 2	5'-GGX AX ATATAGGC 3'-CCA XA XATATCCG	-280	-397	-423	-128
3 4	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA <sub>P</sub> TCCG	-321	-456	-443	-166
3 2	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA TCCG	-290	-456	-423	-177
1 4	5'-GGX AX ATAT AGGC 3'-CCA XA XATA <sub>P</sub> TCCG	-248	-397	-443	-180
5 6	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATA <sub>P</sub> CCG	-338	-376	-449	-75
5 2	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATATCCG	-268	-376	-423	-119
1 6	5'-GGX AX ATATAGGC 3'-CCA XA XATA <sub>P</sub> CCG	-177	-397	-449	-257
5 4	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATA <sub>P</sub> TCCG	-380	-376	-443	-27
3 6	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA P CCG	-278	-456	-449	-215

<sup>a</sup>  $\Delta H_{rec} = \Delta H(5'\text{-Inv:cDNA}) + \Delta H(3'\text{-Inv:cDNA}) - \Delta H(\text{Invader probe}) - \Delta H(\text{dsDNA})$ .  $\Delta H$  for unmodified DNA duplex = -412 kJ/mol. For experimental conditions see Table A.2-1.

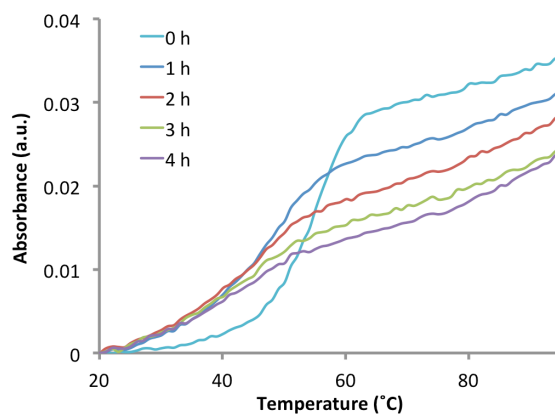
**Table A.4-4.** Change in entropy ( $-T^{298}\Delta S$ ) at 298 K for formation of probe duplexes (Invader probe), and duplexes between individual probe strands and complementary DNA (5'-Inv:DNA and 3'-Inv:DNA), and change in entropy upon Invader-mediated recognition of isosequential dsDNA targets ( $-T^{298}\Delta S_{rec}$ ).<sup>a</sup>

Invader probe	Sequence	$-T^{298}\Delta S$ (kJ/mol)			$-T^{298}\Delta S_{rec}$ (kJ/mol)
		Probe duplex	5'-Inv: cDNA	3'-Inv: cDNA	
1 2	5'-GGX AX ATATAGGC 3'-CCA XA XATATCCG	229	328	348	88
3 4	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA <sup>P</sup> TCCG	261	383	366	129
3 2	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA TCCG	239	383	348	133
1 4	5'-GGX AX ATAT AGGC 3'-CCA XA XATA <sup>P</sup> TCCG	198	328	366	137
5 6	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATAPCCG	277	316	379	59
5 2	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATATCCG	221	316	348	84
1 6	5'-GGX AX ATATAGGC 3'-CCA XA XATAPCCG	196	328	379	152
5 4	5'-GGX AX ATAT <sup>P</sup> GGC 3'-CCA XA XATA <sup>P</sup> TCCG	324	316	366	-1
3 6	5'-GGX AX ATAT <sup>P</sup> AGGC 3'-CCA XA XATA <sup>P</sup> CCG	234	383	379	169

<sup>a</sup>  $-T^{298}\Delta S_{rec} = -T^{298}\Delta S(5'-Inv:cDNA) + -T^{298}\Delta S(3'-Inv:cDNA) + T^{298}\Delta S(\text{Invader probe}) + T^{298}\Delta S(\text{dsDNA})$ .  $-T^{298}\Delta S$  for unmodified DNA duplex = 359 kJ/mol. For experimental conditions see Table A.2-1.



**Figure A.4-2.** Representative electrophoretograms for data shown in Figure A.2-1.



**Figure A.4-3.** Thermal denaturation curve of ON4:cDNA with increasing exposure of UV-light (365 nm).

#### A.4 References

1. Peng, X.; Greenberg, M. M. *Nucleic Acids Res.* **2008**, *5*, e31.
2. Johnson, M. D.; Fresco, J. R. *Chromosoma* **1999**, *108*, 181.
3. Singer, M. J.; Podymingon, M. A.; Metcalf, M. A.; Reed, M. W.; Brown, D. A.; Gamper, H. B.; Meyer, R. B.; Wydro, R. M. *Nucleic Acids Res.* **1999**, *27*, e38.

4. Cimino, G. D.; Gamper, H. B.; Isaacs, S. T.; Hearst, J. E. *Ann. Rev. Biochem.* **1985**, *54*, 1151.
5. Serrano-Perez, J. J.; Merchan, M.; Serrano-Andres, L. *J. Phys. Chem. B* **2008**, *112*, 14002.
6. Guenther, D. C.; Anderson, G. H.; Karmakar, S.; Anderson, B. A.; Didion, B. A.; Guo, W.; Verstegen, J. P.; Hrdlicka, P. J. *Chem. Sci.*, **2015**, *6*, 5006.
7. Guenther, D. C.; Karmakar, S.; Hrdlicka, P. J. *Chem. Comm.* **2015**, *51*, 15051.
8. Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem.* **2011**, *76*, 7119.
9. Extinction coefficient at 260 nm for psoralen monomer obtained from Glen Research.
10. Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515.

## **APPENDIX B: Inhibition of in vitro transcription by Invader probes**

Dale C. Guenther, Saswata Karmakar, Patrick J. Hrdlicka

### **B.1 Introduction**

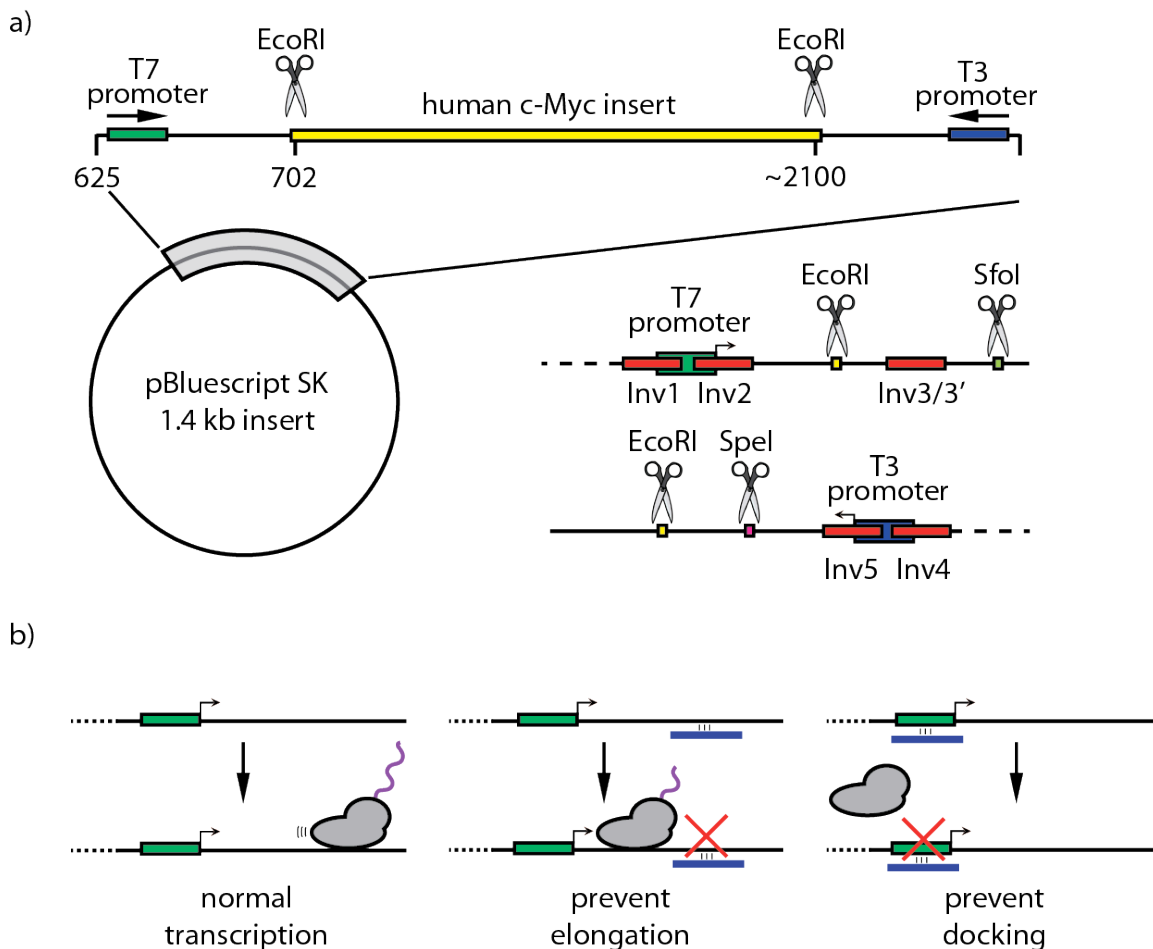
Invader probes have been shown to be able to target dsDNA hairpins and chromosomal DNA for diagnostic applications.<sup>1</sup> We aimed to evaluate these probes for use as antigene agents, which inhibit gene expression at the transcription level. As a proof-of-concept, we sought to use Invader probes to inhibit RNA synthesis in in vitro transcription assays. Double-stranded DNA templates contain a promoter region that RNA polymerases (RNAPs) bind to, resulting in initiation of RNA synthesis. Binding of Invader probes to the template DNA was hypothesized to interrupt this process.

### **B.2 Results and discussion**

We decided to use a plasmid for the double-stranded template DNA for the in vitro transcription assay. We initially utilized plasmid 14971 (addgene.org), which contains a human c-Myc insert (~1.4 kb) cloned into the pBlueScript SK vector at the EcoRI restriction site (Figure B.2-1). The plasmid contains two phage promoters, T7 and T3, facilitating RNAP-catalyzed RNA synthesis. The first 700 bp of the insert have been sequenced and are available at addgene.org. However, the precise length of the insert is not specified. I found the known 700 bp sequence in the human c-Myc gene and assumed the following 700 bp constituted the insert sequence. In hindsight, the plasmid should



have been sequenced for identity verification. For the in vitro transcription assay, RNA synthesis is terminated where the template dsDNA ends (run-off assay), which necessitates that the template DNA is cleaved or digested. Initial attempts at digestion to obtain appropriate length RNA fragments for transcription from the T3 promoter, assuming a 1400 bp insert, failed. Restriction sites that should be present in the 1400 bp insert did not result in digestion, even under optimized conditions, for BbvC and HpaI, corresponding to 1293 bp and 1079 bp of the insert, respectively. This indicates that either the assumption of the sequence is incorrect or that the insert is not 1400 bp. To address the latter, we used the EcoRI restriction site that was used for cloning, and not destroyed in the process, to digest the plasmid and isolate the fragment that was inserted. Figure B.4-1 shows the excised fragment, and it is approximately ~1400 bp, as reported. However, we cannot be certain of the sequence identity of the insert after the first 700 bp. This influenced the experimental design by limiting the target regions to sequences that we were confident were present (i.e. from the pBlueScript SK plasmid or the first 700 bp of insert).



**Figure B.2-1.** Architecture of plasmid 14971 (~4.4 kb), which contains a human c-Myc insert (~1.4 kb) in the pBluescript SK vector at the EcoRI restriction site (a). Red bars indicate Invader probe binding sites associated with either the T7 template or T3 template. (b) Possible mechanisms of Invader probe-mediated inhibition of in vitro transcription.

Invader probes were designed to target different regions involved in the in vitro transcription process, namely the promoter (Inv1 and Inv4), the initiation start site (Inv2 and Inv5), or mid template (Inv3) under control of T7 RNA polymerase (Inv1 – Inv3) or T3 RNA polymerase (Inv4 and Inv5). These regions were chosen due to potentially different mechanisms of inhibition (Figure B.2-1), where targeting the promoter region could prevent docking of the RNA polymerase (RNAP), which is the rate-limiting step for transcription,<sup>2</sup> and prevent initiation of transcription. Targeting the initiation start,

with concomitant overlap of the promoter region, is hypothesized to result in inhibition via a similar mechanism. This is an interesting region because the RNAP unwinds this region of the dsDNA (transcription bubble) allowing for elongation of the transcript that is complementary to the template strand of the dsDNA. The transcription bubble has previously been used to gain access to the Watson-Crick face of the dsDNA to allow for binding of high-affinity oligonucleotides,<sup>3</sup> and may facilitate binding of Invader probes as well. Recognition of a transcribed down-stream region, on the other hand, is hypothesized to result in inhibition of RNA synthesis by creating a 'road-block', where the bound Invader probe prevents the progression of the RNAP, resulting in dissociation of the polymerase and formation of a truncated transcript.<sup>2,4,5</sup>

Invader probes containing three energetic hotspots based on 2'-*O*-(pyren-1-yl)methyl RNA monomers were designed to recognize plasmid dsDNA target regions for use as antigene agents (Table B.2-1). Thermal denaturation temperatures of individual Invader probe strands with cDNA reveal highly stabilized duplexes ( $\Delta T_m = +6.5$  to  $+25.5$  °C) relative to corresponding unmodified dsDNA duplexes. Probe duplexes have lower  $T_m$ 's than the Inv:cDNA duplexes, but higher  $T_m$ 's than the corresponding unmodified dsDNA. Surprisingly, Inv3 does not result in a change in absorption at 260 nm within the temperature range of 20 – 85 °C. To understand if this is due to a lack of duplex formation or persistence of a very highly stable duplex, the probe duplexes were examined using non-denaturing PAGE. The assay revealed that Inv3 most likely exist as a duplex, although duplex formation is dependent on the salt concentration of the buffer that it is annealed in. The probe strands are dissociated in water but appear to be duplexed at 110 mM Na<sup>+</sup> (data not shown). The differences in  $T_m$  between the probe duplex and

probe:cDNA duplexes represents the thermal advantage ( $TA$ ), an indicator of the thermodynamic driving force for recognition of dsDNA. All of the studied probes show favorable  $TA$ .

**Table B.2-1.** Thermal denaturation temperatures ( $T_m$ 's) of Invader probes evaluated for inhibition of in vitro transcription.<sup>a</sup>

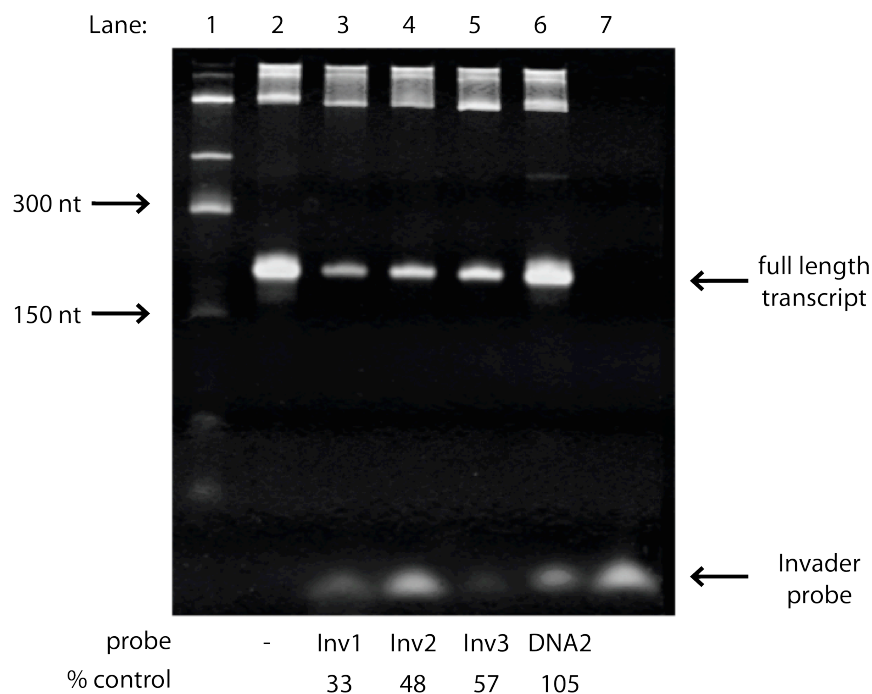
Invader probe	Sequence	$T_m$ [ $\Delta T_m$ ] (°C)			dsDNA	$TA$ (°C)
		5'-Inv: cDNA	3'-Inv: cDNA	Probe duplex		
1	5'-ATUG <u>UA</u> AATACGACTC 3'-TAACA <u>U</u> TATGCTGAG	57.0 [+12.0]	58.0 [+13.0]	30.0* [-15.0]	45.0	-
2	5'-ACU <u>U</u> AAGGGCGAATT 3'-TGAU <u>A</u> UCCCGCTTAA	67.0 [+18.0]	67.0 [+18.0]	52.5 [+3.5]	49.0	+32.5
3	5'-CCG <u>U</u> ATTTCTACUGC 3'-GGCA <u>U</u> AAAGATGACG	60.0 [+9.5]	57.0 [+6.5]	nt	50.5	-
3'	5'-CU <u>A</u> UGACCTCGACUAC 3'-GAU <u>A</u> CTGGAGCTGAUG	67.5 [+13.0]	67.5 [+13.0]	65.0 [+10.5]	54.5	+15.5
4	5'-GUG <u>A</u> GGGTUAAATTT 3'-CACTCCCAA <u>U</u> TAAA	58.5 [+14.0]	57.5 [+13.0]	48.0 [+3.5]	44.5	+23.5
5	5'-TUGTTCCCTT <u>U</u> AGTGA 3'-AAC <u>A</u> AGGGAAAUC <u>A</u> CT	60.5 [+12.5]	64.5 [+16.5]	52.0 [+4.0]	48.0	+25.0
1 MM <sup>b</sup>	5'-ATUG <u>U</u> ATTTCGAGTC 3'-TAACA <u>U</u> AAAGCTCAG	59.0 [+12.0]	65.5 [+18.5]	39.0 [-8.0]	47.0	+38.5
6	5'-GGU <u>A</u> UATAUAGGC 3'-CCA <u>U</u> AUATAUCCG	61.0 [+23.5]	63.0 [+25.5]	51.0 [+13.5]	37.5	+35.5
7	5'-UGCACAGGU <u>A</u> UATATAGGC 3'-CCA <u>U</u> AUATATCCGGCGTAU			49.0		

<sup>a</sup> Thermal denaturation curves were recorded in medium salt buffer ([Na<sup>+</sup>] = 110 mM, [Cl<sup>-</sup>] = 100 mM, pH 7.0 (NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>), [EDTA] = 0.2 mM) and [ON] = 1.0 μM. A = adenin-9-yl, C = cytosin-1-yl, G = guanin-9-yl, T/U = thymin-1-yl/uridin-1-yl DNA or 2'-O-(pyren-1-yl) RNA (bold and underlined) monomers.  $TA = T_m$  (5'-Inv:cDNA + 3'-Inv:cDNA) -  $T_m$  (probe duplex + dsDNA) <sup>b</sup> Italicized bases represent mismatched nucleotides with respect to the Inv1 target region. Note: Inv1 MM is not activated for dsDNA recognition of Inv1 target site ( $TA = -38.5$  °C). \* Two transitions observed. nt = no transition.

Invader probes were evaluated for inhibition of *in vitro* transcription by incubating the probes with the template DNA (SfoI-digested plasmid 14971) followed by addition of transcription reagents, RNAP and RNA triphosphates (rNTPs). The synthesized RNA fragments were then separated using denaturing PAGE and visualized with Sybr Gold stain. Initially, a screen was performed for Invader probes targeting the T7 template (Inv1 – Inv3; Figure B.2-2). The template DNA was cut by SfoI at 208 bp from the T7 initiation start site, which means the full-length transcript should be 208 bp.

Invader probes (Inv1 – Inv3) do appear to inhibit transcription (33 – 57 % of control signal) compared to the control, where no Invader probe was added (lane 2). We also looked at the corresponding unmodified dsDNA of Inv2 (DNA2), which shows no reduction in signal (lane 5). This demonstrates that the observed activity is facilitated by the modifications of the Invader probe (only difference between Inv2 and DNA2). These initial results were very encouraging, though we needed to determine whether the observed inhibition is sequence-specific (i.e. resulting from binding of Invader probes to the target DNA region). Another observation warranting further inquiry is the lack of truncated product formation that was expected for the probe (Inv3), targeting a mid-template region. To determine if the excess probes in the reaction mixture were responsible for the observed non-specific inhibition of RNAP, they were removed by using a spin column (isolated species >100 bp) after incubation with plasmid DNA to allow for binding but before addition of RNAP and rNTPs. A slight reduction of signal was observed compared to the control lane where no probe was added, but the inhibition is not as great as when excess probe is present during transcription (compare lane 3 in Figure B.2-2 and lane 3 in Figure B.4-2). One possible explanation could be the existence

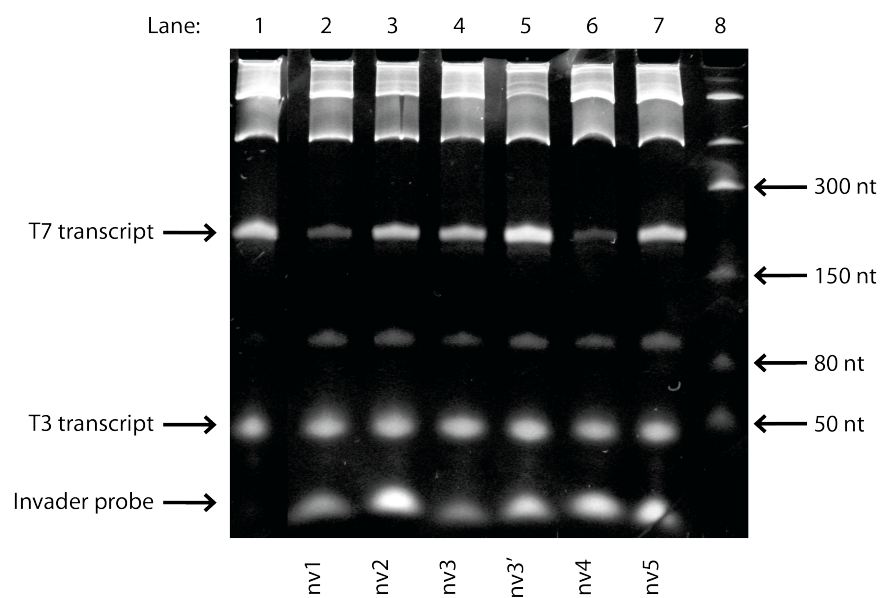
of non-specific interaction between the pyrene and RNAP, resulting in inhibition of RNA synthesis.



**Figure B.2-2.** Inhibition of in vitro transcription by Invader probes. Linearized plasmid DNA (0.5  $\mu$ g, SfoI digest of plasmid 14971) incubated with Invader probes (1000-fold molar excess) for 17 h at 37 °C in transcription buffer (contains  $Mg^{2+}$  and DTT, concentration and other components not published- Ambion) followed by addition of rNTPs (0.25 mM each) and T7 RNAP (15 U) and incubation for 1 h at 37 °C. Separation of fragments was achieved using 5% denaturing-PAGE and visualized with Sybr Gold stain. Quantification of signal was determined using ImageJ software.

Another experimental design was used to understand more about the inhibition of in vitro transcription by Invader probes. A double-digest of plasmid 14971 (SfoI and SpeI) was used, which results in two cuts, one near the T7 promoter and the other near the T3 promoter and results in transcripts using the respective RNAP, of 208 bp and 47 bp. The RNAPs are highly specific for their promoters, which can be used to determine if the given probe is specific for the intended target, thus lending support that this process is indeed a result of the hypothesized mechanism. A screen of the probes was conducted

(Figure B.2-3) to evaluate this. Probes expected to inhibit T7 transcription, do appear to result in a reduction of signal corresponding to the full-length T7 transcript (Inv1 – Inv3), while the T3 transcript signal is similar across lanes 2 – 4, suggesting that this inhibition is specific to the T7 RNAP. A new probe (Inv3') was designed to target a different region of the T7 template to determine whether the lack of truncated product formation is a result of the probe design, which showed no transition in thermal denaturation experiments, or inherent to the mechanism of inhibition. This probe did not appear to inhibit the T7 RNAP and a signal similar to the control (Figure B.2-3, lane 1) is observed. Although Inv4 and Inv5 target the T3 promoter region, they do not appear to inhibit transcription by T3 RNAP, but instead reduced signal corresponding to the signal of the full-length transcript from the T7 RNAP is observed in the case of Inv4 (Figure B.2-3, lane 6). This could indicate non-specific binding of the probe with the T7-promoter region or interactions between the probe and T7 RNAP that does not occur with the T3 RNAP.

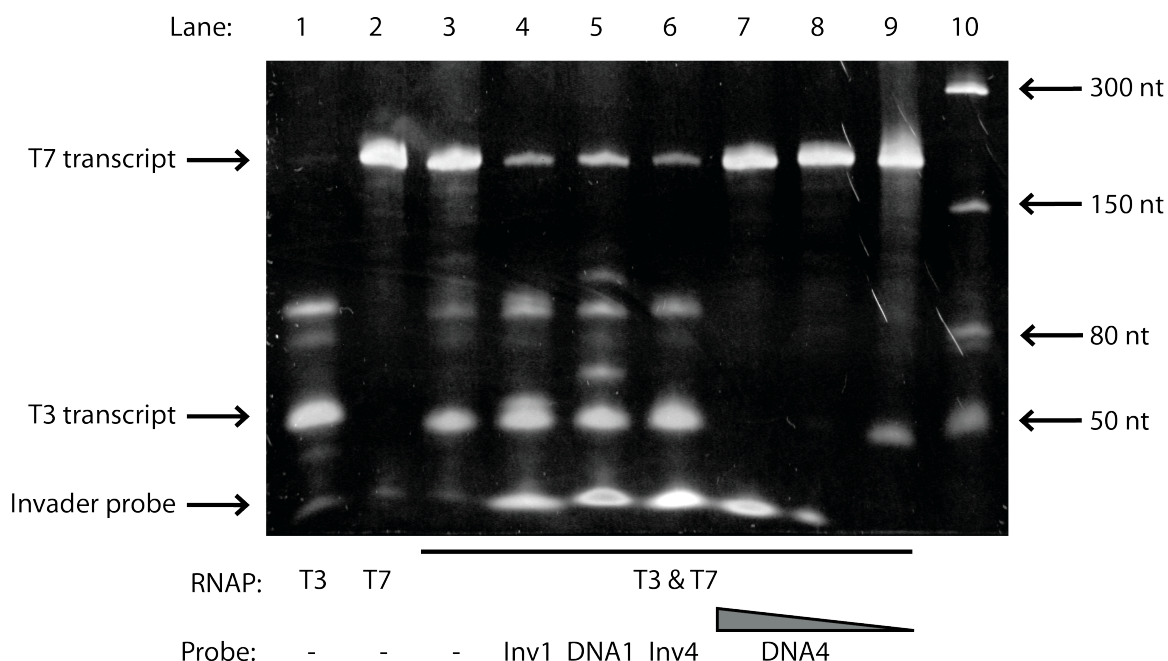


**Figure B.2-3.** Linearized plasmid DNA (SfoI and SpeI double-digest of plasmid 14971; 0.5  $\mu$ g) incubated with Invader probes (1000-fold molar excess, 17 h, 37  $^{\circ}$ C, in transcription buffer) followed by addition of rNTPs (0.25 mM each), T3 and T7 RNAP (30 or 15 U, respectively; lanes 1 - 7) and incubation at 37  $^{\circ}$ C for 1 h. Invader probes expected to inhibit T7 (lane 2 – 5) or T3 (lane 6 and 7) catalyzed RNA synthesis. Lane 8: ssRNA ladder. Separation achieved using 5% denaturing-PAGE and visualized with Sybr gold stain. Note that lack of band  $\sim$ 90 nt in lane 1 is likely due to loading differences, as other experiments showed this band. These experiments were repeated in a HEPES buffer with similar results.

Using the same experimental design, with both phage promoters on one template DNA (plasmid 14971), and simultaneous transcription by the corresponding RNAPs (T7 and T3), we further evaluated the activity observed for Inv1 and Inv4 and the corresponding unmodified DNA (DNA1 and DNA4, respectively) in Figure B.2-4. Both of these probes target a similar region of their corresponding promoters (T7 or T3; Figure B.4-4), namely the RNAP recognition element of the minimal promoter sequences used. Inv1 again shows reduced intensity of the band corresponding to the full length T7 transcript, however the corresponding unmodified DNA (DNA1) also shows reduction of signal. Since we assume that unmodified double-stranded DNA probes do not possess the favorable driving force for targeting dsDNA, DNA1 is likely acting as a decoy substrate.



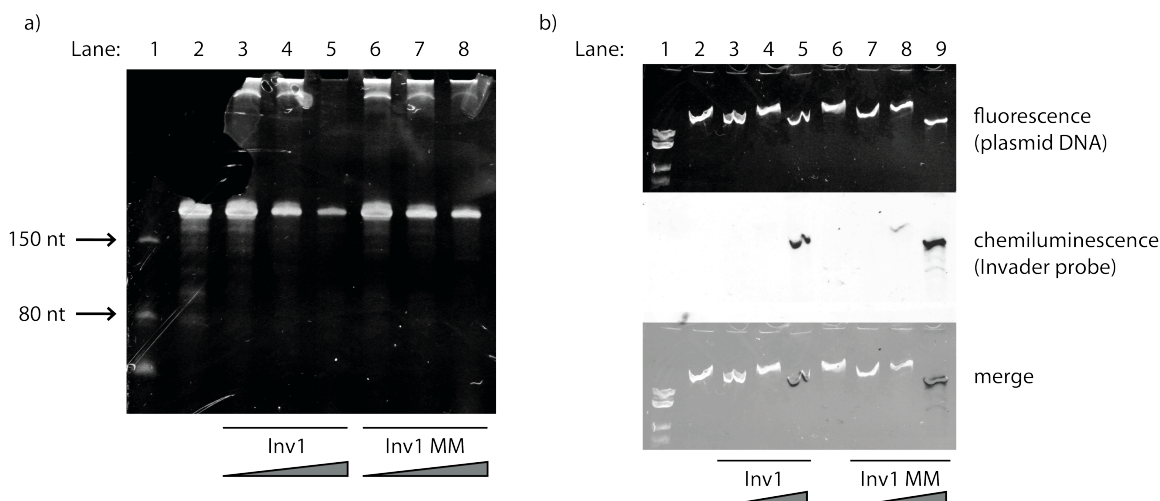
If the short double-stranded DNA probe contains part of the promoter region, it can be recognized by RNAP but does not result in RNA synthesis because of a very short template, thus sequestering the RNAP. This effectively reduces the signal of full-length product because of the lower concentration of available RNAP. This effect is also observed for DNA4, the corresponding unmodified DNA for Inv4, which has the same sequence as a portion of the T3 promoter. DNA4 results in complete inhibition of transcription from the T3 promoter, at high doses (1000- and 100-fold molar excess; lanes 8 and 9 of Figure B.2-4), with a small amount of full-length T3 transcript observed at the lowest dose (10-fold molar excess, lane 9). Interestingly, Inv4 does not result in reduced signal corresponding to the T3 transcript, and coupled with the inhibition observed for DNA4, suggests that T3 RNAP does not recognize the Invader probe duplex as a decoy substrate. This does not preclude probe:DNA duplexes being recognized by T3 RNAP, however. The implications being that even if the probe binds to the promoter region, it may be treated by the RNAP like natural dsDNA and recruits the RNAP to the polymerase in the case of promoter bound probe or when RNAP encounters such a duplex in the template region, it is simply dissociated to reveal the single-stranded template.



**Figure B.2-4.** Controls for observed inhibition of in vitro transcription of template plasmid DNA (SfoI and SpeI digest of plasmid 14971). Addition of T3 RNAP (lane 1), T7 RNAP (lane 2) or both T3 and T7 RNAP (lane 3 – 9). Invader probe or corresponding unmodified dsDNA (500-fold for lane 4 - 7; 100-fold or 10-fold for lane 8 and 9, respectively) was added to template DNA (0.25  $\mu$ g) and incubated for 17 h at 37 °C in transcription buffer followed by addition of rNTPs (0.42 mM) and RNAP (7.5 U) and incubation for 1 h at 37 °C. Lane 10: ssRNA ladder. Separation of fragments was achieved using 5% denaturing-PAGE and visualized with Sybr stain.

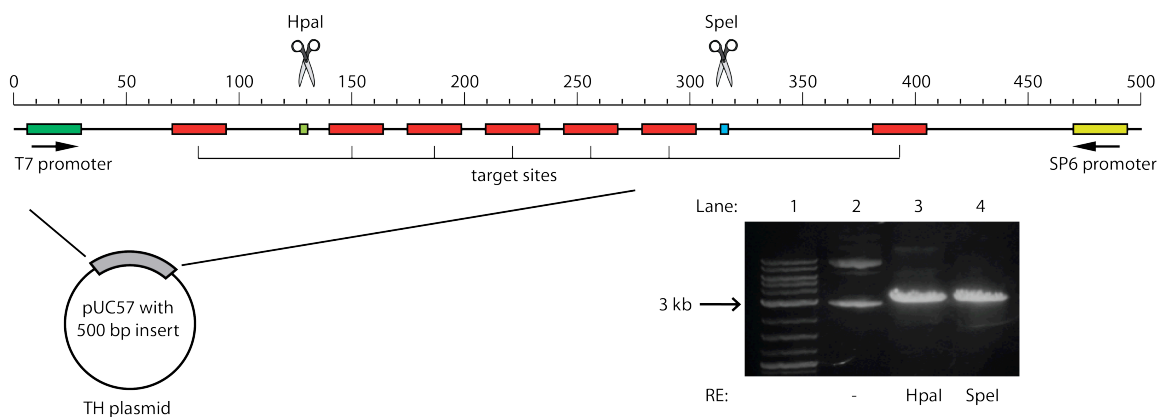
In another experiment to probe the sequence-specificity of the observed inhibition by Inv1, a mismatched probe (Inv1 MM) was used, which is fully complementary but contains three base pair mismatches with respect to the target sequence of Inv1. According to the thermal denaturation analysis (Table B.2-1), this probe has strongly unfavorable energetics for recognition of the T7 promoter. Varying the concentration of Inv1 and Inv1 MM (Figure B.2-5a), revealed a dose-dependent reduction of signal corresponding to the full length T7 transcript. To elucidate whether this is an effect of the non-specific inhibition or if indeed both probes are binding and resulting in the observed inhibition, a colocalization experiment was performed (Figure B.2-5b). The Invader

probes were DIG-labeled and incubated with the linearized plasmid DNA, then loaded onto non-denaturing PAGE. Plasmid DNA was visualized using Sybr stain (upper panel) and Invader probe was visualized by the chemilumnescent signal resulting from the DIG-label after processing (i.e. incubation with anti-DIG alkaline phosphatase and chemiluminescent substrate). This revealed that at high doses both Inv1 and Inv1 MM signal colocalize with plasmid DNA, suggesting that both are binding to the plasmid. To see if experimental conditions could favor binding of Inv1 over the mismatched probe, Inv1 MM, the temperature was varied (Figure B.4-3a) as well as the buffer used during incubation (Figure B.4-3b), which did not result in improved specificity. The binding of Inv1 MM could be due to sequence complementarity with the portion of the insert that is unknown to us, though that does not explain the observed inhibition. Because of this uncertainty, we decided to use a different system to evaluate the potential for Invader probes to sequence-specifically inhibit *in vitro* transcription.



**Figure B.2-5.** Evaluation of sequence-specificity of Invader probes for inhibition of in vitro transcription. (a) Electrophoretogram of RNA transcripts after incubation of plasmid 14971 (SfoI digest;  $0.25 \mu\text{g}$ ) with increasing concentrations of Invader probe (10-, 100-, or 1000-fold molar excess) incubated at  $37^\circ\text{C}$  for 17 h in transcription buffer, followed by addition of rNTP ( $0.25 \text{ mM}$  each) and T7 RNAP ( $7.5 \text{ U}$ ) and incubation for 1 h at  $37^\circ\text{C}$ . Separation achieved using 5% denaturing PAGE and visualized with Sybr stain. (b) Colocalization of increasing concentrations (10-, 100-, or 1000-fold molar excess) of Inv1 and Inv1MM with plasmid 14971 (SfoI digest) incubated for 17 h at  $37^\circ\text{C}$  in CutSmart buffer (1.43x:  $71.5 \text{ mM}$  potassium acetate,  $28.6 \text{ mM}$  tris-acetate,  $14.3 \text{ mM}$  magnesium acetate,  $143 \mu\text{g/mL}$  BSA, pH 7.9). Plasmid DNA visualized using Sybr stain (upper panel) after electrophoresis (4% nd-PAGE), which was subsequently processed for chemiluminescent detection of DIG-labeled probe (middle panel). The lower panel is a merger of these two images, which represents colocalization of Invader probe and plasmid DNA.

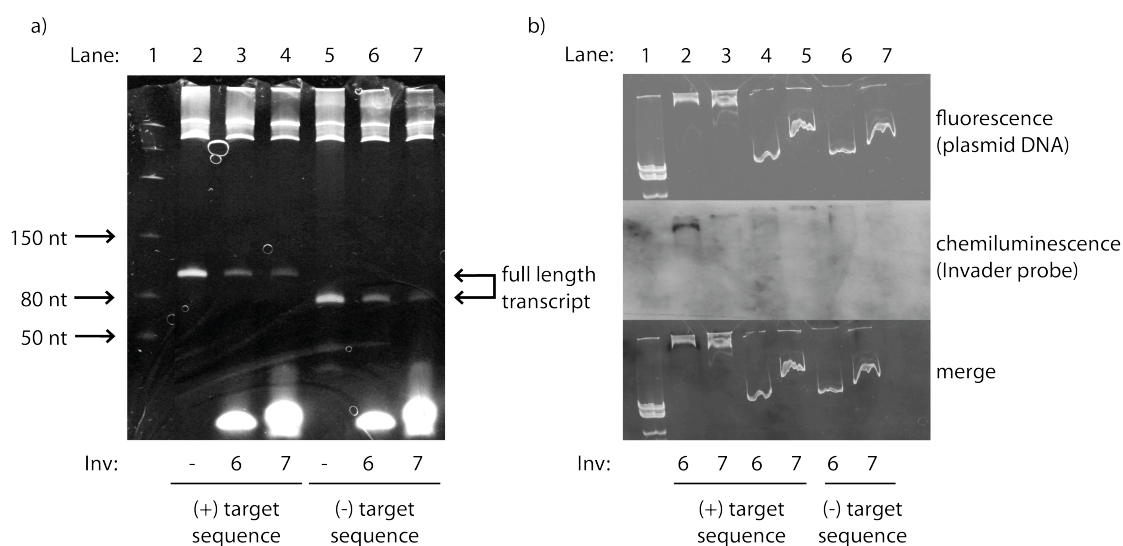
We decided to design and purchase our own plasmid construct (TH plasmid), which includes seven repeats of the same Invader probe binding site, and had it cloned into the pUC57 plasmid (Figure B.2-6). This allowed us to a) be certain of the sequence, b) conveniently design restriction sites and binding regions with respect to the phage promoters, and c) design optimized target sites. We also designed a control plasmid that lacks the target sites (TH control), but which otherwise has an identical sequence.



**Figure B.2-6.** Illustration of the TH plasmid construct (3.2 kb). Synthetic 500 bp insert with 7 identical target sites was cloned into the pUC57 vector. Digestion with HpaI (lane 3) or SpeI (lane 4) yields templates appropriate for in vitro transcription under T7 control with 1 or 6 binding sites, respectively. Electrophoretogram shows complete digestion of plasmid DNA with HpaI or SpeI after separation on 1% agarose gel and ethidium bromide staining. Lane 1: 1 kb ladder. A control plasmid (TH control plasmid) was also cloned, where the target sites were excluded but otherwise the sequence is identical (325 bp insert).

In vitro transcription using T7 RNAP and the TH plasmid linearized with HpaI results in 100 bp or 75 bp transcripts for the TH plasmid or TH control plasmid, respectively (Figure B.2-7). Inv6 and Inv7 (Table B.2-1) were designed to bind to the target sites located mid-template (55 nt from initiation start site). Inv7 contains 5'-single-stranded overhangs, or toe-holds, which has emerged as a promising new probe architecture. Both of these probes resulted in significant reduction in full-length transcript formation in the in vitro transcription assay (Figure B.2-7a, lanes 2 - 4), but again no traces of truncated product was observed. Unfortunately, this cannot be specific inhibition, as incubation of Inv6 or Inv7 with the linearized control plasmid (lanes 5 - 7) also results in decreased signal of the band corresponding to the full-length transcript. We also attempted inhibition of the longer transcript that includes six target sites (digestion with SpeI), but similar results were observed (data not shown). A colocalization assay was performed (Figure B.2-7b), in which Inv6 and Inv7 were incubated with either supercoiled (lanes 2

and 3) or linearized (lanes 4 and 5) TH plasmid, as well as the control plasmid lacking binding sites (lanes 6 and 7). In this case, no colocalization of the Invader probes was observed for linearized plasmid DNA, even though seven target sites were present. The supercoiled plasmid, however, does seem to result in some binding of Inv6 under these conditions. The differences in binding relative to the other plasmid, where linearized targets were recognized more efficiently than supercoiled, are perplexing. Moreover, the apparent



**Figure B.2-7.** Evaluation of Invader probes for inhibition of in vitro transcription using HpaI digests of TH plasmid as DNA template, which contains one copy of the Invader target sequence under T7 RNAP control, and comparison to TH control template, which lacks the target sequence (a). Electrophoretogram of RNA transcript after incubation with template (0.25  $\mu\text{g}$ ) with Invader probe (1000-fold molar excess) incubated at 37  $^{\circ}\text{C}$  for 17 h in transcription buffer, followed by addition of rNTP (0.25 mM each) and T7 RNAP (3.75 U) and incubation for 4 min at 37  $^{\circ}\text{C}$ . (b) Colocalization of Inv6 and Inv7 (1000-fold molar excess) with supercoiled (lane 2 and 3) or linearized TH plasmid (7 binding sites; lanes 4 and 5) or TH control plasmid (no binding sites; lanes 6 and 7) incubated for 17 h at 37  $^{\circ}\text{C}$  after annealing for 2 min at 95  $^{\circ}\text{C}$  in CutSmart buffer (1.43x: 71.5 mM potassium acetate, 28.6 mM tris-acetate, 14.3 mM magnesium acetate, 143  $\mu\text{g}/\text{mL}$  BSA, pH 7.9).

### **B.3 Conclusions**

In conclusion, this project was suspended because of the lack of specific inhibition of T7 RNAP by Invader probes. In one system we observed non-specific binding of Invader probes (Inv1 MM binding) and in the other system we do not observe binding of the probes to linear templates, however in both cases we see inhibition of in vitro transcription. Non-specific interactions with RNAP could result in inhibition of transcription, however experiments where the probes were removed prior to the transcription reaction also resulted in reduced signal corresponding to the full-length transcript. There appears to be a difference specificity and activity of the different RNAPs, so a deeper understanding of this may guide us to using an RNAP that would more sensitive to minor perturbations of the promoter duplex to prevent docking or less likely to dislodge probe bound to the template. Also, a better understanding of the design requirements for specific binding of Invader probes to long dsDNA would benefit this project.

### **B.4 Supporting information**

#### *Protocol - synthesis and purification of Invader probes*

ONs modified with 2'-*O*-(pyren-1-yl) RNA monomers were synthesized on an automated DNA synthesizer (0.2  $\mu$ mol scale) using a long chain alkyl amine controlled pore glass (LCAA-CPG) solid support with a pore size of 500 Å. The corresponding 2'-*O*-(pyren-1-yl) RNA phosphoramidites U/C were prepared as previously described<sup>6</sup> and incorporated into ONs via hand-couplings (0.05 M in acetonitrile, using 0.01 M 4,5-dicanoimidazole as the activators (15 min)) with extended oxidation (45 s). Treatment with 32% ammonia

(55 °C, 17 h) facilitated deprotection and cleavage from solid support. DMT-protected ONs were purified via ion-pair reverse phase HPLC (XTerra MS C18 column: 0.05 M triethyl ammonium acetate and acetonitrile gradient) followed by detritylation (80% acetic acid, 20 min) and precipitation (NaOAc, NaClO<sub>4</sub>, acetone, -18 °C, 16 h). The purity and identity of synthesized ONs were verified using analytical HPLC (>85% purity) and MALDI-MS analysis (Table B.4-1) recorded on a Quadrupole Time-of-Flight (Q-TOF) mass spectrometer using THAP and ammonium citrate as the matrix.

#### *Protocol - thermal denaturation experiments*

The concentrations of ONs were estimated using the following extinction coefficients (OD<sub>260</sub>/μmol): G (12.01), A (15.20), T (8.40), C (7.05), pyrene (22.4). Thermal denaturation temperatures were calculated as the first-derivative maximum of the A<sub>260</sub> vs. *T* curve. ONs (1.0 μM) were annealed (85 °C for 2 min) in medium salt buffer ([Na<sup>+</sup>] = 110 mM, [Cl<sup>-</sup>] = 100 mM, pH 7.0 (NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>), [EDTA] = 0.2 mM) and subsequent cooling to the starting temperature. The experimental temperature ranged from at least 15 °C below *T*<sub>m</sub> to 15 °C above *T*<sub>m</sub>, with the *T*<sub>m</sub> being determined as the average of two experiments within ±1.0 °C.

#### *Protocol - transformation of cloned plasmids and plasmid purification*

Plasmid 14971 was obtained from addgene.org as a bacterial stab. Single colonies were obtained after streaking on LB agar plates and incubation at 37 °C overnight, and amplified overnight while shaking at 37 °C in liquid culture (LB broth) all in the presence of ampicillin (100 μg/mL), per recommendations from the supplier.



TH plasmid and TH cloned plasmid were purchased from GeneScript and received as isolated plasmids. Prior to amplification, *E. coli* was transformed by adding 25 ng plasmid to barely thawed *E. coli* cells (100  $\mu$ L) in chilled culture tubes. Tubes were placed on ice again for 10 min, after which they were incubated at 42 °C for 2 min using a water bath then returned to ice. Heat-shocked cells were then diluted with LB broth (900  $\mu$ L) and incubated at 37 °C for 1 h. Tubes were further diluted with LB broth (1:10 or 1:100) and ampicillin was added prior to incubation overnight at 37 °C with agitation.

All plasmids were isolated from the liquid cultures using QIAprep spin mini prep per manufacturer's recommendations. Briefly, bacteria was pelleted at 10 krpm for 3 min and the supernatant was removed. Pellet was suspended in P1 buffer, to which P2 buffer was added and the solution was inverted until turning clear. N3 buffer was then added and the solution was centrifuged (13 krpm). The supernatant was then applied to the spin column and washed with PB and PE buffer. The purified plasmid was then eluted using water. Plasmid was quantified using Qubit fluorometric quantitation, using the BR dsDNA reagent. An average of three independent measurements were used in subsequent experiments. Plasmid purity was assessed using 1% TBE-agarose electrophoresis after ethidium bromide staining.

#### *Protocol – linearization of plasmid via restriction enzyme*

Plasmids were digested using the appropriate restriction enzyme according to the supplier's recommendations (NEB). Briefly, the plasmid and restriction enzymes were incubated in CutSmart buffer (1x: 50 mM potassium acetate, 20 mM tris-acetate, 10 mM magnesium acetate, 100  $\mu$ g BSA, pH 7.9) for 1 h at 37 °C followed by ethanol

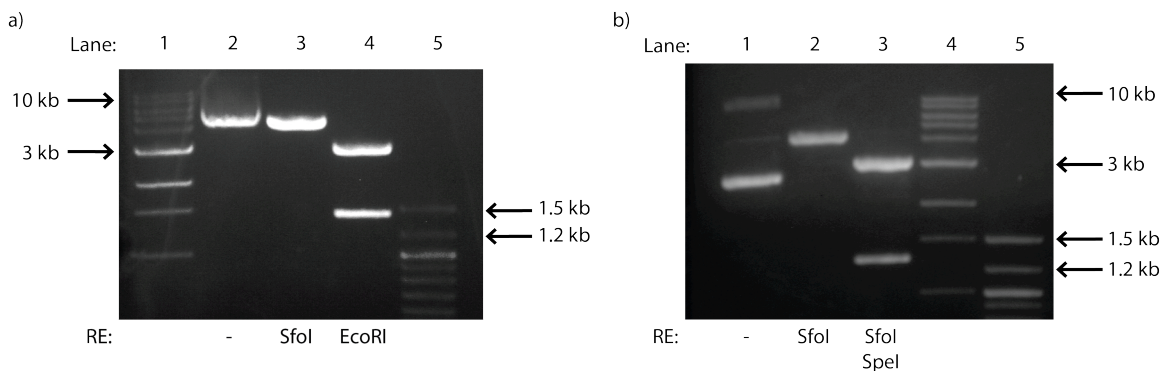
precipitation. EDTA (0.2 M; 1:20 reaction volume) and sodium acetate (3 M; 1:10 reaction volume) were added to the reaction mixture, mixed, and the DNA was then precipitated out by addition of abs. ethanol and storage at -20 °C (overnight). After pelleting (10 krpm, 10 min) and removal of supernatant, DNA was washed with ice-cold 70% ethanol and pelleted again. DNA was then dissolved in water and used in experiments with out further purification.

#### *Protocol- in vitro transcription*

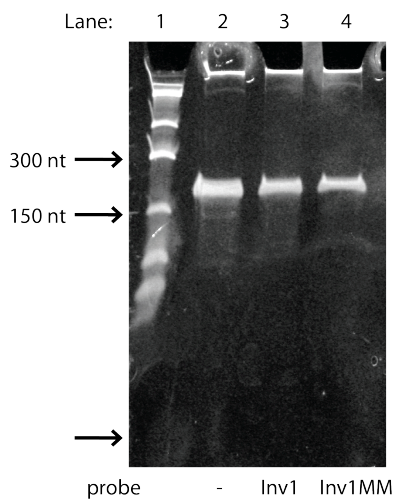
Invader probes were incubated with template DNA in transcription buffer (Ambion- does not specify formulation but does say that RNA polymerase requires magnesium ions and dithiothreitol) for 17 h at 37 °C prior to in vitro transcription to allow for binding. In vitro transcription reactions are generally performed according to manufacturing recommendations (Ambion), but exact condition are specified in figure captions. RNA polymerase (30 U per 1  $\mu$ g DNA) and RNA triphosphates (0.25 – 0.42 mM) were added, followed by incubation at 37 °C for 1 h, after which EDTA (2  $\mu$ L, 0.5 M) and RNA loading dye were added (2x) and denaturation was performed by heating to 95 °C for 2 min then immediately placing on ice. Samples were then loaded onto 4% denaturing PAGE (7 M urea) and electrophoresis was performed using constant voltage (80 V, RT, ~ 1.5 h). Excess probe was removed in select samples via Illustra S-300 HR microspin columns that isolates fragments >100 bp using protocols as suggested by the manufacturer.

#### *Protocol – colocalization experiments*

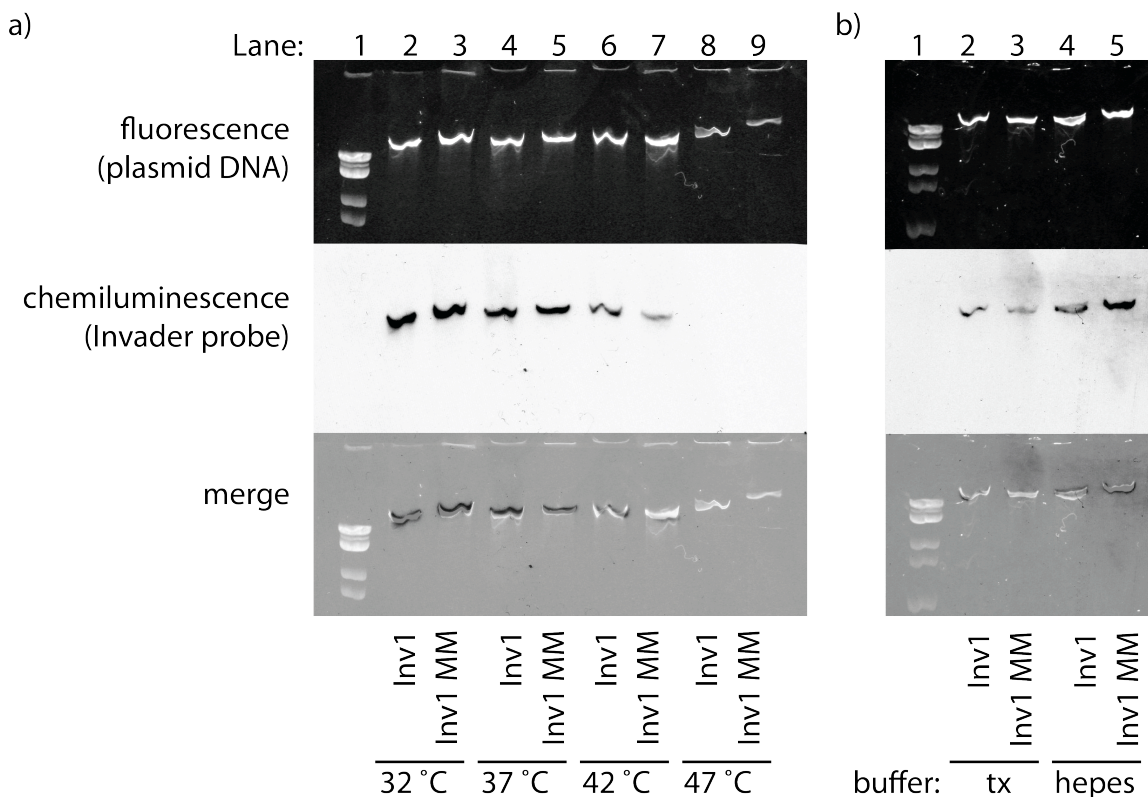
Invader probes were DIG-labeled using the 2<sup>nd</sup> generation DIG Gel Shift Kit (Roche Applied Bioscience). Briefly, 11-digoxigenin-ddUTP was incorporated at the 3'-end of the probe (100 pmol) using a recombinant terminal transferase. The reaction mixture was quenched through addition of EDTA (0.05 M), diluted, and used without further processing. Invader probes (DIG-labeled:unlabeled 1:10) was incubated with plasmid DNA at specified conditions, after which loading dye containing Sybr gold (1:1000) was added and the reaction mixtures were loaded onto 4% non-denaturing TBE-PAGE (45 mM tris-borate, 1 mM EDTA; acrylamide:bisacrylamide (19:1)). Electrophoresis was performed using constant voltage (70 V) at ~4 °C for 1.5 h. Bands were blotted onto positively charged nylon membranes (100 V, 30 min, ~4 °C) and cross-linked through exposure to UV (254 nm, 5 × 15 watt bulbs, 3 min). Membranes were incubated with anti-digoxigenin-alkaline phosphatase F<sub>ab</sub> fragments as recommended by the manufacturer, and transferred to a hybridization jacket. Membranes were incubated with the chemiluminescence substrate (CSPD) for 10 min at 37 °C, and chemiluminescence was captured on X-ray films.



**Figure B.4-1.** (a) Digestion of plasmid 14971 (~4.3 kb) by SfoI (lane 3) and EcoRI (lane 4), visualized by ethidium bromide after electrophoresis with 1% agarose gel (70 V, 1 h). Lane 1: 1 kb ladder. Lane 5: 100 bp ladder. SfoI (1 cut - 852 bp) and EcoRI (2 cuts - ~1400 bp). (b) Double-digest of plasmid 14971 with SfoI and SpeI (lane 3). Lane 4: 1 kb ladder and lane 5: 100 bp ladder.



**Figure B.4-2.** Evaluation of inhibition of in vitro transcription after incubation of Invader probe (1000-fold molar excess) with plasmid DNA (0.5  $\mu$ g, SfoI digest of plasmid 14971) for 17 h at 37 °C in transcription buffer. Prior to addition of rNTPs (0.25 mM each) and T7 RNAP (15 U), excess probes were removed from the reaction mixture using a spin column (isolates species >100 bp). The reaction mixture was then incubated for 1 h at 37 °C, followed by separation of fragments using 5% denaturing PAGE, and visualized with Sybr Gold stain.

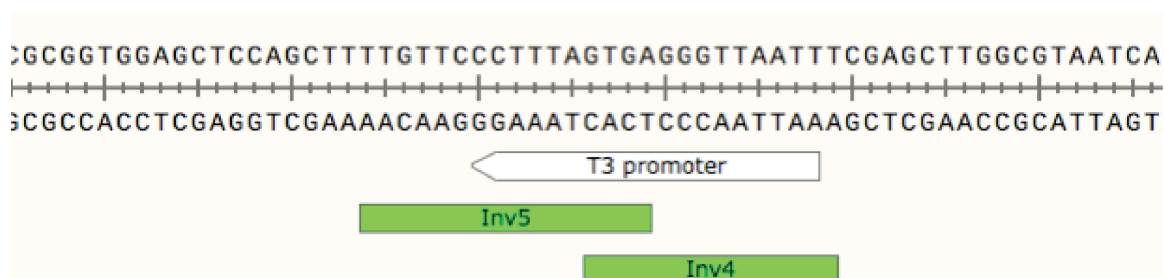
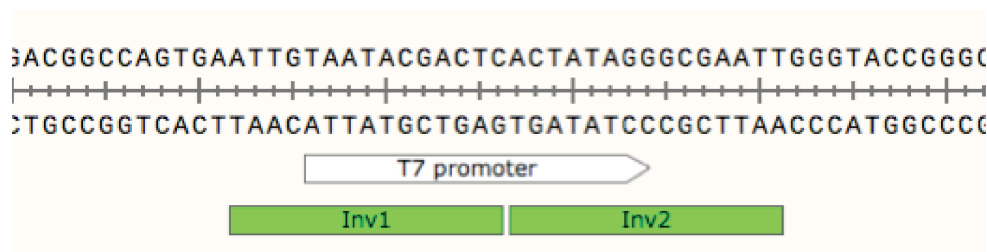


**Figure B.4-3.** (a) Colocalization of Inv1 and Inv1MM (1000-fold molar excess) with plasmid 14791 (SfoI digest) incubated for 17 h at specified temperature in CutSmart buffer (1.43x: 71.5 mM potassium acetate, 28.6 mM tris-acetate, 14.3 mM magnesium acetate, 143  $\mu\text{g}/\text{mL}$  BSA, pH 7.9). (b) Invader probes (1000-fold molar excess) incubated with linearized plasmid 14791 in different buffers (tx: Ambion transcription buffer; HEPES: 71.5 mM HEPES, 143 mM NaCl, 7.15 mM  $\text{MgCl}_2$ , 14.3% sucrose, pH 7.2, 2.06 mM spermine). Plasmid DNA visualized using Sybr stain (upper panel) after electrophoresis (4% nd-PAGE), which was subsequently processed for chemiluminescent detection of DIG-labeled probe (middle panel). The lower panel is a merge of these two images, which represents colocalization of Invader probe and plasmid DNA. For additional data see Figure B.2-3.

**Table B.4-1.** MALDI-MS of ONs modified with 2'-*O*-(pyren-1-yl)methyl-RNA monomers.<sup>a</sup>

Inv	Sequence	Observed <i>m/z</i> [M+H] <sup>+</sup>	Calculated <i>m/z</i> [M+H] <sup>+</sup>
1	5'-ATUG <u>U</u> AATACGACTC	5212.6	5211.0
	3'-TAACA <u>U</u> TATGCTGAG	5266.6	5265.0
2	5'-ACU <u>A</u> UAGGGCGAATT	5277.6	5276.0
	3'-TGA <u>U</u> AUCCCGCTTAA	5188.5	5187.0
3	5'-CCG <u>U</u> ATTTCTACUGC	5155.6	5154.0
	3'-GGC <u>A</u> UAAAGATGACG	5325.5	5324.0
3'	5'-CUAUGACCTCGACUAC	5447.9	5448.0
	3'-GAU <u>A</u> CTGGAGCTGA <u>U</u>	5622.1	5622.1
4	5'-GUG <u>A</u> GGGTUAATTT	5009.9	5009.9
	3'-CACTCCCAA <u>U</u> TAAA	4850.9	4851.0
5	5'-TUGTTCCTTUAGUGA	5500.7	5499.0
	3'-AAC <u>A</u> AGGGAAAUC <u>A</u> CT	5582.7	5581.1
1 MM	5'-ATUG <u>U</u> AATTTCGAGTC	5233.9	5233.0
	3'-TAACA <u>U</u> AAAAGCTCAG	5243.2	5243.0
6	5'-GGU <u>A</u> UATAUAGGC	4660.8	4660.9
	3'-CCA <u>U</u> AUATAUCCG	4540.8	4540.9
7	5'-UGCACAGGTAU <u>A</u> UATAGGC	6727.8	6728.3
	3'-CCATAU <u>A</u> UATCCGGC <u>G</u> TAU	6638.7	6639.3

<sup>a</sup> A, C and U denote 2'-*O*-(pyren-1-yl)methyladenosine, 2'-*O*-(pyren-1-yl)methylcytidine and 2'-*O*-(pyren-1-yl)methyluridine.



**Figure B.4-4.** Expected binding by Invader probes to the promoter regions for T7 and T3 RNAPs.

*Sequences of plasmid inserts*

Sequence of TH plasmid insert:

GGATCCGTAGATTGTAATACGACTCACTATAGGGCGAATTGCCACTCGAGAC  
 GATGCCGTATCGCTACTATGACTTTCACCATGCACAGGTATATATAGGCCGCA  
 TAACTGCGAACAGTCGAGGGCTAGCCAGCGTTAACAGCACTCGAGGCAGTGC  
 ACAGGTATATATAGGCCGCATAACAACGTGCACAGGTATATATAGGCCGCATA  
 GCCCATGCACAGGTATATATAGGCCGCATAGCGAGTGCACAGGTATATATAG  
 GCCGCATAGCTCATGCACAGGTATATATAGGCCGCATAAAGAAACCGCGGTTT  
 GACTAGTGCTGGCTAGCACGGCGGTGGCTGGAGATGGTGACCGAGCTGCTGA  
 GACAACTGCACAGGTATATATAGGCCGCATAGTTTCATCAGGTAACATCATC  
 ATCCACTATTCAAATCCTTGCGGACCGTATGCTCGGCCCGCGGCGCCGTGAA  
 GATATCACAGTGGATTTAACAGTAGAAGCTT

Sequence of TH control plasmid insert:

GGATCCGTAGATTGTAATACGACTCACTATAGGGCGAATTGCCACTCGAGAC  
 GATGCCGTATCGCTACTATGACTTTCACCAACTGCGAACAGTCGAGGGCTAG  
 CCAGCGTTAACAGCACTCGAGGCAGCAACGGCCAGCGAGGCTCAAGAAAC  
 CGCGGTTTCTAGTGTGGCTAGCACGGCGGTGGCTGGAGATGGTGACCGA  
 GCTGCTGAGACAACGTTTCATCAGGTAACATCATCATCCACTATTCAAATCCT  
 TGCGGACCGTATGCTCGGCCCGCGGCGCCGTGAAGATATCACAGTGGATTTA  
 ACAGTAGAAGCTT

## B.5 References

1. Guenther, D. C.; Anderson, G. H.; Karmakar, S.; Anderson, B. A.; Didion, B. A.; Guo, W.; Versteegen, J. P.; Hrdlicka, P. J. *Chem. Sci.*, **2015**, *6*, 5006.
2. Hao, N.; Krishna, S.; Ahlgren-Berg, A.; Cutts, E. E.; Shearwin, K. E.; Dodd, I. B. *Nucleic Acids Res.*, **2014**, *42*, 8861.
3. Larsen, H. J.; Nielsen, P. E. *Nucleic Acids Res.*, **1996**, *24*, 458.
4. Milne, L.; Xu, Y.; Perrin, D. M.; Sigman, D. S. *Proc. Natl. Acad. Sci.*, **2000**, *97*, 3136.
5. Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, F. C.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science*, **1992**, 1481.
6. Karmakar, S.; Anderson, B. A.; Rathje, R. L.; Andersen, S.; Jensen, T.; Nielsen, P.; Hrdlicka, P. J. *J. Org. Chem.* **2011**, *76*, 7119.



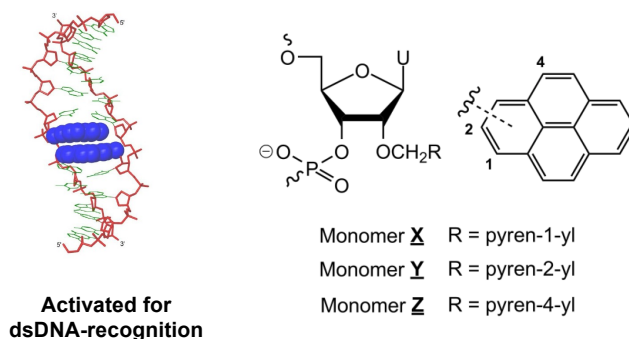
## APPENDIX C: Auxiliary projects

### C.1 Recognition of double-stranded DNA using energetically activated duplexes with interstrand zippers of 1-, 2- or 4-pyrenyl-functionalized O2'-alkylated RNA monomers

Saswata Karmakar, Andreas S. Madsen, Dale C. Guenther, Bradley C. Gibbons and Patrick J. Hrdlicka

Published in: *Org. Biomol. Chem.* **2014**, *12*, 7758.

Published by The Royal Society of Chemistry.



**Abstract.** Despite advances with triplex-forming oligonucleotides, peptide nucleic acids, polyamides and – more recently – engineered proteins, there remains an urgent need for synthetic ligands that enable specific recognition of double-stranded (ds) DNA to accelerate studies aiming at detecting, regulating and modifying genes. Invaders, i.e., energetically activated DNA duplexes with interstrand zipper arrangements of intercalator-functionalized nucleotides, are emerging as an attractive approach toward this goal. Here, we characterize and compare Invaders based on 1-, 2- and 4-pyrenyl-functionalized O2'-alkylated uridine monomers X–Z by means of thermal denaturation

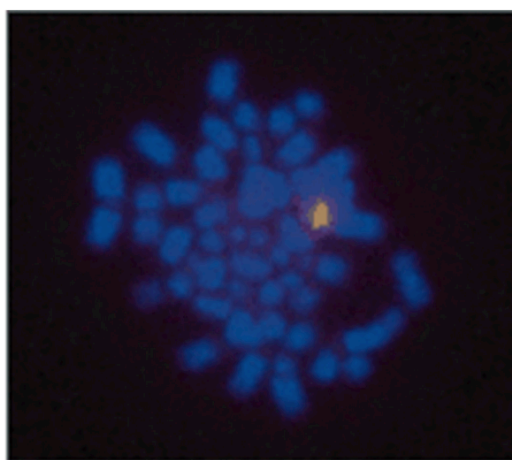
experiments, optical spectroscopy, force-field simulations and recognition experiments using DNA hairpins as model targets. We demonstrate that Invaders with +1 interstrand zippers of X or Y monomers efficiently recognize mixed-sequence DNA hairpins with single nucleotide fidelity. Intercalator-mediated unwinding and activation of the double-stranded probe, coupled with extraordinary stabilization of probe–target duplexes ( $\Delta T_m$ /modification up to +14.0 °C), provides the driving force for dsDNA recognition. In contrast, Z-modified Invaders show much lower dsDNA recognition efficiency. Thus, even very conservative changes in the chemical makeup of the intercalator-functionalized nucleotides used to activate Invader duplexes, affects dsDNA-recognition efficiency of the probes, which highlights the importance of systematic structure–property studies. The insight from this study will guide future design of Invaders for applications in molecular biology and nucleic acid diagnostics.

**C.2 Invaders: Recognition of double-stranded DNA by using duplexes modified with interstrand zippers of 2'-O-(pyren-1-yl)methyl-ribonucleotides**

Dale C. Guenther,\* Bradley A. Didion,\* Saswata Karmakar,\* Sujay P. Sau, John P. Verstegen, and Patrick J. Hrdlicka

\* Joint first authorship

Published in: *ChemBioChem* **2013**, *13*, 1534.

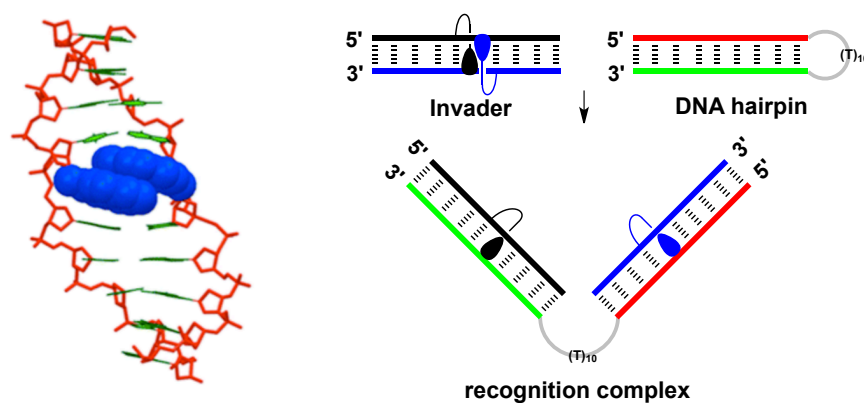


**Abstract.** Invaders are shown to recognize DNA hairpins in cell-free assays and chromosomal DNA during non-denaturing fluorescence in situ hybridization (nd-FISH) experiments. As Invaders are devoid of inherent sequence limitations, many previously inaccessible DNA targets could become accessible to exogenous control with important ramifications for karyotyping, in vivo imaging, and gene regulation.

### C.3 Identification and characterization of second-generation Invader Locked Nucleic Acids (LNAs) for mixed-sequence recognition of double-stranded DNA

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

Published in: *J. Org. Chem.* **2013**, 78, 9560.



**Abstract.** The development of synthetic agents that recognize double-stranded DNA (dsDNA) is a long-standing goal that is inspired by the promise for tools that detect, regulate, and modify genes. Progress has been made with triplex-forming oligonucleotides, peptide nucleic acids, and polyamides, but substantial efforts are currently devoted to the development of alternative strategies that overcome the limitations observed with the classic approaches. In 2005, we introduced Invader locked nucleic acids (LNAs), i.e., double-stranded probes that are activated for mixed-

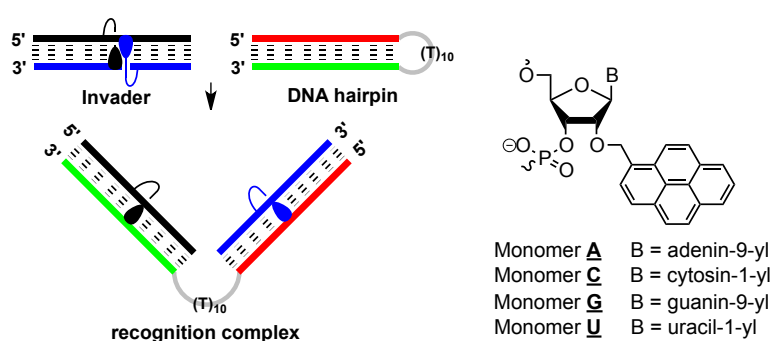
sequence recognition of dsDNA through modification with “+1 interstrand zippers” of 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA monomers. Despite promising preliminary results, progress has been slow because of the synthetic complexity of the building blocks. Here we describe a study that led to the identification of two simpler classes of Invader monomers. We compare the thermal denaturation characteristics of double-stranded probes featuring different interstrand zippers of pyrene-functionalized monomers based on 2'-amino- $\alpha$ -L-LNA, 2'-*N*-methyl-2'-amino-DNA, and RNA scaffolds. Insights from fluorescence spectroscopy, molecular modeling, and NMR spectroscopy are used to elucidate the structural factors that govern probe activation. We demonstrate that probes with +1 zippers of 2'-*O*-(pyren-1-yl)methyl-RNA or 2'-*N*-methyl-2'-*N*-(pyren-1-yl)methyl-2'-amino-DNA monomers recognize DNA hairpins with similar efficiency as original Invader LNAs. Access to synthetically simple monomers will accelerate the use of Invader-mediated dsDNA recognition for applications in molecular biology and nucleic acid diagnostics.

## C.4 Recognition of mixed-sequence DNA duplexes: Design guidelines for Invaders based on 2'-O-(pyren-1-yl)methyl-RNA monomers

Saswata Karmakar, Dale C. Guenther, and Patrick J. Hrdlicka

Published in: *J. Org. Chem.* **2013**, 78, 12040.

Reproduced by permission of The Royal Society of Chemistry



**Abstract.** The development of agents that recognize mixed-sequence double-stranded DNA (dsDNA) is desirable because of their potential as tools for detection, regulation, and modification of genes. Despite progress with triplex-forming oligonucleotides, peptide nucleic acids, polyamides, and other approaches, recognition of mixed-sequence dsDNA targets remains challenging. Our laboratory studies Invaders as an alternative approach toward this end. These double-stranded oligonucleotide probes are activated for recognition of mixed-sequence dsDNA through modification with +1 interstrand zippers of intercalator-functionalized nucleotides such as 2'-O-(pyren-1-yl)methyl-RNA monomers and have recently been shown to recognize linear dsDNA, DNA hairpins, and chromosomal DNA. In the present work, we systematically studied the influence that the nucleobase moieties of the 2'-O-(pyren-1-yl)methyl-RNA monomers have on the recognition efficiency of Invader duplexes. Results from thermal denaturation, binding

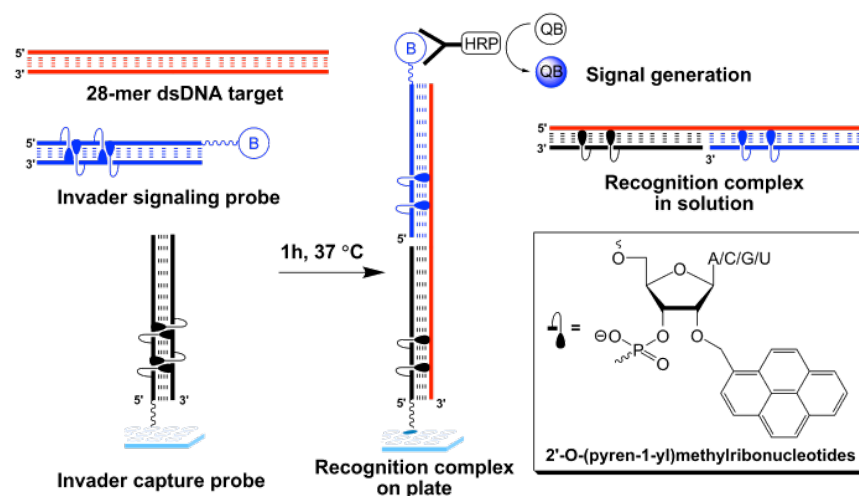
energy, and recognition experiments using Invader duplexes with different +1 interstrand zippers of the four canonical 2'-*O*-(pyren-1-yl)methyl-RNA A/C/G/U monomers show that incorporation of these motifs is a general strategy for activation of probes for recognition of dsDNA. Probe duplexes with interstrand zippers comprising C and/or U monomers result in the most efficient recognition of dsDNA. The insight gained from this study will drive the design of efficient Invaders for applications in molecular biology, nucleic acid diagnostics, and biotechnology.

## C.5 Sandwich assay for mixed-sequence recognition of double-stranded DNA:

### Invader-based detection of targets specific to foodborne pathogens

Benjamin Denn, Saswata Karmakar, Dale C. Guenther and Patrick J. Hrdlicka

Published in: *Chem. Commun.* **2013**, 49, 9851.



**Abstract.** A 96-well plate sandwich assay based on Invader capture/signalling probes is used to recognize 28-mer mixed-sequence dsDNA targets specific to *Salmonella enterica*, *Campylobacter jejuni*, *Escherichia coli*. Targets are detected down to 20 – 55 pM concentration with excellent binding specificity.



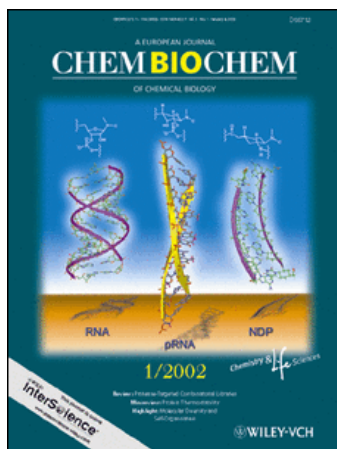
## APPENDIX D: Copyright permissions

11/11/2015

Rightslink® by Copyright Clearance Center



# RightsLink®

[Home](#)
[Account Info](#)
[Help](#)


**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. Versteegen, Patrick J. Hrdlicka

**Publication:** ChemBioChem

**Publisher:** John Wiley and Sons

**Date:** Aug 23, 2013

Copyright © 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Logged in as:  
Dale Guenther  
University of Idaho  
Account #:  
3000971108

[LOGOUT](#)

### Order Completed

Thank you for your order.

This Agreement between University of Idaho -- Dale Guenther ("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	3746271224181
License date	Nov 11, 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. Versteegen, Patrick J. Hrdlicka
Licensed Content Date	Aug 23, 2013
Licensed Content Pages	5
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Abstract
Will you be translating?	No
Title of your thesis / dissertation	Exploiting the energy of intercalation for sequence specific-specific recognition of double stranded DNA
Expected completion date	Dec 2015
Expected size (number of pages)	275
Requestor Location	University of Idaho 875 Perimeter Dr.

<https://s100.copyright.com/AppDispatchServlet>

1/2

11/11/2015

Rightslink® by Copyright Clearance Center

	MS2343
	MOSCOW, ID 83844
	United States
	Attn: Dale Guenther
Billing Type	Invoice
Billing address	Dale Guenther
	875 Perimeter Dr.
	MS2343
	Moscow, ID 83844
	United States
	Attn: Dale Guenther
Total	0.00 USD

**Would you like to purchase the full text of this article? If so, please continue on to the content ordering system located here: [Purchase PDF](#)**

**If you click on the buttons below or close this window, you will not be able to return to the content ordering system.**

[ORDER MORE](#)[CLOSE WINDOW](#)

Copyright © 2015 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#).  
Comments? We would like to hear from you. E-mail us at [customercare@copyright.com](mailto:customercare@copyright.com)

11/11/2015

Rightslink® by Copyright Clearance Center



RightsLink®

Home

Account  
Info

Help

ACS Publications  
Most Trusted. Most Cited. Most Read.**Title:** Identification and  
Characterization of Second-  
Generation Invader Locked  
Nucleic Acids (LNAs) for Mixed-  
Sequence Recognition of  
Double-Stranded DNA**Author:** Sujay P. Sau, Andreas S.  
Madsen, Peter Podbevsek, et al**Publication:** The Journal of Organic  
Chemistry**Publisher:** American Chemical Society**Date:** Oct 1, 2013

Copyright © 2013, American Chemical Society

Logged in as:  
Dale Guenther  
University of Idaho

LOGOUT

**PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE**

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK

CLOSE WINDOW

Copyright © 2015 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement.](#) [Terms and Conditions.](#)  
Comments? We would like to hear from you. E-mail us at [customercare@copyright.com](mailto:customercare@copyright.com)